A Genetic Perspective Into Impulsive and Compulsive Behaviours

Nuno R. Zilhão

A Genetic Perspective Into Impulsive and Compulsive Behaviours

Nuno R. Zilhão

Reading committee

Dr. J.H.D.A. van Beek	Open Universiteit Nederland
Prof. dr. P. Cuijpers	Vrije Universiteit Amsterdam
Dr. D. van 't Ent	Vrije Universiteit Amsterdam
Prof. dr. P.J. Hoekstra	Universitair Medisch Centrum Groningen & Rijksuniversiteit Groningen
Prof. dr. P.E. Slagboom	Leids Universitair Medisch Centrum
Prof. dr. A.H. Zwinderman	Academisch Medisch Centrum & Universiteit van Amsterdam

Paranymphs

Hrafnhildur Jóhannesddóttir

Nils von der Assen

Acknowledgements

This project was funded by FP7-People-2012-ITN, project: TS-Eurotrain, grant number 316978. I also acknowledge BBR Foundation (NARSAD) 21668; ZonMW (Addiction) 31160008; European Research Council (ERC-230374); NWO 480-15-001/674: Netherlands Twin Registry Repository: "Researching the interplay between genome and environment"; the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health, Rutgers University Cell and DNA Repository (NIMH U24 MH068457-06), the Avera Institute, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH R01 HD042157-01A1, MH081802, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995); Tourette syndrome association USA research award program: "The genetic epidemiology of tics"; and the Nederlandse Hersenstichting 11F03.14: "The search for endophenotypes in OCD".

ISBN: 978-9935-24-295-2

Print: Pixel ehf (www.pixel.is)

Cover design: Hrafnhildur Jóhannesddóttir (hrafnhildurj@gmail.com)

Artwork and layout: Hrafnhildur Johannesddottír

VRIJE UNIVERSITEIT

A Genetic Perspective Into Impulsive and Compulsive Behaviours

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Gedrags- en Bewegingswetenschappen
op donderdag 18 januari 2018 om 11.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Nuno Rodrigues Zilhão Nogueira

geboren te Porto, Portugal

promotor: prof.dr. D.I. Boomsma

copromotoren: dr. D.C. Cath
dr. D.J.A. Smit

The great Oz has spoken,
pay no attention to the man behind the curtain!

– Wizard of Oz, L. Frank Baum

Table of Contents

Chapter 1
Introduction

Review:
Genetics of Obsessive-Compulsive Disorder, Tourette's Disorder and Hoarding Disorder

This chapter is based on: Zilhao, N. R., Boomsma, D. I., Smit, D. J. A, Cath, D. C. Genetics of Obsessive-Compulsive Disorder and Tourette's Syndrome. This chapter is (with the exception of the paragraphs on hoarding disorder) to be published (accepted) in: Genes, Brain and Emotions, Oxford University Press, eds. Klaus-Peter Lesch, Judith Homberg and Andre Miu; 2018.

Abstract

Obsessive Compulsive Disorder (OCD), Tourette's Disorder (TD) and Hoarding disorder (HD) are common psychiatric disorders affecting approximately between 1-3% of the population. Since their first descriptions in the nineteenth century to today, we have witnessed considerable progress in the way these disorders are conceptualized, both clinically and genetically - from simple manifestations of psychological disturbances to the notion of highly polygenic and heritable disorders. Family and twin studies have shown that both disorders run in families, and that this familiarity is in part due to genetic factors, with heritability estimates ranging between 0.26-0.60 for OCD, 0.25-0.54 for TD and 0.35-0.50 for HD. Over the years, advances in the fields of psychiatric genetics have sought to uncover specific genetic variants underlying the etiology of these disorders, and although there has been a steady progress, results so far remain inconclusive. In the recent decades, a remarkable increase in international collaborative efforts within the field of psychiatric genetics has opened new possibilities of research, with larger datasets and multiple sources of genetic information available. In this chapter, we provide a broad overview into the genetics of OCD, TD and HD and on their genetic association over the years until the present. We summarize historical evidence from the literature over the years, and the recent results in the light of current research.

The first scientific descriptions of Obsessive Compulsive disorder (OCD) and Tourette's Disorder (TD) as psychiatric disorders stem from medical research in the nineteenth century, starting with the works of Jean Itard (1825) and George Gilles de la Tourette (1885) (on tic disorders), and of Pierre Janet (1900) (on OCD). Hoarding disorder (HD), has only recently sprouted attention within the research community. It still remains, actually, as the focus of a debate of whether it constitutes a separate mental disorder rather than a symptom of OCD. Conceptualization of these disorders over time reflected the evolving views regarding underlying causes and mechanisms. The Freudian psychogenic view prevailed for many decades, explaining both TD/tics and OCD symptoms as manifestations of inner psychological or psychosexual conflicts as a result of environmental adversities (predominantly stemming from early life traumatic events and disadvantageous upbringing). From the second half of the twentieth century onwards, advances in pharmacotherapeutics and new neurobiological insights (Pauls & Leckman, 1986; Peterson et al., 1993; Shapiro, Shapiro, & Feinberg, 1988), questioned the psychoanalytical explanations of tics and OC symptoms, and provided new insights into more neurobiological genesis of these diseases. Developments within the field of twin and family studies, following applications of structural equation modelling (Martin, Eaves, & Davies, 1978), have advanced the field of psychiatric genetic research. Both family-based and epidemiological twin studies have sought to estimate the contribution of genetic and environmental influences to variation in the phenotype, often specified as the heritability of the liability to develop the disease. Collaborative psychiatric genetic research further contributed to the insight into the genetic basis of both OCD and TD, as well as their shared underlying etiology (den Braber et al., 2016; Mattheisen et al., 2014; Scharf et al., 2013; Stewart et al., 2013). This rapidly growing field now also aims at finding the genetic DNA variants that help to identify pathways of disease development. These efforts, when successful will aid in understanding the biological mechanisms underlying the disorders. Here, we provide an overview of the genetic studies into OCD, TD and HD, and of the genetic comorbidity between these disorders.

Symptom Characteristics

Clinical Definition and Symptomatology

The current definition of OCD was introduced in the 1980s in the Diagnostic and Statistical Manual of Mental Disorders, third edition (American Psychiatric Association [APA], 1980). Until then, OCD was commonly referred to as "obsessive compulsive neurosis", or psychasthenia. OCD had been considered a sub-dimension of depression (World Health Organization

[WHO], 1992). According to the fourth edition of the Diagnostic and Statistical Manual of Mental disorders (American Psychiatric Association [APA], 1994), OCD was classified as an anxiety disorder, characterized by persistent, repetitive and intrusive thoughts/images (obsessions) that cause anxiety and or tension, and are in >80% of the cases followed by anxiety/tension reducing actions and ritualized behaviours (compulsions). Relief from carrying out compulsions is usually brief and the repetitive behaviour by and in itself fuels subsequent repetitive behaviour (van den Hout & Kindt, 2003), with the person getting caught in an on-going cycle of fear, doubt, worry and distress, leading to high levels of social impairment: Global Burden of Disease (GBD) estimates have found that OCD accounted for 2.2% of all years lost to disability (YLD) (Ustun, 2004), placing it at the 11th position of causes of non-fatal disease burden worldwide (World Health Organization [WHO], 2007). Whereas in DSMIV OCD was classified among the anxiety disorders, the new edition (DSM-5) includes OCD in its own category ('OCD and related disorders') alongside HD, excoriating disorder, body dysmorphic disorder and trichotillomania (American Psychiatric Association [APA], 2013).

The course of the disease is diverse: around 50% of cases remit with time, and 50% of persons with OCD run a chronic course (Skoog & Skoog, 1999). Recent studies from the Netherlands Twin Register showed that OCD diagnosis as well as symptom severity on average decreases from age 35 onwards, with an increase after age 60 that is predominantly driven by checking behaviour (Cath, Nizar, & Mathews, 2016). In all, OCD is considered as a lifelong disorder, which varies with regard to symptomatology. The most frequently reported obsessions are fear of contamination, a need for order and symmetry, and persistent and unwanted aggressive thoughts of causing harm to oneself or others (APA, 2013). Compulsions appear as repetitive behaviours or rituals in an attempt to decrease anxiety or tension as a result of the obsessions. The most frequently reported compulsions include washing and cleaning, checking, counting, symmetry and hoarding behaviour (Katerberg et al., 2010; Mataix-Cols, Rosario-Campos & Leckman, 2005).

According to the 5th edition of the DSM (2013), three categories of tic disorders are recognized: TD, chronic tic disorder (motor or vocal) and provisional tic disorder (APA, 2013). They are classified as neurodevelopmental disorders according to DSMIV and DSM5, and characterized by sudden, repetitive and unwanted motor movements or sounds (tics) (Cath et al., 2011). Tics can be either motor/vocal or simple/complex. Simple motor tics include brief, abrupt movements, involving only single muscle groups (e.g.: eye blinking and rolling, nose wrinkling, head jerk/nodding, shoulder shrugs). Complex motor

tics are performed as a sequence of simpler movements (e.g.: touching objects, hopping/jumping, squatting). Simple vocal tics include simple meaningless sounds and noises such as grunting, sniffing, throat clearing, coughing and snorting. Complex vocalizations entail words/syllables or making of animal sounds, echolalia (repeating another person's words), or coprolalia (obscene words) (Cath et al., 2011).

Tics have a fluctuating course with time. Strikingly, over 70% of TD children experience a decrease in tic symptoms in adolescence, starting at age 12, most likely as a result of maturation of the frontal lobes (Felling & Singer, 2011). Tics are usually more acute and expressive in periods of stress and anticipation, and tend to be reduced when the individual is mentally absorbed, focused and concentrated in activities. During adolescence, in up to 80% of tic-affected children, symptom severity decreases to an extent that the individuals no longer experience any distress and run a course that is indistinguishable from individuals with normal development (Burd, 2001; Leckman et al., 1998). However, in a small number of individuals, tics persist into adulthood and sometimes, even increase in severity. To what extent this differential course (remitting versus persistent) reflects differential genetic architecture, differences in environmental influences or differences in both has not been studied to date.

HD is defined as the pathological acquisition of (and failure to discard) large amounts of possessions and items leading to severe clutter, precluding the normal activities for which living spaces were designed (Frost & Hartl, 1996; APA, 2013). HD was for long considered a mere sub dimension of OCD or one of the criteria of Obsessive-compulsive personality disorder (APA, 1994). In DSM-5 HD has now been defined as a separate disorder as part of obsessive-compulsive spectrum disorders according to DSM5 (APA, 2013; Pertusa et al., 2008; Mataix-Cols et al., 2010). The arguments to grant HD a separate classification are based on HD clinical research showing that up to 80% of HD patients is without concurrent OCD symptoms, and on HD potential neurobiological distinctions from OCD (Tolin et al., 2014). This new diagnosis has been empirically validated, and deemed clinically reliable (Pertusa et al., 2011), following the core characteristics described by Frost and Hartl (1996). These recent conceptualizations have been corroborated by twin-based genetic studies (Iervolino et al., 2011; Ivanov et al., 2013).

Comorbidity

OCD is frequently accompanied by psychiatric comorbidities, mostly with concurrent major depression (31%), social or specific phobias (11%), anxiety disorders (25%), or bipolar disorder (7%) and 20% has co-morbid TD or tics (Angst et al., 2005; Canavera, Ollendick, & Pincus, 2010; Geller, 2006; Langley, Lewin, & Piacentini, 2010; Pallanti, Grassi, & Pellegrini, 2011; Storch, Lewin, & Murphy, 2010).

For people with TD, most frequently occurring co-morbid psychiatric disorders include ADHD and OCD (Freeman et al., 2000; Freeman et al., 2007). Between 20-89% of individuals with TD exhibit OC behavior (Karno, Golding, & Burnam, 1988; Weissman et al., 1994), and between 50-60% are diagnosed with ADHD (Polanczyk, de Lima, & Rohde, 2007). Furthermore, increased rates of anxiety and depression have been reported in individuals with TD (Cath & Ludolph, 2013). Hirschtritt et al. (2015) published a comprehensive analysis on the characteristics of ADHD and OCD as comorbid disorders in TD, reporting a mean of 2.1 comorbid disorders in TD patients and 1.6 when excluding OCD and ADHD (Hirschtritt et al., 2015).

The most commonly co-occurring conditions with HD include ADHD, Major Depressive Disorder and Generalized Anxiety Disorder (Mataix-Cols et al., 2012). Regarding co-morbid rates with OCD, this is estimated to be between 12-20% of HD patients (Frost et al., 2011; Ivanov et al., 2013).

Population Prevalence and Clinical Course

Lifetime and one-year prevalence of OCD is estimated at respectively 2.3% and 1.2% (Ruscio, Stein, & Kessler, 2010). Moreover, many individuals experience OC symptoms without the full diagnosis. Hence, the prevalence of OC symptoms and of subthreshold OCD (operationalized as having OC symptoms, but either less than 1 hour a day with significant suffering or distress, or > 1 hour per day without suffering or distress) is likely to be much higher (up to 6%) in the general population (Adam, Meinlschmidt, & Lieb, 2012; de Bruijn, Beun, & Denys, 2010; Fineberg et al., 2013). OCD shows a slight preponderance for women and a bimodal pattern for the age of onset. Boys typically experience an onset before the age of 10 and more often have tic-related symptoms, while girls experience a later onset (Nestadt, 2000; Pauls, Alsobrook, & Leckman, 1995).

For TD, Robertson et al. (Robertson et al., 2008) suggested an overall prevalence estimate of 1% in the general population, and Scahill, Dalsgaard,

and Bradbury (2013) reported prevalence rates between 0.5-0.7% in children between age 6 and 18 (Scahill, Dalsgaard, & Bradbury, 2013). For tic disorders, prevalence rates range between 0.3% and 0.8% for chronic motor tic disorder, whereas for vocal tic disorder no reliable estimate is currently available. In adults, tic prevalence rates are considerably lower than in children. Males are more commonly diagnosed with TD than females in ratios of 3-4:1 (Cavanna & Rickards, 2013; Robertson, 2000). The mean age of onset of TD is estimated at 7 years, ranging from 2 to 21 years of age (Bloch & Leckman, 2009). Typically, tics start at 5-7 years of age (range between 3 and 8 years), and worsen at the age of 12, followed by a decline in severity during adolescence. By adulthood, it is estimated that roughly one-third of children with TD may be tic-free. The onset of vocal tics usually occurs around 2 years later than motor tics.

With respect to HD, current estimates on the population and lifetime prevalence of hoarding symptoms are between 2-6% (Samuels et al., 2008; Mueller et al., 2009, 2008; Iervolino et el., 2009; Timpano et al., 2011). HD is a condition typically associated with old age (Cath et al., 2016), although it is suggested that hoarding-specific behaviour can originate in childhood (Grisham et al., 2006; Tolin et al., 2010). Ivanov and colleagues (2013) examined prevalence and heritability estimates of hoarding symptoms associated with young age and gender; this study was performed in a population-based sample of 15-year-old twins (N=3,974) in the Swedish Twin Register. It was reported prevalence rates of 2%, with a significantly higher preponderance for girls over boys, and suggested that sex-specific underlying etiological factors to hoarding symptoms may operate early in adolescence (Ivanov et al., 2013). No other study suggested quantitative or qualitative sex differences regarding influences from genetic vs. environmental factors in HD.

Genetic Studies

Family and Twin Studies: Heritability

Family studies investigate clustering of a trait within families, i.e. in biological relatives, and provide a first insight into a possible contribution of genetic factors to the etiology of complex disorders. However, increased familial aggregation is in itself not sufficient to prove a genetic origin of a disorder, because phenotypic resemblance between family members may also be due to family members sharing a common environment.

The first family studies reported on OCD were based solely on family history data without directly interviewing the relatives of patients themselves. Although these studies tend to underestimate the real disease rate within families, they

have generated consistent conclusions on the familiality of OCD (Eley et al., 2003; Hudziak et al., 2004; van Grootheest, Cath, & Boomsma, 2005). A total of 15 family study reports have appeared in the literature, of which 8 in adults and 7 in children/adolescents, all using direct interviews with at least one family member, and in some instances combined with family history data (Grados et al., 2001; Horwath & Weissman, 2000; Nestadt, 2000). Overall, the results showed an up to 10-fold increased rate of OCD in first-degree relatives of children/adolescents with OCD, and a 2-fold increase in relatives of adults with OCD. Interestingly, within families of affected male probands, male relatives tend to report more tic symptoms, while female relatives more often present obsessive-compulsive behaviour (Pauls et al., 1995).

In contrast to family studies, twin studies can make a distinction between genetic and shared environmental factors by modelling the differences in correlations between monozygotic and dizygotic twins (Boomsma, Busjahn, & Peltonen, 2002). Twin studies of OCD go back to 1929, with the first report on OC symptom pathology of twins by Lange (1929). In the following 40 years a number of twin studies were published in OCD (for a historical overview, see van Grootheest et al. (2005). Two of the largest clinical twin studies performed within the period of the DSMIII diagnostic system (Carey & Gottesman, 1981; Skre, Onstad, & Kringlen, 1993) reported concordance rates for OCD of 45% in MZ pairs and 15% in DZ pairs (Carey et al. 1981), and of 33% in MZ pairs and 7% in DZ pairs (Skre et al. 1993) respectively, indicating a genetic background for OCD.

Other studies on the heritability of OCD focused on OC symptoms (OCS) instead of OC disorder. By using continuous data from self-report questionnaires of a variety of OC symptoms, and population-based epidemiological samples as opposed to clinically-based samples, larger twin samples could be recruited (Macdonald, Murray & Clifford, 1991). The first relatively large study performed on OC symptoms was conducted in an epidemiological sample of 419 twin pairs by Clifford et al. (1984). Using genetic structural equation modeling (SEM), the study obtained heritability estimates for obsessional traits and OC symptoms of respectively, 44% and 47%. Subsequent studies, using larger epidemiological sample sets, provided heritability estimates ranging between 26% and 65%, depending on age of the sample and sex (Eley et al., 2003; Hudziak et al., 2004; Jonnal, Gardner, & Kendler, 2000; van Grootheest, Bartels, & Boomsma, 2007). Bolton et al. (2007) estimated the heritability of OC symptoms in a population-based sample of 854 6-year old twins to be 29% (Bolton, Rijsdijk, O'Connor, Perrin, & Eley, 2007). In recent years two studies have added a longitudinal

perspective to the twin design in cohorts of children, adolescents and adults (Bolhuis et al., 2014; van Grootheest, Cath, & Boomsma, 2009). Stability as expressed by longitudinal phenotypic correlations was between 0.50-0.63, with substantial genetic contributions to symptom persistence (34%-56%), indicating not only that OC symptoms are influenced by genetic factors, but also that these factors are highly stable over time in adults. In children, stability of OC symptoms was estimated at between 35%-51% for boys and between 28%-34% for girls (van Grootheest et al., 2007).

The first studies of familial effects in TD (Shapiro et al., 1988) described an increased risk of tics in families with TD probands. However, no distinction was made between recurrent childhood tics or newly onset adult tics in these reports. The first studies on newly developed tics in relatives of TD probands reported increased frequencies of tics/TD in these families when compared to the general population (Pauls, Cohen, & Kidd, 1981; Price, Kidd, & Leckman, 1985). The first large-scale family study utilizing direct interviews with patients and family members and employing a control group was published in 1991, showing a 10 to 100-fold increase among first-degree relatives of affected family members (Pauls, Raymond, & Leckman, 1991). Further, increased prevalences of tic disorders, ADHD, and OCD were found in parents and siblings of 4479 school-age children who were assessed for TD and tic disorders (Khalifa & von Knorring, 2005). Stewart et al. (2006) assessed (direct-interview) 692 relatives of 239 patients diagnosed for TD, ADHD/TD and ADHD, with the aim to explore familial associations between TD and ADHD. This case-control study found an increased risk of comorbid TD and ADHD among relatives of cases diagnosed for one of these two disorders (Stewart et al., 2006). A large multi-generational family sample from Sweden assessed 4826 individuals for TD and chronic tic disorder between the periods of 1969-2009. There was an increased risk for tic disorders proportional to the degree of genetic relatedness, with first-degree relatives having a risk of 1:18.69 (odds ratio), and heritability of tic disorder estimated at 0.77 (Mataix-Cols et al., 2015). A slightly different approach was taken by de Haan and colleagues (de Haan et al., 2015) who estimated heritability of different tic symptom factors, carrying out a factor analysis in a sample of probands (N=494) and their family members (N=351). Three factors (complex vocal tics and obscene behavior; body tics; head/neck tics) formed a core etiological entity, with heritabilities between 0.19-0.25. Recent factor analysis studies of symptoms tried to uncover tic-based subtypes or TD-related phenotypes including both tics and co-morbidities. Hirschtritt et al. (2016) analysed item-level data from 1191 TD probands and 2303 first-degree relatives and identified a sub-phenotype characterized by high rates

of social disinhibition, with a heritability of 0.53 (SE=0.08, P=1.7x10-18) (Hirschtritt et al., 2016). Darrow et al. (2016) reported similar results for 3494 individuals assessed for TD, OCD and ADHD symptoms – two cross-disorder (TD-related) phenotypes were identified: disinhibition (heritabilities 0.35, SE=0.03, p= 4.2x10-34) and symmetry (heritability 0.39, SE=0.03, p= 7.2x10-31) (Darrow et al., 2016).

Fewer twin studies have been performed in TD than in OCD. The available evidence points to a genetic basis of the disorder. The largest clinical study (MZ twins: N=60, DZ twins: N=26) found concordance rates in MZ and DZ twins of 77% and 23% (Price, Kidd, Cohen, Pauls, & Leckman, 1985). Another clinical study analysed the concordance rates for TD and tic disorders among MZ pairs (N=16 pairs) in relation to intra-pair differences in birth weight. Results showed 56% concordance rates of TD and 94% of tic disorder. The results suggested that prenatal events could be affecting different phenotypic expression for TD – i.e. intra-pair birth-weight difference was found to be a predictor of the intra-pair difference for tic scoring (Hyde, Aaronson, & Weinberger, 1992). Heritability estimates from epidemiological studies range between .28 and .56 (Bolton et al., 2007; Anckarsäter et al., 2011; Lichtenstein et al., 2010; Pinto et al., 2016).

As a result of hoarding disorder's past definitions in DSM-III and IV, the majority of previous published family studies on HD have used clinical samples of OCD and estimated frequencies of hoarding symptoms from the hoarding dimension of OCD. With respect to those studies specifically targeting hoarding symptoms, the numbers of non-OCD hoarding cases were usually low. Still, evidence gathered from somewhat earlier family studies has been suggestive towards the familiality of HD (Pertusa et al., 2008; Seedat & Stein, 2002; Winsberg et al., 1999).

Equally, only a handful of population-based heritability studies have been published so far. To date, a total of seven studies have estimated the genetic contributions to HD, and results range between 0.33-0.50 (Iervolino et al., 2009, 2011; Taylor et al., 2010; Ivanov et al., 2013; Nordsletten et al. 2013; Mathews et al., 2014). Taylor (2010) studied a Canadian-based sample of twin pairs (N=167 MZ; N=140 DZ) and reported heritability proportions of 0.42 (Taylor et al., 2010). Three studies have emerged from the TwinsUk twin register, analysing a sample consisting of adult females (N=5022); heritability estimates were reported at around 0.50 (Iervolino et al, 2009; 2011; Nordsletten et al., 2013). More recently, Mathews (2014) analysed a population-based family sample from the Netherlands Twin Register (NTR). This study included 7,906 twins and respective family members (totalling

15,914 individuals) and heritability was estimated at 0.36. The study from Ivanov et al. (2013) reported heritability of 0.32 for boys only, whereas for girls the role of additive genetic factors was negligible (Ivanov et al., 2013). It is noteworthy that these studies provided supportive evidence for the genetic contribution regarding HD symptomatology to be separate from OCD. The studies by Iervolino et al. (2009; 2011) estimated genetic correlations for OCD and HD at 0.45, within the range of what is reported for any other two internalizing disorders (Tambs et al., 2010).

Linkage Studies

Linkage studies seek to find genetic markers co-segregating with the disease/phenotype of interest that are possibly 'linked' to the causal genetic variants. The linkage approach involves scanning large family pedigrees with known disease transmission patterns, and narrowing down regions of the genome that segregate in association with the disorder.

So far, five studies on OCD have been performed (Hanna et al., 2002, 2007; Shugart et al., 2006). Results implicated genomic regions on chromosomes 1q, 3q, 6q, 6p, 6q, 7p, 9p, 10p, 11p, 14q, 15q and 19q. However, none of these results reached genome-wide significance. Hanna et al. in 2002, conducted a study in 56 individuals from seven families scanned for childhood OCD, and found region 9p24 to be a candidate region with linkage (Hanna et al., 2002), a finding that was replicated by Willour et al. (2004). A suggestive signal (LOD score= 3.13 - usually the threshold for genome-wide linkage significance is 3.6) was reported in three Costa Rican families in region 15q14 (Ross et al., 2011), overlapping with findings from a previous study (Shugart et al., 2006). The strongest signal so far (LOD score = 3.77) was recently found on the 1p36 region, by Mathews et al. (2012) in 33 Caucasian families (n= 245 individuals) with childhood onset OCD (Mathews et al., 2012).

In TD, five genome-wide linkage studies have been performed with inconsistent results (Barr, Wigg, & Tsui, 1997; Mérette et al., 2000; Simonic, Gericke, & Weber, 1998; The Tourette Syndrome Association International Consortium for Genetics [TSAICG], 1999, 2007). In 1997, the Tourette Syndrome Association International Consortium for Genetics (TSAICG), published a linkage scan on 110 sib pairs, and found suggestive evidence for linkage in regions 4q and 8p (TSAICG, 1999). Simonic et al. (2001) studied an Afrikaner population using linkage transmission disequilibrium tests (TDT), and replicated previous findings for linkage in three previously identified regions, 2p11, 8q22, and 11q23-24 (Simonic et al., 1998; Simonic et al., 2001). Interestingly, for this last region (11q23-24), another study from

a French-Canadian population of 127 families found the marker D11S1377 to be statistically significant for linkage (LOD score=3.18) (Mérette et al., 2000). The largest linkage study so far, was from TSAIC (2007) within a collaborative effort comprising 2040 individuals from 238 nuclear families and 18 multigenerational families. Statistically significant linkage was found for marker D2S144 on chromosome 2p32.2 (TSAICG, 2007).

Zhang and colleagues (2002) performed a genome-wide scan for HD in 77 sibling pairs diagnosed for TD, within the scope of TSAICG mentioned above. The top hits were in the regions 4q34-35 (P=.0007), 5q35.2-35.3 (P=.000002) and 17q25 (P=.00002). Despite only nominally significant, its noteworthy that the 4q region had been previously identified by the same group as possibly linked to TD (TSAICG, 1999). Similarly, another study attempted to find chromosomal regions specific to compulsive hoarding, scanning a total of 219 families with OCD, as part of the OCGS consortium. A region in chromosome 14 (marker D14S588) was found suggestively linked with HD (LOD=2.9) (Samuels et al., 2007) . Following this, Liang et al. suggested that the regions 9q (harbouring the SLC1A1 gene) and 14q interacted, when considering compulsive hoarding within OCD-diagnosed individuals (Liang et al., 2008). Importantly, a study by Mathews et al. (2007) provided a comprehensive analysis over a set of 11 multigenerational families with OCD, that had been gathered over the years. The goal was to phenotypically characterize this dataset regarding its power and other clinical characteristics, to detect linkage in the context of HD. Indeed in this report, Mathews (2007) showed that the total of 92 individuals (44 with OCD) that comprised this dataset had a number of characteristics (namely adequate power) to detect linkage within the scope of OCD, but a higher number of families was still needed for reasonably powered linkage studies in HD (Mathews et al., 2007).

Linkage studies were the primary method for genetic analysis until the era of large-scale genome-wide association studies. Because linkage studies require large family pedigrees and several affected generations with a disease of interest, or large numbers of affected sibling pairs, they became less popular over the years. When dense genotyping became feasible at reasonable costs, genome-wide association studies in unrelated subjects quickly lead to successes for numerous complex traits and diseases, including genetic disorders (Ripke et al., 2013).

Genetic Association Studies

Candidate Genes

Genetic association studies can be performed at a population level, in large samples of unrelated persons, and done either for candidate genes or for genome-wide SNP markers. Candidate gene studies require prior (biological) hypotheses, either from positional information of a gene(s), or on a gene(s) that has a potential functional relevance to the disease under study (gathered either from pharmacological studies, animal/theoretical models or linkage studies). With the advent of genome-wide association studies (GWAS) and the decreasing costs in measuring of genetic variants, candidate gene studies have become less important. Here we summarize the genes found so far through candidate gene studies, i.e. reported more than once to be implicated in either OCD, TD/tic disorders or HD. Most studies searching for genes associated with OCD, TD and HD focused on central nervous system (CNS) neurotransmitters, or CNS neurodevelopmental pathways. Particularly, genes in the serotonergic, dopaminergic and glutamatergic systems, that are also often the pharmacological targets of prescribed drugs, have been studied. Results have implied the glutamate transporter SLC1A1, the serotonin transporter SLC6A4, the MAOA gene encoding the enzyme monoamine oxidase A, the COMT gene encoding catechol-O-methyl transferase affecting dopaminergic functioning, the DRD4 gene, and the NTRK3 gene.

A de novo mutation in the serotonin transporter SLC6A4 gene has been implicated in OCD and comorbid disorders (anorexia nervosa, autism) in a study in two unrelated families (Ozaki et al., 2003). This finding was replicated in two family studies in OCD (Hu et al., 2006, Voyiaziakis et al., 2011). A recent meta-analysis of a total of 113 genetic association studies in OCD identified COMT and MAO-A polymorphisms associated with OCD in males (Taylor et al., 2013). Still, the only consistent replication so far, has been associated with the glutamate transporter gene SLC1A1 (Arnold et al., 2006, Wendland et al., 2009, Stewart et al., 2007).

For TD, attention has largely focused on SLITRK1 gene, a member of the SLIT and TRK family proteins, involved in the control of neurite outgrowth. In 2005, Abelson et al. mapped an inversion and two subsequent mutations at 13q33.1, close to the gene SLITRK1 (13q31.1), in two patients with TD (Abelson et al., 2005). Subsequently, a Canadian based family study identified one polymorphism tagging the major haplotype of SLITRK1, and these findings were later replicated (Karagiannidis et al., 2012; Miranda, Wigg, & Barr, 2008). With respect to the dopamine receptor gene DRD4, Liu et al. (2014) studied a tandem repeat polymorphism in DRD4 in a population-

based Han Chinese sample. Results suggested a protective role for the 2-repeat allele and a deleterious role for the 4-repeat allele (Liu et al., 2014). Lastly, a mutation in the HDC gene (involved in the neuronal histaminergic pathways) was described in a multi-generational linkage family-study where the father and all his eight children were diagnosed with TD (Ercan-Sencicek et al., 2010). A subsequent study in 520 nuclear families with TD, found a significant association of SNPs rs854150 and rs1894236 (intronic SNPs in gene HDC), providing supportive evidence for the role of this gene in the development of TD (Karagiannidis et al., 2013).With respect to HD, Alonso et al. (2008) specifically investigated whether the NTRK3 gene increases the susceptibility to hoarding in OCD, by studying 52 tag SNPs in a sample of 120 OCD patients and 342 controls. This gene was investigated following its implication in, both neuronal proliferation and differentiation during embryonic development stages, and neuronal growth and survival in the adult developed nervous system. This study identified the SNPs rs1017412 (OR=2.16, P=0.001) and rs7176429 (OR=2.78; P=0.0001), in the NTRK3 gene, associated with compulsive hoarding (Alonso et al., 2008).

GWAS

The focus of genetic research has recently shifted from candidate gene studies to GWAS that seek to establish an association between common genetic variants of small effect and a trait of interest. Through large international collaborations (e.g., the Obsessive-Compulsive Foundation International Genetic Consortium [OCF-genetics consortium], the OCD Collaborative Genetics Association Study Group [OCGAS consortium], the Tourette Syndrome Association International Consortium for Genetics [TSA genetic consortium], and TIC-genetics, all collated within the Psychiatric Genomics Consortium), and high-throughput analysis techniques, very large datasets of cases and controls have become available. Participants in these studies are genotyped, and the genotype data imputed based on a common reference set, allowing for meta-analyses in much larger numbers of subjects. The genomic era has allowed scientists to measure and impute millions of common genetic variants in the human genome, called SNPs (single nucleotide polymorphisms), which represent a form of common inter-individual genetic variation across the human genome.

GWAS studies are largely hypothesis-free and exploratory. They entail scanning the entire genome for SNPs that are associated with the phenotype of interest, i.e., that vary in allele frequency between cases and controls. The rationale behind these studies is that the collective variation embedded in all these SNPs can account for a statistically significant effect and play an

important role in the development of disorders. The exploratory nature of GWAS, involving large numbers of statistical tests, has forced a stringent control of type I error rate by setting very low p-value thresholds (the generally accepted genome-wide threshold is 5.0x10-8, (Hoggart, Clark, & Balding, 2008; Risch & Merikangas, 1996), prompting the need for large datasets. Unlike candidate gene studies, genome-wide data in GWAS additionally allow controlling for population stratification effects – the systematic difference in allele frequencies between subpopulations (or cases/controls). Here we briefly summarize the knowledge gathered so far in this field.

Four GWAS in total have so far been conducted for OCD, TD and HD. The International OCD Foundation Genetics Consortium (IOCDF-GC) performed a first GWAS in OCD, involving 1465 cases, 5557 controls and 400 trios (Stewart et al., 2013). Two separate case-control analyses were performed with and without the trios with a total of 469,410 autosomal SNPs and 9657 X-chromosome SNPs. The study found two trend-significant SNPs ($P = 2.49 \times 10\text{-}6$, $P = 3.44 \times 10\text{-}6$) located in the DLGAP1 gene in the case-control analysis, and a significant SNP ($P = 3{,}84 \times 10\text{-}8$) near the BTBD3 gene. However, in the subsequent meta-analysis combining the trios and the case-control samples, no genome-wide significant SNPs were obtained. A second GWAS was published recently by the OCGAS, consisting of 1406 individuals with OCD among a total sample of 5061 individuals from 1065 families (Mattheisen et al., 2014). The authors failed to find any genome-wide significant result, but their follow-up gene-based analysis revealed a significant association of OCD with IQCK, C16orf88 and OFC11, all protein-coding genes.

In TD, one GWAS study has been performed so far, published in 2012 by the TSAICG, in a sample of 1285 cases and 4964 ancestry-matched controls (Scharf et al., 2013). The top signal, a SNP located in the COL27A1 gene, did not reach genome-wide significance ($P = 1.85 \times 10\text{-}6$), which might be the result of power issues. Currently, several GWAS replication efforts are underway through world-wide consortia (TSAICG replication effort, EMTICS, TS-EUROTRAIN).

Finally, for HD one GWAS was published in 2010 in a sub-sample of genotyped individuals (N=3410) from the TwinsUK, consisting of predominantly female (91.8%) singletons (Perroud et al., 2010). Two loci showed suggestive evidence for association with HD, in chromosome 5 and 6. Noteworthy, these results remained after correcting for other OCD traits, suggesting specific effects of these SNPs separate from any other OCD sub-dimensions.

Copy Number variants (CNV) entail a form of variation where segments of DNA may be duplicated or deleted and people differ in the length of these segments. Following this approach, a study by Fernandez et al. (2012) analysed rare CNVs (population frequency <1%) in a case-control study of 460 individuals with TD. Pathway analysis revealed the enrichment of genes involved in histaminergic pathway, mapping within the region of these rare CNVs. Interestingly, these rare CNVs for TD overlapped with previous findings for autism spectrum disorder (Fernandez et al., 2012). The first GWAS on copy number variants in TD was performed in an ascertained sample from Latin America (210 cases and 285 controls) (Nag et al., 2013). Results showed an increase of large CNVs among cases compared to controls with four duplications and two deletions were located in the COL8A1 and NRXN1 gene, respectively. McGrath and colleagues performed a cross-disorder genome-wide CNV analysis on TD and OCD (McGrath et al., 2014). Large CNVs (>500kb) with frequency <1% were assessed in a case-control design (2,699 cases; 1,789 controls), ascertained for OCD (N=1,613) and TD (N=1,086). A 3.3-fold increase of large deletions was observed among OCD/TS cases compared to controls, mostly located in the 16p13.11 locus, a region previously linked to other neurodevelopmental disorders, with weaker evidence in the TD cases than in the OCD cases (McGrath et al., 2014). Bertelssen et al. (2014) screened 188 TD cases from a Danish cohort, and identified seven patients with intronic deletions in the IMMP2L gene, with a significant higher frequency (p=0.045)) when compared to both the Danish control population and the genotyping reference panel (Affymetrix) cohort.

SNP Heritability and Genetic Correlations

GWAS studies focus on the identification of common genetic variants. Other techniques focus on the estimation of the heritability of a trait as a result of all measured and imputed SNPs (referred to as SNP-based heritability), i.e. the amount of heritability explained by the joint effect of common variants across the genome. This method – often referred to as Genome-Wide Complex Trait Analysis (GCTA) - calculates the genetic similarity between individuals based on the measured SNPs, and estimates how much of this similarity explains their phenotypic resemblance. In the bivariate situation, genetic relatedness can be used to explain the genetic correlation between the phenotypes. In an attempt to assess the degree of shared heritability between OCD and TD (Davis et al., 2013), Davis and colleagues performed GCTA with data were collected in the scope of the consortia mentioned above—TSAICG and IOCFGC. In this study SNP heritabilities for OCD and TD were reported at 0.37 and 0.58, respectively. Interestingly, it was found that

for TD 21% of the heritability was explained by rare SNPs (with a frequency in the population of less than .05), whereas for OCD the entire heritability was attributable to common SNPs. The results of the bivariate analysis combining both datasets (OCD and TD) revealed a genetic correlation of 0.41 (se=0.15, p=0.002). This reflects the genetic overlap, i.e., the degree of shared heritability between these two disorders. Following on their first collaborative effort on GWAS in TD and OCD (Scharf et al., 2013; Stewart et al., 2013), a PRS analysis was carried out in the same groups. The combined GWAS analysis, as described above, sought to unravel functional variants shared between these two disorders (Yu et al., 2014). The aggregated risk score in the discovery sample was then evaluated for predicting the disease status on the target sample. This analysis failed to detect a polygenic signal, and the overall conclusion from this work pointed towards specific, rather than shared, genetic effects underlying the etiology of OCD and TD.

Conclusions and Directions for Future Research

Total Heritability estimates of OCD, TD/tic and HD disorders range between .27-.65 for OCD, between .25-.77 for TD and between 0.35-0.50 for HD. With respect to genetic factors underlying the etiology for these disorders, we expect the genetic liability to be highly polygenic with small effect sizes for the individual SNPs. With the increase in sample size and the continued collaborations in the field of psychiatric genetics, we also expect that new genetic variants will soon be uncovered.

An open issue is the discrepancy between heritability estimates as found in epidemiological twin studies (between 0.25-0.37) versus clinical TD samples/ TD families using either SEM or SNP-based methods (between 0.25-0.56). Explanations might be 1) a larger amount of measurement error in the epidemiological samples (as a result of retrospective self-report measures used on lifetime tic occurrence). Generally, measurement error adds to inflated estimates of unique environmental influence, and lower heritability estimates. Alternatively, 2) the discrepancy between clinical and epidemiological samples might reflect phenotypic differences. TD family studies might have an over-representation of participants with persistent TD, which might be associated with increased genetic load. By contrast, the epidemiological studies also included subjects with childhood-only tic disorders and may be genetically more heterogeneous, or may represent a group in which environmental factors play a larger role.

To conclude, we have witnessed a remarkable progress in our conceptualization of TD, OCD and HD in the recent decades. By 1986 TD was conceptualized

as an autosomal dominant genetic disorder (Pauls and Leckman, 1986), with only a few genes with a large effect contributing to its expression. Moreover, within TD families co-morbid OCD was then considered as an alternate expression of the same phenotype (Pauls and Leckman, 1986). It is largely consensual now that TD is highly heterogeneous, mediated by large numbers of SNPs of small effect, with influences of CNVs and rare variants, and that TD and OCD only partly share genetic architecture. HD was diagnosed as mental disorder not until as recent as 2013. Current technology on genetic research offers interesting perspectives for new discoveries and understandings of these three disorders in the near future. First, a number of international consortia and collaborative efforts aim to standardize phenotyping and diagnostic criteria. This will allow researchers to increase sample sizes, and combine information from different populations worldwide while properly accounting for population stratification and cultural differences.

Second, we expect that OCD, TD and HD will undergo a similar scenario to that of other psychiatric disorders such as schizophrenia, for which, within less than a decade after the first GWAS published, more than 100 loci were identified as a result of huge sample sizes (n= 37,000 cases and 113,000 controls) in collaborative GWAS meta-analytic efforts (Ripke et al., 2014). Shortly following this report, the involvement of the Major histocompatibility complex (MHC) locus was confirmed, specifically the role of complement component 4 (C4) genes in the reduced number of synapses in the brains of schizophrenic individuals (Sekar et al., 2016). These genes regulate the expression of the human C4 protein, which has a central role in the Classical Complement pathway, part of the innate immune system. This seems to confirm hypotheses that SCZ may in part be mediated by an affected immune system. In the near future, largely increased sample sizes due to worldwide consortia might equally yield results with respect to gene pathways involved in other disorders.

Third, the fast development of other genetic techniques in targeting other genetic variation than SNPs, including CNVs, de novo mutations or rare variants and micro RNAs supplemented by pathway and network approaches suggests that new insightful results will soon be generated. There is (overall) consensus that the emerged paradigm of 'missing heritability' should be interpreted as 'hidden' rather than 'missing'. Part of this (so far) un-captured genetic variation may indeed be obscured in other forms of genetic variation than the currently measured SNPs.

Lastly, it is expected that similar developments in other fields such as epigenetics, animal models, systems biology (proteomics, metabolomics and

genomics) and imaging genetics lead to new insights. In sum, there are many scientific roads that can be followed in understanding how OCD, TD and HD develop, how genes express themselves to cause liability for these diseases, and possibly provide new drug targets or treatment options.

Aims and Outline

From their early conceptions in the nineteenth century until modern times, the field of Psychiatric Genetics has undergone substantial transformation. Sir Francis Galton, widely regarded as the founder of twin studies and behavioral genetics, was the first to study variation in human populations, publishing his highly acclaimed essay 'Hereditary Genius' in 1869, ten years following Darwin's revolutionary work 'The Origin of Species' (Darwin, 1859; Galton, 1869). Along the decades, advances in this field have come from renowned figures in science, namely Gregor Mendel (postmortem), Ronald Fisher and Sigmund Freud. Equally, major scientific hallmarks have impacted the field of (psychiatric) genetics. The most notable developments including the statistical methodology underlying modern Twin studies and the later emergence of Quantitative genetics (Martin & Eaves, 1977; Rao & Morton, 1974; Wright, 1949).

Fast-forwarding to today, we are now witnessing a formidable age in genetics. In the past two decades, the advent of large-scale high-throughput DNA technologies have allowed scientists to scan through millions of genetic variants in thousands of genomes, providing us with unique insights into the basis of individual differences in behavior. It is at this juncture that this dissertation stands. From Structural Equation Modeling and Twin studies, to molecular and statistical genetics, the field of Psychiatric Genetics (and more broadly Behavioral Genetics) brings together these complementary but separate disciplines with the goal of understanding the nature and causes (etiology) of psychiatric conditions.

The work within this dissertation focuses on obsessive-compulsive disorder (OCD), Tourette Disorder (TD) and tic disorders, and hoarding disorder (HD), conditions at the core of obsessive-compulsive behavior. The three main objectives are 1) to identify the relative contributions of genetic and environmental factors underlying these phenotypes; 2) to quantify the degree to which these factors are shared across these phenotypes and 3) characterize the genetic (and epigenetic) architecture underlying the etiology of these phenotypes.

This dissertation is organized as follows:

Chapters 2-4 provide a series of genetic epidemiological analyses on these phenotypes based on the classical twin design. The twin method allows us to disentangle the relative contributions genetics and environmental factors to the variability of a given phenotype, by comparing the relatedness across

different family members. The Netherlands Twin Register (NTR) contains a large population-based twin-family sample, providing a unique opportunity for twin studies. In particular, Chapter 2 presents longitudinal genetic epidemiological analyses on the stability in time of Obsessive-compulsive symptoms, and the relative contribution of environmental or genetic factors to this stability. To this end, we had at our disposal self-report longitudinal data collected at the NTR at two different time points corresponding to the 6th (2006) and 8th (2008) waves of collection. Chapter 3 presents a comprehensive analysis on the heritabilities of four different tic phenotypes. Here, the most reliable phenotype is described considering the most reliable heritability estimate. Four different tic phenotypes (according to DSM IV) were analysed with respect to the contribution of genes, and environmental factors, and their impact on future genetic studies is discussed. Chapter 4 explores the (co)-variation of OCD, TD/tic disorders and HD, regarding their shared underlying genetic and environmental factors.

