Nordic OCD & Related Disorders Consortium (NORDiC): Rationale, Design and Methods

Journal:	*American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*
Manuscript ID	NPG-19-0023
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	24-Feb-2019
Complete List of Authors:	Mataix-Cols, David; Karolinska Institutet, Department of Clinical Neuroscience, Division of Psychiatry Hansen, Bjarne; Haukeland Universitetssjukehus, OCD-team; University of Bergen, Clinical Psychology Mattheisen, Manuel; Life and Brain GmbH, Genomics; University of Bonn, Institute of Human Genetics; Aarhus University, Biomedicine; University of Wurzburg Physiological Institute, Psychiatry, Psychosomatics and Psychotherapy Karlsson, Elinor ; Broad Institute Addington, Anjene; National Institute of Mental Health, Boberg, Julia; Karolinska Institute, Clinical Neuroscience Djurfeldt, Diana; Karolinska Institute Halvorsen, Matthew; University of North Carolina at Chapel Hill, Genetics Lichtenstein, Paul; Karolinska Institute, Medical Epidemiology and Biostatistics Solem, Stian; Haukeland University Hospital, OCD-team; Norwegian University of Science and Technology, Psychology Lindblad-Toh, Kerstin; Broad Institute; Uppsala University, Science for Life Laboratory, IMBIM none, Nordic OCD and Related Disorders Consort; University of North Carolina at Chapel Hill Haavik, Jan; University of Bergen, Department of biomedicine; University of Bergen, K. G. Jebsen Centre for Research on Neuropsychiatric Disorders Kvale, Gerd; Haukeland University Hospital, OCD-team; University of Bergen, Clinical Psychology Rück, Christian; Karolinska Institutet, Department of Clinical Neuroscience, Division of Psychiatry Crowley, James; UNC Chapel Hill, Genetics and Psychiatry
Keywords:	GWAS, Obsessive-compulsive disorder, Norway, Sweden, Denmark

NORDIC OCD & RELATED DISORDERS CONSORTIUM (NORDIC): RATIONALE, DESIGN AND METHODS

David Mataix-Cols[1,2], Bjarne Hansen[3,4], Manuel Mattheisen[5-8], Elinor K. Karlsson[9-10], Anjené M. Addington[11], Julia Boberg[1,2], Diana R. Djurfeldt[1,2], Matthew Halvorsen[12], Paul Lichtenstein[13], Stian Solem[3,14], Kerstin Lindblad-Toh[9,15], Nordic OCD and Related Disorders Consortium (NORDiC), Jan Haavik[16,17], Gerd Kvale[3,4], Christian Rück[1,2], James J. Crowley[1,12,18]

[1] Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[2] Stockholm Health Care Services, Stockholm, Sweden
[3] Haukeland University Hospital, OCD-team, Bergen, Norway
[4] Department of Clinical Psychology, University of Bergen, Bergen, Norway
[5] Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany
[6] Institute of Human Genetics, University of Bonn, Bonn, Germany
[7] Center for Integrative Sequencing, iSEQ, Department of Biomedicine, Aarhus University, Denmark
[8] Department of Psychiatry, Psychosomatics, and Psychotherapy, University of Würzburg, Germany
[9] Broad Institute of MIT and Harvard, Cambridge, MA, USA
[10] Program in Bioinformatics & Integrative Biology and Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA, USA
[11] Genomics Research Branch, National Institute of Mental Health in Bethesda, Bethesda, Maryland, USA
[12] Department of Genetics, University of North Carolina at Chapel Hill, NC, USA
[13] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[14] Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway
[15] Science for Life Laboratory, IMBIM, Uppsala University, Uppsala, Sweden
[16] Department of Biomedicine, University of Bergen, Bergen, Norway
[17] Division of Psychiatry, Haukeland University Hospital, Bergen, Norway
[18] Department of Psychiatry, University of North Carolina at Chapel Hill, NC, USA

Correspond with Dr. James J. Crowley (crowley@unc.edu)

DOCUMENT INFORMATION

Submitting to:	*American Journal of Medical Genetics Part B*
Running head:	NORDiC OCD study
Document:	NORDiC_overview_manuscript.docx
Date:	2/22/2019 3:37 PM
Abstract word count:	195 words
Body word count:	4,118 words
Additional:	3 tables, 3 figures, 0 supplemental file

Grant numbers: NORDiC is funded by NIMH R01 MH110427 (PI Crowley), NIMH R01 MH105500 (PI Crowley) and the Swedish Research Council grant # 2015-02271 (PI Mataix-Cols).

ABSTRACT

Obsessive-compulsive disorder (OCD) is a debilitating psychiatric disorder, yet its etiology is unknown and treatment outcomes could be improved if biological targets could be identified. Unfortunately, genetic findings for OCD are lagging behind other psychiatric disorders. Thus, there is a pressing need to understand the causal mechanisms implicated in OCD in order to improve clinical outcomes and to reduce morbidity and societal costs. Specifically, there is a need for a large-scale, etiologically informative genetic study integrating genetic and environmental factors that presumably interact to cause the condition. The Nordic countries provide fertile ground for such a study, given their detailed population registers, national healthcare systems and active specialist clinics for OCD. We thus formed the Nordic OCD and Related Disorders Consortium (NORDiC, www.crowleylab.org/nordic), and with the support of NIMH and the Swedish Research Council, have begun to collect a large, richly phenotyped and genotyped sample of OCD cases. Our specific aims are geared toward answering a number of key questions regarding the biology, etiology and treatment of OCD. This paper describes and discusses the rationale, design and methodology of NORDiC, including details on clinical measures and planned genomic analyses.

Key words: Obsessive-compulsive disorder, OCD, genetic, genomic, GWAS, Sweden, Norway, Denmark

RATIONALE

OCD is a neuropsychiatric disorder characterized by recurrent, unwanted thoughts (obsessions) and repetitive behaviors (compulsions), which are performed to alleviate the anxiety caused by the obsessions (American Psychiatric Association, 2013; World Health Organization, 1992). The lifetime prevalence of OCD is 1-3% (Karno, Golding, Sorenson, & Burnam, 1988; Ruscio, Stein, Chiu, & Kessler, 2010), with onset often in childhood and similar prevalence by sex. Most OCD cases have a comorbid psychiatric disorder (e.g., tic disorders, mood, anxiety disorders) (Fullana et al., 2009; Ruscio et al., 2010) and though medication and behavioral therapy are useful (Ost, Havnen, Hansen, & Kvale, 2015; Ost, Riise, Wergeland, Hansen, & Kvale, 2016), symptom control is imperfect, and the course is often chronic (Skoog & Skoog, 1999). OCD is a multidimensional disorder shown to consist of roughly four primary symptom dimensions (symmetry, forbidden thoughts, contamination, and hoarding) (Bloch, Landeros-Weisenberger, Rosario, Pittenger, & Leckman, 2008; Mataix-Cols, Rosario-Campos, & Leckman, 2005) which may have distinct neural circuitry (Mataix-Cols et al., 2004), genetic (Hasler et al., 2007) and etiological origins (Iervolino, Rijsdijk, Cherkas, Fullana, & Mataix-Cols, 2011b). Early onset and tic-related OCD may also be etiologically meaninful subtypes of the disorder (Leckman et al., 2010).

The causes of OCD have so far remained elusive. A range of perinatal risk factors are associated with a higher risk for OCD independent of shared familial confounders, suggesting that perinatal risk factors may be in the causal pathway to OCD (Brander, Perez-Vigil, Larsson, & Mataix-Cols, 2016; Brander, Rydell, et al., 2016). OCD is also highly heritable (~50%) (Bolton, Rijsdijk, O'Connor, Perrin, & Eley, 2007; Iervolino, Rijsdijk, Cherkas, Fullana, & Mataix-Cols, 2011a; Mataix-Cols et al., 2013; Pauls, 2010; van Grootheest, Cath, Beekman, & Boomsma, 2005). First-degree relatives of affected individuals have a 4-8x increased risk of OCD (Insel, Hoover, & Murphy, 1983; Mataix-Cols et al., 2013; Nestadt et al., 2000; Rasmussen & Tsuang, 1986; Rosenberg, 1967). The OCD recurrence risk in first-degree relatives is ~5 in Sweden, and ~6.5 in Denmark (Browne et al., 2015). As with other psychiatric disorders, OCD linkage studies (Hanna et al., 2002; Hanna et al., 2007; Liang et al., 2008; Mathews et al., 2012; Ross et al., 2011; Samuels et al., 2007; Shugart et al., 2006; Willour et al., 2004) (reviewed by Pauls *et al.* (Pauls, Abramovitch, Rauch, & Geller, 2014)) and >100 candidate gene studies (meta-analyzed by Taylor *et al.* (Taylor, 2013)) have produced inconsistent results. There is support for

DLGAP3, including rare and common variants in OCD and trichotillomania subjects (Bienvenu et al., 2009; Boardman et al., 2011; Zuchner et al., 2009) and excessive grooming in mice lacking expression of the ortholog *Sapap3* (Welch et al., 2007). To date, several copy number variants (CNVs) have been putatively associated with OCD (Gazzellone et al., 2016; Grunblatt et al., 2017; L. M. McGrath et al., 2014) and exome sequencing studies are underway, with the first showing an elevated rate of *de novo* mutations in OCD (Cappi et al., 2016). A recent targeted resequencing study (Noh et al., 2017) of ~600 candidate genes identified four notable genes, with *NRXN1* and *HTR2A* enriched for coding mutations and *REEP3* and *CTTNBP2* enriched for regulatory mutations in OCD patients.

The first OCD GWAS (Stewart et al., 2013) had 1,465 cases, 5,557 controls, and 400 trios. No genome-wide significant loci were identified, but polygenic risk analysis revealed overlap with Tourette's syndrome and indicated that increased sample size should reveal significant loci (Davis et al., 2013). A second OCD GWAS (Mattheisen et al., 2014) had 1,406 cases and 3,655 controls and a meta-analysis of these two studies (Collaborative, 2018) has yet to identify a genome-wide significant locus. Without a marked increase in sample size, little progress will be made (Sullivan, Daly, & O'Donovan, 2012). Since sample size is a major limiting factor (Mattheisen et al., 2014; Stewart et al., 2013), we have designed a strategy to markedly increase the worldwide numbers of genotyped OCD cases. Here we describe the Nordic OCD and Related Disorders Consortium (NORDiC), a psychiatric genetic and treatment outcome study funded by the NIMH (R01 MH110427) and the Swedish Research Council (2015-02271).

DESIGN

Figure 1 summarizes our study. In Aim 1, we will collect a large sample of OCD cases (N=10,000) in a cost-effective manner via an ongoing nationwide OCD treatment study in Norway and specialist OCD clinics in Sweden (details below). We will also have systematic treatment response and long-term follow-up data. In Aim 2, 10,000 cases will be GWAS genotyped and meta-analyzed with all external OCD GWAS data, including those from the Danish OCD and Tourette Study (DOTS) as well as the Psychiatric Genomics Consortium (PGC, www.med.unc.edu/pgc) OCD and Tourette Syndrome Working Group. We will also include a novel comparative genomic approach by taking advantage of genetic data from a canine model for OCD, canine compulsive disorder (CCD). These analyses will yield novel, robust OCD

associations and specific hypotheses about the biology of OCD. In Aim 3, we will integrate the genetic and national register data by analyzing genetic risk scores in conjunction with environmental risk factors, followed by replication across multiple Nordic countries. These analyses will yield replicated knowledge about how genes and environment mediate risk for OCD and influence treatment response.

METHODS

Aim 1

OCD definition. Cases will have a primary ICD-10 and/or DSM-5 diagnosis of OCD from a multidisciplinary specialist OCD team (established with a semi-structured instrument such as the MINI or the SCID). All patients will be included in the study regardless of psychiatric comorbidity, as long as they fulfill strict diagnostic criteria for OCD. All comorbidity will be registered. Patients will be excluded in cases of diagnostic uncertainty, such as OCD secondary to a neurological disorder or CNS insult, or where the differential diagnosis between OCD and an alternative condition is unclear.

Control definition. Unrelated to any OCD case to the third degree and unaffected with any ICD-10 / DSM-5 psychiatric disorder. Exclusion criteria are as for cases. We have maximized comparability between cases and controls, while minimizing cost. First, regarding ancestry, we will inherit >50,000 genotyped controls from Nordic countries. As with the cases, these controls were drawn from the major population centers and cover the north-south axis for each country. Second, regarding chip type, all control samples will have modern Illumina content, which is critical for proper GWAS imputation and CNV analysis.

Consent. Each subject will provide informed consent (or verbal assent from the patient and written consent from parents/legal guardians). Subjects will be asked for permission to use their personal ID numbers to perform a linkage with the various population-based health and administrative registers. They will also be asked for permission for re-contact in the future (not proposed here, but for example for future induced pluripotent stem cell work). All subjects will give permission for their DNA to be included in the Karolinska Institutet (KI) Biobank (in Sweden) or the Haukeland University Hospital (HUH) Biobank (in Norway) with indefinite storage and for their de-identified data to be made available to the scientific community, by depositing genomic and phenotype data in repositories (e.g., EU Genomics Repository or US

dbGaP) and sharing with international consortia (e.g., the Psychiatric Genomics Consortium, PGC).