Chapters 5-8 make use of large SNP-array datasets made available from the field of genetic association analysis. We explore the genetic variance captured within these SNP-arrays, in search for variants associated with these disorders. Specifically, Chapter 5 describes a series of exploratory genetic analyses into the genetic basis of OC symptoms, and the informative value of a population-based twin-family sample in the context of GWASs. In Chapter 6, a GWAS on TD within the NTR is described, and the results of a meta-analysis with the results obtained in a GWAS performed by TSAIGC – an international consortium for the study of TD, combining a population-based cohort from the NTR with a clinically-based sample. Chapter 7 presents the results of a genetic meta-analysis for HD, the first ever published, and only the second ever performed GWAS on HD. Chapter 8 presents the results of a polygenic risk score analysis of OC symptoms.

Finally, in chapter 9 an epigenome-wide association study on tic disorders is presented.

In conclusion, chapter 10 summarizes and discusses these findings in concert. The work ends with future perspectives within the field.

.

Chapter 2
Genetic and Environmental Contributions to Stability in Adult Obsessive Compulsive Behaviour

This chapter is based on: Zilhao, N. R., Smit, D. J. A, den Braber, A., Dolan, C. V, Willemsen, G., Boomsma, D. I., & Cath, D. C. (2014). Genetic and Environmental Contributions to Stability in Adult Obsessive Compulsive Behavior. Twin Research and Human Genetics, 18(1), 52–60.

Abstract

This study investigates the relative contribution of genetic and environmental factors to the stability of obsessive-compulsive (OC) symptoms in an adult population-based sample. We collected data in twin pairs and their siblings using the Padua Inventory Revised Abbreviated, from the population-based Netherlands Twin Register in 2002 (n=10.134) and 2008 (n=15.720). Multivariate twin analyses were used to estimate the stability of OC symptoms as a function of genetic and environmental components. OC symptoms were found to be highly stable, with a longitudinal phenotypic correlation of .63. Longitudinal broad sense heritability was found to be 56.0%. Longitudinal correlations for genetic (r=.58 for additive, r=1 for non-additive genetic factors) and non-shared environment (r=.46) reflected stable effects, indicating that both genes and environment are influencing the stability of OC symptoms in adults. For the first time evidence is reported for non-additive genetic effects on the stability of OC symptoms. In conclusion, this study showed that OC symptoms are highly stable across time in adults, and that genetic effects contribute mostly to this stability, both in an additive and non-additive way, besides non-shared environmental factors. These data are informative with respect to adult sample selection for future genetic studies, and suggest that gene-gene interaction studies are needed to further understand the dominance effect found in this study.

Introduction

Obsessive Compulsive Disorder (OCD) is characterized by intrusive, unwanted thoughts, and repetitive behaviours performed in a ritualized fashion (APA, 1994). It has a lifetime prevalence of 0.5 – 2.0 % (APA, 1994), and it is, among all anxiety disorders, recognized by the World Health Organization (WHO) as a leading cause for non-fatal illness-related disability, affecting mostly individuals between 15 and 44 years of age (WHO, 2007). Quality of life is seriously impaired in OCD, more so than for instance in depression (Srivastava et al., 2011).Various longitudinal clinical studies have established that, in contrast to children/ adolescents in whom OC symptoms seem to remit somewhat more often (Fernández de la Cruz et al., 2013; Leckman et al., 2009; Micali et al., 2010; Stewart et al., 2004), in adults OC symptoms tend to be more stable with respect to symptom dimensions, with probability estimates of full remission between 17% and 27%, and of partial remission of between 22% and 53% in course up to 40 years, depending on study methodology (prospective versus retrospective), and country of origin (Alonso et al., 2001; Eisen et al., 2013; Mancebo et al., 2014; Orloff et al., 1994; Reddy et al., 2005; Skoog & Skoog, 1999; Steketee et al., 1999). Specifically in patients who only experience partial treatment response, symptoms recur in up to 70% within 5 years of follow-up (Eisen et al., 2013). Longitudinal studies on OC symptoms are far more scarce in epidemiological than in clinical cohorts, both in children and in adults (Angst et al., 2004; Fineberg et al., 2013). Only two studies examined OCD and OC symptoms in a community cohort. The Zurich community cohort study followed a group of adolescents for 30 years (Angst et al., 2004; Fineberg et al., 2013). This study indicated that over these 30 years, at age 50 in more than one third of the sample OC symptoms had not remitted. Fullana et al. (2007) examined the temporal stability of obsessive-compulsive dimensions over a period of 2 years in undergraduate students, and found no significant changes in OC symptom scores between baseline and follow-up, except for the occurrence of obsessions (Fullana et al., 2007). These data provide a longitudinal perspective on OCD and OC symptoms in adults, but do not address the etiology for the observed stability, of which still little is known.

Several family-based studies have indicated a familial basis for OCD, with increased frequencies of OC symptoms in first degree relatives (Rosario-Campos et al., 2005; Pauls et al., 1995). Population based twin family studies have shown that variation in OC symptoms is heritable (Rosario-Campos et al., 2005; Hudziak et al., 2004; Iervolino et al., 2011; Pauls et al., 1995) with somewhat lower heritabilities in adults (h2 ranging between 0.30-0.40) than in children (h2 between 0.45-0.58 for 12 year old twins; (Hudziak et al., 2004),

and 55% for 6 year old twins (Bolton et al., 2009), and little support for common environment and non-additive genetic effects.

Some recent studies have extended these findings by introducing a longitudinal twin design. A first longitudinal study at the Dutch Twin register (van Grootheest et al., 2007) examined the stability of OC symptoms in twins aged 7, 10 and 12 years old. The twins were measured on OC symptoms using the Obsessive Compulsive Scale of the Child Behavior Checklist (CBCL-OCS). The twin design allows stability to be attributed to either genetic or environmental factors. Van Grootheest et al. (2007) reported an average phenotypic longitudinal correlation of 0.5 across a five-year period in children between ages 7, 10 and 12. Genetic factors explained a substantial part (between 35% (father ratings) and 51% (mother ratings)) of this OC symptom stability. Fully in line with these findings, Bolhuis et al. (2014) used OC symptom data from a longitudinal cohort of adolescent twins and siblings (n=2651; the Genesis 12-19 study; mean age 15yrs), and a cross-sectional sample of adult twins (n=4920; mean age=55years) to explore genetic and environmental relationships between OC and depressive symptoms both cross-sectionally and longitudinally (Bolhuis et al., 2014). Within the adolescent sample, covariance (ß) between OC symptoms at timepoints 1 and 2 (with mean 25 months interval between measurements) was 0.48 (CI between 0.42-0.56), indicating substantial stability over time.

A previous genetic epidemiological study has been conducted in adult twins within the Dutch Twin Register on OC symptoms (van Grootheest et al., 2009). Twins (van Grootheest et al., 2009) (average age at baseline: 17.8 years) were measured at four time-points between 1991 and 2002, using two different measurement scales—the YASR-OCS (Nelson et al., 2001) at the first 3 timepoints and the Padua Inventory Revised Abbreviated (PI-R ABBR; (Cath et al., 2008) at timepoint 4. The correlation across time ranged between 0.39 and 0.60 between sequential measurement occasions, and the longitudinal heritability was calculated to be around 40%. However, the stability measures of OC symptoms and the etiology of that variance may have been obscured by the use of different measurement instruments at different time-points. It is reasonable to expect that different scales tag different sources of variation, thus leading to an underestimation of stability and the proportion of that estimation to be attributed to genetic factors.

Therefore, the aim of the present study was to extend the findings of the van Grootheest et al. (2009) study, overcoming the methodological weakness of the previous study (different instruments used at different time points to measure OC symptoms), moreover including siblings in the analyses, to

explore OC symptoms over a period of 6 years, and to determine the genetic and environmental contributions to stability of OC symptoms in a large sample of adult twins and their siblings. OC symptoms were measured in 2002 and in 2008, at a subsequent wave of collection within the Netherlands Twin Register (NTR) with a large proportion of the subjects overlapping with the 2002 wave.

Materials & Methods

Participants and Procedure

The data for this study were collected as part of longitudinal survey studies of the NTR. Since 1991, every two to three years, twins and their families are assessed and receive surveys by mail, with questionnaires about health, personality and lifestyle. For the present analysis, we analysed the Obsessive Compulsive (OC) data collected at the 2002 and 2008 wave of collection, corresponding to survey 6 and 8. The total sample consisted of 20.376 individuals from 7812 different families. To survey 6, 10.134 individuals responded, and 15.720 responded to survey 8. Longitudinal data were available for 5478 individuals. Twin participants with incomplete data on zygosity (N=41 for complete pairs, and N=170 for incomplete pairs) were excluded from the analysis. This study has been approved by the Medical Ethical Committee of the VU Medical Centre Amsterdam. In the genetic analysis, we included a maximum of four siblings (two brothers and two sisters) per family. Table 1 gives an overview of the total number of participants. The number of complete twin pairs included in the study by zygosity, is given in Table 2. Zygosity was assessed by a questionnaire using items on physical similarity, blood group and DNA polymorphisms (Willemsen et al., 2013).

Table 1. Number of participants included in genetic analysis.

	Survey 6			Survey 8		
	Male	Female	Male + Female	Male	Female	Male + Female
Twins						
N	1398	2987	4385	2286	5163	7449
age	32.27	32.46		33.65	34.05	
age SD	11.48	11.15		14.92	14.51	
Siblings						
N	552	843	1395	629	1069	1698
age	35.01	34.62		38.97	37.77	
age SD	13.06	11.13		14.72	13.59	

Table 2. Number of twins by zygosity.

Twin zygosity	**Survey 6** N twins (N complete pairs)	**Survey 8** N twins (N complete pairs)	Total number of unique Twins
Monozygotic Males (MZM)	635 (231)	1017 (354)	1306
Dizygotic Males (DZM)	360 (102)	623 (182)	800
Monozygotic Females (MZF)	1588 (625)	2714 (1047)	3260
Dizygotic Females (DZF)	847 (298)	1412 (458)	1763
Dizygotic Opposite Sex (DOS)	955 (287)	1683 (418)	2151

Phenotype Measures

For both survey 6 and survey 8, OC symptom scores were measured with the Padua Inventory abbreviated (PI-R ABBR) (Cath et al., 2008). This 12 item questionnaire has been derived from the Padua Inventory-Revised (PI-R), a widely used 41 item self-report inventory on obsessive-compulsive symptoms with item ratings between 0-4, and 5 subscales (washing, checking, rumination, precision and impulses) (Sanavio, 1988; van Oppen, 1992). The 12 item PI-R ABBR contains 2 items on each subscale, items have been chosen based on the highest factor loadings in a previous validation study (Van Oppen et al., 1995), and one item has been added to the rumination and impulses subscales.

Genetic Modelling and Testing

Genetic epidemiological twin studies are based on comparing the different degrees of family relatedness between family members, to estimate the relative contribution of the genetic and the environmental components to a trait. Monozygotic (MZ) twins share (nearly) all their genes, whereas Dizygotic (DZ) twins, just like non-twin siblings, share on average half of their segregating genes. In quantitative genetics, the total phenotypic variance is decomposed into variance components due to genetic (G), shared environmental (C) and non-shared environmental (E) factors. C reflects the common environmental effects shared by family members and E the non-shared environmental

influences i.e. the unique environmental component for an individual. The genetic variance can be additive (A), reflecting the additive effect of different alleles, or non-additive (dominance; D), indicating interaction between alleles. The comparison between MZ and DZ twin and sibling correlations provides a first impression to the relative contributions of each component. The greater the phenotypic similarity between MZ twins when compared to DZ twins and non-twin siblings, the more of the variance of the trait is explained by genetic factors. An MZ correlation that is the double of DZ correlation is indicative of additive genetic influences. DZ correlations higher than half of the MZ correlations indicate the role of shared environmental factors, while DZ correlations that are less than half the MZ correlations indicate genetic non-additive effects (dominance). All genetic analyses were carried out with the use of Structural Equation Modelling as implemented in OpenMx (Boker et al., 2011). A saturated model (with input of a 6 x 6 data matrix of 2 twins, 2 brothers, and 2 sisters for questionnaire data) was used to estimate familial correlations, to test for sex differences and for the effect of age on the OC symptom scores. Next, to evaluate the stability of the PI-R ABBR scores between the two time points, a bivariate saturated model (with input of a 12 x 12 data matrix of 2 twins, 2 brothers and 2 sisters, each with two time-point measures), was fitted. These models estimate variances and co-variances within and between MZ and DZ twins and between twins and siblings, and specify all correlations between family members. Next, a genetic model (ADE model) was fitted to the data to estimate the relative proportion of A, D and E in OC symptom scores, as the twin correlations indicated an ADE, rather than an ACE model. The comparison between different models (e.g. ADE versus AE) was done by means of likelihood-ratio tests. The negative log likelihood (-2LL) of a more restricted model is subtracted from the -2LL of a more general model. This generates a statistic that is distributed as a X^2 distribution with degrees of freedom (df) equal to the difference in the number of estimated parameters in the two models. The more parsimonious constrained model is selected if the X^2 yields a nonsignificant p-value. An alpha of 0.01 was set as threshold for significance.

Results

Longitudinal data were available for 5478 individuals (2048 males, 3430 females). Their average age was 33.0 years (SD = 11.5) at survey 6, and 34.7 years (SD = 14.6) at survey 8. The correlation between the two surveys was 0.63. The effect of age on OC symptom scores was a drop of 0.03 per year for both Survey 6 and 8.

Since the distribution of OC symptom scores was skewed (skewness =1.10 in survey 6, and skewness =1.22 in survey 8 (Figure 1)), scores were transformed using square root transformation. The graph in Figure 2 plots the distribution of OC symptom scores, after transformation of the data (skewness= -0.315 in survey 6, and skewness= -0.254 in survey 8).

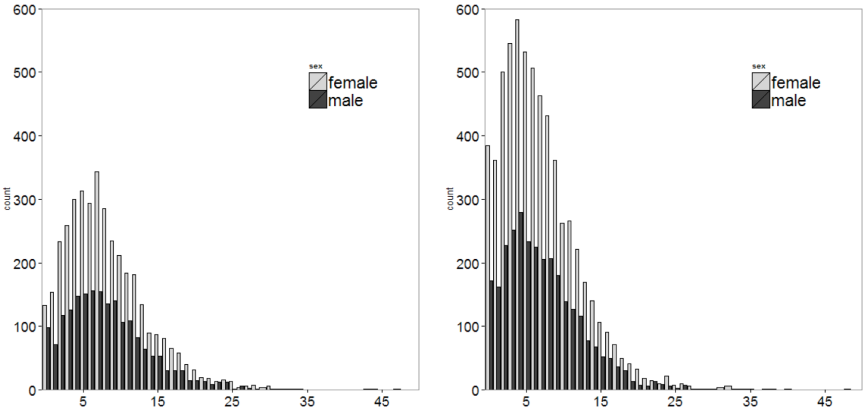


Figure 1. Distribution of OC symptom scores, before transformation of the data, in both survey 6 (left), and 8 (right).

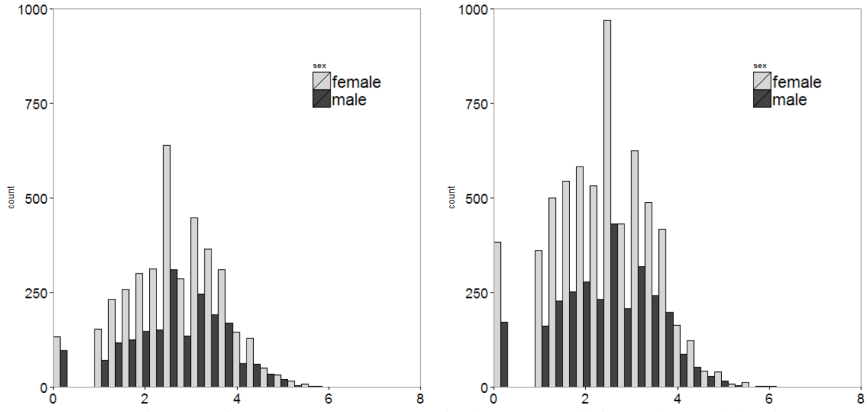


Figure 2. Distribution of OC symptom scores in both survey 6 (left), and 8 (right), after square-root transformation of the data.

The estimates for familial correlations are given in Tables 3 and 4, for surveys 6 and 8 respectively. These are given for MZ and DZ twins, sib-sib and twin-sib pairs, and conditional on sex. Monozygotic males (MZM) and monozygotic females (MZF) correlations are moderate, and both sib-sib and sib-twin correlations, at survey 8, are smaller than half the MZ correlations. The observed patterns in familial correlations, when comparing MZ correlations with all other first-degree relatives, suggest that genetic factors may play a role in individual differences for OC symptoms. We observed a high stability for OC symptoms across time with correlations of 0.629 for within individual measurements. Cross Twin-Cross Time correlations (Table 5) represent the correlations between the scores in one twin at one time-point, and its co-twin at another time-point. First and second-born twin correlations are constrained to be equal to second-born and first-born correlations. Cross Twin-Cross Time correlations were estimated for males and females separately, being slightly higher in males (MZ CTCT: Rmale = 0.400, Rfemale = 0.359; DZ/Sibling CTCT: Rmale = 0.130, Rfemale = 0.084). When estimated and compared between MZ and DZ twin pairs, the results suggest that the stability in OC symptom scores is predominantly due to genetic factors.

Table 3. Familial correlations estimated from Maximum likelihood, in survey 6.

Type of relation	*r*	*95% CI*	*n* pairs
MZ twins	**0.398**	**0.371 – 0.471**	856
MZ males	0.414	0.305 – 0.507	231
MZ females	0.400	0.335 – 0.460	625
Male relatives	**0.227**	**0.125 - 0.321**	**446**
DZ male	0.253	0.084 – 0.400	102
Brother - male twin	0.196	0.062 - 0.316	292
Brother - brother	0.296	-0.014 – 0.524	52
Female relatives	**0.150**	**0.079 - 0.220**	**1135**
DZ female	0.207	0.084 - 0.318	298
Sister - female twin	0.134	0.0176 - 0.191	732
Sister - sister	0.058	-0.121 - 0.231	105
Female -Male	**0.106**	**0.038 – 0.171**	**456**
DOS	0.228	0.099 – 0.342	287
Sister- brother	-0.025	-0.213 – 0.168	169

Table 4. Familial correlations estimated from Maximum likelihood, in survey 8.

Type of relation	*r*	*95% CI*	*n* pairs
MZ twins	**0.396**	**0.353 – 0.436**	**1385**
MZ males	0.385	0.301 – 0.460	354
MZ females	0.407	0.356 - 0.455	1047
Male relatives	**0.124**	**0.027 – 0.219**	**491**
DZ male	0.209	0.060 – 0.337	182
Brother - male twin	0.041	-0.092 – 0.172	256
Brother - brother	0.278	-0.201 - 0.564	53
Female relatives	**0.106**	**0.048 – 0.164**	**1337**
DZ female	0.114	0.026 – 0.198	458
Sister - female twin	0.091	0.012 - 0.167	777
Sister - sister	0.144	-0.080 - 0.345	102
Female -Male	0.168	**0.110 – 0.222**	**587**
DOS	0.222	0.125 – 0.311	418
Sister- brother	-0.013	-0.219 – 0.199	169

Table 5. Cross Twin - Cross Time correlations.

Type of relation	*r*	*95% CI*
MZ twins	**0.367**	**0.327 – 0.406**
MZ males	0.400	0.324 – 0.468
MZ females	0.359	0.308 – 0.400
DZ twins/siblings	**0.119**	**0.867 – 0.151**
DZ male/brothers	0.130	0.090 – 0.169
DZ female/sisters	0.084	0.036 – 0.131
DOS/sister-brothers	0.157	0.083 – 0.228

Model fitting analyses are displayed in Table S1 (Supplementary Table S1) for surveys 6 and 8. There were no differences in variances across zygosity or between all pairs, and no differences in variances and correlations between the sexes. In both surveys, for twin-sib pairs, correlations could be constrained to be equal; equality assumptions were tested separately for males (Table S1, model 3 compared to model 2), for females (Table S1, model 4 compared to model 3), and for opposite-sex relatives (Table S1 , model 5 compared to model 4). Gender differences were also tested for equality and constrained to be equal (Table S1, model 6 compared to model 5). The correlation between MZM and MZF could also be constrained to be equal, estimated at 0.398 (95% CI = 0.371 - 0.471) for survey 6, and 0.399 (95% CI = 0.356 – 0.440), for survey 8 (Table S1, model 7 compared to model 6).

Working from a fully saturated model, considering all results from model comparisons and the correlations between MZ and DZ twins, we decided to

next fit a genetic ADE model to the data. The relative contributions D could not be constrained to zero (X^2=11.464, df=3, p=0.009), and consequently the best fitting model was the ADE. Figure 3 shows the unstandardized estimates of the ADE model. This allows to compute the relative contributions of both A, D and E for each time-point, given on the diagonal of Table 6.

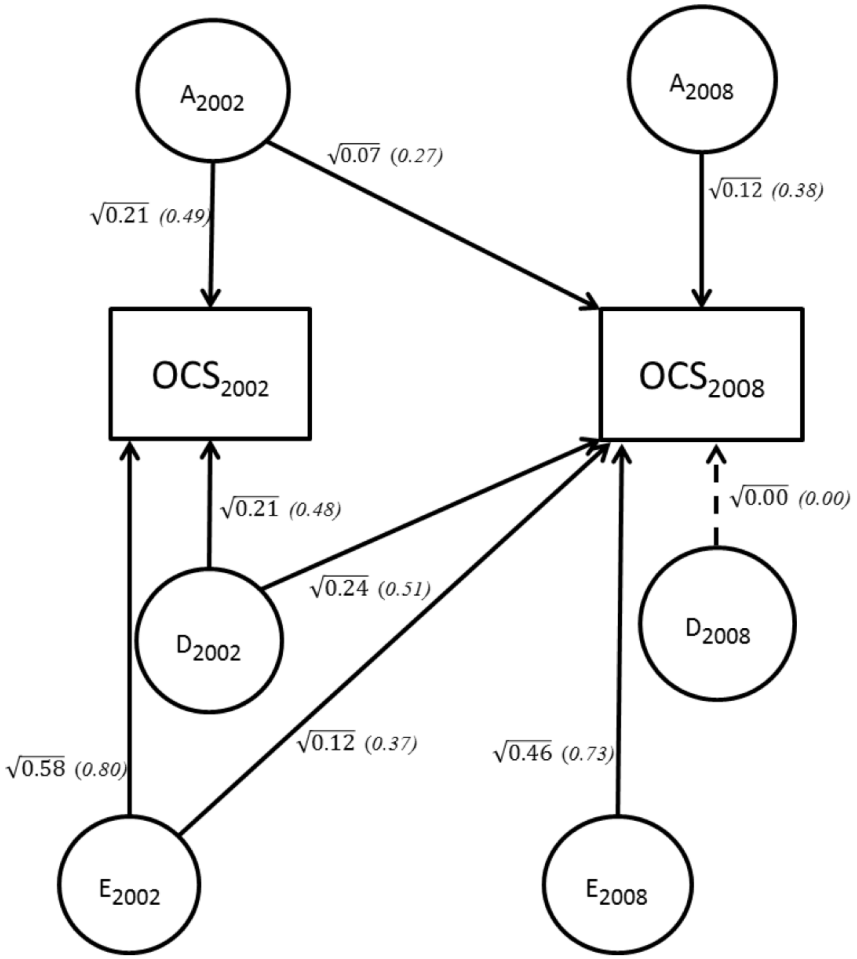


Figure 3. Standardized and unstandardized (italics) estimates of the final ADE model for OC symptoms at the two time-points. Rectangles represent observed variables at the two time-points, OCS 2002 (survey 6) and OCS 2008 (survey 8). Circles represent the latent factors A (additive genetic influences), D (dominant genetic influences) and E (non-shared environmental influences). The dashed line represents a non-significant path. The values represent the loadings of each observed variable in the latent factors.

Table 6. Relative contributions of additive genetic and non-shared environmental influences within time (diagonal) and across time (off-diagonal) for PI-R ABBR.

	Additive genetic effects (A)		Non-additive genetic effects (D)		Non-shared environmental effects (E)	
	Survey 6	Survey 8	Survey 6	Survey 8	Survey 6	Survey 8
Survey 6	0.212	-	0.208	-	0.580	-
Survey 8	0.194	0.187	0.365	0.231	0.441	0.582

The results show that, for both time-points, the contributions of A and D to OC symptoms are around 0.2. Because the contribution of D could not be dropped from the model, this results in a broad sense heritability estimate (both A and D component) of 0.420 and 0.418 for survey 6 and survey 8, respectively. For non-shared environmental influences this was around 0.58 for both time-points. The off-diagonal estimated in Table 6 gives the results of the decomposition of the phenotypic stability between the two-time points. The result show that 56% of the stability for OC symptoms is due to genetic factors (both A and D component). Non-shared environmental factors had a moderate contribution to the stability of 44%. Finally, longitudinal correlations were calculated, indicating the relative overlap of genetic and environmental effects between the two time-points. The additive genetic correlation was estimated at .58 from the ADE model. This value indicates a moderately high overlap for the genetic influences between both time-points. For the D component a correlation of 1 was obtained, indicating a perfect overlap for non-additive effects between the two time-points, and that no new 'D' is involved at a later stage. For non-shared environmental factors the correlation was 46.

Discussion

To our knowledge, this is the first genetic epidemiological study in twin adults assessing the longitudinal genetic and environmental contributions to the stability of OC symptoms using the PI-R ABBR questionnaire at multiple time points. First, in slight contrast to the previous longitudinal study in children (van Grootheest et al., 2007), we found no quantitative sex differences in average OC scores or heritability estimates in this adult sample, with the same additive genetic factors influencing both males and females. In line with previous cross-sectional studies in adolescents at different time-points and in adults using the YASR-OCS (van Grootheest et al., 2007) and 2008, no evidence for a special twin environment was found, since correlations could

be equated across zygosities and between twins and siblings. Heritability estimates at each time-point were also in line with what is described in the literature. The values for the broad sense heritability (both A and D component), at both time points, were at around 42%, and the remaining variance was due to unique environment. In a recent paper, addressing the shared genetic and environmental contributions to both OC symptoms and hoarding, with the same adult twin data from NTR 2008 wave of collection as used in these analyses, the authors found heritabilities of 40% for OC symptoms (Mathews et al., 2014). In an earlier report on the 2002 wave of data collection, Grootheest et. al found heritability rates of 38% and 44%, for males and females respectively, (van Grootheest et al., 2009). All these studies in adults found no contribution of shared environment. Only in one study in children, a small contribution of shared environment has been found at age 12 (van Grootheest et al., 2008), suggesting differences in environmental architecture underlying stability of OC symptoms in children versus adults.

We found the OC symptoms to be highly stable with a longitudinal phenotypic correlation of 0.63 across a 6 year time interval. Van Grootheest et. al (van Grootheest et al., 2009), using the same 2002 wave of data collection with the Padua-R-ABBR in 2002 as an end point, and the 1991, 1995 and 1997 data waves that used the Young Adult Self Report Obsessive Compulsive Scale (YASR-OCS) as starting points, found longitudinal phenotypic correlations of around 0.2 (for the 11-year time interval) and 0.4 (for a 5 year time interval). These different results could be explained by the fact that in that study the stability of OC symptoms was calculated across different measurement scales. Furthermore, the patterns in Cross Twin-Cross Time correlations (r=0.367 for MZ twins, r=0.119 for DZ twins/siblings) show that the stability is due to both genetic and non-shared environmental factors (Table 6). This suggests that the same genes are expressed across time, and are influencing OC symptoms and also, that a substantial amount of E reflects stable effects. Collectively, these set of results show that in adults, data from men and women, as well as twins and non-twins of different ages, can be combined, which is particularly beneficial in molecular genetic studies, where combining data will result in an increase in power to detect underlying genetic effects.

Of interest are the results regarding the decomposition of the longitudinal phenotypic variance. Previously, it was reported that for children common environmental influences explain part of the stability (around 40%) (van Grootheest et al., 2007), unlike what happens in adult family members, who generally do not share the same household any more with their respective

co-twin. Instead, it was shown that in adults about 70% of the stability was due to additive genetic factors (van Grootheest et al., 2009). Here, we found no common environmental influences for the stability of OC symptoms but we have, however, presented for the first time new evidence for contributions of unique environmental influences in adults, and these influences correlate substantially across time (r=0.46, Table 11). This relatively high contribution of non-shared environment to the stability of OC symptoms (off-diagonals in Table 6), indicates that unstable variance such as measurement error (time-point specific variance) cannot account for its variation. Therefore, individual experiences may have a relevant impact on the stability of OC symptom in adults. The detrimental influence of a early-life experiences might persist and influence OC symptomatology into adulthood. One study specifically addressed the influence of unique and shared environmental factors in developing (or protecting against) OC symptoms, by comparing scores between highly concordant and highly discordant MZ twins (Cath et al., 2008). The comparison of highly within-discordant pairs indicated some important influencing life-events, among which the most relevant appeared to be past experiences of sexual assault. These possible risk factors were more highly associated with OC symptoms in high-scoring discordant MZ pairs than in high-scoring concordant MZ pairs, thus pointing to the relevance of individual experiences in childhood in developing OC symptomatology. It appears then, that, once OC symptoms are acquired in early adulthood, stability tends to be higher and more genetically mediated than in children.

For the first time, we observed significant non-additive genetic effects. Dominant effects have not been observed before for OC symptoms, even in the same subject sample. Non-additive variance encompasses all forms of genetic factors with a non-linear effect such epistasis/gene-gene interaction and dominant effects (Mather, 1974). This may have unexpected consequences for molecular genetics approaches, because most genome-wide technologies such as Genome Wide Association Analysis (GWAS) and Genome-wide complex trait analysis (GCTA) (Yang et al., 2011), primarily assume additive genetic variance. These studies may have been underpowered by not including dominance in the linear regression models. Random effects modelling using GCTA to explain phenotypic variance (Yang et al., 2010) may not be able to provide SNP based heritability estimates for all genetic effects and lead to "missing heritability", since modelling non-additive effects requires very large sample sizes (Yang et al., 2013).

In sum, although our results for the broad sense heritability estimates (all genetic effects involved) are in line with what has been described before for

narrow sense heritability (Davis et al., 2013; van Grootheest et al., 2009; van Grootheest et al., 2007), they do indicate that the genetic etiology of OC symptoms and their stability in time may be more complex than previously thought, and that some differences occur with respect to its underlying etiologies between children and adults. Molecular and genome-wide studies as well as twin studies could, in the future, include dominance effects in the linear regression, and/or take epistatic effects/gene-gene interactions into account.

The Supplementary material can be found in the online version of this manuscript

Chapter 3
Heritability of Tic Disorders: A Twin-Family Study

This chapter is based on: Zilhão, N. R., Olthof, M. C., Smit, D. J. A., Cath, D. C., Ligthart, L., Mathews, C. A., …, Dolan, C. V. (2016). Heritability of tic disorders: a twin-family study. Psychological Medicine 47(6), 1085–1096.

Abstract

Genetic-epidemiological studies that estimate the contributions of genetic factors to variation in tic symptoms are scarce. We estimated the extent to which genetic and environmental influences contribute to tics, employing various phenotypic definitions ranging between mild and severe symptomatology, in a large population-based adult twin-family sample. In an extended twin-family design, we analysed lifetime tic data reported by adult mono- and dizygotic twins (n= 8,323) and their family members (n=7,164; parents and siblings) from 7,311 families in the Netherlands Twin Register (NTR). We measured tics by the abbreviated version of the Schedule for Tourette and Other Behavioral Syndromes (STOBS) (TSAICG, 2007). Heritability was estimated by genetic Structural Equation Modelling (SEM) for four tic disorder definitions: three dichotomous and one trichotomous phenotype, characterized by increasingly strictly defined criteria. Prevalence rates of the different tic disorders in our sample varied between 0.3 and 4.5% depending on tic disorder definition. Tic frequencies decreased with increasing age. Heritability estimates varied between .25 and .37, depending on phenotypic definitions. None of the phenotypes showed evidence of assortative mating, effects of shared environment, or non-additive genetic effects. Heritabilities of mild and severe tic phenotypes were estimated to be moderate. Overlapping confidence intervals of the heritability estimates suggest overlapping genetic liabilities between the various tic phenotypes. The most lenient phenotype (defined only by tic characteristics, excluding criteria B, C and D of DSMIV) rendered sufficiently reliable heritability estimates. These findings have implications in phenotypic definitions for future genetic studies.

Introduction

Tics are defined as involuntary sudden, recurrent, non-rhythmic, stereotypical motor movements or vocalizations (DSM-IV-TR) varying from almost indiscernible eye-blinking to complex motor movements involving multiple muscle systems. The Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; APA 2000) distinguishes four categories of tic disorders: Tourette's Disorder, also called Tourette Syndrome (TS), Chronic Motor or Vocal Tic Disorder (CMT/CVT), transient tic disorder, and tic disorder not otherwise specified (NOS). Tic diagnosis depends on age of onset, duration, and type (motor, vocal or both). Tics typically first manifest between age 4 to 6 years, and peak in severity between 10 and 12 years (Erenberg et al. 1987). Over 70% of patients experience significant reduction in tic frequency and intensity by adulthood (Bloch & Leckman 2009; Cath et al. 2011).

Community-based studies have produced disparate TS prevalence estimates in children and adolescents – ranging from 0.5 to 38 cases per 1000 (Apter et al. 1993; Hirtz et al. 2007; Knight et al. 2012; Robertson et al. 2009; Scahill et al. 2005; Mathews et al. 2014; Miller et al. 2014). In a review study Scahill et al. (2013) concluded that the prevalence of TS in children between age 6 and 18 lies between 0.5% and 0.7%. Estimates for chronic motor tics ranged from 0.3% to 0.8% (Kurlan et al. 2001; Khalifa 2006; Scahill et al. 2006; Kraft et al. 2012; Cubo et al. 2011) in several studies in children. Male-to-female ratios varied between 3- 4 to 1, with higher prevalence rates in boys (1.06 to 4.5% in boys, and 0.25% to 1.7% in girls (Lichtenstein et al. 2010; Knight et al. 2012)).

In adults, tic prevalence rates are considerably lower, with estimates of 0.05% to 0.1% for TS, and of 0.08% to 6.7% for any tic disorders (Knight et al. 2012; Apter et al. 1993; Bar-Dayan et al. 2010; Eapen et al. 2001; Robertson et al. 1994; Schlander et al. 2011; Wenning et al. 2005; Zohar et al. 1992). A Swedish population-based twin study (n=21,911) found prevalence rates of 6.7% for having any tic, and 1.4% for having TS (not taking the DSM-IV-TR criterion of presence of at least one vocal tic into account) (Pinto et al. 2016).

The causes of individual differences in tic disorder characteristics and severity are poorly understood. Genetic and environmental factors contribute to phenotypic variation. There is some suggestion of assortative mating (i.e., spousal resemblance) for tics (Hasstedt et al. 1995; Kurlan et al. 1994) but studies have been scarce. The presence of parental data allows us to take into account assortative mating. This is important as assortative mating in the parental generation may in their offspring increase the genetic correlations between siblings including DZ twins (these are on average 0.5 under

random mating). This may bias the results obtained from the classical twin design, with underestimation of heritability and overestimation of shared environmental effects. In family studies, heritability estimates of TS and tic symptoms range from .18 to .77 (Mathews & Grados 2011; Pauls et al. 1991; de Haan et al. 2015; Mataix-Cols et al. 2015; Hirschtritt et al. 2015). Tic risk in first-degree relatives of tic sufferers is high (Mataix-Cols et al. 2015). In one small clinical twin study (Price et al. 1985) of 30 MZ and 13 DZ twin pairs concordance rates were .53 for MZ pairs and .08 for DZ pairs. When criteria were broadened to include any tic, concordance rates were 0.77 for MZ pairs and 0.23 for DZ pairs.

Three population-based twin heritability studies have been performed in children or adolescents, and one in adults. (Bolton et al. 2007; Anckarsäter et al. 2011; Lichtenstein et al. 2010; Pinto et al., 2016). The longitudinal Child and Adolescent Twin Study in Sweden (CATSS) assessed tic disorders in 17,000 twins aged 9 to 12 years. The assessment consisted of three questions on tic occurrence which the parents' twins answered during a telephone interview (Anckarsäter et al. 2011). Tics were further assessed using the 'Autism — Tics, ADHD and other Comorbidities inventory' (A-TAC; Hansson & Svanstro 1994; Larson et al. 2010). Correlations for tic disorder were .38 in MZ and .11 in DZ twins, and heritability estimates were .26 in girls and .39 in boys. The heritability estimate of a binary TS diagnosis based on these data (3.1% diagnosed as affected) equalled .56 (Lichtenstein et al. 2010). Furthermore, the Genetic and Environmental Effects on Emotion study (GEMS) estimated the heritability of tic disorders based on a binary diagnosis in 4662 4-6-year old twin pairs (Bolton et al. 2007). Mothers were interviewed in a two stage telephone screen with questions on tic occurrence in their 4-year-old twins. The high scoring sample from stage 1 was selected for stage 2 (n=854 pairs) and re-interviewed. Using a liability threshold model, the heritability estimate was .5. A Japanese twin study employed a liability threshold model to assess the heritability of mother-rated tics in a sample of 1896 twin pairs between 3 and 15 years (Ooki, 2005). The mothers rated their twins with respect to the frequency of tic behaviors. Tic heritabilities estimates were .28 (boys) and .29 (girls), with shared environmental effects explaining 41% of the variance in boys and 32% in girls. Finally, Pinto et al (2016) studied the co-variation of tics, OC symptoms and ADHD in adult twins (n=20.821). The tic heritability estimate based on liability threshold modelling was 0.33 (Pinto et al., 2016). In sum, heritability estimates from epidemiological studies vary between ~.28 and ~.56, with different tic definitions and rating methods used and most studies estimating tic heritability of a single phenotypic operationalization.

The aim of the present study is to examine the genetic and environmental contributions to tic symptoms using different DSM-IV-TR derived tic phenotypic definitions in a population-based adult twin-family sample. As different studies use different measures of tics, it is highly useful to explore the influence of these varying measurement methods on variation in tic heritability. In addition, for future GWAS or other studies using genetic variants, it seems paramount to use those phenotypic tic definitions that capture the most optimal heritability estimates; "most optimal" meaning a combination of significant nonzero heritability and narrower CIs, reflecting the largest information content.

In addition, an adult twin-family sample has the advantage that lifetime tics are taken into account, allowing tics to be included that develop in adolescence. An extended twin design was used, including twins, siblings and their parents. The presence of parental data allowed us to further study the influence of assortative mating, a topic that has been scarcely addressed in TS (Hanna, Janjua, Jankovic, 1999; Kurlan, Eapen, & Robertson, 1994; McMahon et al., 1996). This is important as assortative mating may increase the additive genetic correlation among DZ twins (i.e., .5 given random mating), which may bias the results obtained analyzing only data from twins. Specifically assortative mating results in overestimation of shared environmental effects, and underestimation of genetic effects. In addition, an extended twin design confers greater power than the classical twin design (Posthuma & Boomsma, 2000). Our aims were to 1) quantify the genetic contributions to the various tic phenotypes, using both lenient and strict phenotypic definitions of tics and tic severity; 2) explore the role of assortative mating and dominance effects; and 3) determine how much the heritability estimates vary with phenotypic definition.

Method

Participants

This study is part of an ongoing longitudinal study of twins and families registered in the Netherlands Twin Register (NTR), in which participants complete a series of questionnaires on health and behavior every two to four years. A tic questionnaire was included in the 2008 survey (see Willemsen et al. 2013 for a more detailed description of the data collection). Data from 8,323 adult twins and 7,164 family members (clustered in 7,311 families) were available. Family members included twins, and parents and siblings of twins. From each family, data from two twins, two additional siblings, and their parents, if available, were selected. Non-biological parents and non-full siblings were excluded. In cases of triplets or higher-order multiples,

the first- and second-born twins were included. In cases of more twin pairs per family, one twin pair was included. Online Supplementary table 1 gives the number of family members. Data from both twins were present in 2748 families (38%), and data from twins as well as parents were present in 804 (11%) of the families. Zygosity of same-sex twins was determined by blood type, DNA markers, or questionnaire (Rietveld et al. 2000). There were 2,714 complete twin pairs with known zygosity (98.8% of all complete twin pairs): 388 MZ and 200 DZ male pairs, 1129 MZ and 507 DZ female pairs, and 490 DZ pairs of opposite sex (DOS). The age of twins ranged from 17 to 97 years (mean=33.1, SD=14.5), and the age of siblings from 11 to 88 years (mean=37.1 and SD=13.8) and of the 5,441 parents from 37 to 94 (mean=54.9 and SD=8.6). Ethical approval for the study was obtained from the Medical Ethical Committee of the VU University Medical Centre.

Measures

Data on tics from NTR Survey 8 were collected using the abbreviated Schedule for Tourette and Other Behavioral Syndromes (STOBS-ABBR) that provides a semi-structured assessment on tics, OC, and ADHD symptoms (Pauls & Hurst, 1996). This scale has been used widely by the Tourette Syndrome Association International Consortium for Genetics (TSAICG), both as interview and as self-report measure. For the NTR 2008 survey, the STOBS was abbreviated to include 9 items on the most frequent tics occurring in clinical samples (Cath et al. 2011; Freeman et al. 2000); see online supplementary table 2 for the STOBS-ABBReviated. Participants indicated for each tic type whether they ever/never experienced it. When given the response 'ever', they indicated whether the tic had occurred 0-1 year ago, 1-5 years ago, or more than 5 years ago. Subsequently, given a positive response on tic presence, items were filled in on age at onset, duration of tics (<1 year versus >1 year), and tic frequency/severity in three additional self-report items. A paper version of the questionnaire was completed by 7,028 participants (45%), and an online version was completed by 8,459 participants.

Using the STOBS-ABBR, all participants were classified according to DSM-IV-TR criteria (APA, 2000) into the following mutually exclusive categories: probable TS, probable chronic (motor or vocal) tic disorder, probable transient tic disorder, or probable tic disorder NOS (see table 1 for a summary of tic definitions). We added the term "probable", since subjects were classified based on self-report, whereas a tic diagnosis is usually established through interview and observation by experienced clinical experts, a requirement that we were unable to fulfill in this large population-based study. The DSM-IV-TR requires an age at onset before 18 years to fulfill criteria for a tic disorder

diagnosis. However, in view of the age at onset distributions of our data (figure 1) and as used by the Tourette Syndrome Study Group (Anon 1993), we adopted an age of onset≤21 years as a requirement for the definitions of "probable TS", "probable chronic (motor or vocal) tic disorder", and "probable transient tic disorder". "Probable tic disorder NOS" did not require the age-of-onset criterion.

Table 1. DSM-IV-TR-criteria for the different tic disorders

	Motor tic(s)		Vocal tic(s)	> 4 weeks	> 1 year	Age of onset before adulthood	Many tics a day	Other requirements
Probable TS	yes >1	AND	yes	yes	yes	yes	yes	-
Probable Chronic td	yes	OR	yes	yes	yes	yes	yes	-
Probable Transient td	yes	AND/OR	yes	yes	no	yes	yes	-
Probable td-NOS	yes	AND/OR	yes	-	-	-	-	no other tic disorder

To classify as probable TS, the following was required: 1) positive responding ('ever') to at least two motor and one vocal tics; 2) age of onset≤21 years; and 3) a tic duration of ≥ one year. The same criteria were used to classify as a probable chronic tic disorder, except that either one vocal or one motor tic was required. These subjects were further subdivided based on the nature of their tics (motor / vocal). "Probable transient tic disorder" required: 1) one or more motor and/or vocal tics, 2) age at onset≤21, and 3) tic duration of <1 year. Participants who reported at least one tic, but without an age at onset≤21, and/or with a tic duration of <1 year were categorized as a probable tic disorder NOS.