Sweden: ascertainment. In both Sweden and Norway, participants will be recruited from a network of specialist OCD clinics (see ***Figure 2***) that have highly standardized assessment and treatment protocols. In Sweden's largest city, Stockholm, there are 4 large specialist OCD clinics (3 adult, 1 pediatric). Additional specialist OCD teams are scattered around the country. Ethical approvals in Sweden allow us to recontact OCD cases previously treated at these clinics, increasing the pool of potential deep-phenotyped participants. In another attempt to increase sample size, and to reach as-yet undiagnosed OCD cases, we have created a self-referral website (https://ocdgenetik.se). Participants log-in securely and fill in detailed demographic and phenotypic information before booking a brief telephone interview with a clinician to confirm diagnosis; DNA is gathered by mailing a saliva kit. In total, we plan to collect 5,000 DNA samples from Sweden over the course of this project.

Norway: ascertainment. In Norway, the government initiated a national implementation project in 2012 (led by co-investigators Drs Hansen and Kvale) to ensure the delivery of evidence-based treatment to all patients with OCD (Kvale & Hansen, 2014). Thirty OCD teams (15 adult, 15 pediatric) exist across Norway (see ***Figure 2***), and all use an identical cognitive-behavioral therapy (CBT) treatment paradigm in addition to identical assessment, diagnostic, and follow-up procedures. These assessment procedures are nearly identical to those employed in Sweden. This creates an opportunity for large-scale DNA collection from patients with deep phenotyping. In total, we plan to collect 5,000 DNA samples from Norway over the course of this project.

Denmark: ascertainment. The Danish OCD and Tourette Study (DOTS, www.crowleylab.org/dots) is an ongoing NIMH funded study (R01 MH105500, *Genetic & Environmental Predictors of OCD & Tourette's Syndrome in Denmark*) that forms the Danish arm of NORDiC (www.crowleylab.org/nordic). Its primary aim is a GWAS of ~6,500 OCD and ~4,000 Tourette syndrome cases where the DNA is gathered from neonatal blood spots (Norgaard-Pedersen & Hougaard, 2007). The genetic data are linked to Danish national medical registry data, allowing integration to identify gene-by-environment (GxE) interactions. Since DOTS uses neonatal blood spots, rather than ascertainment from clinics as done in Sweden and Norway, we do not have detailed clinical phenotypes for Danish subjects. Nonetheless, since DOTS is a population-based study of all diagnosed OCD cases for a given age group (born 1990-

present), it will provide an interesting contrast to clinic-derived OCD cases in Sweden and Norway. We have validated the accuracy of register OCD diagnoses in Denmark (Nissen et al., 2017).

Treatment. Most patients in Sweden and Norway receive specialist CBT and/or serotonin reuptake inhibitors (SRIs). Cases will receive protocol-driven CBT with 14-16 sessions delivered flexibly and supplemented with booster sessions if necessary. In addition to standard CBT treatment, at some Norwegian sites, patients may receive intensive treatment (concentrated exposure and response prevention) over the course of 4 days (Hansen, Hagen, Ost, Solem, & Kvale, 2018; Hansen, Kvale, Hagen, Havnen, & Ost, 2018; Kvale et al., 2018). Patients will be treated by psychologists, psychiatrists or psychiatric nurses specialized in OCD at one of the clinics mentioned above. In both countries, cases may also be prescribed medication (usually an SRI, alone or in combination with augmentation strategies). These treatments are recorded and drug dispensations can be further followed-up with the national prescription registers to assess adherence. About 40% of patients do not achieve sufficient symptom relief following CBT ± medication (Ost et al., 2015). Thus, our study will be powered to examine genetic and environmental predictors of short- and long-term treatment improvement under naturalistic, real-world conditions.

Clinical phenotype and treatment outcome measures. A strength of this study is the depth and consistency of phenotyping and treatment across sites, which will yield a GWAS powered to analyze comorbidity, potential subtypes, symptom dimensions and treatment response. The specialist OCD clinics across Sweden and Norway have harmonized their diagnostic and outcome assessment protocols and are now using the same set of outcome measures at baseline, post-treatment and follow-up (see ***Table 1***). These clinics employ a similar set of standardized, internationally recognized measures. This detailed phenotypic information will allow genomic and gene by environment (GxE) analyses based on particular subgroups (e.g., tic-related OCD, symptom dimensions, familial vs sporadic, etc.) as well as the use of genetic risk scores (GRS) to predict treatment outcomes and long-term follow-up. ***Table 2*** summarizes how we will measure the various sources of heterogeneity in our cohort.

We will measure symptom dimensions using both the clinician-administered Y-BOCS symptom checklist, which includes 13 major categories as well as the self-administered OCI-R (Foa et al., 2002) (for adults) or OCI-CV for young people (***Table 1***) which provides a dimensional measure

of the major symptom clusters. The OCI-R is a psychometrically sound, self-assessment instrument consisting of 18 items rated on a five-point distress scale, with six subscales (washing, checking, ordering, obsessing, hoarding and neutralizing). Similarly, the OCI-CV assesses the frequency of individual OCD symptoms as well as the severity of six correlated symptom domains (doubting/checking, obsessing, hoarding, washing, ordering and neutralizing). Each of the 21 items on the OCI-CV is rated on a 3-point scale ranging from 0 = *Never* to 2 = *Always*. With the use of two different measures of OCD symptom dimensions, we should be able to provide robust data that is replicable across different instruments and is independent of the administration modality (clinician vs self-report).

Environmental measures. Similar registers in Sweden, Norway, and Denmark form the basis of a detailed, longitudinal dataset including nearly all health care contacts since 1970 (see ***Table 3***). These registers provide an opportunity to capture a wide range of OCD-relevant environmental factors for cases presenting for care with varied symptom severity and treated at in- and outpatient settings. These registers have been used for numerous prior publications in psychiatric epidemiology and a fair fraction of what we know about the epidemiology of OCD is from these registers. Each country has a central register containing data on place of birth, historical addresses, and links to 1st, 2nd and 3rd degree relatives. For example, cases will have gestational age and size at birth, maternal smoking during pregnancy, maternal pre-pregnancy BMI, inter-pregnancy interval and maternal age at childbearing (Medical Birth Registries). From the National Patient Registries, we will get information on maternal infection in pregnancy (viral and bacterial). Paternal age at childbearing will be available for the Multi-Generation Registries. We will also collect data on family socioeconomic status (occupation, education, social welfare status, disability pension, income), marital status (cohabiting at childbirth, annual marital status), and family mobility (residential moves).

Aim 2

Genotyping. The current plan is to genotype samples on the Illumina Global Screening Array (GSA). This is Illumina's most recent array containing ~700,000 markers and is expected to perform well in a European population. As genotyping technology evolves, this selection may change.