For genetic modelling we classified subjects as affected or non-affected according to different inclusion criteria: 1) all subjects who scored any tic at any age of onset for any period of time included as affected ('any probable tic' - the most lenient phenotype); 2) subjects with "probable TS", "probable chronic tic disorder--motor", or "probable tic disorder--vocal" were classified as affected; 3) all subjects with probable TS and probable chronic tic--motor tics as were classified affected. One additional definition was considered with 3 categories: "no tic disorder" (unaffected), "probable tic disorder NOS" combined with "probable transient tic disorder"(affected), and "probable chronic tic disorder—motor", "probable chronic tic disorder—vocal" and "probable TS" (affected).

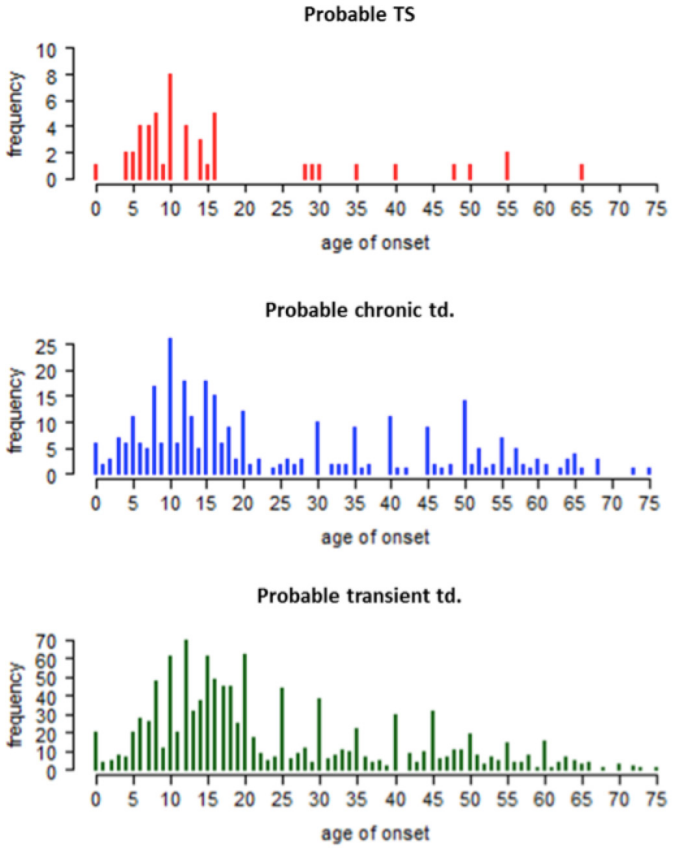


Figure 1. Reported age of onset of tics for participants that fulfil criteria of Tourette syndrome, chronic tic disorder and transient tic disorder (including the age of duration criterium but without the age of onset-criterium).

Statistical Analyses

Population prevalence of the different tic disorder definitions were estimated in the entire sample of 15,487 individuals. Fitting the genetic models and calculating correlations between family members was done by assuming that a normally distributed liability underlies the discrete phenotypes (Falconer, 1965; Falconer, 1967). In the case of a dichotomous phenotype, the threshold separates the two classes of subjects, namely the "affected" and "unaffected". For the trichotomous phenotype, two thresholds were estimated, separating three classes following the definition described above.

To assess the significance of covariates prior to the genetic modeling, we performed logistic regression; examining the covariates sex, age at filling in the

questionnaire, method of reporting (paper versus internet survey) and their interactions. To correct for family clustering, we used generalized estimation equation (GEE; Dobson & Barnett 2008) with the R-package 'GEE' and the logistic link function. For all analyses and model fitting procedures the threshold for significance was set at alpha=.05. We obtained initial estimates of familial resemblance by estimating tetrachoric and polychoric correlations between the liabilities of the family members using the R-package 'polychor' (https://cran.r-project.org/web/packages/polycor/index.html).

Genetic model fitting was conducted in R-package OpenMx version 2.2.4 (Boker et al., 2011). Parameters were estimated by raw data maximum-information likelihood. We first tested whether parent-offspring correlations were equal to DZ and sibling correlations (as all share on average 50% of their segregating genes). For the two strictest variable definitions, we encountered computational problems due to the low prevalence, giving rise to empty cells in the tables. We therefore excluded data from siblings and parents. Next, we fitted genetic variance decomposition models. These decompose variance in the liability to have tics into additive genetic (A), unique environmental (E), common environmental (C), and/or dominant genetic factors (D). Since C and D cannot be estimated together, we included C if the MZ correlation was less than twice the DZ twin correlation. If the MZ correlation was larger than twice the DZ correlation, we included D. The influence of common environmental factors and of genetic dominance was tested by comparing a nested AE model with either the ACE or the ADE model using likelihood-ratio tests. The AE models are depicted in online supplementary Figure 1 and 2.

Results

Descriptive Statistics

Prevalence rates of STOBS-ABBR tic items are summarized in online supplementary table 3. Given these symptoms we derived four probable tic disorder diagnoses in accordance with the DSM-IV-TR (i.e.: TS, Chronic motor/vocal tic disorder, transient tic disorder and tic disorder NOS). Prevalence rates of these disorder diagnoses varied from 0.3% (probable TS) to 4.5% (probable transient tic disorder; Table 2).

Genetic analyses were performed on the four tic phenotypes grouped together in different ways for the various genetic analyses as described above in the methods section. Figure 2 shows the prevalence rates for each of these four (three dichotomous and one trichotomous). Depending on the strictness of the phenotype definition, prevalence rates varied from 1.3% ('probable TS or

probable chronic motor tic disorder') to 12.5% ('any probable tic', the most lenient phenotype).

Table 2. Descriptive statistics: number of participants and prevelance rate of DSM-IV probable tic disorders

	Probable tic disorder	N (%)	Male/Female
Participants classified with a tic disorder	Probable TS	44 (0.3%)	23/21
	Probable Chronic tic disorder (motor)	150 (1.0%)	71/79
	Probable Chronic tic disorder (vocal)	42 (0.3%)	22/20
	Probable Transient tic disorder	658 (4.5%)	264/394
	Probable Tic disorder NOS	637 (4.4%)	316/321
Total		1531 (10.5%)	696/835

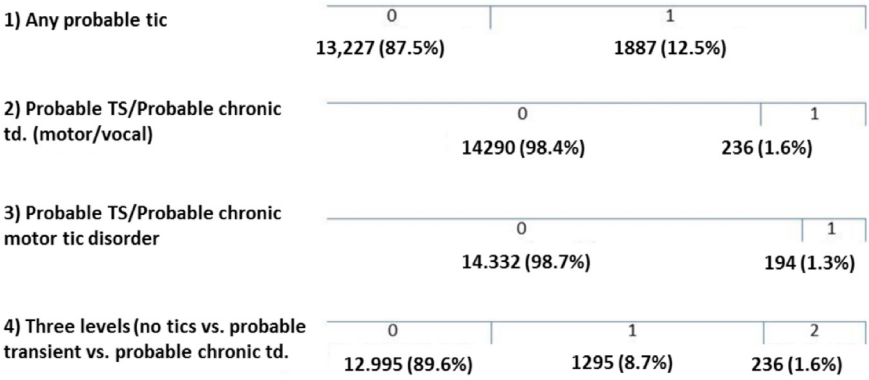


Figure 2. Descriptive statistics: Number of unaffected ('0') and affected ('1' and '2') participants, according to each of the four phenotypes. The fourth phenotype has two thresholds and is a combination of the first and second dichotomous phenotype.

Thresholds and Covariate Effects

The locations of the thresholds and effects of covariates are shown in online supplementary figure 3. In all baseline models, separate thresholds were estimated for offspring and parents (e.g., for the 'any probable tic' phenotype when only twins and parents were included, one threshold estimate instead of four resulted in a significantly worse fit: $X^2(3)=15.25$, p=.002; two threshold estimates, one for parents and one for offspring, did not significantly reduce the fit: $X^2(2)=3.07$, p=.22). Thus, for the 'any probable tic' phenotype, threshold estimates for parents were higher than for offspring, indicating that parents reported less tics. This was not seen in the more strict dichotomous phenotypes, indicating that the frequency of more severe tic disorders, based on self-report, and after correction for age, did not differ between parents and offspring.

Covariate effects were similar for the dichotomous variables (with the 'any probable tic' phenotype corresponding to the lowest threshold and the 'probable TS or probable chronic tic disorder' and 'probable TS or probable chronic motor tic disorder' phenotypes corresponding to the second threshold). Males were affected more often than females (e.g., for the 'any probable tic' dichotomous variable: b=-0.38, se=0.07, p<.001). We observed a decrease in the reporting of tics with increasing age ('any probable tic' phenotype, with standardized age: b=-0.17, se=0.05, p<.001). Participants who answered the paper-version of the questionnaire (instead of the online-version) reported more tics (using the 'any probable tic' phenotype: b=-0.16, se=0.08, p=.045; for the stricter tic disorder phenotypes this was not significant; online supplementary table 4). The interaction between age and method of reporting for the second dichotomous phenotype (probable chronic tic disorder and TS versus mild or no tic disorder) was found to be significant (b=-.33, se=0.15, p=.03).

Familial Correlations and Assortative Mating

Familial correlations are shown in Table 3. MZ twin correlations were higher than DZ twin correlations and correlations in other first-degree family members. Online Supplementary Table 5 summarizes the correlations between other family members. Since the correlation structure among relatives did not provide consistent evidence for either dominant genetic or common environmental effects, models with both dominance (ADE) and common environmental effects (ACE) were considered.

With respect to exploration of the influence of assortative mating: our data do not support evidence for assortative mating using any of the phenotypic

definitions.

Table 3. MZ and DZ twin polychoric correlations (and standard errors) for each phenotype

Phenotype	MZ twin corr. (se)	DZ twin corr. (se)
1) Any probable tic	.37 (0.05)	.18 (0.07)
2) Probable TS/probable chronic tic disorder	.24 (0.21)	.15 (0.21)
3) Probable TS/ probable chronic motor tic disorder.	.32 (0.21)	.19 (0.21)
4) Three levels (no tics vs. probable transient/NOS tic disorder. vs. TS/probable chronic tic disorder	.37 (0.06)	.17 (0.07)

Table 4. Estimated parameters, and fit indices of genetic analysis for each of the four phenotypes

	family members included	baseline model	a^2 (95%CI) *	fit (AE compared to baseline model)			thr 1 offspr/ parents	thr 2 offspr/ parents
				diff -2LL	Diff df	p		
1) any probable tic	tw+par	sat.	.31 (.23-.40)	11.98	11	.63	.92/1.06	-
	tw+par+sibs	constr.sat	.30 (.23-.38)	8.23	7	.69	.93/1.06	-
2) Probable TS / chr. TD	tw+par	-	.37 (.08-.61)				1.99/1.95	-
	tw+par+sibs	-	.31 (.04-.55)				2.04/1.95	-
	tw	sat.	.25 (.02-.60)	.008	1	.93	2.03	-
3) Probable TS /chr. motor TD.	tw+par	-	.32 (.02-.61)				2.10/2.07	-
	tw+par+sibs	-	.28 (.02-.56)				2.16/2.06	-
	tw	sat.	.34 (.02-.68)	.018	1	.89	2.12	-
4) three levels**	tw+par	sat.	.34 (.24-.44)	6.45	11	.84	.98/1.09	1.93/2.04
	tw+par+sibs	constr.sat	.33 (.24-.42)	3.36	7	.85	1.00/1.08	1.97/2.06

*Note. * c2 and d2 were tested but never significant; the remainder of the variance comes from unique environmental effects (e2); ** no tics vs. probable transient tic disorder/probable tic disorder NOS vs. probable TS/probable chronic tic disorder; 'TS/chr. TD' = Tourette Syndrome / Chronic tic disorder; 'thr' = threshold; 'tw' = twins; 'par' = parents; 'sibs' = siblings; 'sat.' = saturated model; 'constr.sat.' = constrained saturated model (i.e. parent-offspring correlations are set equal and full sibling correlations are setequal.*

Genetic Model fitting

Table 4 shows the results of genetic model fitting, where ACE and ADE differ in that the first model attributes familial resemblance to additive genetic and common environmental factors, and the second model attributes resemblance to additive and non-additive (dominance) genetic factors. In all models the C and D parameters were not significant: comparison with the more parsimonious AE model did not show a significant drop in the fit (e.g. for the first dichotomous phenotype, when twins, parents and siblings were included; AE vs. ACE: $x^2(1)<.001$, $p>.99$, and AE vs. ADE: $x^2(1)=2.59$, $p=.11$). Heritability point estimates ranged from .25 to .37. Thus, familial resemblance can be explained solely by additive genetic factors. The 95% confidence intervals were wide and all overlapping. The 'any probable tic' phenotype showed the narrowest confidence interval (0.31, 95% CI [0.23, 0.38]).

Discussion

The aim of this study was to estimate the heritability of increasingly strict phenotypic definitions of lifetime tic disorders that were mostly in line with current DSMIV and DSM 5 criteria for tic disorders, in a large adult population-based sample Further, using an extended twin design, we estimated the relative contribution of additive and non-additive genetic effects, effects of common and unique rearing environment, and the role of assortative mating. In line with Walkup, Ferrao, and Singer (2010), we were specifically interested in obtaining a clear understanding of the core phenomenological features of tics, taking one step further, i.e. by investigating whether and to what extent the various phenotypic definitions influence estimates of genetic and environmental contributions to tics (Walkup, Ferrao, & Singer, 2010). The abbreviated STOBS that we used is in line with both DSMIVTR and DSM5 criteria of the various tic disorders with respect to their core criteria of tic characteristics, duration, and age at onset, except for criterion D (i.e.: the disturbance is not attributable to a medical condition). In sum, the first 9 items of the abbreviated STOBS asked about tic characteristics (pertaining to criterion A), one additional item asked about age at onset (before versus after age 18) and one item asked about duration of tics (<1 year or > 1 year).

Our prevalence rates of a tic disorder are in the expected range (i.e. between 0.3-4.5%). In epidemiological studies in children, prevalence rates are between 3-8 cases per 1000 between ages 6 and 18 (Scahill et al. 2013). Our rates are higher than reported in most epidemiological studies in adults (0.001-0.05%), but in line with the other tic twin study in adults using self-reports by Pinto et al. (2016), who reported prevalence rates of TS between 0.4-1.4% depending

on strictness of phenotypic definition. Also, the prevalence rates of the most lenient definition of "any probable tic" of 12.5% in our sample is by and large in accordance (although somewhat higher) with the rates reported by Pinto et al. (2016) of 7.2% for any tic in men and 6.5% in women (Pinto et al. 2016). An explanation for the somewhat higher rates in our twin study and previous epidemiological studies (Pinto et al., 2016) might be that, using DSM-III and IV criteria (APA, 1994), earlier studies included an impairment/disability criterion for TS which has been subsequently removed from the DSM-IV-TR and DSM-V (APA, 2000). This may have caused a relative underestimation of tic prevalence. In our cross-sectional sample self-reported tic frequencies tended to decrease during adolescence and throughout adulthood, which is fully in line with interview-based epidemiological and clinical studies across the lifespan indicating that self-reported tic measures can be reliably used in large scaled studies. The decrease in tic frequencies with age might be the result of maturation of the frontal lobes, and -as a result- increased inhibitory efficiency of the cortico-striato-thalamo-cortical (CSTC) circuitry (Felling & Singer, 2011). However, recall bias, resulting in under-reporting of milder tics with age, should also be taken into account as a contributor to decrease of tic severity and frequency with age.

In this study, (narrow-sense) heritability estimates ranged from 0.25-0.37, with large confidence intervals that overlapped across the phenotypic definitions and that were in line with the other tic twin study in adults (Pinto et al., 2016) and somewhat lower than some family studies and twin-family studies in children (Bolton et al., 2007; Anckarsäter et al., 2011; Lichtenstein et al., 2010; Mathews & Grados, 2011). Possibly, with time, unique environmental mediators become increasingly important in the expression of these complex disorders. To conclude, in the present study heritability estimates for both mild and severe tic phenotypes were consistent, ranging from 0.25 to 0.37. However, the prevalences of the severe tic phenotypes were low, resulting in relatively low power to estimate the most strict thresholds models. As a consequence, confidence intervals of the heritabilities for the severe tic disorders are wide, and the narrow-sense heritability for severe tic disorders might be as large as 56% (the upper border of the confidence interval when siblings are included). Family-based studies specifically ascertaining probands with TS corroborate the data provided by this study, and suggest that heritability estimates might actually be on the high end of this estimate (58%-77%) (Hirschtritt et al., 2015). Interestingly, we found that heritability estimates of the "any tic" definition showed the narrowest Confidence Intervals, yielding moderate heritability estimates. Thus, the "core" tic phenotype that included only DSMIVTR and 5 criterion A of the various tic

disorders definitions (i.e. presence of a tic), seems to render the most reliable heritability estimates. In our opinion, in line with Walkup et al., 2010, this pleas for a relatively clear and simple phenotypic definition of tics in future data collection efforts for genetic studies, provided that the core phenotypic characteristics of tic disorders have been met, i.e. presence of tics, defined as "sudden, rapid, recurrent, non-rhythmic, stereotyped motor movements or vocalisations".

We found no evidence for assortative mating with respect to any of the tic phenotypes. In addition, we found no evidence of a contribution of common environment (C) or non-additive genetic effects (D), implying that all phenotypic definitions of probable tic disorders (mild/severe) are influenced by additive genetic factors and unique environmental factors. The absence of C is consistent with the Swedish twin study in children (Lichtenstein et al., 2010; Anckarsäter et al., 2011; Pinto et al., 2016). The discrepant findings by a Japanese twin study (Ooki, 2005) who found a large contribution of shared environmental effects on tics, might be due to cultural differences; i.e. cultural adaptations reflect differences in shared environmental contributions to heritability estimates in cross-groups comparisons.

The heritability estimates mentioned so far were estimated using the twin method. Davies et al. (2013) used SNP data from a GWAS of clinical TS cases to estimate the heritability attributable to the contribution of SNPs (e.g., GCTA; (Yang et al., 2011)). In contrast to our findings, Davis et al. found a high chip-based heritability estimate of .58, which is remarkably high compared to most SNP-based studies of complex disease (Wray & Maier 2014). We do not have a clear explanation for these divergent findings, although power issues and sample selection (clinical versus epidemiological) might play a role.

We did not attempt to model all phenotypic operationalizations simultaneously, as this is impossible due to the many empty cells in the cross tables. However, we assume that the variation in scoring of the various phenotypic definitions has a direct bearing on the diagnostic threshold, but not on the underlying liability. This implies that the estimates of the genetic and environmental contributions to individual differences in the liability should be equal. Our results are consistent with this, as the confidence intervals largely overlap across phenotype definitions, suggesting a continuous normally distributed liability for having a mild or severe tic phenotype. However, as the different phenotypic definitions did yield small but significantly different heritability estimates, this suggests that small (but significant) quantitative differences exist in genetic liability to tics.

The relatively modest heritabilities as found in this study, coupled with relatively large contribution of unique environmental influences, are consistent with the conceptualization of TS as a complex disorder like other complex psychiatric disorders, such as OCD and Anxiety Disorders (Zilhao et al. 2014; Pauls, 2010; van Grootheest et al. 2005; Shimada-Sugimoto et al. 2015; Hettema et al. 2001; Van Grootheest et al. 2007). In line with this, various environmental factors (such as stress, fatigue and life events) have been found to be relevant to the expression of tics (Findley et al. 2003; Swain & Leckman 2005). Importantly, this study has relevance for molecular genetics and GWASs. GWASs in complex traits have not been very successful to date, partly as a consequence of difficulties in defining and standardizing phenotypes (Sabb et al., 2009; Smith et al., 2013; Wray et al., 2012; Wray & Maier, 2014). Our work indicates that the heritability estimates from multiple tic phenotypic definitions largely overlap, strongly suggesting that future studies may use lower thresholds for tic classification, hence taking advantage of the increased power due to the higher number of cases that can be included in GWASs.

Results from this study should be interpreted in light of some limitations. The data collected are based on self-report measures (as this is a population-based study) and not on clinician-administered structural interviews, which might have led to misclassification. Additionally, since lifetime tics have been reported retrospectively, recall bias might have caused inaccuracy in recollecting past occurrences of tics.

In conclusion, our results indicate that genetic and unshared environmental factors contribute to the phenotypic variability across the full range of tic disorders. No shared environmental or genetic dominance effects were found to contribute. Finally there was no/little evidence for assortative mating. Our findings replicate and extend previous work in adults (Pinto et al., 2016), suggesting a relatively large contribution of environmental factors to the phenotype. However, these environmental influences might also include epigenetic or even genetic effects (private mutations). The heritability estimates of the different phenotypic definitions estimates are comparable (considering the confidence intervals), which is consistent with the liability threshold model, in which alternative scoring has a bearing on the threshold(s), but much less on the contributions of genetic and environmental factors to individual differences in the liability.

The Supplementary material can be found in the online version of this manuscript

Chapter 4
Cross-Disorder Genetic Analysis of Tic Disorders, Obsessive Compulsive and Hoarding Symptoms

This chapter is based on: Zilhao, N. R., Smit, D. J., Boomsma, D. I., & Cath, D. C. (2016). Cross-disorder genetic analysis of tic disorders, obsessive-compulsive, and hoarding symptoms. Frontiers in Psychiatry, 7.

Abstract

Hoarding, Obsessive-Compulsive Disorder (OCD) and Tourette's Disorder (TD) are psychiatric disorders that share symptom overlap, which might partly be the result of shared genetic variation. Population-based twin studies have found significant genetic correlations between hoarding and OCD symptoms, with genetic correlations varying between 0.1 and 0.45. For tic disorders, studies examining these correlations are lacking. Other lines of research including clinical samples and GWAS or CNV data to explore genetic relationships between tic disorders and OCD have only found very modest if any shared genetic variation. Our aim was to extend current knowledge on the genetic structure underlying hoarding, OC symptoms (OCS) and lifetime tic symptoms and, in a trivariate analysis, assess the degree of common and unique genetic factors contributing to the etiology of these disorders. Data have been gathered from participants in the Netherlands Twin Register comprising a total of 5293 individuals from a sample of adult monozygotic (n=2460) and dizygotic (n=2833) twin pairs (mean age 33.61 years). The data on Hoarding, OCS and tic symptoms were simultaneously analysed in Mplus. A liability threshold model was fitted to the twin data, analysing heritability of phenotypes and of their co-morbidity. Following the criteria for a probable clinical diagnosis in all phenotypes, 6.8% of participants had a diagnosis of probable hoarding disorder (HD), 6.3% of OCS, and 12.8% of any probable lifetime tic disorder. Genetic factors explained 50.4%, 70.1% and 61.1% of the phenotypic covariance between hoarding-OCS, hoarding-tics and OCS-tics respectively. Substantial genetic correlations were observed between hoarding and OCS (0.41), hoarding and tics (0.35) and between OCS and tics (0.37). These results support the contribution of genetic factors in the development of these disorders, as well as their co-morbidity. Furthermore, tics were mostly influenced by specific environmental factors unshared with OCS and HD.

Introduction

Current classification systems of psychiatric disorders are primarily based on consensus statements with respect to clinical symptom diagnostics by physicians. These classification systems, i.e. the International Classification of Diseases (ICD) ("The ICD-10 Classification of Mental and Behavioural Disorders," n.d.) and the Diagnostic and Statistical Manual of Mental Disorders (DSM) (APA, 2013), have rendered the separate and categorical entities we know as disorders – including Obsessive-Compulsive Disorder (OCD), Tourette's Disorder (TD) and (starting from DSM5) Hoarding Disorder (HD).

More specifically, OCD, HD and tic disorders/Tourette's Disorder are complex neuropsychiatric disorders all characterized by repetitive behaviours that show substantial co-morbidity, i.e. that co-occur more often than expected by chance (Frost et al., 1999; Horwath & Weissman, 2000; Scahill, Sukhodolsky, & Leckman, 2005; Tolin, Frost, & Fitch, 2008). OCD is a neurodevelopmental disorder characterized by recurrent intrusive thoughts (obsessions) and repetitive behaviours (compulsions) designed to relieve either tension or anxiety stemming from the obsessions (APA, 1994; WHO, 2007). HD has since long been classified as a symptom dimension of OCD, and -to a lesser extent- as a characteristic of Obsessive-Compulsive Personality disorder (APA, 1994). It was later suggested, however, that 1) HD presents mostly (in up to 80% of cases) without concurrent OCD (Pertusa et al., 2008) and that 2) the neurological mechanisms underlying hoarding might be distinct from OCD (Mataix-Cols et al., 2010; Pertusa et al., 2008). Therefore, it was included in DSM-5 as a distinct disorder in the category of OCD spectrum disorders, and characterized by the inability to discard an excessive amount of items of no significant value, combined with excessive acquisition and clutter to such an extent that living spaces of an individual are occupied (APA, 2013). Tic disorders are characterized by recurrent motor and/ or vocal tics that occur in a stereotypical fashion against a background of normal motor/phonic activity, with onset in childhood and tendency to decrease in intensity and frequency during adolescence (Cath et al., 2011).

Prevalence rates for these disorders range between 0.1-0.8% for TD (Apter et al., 1993; Bar-Dayan, Arnson, & Elishkevits, 2010; Eapen, Laker, & Robertson, 2001; Robertson, Verril, & Pauls, 1994; Scahill, Dalsgaard, & Bradbury, 2013; Schlander, Schwarz, & Roessner, 2011; Wenning et al., 2005; Zohar et al., 1992), 2-6% for compulsive hoarding (Iervolino et al., 2009; Timpano et al., 2011) and 0.5-2.0% for OCD (APA, 1994; Ayuso-mateos, 2001).

With respect to co-morbidity rates between HD and OCD, in clinical and epidemiological studies of OCD between 18-42% of patients report hoarding behaviors, depending on phenotypic definition (Frost & Hartl, 1996; Lochner et al., 2005; Rasmussen & Eisen, 1989; Samuels, Bienvenu, & Cullen, 2002), and reversely, in 12-20% of HD patients OCD is reported (Frost, Steketee, & Tolin, 2011; Ivanov et al., 2013; Mathews, Delucchi, & Boomsma, 2014). In TD/chronic tic disorders, OCD is very common, with estimates ranging from 28-49% of OCD/ OCS in TD, and reversely, of 10-20% of tics in OCD (Como, LaMarsh, & O'Brien, 2005; Rosario-Campos et al., 2005). In sum, these comorbidity estimates are well above expected comorbidity rates if the three disorders would be etiologically distinct. Finally, in tic disorders no studies on hoarding co-morbidity have been performed nor have studies been performed on tic co-morbidity in HD.

Family studies and genetic epidemiological twin studies on each separate disorder have shown substantial genetic contribution to each separate phenotype, with heritability estimates from twin studies ranging between .30-.58 (OCD) (Rosario-Campos et al., 2005; Hudziak et al., 2004; Iervolino, Rijsdijk, & Mataix-Cols, 2011; Pauls, Alsobrook, & Leckman, 1995; Zilhao et al., 2014), .35-.50 (HD) (Iervolino et al., 2009, 2011; Ivanov et al., 2013), and .25-.58 (tic disorders) (de Haan, Delucchi, & Cath, 2015; Hirschtritt et al., 2015; Mataix-Cols et al., 2015; Mathews & Grados, 2011; Pauls, Raymond, & Leckman, 1991). A next question is whether the high proportions of co-occurrence between the three phenotypes reflects overlap in genetic or environmental contributions between OCD, HD and tics. Multivariate twin/ family studies are particularly suitable for this, making use of correlations between MZ and DZ twins on the various traits to partition the relative contribution of shared versus unique genetic and environmental factors that influence multiple traits (Dongen, Slagboom, & Martin, 2012).

Despite recent advances in psychiatric genetics, twin studies specifically investigating shared genetic and environmental influences between OC symptoms (OCS), hoarding behaviour and tics are scarce. Two studies by Iervolino et al. in a sample consisting predominantly of female twins from the TwinsUk twin registry (4459 female twins, mean age of 55.0 years) have specifically examined the genetic and environmental overlap between OCS and HD behaviour (Iervolino et al., 2009, 2011). It was found that 45% of the genetic variance was shared between HD and OCS dimensions. Further, hoarding had the lowest loading on the common factor with only 55% of the total variance in OC symptom dimensions being hoarding-specific. A recent twin study of our group within the Netherlands twin Register, which overlaps

with our sample, assessed the unique and shared genetic contributions for HD and OCS in a sample of 7567 twins (2270 males, 5297 females, mean age of 33.2 years) (Mathews, Delucchi, & Boomsma, 2014). The authors found significant genetic contributions to the co-morbidity across both traits, although a low genetic correlation (0.10) was found. Finally, a recently population-based twin family study with data from the Swedish Twin Register (n=20,821), specifically addressed the proportion of shared genetic and environmental factors underlying the liability to chronic tics, ADHD and OCS (Pinto et al., 2016). Tics were broadly defined based on the number of total tics ('no tic score', 'tic score=1', 'tic score > 1'). A substantial correlation of 0.45 between tics and OCS was found.

From another line of research Genome-wide association study (GWAS) data from samples of TD and OCD patients were analysed to find a genetic correlation between OCS and TD of 0.41 (Davis et al., 2013), which was relatively high in light of what has been described for other complex disorders (Wray & Maier, 2014). However, this correlation might have been an overestimation, as the standard error of this estimate was large (se=0.15) and, in addition, the co-occurrence between tics and OCD appeared relatively high (13% of OCD had co-occurring tics/TD and reversely 43% of TD had OCD) Further, in this same sample, Yu et al. (2014), sought to characterize common genetic variants shared among TD and OCD. Although no specific variants were identified, the combined GWAS signals were significantly enriched for functional alleles, suggesting that there is some proportion of TD-OCD shared genetic risk variants (Yu et al., 2014).

So far, genetic-epidemiological twin-family studies to estimate the shared respective unique contributions of genetic and environmental factors between tic- HD symptoms and between tic-HD-OCS are lacking, as are molecular genetic studies to estimate shared genetic contributions from SNPs across TD, OCS and HD phenotypes.

Therefore, the main aim of the present study was to extend the available data so far with respect to shared etiology between OC symptoms and hoarding behaviour (Mathews et al., 2014) by expanding with the tic phenotype, in a large population-based twin sample that includes male, female and opposite sex twin pairs using diagnostic methods that assess the full range of the symptomatology of these disorders to better address their shared underlying etiology. Specifically, we aimed at 1) replicating previous quantifications of shared and independent genetic contributions to OCS-hoarding behaviour; 2) quantifying shared and independent genetic contributions to hoarding behaviour and tics; 3) quantifying shared and independent genetic contributions

to OC symptoms and tics; 4) quantifying shared and independent genetic contribution to OCS-hoarding behaviour and tics.

Materials & Methods

Subjects

Participants included in this study are registered with the Netherlands Twin Register (NTR). Since 1991, twins and their family members receive surveys by mail, and are assessed with questionnaires about health, personality and lifestyle (Boomsma et al., 2006; Willemsen et al., 2013). For these analyses, we used data collected in 2008, corresponding to the survey 8 wave of collection, on Obsessive Compulsive Symptoms, Hoarding and tic symptoms (henceforth named as 'tics'). A total of 16.930 participants from 7400 different families completed the questionnaires. Twins encompassed 8047 individuals (2511 males, 5536 females). This study has been approved by the Medical Ethical Committee of the VU Medical Centre Amsterdam.

Measurements

The assessment instruments used were the Hoarding Rating Scale-Self-Report (HRS-SR) for Hoarding, the Padua Inventory Abbreviated Revised (PI-ABBR) for OCS and an abbreviated self-report questionnaire (the Schedule for Tourette and Other Behavioral Syndromes - STOBS-ABBR) based on the Schedule for Tourette and Other Behavioral Syndromes (STOBS) for tics. The HRS-SR questionnaire consists of five items each scoring on a 0-8 scale, that assess cluttering, difficulty in discarding items, excessive acquisition or collecting, distress derived from hoarding symptoms and functional impairment (Tolin, Frost, & Steketee, 2010). The distress item was discarded due to approval restriction on the items to be included in the larger questionnaire. The PI-ABBR questionnaire has been derived from the Padua Inventory-Revised, a 41 item self-report instrument that measures OC symptoms on a scale from 0-4, and 5 subsequent subscales (washing, checking, rumination, precision and impulses). The PI-ABBR has been abbreviated to 12-items that include two to three items from each of the five OCS dimensions above mentioned (Cath, van Grootheest, & Boomsma, 2008). These subscales refer to four main factors of obsessions and compulsions – 'impaired control', 'fear of contamination', 'checking behaviour', 'urge/worry of losing control' (Sanavio, 1988).

The STOBS consists of a semi-structured assessment on tics, and has been widely used in data collections by the Tourette Syndrome Association International Consortium for Genetics (TSAICG). It consists of 36 tic items

(rated as: current/lifetime, not present), generating lifetime tic information (Pauls & Hurst, 1993). For the NTR 2008 survey, the STOBS was abbreviated to a 12-item tic questionnaire on the 9 most frequent tics occurring in clinical samples (Cath et al., 2011; Freeman et al., 2000). Additionally, 3 items were added on age at onset of symptoms, tic severity and whether the tic persisted for more than a year. Using the STOBS-ABBR, a diagnosis of probable chronic tic disorder was established if the person had 1) one or more chronic motor or one or more vocal tic, that 2) occurred before age 21 and 3) had been present for >1 year. Probable TD diagnosis was established when 2 or more motor and 1 or more vocal tics were reported that occurred before age 21 and had lasted for > 1 year, and probable transient tic disorder was established when motor and/ or vocal tics had occurred before age 21 for less than one year. Participants who reported at least one tic, but without an age at onset≤21, and/ or with a tic duration of <1 year were categorized as a probable tic disorder NOS. We use the term "probable" since tic diagnoses were not confirmed by a face-to-face interview by an experienced clinician.

We fitted a liability threshold model, using for each phenotype a categorical variable derived from several cut points applied to the full distribution of sum scores (for OCS and HD), and defining the presence/absence of a tic disorder (for tics). The liability threshold model assumes an un-observed (and not measured) liability (or risk) to disease, normally distributed in the population (Falconer, 1965, 1967). The categories function as a (indirect) measure of this liability, representing the susceptibility to the true underlying distribution of the disease. Four categories were used for both the HRS-SR and PI-ABBR. The HRS-SR was divided into categories that more closely resemble the clinical patterns of symptomatology (no hoarding symptoms, mild symptoms, subclinical hoarding and clinically significant hoarding or probable hoarding disorder), having unequal distributions in each category (scores of 0, 1-5, 6-16, and ≥17) (Iervolino et al., 2009). For a probable HD diagnosis, we used the cut-off proposed by Tolin to define caseness (Tolin et al., 2010). In this work, a receiver operating characteristic (ROC) analysis determined that the best threshold separating HD from non-HD cases was a sum-score over the cut-off of 17 with a sensitivity and a specificity of 0.95. The scores for PI-ABBR (0, 1-6, 7-15, and ≥16) have been previously described in the literature (Cath et al., 2008). In brief, ROC determined that the best threshold separating OCD from non-OCD cases was a sum-score over the cut-off of 16 with a sensitivity of 0.74 and a specificity of 0.72. For tics, we derived a dichotomous variable defining the presence or absence of any of the tic disorders described here-above, according to a definition of 'probable tic disorder' as defined by the STOBS-ABBR. For further details

on the phenotype definition for tics, please refer to (Zilhao et al, 2016). Briefly, the probable tic disorder dichotomous variable consists of the most lenient definition defined for caseness, in which lifetime probable chronic tic disorder, probable TD, and probable transient tic disorder are included.

Statistical Analyses

Univariate Twin Analyses

Prevalences, means and distributions for the three phenotypes were calculated in the entire sample of 16.930 individuals. Performing these analysis on clinically defined significant symptoms has the advantage of increasing the generalizability of the results. Polychoric correlations (correlations on the liability scale) were calculated in Mplus (Prescott, 2004) for the PI-ABBR, HRS-SR, and STOBS- ABBR, both in MZ and DZ twin pairs by sex, and in all twins for both sexes. Data from both complete and incomplete twin pairs were included in the analysis. Univariate analyses for each phenotype were performed separately using the software OpenMx (Boker et al., 2011) to estimate the relative contributions from additive genetic (A), shared environment (C) and non-shared environment (E) to each phenotype. Maximum-likelihood model-fitting procedures were carried out, as is standard in structural equation modelling, in which the phenotype was a function of the A, C and E factors and polychoric correlations, according to the liability threshold model described above. We investigated the potential influence of twin-specific and gender-specific (sex differences) environment by constraining correlations across zygosity groups to be equal, for all three phenotypes. The effect of covariates (age and sex) on the thresholds was univariately assessed for each phenotype.

Multivariate Twin Analyses

Using the Mplus software we then fitted a Trivariate genetic model to the data with the weighted least square mean and variance adjusted estimation option (WLSMV) (Prescott, 2004), using the described liability threshold models. Covariances between the three phenotypes were partitioned into the relative contributions of shared additive genetic (A), common environmental (C) and non-shared environmental (E) influences to the etiology of the three phenotypes. The influence of common environmental factors and of genetic dominance were tested by comparing a nested AE model with either the ACE or the ADE model using the Chi-square difference test.

Lastly, we performed a single factor analysis on the covariance matrices partitioned between the phenotypes. This analysis gives a representation in

terms of the components shared by the three phenotypes.

Results

Descriptives

Means and distributions

The mean age of the entire sample was 33.61 years (SD= 14.56); for males the mean age was 33.11 (SD=14.66) years and for females 33.84 (SD=14.51) years. The mean average score for HRS-SF was 5.74 (SD=5.6) and for the PI-ABBR was 6.89 (SD=5.2). Males had on average higher scores than females on both the HRS-SF and the PI-ABBR. Also for tics, the prevalence rates were higher in males (13.0%) than in females (12.6%). Table 1 summarizes the demographics in males and females for the PI-ABBR, HRS-SF and STOBS-ABBR.

Table1. Sample demographics for the data included in the analysis

	MZ twins		DZ twins	
	Male	Female	Male	Female
Mean age	35,09 (15,27)	35,63 (15,20)	31,57 (13,98)	31,88 (13,45)
Mean HRS	5,85	5,5	6,08	5,79
Mean PADUA	6,99	6,7	7,04	6,99
Tics (prevalence)	192	188	175	162

Prevalence and phenotype 'overlap'

Table 2 shows prevalence rates for the three phenotypes for MZ and DZ twins, as estimated according to the diagnostic criteria. Of the entire sample, 5.0% had clinically significant HD, 6.0% had clinically significant OCS, and 13.5% had any probable tic disorder according to the STOBS-ABBR. The threshold used to determine caseness in a probable HD disorder diagnosis rendered population prevalence rates that closely resemble previous estimates for clinical HD (Iervolino et al., 2009; Timpano et al., 2011). Furthermore, among individuals with OCS, 18.0% had co-occurring HD and 12.1% had tics; among individuals with HD, 15.0% had OCS and 8.72% had tics; among individuals with tics, 27.1% had OCS and 23.3% had HD. Lastly, in the entire sample, 0.31% (n=25) of individuals had the co-occurrence of all three disorders.

Table 2. Prevalence rates for HD (HRS-SR, OCS (PI-R-ABBR), and tics (YGTSS), for the total sample included in the analysis.

	Category	MZ (n=3990) N(%)	DZ (n=4057) N(%)
HD (n=5221) symptom scores	0	673 (22.8%)	435 (19.1%)
	1 - 5	1059 (36%)	826 (36.2%)
	6 - 16	1079 (36.6%)	880 (38.6%)
	> 16	137 (4.6%)	132 (5.7%)
OCS (n=5167) symptom scores	0	190 (6.5%)	140 (6.3%)
	1 - 4	1447 (49.6%)	1077 (48.0%)
	5 - 15	1107 (37,9%)	898 (39.9%)
	> 15	175 (6.0%)	133 (5.9%)
Probable tic disorder (n=5297) affected/non-affected	TD	15 (0.5%)	14 (0.6%)
	Chr motor tic	39 (1.3%)	35 (1.5%)
	Chr vocal tic	17 (0.6%)	16 (0.7%)
	Transient tic disorder	159 (5.4%)	138 (5.9%)
	Tic disorder NOS	150 (5.0%)	134 (5.7%)

Univariate results

Twin correlations

Table 3 shows the polychoric correlations as calculated on the observed data for the five zygosity groups, on the HRS-SR, PI-ABBR and the STOBS-ABBR. Overall, when comparing MZ and DZ pairs on the three phenotypes, an average two-fold increase for MZ twins when compared to DZ twins is observed. The greater similarity for MZ twins the moderate MZ correlations suggest the influence of non-shared environmental factors for all three phenotypes.

Specific gender/twin environments were tested univariately for each phenotype. As expected from the twin correlations across all zygosities, the fit statistics results show that correlations could be equated across twins and sex, with no twin-specific or sex-specific environments observed (table 4).

Table 3. Polychoric twin correlations for observed data for HD, OCS and tics

	MZ	MZM	MZF	**DZ**	DZM	DZF	DOS
HD (HRS-SR)	**0.336**	0.379	0.325	**0.177**	0.247	0.151	0.048
OCS (PI-R-ABBR)	**0.384**	0.379	0.386	**0.177**	0.197	0.139	0.214
Tics (STOBS)	**0.37**	0.242	0.414	**0.19**	0.238	0.172	0.114

Table 4. Model fit indices for the univariate models, examining the role of sex and zygosity of each phenotype separately.

Model	NP	-2LL	Versus model	χ^2	df	p
1. Hoarding, saturated	10	-	-	-	-	-
2. Hoarding, equal sex and zygosities	7	3250.51	Hoarding, Saturated	3.83	3	0.28
3. OCS, Saturated	10	-	-	-	-	-
4. OCS, equal sex and zygosities	7	840.18	OCS, Saturated	1.07	3	0.78
5. Tics, saturated	10	-	-	-	-	-
6. Tics, equal sex and zygosities	7	16091.12	Tics, Saturated	5.45	3	0.14

*Note: NP, Number of parameters; -2LL, -2*Log-Likelihood; df, degrees of freedom for χ2 test*

Heritabilities and fit statistics

The total heritability estimates were 0.33 (SE=0.05, p<0.001) for clinically significant HD, 0.38 (SE=0.05, p<0.001) for OCS and 0.37 for any tic disorder (SE=0.05, p<0.001) (off-diagonal in table 6). For non-shared environment the estimates were 0.67 (SE=0.05, p<0.001) for clinically significant HD, 0.62 for OCS (SE=0.05, p<0.001) and 0.63 (SE=0.05, p<0.001) for tics. No evidence was found for an effect of common environment.

Cross-disorder correlations

Examining the cross-disorder correlations (cross-twin cross-trait) again suggests that genetic factors are involved in the correlations between traits (table 5). The MZ cross-twin cross-trait correlations were 0.14 (HD vs. OCS), 0.12 (HD vs. tics) and 0.16 (OCS vs. tics), while the DZ correlations were 0.07 (HD vs. OCS), 0.05 (HD vs. tics) and 0.02 (OCS vs. tics). The within-person cross-trait correlations (phenotypic correlation) were 0.30 (HD vs. OCS), 0.15 (HD vs. tics) and 0.25 (OCS vs. tics) (table 5).

A trivariate ACE model was fitted to the data in order to examine the relative contributions from shared genetic and environmental contributions to the covariance among the traits. Again, as suggested by patterns of twin correlations, no evidence for common environment was found and the C parameter could be dropped when compared to the more parsimonious AE model (AE vs. ACE model: x^2 (6) =0.876, p=0.99 and AE vs. ADE model: x^2 (6) =2.994, p=0.81). Hence, the best fitting model to the data was one in which the covariation between the three phenotypes can be explained by a set of common A and E factors. Table 5 and figure 1 show the estimates of the relative contributions of genes and non-shared environment factors, calculated from the best-fitting model. The total variance for each variable

was constrained to 1, in order to estimate the proportion of individual liability due to shared vs. common genetic/environmental factors. Bivariate heritability results (table 5), show that 50% of the covariance between HD and OCS, 70% of the covariance between HD and tics, and 61% of the covariance between OCS and tics are due to genetic factors. The remaining variance is accounted for by non-shared environmental factors. Furthermore, the genetic correlations were 0.41 (HD vs. OCS), 0.35 (HD vs. tics) and 0.37 (OCS vs. tics). Figure 1 depicts the path diagram in terms of correlated A and E factors.

Table 5. Relative contributions of additive genetic and non-shared environmental influences on the trait variance (diagonal) and covariance cross-trait (off-diagonal) for HD (HRS-SR), OCS (PI-R-ABBR) and tics (YGTSS).

	Phenotypic correlation		CTCT (MZ below, DZ above diagonal)			Additive genetic effects (A)			Non-shared environmental effects (E)		
	HD	OCS	HD	OCS	tics	HD	OCS	tics	HD	OCS	tics
HD	-		-	0.07	0.05	0.33	-	-	0.67	-	-
OCS	0.30	-	0.14	-	0.02	0.50	0.38	-	0.50	0.63	-
Tics	0.15	0.25	0.12	0.16	-	0.70	0.61	0.37	0.30	0.39	0.63

Note: CTCT= cross-twin-cross-trait correlations.

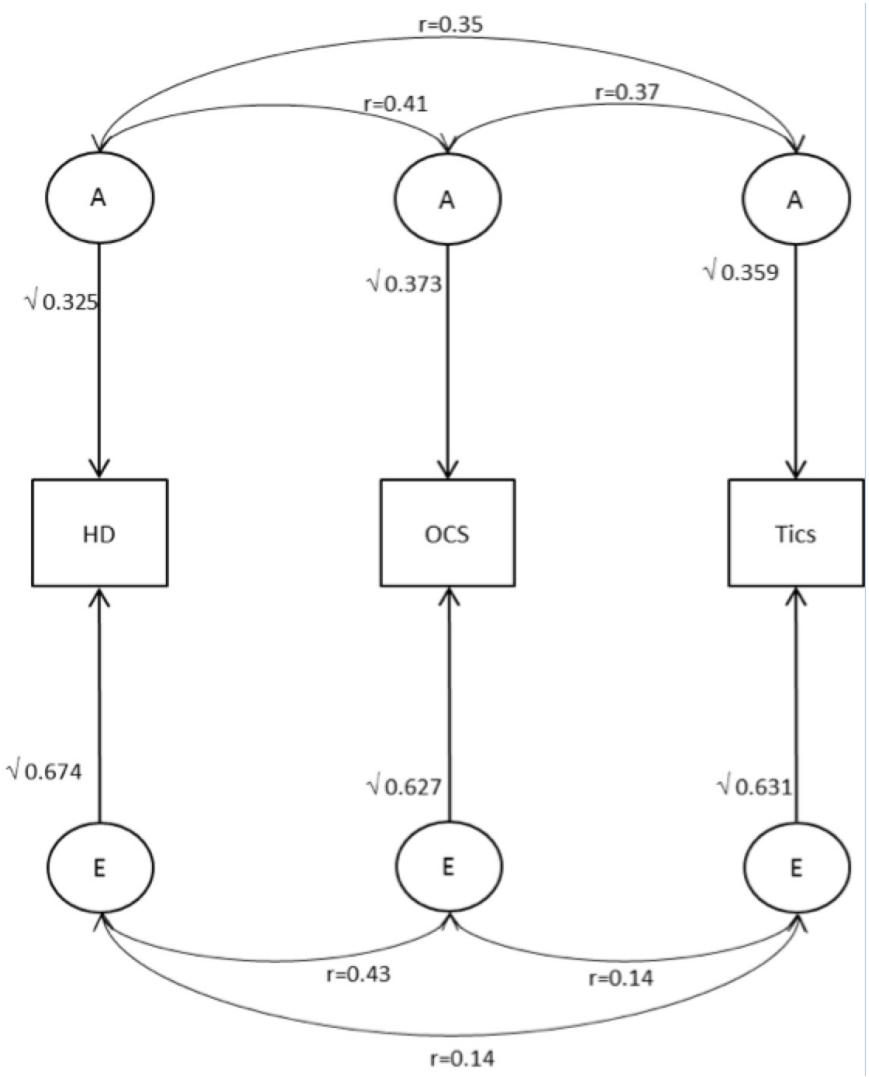


Figure 1. Path diagram for the best fitting model. Squaring these paths gives the proportion of variance accounted by each of the A and E components. Also indicated are the correlations among each A and E component for each of the three phenotypes. A indicates additive genetic factors; E indicates non-shared environmental factors.

Lastly, single factor analysis for the A and E component revealed the degree of genetic and environmental overlap shared by the three phenotypes (Figure 2). As shown, between 31.5% and 43% of the total genetic variance of each phenotype is due to genetic factors shared among all three phenotypes. Specific genetic variance unshared with other phenotypes was 60.7% (HD), 57.0% (OCS) and 68.5% (tics). Further, 43.2% and 41.8% of the total environmental variance is due to unique environmental factors shared between HD and OCS, respectively, whereas for tics this amounts only to 4.4% of the total environmental variance – in other words, tics had the lowest loading on the common factor and were mostly influenced by tic-specific environmental effects (Figure 2).

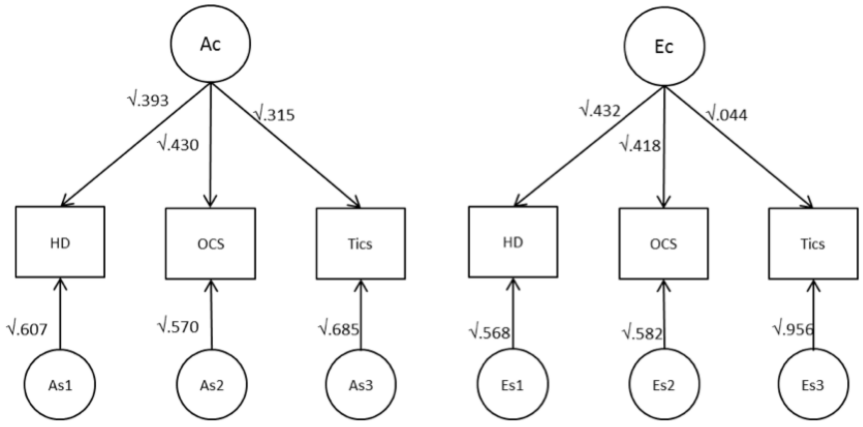


Figure 2. Single factor representation for the each of the A and E component for the best-fitting model. Numbers indicate the proportion (for both A and E components) shared by the three phenotypes. Ac indicates common additive genetic factors; Ec indicates common non-shared environmental factors.

Discussion

In this study we sought to examine the extent to which shared genetic and environmental factors contribute to clinically significant OCS, HD and tic symptomatology. We had at our disposal the largest twin pair sample available to date in which these three phenotypes were measured at the same wave of data collection. The present results extend previous work in the same NTR sample on shared genetic contributions to OCS and HD (Mathews et al., 2014).

Our univariate prevalence rates for clinical significant HD symptoms and OCS are in the expected range when compared to the literature (Cath et al., 2008; Samuels et al., 2008). For tics we note that our somewhat higher prevalence rates than described in the literature might be due to the fact that they reflect lifetime tic disorders, and therefore a somewhat lenient definition for caseness, reflecting our approach to generate optimal results with respect to phenotypic validity, in light of the self-report measures used in the NTR.

Our comorbidity prevalence rates - 18% of OCS patients reported co-occurring HD, and reversely, 15.0% of HD patients reported co-occurring OCS; 12.1% of OCS patients reported co-occurring tics, and reversely, 27.1% of TD/chronic tic disorders reported co-occurring OCS), are within the expected range when compared with the epidemiological literature (Como et al., 2005; Rosario-Campos et al., 2005; Frost & Hartl, 1996; Frost et al., 2011; Ivanov et al., 2013; Lochner et al., 2005; Mathews et al., 2014; Rasmussen & Eisen, 1989; Samuels et al., 2002). For HD/tics, to the best of our knowledge, we report here the first co-morbidity prevalence rate estimate – 8.72% of HD individuals having co-occurring tics and 23.3% of tic individuals having HD.

The univariate model fitting results

Previous results with data from the NTR, using by and large the same sample, have yielded heritabilities of 0.40-0.50 for OCS (Cath et al., 2008; van Grootheest et al., 2009), 0.36 for HD (Mathews et al., 2014) and 0.30 for tics (Zilhao et al., 2016). Other previous twin/family studies have rendered comparable estimates (0.26-0.55 for OCS, 0.35-0.50 for hoarding, and 0.28-0.56 for tics) (Pinto et al., 2016; Taylor, Jang, & Asmundson, 2010; van Grootheest et al., 2005). We found no evidence for sex differences in twin correlations for any of the phenotypes. Similar findings have been reported for OCS (Bolton et al., 2007; Hudziak et al., 2004; Iervolino et al., 2011), whereas for HD results have been mixed (Iervolino et al., 2009; Ivanov et al., 2013); for tics, to the best of our knowledge, the issue of sex differences in twin correlations has not yet been addressed. Our results here show that the

genetic contributions to these phenotypes are consistent across both sexes.

Bivariate analyses

Second, our results provide evidence for shared genetic variation between the phenotypes. The phenotypic correlation between OCS and HD was of 0.30. As expected, we observed a higher phenotypic correlation between OCS and tics (0.25) than between HD and tics (0.15). The genotypic correlations also mirrored this – there was higher shared genetic variance between OCS and HD (0.41), than both OCS or HD with tics (0.37 and 0.35, respectively). Interestingly, a relatively high proportion of the phenotypic correlations were attributable to genetic factors. In other words, although the genetic overlap (expression of same genes) between tics and both OCS and HD is moderate, a substantial proportion of the phenotypic correlation is mediated by their shared genetic variance (61% and 70%, respectively).

Importantly, Iervolino et al. (2011) recently reported a genetic correlation between OCS and HD of 0.45, combined with their data suggesting that HD was mostly influenced by specific genetic effects (54.5% specific) (Iervolino et al., 2011). The authors argued that this supports the notion of these disorders constituting two etiologically distinct, although related, entities (Iervolino et al., 2009, 2011). Further, Mathews et al. reported a substantially lower genetic correlation of 0.10 (Mathews et al., 2014). Our current findings of all cross twin cross trait genetic correlations being below 0.2, and the within person cross trait correlations being all below 0.35 are mostly in line with those in the study by Iervolino et al. (2011). Iervolino et al argue that the magnitude of these genetic correlations is lower than the shared genetic variance of 0.55 between OCD and other internalizing disorders. i.e. panic disorder, generalized anxiety, phobias, and PTSD (Iervolino et al., 2011; Tambs et al., 2009). They reason that a genetic overlap just under .50 argues in favor of HD being a separate, but related entity as it is currently defined in DSM-5. Our data on the relationship between HD and OCS are in support of this view.

Our estimates of genetic correlations between OCS and tics (0.37) are somewhat lower than the genetic correlations (0.45) as found by Pinto et al. (Pinto et al., 2016). The differences in estimates might be explained by the different phenotypic tic definitions requiring an age of onset before 21 resulting in a prevalence of 13.5%, whereas their multinomial definition of lifetime tics into categories 'no tic', 'one tic', and 'two or more tics' resulted in prevalences of 16% at the first, and 6% at the second threshold. Further, our results are not fully in line with tic/ OCS enriched clinical family studies

reporting very high genetic correlations between TD and OCS (genetic correlation= 0.92), although the standard errors in this study were high (se=0.42) (Mathews & Grados, 2011).

With respect to the shared genetic and environmental contributions to HD and tics, to our knowledge, this is the first twin- family study partitioning the covariance between tics and HD in its relative genetic and environmental components. Our moderate correlation estimate (0.35) support the argument of viewing TD as distinct from HD.

Third, the common factor model further supports the view of shared genetic etiology between the three phenotypes. Neuroimaging studies have reported structural and functional dysfunctions in the cortico-striato-thalamo-cortical (CSTC) circuitries across all three disorders that have negative implications for motor response inhibition and interference control in these disorders, which might underlie the phenotypic behaviors of all these three disorders (Posner, Marsh, & Simpson, 2014; Velzen, Vriend, & Heuvel, 2014; Wang et al., 2011). Our results raise the interesting possibility that a common genetic architecture defines underlying CSTC dysfunctions across the three disorders. Follow-up genome-wide studies may investigate whether specific genetic variants involved in all three disorders are differentially expressed in these brain areas as a result of non-shared environmental influences. In support of this, interestingly, OCS and HD showed low environmental correlations with tics, suggesting that tic disorders have specific environmental contributors invoking tic symptoms. In other words, non-familial (unique) environmental experiences may determine the development of tics, separately from the broader obsessive-compulsive related disorders, as currently defined in DSM-5 (APA, 2013).

Finally, our results are relevant for the field of molecular genetics. The lack of power to detect specific genetic risk variants is a recurrent issue in genome-wide studies. One way to overcome this limitation is to combine related phenotypes therefore increasing sample sizes, with consequent power gains. A crucial point here is the balance between power gains from increased sample sizes and power losses from increased heterogeneity (Manchia et al., 2013; Wray & Maier, 2014). Our results suggest that although these disorders share substantial genetic overlap, a substantial proportion of the genetic risk variance contributing to the liability to each disorder is independent from each other, and care should be taken when combining the phenotypes as studied in this paper.

Limitations

These results should be considered in the light of some limitations, mainly considering the phenotypes. Because this is a population-based study, the data collected are based on self-report measures, rather than on clinician-administered structural interviews. The cut-offs have been empirically derived, and are therefore somewhat arbitrary. The cut-offs to determine symptom thresholds (in the case of OCS and HD), by considering the entire range of age available, may have rendered different prevalence estimates, which might have affected estimations of genetic and environmental effects. We note, however, that although these threshold cut-offs do not represent definite clinical diagnoses, they do correspond to clinical significant symptoms. Moroever, investigation of dimensions rather than true/false categorical diagnosis is consistent with the ideas forwarded in the NIMH Research Domain Criteria (RDoC; Patrick & Hajcak, 2016).

To conclude, OCS, HD and tics share etiologic variance that can be explained by substantial genetic correlations. Tics are mostly influenced by specific environmental effects unshared with neither OCS or HD, suggesting that specific environmental stressors might cause the development of tics separate from OCS and HD. Our results are in line with the literature supporting the current definition in DSM-5 of separating these disorders into different, although related, entities.

The Supplementary material can be found in the online version of this manuscript

Chapter 5

Obsessive-Compulsive Symptoms in a Large Population-Based Twin-Family Sample are Predicted by Clinically-Based Polygenic Scores and Genome-Wide SNPs

This chapter is based on: den Braber A, Zilhão NR, Fedko IO, Hottenga JJ, Pool R, Smit DJ, Boomsma DI. (2016). Obsessive-compulsive symptoms in a large population-based twin-family sample are predicted by clinically based polygenic scores and by genome-wide SNPs. Translational Psychiatry, 6(2).

Abstract

The heritability of Obsessive-Compulsive symptoms (OCS) is significant, but genetic association studies have thus far not yielded consistent findings of variants influencing OCS. We aimed to contribute further insights into the genetic basis of OCS by performing a series of analyses in a homogeneous, population-based sample from the Netherlands. First, phenotypic and genetic longitudinal correlations over a 6-year period were estimated by modeling of OCS data from twins and siblings. Second, polygenic scores (PRS) for 6931 subjects with genotype and OCS data were calculated based on meta-analysis results from Stewart et al. (2013), to investigate their predictive value. Third, we used random-effects modeling to estimate the contribution of measured SNPs to the heritability. Lastly, we performed an exploitative GWAS of OCS, testing for SNP- and gene-based associations. Stability in OCS (0.63) was mainly explained by genetic factors. PGS obtained from a European case-control GWAS predicted OCS in the population-based sample. Common genetic variants explained 14% of OCS. GWAS showed one SNP (rs8100480), located within the MEF2BNB gene, to be associated with OCS (P=2.56x10-8). Additional gene-based testing resulted in 4 significantly associated genes, located in the same chromosomal region (19p13.11): RFXANK, MEF2BNB, MEF2BNB-MEF2B and MEF2B. Common genetic variants explained a significant proportion of OCS trait variation. Genes significantly associated with OCS are expressed in the brain and involved in development and control of immune system functions (RFXANK) and the regulation of gene expression of muscle specific genes (MEF2BNB). MEF2BNB also showed a suggestive association with OCD in an independent case-control study, suggesting a possible role for these genes in the development of this disorder.

Introduction

Obsessive–Compulsive Disorder (OCD) is a neuropsychiatric disorder characterized by recurrent, persistent and intrusive anxiety-provoking thoughts or images (obsessions) and subsequent repetitive behaviours (compulsions) (APA, 2000). The lifetime prevalence of OCD has been estimated between 0.5-2.0%, and among all anxiety disorders, it is known as a major cause of social impairment and a leading cause of non-fatal disease burden worldwide (Ustun, 2004).

It is clear that genetic factors are important in the etiology of OC symptoms, with heritability estimated at 40% (Mathews et al., 2014; Taylor et al., 2011). These results did not vary with sex or symptom severity. Consistent with what is expected for individual SNP effect sizes in highly polygenic traits, the first molecular association studies for OCD have not identified large effect variants (Pauls, 2010; Stewart et al., 2013; Taylor, 2013). Compared to genetic association studies for other complex human traits, including psychiatric disorders, sample sizes and thus statistical power were not large. Therefore, Taylor and colleagues (Taylor, 2013) performed a comprehensive meta-analysis of the literature of genetic association studies in OCD, including 113 relevant studies. Their main meta-analysis showed that OCD was associated with serotonin-related polymorphisms (5-HTTLPR and HTR2A) and, in males only, with polymorphisms involved in catecholamine modulation (COMT and MAOA). In addition, a secondary meta-analysis by this group, which analysed polymorphisms that were investigated in fewer than five datasets, identified another 18 polymorphisms with significant odds ratios. These polymorphisms might be useful candidates for further investigation, although most included results were based on candidate gene studies and must be treated with great care due to the possible confounding effects of population stratification.

As a first attempt to identify the genetic variation predisposing to OCD at genome-wide level, The international OCD foundation Genetics Collaborative (IOCDF-GC; Stewart et al., 2013) (Stewart et al., 2013; Taylor, 2013), conducted an ancestry-stratified case-control genome-wide association (GWA) analyses, containing 1465 cases, 5557 ancestry-matched controls, and 400 trios, consisting of one affected offspring with two parents. The trio analysis revealed a significant single nucleotide polymorphism (SNP) near the BTBD3 gene (rs6131295). However, this variant lost genome-wide significance when meta-analysed with the case-control data,. In addition, meta-analysis showed a significant enrichment of methylation quantitative trait loci (QTLs) ($P<0.001$) and frontal lobe expression for the top-ranked

SNPs ($P<0.01$). Recently, a second multinational collaboration (OCD Collaborative Genetics Association Study; OCGAS; Mattheisen et al., 2014) performed a GWAS in 1598 patients and 3473 controls. None of the SNPs tested were significantly associated with OCD, with the smallest p-value observed for a marker near the PTPRD gene on chromosome 9. Gene-based testing for associations at the gene instead of SNP level, revealed a significant association of IQCK, C16orf88 and OFC11 (Mattheisen et al., 2014). These findings await replication.

With this study, we aim to gain further insights into the genetic basis of OC symptoms by performing a series of analyses in a homogeneous, population based sample of twin families, registered with the Netherlands Twin Registry (NTR). Well-phenotyped population cohorts may contribute to the understanding of the underlying architecture of common complex disorders. To ensure a sound OC symptom phenotype, we first performed genetic structural equation modelling (SEM) (Boker et al., 2011) to estimate twin–twin and twin–sibling correlations and heritability for OC symptoms in adults as measured with the Padua-Inventory-Revised-Abbreviated (PI-R-ABBR) (den Braber et al., 2010). Information on the PI-R-ABBR was available for two time-points (2002 and 2008), which allowed calculation of the stability of OC symptoms across a six-year period. The long-term stability of the phenotype puts an upper limit to heritability – i.e. reveals the proportion of total variation across time that is due to differences among individuals, and puts genetic association studies into perspective. Next, we investigated whether polygenic scores (PRS) based on a GWA analysis of clinical OCD cases (Stewart et al., 2013) significantly predict OC symptoms in our population-based sample. If so, this indicates genetic overlap between the two samples, and suggests that future GWA studies can benefit from the inclusion of both population-based and case-control samples. We estimated the proportion of phenotypic variance explained by all autosomal SNPs in a sample of unrelated individuals whose Genetic Relationship Matrix (GRM) was obtained in the Genome-wide Complex Trait Analysis (GCTA) software. Further, we performed an explorative GWAS on OC symptom scores from 6931 subjects and entered the GWAS output into a gene-based analysis, to test for associations at the gene rather than the single SNP level, in order to obtain more information on the biological meaning of the results (Li, Gui, & Sham, 2011).

Materials & Methods

Participants and measures

Study data were collected in participants of the NTR (Boomsma et al., 2006; Willemsen et al., 2013). We analysed data from the PI-R-ABBR collected in 2002 and 2008. The total sample with phenotype data from either the 2002 and/or 2008 data collection contained 20.376 individuals from 7.812 families. A total of 10.134 individuals had sum scores available for the PI-R-ABBR collected in 2002 and 15.720 individuals had sum scores for the PI-R-ABBR collected in 2008. Longitudinal data were available for 5.478 individuals.

The distribution of the OC symptom scores, from the PI-R-ABBR collected in 2002 and 2008, is provided in Supplementary figure 1. For more information on the PI-R-ABBR OCS measures, please see supplementary methods "participants and measures". The number of participants differed between analyses. Table 1 gives an overview of subjects per analysis and their demographic data (see also supplementary methods "Subjects entered into different genetic analyses" for more details and Zilhao et al. (2014)).

Table 1. Subjects (including age and sex) entered per analysis

	N		Sex (M/F (%))		Age (Mean (SD))	
	PI-R-ABBR 2002	PI-R-ABBR 2008	PI-R-ABBR 2002	PI-R-ABBR 2008	PI-R-ABBR 2002	PI-R-ABBR 2008
SEM (twins and sibs)	5780	9147	33.7/66.3	31.9/68.1	33.0(11.5)	34.7(14.6)
PRS (genotyped sample)	6931		35.7/64.3		42.8 (15.7)	
GCTA (genotyped sample)	6881		38.0/62.0		46.7 (15.4)	
GWAS (genotyped sample)	6931		35.7/64.3		42.8 (15.7)	

Note: Table 1. PI-R-ABBR =Padua Inventory Revised Abbreviated; PI-R-ABBR 2002 = PI-R-ABBR collected in 2002; PI-R-ABBR 2008 = PI-R-ABBR collected in 2008; SEM = structural equation modeling; PRS = Polygenic score; GCTA = Genome-wide Complex Trait Analysis ; GWAS = Genome-wide association study.

This study has been approved by the Central Ethics Committee on Research involving human subjects of the VU University Amsterdam.

Genotyping and Imputation

DNA was extracted from either blood or buccal cell samples that were collected for various projects done by the NTR (Boomsma et al., 2006; Willemsen et al., 2013). For further details on genotyping, quality control and imputation methods see supplementary methods "genotyping and imputation".

Heritability Estimates from Structural Equation Modelling

To estimate the total contribution of genetic factors to trait variance and to the longitudinal covariance, the resemblance among twins and siblings was compared. Monozygotic (MZ) twins share (nearly) all their genes, whereas Dizygotic (DZ) twins, just like non-twin siblings, share on average half of their segregating genes. In quantitative genetics, this information is used to decompose the total variance of a trait into additive genetic (A), non-additive genetic (D=dominance), and unshared environmental variance (E). The greater the phenotypic similarity between MZ twins, when compared to DZ twins and non-twin siblings, the more of the variance of the trait is attributed to genetic factors. Genetic analyses were carried out by Structural Equation Modeling (SEM) as implemented in OpenMx (Boker et al., 2011). For further details on SEM analysis may be found in supplementary methods "SEM methods".

Polygenic risk scores

To examine shared polygenic risk at an aggregate level between two independent GWAS samples, we used genetic risk-score profiling as described by Purcell and colleagues (Purcell et al., 2009). GWAS results (SNP, effect allele, effect size as represented by a Beta-value and p-value) obtained from the analysis in the European case-control sample by Stewart et al. (2013) was used as a discovery dataset for calculating polygenic scores (PGS) within our NTR target sample in PLINK. PGS were then regressed against the OCS scores from the NTR dataset (n=6931) in order to calculate the proportion of variance in this target set explained by PGS obtained from the discovery set, with 15 statistical cutoffs for SNP inclusion in the score (cutoffs: $p<0.00001$, $p<0.0001$, $p<0.001$, $p<0.01$, $p<0.05$, $p<0.1$, $p<0.2$, $p<0.3$, $p<0.4$, $p<0.5$, $p<0.6$, $p<0.7$, $p<0.8$, $p<0.9$, $p<=1$). To correct for family relatedness in the NTR sample generalized estimating equations in SPSS with the exchangeable and robust function were used (Minica, Dolan, & Vink, 2014). For Sex, age, principal components to correct for the population substructure (Abdellaoui et al., 2013) and genotyping platform were included as covariates. As an additional test, we regressed the PGS on a non-psychiatric, and OCS-uncorrelated trait (height), from the NTR (n=6715). We sought to rule out a possible spurious association that can arise as a consequence of incomplete correction for population stratification and/or cryptic relatedness between the discovery and target samples. Finally, to test for concordance of effect directions across both datasets, we performed Fisher's exact statistical test using the online-based application SECA (SNP effect concordance analysis) - http://neurogenetics.qimrberghofer.edu.au/SECA/ (Nyholt, 2014). For 12

subsets of SNPs with P < {0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.8, 0.9, 1.0} in both datasets, Fisher's exact tests were performed to evaluate if there is an excess of SNPs in the first dataset (European case-control study) with same direction of effect in the NTR dataset across the total 144 SNP subsets. An empirical P-value was generated by permutations (1000) for observing the number of SNP subsets with nominally significant concordance.

Estimations of variance explained by common SNPs (GCTA)

The variance of OC symptoms explained by measured and imputed SNPs was estimated with Genome-wide Complex Trait Analysis (GCTA) using the Restricted Maximum Likelihood analysis procedure (Yang, Lee, & Visscher, 2011, 2013). This method gives insight into the contribution to the additive genetic variance of all SNPs tagged by all genotyped and imputed SNPs. This provides an upper bound on the variance that can be explained by the set of SNPs. To analyse all available data we followed the method proposed by Zaitlen et al (2013). In this method, the SNP based and kinship-based heritability can be estimated simultaneously. For this, a second variance component is included in the model, which only considers closely related individuals for a certain predefined threshold of genetic relatedness (here defined only those with Identity by descent values greater than 0.05). For further details on the method followed and quality control, refer to supplementary methods "estimations of variance explained by common SNPs (GCTA)".

GWAS

Genome-wide association (GWA) analysis was conducted with linear regression under an additive model with adjustment for age, age-squared, principal components of genetic ancestry, genotyping platform and sex. SNPs with values of p<5.00E-08 were declared genome-wide significant. For further details on this analysis, refer to supplementary methods "GWAS".

Gene-based Analysis

GATES, as implemented in the open-source tool Knowledge-Based Mining System for Genome-Wide Genetic Studies (KGG, version 3.0), was used to perform a gene-based genome-wide analysis (Li et al., 2011). GATES employs an extension of the Simes procedure (Benjamini & Hochberg, 1997) to assess the significance of a statistical association at the gene level, by combining the individual genotype-phenotype association tests applied at each single marker. In short, it sums all the individual SNP p-values, available from GWAS summary data, within a gene to output a gene-based p-value. The effective number of independent p-values is given by appropriately controlling for the

Linkage Disequilibrium structure on the specified SNPs. SNPs were allocated to genes including gene boundaries of +/- 5 kb from the 5'and 3'UTR. To correct for multiple testing false discovery rate (FDR) was set at $q<0.05$.

Results

Demographics

For subjects who completed the PI-R-ABBR in 2002 and in 2008 the average age was 33.0 (SD=11.5), and 34.7 (SD=14.6). Retest stability for OCS scores over a timespan of ~6 years was 0.63. The effect of age on OCS scores was a drop of 0.03 per year for both PI-R-ABBR collected in 2002 and 2008.

Genetic Modelling (SEM)

MZ, DZ/sibling, test-retest and cross-correlations are summarized in table 2. Twin correlations for MZ males and MZ females were equal, as were twin/sibling correlations for DZ males, DZ females, DZ opposite-sex twins and siblings. This indicates that there is no evidence for qualitative sex differences in the heritability of OC symptoms and that to a large extent the same genes influence these symptoms in males and females. Twin correlations were more than twice as large in MZ as compared to DZ/sib pairs, indicating that phenotypic similarity is predominantly accounted for by genetic effects rather than shared environment. The same pattern was observed for cross-twin–cross-time correlations, indicating that the observed stability is also mainly caused by genetic factors. Structural equation modelling showed a significant heritability ($p < 1.0 \times 10\text{-}10$) for OC symptoms at both time points of 0.42 (95% CI PI-R-ABBR 2002 = 0.371 – 0.467, and 95% CI PI-R-ABBR 2008 =0.379 – 0.456). The estimation of the bivariate (broad sense) heritability found that 56% (95% CI = 0.497 – 0.619) of the stability of OC symptom was due to genetic factors (both additive and dominant components), and the longitudinal additive genetic correlation, that is, the degree of overlap between additive genetic influences at both time-points was 0.58.

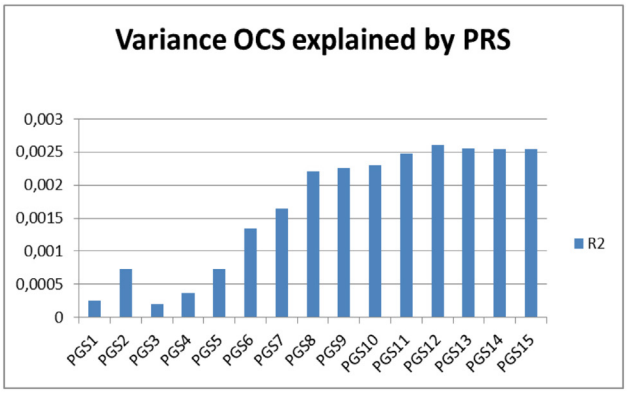
Table 2. Familial correlations estimated from Maximum likelihood for OC symptoms measured over 2 points in time.

	Twin correlation	Cross-twin - cross-time correlation	Retest stability (Within person)
MZ	0.41	0.36	0.63
DZ/siblings	0.13	0.11	

Note: Table 2. MZ = monozygotic; DZ = Dizygotic.

PRS regressed on PI-R-ABBR sum scores

The proportions of variance of OC symptoms in the NTR sample explained by polygenic scores obtained from the discovery set (European case control sample Stewart et al. 2013), using the 15 statistical cut-offs for SNP inclusion in the score, are summarized in figure 1. Results show that when including more SNPs in the analyses an increasing amount of variance in OC symptoms is explained; reaching a plateau of 0.2% explained variance ($p<0.001$) at PRS12 (includes all SNPs from Stewart et al., 2013 with a $p<0.7$ [N=4.288.152]). Supplementary figure 2 provides the results obtained when using height as the outcome phenotype for the target set for the same PGS. For the same 15 statistical cut-offs, no significant result was observed ($p < 0.01$), i.e. the PGS did not explain any variance in height, hence excluding the confounding effect due to cryptic relatedness across sample sets and possible residual genetic stratification effects present in both populations. Finally, supplementary figure 3 presents the results in a heat map plot from analysis of concordance using SECA. The permuted P-value for the number of SNP subsets nominally significant was P=0.002 thus indicating significant concordance of genetic risk across the datasets.



PRS	PRS cutoff	P-value
1	0 - 0,00001	0,32315
2	0 - 0,0001	0,05815
3	0 - 0,001	0,50025
4	0 - 0,01	0,29122
5	0 - 0,05	0,03754
6	0 - 0,1	0,00426
7	0 - 0,2	0,00152
8	0 - 0,3	0,00026
9	0 - 0,4	0,00025
10	0 - 0,5	0,00022
11	0 - 0,6	0,00013
12	0 - 0,7	0,00010
13	0 - 0,8	0,00011
14	0 - 0,9	0,00012
15	0 - 1	0,00012

*Figure 1. Proportion of variance in OC symptom scores, as measured in the NTR sample, explained by polygenic scores (PRS) obtained from European case-control sample by Stewart et al. 2013, with a range of 15 statistical cut-offs for SNP inclusion in the score (PRS1; p<0.00001, PRS2; p<0.0001, PRS3; p<0.001, PRS4; p<0.01, PRS5; p<0.05, PRS6; p<0.1, PRS7; p<0.2, PRS8; p<0.3, PRS9; p<0.4, PRS10; p<0.5, PRS11; p<0.6, PRS12; p<0.7, PRS13; p<0.8, PRS14; p<0.9, PRS15; p<=1). *=p<0.05; **=p<0.01.*

SNP-based Heritability

GCTA results showed a significant SNP-based heritability estimate of 0.14 (SE=0.05, p=0.003), and a total narrow-sense heritability of 0.338 (SE=0.02, p<0.001). Thus 14% of the total phenotypic variance in OCS can be accounted for by the SNPs in the genotyping platform, and the total set of SNPs included in an additive genetic model account for 33.8% of the total heritability.

Genome-wide Association Analysis

Top associated variants in GWAS analysis (top 20 SNPs) are summarized in table 3 (a more comprehensive overview of these results is present in supplementary table 1). The Manhattan plot, showing the -log(P) plotted against genomic location, and QQ plot of observed versus expected -log(P) statistics for the OC symptom GWAS, are illustrated in figure 2 & supplementary figure 4, respectively. Of the top associated variants, one SNP, rs8100480 (19299079, 19p13.11 (hg19), P=2.56x10-8), exceeded the threshold for genome-wide significance and showed a positive association with OCS. This SNP is located within the MEF2BNB gene. A more detailed look into this region is provided by the regional association plot in Supplementary Figure 5.

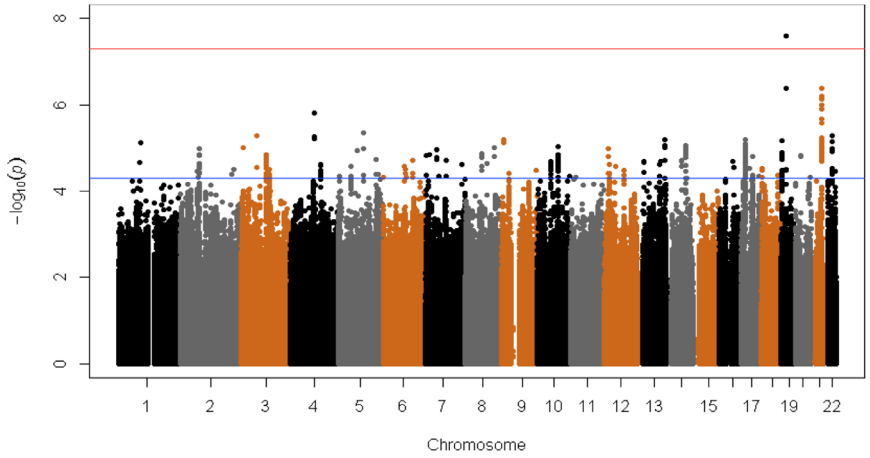


Figure 2. Manhattan plots of all genotyped single-nucleotide polymorphisms (SNPs). Red and blue lines indicate significance thresholds of 5.00E-08 and 1.00E-05, respectively

Table 3. Top associated variants in NTR-GWAS analysis.

SNP	CHR	BP(hg19)	P-value	A1/A2	BETA	Left gene(kb)	Intragenic location	Right gene(kb)
rs8100480	19	19299079	2.56E-08	C/T	1.4095		*MEF2BNB*	
rs11671119	19	19286077	4.11E-07	C/T	1.2036		*MEF2BNB-MEF2B*	
rs4818048	21	40908952	4.24E-07	C/G	0.8956	*LOC729056 (7173)*		*B3GALT5(19417)*
rs4818050	21	40910600	6.37E-07	C/T	0.8725	*LOC729056 (8821)*		*B3GALT5(17769)*
rs77959192	21	40911027	7.36E-07	C/A	0.8707	*LOC729056 (9248)*		*B3GALT5(17342)*
rs4818049	21	40910464	1.03E-06	G/A	0.8557	*LOC729056 (8625)*		*B3GALT5(17965)*
rs77615161	21	40911050	1.30E-06	G/A	0.8541	*LOC729056 (9271)*		*B3GALT5(17365)*
rs17384439	4	96424680	1.59E-06	T/C	0.8323		*UNC5C*	
rs2837096	21	40978013	2.18E-06	G/A	0.747		*C21orf88*	
rs4818052	21	40912745	2.64E-06	A/G	0.8475	*LOC729056 (10966)*		*B3GALT5(15670)*
rs77460585	5	101123995	4.45E-06	G/A	0.8668	*LOC100420593*		*OR7H2P*
rs581043	3	62830115	5.26E-06	C/T	0.4358		*CADPS*	
rs999719	22	34264838	5.33E-06	T/A	0.4442		*LARGE*	
.	4	96399160	5.57E-06	R/D	0.7201			
rs74276709	21	40913995	5.85E-06	A/G	0.8226	*LOC729056*		*B3GALT5*
rs17024030	4	96399606	6.07E-06	G/A	0.7155		*UNC5C*	
rs9520326	13	107865442	6.37E-06	T/C	-0.4326		*FAM155A*	
rs79219884	21	40899981	6.40E-06	A/T	0.8467		*LOC729056*	
rs60588302	9	7900777	6.44E-06	C/T	1.1278	*C9orf123* (100971)		*TPRD* (413469)
rs11658311	17	17470526	6.50E-06	C/T	0.7719		*PEMT*	

Note: Table 3. Single Nucletotide Polymorphisms (SNP) listed include the top 20 P-values for the GWAS results. Chromosome (Chr) and base pair position (BP), based on hg19 build, are also given. The beta indicates the effect size, and the direction of the association is given by its positive or negative value. The location of each SNP is given in the last column; when located in non-intronic locus, the left and right closest flanking genes are additionally noted. A1 and A2 indicate the effect allele and the non-effect allele, respectively.

Further, we searched in our results whether they replicate top SNPs reported by Stewart et al., 2013 and Mattheisen et al., 2014. Supplementary table 2 summarizes results for the strongest associated GWAS variants from Stewart et al, 2013. From their table of 43 SNPs (with P<1.0x10-5), 16 were found to be independent (Mattheisen et al., 2014). None of these SNPs were significantly replicated in our sample when correcting for multiple comparison (0.05/16=0.003). However, a trend (p=0.0049) for replication was found for rs4868342, located on chromosome 5. This SNP is located within the HMP19 gene.

Supplementary table 3 summarizes results for the strongest associated GWAS variants from Mattheisen et al. (2014). None of the 32 suggestive associations (with P<1.0x10-4) were replicated in our sample after correcting for multiple comparison (0.05/32=0.001). Neither were the significant results obtained in our study replicated in the Stewart sample (data not shown).

Gene-based analysis

A total of 2.644.694 SNPs were mapped to 22.759 genes. The QQ plots with the observed versus expected –log(P) of the association tests are presented in supplementary figure 6. Table 4 depicts all the genes with a significant association. Although these are all nominal significant, the

Benjamin-Hochberg procedure was set to control for the q<0.05 (Benjamini & Hochberg, 1997). After correction, four genes remained significant, the regulatory factor X-associated ankyrin-containing protein (RFXANK), the myocyte enhancer factor 2B (MEF2B), the MEF2B neighbor (MEF2BNB), and the MEF2BNB-MEF2B read through (MEF2BNB-MEF2B). All these genes are located in the same chromosomal region (19p13.11), and they all share the SNP with the lowest P-value annotated to the gene (rs8100480, P=2.56E-08). The Manhattan plot for the gene association P-values is present in supplementary figure 7.

Table 4. List of all nominal significant (genes α =0.05). Genes significant after FDR corrections are depicted with a (*)

Gene	Chromosome	Start position	Length	Number of SNPs	Gene *P* value
RFXANK *	19	19303574	9104	10	5.60E-07
MEF2BNB *	19	19292684	10716	40	9.72E-07
MEF2BNB-MEF2B *	19	19256375	47025	73	1.29E-06
MEF2B *	19	19256375	24723	41	8.08E-06
C21orf88	21	40969074	15675	14	8.22E-05
SAFB2	19	5587009	35929	22	8.97E-05
PEMT	17	17408876	71903	73	1.02E-04
SIAE	11	12450568	40515	1	2.94E-04
ATP5G1	17	46970147	3085	18	3.47E-04
CCL2	17	32582295	1925	1	3.48E-04
GOPC	6	11788143	42273	17	3.61E-04
TMEM63C	14	77648101	77737	90	4.25E-04
DKFZP434H168	16	56226528	1909	1	4.51E-04
UNC5C	4	96083655	38670	43	5.20E-04
UBE2Z	17	46985730	20692	70	5.43E-04
SMCR5	17	17679999	2844	8	5.58E-04
C19orf70	19	5678432	2479	1	6.01E-04
SNF8	17	47007458	14696	59	6.99E-04
SAFB	19	5623045	45444	4	7.25E-04
LOC101929154	5	77180479	74441	20	7.27E-04
MIR602	9	14073287	98	2	7.34E-04
MIR128-2	3	35785967	84	1	7.85E-04
NCAN	19	19322781	40280	25	7.89E-04
RPL36	19	5690271	1407	3	8.49E-04
CECR1	22	17659679	20994	28	8.61E-04
HSD11B1L	19	5680775	7759	4	9.12E-04
RP11-461O7.1	16	56126898	98108	10	9.98E-04

Note: Table 4. Genes listed include all the nominal significant genes (α= 0.05). Significant genes after FDR correction are depicted with (*). For each gene, the chromosome, start position and respective gene length (in base pairs) are given. The number of SNPs from the GWAS within each gene are also present. The last column represents the calculated gene-based test P-value.

Discussion

This study aimed at getting a better insight into the genetic basis of OC symptoms in a population-based sample.

First, in line with previous studies, we estimated the heritability for OC symptoms at 0.42 (S Taylor, 2013). Stability of OCS over a 6 year time period was 0.63, and cross-twin–cross-time correlations were found to be twice as high in MZ compared to DZ/sib pairs, indicating that the observed stability is mainly caused by genetic factors. Bivariate analyses showed a longitudinal genetic correlation of 0.58.

Second, polygenic scores based on a GWA analysis of clinical OCD cases significantly predicted OC symptoms in the independent population-based sample. Also, analysis of concordance results indicates that the genetic risk between the two datasets are concordant. This suggests a high polygenic contribution to the trait and indicates genetic overlap between OCD assessed as a categorical disorder and OC symptoms assessed as a continuous trait. These results indicate that future GWA studies can benefit from the inclusion of both population-based and case-control studies, and by analysing OCD as a quantitative rather than a categorical trait.

Third, we estimated the SNP based heritability in the sample of genotyped individuals, In a previous study by Davis and colleagues (Davis et al., 2013) the genetic variance in OCD explained by SNPs was estimated at 0.37 In accordance with this, we were able to find significant explained variance at 0.34. More importantly we were able to partition the heritability into two components, revealing a SNP heritability of 0.14 and the proportion of the heritability not accounted for by SNPs (so called 'missing heritability') at around 0.20. These results corroborate well with our findings from our twin model (broad heritability at 0.42).

Fourth, the GWAS performed on our continuous OC symptom scores resulted in one significant SNP (rs8100480), which is located in the MEF2BNB-MEF2B gene. In addition, we sought to calculate gene-based p-values, to determine whether there are genes with associated SNPs, which can collectively achieve statistical significance. Implementing a gene-based analysis as a follow-up complementary analysis seems to be of additional value over traditional GWAS results, since gene-based analysis takes the number of SNPs in each gene and gene size into account, considering genes as functional units informing on the underlying genetic architecture of the phenotype under review. Furthermore, since for most psychiatric disorders we do not expect large effect causal variants, replication analysis from

underpowered GWAS would not reflect these findings. Gene-based tests, by providing an aggregated analysis, may successfully capture those combined effects. We found significant associations for four genes. RFXANK encodes a protein that belongs to the Major histocompatibility (MHC) class II molecules, which has an important role in the development and control of the immune system. Bare lymphocyte syndrome 2, an immunodeficiency disorder, has been linked with mutations in this gene. Although RFXANK has not been identified in OCD previously, several lines of research have indicated a role of immune system alterations and of immune system genes, including TNF**α**, and SLC1A1, in combination with cerebral immunopathological reactions to amongst others group A ß-hemolytic streptococcus infections (Swedo et al., 1998) and Bornea virus infections (D Marazziti et al., 2009; Marazziti et al., 1999) in the onset (and expression) of OCD. This has led to the concept of "PANDAS" and "non-PANDAS" OCD (Cappi et al., 2012; da Rocha, Correa, & Teixeira, 2008; Rotge et al., 2010). However, very little research has been performed on direct gene*gene interactions/pathways or on gene*environment interactions to better understand immunopathological pathways related to OC symptoms. MEF2B is a protein-coding gene belonging to the DNA binding protein family MADS/MEF2, that regulates gene expression, specifically in the smooth muscle tissue. Both MEF2BNB and MEF2BNB-MEF2B are closely related to MEF2B, and have mostly regulatory functions. Additional support for a relation between MEF2BNB and OCD comes from a gene-based analysis in a recently published GWAS of OCD (Mattheisen et al., 2014) where the gene was ranked 21st of the 21.567 genes tested (gene-based p = 8.09E-04, N.S. after correction for multiple testing). All four genes are overlapping between themselves, and span a region of 56302 bp located in the 19p13.11 cytogenetic band, on the short arm of chromosome 19. Perhaps the best interpretation for these results, therefore, is in the implication of this genomic location (19p13.11) as a susceptibility locus for OC symptoms. Several SNP-trait associations have been linked to this locus. For example, rs1064395 (NCAN) has been reported as a susceptibility factor for bipolar disorder in a genome-wide association study. NCAN gene is located just 10095 bp at 5' from RFXANK, and is one of the few genetic variants that has been genome-wide replicated as a risk factor in both bipolar disorder and schizophrenia (Cichon et al., 2011; Mühleisen et al., 2012). In addition, a recent study focusing on the cortical thickness and folding in schizophrenia patients found evidence for association of the NCAN genetic variant in the occipital and prefrontal cortex (Schultz et al., 2014). The SNP rs874628 (MPV17L2 gene), located in this locus, was implicated in multiple sclerosis, an inflammatory disease with disruptions in the nervous system

(Sawcer et al., 2012).