GWAS: case-control. We will use a GWAS pipeline established by the PGC ("ricopili"). We will first examine OCD as a categorical variable (affected or unaffected) and perform disease

association testing using ricopili. We will use PLINK to analyze imputed SNP dosages with the inclusion of principal components (PCs) to control for population stratification. The results will be combined using an inverse-weighted fixed effects model (Devlin & Roeder, 1999). For chrX, we will use an additive logistic regression model with the same covariates. We will test the genome-wide distribution of the test statistic in comparison with the expected null distribution using λ_{GC} and QQ plots. λ_{GC} quantifies the extent of the bulk inflation resulting from a combination of true polygenic signal, systematic technical bias and population stratification (Bulik-Sullivan et al., 2015; Devlin & Roeder, 1999). In order to quantify the contribution of these factors, we will use LD Score regression (Bulik-Sullivan et al., 2015), where the intercept estimates the inflation in the mean chi-square that results from confounding biases, such as cryptic relatedness or population stratification. To declare genome-wide significance, we will strictly adhere to a P-value threshold of $5x10^{-8}$ (Dudbridge & Gusnanto, 2008; Pe'er, Yelensky, Altshuler, & Daly, 2008).

GWAS: phenotypic heterogeneity. Consistent phenotyping across sites (***Table 1***), allows us to analyze the genetics of biologically plausible subtypes (early onset, tic-related) and symptom dimensions (***Table 2***). As opposed to the case-control GWAS, these analyses will consider both quantitative, as well as categorical, data and are restricted to cases only. It is well recognized that reducing quantitative data to univariate surrogates generally results in a substantial loss of statistical power to detect genetic loci (Medland & Neale, 2010; Minica, Boomsma, van der Sluis, & Dolan, 2010; van der Sluis, Posthuma, Nivard, Verhage, & Dolan, 2013; van der Sluis, Verhage, Posthuma, & Dolan, 2010). One strategy to prevent this loss of power is to use a multivariate method. A number of strategies have been developed and are still being developed. We will evaluate the available options when the data are available. We will analyze baseline Y-BOCS checklist and OCI-R/OCI-CV data using the aforementioned approaches to identify genes associated with particular symptom dimensions as well as overall score. In a similar manner, we will examine additional baseline diagnostic and clinician-administered tests listed in ***Table 1***.

GWAS: meta-analysis. Following the initial GWAS, our immediate goals will be to meta-analyze the case-control results with all available OCD GWAS data, including DOTS and the latest results from the PGC OCD and Tourette Syndrome Working Group. Including NORDiC, the anticipated total number of OCD cases with GWAS data in 2022 will be at least 20,000. Genotype data from the individual studies will be subjected to unified QC, imputation and

association testing using ricopili. We will test all GWAS datasets separately for association with OCD (allowing cross-study reliability measures) and then conduct a meta-analysis of the result sets using an inverse-weighted fixed effects model (de Bakker et al., 2008).

Statistical power. NORDiC is expected to markedly increase the expected sample size for OCD GWAS by 2022. In ***Figure 3***, we show the critical role these samples will play in GWAS discovery. We assume a log additive model, lifetime risk of 0.02, $\alpha=5x10^{-8}$, and compute the minimum detectable genotypic relative risk with 80% power. The current published OCD GWASs (black) were underpowered to detect realistic effects for the majority of common complex traits (Sullivan et al., 2012). For example, the top 20 schizophrenia GWAS hits have a mean relative risk of 1.16 (range: 1.11–1.33). The addition of our Danish DOTS samples (orange) will push OCD closer to this "discovery zone", but likely not close enough. It is critical to increase power even further through NORDiC (pink) and other efforts, such as those led by the PGC OCD and Tourette Syndrome Working Group.

Comparative genomic analysis. Canine Compulsive Disorder (CCD) is a naturally occurring, face-valid model for human OCD (Overall, 2000). CCD manifests as repetition of normal canine behaviors such as grooming (lick dermatitis), predatory behaviour (tail chasing), and suckling (flank and blanket sucking) (Luescher, 2004; Moon-Fanelli, Dodman, & Cottam, 2007; Ogata et al., 2013; Overall, 2000; Overall & Dunham, 2002). As in human cases, these dogs have structural abnormalities in cortico-striato-thalamic loops (Ogata et al., 2013) and about half respond to treatment with SRIs or clomipramine (Overall & Dunham, 2002). Certain breeds have exceptionally high rates of CCD (e.g., Doberman Pinschers and German Shepherds) (Luescher, 2004; Moon-Fanelli et al., 2007; Ogata et al., 2013; Overall, 2000; Overall & Dunham, 2002). The high disease rates and limited genetic diversity of dog breeds suggest that CCD may be less complex than OCD, facilitating mapping and functional testing of associated variants (E. Axelsson et al., 2013; Karlsson et al., 2013; Tengvall et al., 2013). For example, members of NORDiC have discovered genes and biological pathways that regulate CCD risk (Dodman et al., 2010; Noh et al., 2017; Tang et al., 2014) in different breeds, and through genes nominated by this work, were able to identify *NRXN1* as the first genome-wide significant gene for OCD (Noh et al., 2017).

Through integrating genetic results for CCD with those for human OCD, we expect increased power to identify core genes regulating these phenotypes. CCD is not limited to pure-bred dogs –

mixed breeds also show various degrees of repetitive behavior, with great inter-individual variability, as seen in humans. NORDiC is partnered with a citizen-science initiative called Darwin's Dogs (https://darwinsdogs.org) that is collecting DNA and CCD-relevant behavioural phenotypes in thousands of pet dogs. To compare CCD and human OCD at the gene level, we will define canine and human genomic regions that are strongly associated with OCD using LD-based clumping (r^2>0.8) on the published canine CCD GWAS data (Dodman et al., 2010) and our human OCD GWAS. We will then identify genes that are within the strongly associated canine and human genomic regions based on the most up-to-date genome builds. Combined with our CCD sequencing studies (Noh et al., 2017; Tang et al., 2014) as well as ongoing human OCD sequencing results, we will make a list of candidate genes for canine CCD and human OCD separately, and compare to find overlapping genes. We will also perform a series of pathway analyses for CCD using the same tools and public databases that will be used for the human OCD analysis. We will then compare the lists of associated pathways in canine and human OCD to find common OCD-associated pathways across dogs and humans.

Copy number variant (CNV) analysis. Disease-associated CNVs are attractive causative mutations since, by altering gene dosage or structure, they provide both a direction of effect and molecular mechanism. Indeed, cases with autism and schizophrenia have a greater burden of large (>500 kb) rare (<1%) CNVs, particularly genic CNVs. To date, ~25 large rare recurrent CNVs of strong effect (genotypic relative risks 4-20) with consistent replication have been identified (e.g., 16p11.2 and 22q11.21) although most are multi-genic and pleotropic (Guha et al., 2013; Levinson et al., 2011; Liao et al., 2012; Malhotra & Sebat, 2012; Sullivan et al., 2012). To date, several CNVs have been putatively associated with OCD (Gazzellone et al., 2016; Grunblatt et al., 2017; L. M. McGrath et al., 2014) but larger studies are needed to see if these findings replicate. Therefore, we will use our GWAS array genotype and intensity data for CNV calling. We will evaluate CNV burden, recurrent individual CNVs, and overlap with existing psychiatric CNVs. We will validate putative OCD CNVs with an independent technology and test for replication in independent OCD samples. If we can identify even one new gene-level CNV for OCD, it will represent an important advance.