Overall, our results, combined with previous genetic studies in OCD, suggest a possible role for the 19p13.11 region (MEF2BNB gene) in OC symptoms. It might be of interest for future genetic studies to investigate this area in association with OCD into depth. Further, our data shows that well phenotyped population cohorts could contribute to the understanding of the underlying architecture of common complex disorders such as OC symptoms, and that these partly overlap with results from case-control studies. Therefore, future studies could benefit from combining case-control and population-based samples.

The Supplementary material can be found in the online version of this manuscript

Chapter 6
Genome-Wide Association Study of Tics and Tourette Disorder Symptoms

This chapter is based on: Zilhao, N. R., Smit, D. J. A., Cath, D., Hottenga J. J., Boomsma, D. I. (2017). Genome-wide association study of tics and Tourette Disorder Symptoms. European Journal of Human Genetics (under review).

Abstract

Tic disorders are common neuropsychiatric diseases with a familial inheritance, but specific genetic causal variants remain elusive. Here, we present results from a genome-wide analysis on Tourette Disorder (TD) from a population-based family sample from the Netherlands (NTR-TD: N=88 cases; 6381 controls), and meta-analyzed these with those from the Psychiatric Genomics Consortium Tourette's workgroup (PGC-TS: N=778 cases; 4414 controls). NTR top-ranking SNP (rs73193806) (p=2.13e-08) is located in the same region (12q23) as reported among the top results from the PGC-TS. LD score regression rendered a SNP-based heritability of 0.146 (se=0.066) for the NTR-TD, in line with what has been reported for tic phenotypes. In the meta-analysis the top SNP was rs7783290 (p=1.49e-07), located on chromosome 7p21.3, 600Kbp from the closest gene, NDUFA4. In conclusion, we estimated a SNP-based heritability of 14.6% indicating that a substantial part of the total TD heritability can be attributed to SNPs that are relatively common. We have extended previous work on the genetics of TD and shown that combining clinical- and population-based samples is of added value.

Introduction

Tics are defined as involuntary, sudden, recurrent, non-rhythmic and stereotypical motor movements or vocalizations. The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) (American Psychiatric Association, 2013), recognizes three types of tic disorders: Tourette's Disorder (TD), Chronic Motor or Vocal Tic Disorder (CMT/CVT) and Provisional tic disorder, with the diagnosis depending on age of onset, duration, and tic type (motor, vocal or both).

Both twin and family studies have indicated that TD is heritable, with first-degree relatives of TD-affected probands having a 5-15 fold increased risk for TD compared to the general population (Pauls, Raymond, & Leckman, 1991; Robertson, 2008). Epidemiological studies estimated heritability ranging from .28 to .56 (Anckarsäter et al., 2011; Bolton, Rijsdijk, & Eley, 2007; Larson et al., 2010). Yet, studies seeking specific genetic variants for the susceptibility of this disorder have thus far rendered only suggestive results. Candidate gene and CNV studies in TD have mainly been positive on SLITRK1, HDC, DRD4 and AADAC. However, none of these studies were followed by consistent replications, and were hampered by small sample sizes (Abelson et al., 2005; Bertelsen et al., 2014; Ercan-Sencicek et al., 2010; Liu et al., 2014). The only genome wide association study (GWAS) in TD so far, performed by the Tourette Syndrome Association International Consortium for Genetics (TSAICG), from a sample of 1285 cases and 4964 ancestry-matched controls did not identify genome-wide significant results. The top signal (rs7868992) was located in COL27A1 on chromosome 9q32 (p=1.85e-06) (Scharf et al., 2013).

Here, we present results from a GWAS on TD in a population-based family study including screened controls from the Netherlands. For these results, we estimated SNP-based heritability using LD score regression. Finally, we performed a meta-analysis combining results from our study and the study performed by the TSAICG, to investigate the added value of combining clinically-based and population-based datasets.

Materials & Methods

Participants and Measures

Participants in this study are adults from the Netherlands twin Register (NTR) who completed the 8th NTR survey between 2008-2012 (Willemsen et al., 2013). Data on tics were collected using the abbreviated Schedule for Tourette and Other Behavioral Syndromes (STOBS-ABBR - Supplementary Table 1), a self-report that contains 9 tic items and 3 additional questions on age at

onset, duration and tic severity, classifying participants according to DSM-IV-TR and in line with DSM5 criteria (American Psychiatric Association, 2013). Individuals identified as cases were diagnosed as having either 'probable CMT/CVT' or 'probable TD' according to the STOBS-ABBR. These individuals constitute a sample of screened controls as we only included those individuals with the most severe manifestation of a tic disorder, excluding those classified with a milder form of the phenotype (N=173) (Zilhão et al., 2016). A final total of 88 cases and 6381 controls (from 3338 different families,) were included in the analysis (see Supplementary Table 2 for an overview on the sample demographics). The study was approved by the Central Ethics Committee on Research involving human subjects of the VU University Amsterdam. Informed Consent was obtained from all subjects.

Genotyping

Genotyping and genotype calling was performed on multiple array platforms, and respective platform specific software, for a total of 21,775,581 SNPs included in the analysis. Imputation was performed from the HRC and 1000 Genomes Project Phase 3 (v5) reference panels (for information on Genotype, Imputation and Quality control steps please view Supplementary Methods). Post-analysis quality control steps involved removing all SNPs with INFO imputation score < 0.8 and SNPs with MAF>0.05. From a total of 21,775,581 SNPs, 6,789,596 were finally included in the analyses.

Association Analysis

For genetic association analyses we employed the method proposed by Chen et al. (Chen et al., 2016), freely available and implemented as the R-package GMMAT (https://www.hsph.harvard.edu/xlin/software.html). This method fits a generalized linear mixed model with random effects (genetic relationship matrix) to account for family relatedness, and subsequently performing a score test (Engle, 1983) for each tested genetic variant. The actual statistical test (score) is based on the slope of the likelihood function under the null hypothesis (H0) of no SNP effect. This slope (or 'score') is used to estimate the improvement in model fit when including the SNP of interest, defined as the (expected) difference in X^2 statistic. This is done using the penalized quasi-likelihood (PQL) method and the average information restricted maximum likelihood (AI-REML) algorithm, to test for its association with the binary phenotype (Gilmour, Thompson, & Cullis, 1995). Importantly, the score test is computationally more efficient than regular statistical techniques that require an estimate of the SNP effect coefficient. However, the score test is only valid for small effect sizes, but this is a reasonable assumption for

testing SNP effects.

SNP Heritability

Based on the GWA results, we estimated the SNP-heritability (probable TD or probable CMT/CVT) using linkage disequilibrium score regression (LDSR; (Bulik-sullivan et al., 2015). Following this approach, we estimated the heritability captured by the set of genotyped SNPs – i.e. the contribution to the additive genetic variance of all genotyped SNPs on the current panel. This was achieved by separately estimating the true polygenic signal and the confounding bias (e.g. population stratification), using Linkage disequilibrium (LD) information from a reference set (European-ancestry) as predictors for the X^2 statistic from the GWAS results (https://data.broadinstitute.org/alkesgroup/LDSCORE). The slope of this regression is an estimator of the contribution from true polygenecity to the inflation factor – heritability.

Meta-Analysis

We performed a meta-analysis, combining our results with results from the PGC-TS using the software METAL (http://genome.sph.umich.edu/wiki/METAL_Program). The summary statistics from the PGC-TS were based on the imputed dataset (N=778 cases; N=4414 controls), derived from a European-ancestry sub-population (Scharf et al., 2013). Following METAL software's approach, p-values (and respective effect direction) from both studies were transformed into signed Z-scores and weighted to their sample size, reflecting higher or lower disease risk. From a combined total of 8,144,975 SNPs, 5,103,655 were meta-analyzed (i.e. in common between datasets).

Results & Discussion

No single SNP reached the genome-wide significance level ($p<5.0e\text{-}08$); we note, however, that at the given sample size genome-wide significance was not expected. The QQ-plot (Figure 1) indicates that there is little inflation in the observed results, and that the model properly accounted for population stratification theoretically present in the data, or any systematic technical artifacts.

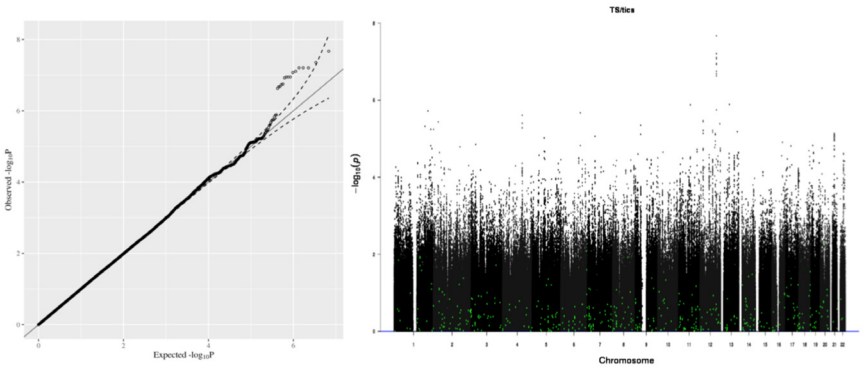


Figure 1. QQ-plot of expected vs observed –log(p-values) from the GWAS analysis (left). Manhattan plot for all 5595601 SNPs included in the analysis, for 88 cases and 6381 controls. Green dots represent the top 55 SNPs from the previous published TD GWAS (Scharf 2013, Supplementary Table S2) (right).

The top SNP rs73193806 (p=2.13e-08) is located on chromosome 12q23.3, 2055 bp from ST13P3 gene. Supplementary Table 3 shows results from the association analysis, with a p<1.0e-05. The top 13 SNPs span a region of 43,236 bp and are in high LD with each other (Figure 2, right). Interestingly, Scharf and colleagues reported a SNP (rs6539267, p=7.41e-06) among its top hits, located within the POLR3B gene in the region 12q23 (Scharf et al., 2013). Our top hit, rs73193806, is located in the same region, 350kb downstream of POLR3B (Figure 2, right).

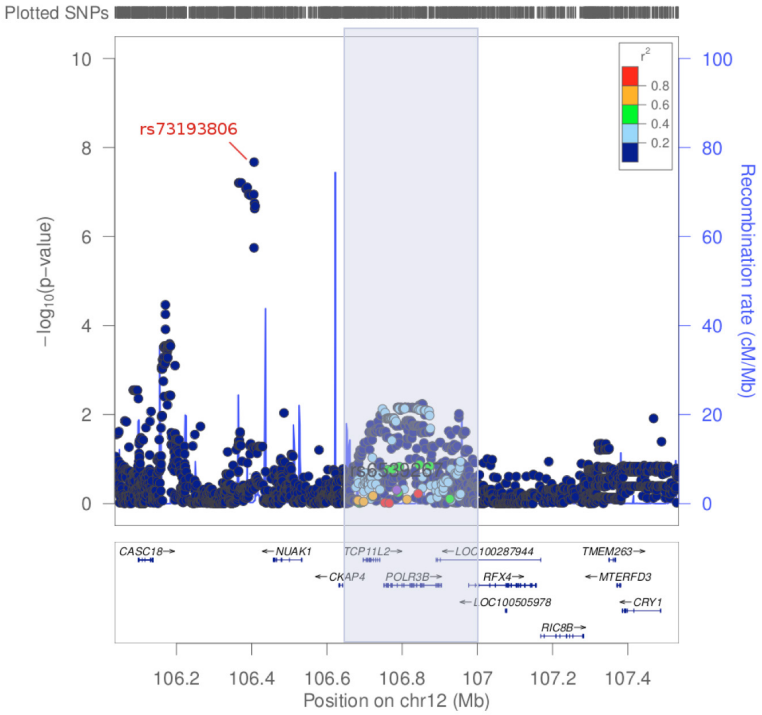


Figure 2. Regional plots showing the regional association for selected SNPs using the tool LocusZoom (http://locuszoom.org/); highlighting the region flanking the top SNP obtained in the study by Scharf (2013). Our top SNP rs73193806 (in red) is located 345240 bp from POLR3B in the highlighted area.

Based on the NTR-TS GWA results, we estimated the SNP-based heritability at 0.146 (se=0.0658). Previously Davis and colleagues estimated SNP-heritability for TD at 0.58 (Davis et al., 2013), which seems to be a high estimate in light of recent genetic epidemiological work estimating TD and Chronic tic disorder (CTD) heritability at 0.25, based on NTR twin-family data (Zilhão et al., 2016). This study showed (narrow-sense) heritability estimates between 0.25-0.37 across a spectrum of severity using 4 different definitions of tic disorders. The lower boundary of 0.25 phenotypically (TS or CTD) matches our current result of 14.6% for SNP-heritability - aligning with the observation that most heritability estimates from twin-family studies are higher than from SNP-based studies (Vinkhuyzen, Wray, Yang, Goddard, & Visscher, 2013). However, our results originate from a population-based cohort, whereas Davis et al. analyzed data from a clinical cohort. The estimate of 0.146 indicates that roughly half of the total heritability can be attributed to SNPs that are relatively common.

Figure 3 shows the results from the meta-analysis. The QQ-plot indicates little inflation in the results. Among the 5,103,655 SNPs included in the

Meta-analysis, 2,548,950 were concordant for effect direction (1,272,166 for increased risk, and 1,276,784 for decreased risk). The summary statistics we utilized included a sub-population of European ancestry from the study presented by Scharf et al (2013). The meta-analysis presented consists of combining a clinical sample with a non-clinical population-based sample of screened-controls. Shown in the Manhattan plot, the top results from the PGC-TS provide expected higher relative weight in the Meta-analysis (Figure 3). The top SNP, rs7783290 located on chromosome 7, has a p-value of 1.49e-07 and a Z-score of -5.25. This SNP was the 3rd top ranking SNP in the GWAS from the PGC-TS (p=7.49e-07), and showed p=0.077 in the NTR-TD GWAS (Supplementary Table 4).

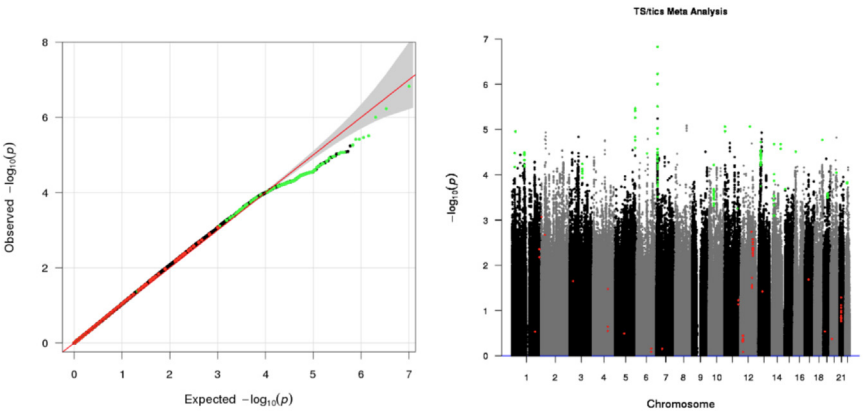


Figure 3. QQ-plot of expected vs observed –log(p-values) from the Meta-analysis between the PGC-TS and the NTR GWAS (left). Manhattan plot for all 5,103,655 SNPs; original results from each independent analysis are highlighted in green (for the TD consortium) and orange (for the NTR) (right).

We presented the second GWAS on probable TD performed to date and meta-analyzed the results with the first GWA carried out by Scharf (2013). We note that the lack of genome wide significant results from our study and the meta-analysis is likely due to the low power, well in line with observations for other traits within this field. The SNP-heritability estimate of 0.146 provides positive expectations for upcoming genetic analyses in TD, for which higher sample sizes may help uncover causal variants underlying the genetic etiology of this disease. Importantly, we show relevance of combining clinical and population-based samples in context of GWA for tic disorders.

The Supplementary material can be found in the online version of this manuscript

Chapter 7
Genome-Wide Association Study of Hoarding Behaviour: A Meta-Analysis in 2 Cohorts

Hoarding disorder (HD) is defined as disorder characterized by a pathological acquisition and difficulty in discarding possessions of relatively low value. This leads to severe clutter and social impairment, precluding normal daily living and activities for which living spaces were originally designed (APA, 2013; Frost & Hartl, 1996). For a long time, HD was considered a form of obsessive-compulsive disorder (OCD), and there is a debate whether HD constitutes a separate mental disorder rather than a symptom of OCD. In both epidemiological and clinical samples, OCD and HD show comorbidity, with up to 16.3% OCD in individuals with suggestive HD, and reversely, up to 24% of individuals with OCD scoring above the cut-point suggestive for HD in the general population (Frost, Steketee, & Tolin, 2011; Ivanov et al., 2013; Mathews et al., 2014). Interestingly, other co-occurring disorders with HD, i.e. ADHD, Major Depressive Disorder and Generalized Anxiety Disorder (between 25-50%) are more common in HD than OCD (Mataix-Cols et al., 2010). Partly as a consequence of these results, in the latest version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), HD is positioned as a separate disorder, at the same time being part of obsessive-compulsive spectrum disorders (APA, 2013). HD has a genetic basis, as established by multivariate twin studies (Iervolino et al., 2011; Ivanov et al., 2013; Mathews et al., 2014; Zilhao et al., 2016). Zilhão, et al. (2016) estimated the genetic correlation between OC symptoms and hoarding symptoms at 0.41. This genetic correlation was significantly lower than 1 and corroborates the notion of HD being (at least partly) distinct from OCD (Zilhao et al., 2016).

Possibly as a consequence of the past positioning of HD within the framework of OCD in DSM-III and IV classification systems, there has been a scarcity of studies on HD as a separate disorder. With respect to the genetic epidemiology of HD, current evidence points to familial transmission of HD (Lochner et al., 2005; Pertusa et al., 2008; Winsberg, Cassic, & Koran, 1999), with twin studies estimating the heritability between 0.33-0.50 in adults (Iervolino et al., 2009; Ivanov et al., 2013; Nordsletten et al., 2013; Taylor, et al., 2010; Yu et al., 2014; Zilhao et al., 2016).

The first GWAS of hoarding behavior was published in 2010, in a population based sample (N=3410 independent persons) from the TwinsUK study, consisting of predominantly female participants (91.8%) (Perroud et al., 2011). Although no SNPs with genome-wide significance were reported, results revealed two loci with suggestive evidence for association with HD. Here, we present the second genetic association study, which entails a meta-analysis on HD, combining two population-based samples; one from the Netherlands

Twin Register (NTR), and the other one from the previously mentioned TwinsUK sample partly overlapping with the published GWAS (Spector & Williams, 2006; Perroud et al., 2011). In both samples, we allow data from related individuals to be analysed, using appropriate methods (Purcell et al., 2007), thereby increasing the sample size for TwinsUK as compared to the previous GWAS. The studies included 6521 participants from NTR and 5190 from TwinsUK, who were phenotyped by the Hoarding Rating Scale-Self-Report (HRS-SR; (Tolin, et al., 2010) and who had genome wide genotyped data, imputed on the 1000G, phase 3 (v5).

Participants in this study are adults from the Netherlands twin Register (NTR) who completed the 8th NTR survey between 2008-2012 (Geels et al., 2013; Willemsen et al., 2013) and from the TwinsUK as part of a larger wave of data collection (Moayyeri, Hammond, Hart, & Spector, 2013). Data on hoarding were collected using the Hoarding Rating Scale-Self-Report (HRS-SR), a self-report assessment instrument consisting of five items each scoring on a 0-8 scale, producing a total score range between 0 and 40. It assesses cluttering, difficulty in discarding items, excessive acquisition or collecting, distress derived from hoarding symptoms and functional impairment (Tolin et al., 2010). For both samples the distress item (item 4) was discarded, due to approval restriction on the items to be included in the larger questionnaire (Cath, Nizar, & Mathews, 2017). Therefore, phenotypic scores were based on a 4-item scale. The GWAS analyses were performed on a continuous score for hoarding ranging from 0 to 32. In total, data from 6521 (N=2237 males; N=4284 females) individuals from NTR and 5190 from TwinsUK (N=596 males; N=4594 females) were included in the analysis.

Genotyping and genotype calling was performed on multiple array platforms, and respective platform specific software. Imputation was performed from the HRC and 1000 Genomes Project Phase 3 (v5) reference panels (for information on Genotype, Imputation and Quality control steps please view Supplementary Methods). After all Quality control steps were performed, a final total of 7,666,767 and 9,618,499 SNPs were included in the analyses. The GWAS analyses of each separate cohort were carried out in PLINK 2.0 (Purcell et al., 2007). For both analyses, age, sex and 20 principal components were used as covariates to correct for population stratification. A linear regression was used for the association analysis, and the --family option in PLINK was used for correcting for the family structure.

The meta-analysis was performed using the METAL software by weighting the effect size estimates to their inverse of the standard errors (http://genome.sph.umich.edu/wiki/METAL_Program). From its combined total,

a final set of 6,767,570 SNPs were meta-analyzed (i.e. in common between the two datasets).

Figure 1 depicts the association results on the Meta-analysis. The QQ-plot shows that the 20 PCs included were appropriately corrected for population stratification. No genome-wide level of significance was found for any of the SNPs under investigation. The two top SNPs from the meta-analysis were rs139052 (p=8.30x10-07) and rs12873866 (1.32x10-06), both with a protective-effect for its major allele (Supplementary Table S1 depicts the top results obtained in the meta-analysis). SNP rs139052 is located in the PNPLA3 gene in 22q13.31, and the top results for chromosome 22 map to this locus within a span of 100kb of high LD region. The PNPLA3 gene encodes the adiponutrina protein - a triacylglycerol lipase involved in balancing the usage and storage of energy in adipocytes. SNP rs12873866 is located in 13q33.1, a large intronic region, forming a large LD block (Figure 2).

This study constitutes the largest sample available to date on GWAS in hoarding symptoms, but we realize that the lack of genome-wide significant results in this study is not unexpected given the still limited sample size combined with the moderate heritability of the HD trait. Only recently have we seen the first robust GWAS findings for e.g. schizophrenia, bipolar disorder, ADHD and major depression disorder (Demontis et al., 2017; Ripke et al., 2013; Ruderfer et al., 2014; Wray & Sullivan, 2017), as a result of much larger samples than described in this meta-analysis. Similar to what has been observed for other traits such as height and blood lipid levels, we expect that for psychiatric traits there will be the critical point (varying for different disorders) in sample sizes above which significant findings become detectable. It is currently widely agreed that until reaching this 'inflection point' GWAS consortia should collaborate to detect individual signals. This study constitutes the largest sample available to date for hoarding symptoms. Given the observed scenario for other psychiatric traits, as well as the recent works dedicated to hoarding from the NTR and the TwinsUK, we expect future genetic studies in HD to gain further relevance, and these results provide a solid starting point and groundwork for the future studies.

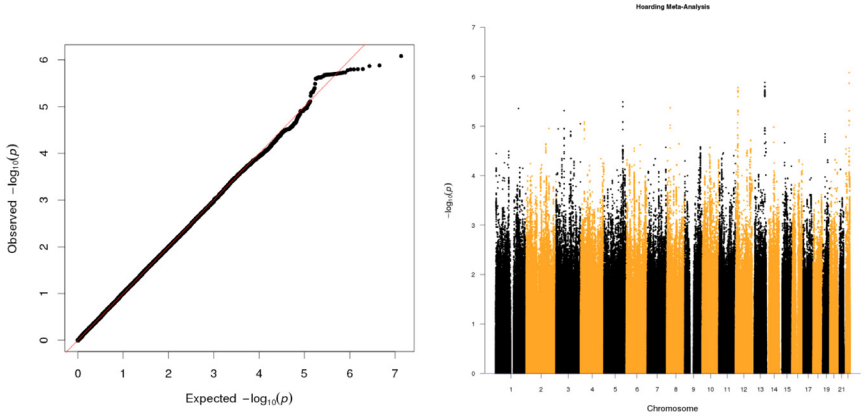


Figure 1. QQ-plot (left) and Manhattan plot (right) from the meta-analysis results from the NTR and the UKTwins for Hoarding Disorder.

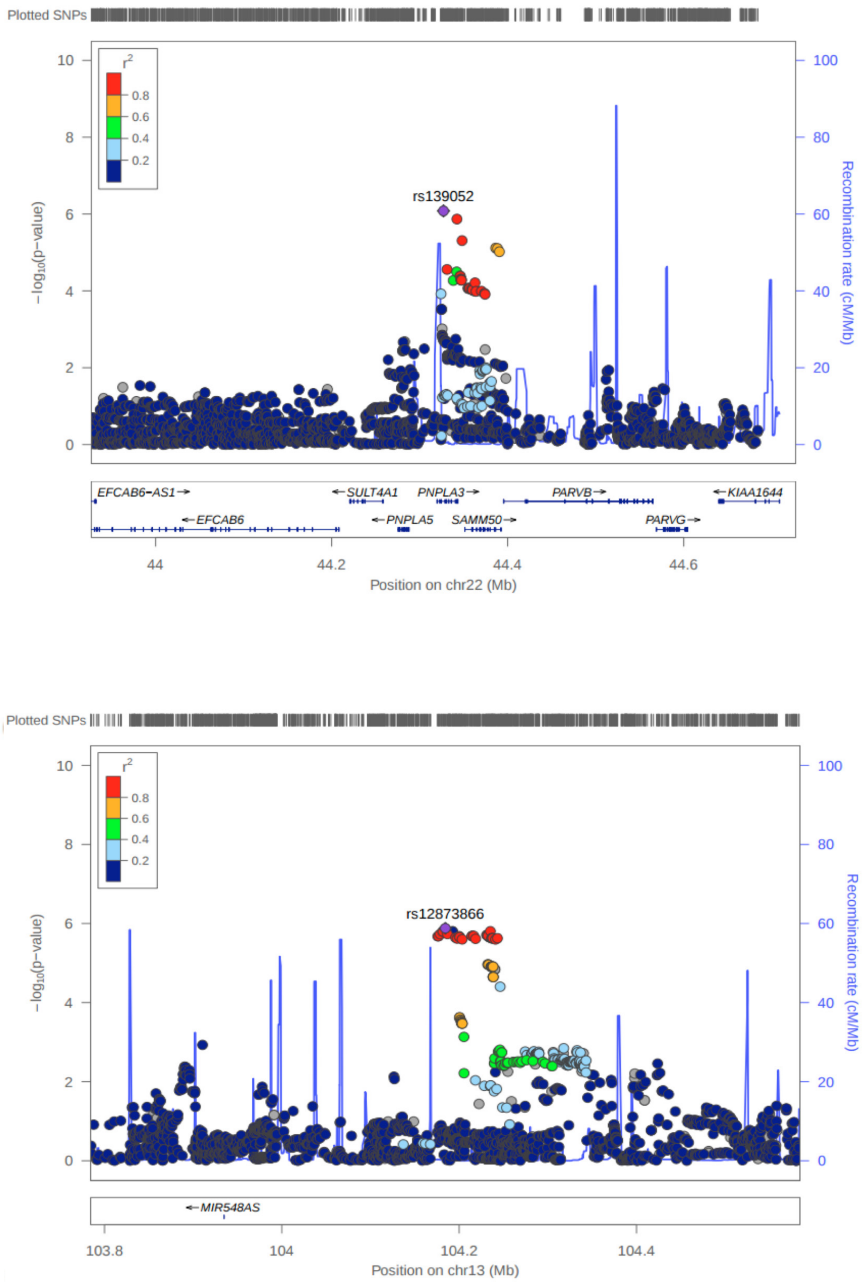


Figure 2. Regional plots showing the regional association for selected SNPs using the tool LocusZoom (http://locuszoom.org/) for the region 22q13.31 harbouring the SNP rs139052 (top), and the region 13q33.1 harbouring the SNP rs12873866

Chapter 8
Polygenic Risk Scores for Schizophrenia and Depression Linked to Obsessive-Compulsive Disorder

This chapter is based on: Zilhao, N. R., Abdellaoui, A., Smit, D. J. A., Cath, D., Hottenga J. J., Boomsma, D. I. (2017). Polygenic prediction of obsessive compulsive symptoms. Molecular Psychiatry (accepted).

Obsessive-compulsive disorder (OCD) is a common neuropsychiatric disorder, causing considerable social impairment and disease burden worldwide (Stewart et al., 2013). OCD is characterized by the presence of a heterogeneous array of 'clinically significant' obsessions and compulsions, in a permanent uncontrollable cycle. It is widely recognized, similarly to what is observed for other psychiatric disorders, that individual differences in the susceptibility to OCD are at least partly genetic (Stewart et al., 2013).

Twin and family studies estimate the heritability of OCD to be between 0.27 and 0.50. Molecular genetic studies in epidemiological samples offer evidence for the highly polygenic nature of OCD, i.e., OCD is influenced by large numbers of (common) genetic variants with small-effects (Stewart et al., 2013). This polygenicity is within the expectations, considering the findings on the genetic architecture of other psychiatric traits.

Polygenic risk scores (PRS) offer an opportunity to test if results from large meta-analyses for psychiatric disorders can predict OCD. PRS are a measure of an individual's genetic risk to develop a certain disorder, calculated by summing all genotype scores for a person after weighting them by their estimated effect size for a trait as obtained from a genetic association analysis. These scores can be used to test the predictive value of a PRS towards the disorder in independent samples, or towards another disorder of interest. This approach allows one to simultaneously 1) tackle the inherent culprit of polygenicity surrounding complex traits – that true genetic variance is captured by the full set of measured SNPs not only significant signals, and 2) assess to which extent the individual genetic risk to a trait has predictive value for another trait – serving as measure of genetic correlation between disorders, even in the absence of phenotype measurements for multiple disorders (Dudbridge, 2013).

Recently, large GWASs for psychiatric traits have been completed and we constructed polygenic scores based on the summary statistics from a series of disorders that are epidemiologically related to OCD, namely: Attention Deficit Hyperactivity Disorder (ADHD), Bipolar Disorder (BD), Schizophrenia (SCZ), major depressive disorder (MDD), Autism, and Migraine. In addition we had at our disposal summary statistics from a meta-analysis performed on clinically-derived OCD samples from the International OCD Foundation Genetic Collaborative (IOCDFGC) and OCD Collaborative Genetics Association Study (OCGAS) (Stewart et al., 2013; Mattheisen et al., 2014). We investigated the power of each of these scores to effectively predict obsessive-compulsive symptoms (OCS) in a population-based sample, considering different values for genetic correlations between the score and

OCS, following the method proposed by Dubridge (Dudbridge, 2013).

Table 1 summarizes the results. The polygenic scores for OCD, SCZ ($p=1.4x10-6$), MDD ($p=5.6x10-5$) and BP-SCZ combined ($p=8.1x10-7$) were significant after correcting for multiple testing ($\alpha=0.006$). This last result is probably driven by the genetic component that Schizophrenia shares with BD. Overall, the polygenic scores explained between 0.38-0.79% of the total variance of OCS - in line with what is observed for most PRS analysis. The score 'BD vs. Schizophrenia' – which includes the genetic component that differentiates between BD and Schizophrenia, did not significantly predict OCS, which indicates that genetic variation unique to SCZ or BIP is not related to OCS. MDD scores also showed a strong genetic association with OCS. Among the most co-occurring disorders with OCD are major depression (31%) and BD (7%), whilst for schizophrenia rates for co-occurrence are rare. These results indicate that the reason for comorbidity partly is genetic, with a shared genetic liability to OCS.

These results show that population-based OCS can be predicted with the help of clinically based GWASs, which in turn reveals that psychiatric genetic risk factors can lead to the presence of sub-clinical OC symptoms. The risk scores also indicate that OCS are clinically relevant markers of other psychiatric disorders related to OCD. In other words, sub-clinical OCS shares genetic variance that underlies OCD, Schizophrenia and MDD.

The relevance of PRS lies within their predictive value. PRS have a better predictive value than the individual GWASs' top hits. Their application may in many cases be clinically useful when applied to population-based samples (as herein shown), and predict the development of psychiatric traits such as OCD. Our results regarding the predictive value of MDD and Schizophrenia for OCS should be looked into in further studies.

Table 1. Results of the generalized estimation equations (GEE) association analyses between OCS and different polygenic scores (N=6,506). The third column also indicates the proportion (in percentage of variance explained in OCS by each polygenic risk score (R2). N-GWAS gives sample size and h2 the, heritability estimates (h2 LDSC) of the GWAS summary statistics from LD score regression. The last 4 columns give the power to detect an association between OCS and the polygenic scores given a genetic correlation (rg) of .2, .4, .6, and .8. Bold font indicates significant p-values after multiple testing correction (α=0.05/9=0.006); grey colour font indicates nominally significant p-values (α=0.05). The reference for the provenance of each summary statistics is given alongside the respective score.

Polygenic score	p-value	R^2 (%)	N GWAS	h^2 LDSC	Power (r_g=0.2)	Power (r_g=0.4)	Power (r_g=0.6)	Power (r_g=0.8)	References
OCD	**3.0×10^{-4}**	**0.57**	**9761**	**0.14**	**0.18**	**0.54**	**0.87**	**0.99**	**Stewart et al., 2013; Mattheisen et al., 2014**
ADHD	0.85	0.39	51306	0.23	0.84	0.84	1	1	Demontis et al., 2017
Bipolar Disorder (BD)	0.015	0.48	63766	0.43	0.29	0.29	0.98	1	Ruderfer et al., 2014
Schizophrenia	**1.4×10^{-6}**	**0.71**	**79845**	**0.45**	**0.97**	**0.97**	**1**	**1**	**Ripke et al., 2014**
BD & Schizophrenia	**8.1×10^{-7}**	**0.79**	**39202**	**0.37**	**0.48**	**0.48**	**0.99**	**1**	**Ruderfer et al., 2014**
BD vs Schizophrenia	0.025	0.44	16381	0.33	0.11	0.11	0.78	1	Ruderfer et al., 2014
MDD (PGC & 23andMe)	**5.6×10^{-5}**	**0.57**	**370,973**	**0.07**	**0.78**	**0.98**	**1**	**1**	**Wray & Sullivan, 2017**
Autism	0.47	0.38	10263	0.46	0.12	0.12	0.81	1	Smoller et al., 2013
Migraine	0.16	0.41	196685	0.04	0.03	0.03	0.31	0.99	Gormley et al., 2016

Chapter 9
Epigenome-Wide Association Study of Tic Disorders

This chapter is based on: Zilhão, N. R., Padmanabhuni, S. S., Pagliaroli, L., Barta, C., Smit, D. J. a., Cath, D., Boomsma, D. I. (2015). Epigenome-Wide Association Study of Tic Disorders. Twin Research and Human Genetics, 18(6), 1–11.

Abstract

Tic disorders are moderately heritable common psychiatric disorders that can be highly troubling, both in childhood and in adulthood. In this study we report results obtained in the first epigenome-wide association study (EWAS) of tic disorders. The subjects are participants in surveys at the Netherlands Twin Register (NTR) and the NTR biobank project. Tic disorders were measured with a self-report version of the Yale Global Tic Severity Scale Abbreviated version (YGTSS-ABBR), included in the 8th wave NTR data collection (2008). DNA methylation data consisted of 411,169 autosomal methylation sites assessed by the Illumina Infinium HumanMethylation450 BeadChip (HM450k array). Phenotype and DNA methylation data were available in 1678 subjects (mean age=41.5). No probes reached genome-wide significance ($p<1.2 \times 10^{-7}$). The strongest associated probe was cg15583738, located in an intergenic region on chromosome 8 ($p=1.98 \times 10^{-6}$). Several of the top ranking probes ($p < 1 \times 10^{-4}$) were in or nearby genes previously associated with neurological disorders (e.g. GABBRI, BLM and ADAM10), warranting their further investigation in relation to tic disorders. The top significantly enriched gene ontology terms among higher-ranking methylation sites included anatomical structure morphogenesis (GO:0009653, $p=4.6 \times 10^{-15}$) developmental process (GO:0032502, $p=2.96 \times 10^{-12}$), and cellular developmental process (GO:0048869, $p=1.96 \times 10^{-12}$). Overall, these results provide a first insight into the epigenetic mechanisms of tic disorders. This first study assesses the role of DNA methylation in tic disorders, and it lays the foundations for future work aiming to unravel the biological mechanisms underlying the architecture of this disorder.

Introduction

Tic disorders form a broad spectrum encompassing four different clinical entities: Tourette Syndrome (TS), chronic (motor or vocal) tic disorder, transient tic disorder and tic disorder not-otherwise specified. Tics are characterized by sudden, rapid, motor movements or vocalizations, performed in a ritualized, recurrent and stereotypical fashion (APA, 2000).

Multiple lines of evidence suggest that both genetic and interacting environmental factors are causes underlying the etiology of these phenotypes (Paschou, 2013). Twin and family studies have shown that the prevalence of Tourette Syndrome or chronic tic disorders among first degree relatives of affected individuals varies between 15% and 53% (Towbin, 2010). Furthermore, heritability estimates on tic disorders or TS range between 0.28 and 0.56 (Anckarsäter et al., 2011; Bolton et al., 2007; de Haan et al., 2015; Lichtenstein et al., 2010; Mathews & Grados, 2011; Ooki, 2005). However, most genetic and clinical studies so far have been hampered by small sample sizes, ambiguity in phenotype definitions and difficulties to disentangle genetic and familial environmental effects. Thus, molecular genetic studies have thus far not yet yielded robust findings (Pauls et al., 2014).

The study of epigenetic mechanisms on a genome-wide scale in humans represents the bridge between disease susceptibility and gene expression variation. It is known that epigenetic mechanisms regulate gene expression, and that these epigenetic mechanisms change during life, up and down-regulating different genes in response to external environmental conditions. DNA methylation at CpG dinucleotides is the most studied epigenetic mechanism in humans. It regulates gene expression by targeting for example promoters and enhancers and its pattern may change as a consequence of external and internal stimuli (Hackett et al., 2013). DNA methylation is also involved in other biological processes such as genomic imprinting where CpG sites are methylated based on their parental origin which are different between the paternal and maternal branch (Liu, Morgan, & Calhoun, 2010), and chromosome X inactivation where one copy of the female X chromosome is inactivated (Heard, Clerc, & Avner, 1997). Genomic imprinting has also been suggested to be involved in TS in a study evaluating parental gender-influenced differences in childhood TS phenotype. Greater motor tic complexity was associated with maternal transmission, whereas higher frequency in vocal tics was associated with paternal transmission (Lichter, Jackson, & Schachter, 1995).

Epigenome-wide association studies (EWAS) have thus far been used to

reveal altered methylation patterns in several complex disorders such as obesity, diabetes, schizophrenia and autism (Dempster et al., 2011; Moore, McKnight, & O'Neill, 2014; Rakyan, Down, & Beck, 2011; Wockner et al., 2014). This approach is becoming increasingly accessible and it is likely that DNA methylation is also involved in other neurological disorders. Currently no epigenome-wide association studies of tic disorders have been performed. These studies may clarify underlying mechanisms that differentially regulate genes in individuals manifesting tics.

Here, we report on the first epigenome-wide association study (EWAS) of tic disorders performed in a population-based sample from the Netherlands Twin Register (NTR).

Materials & Methods

Subject demographics

Subjects in this study are participants in the NTR biobank Project (Willemsen et al., 2010; Willemsen et al., 2013). Since 1991, the NTR has been collecting information on a broad range of phenotypes in twins and family members (Willemsen et al., 2013). In total 3264 peripheral blood samples from 3221 NTR participants have been assessed for genome-wide methylation data. After quality control (QC) on the methylation data, the total final selection comprised 3089 samples, for a total of 3057 individuals (32 subjects had methylation data for two time-points). For a complete description of the entire methylation dataset from the NTR, please see (van Dongen, et al., 2016).

In the present study we analysed tic data from the 2008 wave of collection, in a subset of individuals in whom genome-wide methylation data were available. A total of 1678 individuals (twins, siblings and parents) from 1057 families were included in the analysis. Table 1 gives an overview of the subjects entered in the analysis and demographics. Zygosity was assessed by DNA polymorphisms as described by (van Beijsterveldt et al., 2013). The study was approved by the Central Ethics Committee on Research involving human subjects of the VU University Amsterdam.

Table 1. Number of participants included in the analysis

	MZ twins		DZ/DOS twins		Relatives	
	male *(case/control)*	female *(case/control)*	male *(case/control)*	female *(case/control)*	male *(case/control)*	female *(case/control)*
N	241 *(30/211)*	680 *(74/606)*	213 *(31/182)*	404 *(38/366)*	63 *(8/55)*	77 *(7/70)*
age (se)	42.48(14.00)	40.57(12.08)	37.61(10.11)	38.46(11.04)	42.83 (14.10)	41.96 (12.45)

Abbreviations: MZ, monozygotic twins; DZ, dizygotic twins; DOS, dizygotic opposite-sex twins.

Phenotype

Tics were measured using an abbreviated 12 item self-report version of The Yale Global Tic Severity Scale (YGTSS-ABBR), the latter being a well validated interview with a high internal consistency (Cronbach's alpha > 0.90) (Leckman et al., 1989). The YGTSS-abbreviated (YGTSS-ABBR) contains 12 most frequently occurring tics, assessing their occurrence: never (0), < than one year ago (1), between 1-5 years ago (2) or as a child (3). Three additional questions are asked on age at onset, duration and severity to enable establishing a probable diagnosis according to DSM-IV-TR criteria (APA, 2000). Table 2 shows the YGTSS-abbreviated questionnaire used for measuring tics. A diagnosis of probable chronic tic disorder was established if the person had 1) one or more chronic motor or one or more vocal tics, that 2) occurred before age 21 and 3) had been present for >1 year. Probable TS diagnosis was established when 2 or more motor and 1 or more vocal tics were reported that occurred before age 21 and had lasted for > 1 year, and probable transient tic disorder was established when motor and/ or vocal tics had occurred before age 21 for less than one year. From these categories we derived a dichotomous variable on absence or presence of a probable tic disorder diagnosis - chronic tic disorder, transient tic disorder or Tourette Syndrome (TS), as referenced in the Tourette Syndrome Classification Study Group (TSCG, 1993). An extensive genetic analysis on the heritability of tic disorders has been performed (Zilhão et al., 2016). Since smoking is known to have an effect on DNA methylation (Lee & Pausova, 2013), we controlled for smoking status in the epigenome-wide association analysis. Smoking status was assessed at the moment of blood draw by interview.