Aim 3

Overview. An exceptional feature of our Nordic samples are the range and quality of risk factor data available (***Table 1***). Work in Aim 3 will provide crucial information about how environment

modifies genetic effects. Essentially, we will merge genetic data with epidemiological data. Recent developments in statistical genetics provide a means to leverage epidemiological clues in a new way (J. J. McGrath, Mortensen, Visscher, & Wray, 2013). Genome-wide SNP genotypes can be used to calculate an individual's genetic risk score (GRS). GRS have several attractive features for GxE analyses: they provide a continuous measure with greater power than categories, and it is feasible to generate GRS sub-scores conditioned on prior hypotheses (e.g. only immune-related genes). We will calculate GRS_{OCD} for all Swedish and Norwegian OCD individuals and control samples and test for interactions with a diverse set of epidemiological and genetic epidemiological factors.

Epidemiological risk factors for OCD. A range of perinatal risk factors are associated with a higher risk for OCD independent of shared familial confounders, suggesting that perinatal risk factors may be in the causal pathway to OCD (Brander, Perez-Vigil, et al., 2016; Brander, Rydell, et al., 2016). The literature also reports that advanced paternal (Wu et al., 2012) or maternal (Steinhausen, Bisgaard, Munk-Jorgensen, & Helenius, 2013) age, a family history of autoimmune disease (Mataix-Cols et al., 2018) and childhood *Streptococcal* infections (Murphy, Storch, Lewin, Edge, & Goodman, 2012; Swedo et al., 1998) increase risk for OCD. Members of NORDiC are currently examining additional risk factors using the Swedish registers.

GRS: calculation. We will first select a high-quality set of autosomal SNPs (frequency 0.02-0.98, imputation INFO > 0.9, drop A/T or C/G SNPs, drop indels). These will then be pruned to remove SNPs in high LD (r^2>0.25 in 500kb windows) to yield a relatively independent set of high-quality SNPs (~100K). Using these lists, we will follow standard practice (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013) to compute GRS_{OCD} (using the PLINK --score algorithm) at multiple P-value threshold filters (P_T = 0.0001, 0.001, 0.01, 0.1, and 1.0). We will use ancestry PCs calculated in the GWAS. For GRS_{OCD}, we will select one P_T value for analysis that is uncorrelated with ancestry PCs (Dudbridge, 2013). To validate GRS_{OCD}, we will use all published OCD samples as a discovery set and our pooled Swedish/Norwegian GWAS as a target set. We will determine the proportion of variance in OCD attributable to these GRS.

GRS: interaction analysis. First, we will use logistic regression to evaluate the relationship between risk factor and disease. Second, we will add GRS_{OCD} to these models and empirically determine whether they replace other variables or have improved predictive power. For example, results from other psychiatric disorders suggest that cases and controls with a positive family

history will have a higher GRS (Agerbo et al., 2015; Lu et al., 2018). Therefore, the addition of GRS_{OCD} to the model may improve predictive power or replace family history as a risk factor for OCD. We will test the other risk factors mentioned above and any other environmental factors identified by our colleagues.

GRS: treatment outcome. About 40% of patients do not achieve sufficient symptom relief following CBT ± medication. Any reliable variable that could be used to personalize treatment selection could be immensely useful in increasing efficacy and avoiding treatment failure. Here we examine the utility of GRS_{OCD} for predicting treatment response. As such, we will test for an interaction between GRS_{OCD} and symptomatic change (e.g., decrease in Y-BOCS score), with treatment regimen (e.g., CBT only, CBT + SRI, CBT + SRI + antipsychotic) as a covariate. We may find, for example, that individuals with higher GRS respond more poorly to treatment in general, independent of medication. Or perhaps we may find that individuals with higher GRS respond better, or more quickly, to CBT augmented with medication. Regardless, any trends observed in this study can be examined further in collaboration with groups performing OCD treatment trials.

Statistical power. GxE power calculations depend on the effect size and distribution of the main effects as well as the interaction. As an example of this approach, pregnancy complications are associated with a roughly two-fold increased risk of OCD and ~30% of pregnant women experience complications. Given 10,000 OCD cases, 25,000 controls and $\alpha = 5x10^{-4}$ we have 80% power to detect a 9% difference in GRS on the basis of pregnancy complication (Gauderman, 2002).

Alternative GxE approaches. This is clearly an area of active methods development and we will follow the field closely. One limitation of the GRS approach is that it assumes that G is polygenic. Since this may not necessarily be true, we intend to also investigate pathway-based and full-genome GxE scans as well (Thomas, 2010a, 2010b).

DISCUSSION

The validity of our NORDiC project depends on several assumptions. The first is that a study based on the populations of Sweden, Norway and Denmark will generalize to other populations. The epidemiological and genetic data from many disease studies suggest this assumption is

reasonable. The second assumption is that OCD is a complex polygenic disease for which large samples and high-density genomic data are required for discovery of susceptibility genes. However, given that smaller studies have thus far encountered difficulty in identifying risk genes and the observation that polygenicity is the norm for schizophrenia, bipolar disorder, major depressive disorder, and autism, we believe this assumption is reasonable. A final limitation is that, while the Nordic registers contain high-quality information about many environmental factors of interest to OCD, they certainly do not contain information for all possible risk factors.

Overall, we believe successful completion of these aims will answer a number of key questions regarding OCD biology, etiology and treatment. For example, how does OCD relate to other psychiatric disorders? Do early-onset, tic-related and symptom dimensions (partially) possess distinct genetic origins? Is there a relationship between cumulative genetic risk and response to particular forms of treatment? We hope this study will bring us closer to converting this idiopathic disorder into a pathophysiologically defined disease, nominate potential drug targets and demonstrate the utility of the comparative genomics approach for other complex biomedical traits. NORDiC also provides unique infrastructure for similar data collection efforts in OCD-related disorders, which are often treated in the same clinics. We have already begun collecting DNA samples from patients with Tourette's syndrome, body dysmorphic disorder, hoarding disorder, hair pulling disorder and skin picking disorder.

ACKNOWLEDGEMENTS

NORDiC is funded by NIMH R01 MH110427 (PI Crowley), NIMH R01 MH105500 (PI Crowley) and the Swedish Research Council grant # 2015-02271 (PI Mataix-Cols).

FINANCIAL DISCLOSURES

The authors report no biomedical financial interests or potential conflicts of interest.