Table 2. Yale Global Tic Severity Scale abbreviated

1		**Please indicate whether you have ever had any of the following involuntary sudden nervous tics. (several answers possible)**	Never	0-1 years ago	1-5 years ago	> 5 years ago
	a	Eye movements: e.g. blinking, rolling, squinting				
	b	Nose movements: e.g. nose twitching, broadening nostrils				
	c	Lip or mouth movements: e.g. chewing, licking, pouting				
	d	Head shaking				
	e	Shoulder or neck movements				
	f	Arm or hand movements: e.g. rapid 'purposeless' bending, stretching				
	g	Squeaking or whistling noises				
	h	Growling, throat clearing, coughing, sniffing				
	i	'Purposeless' cursing or utterance of rude or obscene language				
2		**If you have suffered from any of the above tics:**				
	a	At what age did you first exhibit any of these phenomena?	_____ years (age)			
	b	Have you ever suffered from these tics for more than a year at a time?	___ No__ Yes			
	c	How often did you have tics in the period that you suffered from them most?				
			__ not daily			
			__ daily, but tic-free for most of the day			
			__ daily, but tic-free periods of 3 hours not uncommon			
			__ daily, with tic-free periods of at most half an hour			

Infinium HumanMethylation450 BeadChip Data

DNA methylation was assessed with the Infinium HumanMethylation450 BeadChip Kit (Illumina, Inc.) (Bibikova et al., 2011). Genomic DNA from whole blood (500ng) was bisulfite treated using the Zymo Research 96-well plate using the standard protocol, by the department of Molecular Epidemiology from the Leiden University Medical Center (LUMC), The Netherlands. Subsequent steps (i.e. sample hybridization, staining, scanning) were performed by the Erasmus Medical Center micro-array facility, Rotterdam, The Netherlands. QC and processing of the blood methylation dataset has been described in detail previously (Van Dongen, et al., 2016). In short, a number of sample-level and probe-level quality checks were performed. Sample-level QC was performed using MethylAid (van Iterson et al., 2014). Probes were set to missing in a sample if they had an intensity value of exactly zero, or a detection P-value > 0.01, or a bead count < 3. After these steps, probes that failed based on the above criteria in > 5 % of the samples were excluded from all samples (only probes with a success rate >= 0.95 were retained). Probes were also excluded from all samples if they mapped to multiple locations in the genome (Chen et al., 2013), and/or had a SNP within the CpG site (at the C or G position) irrespective of minor allele frequency in the Dutch population (Francioli et al., 2014). Only autosomal methylation sites were analysed in the EWAS. The methylation data were normalized with Functional Normalization (Fortin et al., 2014) and normalized intensity values were converted into beta (β)-values. The β-value represents the methylation level at a site, ranging from 0 to 1 and is calculated as:

$$\beta = \frac{M}{M + U + 100}$$

where M=Methylated signal, U=Unmethylated signal, and 100 represents a correction term to control the β –value of probes with very low overall signal intensity. After QC, from an initial total of 453288 methylation sites, the final total selection of methylation sites was 411,169.

Statistical Analysis

Epigenome-wide association analysis (EWAS) was performed using linear regression under an additive model correcting for Principal Components (PCs) and covariates. Principal components (PCs) were calculated from the methylation data after QC and normalization. The PCA plot calculated can be seen in Supplementary Figure S2. Supplementary Figure S3 provides the correlation plot between the first 20 PCs and the covariates (monocyte count, eosinophils count, neutrophils count, array row number, smoking status, age and sex), in our dataset. Generalized estimation equation (GEE) models were used to test whether tics were associated with DNA methylation. In the final model, DNA methylation level was used as outcome with the following predictors: tics, top 5 PCs, smoking status, eosinophil percentage and monocyte percentage. Additional models were tested to evaluate the inflation factor with different covariates (Supplementary material). Sex, neutrophils count and age were not included due to their correlation with the top 5 PCs (Supplementary Figure S3). Basophil percentage was not included because it showed little variation between subjects, with a large number of subjects having 0% of basophils. Smoking status was coded as '0=non-smoker', '1=former-smoker', '2=current-smoker'. GEE uses cluster robust standard errors with family ID as cluster variable, and thus standard errors are robust for the presence of related individuals in the sample. A permuted p-value was calculated by random sampling of the phenotype (n=10000), defining the proportion of permutations meeting or exceeding our p-value estimate based on the actual data. CpGs with p-values <1.2 x 10-7 (Bonferroni correction of 0.05/411169 – autosomal sites) were considered statistically significant.

Enrichment of Gene Ontology Terms

Enrichment of Gene Ontology (GO) terms among methylation sites having a stronger association with tics was done by ranking all methylation sites based on the EWAS p-value and the resulting ranked gene list was supplied to the online software tool GOrilla (Eden, et al., 2009). The GO tool GOrilla takes

into account the rank of the gene where by p-value cut-off is not required in creating the gene list. A false Discovery Rate (FDR) q-value < 0.05 was considered for a GO term to be statistically significant.

Results

After QC, the final dataset consisted of 411,169 autosomal CpG sites, for 1678 individuals (188 Cases and 1490 Controls). The variation in the data captured by the first two PC is shown in Figure 1. It can be seen that the main contributors to variation in DNA methylation are not associated with the phenotype. The first component represented sex and the second component correlated strongly with neutrophil count. Supplementary Figure S1 shows the distribution of genome-wide methylation level for all individuals included in the analysis.

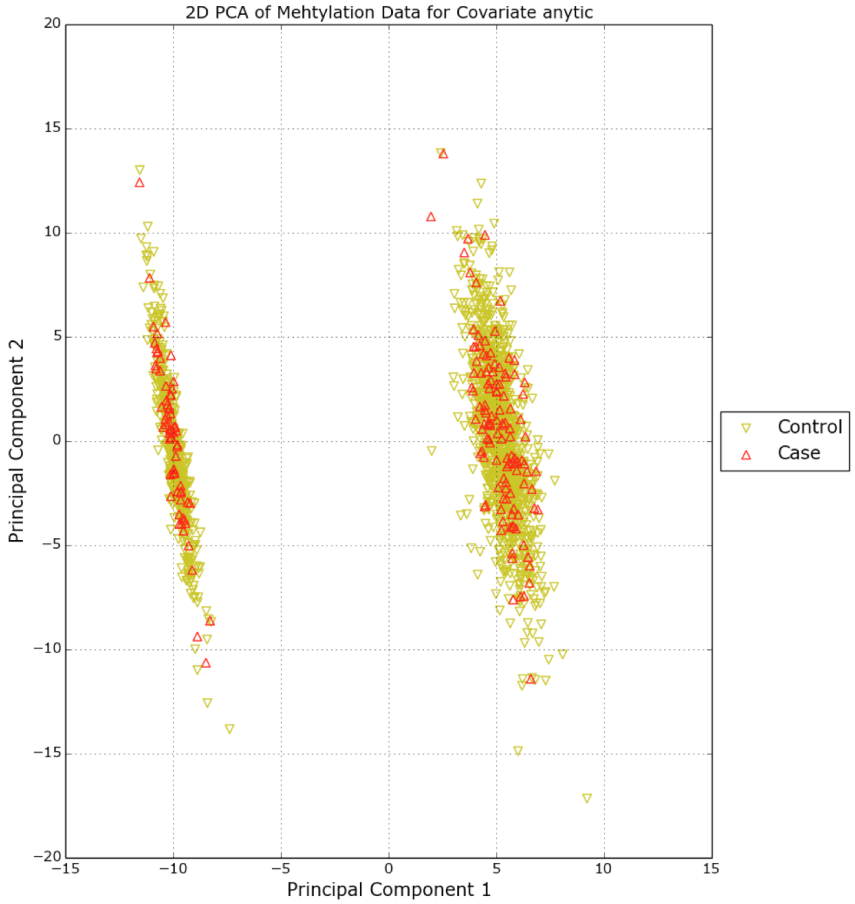


Figure 1. Two dimensional PCA plot labelled by case-control status.

The quantile-quantile (QQ) plot based on the EWAS model is shown in Figure 2. The inflation factor (λ) is 1.03, indicating that the results are not stratified.

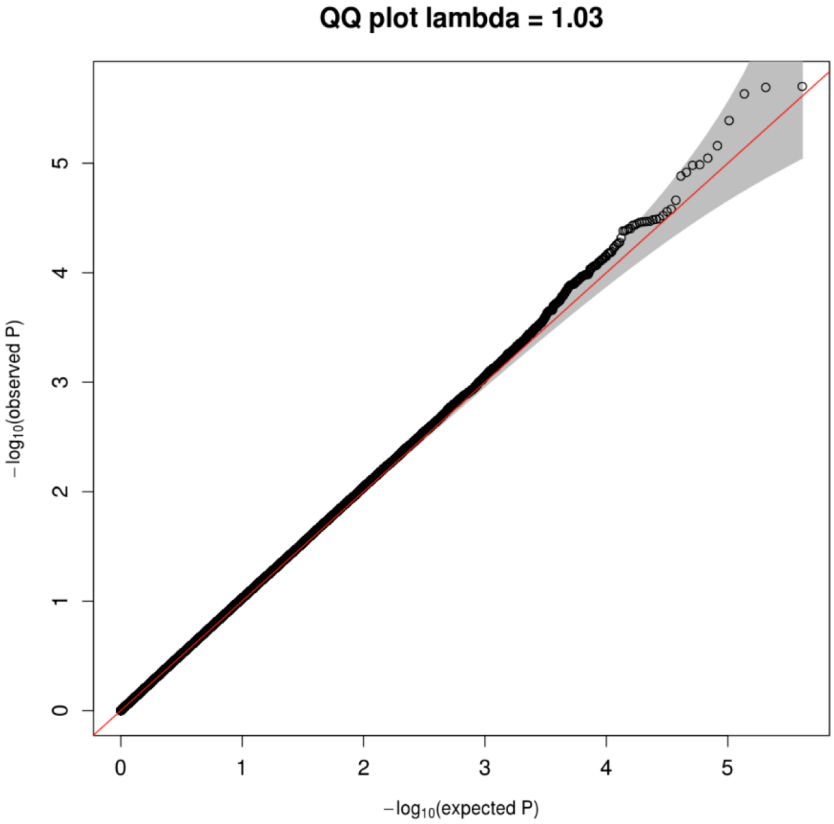


Figure 2. QQ plot of P-values from gee model with top 5 PCs and Covariates Tics, Smoking Status, Eosinophils, Monocytes and Array Row.

None of the CpG sites passed the Bonferroni p-value threshold (p=1.2 x 10-7) for the association with Tics. Table 3 shows all CpG sites with a p-value < 1 x 10-4 (N=57). Permutation tests for all of these 57 CpG sites resulted in a permutation p-value < 0.05 indicating that these CpG sites did not occur by chance. As an example, the distribution of permutation p-values for the top 9 CpGs is shown in Supplementary Figure S7. The Manhattan plot can be seen in Figure 3 with the top 57 CpG sites highlighted in green. Figure 4 shows the methylation level in cases and controls for our top 15 CpGs.

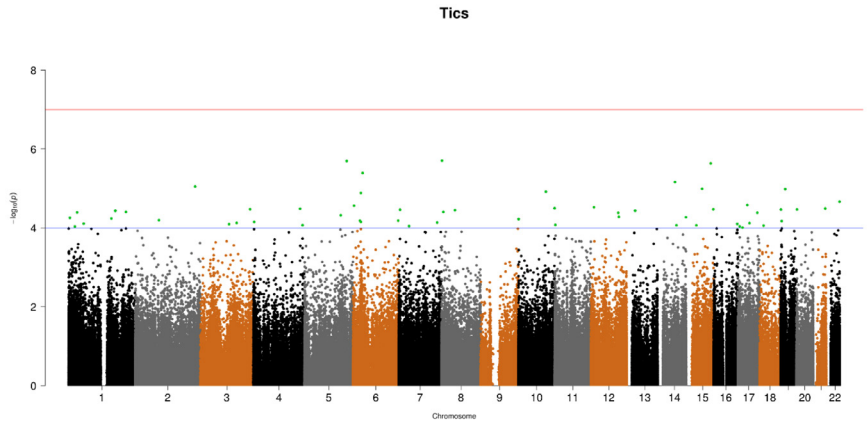


Figure 3. Manhattan plot showing the p-values of genome-wide CpG sites. The red line is the genome-wide threshold and the blue line is the suggestive threshold (P < 1.0 x 10-4). Top CpG sites are highlighted in Green.

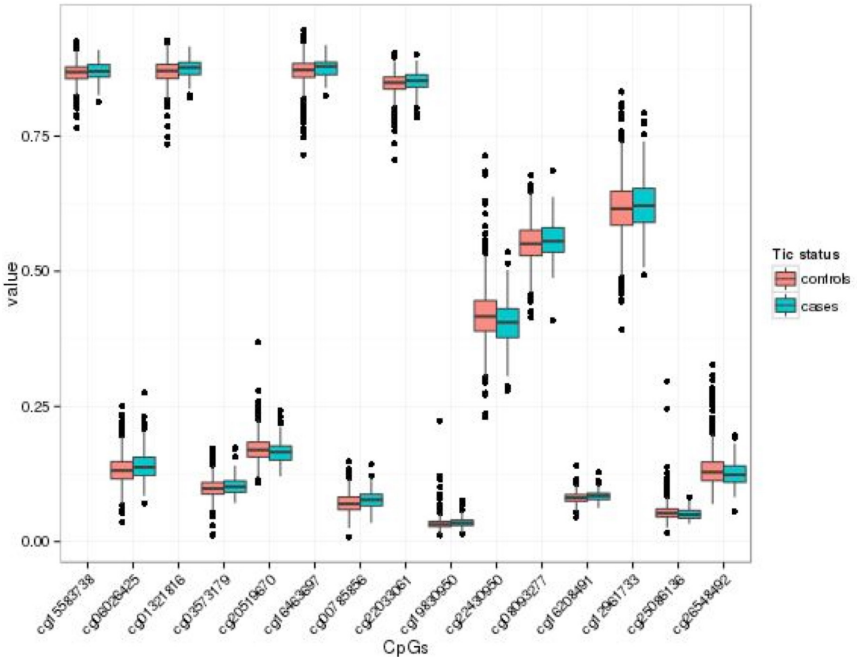


Figure 4. Boxplots of the methylation levels at the top CpGs in cases vs controls

Table 3. Top CpG sites in the EWAS with p-value < 0.0001.

CpG Site	CHR	CpG Location (bp)	P-value	Δβ	[Nearest] gene	Gene Location
cg15583738	8	2176944	1.98E-06	0.00408	--	--
cg06026425	5	157284650	2.03E-06	0.008678	CLINT1	5q33.3
cg01321816	15	91358514	2.32E-06	0.006536	BLM	15q26.1
cg03573179	6	36165382	4.07E-06	0.003791	BRPF3	6p21.31
cg20519670	14	65172006	6.91E-06	-0.00561	PLEKHG3	14q23.3
cg16463697	2	223886480	8.97E-06	0.006082	KCNE4	2q36.1
cg00785856	15	59041883	1.03E-05	0.005749	ADAM10	15q21.3
cg22033061	19	17531746	1.04E-05	0.003424	FAM125A/MVB12A	19p13.11
cg19830950	10	102729375	1.21E-05	0.002377	SEMA4G	10q24.31
cg08093277	6	29595299	1.31E-05	0.006787	GABBR1	6p22.1
cg12961733	22	50165244	2.17E-05	0.009437	BRD1	22q13.33
cg22430950	17	35166190	2.62E-05	-0.0146	--	--
cg16208491	6	4021748	2.76E-05	0.003249	PRPF4B	6p25.2
cg14830166	12	11821908	3.01E-05	-0.01192	ETV6	12p13.2
cg23261919	10	135072960	3.20E-05	0.002744	ADAM8	10q26.3
cg08451325	21	44720984	3.25E-05	0.003124	SNF1LK/SIK1	21q22.3
cg15234400	4	174453166	3.29E-05	-0.01027	NBLA00301/HAND2-AS1	4q34.1
cg21923992	3	185825749	3.38E-05	0.007345	ETV5	3q27.2
cg07394446	15	100881458	3.39E-05	-0.00446	ADAMTS17	15q26.3
cg19391247	20	2360385	3.42E-05	0.00292	TGM6	20p13
cg07975472	19	1503610	3.44E-05	0.003127	ADAMTSL5	19p13.3
cg23572228	7	4923575	3.48E-05	0.006102	RADIL	7p22.1
cg25086136	8	50823124	3.56E-05	-0.00347	SNTG1	8q11.21
cg06872019	13	31588778	3.66E-05	-0.012	C13orf26/TEX26	13q12.3
cg20240091	1	175044916	3.69E-05	0.00516	TNN	1q25.1
cg17750334	1	214776613	3.96E-05	-0.00386	CENPF	1q41
cg14752139	8	7328654	3.96E-05	0.003239	DEFB104B	8p23.1
cg16650530	1	32538413	4.06E-05	0.002929	TMEM39B	1p35.2
cg23221732	17	72383708	4.14E-05	-0.00432	GPR142	17q25.1
cg03704355	12	102270112	4.15E-05	0.002634	DRAM/DRAM1	12q23.2
cg26548492	5	135170171	4.79E-05	-0.00723	LOC153328/ SLC25A48	
cg21651356	12	104685539	5.27E-05	-0.01119	TXNRD1	12q23.3
cg11597832	14	105993747	5.39E-05	0.003214	TMEM121	14q32.33
cg19403269	1	5569798	5.57E-05	-0.01835	--	--
cg24688563	1	160388964	5.77E-05	0.006028	VANGL2	1q23.2
cg18556420	10	864596	5.99E-05	-0.01245	LARP5/LARP4B	10p15.3
cg26116669	2	88654640	6.39E-05	0.007397	--	--
cg01927730	6	168955764	6.56E-05	-0.00385	SMOC2	6q27
cg23066982	6	26204463	6.58E-05	0.002922	HIST1H4E	6p22.2
cg01490283	19	4066033	6.70E-05	0.003977	ZBTB7A	19p13.3
cg21879791	6	29594830	7.07E-05	0.003829	GABBR1	6p22.1
cg15347627	4	2941570	7.08E-05	0.004751	NOL14/NOP14	4p16.3
cg01703966	7	143207845	7.36E-05	-0.00416	LOC285965	
cg19511664	3	135685098	7.53E-05	0.002758	PPP2R3A	3q22.2
cg11155621	17	43238118	7.60E-05	0.003843	HEXIM2	17q21.31
cg07477602	1	56961319	7.82E-05	0.002334	PPAP2B	1p32.2
cg00473985	16	87670568	7.94E-05	0.001967	JPH3	16q24.2
cg04747445	3	107241417	8.09E-05	0.004325	BBX	3q13.12
cg27135510	11	2423571	8.43E-05	0.001627	TSSC4	11p15.5
cg02525719	4	183728549	8.59E-05	0.01153	ODZ3/TENM3	4q34.3
cg19497750	14	70588881	8.63E-05	-0.00479	SLC8A3	14q24.2
cg22112443	15	37393989	8.66E-05	0.002057	MEIS2	15q14
cg21281009	18	14748298	8.79E-05	-0.01545	ANKRD30B	18p11.21
cg01119319	7	38356808	9.04E-05	0.014511	TARP	7p15-p14
cg24419101	17	6484720	9.17E-05	0.002819	KIAA0753/TXNDC17	17p13.1
cg25203007	1	24126017	9.22E-05	0.00304	GALE	1p36.11
cg08490349	17	17086207	9.92E-05	0.002258	M-RIP/MPRIP	17p11.2

Abbreviations: CHR, chromosome; Δβ, Difference in mean methylation β-values between cases and controls; bp, base pairs.

Gene Ontology Enrichment Analysis

Gene ontology enrichment analysis identified a large number of GO terms that were significantly enriched (Supplementary Table T1-T3). Top enriched GO Processes include developmental process (GO:0032502, P=3.98e-16, FDR q-value = 2.69e-12), cellular developmental process (GO:0048869, P=4.35e-16, FDR q-value = 1.96e-12), single-organism developmental process (GO:0044767, P=6.57e-23, FDR q-value = 2.22e-12), and other GO terms related to development. Brain processes, including regulation of neuron projection guidance (GO:0097485, P=8.33e-11, FDR q-value = 5.11e-08), axon guidance (GO:0007411, P=8.33e-11, FDR q-value = 4.89e-08), and neuron differentiation (GO:0030182, P=1.15e-10, FDR q-value = 6.24e-08) were also significantly enriched. The top most enriched GO component was cell junction (GO:0030054, P=6.73e-09, FDR q-value = 1.09e-05), followed by the neuron specific GO components neuron part (GO:0097458, P=4.39*10-07, FDR q-value = 1.78e-04), synapse part (GO: 0044456, P=5.13*10-06, FDR q-value = 1.39e-03), postsynaptic density (GO:0014069, P=1.67e-05, FDR q-value = 3.87e-03) and dendrite (GO:0030425, P=1.04e-04, FDR q-value = 1.05e-02). GO components related to histone modification, including MOZ/MORF histone acetyltransferase complex (GO:0070776, P=4.21e-05, FDR q-value = 6.20e-03) and H3 histone acetyltransferase complex (GO:0070775, P=4.21e-05, FDR q-value = 5.68e-03), were significantly enriched. The top enriched GO function was protein binding (GO:0005515, P=1.10e-17, FDR q-value = 4.55e-14) followed by many functions involving DNA binding.

Discussion

In this work we present the first genome-wide epigenetic analysis on tics/tic disorders. Although no methylation site achieved significance at the genome-wide threshold, gene-ontology analysis of the top hits revealed enrichment in brain-specific and developmental processes. Thus, our findings provide interesting targets for further analysis.

Two of our top CpGs (cg08093277 and cg21879791, ranked 11 and 41) are located near the GABBR1 gene (39670 bp and 39201 bp, at 5', respectively), which represents a very interesting target for further analysis. GABBR1 encodes a subunit of the GABA receptor, the major inhibitory neurotransmitter in the central nervous system. GABA acts at brain inhibitory synapses where it binds transmembrane receptors of both presynaptic and postsynaptic neurons. Notably, GABBR1 has been previously associated with autism, schizophrenia, tremor, and obsessive-compulsive disorder (Fatemi, Folsom, & Thuras, 2009; Hegyi, 2013; Luo, Rajput, & Rajput, 2012; Richter

et al., 2012).

The involvement of GABA in TS is well documented in the extant neurobiological literature. Tics have been associated with reduced basal ganglia volume and reduced cortical thickness in motor and sensorimotor areas controlling the facial, orolingual and laryngeal muscles (Sowell et al., 2008). Studies of brain activity using PET or fMRI in TS subjects compared to controls have implicated dysfunctional striatum, thalamus as well as cortical regions. The medium-sized spiny neurons in the basal ganglia have GABAergic inhibitory projections to the substantia nigra and Globus Pallidus (Ribak, Vaughn, & Roberts, 1979). It is hypothesized that the tonic activity of the striatum acts to inhibit unwanted motor patterns (Albin & Mink, 2006). Treatment strategies have aimed at increasing GABA levels (Awaad, 1999; Mink, 2001) to counteract decreased inhibitory output resulting in excessive activity in frontal cortical areas. Although not all studies showed significant results, one randomized double-blind study administering a GABA ß-receptor agonist in Tourette syndrome resulted in reduced tic severity (Singer, Wendlandt, & Giuliano, 2001).

Importantly, 8 of the top 57 associated CpG sites ($p < 1 \times 10\text{-}4$) mapped to genes that have been previously associated with psychiatric or neurological disorders, some of them sharing neurodevelopmental aspects with tics (OCD, autism, bipolar disorder, schizophrenia, some forms of mental retardation), some disorders involving cortico-striatal pathways similar to tics (as in the case of Parkinson's disease), and some in which the link with tics seems less clear (Alzheimer's disease) (Supplementary table T4 summarizes the genes from our top list that have been previously associated with neurological disorders). CpG site cg06026425, located near CLINT1, has been linked to Schizophrenia (Leon et al., 2011; Wang, Liu, & Aragam, 2010). Moreover, cg01321816 located near BLM and cg00785856 located near ADAM10 involve genes that have been associated with Alzheimer's disease (Schrötter et al., 2013; Vassar, 2013). Another example is cg26548492 which is located near LOC153328/SLC25A48, a gene that has been proposed as candidate for Parkinson's disease (Liu et al., 2011). Lastly, cg2051967 which is located near PLEKHG3, involves a gene that has been linked to mental retardation (Lehalle et al., 2014; Lybaek, Oyen, & Houge, 2008). Furthermore, some of the top hits lead to genes with a brain-specific function. This is the case for cg25086136 located near SNTG1, which is specifically expressed in the central nervous system (Hafner, Obermajer, & Kos, 2010), and cg19830950 located near SEMA4G which might play a role in cerebellar development, the cerebellum being a core structure involved in the precision of motor control

(Maier et al., 2011).

Histone modification might also play a role in the tic phenotype through the MOZ/MORF complex histone acetyltransferase complex (Klein et al., 2014). Genes involved in histone modifications were significantly enriched in the gene ontology analysis. Acetylation represents one of the most frequent post-translational modifications (PTM) and it is catalyzed by lysine acetyl-transferase enzymes (KAT). It neutralizes the positive charge present on the amino group of histone tails allowing the switch from a condensed structure to a more relaxed one which results in an enhance level of transcription. Two of the top CpGs (cg03573179 in the BRPF3 gene and cg12961733 in the BRD1 gene) are pointing to this MOZ/MORF complex which is formed by bromodomain PHD finger protein (BRPF1/2/3), inhibitor of growth 5 (ING5) and homolog of Esa1-associated factor 6 (hEAE6) (Sapountzi & Cote, 2011). BRD1 is associated with schizophrenia and bipolar disorder (Christensen et al., 2012) and the MOZ/MORF complex plays a role in the regulation of the dentate gyrus, which is a brain structure extremely important for learning and memory (You et al., 2015). BRD1, which is a component of the histone H3 acetyltransferase activity within the MOZ/MORF complex, might have a role in gene expression through acetylation of histone H3 and H4 (Doyon et al., 2006). BRD1 has also been associated with schizophrenia and bipolar disorder (Christensen et al., 2012; Severinsen et al., 2006). BRPF3 is another component of the histone H3 acetyltransferase activity within the MOZ/MORF complex (Ullah et al., 2008). CpG site cg23066982 is located near the HIST1H4E gene, which is a member of the histone 4 family and is a part of the nucleosome core. It is known that nucleosomes pack DNA into chromatin to regulate several processes such as transcription, DNA replication and chromosome stability.

These findings support a role for aberrant DNA methylation levels in tic disorders as part of a broader neurodevelopmental dysregulation. It is important to note that our study examined DNA methylation in blood rather than in the central nervous system (CNS). The relationship with DNA methylation in CNS tissue remains unclear. However, it has been suggested that inter-individual variation in DNA methylation is correlated to some extent across blood and brain tissues (Davies et al., 2012). Also, it was observed that exposure to different forms of early life traumas led to similar methylation changes in blood and brain cells (Klengel et al., 2013). It has been proposed that epigenetic changes induced early in development in particular may be present across many different tissues, because they are propagated through cell division. (Feinberg & Irizarry, 2010; Jeffries et al., 2012; Mill & Heijmans,

2013).

Future studies might consider taking a trans-diagnostic neurodevelopmental approach, combining tic disorders and other neurodevelopmental disorders (including OCD, Autism, ADHD and childhood movement disorders), to disentangle common versus disorder-specific underlying methylation patterns. Finally, environmental factors assessed in a longitudinal study design should be incorporated in epigenetic studies, to investigate which environmental stressors / protectors at what age/ stage of development influence DNA methylation.

This study provides a first step to unravel the role of epigenetic mechanisms in tic disorders. It should be noted that we analysed an 'inclusive' tic phenotype definition that may obscure different underlying etiologies. Future studies of larger size or including clinical samples at the more extreme end of the tic phenotypic spectrum are required to improve statistical power. Future studies should also aim to examine different phenotypic tic dimensions (de Haan et al., 2015). Such studies hold the promise to shed light on the complex interaction between environmental and genetic factors leading to development and persistence of neuropsychiatric disorders.

The Supplementary material can be found in the online version of this manuscript

Chapter 10
Summary and General Discussion

Summary

The work presented in this dissertation is largely focused on uncovering and quantifying the genetic factors contributing to the development of disorders within the obsessive-compulsive spectrum. The two main approaches used to analyze data on obsessive-compulsive (OC) symptoms, tics and hoarding were genetic association analysis and twin studies, which intersect within genetic epidemiology. The findings are summarized below.

Chapter 1 provides an overview on the genetics of tic disorders, obsessive-compulsive disorder (OCD), and hoarding disorder (HD).

Chapter 2 describes a genetic epidemiological twin study on OC symptoms, indicating that OC symptoms are highly stable across time (a 6 year time period), in a population of around 5,500 adult twins (age range between 17-90 years). This study modelled the longitudinal phenotypic variance of OC symptoms as a function of genetic and environmental factors. It showed that individual differences in stability are due to a combination of genetic (heritability: h2=56%) and unique environmental factors, with heritability estimated at 56%. Furthermore, it showed that genetic influences on OC symptoms are stable across time with longitudinal genetic correlations of rG=0.58. The longitudinal unique environmental correlation was rE=0.46. This suggests that measurement error alone may not be enough to explain time-point specific variance. It also highlights the role that individual experiences, in childhood and adolescence, may have on OCD far into adulthood. The broad-sense heritability consisted of additive genetic variance, and the bivariate (two time points) model also captured non-additive (dominant) genetic effects contributing to the phenotypic variance. Genetic dominance explained around 22% and additive influence around 36%.

Chapter 3 reports on a heritability analysis using different definitions of tic disorders. A sample of 8,323 mono- and dizygotic adult twins was included, in addition to their 7,164 family members who had been measured on lifetime occurrence and characteristics of tics. This chapter explored the extent to which the contribution from genetic and environmental influences differed across different definitions of tic disorders. The different tic definitions, following the current DSM-5 criteria, represented a range of mild to severe tic symptoms. Heritability was estimated to contribute between 25-37% depending on the phenotype definition. These heritability estimates had overlapping confidence intervals, which suggested a similar genetic liability for the different tic phenotype definitions. Interestingly, the heritability of the most lenient tic definition (including any tic) showed the narrowest

confidence interval (h2=30%; 95% CI= .23-.38), indicating that the core phenomenological characteristics of tics ("sudden, rapid, recurrent, non-rhythmic, stereotyped motor movements or vocalizations") render the highest heritability estimate.

In Chapter 4 we analysed phenotypic data on OC symptoms, hoarding symptoms, and tics, and explored the amount of underlying genetic and environmental influences shared between these three phenotypes, to further explore the common etiology across these three disorders. We had at our disposal population-based data from the NTR for which N=5,293 individuals had phenotypic data available on all three phenotypes. This revealed substantial genetic correlations (between 0.35-0.41), with the highest genetic correlation of 0.41 to be found between OC symptoms-Hoarding symptoms. This specific result is of interest in light of the latest development in DSM-5, in which HD was separated from OCD as a distinct disorder and placed in the category of OC spectrum disorders. Moreover, our findings corroborate the findings by Iervolino (2009; 2011) and Tambs (2009). These results suggest that the symptoms related to OCD and HD share less genetic variance than the shared genetic variance observed between OCD and internalizing disorders such as panic disorder, generalized anxiety, phobias, and PTSD (genetic overlap of 0.55). To conclude, HD can be considered a separate, albeit related, entity to OCD, in line with its current position in DSM-5. Lastly, the results regarding the common factor model that was tested point to shared genetic etiology among the three phenotypes (with genetic correlations between .32-.43). With respect to the total environmental variance, tics had a considerably smaller loading of only 4.4% on the common factor. Based on these results, it can be hypothesized that commonalities in genetic architecture dictate underlying similarities in dysfunction at the structural and functional level in cortico-striato-thalamo-cortical circuitries – the regions so far implicated in these disorders by neuroimaging studies. In view of the lower environmental correlations between tics and both OC symptoms and Hoarding symptoms, it seems that unique environmental experiences that determine the development of tics, are by and large different from the unique environmental factors involved in OC spectrum disorders.

Chapter 5 introduces the use of genome wide array-SNP data in a series of exploratory analysis on the genetics of OC symptoms. It expands the work performed in Chapter 2. Here, 6,931 subjects were included (twins, their siblings, parents and spouses), for whom genetic data were available, i.e. genome wide SNP data, which together with phenotype information were analyzed in genetic association studies (GWAS), with polygenic risk scores,

SNP-based heritability (GCTA), and gene-based testing. For polygenic risk score calculations, GWAS results from a large clinical sample of OCD patients who were of European ancestry (the OCF-GC) were used (Stewart et al., 2013). Polygenic scores summarize genetic effects among a large set of markers that do not individually achieve significance into a single value per subject. These scores significantly predicted OC symptoms in the NTR population-based sample (with 0.2% explained). In the same sample, SNP-based heritability was estimated at 14%. The total variance explained by genetics, i.e. SNP and other heritability, captured using GCTA was of 34%. This means that 14% of the OCS phenotypic variance is attributable to genotyped SNPs, and 20% attributable to genetic variance not captured by the currently used genotyping SNP-arrays. The combined association analysis (GWAS and gene-based test) revealed a significant SNP (rs8100480), located within the MEF2BNB-MEF2B gene ($p=2.56\times10\text{-}8$), and four significant genes (RFXANK, MEF2B, MEF2BNB, MEF2BNB-MEF2B), all located in the chromosomal region (19p13.11).

Chapter 6 reports on a meta-analysis between genome-wide association results on tic disorders performed at the NTR (N=88 cases, using a narrowly defined phenotype; 6,381 controls) and results from a clinically based sample from the Psychiatric Genomics Consortium Tourette's workgroup (PGC-TS) (N=778 cases; 4,414 controls). In line with the results on OCS from Chapter 5, the results showed that polygenic risk scores calculated from the PGC-TS case-control sample significantly predicted tic disorder in the NTR population-based sample. In sum, this chapter showed the added value of combining clinically-based and population-based samples in the context of association analysis. Extending the results on heritability analysis on different tic phenotypes (Chapter 3), screened subjects from the NTR were included in the analysis, consisting of 88 cases diagnosed for the most severe manifestation of a tic disorder ('chronic tic disorder - motor/vocal' or 'Tourette Disorder'), and excluding 173 individuals diagnosed with a milder tic disorder ('transient tic disorder' or 'tic disorder not otherwise-specified'). The GCTA analysis showed that 14.6% of the heritability of the tic phenotype from the NTR is attributable to common SNPs. The top SNP from the meta-analysis, rs7783290, is located on chromosome 7, with a p-value of 1.49x10-7 and a Z-score of -5.25.

In Chapter 7 the first ever-reported meta-analysis on hoarding symptoms is presented. Two population-based samples from the NTR (N=6,521) and TwinsUK (N=5,190) were combined, with genotype data imputed to 21,775,582 and 47,072,643 SNPs, respectively. This study constitutes the

largest sample available to date for hoarding symptoms, and a good solid groundwork for future studies. The two top SNPs from the meta-analysis were rs139052 (p=8.30x10-7) and rs12873866 (1.32x10-6), both with a protective-effect for its major allele. The SNP rs139052 is located in the PNPLA3 gene in 22q13.31, and the SNP rs12873866 is located in 13q33.1, a large intronic region of high LD. Given the observed scenario for other psychiatric traits, future genetic studies in HD will gain further relevance.

Chapter 8 presents a polygenic dissection of OC symptoms, based on data used for the work described in Chapter 5. Building on the recent and ever-growing availability of data from large-scale GWASs, polygenic scores were built for a set of clinically-derived phenotypes chosen for their epidemiological relation to OCS, i.e., Attention Deficit Hyperactivity Disorder (ADHD), Bipolar Disorder (BD), Schizophrenia (SCZ), major depressive disorder (MDD), Autism and Migraine. PRS were also built for clinically-derived OCD samples from the International OCD Foundation Genetic Collaborative (IOCDFGC) and OCD Collaborative Genetics Association Study (OCGAS). A genetic risk score was calculated on these scores and tested for its predictive value for OC symptom. The polygenic scores for OCD (p=3.0x10-4), SCZ (p=1.4x10-6), MDD (p=5.6x10-5) and BP-SCZ combined (p=8.1x10-7) significantly predicted OC symptoms in the population-based sample, accounting for between 0.38-0.79% of its total variance. Following on the increasing value of PRS, these findings show the presence of sub-clinical OC symptoms based on psychiatric genetic risk factors, therefore strengthening the usefulness of using a phenotype derived from clinically significant symptoms. It further extends the work in Chapter 5 in illustrating the polygenicity of OC symptoms and its complex etiology. The growing availability of PRS renders it with a higher predictive value than GWASs, for which epidemiologically-based phenotypes seem to be equally suitable as disorder-based phenotypes.

In Chapter 9 the first Epigenome-wide association study (EWAS) of tics is presented. This study was conducted on 411,169 autosomal methylation sites for 1,678 individuals measured for tic disorders. Of these, all individuals within the NTR with current or retrospectively reported tics ('any probable lifetime tic' as defined in Chapter 3) were included as a case in the analysis for a total of 188 Cases and 1,490 Controls. Gene-ontology analyses for the higher-ranking methylation sites found that the following sites were involved: a methylation site involving anatomical structure morphogenesis (GO:0009653, p=4.6x10-15), one involving developmental process (GO:0032502, p=2.96x10-12), and one involving cellular developmental process (GO:0048869, p=1.96x10-12).

General Discussion

The notion that differences in psychopathology are under genetic control is fundamental to psychiatric genetics. The key component is to understand the genetic basis of individual differences in normal and abnormal behaviour and, with relevance to this dissertation - in the etiology of mental diseases and disorders.

In many aspects, neuropsychiatric disorders still lack acknowledgement as to its importance from the overall society. In Europe, they constitute the third leading cause of disability-adjusted life years (DALYs), following cardiovascular diseases and malignant carcinomas - accounting for around 15.2% of DALYs. Other estimates place neuropsychiatric disorders as the leading cause of years lived with disability (YLD), accounting for 36.1% of all YLD. In recent years, large-scale collaborative research and interdisciplinary training has been carried out worldwide within psychiatry. The work herein presented was developed within the scope of a large scale European Grant obtained in 2013 - the Marie Curie Initial Training Network, funded by the European Union, and motivated by the lack of collaborative infrastructure for the study of psychiatric/neurodevelopmental disorders. Goals were set to unravel the genetic, environmental and neurological basis of Tourette's Disorder and its comorbid disorders such as Obsessive-Compulsive Disorder, using a multidisciplinary perspective, under the guidance of leading experts across Europe.

This dissertation largely focused on uncovering and quantifying the genetic factors playing a role in developmental disorders in the obsessive and impulsive spectrum. The remarkable increase in sample sizes as well as the development of a wide range of techniques within the field of genetics have contributed to a deeper understanding of these disorders – presenting, first and foremost, a solid and robust foundation for future studies. In the research within this thesis, we have used some of the newest tools in large-scale DNA data analysis. Below, the findings are discussed in the context of the larger field of psychiatric genetics.

Insights from the Classical Twin Design

The application of the twin design was an integral part of this dissertation. Going back to almost a century ago, the twin design is based on the observation that comparing phenotypic resemblance between monozygotic (MZ) and dizygotic (DZ) twins can inform on the relative contributions from environmental and genetic factors to the variability within a phenotype – heritability. This relevance of the twin method evolved in parallel to the notion

that many human traits did not follow the monogenic and Mendelian pattern of transmission. Early gene-finding methods – linkage and candidate gene studies, fell short in uncovering causal genetic variation. This reflects a long-standing debate in genetics, traced back as far as 1900s, between Mendelian and Quantitative genetics, that still somewhat endures today. For decades, twin studies played a key role in showing that variation within the human genome was partially responsible for developing psychological and psychiatric traits. In fact, the applicability of twin studies has extended far beyond that of estimating heritability. Within this range is included: assessing the stability of genetic and environmental etiology underlying the development (or change) of behavioural traits, the genetic correlations among multiple phenotypes, direction of causality between different traits, phenotypic assortment, and assessing qualitative and quantitative gender differences in the contribution of genes and environment to psychopathological traits. Being almost a century old, it is remarkable to observe the trust for twin studies emerging anew, prompted by the growing availability of ever-increasing datasets. It is now possible to gather questionnaire data from self-reports, and assess the symptoms distributed at the level of the population, which allows insights into true phenotypic variation underlying the liability to diseases.

As a result of these developments there has been a remarkable paradigm shift regarding how most psychiatric disorders are being conceptualized. Particularly, TD went from being widely viewed as a monogenetic disorder with an autosomal dominant mode of transmission (Pauls et al., 1986) to a highly complex and genetically and phenotypically heterogeneous disorder as it is regarded nowadays, being in fact, within the same range as any other complex trait – schizophrenia, autism and cardiovascular disease, among many more. The work presented in Chapter 3 provides important conclusions on this matter, where TD, conceptualized as the most severe manifestation of tic disorders, was found to represent genetic variability overlapping with other tic disorders; this suggests that a common genetic architecture underlies the core phenomenological aspects of all tic disorders. Moreover, another point derived from Chapter 3 is the added value of using twin studies from a population-based sample, a point which Twin Registers (such as the NTR) are at a unique position to address. Twin Registers also contribute with the extensive collection of longitudinal phenotypic data. All complex disorders are age dependent, and equally the genetic factors underlying them. In Chapter 2 we explored this, and found that the genetic factors contributing to OC symptoms were only moderately stable in time.

As attested for by countless lines written in scientific publications heritability

has undeniably been a crucial parameter in genetics. However, some misconceptions need to be addressed. Heritability can be defined in its simpler form as a ratio of variances – specifically, of genetic variance to total phenotypic population variance. In this way, it constitutes a "measurement". It is merely a descriptive statistic (in its true sense) with no de facto intrinsic value. These are implications that follow from the very definition of heritability, which are outlined below.

Heritability is a parameter of a population and hence specific to the population it derives from. This implies two things. Firstly, that it renders no interpretation on the individual level. Secondly, that heritability does not constitute an immutable property in time. These interpretations are consequential of heritability being a proportion of a variance. A further element that must be grasped is that while being a composite value, heritability refers to non-monomorphic loci, which segregate within a specific population. Formally put, considering the ratio formulating heritability: an (environmental) effect on the mean value of a trait (pertaining to the denominator in the ratio formula) will influence (to greater or lesser extent) that ratio - the statistic value describing heritability. The noteworthy element here is the degree to which this environmental effect can change across time or across populations. This change across time or across populations of an environmental effect pertains to the non-immutability of heritability. Human height stands as one of the clearest example to this: having a very high estimated heritability (around 80%), the observed steady increase in human height across populations around the world reflects improved conditions of nutrition and medical care (environmental factors). This serves as an example of a change in environmental factors with no implications in heritability estimate. Heritability, as estimated in a particular population cannot be generalized - to another population, or to a different point in time - in which a higher variability in an environmental effect (e.g. different standards of access to medical care) can render a higher relative proportion in predicting the total trait variability. Still relating to this point, heritability, having in mind that it refers to inter-individual differences, renders no interpretation on the absolute value of a trait, but rather, is constrained by it. In the example of height – the steady increase (over the years) of the absolute trait value has not been followed by a change in heritability. Moreover, we have the uninformative element of heritability towards specific genes. Heritability can only inform on relative contributions from segregating genes within that same population, therefore ignoring fixed effect alleles that, notwithstanding, are causal for the same trait under study. A further consideration is in the following example of blood pressure: the effect of a gene or a set of genes on blood pressure

in a population consuming a homogeneous diet can be much higher than in a population with diverse dietary habits. This example reflects a situation in which the relative effect of a gene in the variation of a trait is dependent on other (environmental) causes.

It can be argued that the debate on heritability represents two sides of the same coin. On one side, from the standpoint of geneticists it could be argued that as a predictive value heritability estimates performs best when having controlled for known 'fixed' effects within a population, and hence informing on the relative genetic risk to disease independent of known environmental risk factors. Falconer and Mackay (1996) elegantly showed this in their work on how heritability estimates can predict the response to selection (Falconer & Mackay, 1996). On the other side, an evolutionary perspective would state that fixed effects, in its strict sense, are an erroneous assumption. Natural selection acts as environmental pressures on a trait. It does so across the full spectrum of factors influencing inter-individual differences - we are interested in understanding the response to these environmental pressures that shape the totality of a variability of a trait.

In conclusion, with heritability being an aggregate measure, it is uninformative towards individual genes. Heritability is an observational measure, a statistic to be interpreted within specific contextual boundaries. It is descriptive, but not fundamental. The implications, specifically in the study of psychiatric traits, are of particular relevance for its interpretations, for it is not uncommon for such interpretations to be made somewhat blindly. With some of these crucial properties of heritability in mind – non-immutability, change over time, population- and gene-specific, and dependency on environmental effects, examining the oldest standing question in genetics – 'How does genetic variation contribute to phenotypic variation?', one wonders on the implications of heritability findings. The answer is the corollary to everything discussed in this section: understanding the effect that prevalent environmental conditions have on the value of a trait. These environmental conditions stand as potential modifiable risk factors, as opposed to potentially more difficult modifications of biological causes. For the purpose of psychiatric genetics, this teaches us at minimum that environmental interventions can be as successful (if not more) as genetic ones. It may be that there within lies the key to disease prevention, realizing that a less heritable trait is not (necessarily) easier to intervene in than a more heritable one (Haworth & Davis, 2014). Perhaps then, the best lesson from years of genetic research into the nature of psychiatric disorders will be that the key to disease prevention actually lies in managing environmental factors. Ultimately, heritability estimates will reveal the point where all genes

that contribute to a trait have been found. As a final note, the importance of finding heritable influences for psychiatric disorders for patients and their families must be mentioned. Understanding the causes of disorders, even without this understanding leading to immediate treatment, can raise social and personal awareness and lift the stigma often attached to these disorders.

The Role of Genetics in Psychiatric Disorders

The development and contribution from the field of genome-wide data to psychiatric genetics has been extensively outlined in this dissertation. Although this growth is relatively recent, its origins can be traced back to the work of R.A. Fisher, published in 1918, and the so-called "infinitesimal model" (Fisher, 1918). In a nutshell, this model postulates that a quantitative trait will be affected by an ever-decreasing allele effect size, in function of an ever-increasing number of sampled alleles. In other words, a polygenic trait will always be a continuous and normally distributed phenotype given a random sampling of alleles in that population. Proponents of quantitative genetics endorsed this assumption advocating that Mendelian genetics was insufficient to explain the observed normal distribution of several of the studied human traits. Enters the polygenic era in genetics. In fact, the advent of the genome-wide period since 2006 allowed the validation of the polygenic model, whereas before validating the polygenic model could only be done by the use of twin and family studies.

Although the infinitesimal model has largely dominated the view during the past century, and despite its successes, some limitations have surfaced over time. One of the widely-discussed topics in relation to these emerging limitations is the matter of 'missing heritability'. It was first described and named in 2008, after the realization that most findings brought forward by the Human Genome Project did not stand up for replication (Maher, 2008). Most importantly, the observation that the collective proportion of variance explained by significantly associated findings across more and more published studies, still fell short (largely) to the estimates provided by the standard genetic methods of twin- and family-based studies. Several essays have delved into the possible causes and explanations for this so-called paradox. Truly, the discussion around this topic has been evolving during the past decade. Although this seems an over-analyzed topic, it really is the culprit underlying psychiatric genetic research undermining all reflections within it. In summation, the likely reasons put forward to explain the 'missing heritability' have been 1) a possible over-estimation of heritability from twin and family studies, 2) possible genetic variants not tagged by the current genotyping arrays (such as rare and structural variants), 3) lack of power

on current GWASs to capture very small effect sizes from highly polygenic phenotypes, 4) genetic variance hidden as epigenetic variance and possibly not captured in GWAS studies, and lastly 5) issues pertaining to the definition of the phenotype. Following the fast turnaround in this field, this so called 'mystery' has somewhat diminished. The advent of GCTA (Yang et al., 2010) offered strong support to the idea that the highly polygenic nature of most complex traits was the true cause of 'missing heritability'. Their results undeniably showed that – collectively - SNPs explained a substantial proportion of trait variation, but did not, however, explain all of the heritable variance. The imprecision to detect the already inherently small effect SNPs, would be accentuated by the lack of power due to small sample sizes of the ongoing GWASs. This is precisely what we did in Chapter 5, by partitioning the heritability into SNP-heritability and heritability not tagged by the SNP-chip genotyping platform, using the method proposed by Zaitlen (Zaitlen et al., 2013). This premise, although not fully answering the question of missing heritability, did show that part of the heritability had been hidden, rather than missing. The work in Chapter 8 stands on this current prevailing view: polygenic scores seem to offer better predictive value than individual genetic variants obtained in GWASs. Its extensive use, especially in the past 5 years is proof of it. This offers the possibility that, in the near future, genetic datasets with a larger and better tagging variance will offer good hopes for explaining larger proportions of phenotypic variance. In light of the ever-growing notion of the intrinsic complexity of most of these disorders, it is perhaps a better strategy to apply polygenic scores that reflect the cumulative genetic variation underlying several phenotypes (Plomin, Haworth, & Davis, 2009).

Henceforth, two aspects are of relevance. First, the issue of sample sizes - inherently related to statistical power. On this, there is a clear consensus: larger sample sizes will render higher statistical power. The findings from the Schizophrenia working group within the PGC finally provided the long-awaited breakthrough - the identification of 108 loci, and of the C4 gene, has been regarded as proof-of-concept within the genetic community (Ripke et al., 2014; Sekar et al., 2016). The noteworthy point on this is the noticed 'inflection point' in sample size (N~15,000), after which there is a linear relationship between sample size increase and newly discovered loci (estimated at around 4 new SNPs per additional 1,000 cases (Levinson et al., 2014). A number of other complex traits have been following (and verifying) this trend, such as Crohn's Disease, height, Bipolar Disorder, type 2 Diabetes, major depressive disorder, and ADHD (Demontis et al., 2017; Wray & Sullivan, 2017). With this prediction in mind, a similar inflection point is to be expected for TS, OCD and HD. Second, it is increasing likely that effects from genetic variants

not tagged by the current platforms remain to be found. This refers to not only SNPs currently not tagged under LD-based information, but also other forms of genetic variation, including rare variants and structural variants. Rare variants usually refer to rare deleterious mutations of larger effect sizes, and structural variation takes the form of CNVs, tandem repeats, indels (insertions or deletions), duplications and translocations. The advent of Next Generation Sequencing (NGS) technology and Whole Exome Sequencing (WES) offers new windows of possibility in the near future, when the contribution from lower-frequency variants to additive genetic variance are explicitly estimated.

On the long road leading to disease prevention, the role of genetics is to inform translational research. Probably the biggest challenge of today's medical genetics is to demonstrate functional relevance and mechanistic interpretation for genetic variation. This follows on a further paradigm that emerged within the last years of genetic research - the fact that most GWAS findings map to non-coding areas of the genome. This has motivated other works that have dedicated extensive analysis on bridging the gap between biology and GWASs, which, despite being outside the scope of this dissertation, are worth mentioning – PrediXcan, ENCODE, Epigenome RoadMap, REMC, GTEx, fQTL-SCAN, among more (Boyle, Li, & Pritchard, 2017; Gamazon et al., 2015). The rationale underlying these works has been to target the genome for functionally relevant elements that are prioritized from biologically informed frameworks – microRNA expression, gene expression (tissue-dependent), and epigenetic data. On this note, epigenetics processes refer to changes in gene expression levels (rather than protein function directly) by means of modifications to the DNA strand, histones, or the chromatin. DNA methylation is a form of epigenetic modification at level of the DNA strand, associated with both increased and decreased gene expression via methyl-binding proteins. It is currently understood that different epigenomic profiles can influence (silencing or activation) gene expression in aberrant manners. It is therefore unsurprising that the use of quantitative molecular traits has become increasingly popular – expression quantitative trait loci (eQTL) or methylation quantitative trait loci (mQTL). Quantitative trait loci (QTLs) are regions of the DNA associated with variation in a quantitative trait (Lander & Botstein, 1989). These mQTLs or eQTLs refer to genomic loci that influence DNA methylation or mRNA expression, respectively. With the advent of high-throughput DNA technology this has become a growingly popular approach. In chapter 9 we took a first step in attempting to characterize DNA methylation profiles in association with tic disorders.

Notably, on the issue of bridging the gap between biology and GWASs, a recent paper by the group of Pritchard (Boyle et al., 2017) suggested an 'Omnigenic model' of disease. It proposes that there exists a core set of genes with regulatory variants that invariably will affect disease risk, but are vastly outnumbered by peripheral genes that harbour the greater proportion of tagged genetic variance - bringing us to the last point in this section: genetic pleiotropy. As Pritchard et al. elegantly put it, there is strong evidence that several genetic causal variants are the same for different diseases (pleiotropy). This is so when simultaneously in light of Fisher's 'infinitesimal model', and considering that the variation in the human genome is de facto finite (albeit large). Further, the evidence points to a widespread pleiotropy across the human genome. Again, the relevant implications from this is that it is a preferred approach to study diseases in combination, rather than as isolated entities assuming simplistic models of gene->pathway->disease. Leading us, once more, to the use of polygenic scores. Besides its predictive value, a major contribution from this method is that establishes causality and correlations between different diseases.

Interestingly, this provides an opportunity for twin studies in this 'omics era'. First, with respect to the topic of pleiotropy: the two main reasons that can exemplify pleiotropy are genetic correlation and response to selection. In other words, genetic and phenotypic correlations generated by pleiotropy can dictate the response to selection. However, a genetic correlation between two disorders can arise both from pleiotropy or other reasons such as heterogeneity – referring to misclassification of cases from one disease as another. Further, a phenotypic correlation between two disorders is not by itself evidence of pleiotropy as this correlation can be due to environmental factors affecting both disorders. As it has been reviewed in this dissertation, twin studies are uniquely suited to answer these questions. Second, on the topic of the use of molecular traits: the continuing technological advances now make it possible to assess individual differences at the molecular level. This allows us to investigate to what extent these differences (or resemblances) at the molecular level underlie phenotypic similarity for two disorders of interest (van Dongen, Slagboom, Draisma, & Martin, 2012).

Psychiatry: perspectives from genetic epidemiology

Pertaining to all questions in (psychiatric) genetics is the issue that concerns measurement, i.e. the definition and operationalization of the phenotype. This question, though simple is far from trivial. From the perspective of genetics, the definition of the phenotype is of utmost importance, not only for the study design, but also for its interpretation.

It is important to note that most disorders we study were defined in a pre-genetics and pre-neuroscience period; on many occasions this process followed an observational basis and was empirically derived. With time, psychiatry has adopted structured and standardized diagnostic criteria to attempt to surpass some limitations. Published works such as the DSM, for instance, have provided a greater ease of communication amongst clinicians. Following this standardization psychiatric disorders became conceptualized as syndromes, which in the past decades have rapidly increased in number with each successive edition of the DSM. This has undoubtedly provided advantages in both clinical practice and clinical research. However, the progress made in the fields of neuroscience and molecular biology provided for the first time the possibility in psychiatry to have a bottom-up approach in conceptualizing disorders. We now understand that, if on one hand the origin of some disorders is hidden in the functioning of a highly complex circuitry in the mental apparatus, this functioning is, in turn, constrained by all of the genomic machinery preceding it. The term endophenotypes was coined attempting to fill this gap. It is used to describe the search for phenotypes that underlie and precede behavior, and that furthermore can have a stable and clearer genetic relation with the phenotype. The idea is to seek for the neurological underpinnings behind complex disorders. When considering that these biological phenomena are an inheritance from millions of years of evolution framing human behaviour in a continuum along the evolutionary history, it only takes one more step to define evolutionary psychiatry as a field of study, that tries to explain psychiatric disorders through the evolving functional change in the human brain. For the field of genetics, this posed as an additional argument in favour of 'infinitesimally' small effect sizes of some causal genetic variants that avoid being selected against, and thus remain in the population (van Dongen et al., 2013). Indeed, this has put forward the hypothesis that most evolutionary change in complex traits acts through polygenic adaptation.

Realizing the limits of this dissertation for such philosophical endeavours, the relevance of enumerating these topics is to illustrate that at present time, the research approach into the nature of human behaviour and the nature for its inter-individual differences are vastly multidisciplinary. It is within this context, that attempts are already in place to offer a new perspective into the understanding and definition of mental disorders. The Research Domain Criteria (RdoC) project has been such an initiative (Cuthbert, 2015). Presenting itself as a research framework, it attempts to provide a biologically informed approach "by bringing the power of modern research approaches in genetics, neuroscience, and behavioural science to the problem of mental

illness" (Walter, 2017). As mentioned above, the diagnostics and definitions in use today have largely been proposed in an era when knowledge from neuroscience, genetics, and molecular biology was null. The work in this dissertation supports a future translational approach into mental disorders.

The use of genome-wide SNP data within this dissertation has by-and-large supported the polygenic approach - as was more specifically described in Chapters 5 and 8. This greatly contributes to conceptualizing TD/tic disorders, OCD and HD as highly heterogeneous and complex entities. On the other hand however, the genetic overlap showed in Chapter 4, is evidence of pleiotropy, and evidences the limitations of attempting to find causal links between genetic variants and specific clinically-derived syndromes, that as herein suggested are multifaceted and show substantial overlap - both genetic and phenotypic.

Future Perspectives

The field of Psychiatric genetics has seen a remarkable progress over the last decades. Several genetic variants have only recently been identified and robustly associated with disease risk. Technological advances brought computational power, and investments in international collaborations created new frameworks in which these studies are conducted. It is clear that genetics is now living its golden age. One would struggle finding any other domain within our societies that has witnessed the same decrease in cost for the same period of time (Mardis, 2011). Still, in parallel to answers, new questions have arisen. These emerging questions dictate what we can expect for the future of the field.

The steep growth in generating genetic data will most likely continue. On this point, even a conservative prediction would state that the limitation will not be the availability of data, but rather the manpower to analyse it. This extends to other forms of genetic data. It is expected that in the near future we will be able to answer questions such as 'how many distinct genetic variants contribute to this trait's variation'? - or in other words, explain the totality of a trait's additive genetic variation. The increasing interest in rare variants will likely bring about an anew focus on linkage studies, for which the contribution from the classic twin design is unique – e.g. genome-wide linkage scans in DZ twins pairs.

We are still far from incorporating knowledge from the proteome and transcriptome into GWAS findings. Likely, the technological advances and increase in data size will also affect how research is conducted, becoming

increasingly data-driven (as opposed to model-based approaches). Invariably this means a better integration of different fields within biology. We still see developments in our understanding of how cellular networks mediate disease-causing genetic variants, and how different cell types (and different tissues) influence gene expression with effect on disease risk. Likely, advances in algorithm development for multivariate genetic data analysis will be to incorporate additional data from the phenome - transcriptome regulation, post-translation modifications, proteome regulation and cell signalling. In particular machine learning algorithms - that can learn and improve with experience, are already being developed, and are especially promising. Such technologies will assist humans in the analysis of large and highly complex datasets. Although the decrease in cost of DNA-technology and larger sample sizes are impacting the way research is conducted, the better approach may still lie at the phenotyping level. Samples that are phenotypically more informative will maximize gene-discovery. This is to say that the technological advances within DNA-technology cannot invalidate the necessary progress within psychiatry. If we say that genotyping will no longer be a limiting factor, the same is not true for phenotyping. A better understanding on how several disorders are conceptualized is fundamental to our capability to unravel the true genetic architecture of psychiatric disorders.

References

Abelson, J. F., Kwan, K. Y., O'Roak, B. J., Baek, D. Y., Stillman, A. A., Morgan, T. M., …, State, M. W. (2005) Sequence variants in SLITRK1 are associated with Tourette's syndrome. Science 310(5746), 317-20.

Abdellaoui, A., Hottenga, J.-J., De Knijff, P., Nivard, M. G., Xiao, X., Scheet, P., …, Boomsma, D. I. (2013). Population structure, migration, and diversifying selection in the Netherlands. European Journal of Human Genetics 21(10), 1277–1285.

Adam, Y., Meinlschmidt, G., Gloster, A. T., & Lieb, R. (2012). Obsessive-compulsive disorder in the community: 12-month prevalence, comorbidity and impairment. Social Psychiatry and Psychiatric Epidemiology 47(3), 339–349.

Albin, R. L., & Mink, J. W. (2006). Recent advances in Tourette syndrome research. Trends in neurosciences 29(3), 175-182.

Alonso, P., Pujol, J., Ph, D., Cardoner, N., Benlloch, L., Deus, J., … Vallejo, J. (2001). Right Prefrontal Repetitive Transcranial Magnetic Stimulation in Obsessive-Compulsive Disorder: A Double-Blind, Placebo-Controlled Study. American Journal of Psychiatry 158(7), 1143–1145.

Altshuler, D., Brooks, L. D., Chakravarti, A., Collins, F. S., Daly, M. J., & Donnelly, P. (2005). A haplotype map of the human genome. Nature 437, 1299–1320.

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders Third Edition (DSM-III). Washington (DC): American Psychiatric Association.

American Psychiatric Association (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington (DC): American Psychiatric Association.

American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders, Fourth Edition, Text Revision. Washington (DC): American Psychiatric Association.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: American Psychiatric Association.

Anckarsäter, H., Lundström, S., Kollberg, L., Kerekes, N., Palm, C., Carlström, E., … Lichtenstein, P. (2011). The Child and Adolescent Twin Study in Sweden (CATSS). Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies 14, 495–508.

Angst, J., Gamma, A., Endrass, J., Goodwin, R., Ajdacic, V., Eich, D., & Rössler, W. (2004). Obsessive-compulsive severity spectrum in the community: prevalence, comorbidity, and course. European Archives of Psychiatry and Clinical Neuroscience 254(3), 156-64.

Angst, J., Gamma, A., Endrass, J., Hantouche, E., Goodwin, R., Ajdacic, V., … Rössler, W. (2005). Obsessive-compulsive syndromes and disorders: Significance of comorbidity with bipolar and anxiety syndromes. European Archives of Psychiatry and Clinical Neuroscience 255(1), 65–71.

Anon, 1993. Definitions and classification of tic disorders. The Tourette Syndrome Classification Study Group. Archives of neurology 50(10), 1013–6.

Apter, A., Pauls, D L., Bleich, A., Zohar, A H., Kron, S., Ratzoni, G., Dycian, A., Kotler, M.,

Weizman, A., Gadot, N. An epidemiologic study of Gilles de la Tourette's syndrome in Israel. Archives of General Psychiatry 50(9), 734–738.

Arnold, P. D., Sicard, T., Burroughs, E., Richter, M. A., Kennedy, J. L. (2006). Glutamate transporter gene SLC1A1 associated with obsessive-compulsive disorder. JAMA Psychiatry 63(7), 769-76.

Awaad, Y. (1999). Tics in Tourette syndrome: new treatment options. J Child Neurol 14(5), 316-319.

Ayuso-mateos, J. L. (2001). Global burden of obsessive-compulsive disorder in the year 2000, 1–10.

Bar-Dayan, Y., Arnson, Y., & Elishkevits, K. (2010). Screening for Common Neurologic Diseases Among Israeli Adolescents. Journal of Child Neurology 25(3), 348–351.

Barr, C. L., Wigg, K. G., Zovko, E., Sandor, P., & Tsui, L. C. (1997). Linkage study of the dopamine D5 receptor gene and Gilles de la Tourette syndrome. American Journal of Medical Genetics 74(1), 58–61.

Benjamini, Y., & Hochberg, Y. (1997). Multiple Hypotheses Testing with Weights. Scandinavian Journal of Statistics 24(3), 407–418.

Bertelsen, B., Melchior, L., Jensen, L. R., Groth, C., Glenthøj, B., Rizzo, R., … Tümer, Z. (2014). Intragenic deletions affecting two alternative transcripts of the IMMP2L gene in patients with Tourette syndrome.

European Journal of Human Genetics : EJHG 22(11), 1283–9.
Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., . . . Shen, R. (2011). High density DNA methylation array with single CpG site resolution. Genomics 98(4), 288-295.

Bloch, M. H., & Leckman, J. F. (2009). Clinical course of Tourette syndrome. Journal of Psychosomatic Research 67(6), 497–501.

Boker, S., Neale, M., Maes, H., Wilde, M., Spiegel, M., Brick, T., … Fox, J. (2011). OpenMx: An Open Source Extended Structural Equation Modeling Framework. Psychometrika 76(2), 306-317.

Bolhuis, K., McAdams, T. a, Monzani, B., Gregory, a M., Mataix-Cols, D., Stringaris, a, & Eley, T. C. (2014). Aetiological overlap between obsessive-compulsive and depressive symptoms: a longitudinal twin study in adolescents and adults. Psychological Medicine 44(7), 1439–49.

Bolton, D., Rijsdijk, F., O'Connor, T G., Perrin, S., Eley, T C., 2007. Obsessive-compulsive disorder, tics and anxiety in 6-year-old twins. Psychological medicine 37(1), 39–48.

Bolton, D., Rijsdijk, F., Eley, T. C., O'Connor, T. G., Briskman, J., & Perrin, S. (2009). Normative childhood repetitive routines and obsessive compulsive symptomatology in 6-year-old twins. Journal of Child Psychology and Psychiatry, and Allied Disciplines 50(9), 1139–46.

Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. Nat Rev Genet 3(11), 872–882.

Boomsma, D. I., de Geus, E. J. C., Vink, J. M., Stubbe, J. H., Distel, M. a, Hottenga, J.-J., … Willemsen, G. (2006). Netherlands Twin Register: from twins to twin families. Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies (9), 849–857.

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An Expanded View of Complex Traits: From Polygenic to Omnigenic. Cell 196(7), 1177-1186.

Bulik-sullivan, B. K., Loh, P., Finucane, H., Ripke, S., Daly, M. J., Price, A. L., … Alexandra, P. (2015). LD Score Regression Distinguishes Confounding from Polygenicity in Genome-Wide Association Studies 47(3), 291–295.

Burd, L. K. (2001). Long-term follow-up of an epidemiologically defined cohort of patients with Tourette syndrome. Journal of Child Neurology 16, 431–437.

Canavera, K. E., Ollendick, T. H., May, J. T., & Pincus, D. B. (2010). Clinical correlates of comorbid obsessive-compulsive disorder and depression in youth. Child Psychiatry and Human Development 41(6), 583–594.

Cappi, C., Muniz, R. K., Sampaio, A. S., Cordeiro, Q., Brentani, H., Palácios, S. A., … Hounie, A. G. (2012). Association study between functional polymorphisms in the TNF-alpha gene and obsessive-compulsive disorder. Arquivos de Neuro-Psiquiatria 70(2), 87-90.

Carey, G., & Gottesman, I. I. (1981). Twin and familiy studies of anxiety, phobic, and obsessive disorders. In In D. F. Klein & J. Rabkin (Eds.) (Ed.), Anxiety: New research and changing concepts (pp. 117–136). New York: Raven Press.

Cath, D. C., Nizar, K., Boomsma, D., & Mathews, C. A. (2017). Age-Specific Prevalence of Hoarding and Obsessive Compulsive Disorder: A Population-Based Study. American Journal of Geriatric Psychiatry 25(3), 245–255.

Cath, D. C., van Grootheest, D. S., Willemsen, G., van Oppen, P., & Boomsma, D. I. (2008). Environmental factors in obsessive-compulsive behavior: evidence from discordant and concordant monozygotic twins. Behavior Genetics 38(2), 108–20.

Cath, D. C., Hedderly, T., Ludolph, A. G., Stern, J. S., Murphy, T., Hartmann, A., … Wolanczyck, T. (2011). European clinical guidelines for Tourette Syndrome and other tic disorders. Part I: Assessment. European Child and Adolescent Psychiatry 20(4), 155-71.

Cath, D. C., & Ludolph, A. G. (2013). Other Psychiatric Comorbidities in Tourette Syndrome. In D. Martino & J. F. Leckman (Eds.), Tourette Syndrome (pp. 74–106). Oxford University Press.

Cath, D. C., Nizar, K., Boomsma, D. I., Mathews, C. A. (2016). Age-Specific Prevalence of Hoarding and Obsessive Compulsive Disorder: A Population-Based Study. American Journal of Geriatric Psychiatry 25(3), 245-255.

Cavanna, A. E., & Rickards, H. (2013). The psychopathological spectrum of Gilles de la Tourette syndrome. Neuroscience and Biobehavioral Reviews 37(6), 1008-1015.

Chen, H., Wang, C., Conomos, M. P., Stilp, A. M., Li, Z., Sofer, T., … Lin, X. (2016). Control for Population Structure and Relatedness for Binary Traits in Genetic Association Studies via

Logistic Mixed Models. American Journal of Human Genetics 98(4), 653–666.

Chen, Y. A., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., ... Weksberg, R. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics 8(2), 203-209.

Christensen, J. H., Elfving, B., Muller, H. K., Fryland, T., Nyegaard, M., Corydon, T. J., ... Borglum, A. D. (2012). The Schizophrenia and Bipolar Disorder associated BRD1 gene is regulated upon chronic restraint stress. Eur Neuropsychopharmacol 22(9), 651-656.

Cichon, S., Mühleisen, T. W., Degenhardt, F. a, Mattheisen, M., Miró, X., Strohmaier, J., … Nöthen, M. M. (2011). Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. American Journal of Human Genetics 88(3), 372–381.

Clifford, C. A., Murray, R. M., & Fulker, D. W. (1984). (1984). Genetic and environmental influences on obsessional traits and symptoms. Psychological Medicine 14(4), 791–800.

Como, P., LaMarsh, J., & O'Brien, K. (2005). Obsessive-compulsive disorder in Tourette's syndrome. Advances in Neurology 96, 249–261.

Cubo E., Gabriel y Galán JM., Villaverde V A., Velasco S S., Benito V D., Macarrón J V., Guevara J C., Louis E D., Benito-León J. (2011). Prevalence of tics in schoolchildren in central Spain: a population-based study. Pediatric Neurology 45(2), 100–108.

Cuthbert, B. N. (2015). Research Domain Criteria: Toward future psychiatric nosologies. Dialogues in Clinical Neuroscience 17(1), 89–97.

Davis, L. K., Yu, D., Keenan, C. L., Gamazon, E. R., Konkashbaev, A. I., Derks, E. M., … Scharf, J. M. (2013). Partitioning the heritability of Tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. PLoS Genetics 9(10).

Davies, M. N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., ... Mill, J. (2012). Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. Genome Biol 13(6).

Darrow, S. M., Hirschtritt, M. E., Davis, L. K., Illmann, C., Osiecki, L., Grados,..., Mathews, C. A., Tourette Syndrome Association International Consortium for Genetics. (2016). Identification of Two Heritable Cross-Disorder Endophenotypes for Tourette Syndrome. American Journal of Psychiatry 174(4), 387-396.

Darwin, C. (1859). On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life. London: J. Murray.

de Bruijn, C., Beun, S., de Graaf, R., ten Have, M., & Denys, D. (2010). Subthreshold symptoms and obsessive-compulsive disorder: evaluating the diagnostic threshold. Psychological Medicine 40(6), 989–997.

de Haan, M. J., Delucchi, K. L., Mathews, C. M., & Cath, D. C. (2015). Tic symptom dimensions and their heritabilities in Tourette's syndrome. Psychiatric genetics 25(3), 112-118.

Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., … Neale, B.

M. (2017). Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. bioRxiv.

Dempster, E. L., Pidsley, R., Schalkwyk, L. C., Owens, S., Georgiades, A., Kane, F., . . . Toulopoulou, T. (2011). Disease-associated epigenetic changes in monozygotic twins discordant for schizophrenia and bipolar disorder. Human molecular genetics 20(24), 4786-96.

den Braber, A., van 'T Ent, D., Cath, D. C., Wagner, J., Boomsma, D. I., & de Geus, E. J. C. (2010). Brain activation during cognitive planning in twins discordant or concordant for obsessive-compulsive symptoms. Brain 133(10), 3123–3140.

den Braber A., Zilhão NR, Fedko IO, Hottenga JJ, Pool R, Smit DJ, … Boomsma DI. (2016). Obsessive-compulsive symptoms in a large population-based twin-family sample are predicted by clinically based polygenic scores and by genome-wide SNPs. Translational Psychiatry 6.

Dobson, A.J. & Barnett, A. (2008). An introduction to generalized linear models 3rd ed., Chapman and Hall/CRC.

Dongen, J. van, Slagboom, P. E., Draisma, H. H. M., & Martin, N. G. (2012). The continuing value of twin studies in the omics era. Nature Publishing Group 13(9), 640–653.

Doyon, Y., Cayrou, C., Ullah, M., Landry, A.-J., Côté, V., Selleck, W., . . . Côté, J. (2006). ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. Molecular cell 21(1), 51-64.

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. PLoS Genet 9(3).

Eapen, V., Laker, M., Anfield, A., Dobbs, J., & Robertson, M. M. (2001). Prevalence of tics and Tourette syndrome in an inpatient adult psychiatry setting. Journal of Psychiatry & Neuroscience 26(5), 417–20.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., & Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics 10, 48.

Eichler, E. E., Flint, J., Gibson, G., Kong, A., Leal, S. M., Moore, J. H., & Nadeau, J. H. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. Nature Reviews. Genetics 11(6), 446–450.

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 and Allied Disciplines 44, 945–960.

Engle, R. F. (1983). Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics. In Z. Griliches & M. D. Intriligator (Eds.), Handbook of Econometrics (pp. 796–801).

Ercan-Sencicek, A. G., Stillman, A. A., Ghosh, A. K., Bilguvar, K., O'Roak, B. J., Mason, C. E., … State, M. W. (2010). L-histidine decarboxylase and Tourette's syndrome. The New England Journal of Medicine 362, 1901–1908.

Erenberg, G., Cruse, R. & Rothner, A., 1987. The natural history of Tourette syndrome: a follow-up study. Annals of neurology 22(3), 383–385.

Eisen, J. L., Sibrava, N. J., Boisseau, C. L., Mancebo, M. C., Stout, R. L., Pinto, A., & Rasmussen,

S. A. (2013). Five-year course of obssesive-compulsive disorder: predictors of remission and relapse Journal of Clinical Psychiatry 74(3), 233–239.

Evangelou, E., Fellay, J., Colombo, S., Martinez-Picado, J., Obel, N., Goldstein, D. B., … Ioannidis, J. P. a. (2011). Impact of phenotype definition on genome-wide association signals: empirical evaluation in human immunodeficiency virus type 1 infection. American Journal of Epidemiology 173, 1336–42.

Falconer, D. S. (1965). The inheritance of liability to certain diseases, estimated from the incidence among relatives. Annals of Human Genetics 29, 51–76.

Falconer, D. S. (1967). The inheritance of liability to diseases with variable age of onset, with particular reference to diabetes mellitus. Annals of Human Genetics 31, 1–20.

Falconer, D. S., & Mackay, T. F. C. (1996). Introduction to Quantitative Genetics. Longman.

Fatemi, S. H., Folsom, T. D., Reutiman, T. J., & Thuras, P. D. (2009). Expression of GABAB receptors is altered in brains of subjects with autism. The Cerebellum 8(1), 64-69.

Feinberg, A. P., & Irizarry, R. A. (2010). Stochastic epigenetic variation as a driving force of development, evolutionary adaptation, and disease. Proceedings of the National Academy of Sciences 107, 1757-1764.

Felling, R. J., & Singer, H. S. (2011). Neurobiology of tourette syndrome: current status and need for further investigation. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience 31(35), 12387–12395.

Fernández de la Cruz, L., Barrow, F., Bolhuis, K., Krebs, G., Volz, C., Nakatani, E., … Mataix-Cols, D. (2013). Sexual obsessions in pediatric obsessive-compulsive disorder: clinical characteristics and treatment outcomes. Depression and Anxiety 30(8), 732–40.

Fernandez, T., Sanders, S., Yurkiewicz, I., Ercan-Sencicek, A., Kim, Y., Fishman, D., … State, M. (2012). Rare Copy Number Variants in Tourette Syndrome Disrupt Genes in Histaminergic Pathways and Overlap with Autism. Biological Psychiatry 71(5), 392–402.

Findley, D B., Leckman, J F., Katsovich, L., Lin, H., Zhang, H., Grantz, H., Otka, J., Lombroso, P J., King, R a., 2003. Development of the Yale Children's Global Stress Index (YCGSI) and its application in children and adolescents ith Tourette's syndrome and obsessive-compulsive disorder. Journal of the American Academy of Child and Adolescent Psychiatry 42(4), 450–457.

Fineberg, N. a, Hengartner, M. P., Bergbaum, C., Gale, T., Rössler, W., & Angst, J. (2013). Lifetime comorbidity of obsessive-compulsive disorder and sub-threshold obsessive-compulsive symptomatology in the community: impact, prevalence, socio-demographic and clinical characteristics. International Journal of Psychiatry in Clinical Practice 17(3), 188–96.

Fisher, R. A. (1918). The Correlation between Relatives on the Supposition of Mendelian Inheritance. Philosophical Transactions of the Royal Society of Edinburgh 52(2), 399–433.

Fortin, J. P., Labbe, A., Lemire, M., Zanke, B. W., Hudson, T. J., Fertig, E. J., . . . Hansen, K. D. (2014). Functional normalization of 450k methylation array data improves replication in large cancer studies. Genome Biol 15(12), 503.

Francioli, L. C., Menelaou, A., Pulit, S. L., van Dijk, F., Palamara, P. F., Elbers, C. C., … Wijmenga, C. (2014). Whole-genome sequence variation, population structure and demographic history of the Dutch population. Nature Genetics 46(8), 818–825.

Freeman, R. D., Aabech, H., Allaghband-Rad, J., Berlin, C., Bruun, R., Budman, C., … Zinner, S. (2007). Tic disorders and ADHD: Answers from a world-wide clinical dataset on Tourette syndrome. European Child and Adolescent Psychiatry 1, 15-23.

Freeman, R. D., Fast, D. K., Burd, L., Kerbeshian, J., Robertson, M. M., & Sandor, P. (2000). An international perspective on Tourette syndrome: selected findings from 3,500 individuals in 22 countries. Developmental Medicine and Child Neurology 42(7), 436–447.

Frost, R. O., & Hartl, T. L. (1996). A cognitive-behavioral model of compulsive hoarding. Behaviour Research and Therapy 34(4), 341–350.
Frost, R. O., Steketee, G., Williams, L. (1999). Hoarding : a community health problem 8(4), 229–234.

Frost, R. O., Steketee, G., & Tolin, D. F. (2011). Comorbidity in hoarding disorder. Depression and Anxiety 28(10):876-84.

Fullana, M. a, Tortella-Feliu, M., Caseras, X., Taberner, J., Torrubia, R., & Mataix-Cols, D. (2007). Temporal stability of obsessive-compulsive symptom dimensions in an undergraduate sample: a prospective 2-year follow-up study. Behavior Modification 31(6), 815–24.

Galton, F. (1869). Hereditary Genius: An Inquiry into its Laws and consequeunces. Macmillan and Co. London.

Gamazon, E. R., Wheeler, H. E., Shah, K., Mozaffari, S. V, Aquino-Michaels, K., Carroll, R. J., … Consortium, Gte. (2015). PrediXcan: Trait Mapping Using Human Transcriptome Regulation. bioRxiv.

Geels, L. M., Vink, J. M., van Beek, J. H., Bartels, M., Willemsen, G., & Boomsma, D. I. (2013). Increases in alcohol consumption in women and elderly groups: evidence from an epidemiological study. BMC Public Health 13(1), 207.

Geller, D. A. (2006). Obsessive-compulsive and spectrum disorders in children and adolescents. The Psychiatric Clinics of North America 29(2), 353–70.

Gilmour, A. R., Thompson, R., & Cullis, B. R. (1995). Average Information REML: An Efficient Algorithm for Variance Parameter Estimation in Linear Mixed Models. Biometrics 51(4), 1440.

Gormley, P., Anttila, V., Winsvold, B. S., Palta, P., Esko, T., Pers, T. H., … Palotie, A. (2016). Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. Nature Genetics 48(8), 856–866.

Grados, M. A., Riddle, M. A., Samuels, J. F., Liang, K. Y., Hoehn-Saric, R., Bienvenu, O. J., … Nestadt, G. (2001). The familial phenotype of obsessive-compulsive disorder in relation to tic disorders: The Hopkins OCD family study. Biological Psychiatry 50, 559–565.

Hackett, J. A., Sengupta, R., Zylicz, J. J., Murakami, K., Lee, C., Down, T. A., & Surani,

M. A. (2013). Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. Science 339(6118), 448-452.

Hafner, A., Obermajer, N., & Kos, J. (2010). gamma-1-syntrophin mediates trafficking of gamma-enolase towards the plasma membrane and enhances its neurotrophic activity. Neurosignals 18(4), 246-258.

Hanna, G. L., Veenstra-VanderWeele, J., Cox, N. J., Boehnke, M., Himle, J. a, Curtis, G. C., … Cook, E. H. (2002). Genome-wide linkage analysis of families with obsessive-compulsive disorder ascertained through pediatric probands. American Journal of Medical Genetics 114(5), 541–52.

Hanna, G. L., Veenstra-Vanderweele, J., Cox, N. J., van Etten, M., Fischer, D. J., Himle, J. a, … Cook, E. H. (2007). Evidence for a susceptibility locus on chromosome 10p15 in early-onset obsessive-compulsive disorder. Biological Psychiatry 62(8), 856–62.

Hansson, S. L., Röjvall, A. S., Rastam, M., Gillberg, C., Gillberg, C., & Anckarsäter, H. (2005). Psychiatric telephone interview with parents for screening of childhood autism - Tics, attention-deficit hyperactivity disorder and other comorbidities (A-TAC): Preliminary reliability and validity. British Journal of Psychiatry 187, 262-7.

Hasstedt, S J., Leppert, M., Filloux, F., van de Wetering, B J., McMahon, W M. (1995). Intermediate inheritance of Tourette syndrome, assuming assortative mating. American journal of human genetics 57(3), 682-9.

Haworth, C. M. A., & Davis, O. S. P. (2014). From observational to dynamic genetics. Frontiers in Genetics 6.

Heard, E., Clerc, P., & Avner, P. (1997). X-chromosome inactivation in mammals. Annual review of genetics 31(1), 571-610.

Hegyi, H. (2013). GABBR1 has a HERV-W LTR in its regulatory region–a possible implication for schizophrenia. Biol. Direct 8(5).

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.

Hirschtritt, M E., Lee, P C., Pauls, D L., Dion, Y., Grados, M a., Illmann, C., King, R a., Sandor, P., McMahon, W M., Lyon, G J., Cath, D C., Kurlan, R., Robertson, M M., Osiecki, L., Scharf, J M., Mathews, C a. (2015). Lifetime Prevalence, Age of Risk, and Genetic Relationships of Comorbid Psychiatric Disorders in Tourette Syndrome. JAMA Psychiatry 72(4), 325-33.

Hirschtritt, M. E., Darrow, S. M., Illmann, C., Osiecki, L., Grados, M., Sandor, P., … Mathews, C. A. (2016). Social disinhibition is a heritable subphenotype of tics in Tourette syndrome. Neurology, 87(5), 497–504.

Hirtz, D., Thurman, D. J., Gwinn-Hardy, K., Mohamed, M., Chaudhuri, a. R., Zalutsky, R., 2007. How common are the "common" neurologic disorders? Neurology 68, 326–337.

Hoggart, C. J., Clark, T. G., De Iorio, M., Whittaker, J. C., & Balding, D. J. (2008). Genome-wide significance for dense SNP and resequencing data. Genetic Epidemiology 32(2), 179–185.

Horwath, E., & Weissman, M. M. (2000). The epidemiology and cross-national presentation of obsessive-compulsive disorder. The Psychiatric Clinics of North America 23, 493–507.

Hu, X., Lipsky, R. H., Zhu, G., Akhtar, L. A., Taubman, J., Greenberg, B. D., Xu, K., Arnold, P. D., Richter, M. A., Kennedy, J. L., Murphy, D. L., Goldman, D. (2006) Serotonin Transporter Promoter Gain-of-Function Genotypes Are Linked to Obsessive-Compulsive Disorder. American Journal of Human Genetics 78(5), 815-526.

Hudziak, J. J., Van Beijsterveldt, C. E. M., Althoff, R. R., Stanger, C., Rettew, D. C., Nelson, E. C., … Boomsma, D. I. (2004). Genetic and environmental contributions to the Child Behavior Checklist Obsessive-Compulsive Scale: a cross-cultural twin study. Archives of General Psychiatry 61, 608–616.

Hyde, T. M., Aaronson, B. A., Randolph, C., Rickler, K. C., & Weinberger, D. R. (1992). Relationship of birth weight to the phenotypic expression of Gilles de la Tourette's syndrome in monozygotic twins. Neurology 42, 652–658.

Iervolino, A. C., Perroud, N., Fullana, M. A., Guipponi, M., Cherkas, L., Collier, D. a., & Mataix-Cols, D. (2009). Prevalence and heritability of compulsive hoarding: A twin study. American Journal of Psychiatry 166, 1156–1161.

Iervolino, A. C., Rijsdijk, F. V, Cherkas, L., Fullana, M. a, & Mataix-Cols, D. (2011). A multivariate twin study of obsessive-compulsive symptom dimensions. Archives of General Psychiatry 68(6), 637–644.

Ivanov, V. Z., Mataix-Cols, D., Serlachius, E., Lichtenstein, P., Anckarsäter, H., Chang, Z., … Rück, C. (2013). Prevalence, comorbidity and heritability of hoarding symptoms in adolescence: a population based twin study in 15-year olds. PloS One 8.

Jeffries, A. R., Perfect, L. W., Ledderose, J., Schalkwyk, L. C., Bray, N. J., Mill, J., & Price, J. (2012). Stochastic choice of allelic expression in human neural stem cells. Stem Cells 30(9), 1938-1947.

Jonnal, A. H., Gardner, C. O., Prescott, C. A., & Kendler, K. S. (2000). Obsessive and compulsive symptoms in a general population sample of female twins. American Journal of Medical Genetics - Neuropsychiatric Genetics 96, 791–796.

Karagiannidis, I., Dehning, S., Sandor, P., Tarnok, Z., Rizzo, R., Wolanczyk, T., … Paschou, P. (2013). Support of the histaminergic hypothesis in Tourette syndrome: association of the histamine decarboxylase gene in a large sample of families. Journal of Medical Genetics 50(11), 760–764.

Karagiannidis, I., Rizzo, R., Tarnok, Z., Wolanczyk, T., Hebebrand, J., Nöthen, M. M., … Paschou, P. (2012). Replication of association between a SLITRK1 haplotype and Tourette Syndrome in a large sample of families. Molecular Psychiatry 17(7), 665–8.

Karno, M., Golding, J. M., Sorenson, S. B., & Burnam, M. a. (1988). The epidemiology of obsessive-compulsive disorder in five US communities. Archives of General Psychiatry 45(12), 1094–9.

Katerberg, H., Delucchi, K. L., Stewart, S. E., Lochner, C., Denys, D. A. J. P., Stack, D. E., … Cath, D. C. (2010). Symptom dimensions in OCD: Item-level factor analysis and heritability

estimates. Behavior Genetics 40(4), 505–517.

Khalifa, N., & von Knorring, A.-L. (2005). Tourette syndrome and other tic disorders in a total population of children: clinical assessment and background. Acta Paediatrica 94, 1608–14.

Khalifa, N. (2006). Psychopathology in a Swedish Population of School Children With Tic Disorders. Journal of the American Academy of Child & Adolescent Psychiatry 45(11), 1346–1353.

Klein, B. J., Lalonde, M.-E., Côté, J., Yang, X.-J., & Kutateladze, T. G. (2014). Crosstalk between epigenetic readers regulates the MOZ/MORF HAT complexes. Epigenetics 9(2), 186-193.

Klengel, T., Mehta, D., Anacker, C., Rex-Haffner, M., Pruessner, J. C., Pariante, C. M., . . . Binder, E. B. (2013). Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. Nat Neurosci 16(1), 33-41.

Knight, T., Steeves, T., Day, L., Lowerison, M., Jette, N., Pringsheim, T. (2012). Prevalence of tic disorders: A systematic review and meta-analysis. Pediatric Neurology 47, 77–90.

Kraft JT., Dalsgaard S., Obel C., Thomsen PH., Henriksen TB., Scahill L. (2012). Prevalence and clinical correlates of tic disorders in a community sample of school-age children. European Child & Adolescent Psychiatry 21(1) 5–13.

Kurlan R., McDermott MP., Deeley C., Como PG., Brower C., Eapen S., Andresen EM., Miller B. (2001). Prevalence of tics in schoolchildren and association with placement in special education. Neurology 57(8), 1383–1388.

Kurlan, R., Eapen, V., Stern, J., McDermott, M P., Robertson, M M. (1994). Bilineal transmission in Tourette's syndrome families. Neurology 44(12), 2336–42.

Lander, E. S., & Botstein, S. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121(1), 185.

Lange, J. (1929). Leistungen der Zwillingpathologie für die Psychiatrie [The importance of twin pathology for psychiatry]. Allgemeine Zeitschrift Für Psychiatrie Und Psychisch-Gerichtliche Medicin (90), 122–142.

Langley, A. K., Lewin, A. B., Bergman, R. L., Lee, J. C., & Piacentini, J. (2010). Correlates of comorbid anxiety and externalizing disorders in childhood obsessive compulsive disorder. European Child and Adolescent Psychiatry 19(8), 637–645.

Larson, T., Anckarsäter, H., Gillberg, C., Ståhlberg, O., Carlström, E., Kadesjö, B., … Gillberg, C. (2010). The Autism - Tics, ADHD and other Comorbidities inventory (A-TAC): further validation of a telephone interview for epidemiological research. BMC Psychiatry 1–11.

Leckman, J. F., Bloch, M. H., & King, R. A. (2009). Symptom dimensions and subtypes of obsessive-compulsive disorder: a developmental perspective. Dialogues in Clinical Neuroscience 11(1), 21-33.