CONSORTIA

Members of the Nordic OCD & Related Disorders Consortium (NORDiC):

David Mataix-Cols[1,2], Bjarne Hansen[3,4], Manuel Mattheisen[5-8], Elinor K. Karlsson[9-10], Anjené M. Addington[11], Julia Boberg[1,2], Diana R. Djurfeldt[1,2], Matthew Halvorsen[12], Hyun Ji Noh[9], Paul Lichtenstein[13], Stian Solem[3,14], Kerstin Lindblad-Toh[9,15], Anna K. Kähler[13], Elles de Schipper[1], Tetyana Zayats[16], Jan Haavik[16,17], Gerd Kvale[3,4], Christian Rück[1,2], James J. Crowley[1,12,18]

[1] Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
[2] Stockholm Health Care Services, Stockholm, Sweden
[3] Haukeland University Hospital, OCD-team, Bergen, Norway
[4] Department of Clinical Psychology, University of Bergen, Bergen, Norway
[5] Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany
[6] Institute of Human Genetics, University of Bonn, Bonn, Germany
[7] Center for Integrative Sequencing, iSEQ, Department of Biomedicine, Aarhus University, Denmark
[8] Department of Psychiatry, Psychosomatics, and Psychotherapy, University of Würzburg, Germany
[9] Broad Institute of MIT and Harvard, Cambridge, MA, USA
[10] Program in Bioinformatics & Integrative Biology and Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, MA, USA
[11] Genomics Research Branch, National Institute of Mental Health in Bethesda, Bethesda, Maryland, USA
[12] Department of Genetics, University of North Carolina at Chapel Hill, NC, USA
[13] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[14] Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway
[15] Science for Life Laboratory, IMBIM, Uppsala University, Uppsala, Sweden
[16] Department of Biomedicine, University of Bergen, Bergen, Norway
[17] Division of Psychiatry, Haukeland University Hospital, Bergen, Norway
[18] Department of Psychiatry, University of North Carolina at Chapel Hill, NC, USA

REFERENCES

Agerbo, E., Sullivan, P. F., Vilhjalmsson, B. J., Pedersen, C. B., Mors, O., Borglum, A. D., . . . Mortensen, P. B. (2015). Polygenic Risk Score, Parental Socioeconomic Status, Family History of Psychiatric Disorders, and the Risk for Schizophrenia: A Danish Population-Based Study and Meta-analysis. *JAMA psychiatry, 72*(7), 635-641. doi:10.1001/jamapsychiatry.2015.0346

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Arlington, VA: American Psychiatric Publishing.

Andersen, J. S., Olivarius Nde, F., & Krasnik, A. (2011). The Danish National Health Service Register. *Scandinavian journal of public health, 39*(7 Suppl), 34-37. doi:10.1177/1403494810394718

Axelsson, E., Ratnakumar, A., Arendt, M. L., Maqbool, K., Webster, M. T., Perloski, M., . . . Lindblad-Toh, K. (2013). The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature, 495*(7441), 360-364. doi:10.1038/nature11837

Axelsson, O. (2003). The Swedish medical birth register. *Acta Obstet Gynecol Scand, 82*(6), 491-492.

Baadsgaard, M., & Quitzau, J. (2011). Danish registers on personal income and transfer payments. *Scandinavian journal of public health, 39*(7 Suppl), 103-105. doi:10.1177/1403494811405098

Bienvenu, O. J., Wang, Y., Shugart, Y. Y., Welch, J. M., Grados, M. A., Fyer, A. J., . . . Nestadt, G. (2009). Sapap3 and pathological grooming in humans: Results from the OCD collaborative genetics study. *American journal of medical genetics. Part B, Neuropsychiatric genetics, 150B*(5), 710-720. doi:10.1002/ajmg.b.30897

Bloch, M. H., Landeros-Weisenberger, A., Rosario, M. C., Pittenger, C., & Leckman, J. F. (2008). Meta-analysis of the symptom structure of obsessive-compulsive disorder. *The American journal of psychiatry, 165*(12), 1532-1542. doi:10.1176/appi.ajp.2008.08020320

Boardman, L., van der Merwe, L., Lochner, C., Kinnear, C. J., Seedat, S., Stein, D. J., . . . Hemmings, S. M. (2011). Investigating SAPAP3 variants in the etiology of obsessive-compulsive disorder and trichotillomania in the South African white population. *Comprehensive psychiatry, 52*(2), 181-187. doi:10.1016/j.comppsych.2010.05.007

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. doi:10.1017/S0033291706008816

Brander, G., Perez-Vigil, A., Larsson, H., & Mataix-Cols, D. (2016). Systematic review of environmental risk factors for Obsessive-Compulsive Disorder: A proposed roadmap from association to causation. *Neuroscience and biobehavioral reviews, 65*, 36-62. doi:10.1016/j.neubiorev.2016.03.011

Brander, G., Rydell, M., Kuja-Halkola, R., Fernandez de la Cruz, L., Lichtenstein, P., Serlachius, E., . . . Mataix-Cols, D. (2016). Association of Perinatal Risk Factors With Obsessive-Compulsive Disorder: A Population-Based Birth Cohort, Sibling Control Study. *JAMA psychiatry, 73*(11), 1135-1144. doi:10.1001/jamapsychiatry.2016.2095

Browne, H. A., Hansen, S. N., Buxbaum, J. D., Gair, S. L., Nissen, J. B., Nikolajsen, K. H., . . . Grice, D. E. (2015). Familial clustering of tic disorders and obsessive-compulsive disorder. *JAMA psychiatry, 72*(4), 359-366. doi:10.1001/jamapsychiatry.2014.2656

Bulik-Sullivan, B. K., Loh, P. R., Finucane, H. K., Ripke, S., Yang, J., Schizophrenia Working Group of the Psychiatric Genomics, C., . . . Neale, B. M. (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics, 47*(3), 291-295. doi:10.1038/ng.3211

Cappi, C., Brentani, H., Lima, L., Sanders, S. J., Zai, G., Diniz, B. J., . . . Fernandez, T. V. (2016). Whole-exome sequencing in obsessive-compulsive disorder identifies rare mutations in immunological and neurodevelopmental pathways. *Transl Psychiatry, 6*, e764. doi:10.1038/tp.2016.30

Collaborative, I. O. C. D. F. G. (2018). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Molecular psychiatry, 23*(5), 1181-1188. doi:10.1038/mp.2017.154

Cross-Disorder Group of the Psychiatric Genomics Consortium. (2013). Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nature genetics, 45*(9), 984-994. doi:10.1038/ng.2711

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), e1003864. doi:10.1371/journal.pgen.1003864

de Bakker, P. I., Ferreira, M. A., Jia, X., Neale, B. M., Raychaudhuri, S., & Voight, B. F. (2008). Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Human molecular genetics, 17*(R2), R122-128. doi:10.1093/hmg/ddn288

Devlin, B., & Roeder, K. (1999). Genomic control for association studies. *Biometrics, 55*(4), 997-1004.