Leckman, J. F., Riddle, M. A., Hardin, M. T., Ort, S. I., Swartz, K. L., Stevenson, J., & Cohen, D. J. (1989). The Yale Global Tic Severity Scale: initial testing of a clinician-rated scale of tic severity. Journal of the American Academy of Child & Adolescent Psychiatry 28(4), 566-573.

Leckman, J. F., Zhang, H., Vitale, A., Lahnin, F., Lynch, K., Bondi, C., … Peterson, B. S. (1998). Course of tic severity in Tourette syndrome: the first two decades. Pediatrics 102, 14–9.

Lee, K. W., & Pausova, Z. (2013). Cigarette smoking and DNA methylation. Front Genet 4, 132.

Lehalle, D., Sanlaville, D., Guimier, A., Plouvier, E., Leblanc, T., Galmiche, L., . . . Amiel, J. (2014). Multiple congenital anomalies-intellectual disability (MCA-ID) and neuroblastoma in a patient harboring a de novo 14q23.1q23.3 deletion. Am J Med Genet A 164A(5), 1310-1317.

Leon, C. A., Schumacher, J., Kluck, N., Herold, C., Schulze, T. G., Propping, P., . . . Jamra, R. A. (2011). Association study of the GRIA1 and CLINT1 (Epsin 4) genes in a German schizophrenia sample. Psychiatr Genet 21(2), 114.

Levinson, D. F., Mostafavi, S., Milaneschi, Y., Rivera, M., Ripke, S., Wray, N. R., & Sullivan, P. F. (2014). Genetic studies of major depressive disorder: Why are there no genome-wide association study findings and what can we do about it? Biological Psychiatry 76(7), 510-2.

Li, M.-X., Gui, H.-S., Kwan, J. S. H., & Sham, P. C. (2011). GATES: a rapid and powerful gene-based association test using extended Simes procedure. American Journal of Human Genetics 88(3), 283–293.

Lichtenstein, P., Carlström, E., Råstam, M., Gillberg, C., & Anckarsäter, H. (2010). The genetics of autism spectrum disorders and related neuropsychiatric disorders in childhood. American Journal of Psychiatry 167(11), 1357-1363.

Lichter, D. G., Jackson, L. A., & Schachter, M. (1995). Clinical evidence of genomic imprinting in Tourette's syndrome. Neurology 45(5), 924-928.

Liu, S., Cui, J., Zhang, X., Wu, W., Niu, H., Ma, X., … Yi, M. (2014). Variable number tandem repeats in dopamine receptor D4 in Tourette's syndrome. Movement Disorders : Official Journal of the Movement Disorder Society 29(13), 1687–91.

Liu, J., Morgan, M., Hutchison, K., & Calhoun, V. D. (2010). A study of the influence of sex on genome wide methylation. PLoS One 5(4).

Liu, X., Cheng, R., Verbitsky, M., Kisselev, S., Browne, A., Mejia-Sanatana, H., . . . Lee, J. H. (2011). Genome-wide association study identifies candidate genes for Parkinson's disease in an Ashkenazi Jewish population. BMC Med Genet 12, 104.

Lochner, C., Kinnear, C., Hemmings, S., Seller, C., Niehaus, D., Knowles, J., … Stein, D. (2005). Hoarding in obsessive-compulsive disorder: clinical and genetic correlates. Journal of Clinical Psychiatry 66(9), 1155–1160.

Luo, C., Rajput, A., Robinson, C., & Rajput, A. (2012). Gamma-aminobutyric acid (GABA)-B receptor 1 in cerebellar cortex of essential tremor. Journal of Clinical Neuroscience 19(6), 920-921.

Lybaek, H., Oyen, N., Fauske, L., & Houge, G. (2008). A 2.1 Mb deletion adjacent but distal to a 14q21q23 paracentric inversion in a family with spherocytosis and severe learning difficulties. Clin Genet 74(6), 553-559.

Macdonald, A. M., Murray, R. M., & Clifford, C. A. (1991). The contribution of heredity to obsessional disorder and personality: A review of family and twin study evidence. In & M. J. L. M. T. Tsuang, K. S. Kendler (Ed.), Genetic issues in psychosocial epi- demiology (pp. 191–212). New Brunswick, NJ: Rutgers University Press.

Maher, B. (2008). Personal genomes: The case of the missing heritability. Nature 456(7218), 18–21.

Maier, V., Jolicoeur, C., Rayburn, H., Takegahara, N., Kumanogoh, A., Kikutani, H., ..., Friedel, R. H. (2011). Semaphorin 4C and 4G are ligands of Plexin-B2 required in cerebellar development. Mol Cell Neurosci 46(2), 419-431.

Mancebo, M. C., Boisseau, C. L., Garnaat, S. L., Eisen, J. L., Greenberg, B. D., Sibrava, N. J., … Rasmussen, S. a. (2014). Long-term course of pediatric obsessive-compulsive disorder: 3years of prospective follow-up. Comprehensive Psychiatry 55(7), 1498–504.

Manchia, M., Cullis, J., Turecki, G., Rouleau, G. a., Uher, R., & Alda, M. (2013). The Impact of Phenotypic and Genetic Heterogeneity on Results of Genome Wide Association Studies of Complex Diseases. PLoS ONE 8, 1–7.

Marazziti, D., Mungai, F., Masala, I., Baroni, S., Vivarelli, L., Ambrogi, F., … Dell'Osso, L. (2009). Normalisation of immune cell imbalance after pharmacological treatments of patients suffering from obsessive-compulsive disorder. Journal of Psychopharmacology (Oxford, England) 23(5), 567–573.

Marazziti, D., Presta, S., Pfanner, C., Gemignani, A., Rossi, A., Sbrana, S., … Cassano, G. B. (1999). Immunological alterations in adult obsessive-compulsive disorder. Biological Psychiatry 46(6), 810–814.

Mardis, E. R. (2011). A decade's perspective on DNA sequencing technology. Nature 470(7333), 198–203.

Martin, N. G., & Eaves, L. J. (1977). The genetical analysis of covariance structure. Heredity 38(1), 79–95.

Martin, N. G., Eaves, L. J., Kearsey, M. J., & Davies, P. (1978). The power of the classical twin study. Heredity 40(1), 97–116.

Mataix-Cols, D., Rosario-Campos, M. C., & Leckman, J. F. (2005). A multidimensional model of obsessive-compulsive disorder. American Journal of Psychiatry 162(2), 228-38.

Mataix-Cols, D., Frost, R. O., Pertusa, A., Clark, L. A., Saxena, S., Leckman, J. F., … Wilhelm, S. (2010). Hoarding disorder: a new diagnosis for DSM-V? Depression and Anxiety 27, 556–572.

Mataix-Cols, D., Isomura, K., Pérez-Vigil, A., Chang, Z., Rück, C., Larsson, K. J., … Lichtenstein, P. (2015). Familial Risks of Tourette Syndrome and Chronic Tic Disorders: A Population-Based Cohort Study. JAMA Psychiatry 72(8), 787-93.

Mather, K. (1974). Non-allelic interaction in continuous variation of randomly breeding populations. Heredity 32, 414–419.

Mathews, C a., Scharf, J M., Miller, L L., Macdonald-Wallis, C., Lawlor, D a., Ben-Shlomo, Y., 2014. Association between pre- and perinatal exposures and tourette syndrome or chronic tic disorder in the alspac cohort. British Journal of Psychiatry 204, 40–45.

Mathews, C. a., & Grados, M. a. (2011). Familiality of tourette syndrome, obsessive-compulsive disorder, and attention-deficit/hyperactivity disorder: Heritability analysis in a large sib-pair sample. Journal of the American Academy of Child and Adolescent Psychiatry 50(1), 46–54.

Mathews, C. a., Delucchi, K., Cath, D. C., Willemsen, G., & Boomsma, D. I. (2014). Partitioning the etiology of hoarding and obsessive–compulsive symptoms. Psychological Medicine 44(13):2867-76.

Mathews, C. a, Badner, J. a, Andresen, J. M., Sheppard, B., Himle, J. a, Grant, J. E., … Hanna, G. L. (2012). Genome-wide linkage analysis of obsessive-compulsive disorder implicates chromosome 1p36. Biological Psychiatry 72(8), 629–36.

Mattheisen, M., Samuels, J. F., Wang, Y., Greenberg, B. D., Fyera, J., McCracken, J. T., … Nestadt, G. (2014). Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. Molecular Psychiatry 20(3), 337-44

McGrath, L. M., Yu, D., Marshall, C., Davis, L. K., Thiruvahindrapuram, B., Li, B., … Scharf, J. M. (2014). Copy number variation in obsessive-compulsive disorder and tourette syndrome: a cross-disorder study. Journal of the American Academy of Child and Adolescent Psychiatry 53(8), 910–9.

Mérette, C., Brassard, A., Potvin, A., Bouvier, H., Rousseau, F., Emond, C., … Caron, C. (2000). Significant linkage for Tourette syndrome in a large French Canadian family. American Journal of Human Genetics 67, 1008–1013.

Micali, N., Heyman, I., Perez, M., Hilton, K., Nakatani, E., Turner, C., & Mataix-Cols, D. (2010). Long-term outcomes of obsessive-compulsive disorder: follow-up of 142 children and adolescents. The British Journal of Psychiatry : The Journal of Mental Science 197(2), 128–34.

Mill, J., & Heijmans, B. T. (2013). From promises to practical strategies in epigenetic epidemiology. Nat Rev Genet 14(8), 585-594.

Miller, L L., Scharf, J M., Mathews, C a., Ben-Shlomo, Y., 2014. Tourette syndrome and chronic tic disorder are associated with lower socio-economic status: Findings from the Avon Longitudinal Study of Parents and Children cohort. Developmental Medicine and Child Neurology 56, 157–163.

Minica, C. C., Dolan, C. V, Kampert, M. M. D., Boomsma, D. I., & Vink, J. M. (2014). Sandwich corrected standard errors in family-based genome-wide association studies. European Journal of Human Genetics : EJHG 23(3):388-94.

Mink, J. W. (2001). Basal ganglia dysfunction in Tourette's syndrome: a new hypothesis. Pediatric neurology, 25(3), 190-198.

Miranda, D. M., Wigg, K., Feng, Y., Sandor, P., & Barr, C. L. (2008). Association study between Gilles de la Tourette Syndrome and two genes in the Robo-Slit pathway located in the chromosome 11q24 linked/associated region. American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics : The Official Publication of the International Society of

Psychiatric Genetics 147B, 68–72.

Moore, K., McKnight, A. J., Craig, D., & O'Neill, F. (2014). Epigenome-wide association study for Parkinson's disease. Neuromolecular medicine 16(4), 845-855.

Moayyeri, A., Hammond, C. J., Hart, D. J., & Spector, T. D. (2013). The UK Adult Twin Registry (TwinsUK Resource). Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies 16(1), 144–9.

Mühleisen, T. W., Mattheisen, M., Strohmaier, J., Degenhardt, F., Priebe, L., Schultz, C. C., … Cichon, S. (2012). Association between schizophrenia and common variation in neurocan (NCAN), a genetic risk factor for bipolar disorder. Schizophrenia Research 138(1), 69–73.

Munafò, M. R. (2006). Candidate gene studies in the 21st century: meta-analysis, mediation, moderation. Genes, Brain, and Behavior 1:3-8
Nag, A., Bochukova, E. G., Kremeyer, B., Campbell, D. D., Muller, H., Valencia-Duarte, A. V., … Ruiz-Linares, A. (2013). CNV Analysis in Tourette Syndrome Implicates Large Genomic Rearrangements in COL8A1 and NRXN1. PLoS ONE 8(3).

Nelson, E. C., Hanna, G. L., Hudziak, J. J., Botteron, K. N., Heath, a. C., & Todd, R. D. (2001). Obsessive-Compulsive Scale of the Child Behavior Checklist: Specificity, Sensitivity, and Predictive Power. Pediatrics 108(1).

Nestadt, G. (2000). A Family Study of Obsessive-compulsive Disorder. Archives of General Psychiatry 7(4), 358-63.

Nordsletten, A. E., Monzani, B., Fern??ndez de la Cruz, L., Iervolino, A. C., Fullana, M. A., Harris, J., … Mataix-Cols, D. (2013). Overlap and specificity of genetic and environmental influences on excessive acquisition and difficulties discarding possessions: Implications for hoarding disorder. American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics 162(4), 380–387.

Nyholt, D. R. (2014). SECA: SNP effect concordance analysis using genome-wide association summary results. Bioinformatics 30(14), 2086–2088.

Ooki, S. (2005). Genetic and environmental influences on stuttering and tics in Japanese twin children. Twin Research and Human Genetics 8(1), 69-75.

Orloff, L. M., Baer, L., Ivanjack, L., Battle, M. A., Buttolph, M. L., & Jenike, A. (1994). Long-Term follow-Up of 85 patients with obsessive-compulsive disorder 441–442.

Ozaki, N., Goldman, D., Kaye, W. H., Plotnicov, K., Greenberg, B. D., Lappalainen, J., Rudnick, G., Murphy, D. L. (2003). Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype Molecular Psychiatry 8(11), 933-6.

Pallanti, S., Grassi, G., Sarrecchia, E. D., Cantisani, A., & Pellegrini, M. (2011). Obsessive-compulsive disorder comorbidity: Clinical assessment and therapeutic implications. Frontiers in Psychiatry 21, 2-70.

Paschou, P. (2013). The genetic basis of Gilles de la Tourette Syndrome. Neurosci Biobehav Rev 37(6), 1026-1039.

Patrick, C. J., & Hajcak, G. (2016). Reshaping clinical science: Introduction to the Special Issue on Psychophysiology and the NIMH Research Domain Criteria (RDoC) initiative. Psychophysiology 53(3), 281–285.

Pauls, D. L., Alsobrook, J. P., Goodman, W., Rasmussen, S., & Leckman, J. F. (1995). A family study of obsessive-compulsive disorder. The American Journal of Psychiatry 152, 76–84.

Pauls, D. L., & Hurst, C. (1993). Schedule for Tourette and Other Behavioral Disorders, Revised. New Haven, CT: Child Study Center, Yale University.

Pauls, D. L., Fernandez, T. V., Mathews, C. A., State, M. W., & Scharf, J. M. (2014). The Inheritance of Tourette Disorder: A review. J Obsessive Compuls Relat Disord 3(4), 380-385.

Pauls, D. & Hurst, C., 1996. The Yale family genetic study self-report questionnaire for tics, obsessive compulsiveness, attentional difficulties, impulsivity and motor hyperactivity. In Schedule for Tourette and Other Behavioural Syndromes. New Haven: New Haven, CT: Child Study Center, Yale University School of Medicine.

Pauls, D. L., Raymond, C. L., Stevenson, J. M., & Leckman, J. F. (1991). A family study of Gilles de la Tourette syndrome. American Journal of Human Genetics 48, 154–163.

Pauls, D. L. (2010). The genetics of obsessive-compulsive disorder: a review. Dialogues in Clinical Neuroscience 12, 149–163.

Pauls, D. L., & Leckman, J. F. (1986). The inheritance of Gilles de la Tourette's syndrome and associated behaviors. Evidence for autosomal dominant transmission. The New England Journal of Medicine 315(16), 993–7.

Perroud, N., Guipponi, M., Pertusa, A., Fullana, M. A., Iervolino, A. C., Cherkas, L., … Mataix-Cols, D. (2011). Genome-wide association study of hoarding traits. American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics 156(2), 240–242.

Pertusa, A., Fullana, M. a., Singh, S., Alonso, P., Menchón, J. M., & Mataix-Cols, D. (2008). Compulsive hoarding: OCD symptom, distinct clinical syndrome, or both? American Journal of Psychiatry 165, 1289–1298.

Peterson, B., Riddle, M. A., Cohen, D. J., Katz, L. D., Smith, J. C., Hardin, M. T., & Leckman, J. F. (1993). Reduced basal ganglia volumes in Tourette's syndrome using three-dimensional reconstruction techniques from magnetic resonance images. Neurology 43(5), 941–9.

Pinto R., Monzani B., Leckman J.F., Rück C., Serlachius E., Lichtenstein P., Mataix-Cols D. (2016). Understanding the Covariation of Tics , Attention-Deficit / Hyperactivity , and Obsessive-Compulsive Symptoms : A Population-Based Adult Twin Study. Am J Med Genet B Neuropsychiatr Genet 171(7), 938-47.

Plomin, R., Haworth, C. M. a, & Davis, O. S. P. (2009). Quantitative Traits. Genetics (Vol. 10).

Posthuma, D. & Boomsma, D.I., 2000. A note on the statistical power in extended twin designs. Behavior Genetics 30, 147–158.

Prescott, C. A. (2004). Using the Mplus Computer Program to Estimate Models for Continuous and Categorical Data from Twins. Behavior Genetics 34(1).

Price, R A., Kidd, K K., Cohen, D J., Pauls, D L., Leckman, J F. (1985). A twin study of Tourette syndrome. Archives of general psychiatry 42,815–20.

Polanczyk, G., De Lima, M. S., Horta, B. L., Biederman, J., & Rohde, L. A. (2007). The worldwide prevalence of ADHD: A systematic review and metaregression analysis. American Journal of Psychiatry 164(6), 942–948.

Posner, J., Marsh, R., Maia, T. V, Peterson, B. S., & Simpson, H. B. (2014). Reduced functional connectivity within the limbic cortico-striato-thalamo-cortical loop in unmedicated adults with obsessive-compulsive disorder 35(6), 2852–2860.

Price, R. A., Kidd, K. K., Cohen, D. J., Pauls, D. L., & Leckman, J. F. (1985). A twin study of Tourette syndrome. Archives of General Psychiatry 42, 815–20.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics 81(3), 559–575.

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

Rakyan, V. K., Down, T. A., Balding, D. J., & Beck, S. (2011). Epigenome-wide association studies for common human diseases. Nature Reviews Genetics 12(8), 529-541.

Rao, D. C., & Morton, N. E. (1974). Path Analysis of Family Resemblance in the Presence of Gene-Environment Interaction. American Journal of Human Genetics 26(118), 767–772.

Rasmussen, S., & Eisen, J. (1989). Clinical features and phenomenology of obsessive compulsive disorder. Psychiatric Annals 19(2), 67–73.

Reddy, Y. C. J., D'Souza, S. M., Shetti, C., Kandavel, T., Deshpande, S., Badamath, S., & Singisetti, S. (2005). An 11- to 13-year follow-up of 75 subjects with obsessive-compulsive disorder. The Journal of Clinical Psychiatry 66, 744–749.

Ribak, C., Vaughn, J., & Roberts, E. (1979). The GABA neurons and their axon terminals in rat corpus striatum as demonstrated by GAD immunocytochemistry. Journal of Comparative Neurology 187(2), 261-283.

Richter, M. A., de Jesus, D. R., Hoppenbrouwers, S., Daigle, M., Deluce, J., Ravindran, L. N., . . . Daskalakis, Z. J. (2012). Evidence for cortical inhibitory and excitatory dysfunction in obsessive compulsive disorder. Neuropsychopharmacology 37(5), 1144-1151.

Rietveld MJ., van Der Valk JC., Bongers IL., Stroet TM., Slagboom PE., Boomsma DI., 2000. Zygosity diagnosis in young twins by parental report. Twin research : the official journal of the International Society for Twin Studies 3, 134–41.

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T., Farh, K., Holmans, P. A., Lee, P., … O'Donovan, M. C. (2014). Biological Insights From 108 Schizophrenia-Associated Genetic Loci. Nature 511(7510), 421–427.

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kähler, A. K., Akterin, S., … Sullivan, P. F. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. Nature Genetics 45(10), 1150–9.

Risch, N., & Merikangas, K. (1996). The future of genetic studies of complex human diseases. Science 273(5281), 1516–1517.

Robertson, M. M. (2000). Tourette syndrome, associated conditions and the complexities of treatment. Brain 123(3), 425–462.

Robertson, M. M. (2008). The prevalence and epidemiology of Gilles de la Tourette syndrome. Part 1: The epidemiological and prevalence studies. Journal of Psychosomatic Research 65(5), 461-72.

Robertson, M. M., Verril, M., Mercer, M., James, B., Pauls, D L., 1994. Tourette's syndrome in New Zealand. A postal survey. The British journal of psychiatry: the journal of mental science 164, 263–266.

Robertson, M. M., Eapen, V. & Cavanna, A.E., 2009. The international prevalence, epidemiology, and clinical phenomenology of Tourette syndrome: A cross-cultural perspective. Journal of Psychosomatic Research 67, 475–483.

Robertson, M. M., Verril, M., Mercer, M., James, B., & Pauls, D. L. (1994). Tourette's syndrome in New Zealand. A postal survey. The British Journal of Psychiatry : The Journal of Mental Science 164, 263–266.

Rocha, F. F., Correa, H., & Teixeira, A. L. (2008). Obsessive-compulsive disorder and immunology: A review. Progress in Neuro-Psychopharmacology and Biological Psychiatry 32(5), 1139-1146.

Rosario-Campos, M. C., Leckman, J. F., Curi, M., Quatrano, S., Katsovitch, L., Miguel, E. C., & Pauls, D. L. (2005). A family study of early-onset obsessive-compulsive disorder. American Journal of Medical Genetics - Neuropsychiatric Genetics 136B(1), 92-7.

Ross, J., Badner, J., Garrido, H., Sheppard, B., Chavira, D. a, Grados, M., … Mathews, C. a. (2011). Genomewide linkage analysis in Costa Rican families implicates chromosome 15q14 as a candidate region for OCD. Human Genetics 130(6), 795–805.

Rotge, J. Y., Aouizerate, B., Tignol, J., Bioulac, B., Burbaud, P., & Guehl, D. (2010). The glutamate-based genetic immune hypothesis in obsessive-compulsive disorder. An integrative approach from genes to symptoms. Neuroscience 165(2), 408-17.

Ruderfer, D. M., Fanous, A. H., Ripke, S., McQuillin, A., Amdur, R. L., Gejman, P. V, … Kendler, K. S. (2014). Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. Molecular Psychiatry 19(9), 1017–1024.

Ruscio, A., Stein, D., Chiu, W., & Kessler, R. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. Molecular Psychiatry 15, 53–63.

Sabb, F W., Burggren, A C., Higier, R G., Fox, J., He, J., Parker, D S., Poldrack, R A., Chu, W., Cannon, T D., Freimer, N B., Bilder, R M. 2009. Challenges in phenotype definition in the whole-genome era: multivariate models of memory and intelligence. Neuroscience 164,

88–107.

Samuels, J., Bienvenu, O. J., Riddle, M. A., & Cullen, B. A. M. (2002). Hoarding in obsessive compulsive disorder : results from a case-control study 40, 517–528.

Samuels, J. F., Bienvenu, O. J., Grados, M. A., Cullen, B., Riddle, M. A., Liang, K., … Nestadt, G. (2008). Prevalence and correlates of hoarding behavior in a community-based sample. Behaviour Research and Therapy 46(7), 836–844.

Sanavio, E. (1988). Obsessions and compulsions: the Padua Inventory. Behaviour Research and Therapy 26, 169–177.

Sapountzi, V., & Cote, J. (2011). MYST-family histone acetyltransferases: beyond chromatin. Cell Mol Life Sci 68(7), 1147-1156.

Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C. C. A., Patsopoulos, N. A., Moutsianas, L., … Wtccc, C. C. (2012). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476(7359), 214–219.

Scahill, L., Dalsgaard, S., & Bradbury, K. (2013). The prevalence of tourette syndrome and its relationship to clinical features. In D. Martino & J. F. Leckman (Eds.), Tourette Syndrome (pp. 121–133). Oxford: Oxford University Press.

Scahill L., Williams S., Schwab-Stone M., Applegate J., Leckman JF., 2006. Disruptive behavior problems in a community sample of children with tic disorders. Advances in Neurology 99, 184–190.

Scahill, L., Sukhodolsky, D. G., Williams, S. K., & Leckman, J. F. (2005). Public health significance of tic disorders in children and adolescents. Adv Neurol 96, 240–248.

Scharf, J. M., Yu, D., Mathews, C. a, Neale, B. M., Stewart, S. E., Fagerness, J. a, … Pauls, D. L. (2013). Genome-wide association study of Tourette's syndrome. Molecular Psychiatry 18(6), 721–8.

Schlander, M., Schwarz, O., Rothenberger, a, & Roessner, V. (2011). Tic disorders: administrative prevalence and co-occurrence with attention-deficit/hyperactivity disorder in a German community sample. European Psychiatry : The Journal of the Association of European Psychiatrists 26, 370–4.

Schrötter, A., Mastalski, T., Nensa, F. M., Neumann, M., Loosse, C., Pfeiffer, K., . . . Theiss, C. (2013). FE65 regulates and interacts with the Bloom syndrome protein in dynamic nuclear spheres–potential relevance to Alzheimer's disease. Journal of cell science 126(11), 2480-2492.

Schultz, C. C., Mühleisen, T. W., Nenadic, I., Koch, K., Wagner, G., Schachtzabel, C., … Schlösser, R. G. M. (2014). Common variation in NCAN, a risk factor for bipolar disorder and schizophrenia, influences local cortical folding in schizophrenia. Psychological Medicine 44(4), 811–20.

Sekar, A., Bialas, A. R., de Rivera, H., Davis, A., Hammond, T. R., Kamitaki, N., …, McCarroll, S. A. (2016). Schizophrenia risk from complex variation of complement component 4. Nature 530(7589), 177–183.

Severinsen, J., Bjarkam, C. R., Kiar-Larsen, S., Olsen, I. M., Nielsen, M. M., Blechingberg, J., ...,Young, B. D. (2006). Evidence implicating BRD1 with brain development and susceptibility to both schizophrenia and bipolar affective disorder. Molecular psychiatry 11(12), 1126-1138.

Shapiro, A. K., Shapiro, E., Gerald Young, J., & Feinberg, T. E. (1988). Gilles De La Tourette Syndrome (2nd ed.). Raven Press Ltd.

Shimada-Sugimoto, M., Otowa, T. & Hettema, J.M., 2015. Genetics of anxiety disorders: Genetic epidemiological and molecular studies in humans. Psychiatry and Clinical Neurosciences 388–401.

Shugart, Y. Y., Samuels, J., Willour, V. L., Grados, M. a, Greenberg, B. D., Knowles, J. a, … Nestadt, G. (2006). Genomewide linkage scan for obsessive-compulsive disorder: evidence for susceptibility loci on chromosomes 3q, 7p, 1q, 15q, and 6q. Molecular Psychiatry 11(8), 763–70.

Simonic, I., Gericke, G. S., Ott, J., & Weber, J. L. (1998). Identification of genetic markers associated with Gilles de la Tourette syndrome in an Afrikaner population. American Journal of Human Genetics 63, 839–846.

Simonic, I., Nyholt, D. R., Gericke, G. S., Gordon, D., Matsumoto, N., Ledbetter, D. H., … Weber, J. L. (2001). Rapid Publication Further Evidence for Linkage of Gilles de la Tourette Syndrome (GTS) Susceptibility Loci on Chromosomes 2p11 , 8q22 , and 11q23-24 in South African Afrikaners 167, 163–167.

Singer, H. S., Wendlandt, J., Krieger, M., & Giuliano, J. (2001). Baclofen treatment in Tourette syndrome A double-blind, placebo-controlled, crossover trial. Neurology 56(5), 599-604.

Smith, S., Hay, El H., Farhat, N., Rekaya, R. (2013). Genome wide association studies in presence of misclassified binary responses. BMC genetics 14, 124.

Smoller, J. W., Kendler, K. S., Craddock, N., & Consortium C-DGotPG. (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. The Lancet 6736(12), 62129-1

Skoog, G., & Skoog, I. (1999). A 40-Year Follow-up of Patients With Obsessive-compulsive Disorder 56, 121–127.

Skre, I., Onstad, S., Torgersen, S., Lygren, S., & Kringlen, E. (1993). A twin study of DSM-III-R anxiety disorders. Acta Psychiatrica Scandinavica 88(2), 85-92.

Sowell, E. R., Kan, E., Yoshii, J., Thompson, P. M., Bansal, R., Xu, D., . . . Peterson, B. S. (2008). Thinning of sensorimotor cortices in children with Tourette syndrome. Nature neuroscience 11(6), 637-639.

Spector, T. D., & Williams, F. M. (2006). The UK Adult Twin Registry (TwinsUK). Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies 9, 899–906.

Srivastava, S., Bhatia, M. S., Thawani, R., & Jhanjee, A. (2011). Quality of life in patients with obsessive compulsive disorder: A longitudinal study from India. Asian Journal of Psychiatry 4, 178–182.

Steketee, G., Eisen, J., Dyck, I., & Warshaw, M. (1999). Predictors of course in obsessive-compulsive disorder. Psychiatry Research 89, 229–238.

Stewart, S. E., Geller, D. A., Jenike, M., Pauls, D., Shaw, D., & Mullin, B. (2004). Long-term outcome of pediatric obsessive–compulsive disorder : a meta-analysis and qualitative review of the literature. Acta Psychiatr Scand 110(1):4-13.

Stewart, S. E., Illmann, C., Geller, D. a, Leckman, J. F., King, R., & Pauls, D. L. (2006). A controlled family study of attention-deficit/hyperactivity disorder and Tourette's disorder. Journal of the American Academy of Child and Adolescent Psychiatry 45(11), 1354–1362.

Stewart, S. E., Yu, D., Scharf, J. M., Neale, B. M., Fagerness, J. a, Mathews, C. a, … Pauls, D. L. (2013). Genome-wide association study of obsessive-compulsive disorder. Molecular Psychiatry 18(7), 788–98.

Stewart, S. E., Fagerness, J. A., Platko, J., Smoller, J. W., Scharf, J. M., Illmann, C., Jenike, E., Chabane, N., Leboyer, M., Delorme, R., Jenike, M. A., Pauls, D. L. (2007) Association of the SLC1A1 glutamate transporter gene and obsessive-compulsive disorder. American Journal of Medical Genetics 144B(8), 1027-33.

Storch, E. A., Lewin, A. B., Geffken, G. R., Morgan, J. R., & Murphy, T. K. (2010). The role of comorbid disruptive behavior in the clinical expression of pediatric obsessive-compulsive disorder. Behaviour Research and Therapy 48(12), 1204–1210.

Swain, J.E. & Leckman, J.F., 2005. Tourette Syndrome and Tic Disorders. Psychiatry (Edgmont) 2(7), 26–36.

Swedo, S. E., Leonard, H. L., Garvey, M., Mittleman, B., Allen, a J., Perlmutter, S., … Dubbert, B. K. (1998). Pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections: clinical description of the first 50 cases. The American Journal of Psychiatry 155(2), 264–71.

Taylor, S., Jang, K., & Asmundson, G. (2010). Etiology of obsessions and compulsions: A behavioral-genetic analysis. Journal of Abnormal Psychology 119, 672–682.

Taylor, S. (2011). Etiology of obsessions and compulsions: A meta-analysis and narrative review of twin studies. Clinical Psychology Review. 31(8), 1361-72.

Taylor, S. (2013). Molecular genetics of obsessive-compulsive disorder: a comprehensive meta-analysis of genetic association studies. Molecular Psychiatry 18(7), 799–805.

Timpano, K., Exner, C., Glaesmer, H., Rief, W., Keshaviah, A., Brahler, E., & Wilhelm, S. (2011). The epidemiology of the proposed DSM-5 hoarding disorder: exploration of the acquisition specifier, associated features, and distress. Journal of Clinical Psychiatry 72(6), 780–786.

The Tourette Syndrome Association International Consortium for Genetics. (1999). A Complete Genome Screen in Sib Pairs Affected by Gilles de la Tourette Syndrome. The American Journal of Human Genetics 65(5), 1428–1436.

The Tourette Syndrome Association International Consortium for Genetics. (2007). Genome

scan for Tourette disorder in affected-sibling-pair and multigenerational families. American Journal of Human Genetics 80(2), 265–72.

The Tourette Syndrome Classification Study Group. (1993). Definitions and classification of tic disorders. Arch Neurol 50(10), 1013-1016.

Toffolo, M. B. J., van den Hout, M. a., Hooge, I. T. C., Engelhard, I. M., & Cath, D. C. (2013). Mild Uncertainty Promotes Checking Behavior in Subclinical Obsessive-Compulsive Disorder. Clinical Psychological Science 1(2), 103–109.

Tolin, D. F., Frost, R. O., & Steketee, G. (2010). A brief interview for assessing compulsive hoarding : The Hoarding Rating Scale-Interview. Psychiatry Research 178(1), 147–152.

Tolin, D. F., Frost, R. O., Steketee, G., & Fitch, K. E. (2008). Family burden of compulsive hoarding : Results of an internet survey 46, 334–344.

Torres, A. R., Prince, M. J., Bebbington, P. E., Bhugra, D., Brugha, T. S., Farrell, M., … Singleton, N. (2006). Obsessive-compulsive disorder: prevalence, comorbidity, impact, and help-seeking in the British National Psychiatric Morbidity Survey of 2000. The American Journal of Psychiatry 163, 1978–1985.

Towbin, K. (2010). Tic disorders. Dulcan's Textbook of Child and Adolescent Psychiatry: American Psychiatric Publishing, Inc.

Ullah, M., Pelletier, N., Xiao, L., Zhao, S. P., Wang, K., Degerny, C., …, Yang, X. J. (2008). Molecular architecture of quartet MOZ/MORF histone acetyltransferase complexes. Mol Cell Biol 28(22), 6828-6843.

Ustun, T. B. (2004). Global burden of depressive disorders in the year 2000. The British Journal of Psychiatry 184(5), 386–392.

van Beijsterveldt, C. E., Groen-Blokhuis, M., Hottenga, J. J., Franic, S., Hudziak, J. J., Lamb, D., …, Boomsma, D. I. (2013). The Young Netherlands Twin Register (YNTR): longitudinal twin and family studies in over 70,000 children. Twin Res Hum Genet 16(1), 252-267.

van den Hout, M., & Kindt, M. (2003). Repeated checking causes memory distrust. Behaviour Research and Therapy 41(3), 301–316.

van Dongen, J., Nivard, M. G., Willemsen, G., Hottenga, J.-J., Helmer, Q., Dolan, C. V., … Boomsma, D. I. (2016). Genetic and environmental influences interact with age and sex in shaping the human methylome. Nature Communications, 7, 11115.

van Grootheest, Bartels, M., Cath, D. C., Beekman, A. T., Hudziak, J. J., & Boomsma, D. I. (2007). Genetic and environmental contributions underlying stability in childhood obsessive-compulsive behavior. Biological Psychiatry 61(3), 308–15.

van Grootheest, D., Bartels, M., Cath, D C., Beekman, A T., Hudziak, J J., Boomsma, D I. (2007). Genetic and environmental influences on obsessive-compulsive symptoms in adults: a population-based twin-family study. Psychological medicine 37, 1635–1644.

van Grootheest, D. S., Cath, D. C., Beekman, A. T., & Boomsma, D. I. (2005). Twin studies on obsessive-compulsive disorder: a review. Twin Research and Human Genetics : The Official

Journal of the International Society for Twin Studies 8(5), 450–458.

van Grootheest, D. S., Bartels, M., van Beijsterveldt, C. E. M., Cath, D. C., Beekman, A. T., Hudziak, J. J., & Boomsma, D. I. (2008). Genetic and environmental contributions to self-report obsessive-compulsive symptoms in Dutch adolescents at ages 12, 14, and 16. Journal of the American Academy of Child and Adolescent Psychiatry, 47(10), 1182–8.

van Grootheest, D., Cath, D C., Hottenga, J. J., Beekman, A. T., Boomsma, D I. (2009). Genetic factors underlie stability of obsessive-compulsive symptoms. Twin Res Hum Genet 2009 12(5), 411-9.

van Iterson, M., Tobi, E. W., Slieker, R. C., den Hollander, W., Luijk, R., Slagboom, P. E., & Heijmans, B. T. (2014). MethylAid: visual and interactive quality control of large Illumina 450k datasets. Bioinformatics 30(23), 3435-3437.

van Oppen, P. (1992). Obsessions and compulsions: dimensional structure, reliability, convergent and divergent validity of the Padua Inventory. Behaviour Research and Therapy 30, 631–637.

van Oppen, P., Hoekstra, R. J., & Emmelkamp, P. M. (1995). The structure of obsessive-compulsive symptoms. Behaviour research and therapy 33(1):15-23.

Vassar, R. (2013). ADAM10 prodomain mutations cause late-onset Alzheimer's disease: not just the latest FAD. Neuron 80(2), 250-253.

Velzen, L. S. Van, Vriend, C., Wit, S. J. De, & Heuvel, O. A. Van Den. (2014). Response inhibition and interference control in obsessive–compulsive spectrum disorders 8, 1–22.

Vinkhuyzen, A. A. E., Wray, N. R., Yang, J., Goddard, M. E., & Visscher, P. M. (2013). Estimation and Partition of Heritability in Human Populations Using Whole-Genome Analysis Methods. Annual Review of Genetics 47(1), 75–95.

Voyiaziakis, E., Evgrafov, O., Li, D., Yoon, H. J., Tabares, P., Samuels, J., Wang, Y., Riddle, M. A., Grados, M. A., Bienvenu, O. J., Shugart, Y. Y., Liang, K. Y., Greenberg, B. D., Rasmussen, S. A., Murphy, D. L., Wendland, J. R., McCracken, J. T., Piacentini, J., Rauch, S. L., Pauls, D. L., Nestadt, G., Fyer, A. J., Knowles, J. A. (2011), Molecular Psychiatry 16, 108–120.

Walkup, J.T., Ferrao, Y., Leckman, J.F., Stein, D.J., Singer, H., 2010. Tic disorders: some key issues for DSM-V. Depression and anxiety 27,600-610.

Walter, H. (2017). Research Domain Criteria (RDoC). Der Nervenarzt 88(5), 538-548.

Wang, K. S., Liu, X. F., & Aragam, N. (2010). A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. Schizophr Res 124(1-3), 192-199.

Wang, Z., Ph, D., Maia, T. V, Ph, D., Marsh, R., Ph, D., … Peterson, B. S. (2011). The Neural Circuits That Generate Tics in Tourette's Syndrome 168(12), 1326–1337.

Weissman, M. M., Bland, R. C., Canino, G. J., Greenwald, S., Hwu, H. G., Lee, C. K., … Wickramaratne, P. J. (1994). The cross national epidemiology of obsessive compulsive disorder. The Cross National Collaborative Group. The Journal of Clinical Psychiatry 55, 5–10.

Wendland, J. R., Moya, P. R., Timpano, K. R., Anavitarte, A. P., Kruse, M. R., Wheaton, M. G., Ren-Patterson, R. F., Murphy, D. L. (2009). A haplotype containing quantitative trait loci for SLC1A1 gene expression and its association with obsessive-compulsive disorder. JAMA Psychiatry 66(4), 408-16.

Wenning, G. K., Kiechl, S., Seppi, K., Müller, J., Högl, B., Saletu, M., … Poewe, W. (2005). Prevalence of movement disorders in men and women aged 50-89 years (Bruneck Study cohort): A population-based study. Lancet Neurology 4, 815–820.

World Health Organization. (2007). World Health Statistics 2007. In World Health Statistics, 1–97

World Health Organization. (1992). The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines. Geneva: World Health Organization.

Willemsen, G., Vink, J. M., Abdellaoui, A., den Braber, A., van Beek, J. H. D. a, Draisma, H. H. M., …, Boomsma, D. I. (2013). The Adult Netherlands Twin Register: twenty-five years of survey and biological data collection. Twin Research and Human Genetics : The Official Journal of the International Society for Twin Studies 16, 271–81.

Willemsen, G., de Geus, E. J., Bartels, M., van Beijsterveldt, C. E., Brooks, A. I., Estourgie-van Burk, G. F., …, Boomsma, D. I. (2010). The Netherlands Twin Register biobank: a resource for genetic epidemiological studies. Twin Res Hum Genet 13(3), 231-245.

Willour, V. L., Shugart, Y. Y., Samuels, J., Grados, M., Cullen, B., Iii, O. J. B., …, Nestadt, G. (2004). Replication Study Supports Evidence for Linkage to 9p24 in Obsessive- Compulsive Disorder. Am J Hum Genet 75(3), 508–513.

Winsberg, M. E., Cassic, K. S., & Koran, L. M. (1999). Hoarding in obsessive-compulsive disorder: a report of 20 cases. The Journal of Clinical Psychiatry 60, 591–597.

Wockner, L. F., Noble, E. P., Lawford, B. R., Young, R. M., Morris, C. P., Whitehall, V. L., & Voisey, J. (2014). Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. Transl Psychiatry 4.

Wray, N.R., Lee, S.H. & Kendler, K.S., 2012. Impact of diagnostic misclassification on estimation of genetic correlations using genome-wide genotypes. European Journal of Human Genetics 20,668–674.

Wray, N. R., & Maier, R. (2014b). Genetic Basis of Complex Genetic Disease: The Contribution of Disease Heterogeneity to Missing Heritability. Current Epidemiology Reports 1(4), 220–227.

Wray, N. R., & Sullivan, P. F. (2017). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. bioRxiv.

Wright, S. (1949). The genetical structure of populations. Annals of Eugenics 15(4), 323-54.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2013). Genome-wide complex trait analysis (GCTA): Methods, data analyses, and interpretations. Methods in Molecular Biology 1019, 215–236.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., … Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. Nature Genetics 42(7), 565–9.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. American Journal of Human Genetics 88(1), 76–82.

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2013). Genome-Wide Association Studies and Genomic Prediction 1019, 215–236.
You, L., Yan, K., Zhou, J., Zhao, H., Bertos, N. R., Park, M., …, Yang, X. J. (2015). The lysine acetyltransferase activator Brpf1 governs dentate gyrus development through neural stem cells and progenitors. PLoS Genet 11(3).

Yu, D., Mathews, C. A., Neale, B. M., Ph, D., Davis, L. K., Gamazon, E. R., … Hoekstra, P. J. (2014). Cross-Disorder Genome-Wide Analyses Suggest a Complex Genetic Relationship Between Tourette's Syndrome and OCD. American Journal of Psychiatry 172(1), 82-93.

Zaitlen, N., Kraft, P., Patterson, N., Pasaniuc, B., Bhatia, G., Pollack, S., & Price, A. L. (2013). Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. PLoS Genetics, 9(5).

Zilhão, N. R., Padmanabhuni, S. S., Pagliaroli, L., Barta, C., Smit, D. J. a., Cath, D., … Boomsma, D. I. (2015). Epigenome-Wide Association Study of Tic Disorders. Twin Research and Human Genetics 18(6), 1–11.

Zilhao, N. R., Smit, D. J. a, Den Braber, A., Dolan, C. V, Willemsen, G., Boomsma, D. I., & Cath, D. C. (2014). Genetic and Environmental Contributions to Stability in Adult Obsessive Compulsive Behavior. Twin Research and Human Genetics 18(1), 52–60.

Zilhao, N. R., Smit, D. J., Boomsma, D. I., & Cath, D. C. (2016). Cross-disorder genetic analysis of tic disorders, obsessive-compulsive, and hoarding symptoms. Frontiers in Psychiatry.

Zilhão N. R., Olthof M. C., Smit D. J., Cath D. C., Ligthart L., Mathews C. A., Delucchi K., Boomsma D. I., Dolan C. V (2016). Heritability of tic disorders: a twin-family study. Psychological Medicine 47(6), 1085-1096.

Zohar, a H., Ratzoni, G., Pauls, D L., Apter, A., Bleich, A., Kron, S., Rappaport, M., Weizman, A., Cohen, D J., 1992. An epidemiological study of obsessive-compulsive disorder and related disorders in Israeli adolescents. Journal of the American Academy of Child and Adolescent Psychiatry 31, 1057–1061.

Dankwoord

I am grateful to my mentors Dorret Boomsma, Danielle Cath and Dirk Smit. They gave me guidance, and I have grown with their supervision. I owe special gratitude as well to Anouk, Jouke-Jan and Conor. I have learned a lot from their teachings and their contributions are present along this entire work. My gratitude also goes to all members of the Doctoral Examination Committee for their time devoted to reviewing this dissertation.

I want to thank everyone in the Biological Psychology Department for having welcomed me and taking me part in this project. I specially thank all my colleagues and friends from the EUROTRAIN for the company and fun times.

In the past four years I have learn more than what is comprehended in these pages. Today I very fondly recall all my memories in The Netherlands, since the first day in Amsterdam, to the last one in Alkmaar. A truthful acknowledgement then could never be comprehensive, neither an exhaustive list of every person I would want to thank to. I leave a very warm thank you to all the friends that have been with me in this period of my life.

Lastly, I want to thank my family. To my parents, Adriano and Alexandra, my brothers Miguel and Luis. And to Hrafnhildur, for your support.

Nuno
Vestmannaeyjar, Iceland, 2017.