Dodman, N. H., Karlsson, E. K., Moon-Fanelli, A., Galdzicka, M., Perloski, M., Shuster, L., . . . Ginns, E. I. (2010). A canine chromosome 7 locus confers compulsive disorder susceptibility. *Molecular psychiatry, 15*(1), 8-10. doi:10.1038/mp.2009.111

Dudbridge, F. (2013). Power and predictive accuracy of polygenic risk scores. *PLoS genetics, 9*(3), e1003348. doi:10.1371/journal.pgen.1003348

Dudbridge, F., & Gusnanto, A. (2008). Estimation of significance thresholds for genomewide association scans. *Genetic epidemiology, 32*(3), 227-234. doi:10.1002/gepi.20297

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychol Assess, 14*(4), 485-496.

Fullana, M. A., Mataix-Cols, D., Caspi, A., Harrington, H., Grisham, J. R., Moffitt, T. E., & Poulton, R. (2009). Obsessions and compulsions in the community: prevalence, interference, help-seeking, developmental stability, and co-occurring psychiatric conditions. *The American journal of psychiatry, 166*(3), 329-336. doi:10.1176/appi.ajp.2008.08071006

Gauderman, W. J. (2002). Sample size requirements for matched case-control studies of gene-environment interaction. *Statistics in medicine, 21*(1), 35-50.

Gazzellone, M. J., Zarrei, M., Burton, C. L., Walker, S., Uddin, M., Shaheen, S. M., . . . Scherer, S. W. (2016). Uncovering obsessive-compulsive disorder risk genes in a pediatric cohort by high-resolution analysis of copy number variation. *J Neurodev Disord, 8*, 36. doi:10.1186/s11689-016-9170-9

Grunblatt, E., Oneda, B., Ekici, A. B., Ball, J., Geissler, J., Uebe, S., . . . Walitza, S. (2017). High resolution chromosomal microarray analysis in paediatric obsessive-compulsive disorder. *BMC Med Genomics, 10*(1), 68. doi:10.1186/s12920-017-0299-5

Guha, S., Rees, E., Darvasi, A., Ivanov, D., Ikeda, M., Bergen, S. E., . . . Kirov, G. (2013). Implication of a rare deletion at distal 16p11.2 in schizophrenia. *JAMA psychiatry, 70*(3), 253-260. doi:10.1001/2013.jamapsychiatry.71

Hanna, G. L., Veenstra-VanderWeele, J., Cox, N. J., Boehnke, M., Himle, J. A., Curtis, G. C., . . . Cook, E. H., Jr. (2002). Genome-wide linkage analysis of families with obsessive-compulsive disorder ascertained through pediatric probands. *American journal of medical genetics, 114*(5), 541-552. doi:10.1002/ajmg.10519

Hanna, G. L., Veenstra-Vanderweele, J., Cox, N. J., Van Etten, M., Fischer, D. J., Himle, J. A., . . . Cook, E. H., Jr. (2007). Evidence for a susceptibility locus on chromosome 10p15 in early-onset obsessive-compulsive disorder. *Biological psychiatry, 62*(8), 856-862. doi:10.1016/j.biopsych.2007.01.008

Hansen, B., Hagen, K., Ost, L. G., Solem, S., & Kvale, G. (2018). The Bergen 4-Day OCD Treatment Delivered in a Group Setting: 12-Month Follow-Up. *Front Psychol, 9*, 639. doi:10.3389/fpsyg.2018.00639

Hansen, B., Kvale, G., Hagen, K., Havnen, A., & Ost, L. G. (2018). The Bergen 4-day treatment for OCD: four years follow-up of concentrated ERP in a clinical mental health setting. *Cogn Behav Ther*, 1-17. doi:10.1080/16506073.2018.1478447

Hasler, G., Pinto, A., Greenberg, B. D., Samuels, J., Fyer, A. J., Pauls, D., . . . Study, O. C. D. C. G. (2007). Familiality of factor analysis-derived YBOCS dimensions in OCD-affected sibling pairs from the OCD Collaborative Genetics Study. *Biological psychiatry, 61*(5), 617-625. doi:10.1016/j.biopsych.2006.05.040

Helweg-Larsen, K. (2011). The Danish Register of Causes of Death. *Scandinavian journal of public health, 39*(7 Suppl), 26-29. doi:10.1177/1403494811399958

Iervolino, A. C., Rijsdijk, F. V., Cherkas, L., Fullana, M. A., & Mataix-Cols, D. (2011a). A multivariate twin study of obsessive-compulsive symptom dimensions. *Archives of general psychiatry, 68*(6), 637-644. doi:10.1001/archgenpsychiatry.2011.54

Iervolino, A. C., Rijsdijk, F. V., Cherkas, L., Fullana, M. A., & Mataix-Cols, D. (2011b). A multivariate twin study of obsessive-compulsive symptom dimensions. *Archives of general psychiatry, 68*(6), 637-644. doi:10.1001/archgenpsychiatry.2011.54

Insel, T. R., Hoover, C., & Murphy, D. L. (1983). Parents of patients with obsessive-compulsive disorder. *Psychological medicine, 13*(4), 807-811.

Johansson, L. A., & Westerling, R. (2000). Comparing Swedish hospital discharge records with death certificates: implications for mortality statistics. *Int J Epidemiol, 29*(3), 495-502.

Juel, K., & Helweg-Larsen, K. (1999). The Danish registers of causes of death. *Danish medical bulletin, 46*(4), 354-357.

Karlsson, E. K., Sigurdsson, S., Ivansson, E., Thomas, R., Elvers, I., Wright, J., . . . Lindblad-Toh, K. (2013). Genome-wide analyses implicate 33 loci in heritable dog osteosarcoma, including regulatory variants near CDKN2A/B. *Genome biology, 14*(12), R132. doi:10.1186/gb-2013-14-12-r132

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), 1094-1099.

Kildemoes, H. W., Sorensen, H. T., & Hallas, J. (2011). The Danish National Prescription Registry. *Scandinavian journal of public health, 39*(7 Suppl), 38-41. doi:10.1177/1403494810394717

Knudsen, L. B., & Olsen, J. (1998). The Danish Medical Birth Registry. *Danish medical bulletin, 45*(3), 320-323.

Kvale, G., & Hansen, B. (2014). Dissemination and intensifying evidence-based treatment for OCD: Norway is in the lead. *The Nordic Psychiatrist*, 14-15.

Kvale, G., Hansen, B., Bjorgvinsson, T., Bortveit, T., Hagen, K., Haseth, S., . . . Ost, L. G. (2018). Successfully treating 90 patients with obsessive compulsive disorder in eight days: the Bergen 4-day treatment. *BMC Psychiatry, 18*(1), 323. doi:10.1186/s12888-018-1887-4

Langhoff-Roos, J., Krebs, L., Klungsoyr, K., Bjarnadottir, R. I., Kallen, K., Tapper, A. M., . . . Gissler, M. (2014). The Nordic medical birth registers--a potential goldmine for clinical research. *Acta Obstet Gynecol Scand, 93*(2), 132-137. doi:10.1111/aogs.12302

Leckman, J. F., Denys, D., Simpson, H. B., Mataix-Cols, D., Hollander, E., Saxena, S., . . . Stein, D. J. (2010). Obsessive-compulsive disorder: a review of the diagnostic criteria and possible subtypes and dimensional specifiers for DSM-V. *Depression and anxiety, 27*(6), 507-527. doi:10.1002/da.20669

Levinson, D. F., Duan, J., Oh, S., Wang, K., Sanders, A. R., Shi, J., . . . Gejman, P. V. (2011). Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications. *The American journal of psychiatry, 168*(3), 302-316. doi:10.1176/appi.ajp.2010.10060876

Liang, K. Y., Wang, Y., Shugart, Y. Y., Grados, M., Fyer, A. J., Rauch, S., . . . Nestadt, G. (2008). Evidence for potential relationship between SLC1A1 and a putative genetic linkage region on chromosome 14q to obsessive-compulsive disorder with compulsive hoarding. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics, 147B*(6), 1000-1002. doi:10.1002/ajmg.b.30713

Liao, H. M., Chao, Y. L., Huang, A. L., Cheng, M. C., Chen, Y. J., Lee, K. F., . . . Chen, C. H. (2012). Identification and characterization of three inherited genomic copy number variations associated with familial schizophrenia. *Schizophrenia research, 139*(1-3), 229-236. doi:10.1016/j.schres.2012.05.015

Lu, Y., Pouget, J. G., Andreassen, O. A., Djurovic, S., Esko, T., Hultman, C. M., . . . Sullivan, P. F. (2018). Genetic risk scores and family history as predictors of schizophrenia in Nordic registers. *Psychological medicine, 48*(7), 1201-1208. doi:10.1017/S0033291717002665

Ludvigsson, J. F., Andersson, E., Ekbom, A., Feychting, M., Kim, J. L., Reuterwall, C., . . . Olausson, P. O. (2011). External review and validation of the Swedish national inpatient register. *BMC Public Health, 11*, 450. doi:10.1186/1471-2458-11-450

Luescher, A. U. (2004). Diagnosis and management of compulsive disorders in dogs and cats. *Clin Tech Small Anim Pract, 19*(4), 233-239. doi:10.1053/j.ctsap.2004.10.005

Lynge, E., Sandegaard, J. L., & Rebolj, M. (2011). The Danish National Patient Register. *Scandinavian journal of public health, 39*(7 Suppl), 30-33. doi:10.1177/1403494811401482

Malhotra, D., & Sebat, J. (2012). CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell, 148*(6), 1223-1241. doi:10.1016/j.cell.2012.02.039

Mataix-Cols, D., Boman, M., Monzani, B., Ruck, C., Serlachius, E., Langstrom, N., & Lichtenstein, P. (2013). Population-Based, Multigenerational Family Clustering Study of Obsessive-Compulsive Disorder. *JAMA psychiatry*, 1-9. doi:10.1001/jamapsychiatry.2013.3

Mataix-Cols, D., Frans, E., Perez-Vigil, A., Kuja-Halkola, R., Gromark, C., Isomura, K., . . . Larsson, H. (2018). A total-population multigenerational family clustering study of autoimmune diseases in obsessive-compulsive disorder and Tourette's/chronic tic disorders. *Molecular psychiatry, 23*(7), 1652-1658. doi:10.1038/mp.2017.215

Mataix-Cols, D., Rosario-Campos, M. C., & Leckman, J. F. (2005). A multidimensional model of obsessive-compulsive disorder. *The American journal of psychiatry, 162*(2), 228-238. doi:10.1176/appi.ajp.162.2.228

Mataix-Cols, D., Wooderson, S., Lawrence, N., Brammer, M. J., Speckens, A., & Phillips, M. L. (2004). Distinct neural correlates of washing, checking, and hoarding symptom dimensions in obsessive-compulsive disorder. *Archives of general psychiatry, 61*(6), 564-576. doi:10.1001/archpsyc.61.6.564

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), 629-636. doi:10.1016/j.biopsych.2012.03.037

Mattheisen, M., Samuels, J. F., Wang, Y., Greenberg, B. D., Fyer, A. J., McCracken, J. T., . . . Nestadt, G. (2014). Genome-wide association study in obsessive-compulsive disorder: results from the OCGAS. *Molecular psychiatry*. doi:10.1038/mp.2014.43

McGrath, J. J., Mortensen, P. B., Visscher, P. M., & Wray, N. R. (2013). Where GWAS and epidemiology meet: opportunities for the simultaneous study of genetic and environmental risk factors in schizophrenia. *Schizophrenia bulletin, 39*(5), 955-959. doi:10.1093/schbul/sbt108

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), 910-919. doi:10.1016/j.jaac.2014.04.022

Medland, S. E., & Neale, M. C. (2010). An integrated phenomic approach to multivariate allelic association. *European journal of human genetics : EJHG, 18*(2), 233-239. doi:10.1038/ejhg.2009.133

Minica, C. C., Boomsma, D. I., van der Sluis, S., & Dolan, C. V. (2010). Genetic association in multivariate phenotypic data: power in five models. *Twin Res Hum Genet, 13*(6), 525-543. doi:10.1375/twin.13.6.525

Moon-Fanelli, A. A., Dodman, N. H., & Cottam, N. (2007). Blanket and flank sucking in Doberman Pinschers. *J Am Vet Med Assoc, 231*(6), 907-912. doi:10.2460/javma.231.6.907

Murphy, T. K., Storch, E. A., Lewin, A. B., Edge, P. J., & Goodman, W. K. (2012). Clinical factors associated with pediatric autoimmune neuropsychiatric disorders associated with streptococcal infections. *The Journal of pediatrics, 160*(2), 314-319. doi:10.1016/j.jpeds.2011.07.012

Nestadt, G., Samuels, J., Riddle, M., Bienvenu, O. J., 3rd, Liang, K. Y., LaBuda, M., . . . Hoehn-Saric, R. (2000). A family study of obsessive-compulsive disorder. *Archives of general psychiatry, 57*(4), 358-363.

Nissen, J., Powell, S., Koch, S. V., Crowley, J. J., Matthiesen, M., Grice, D. E., . . . Parner, E. (2017). Diagnostic validity of early-onset obsessive-compulsive disorder in the Danish Psychiatric Central Register: findings from a cohort sample. *BMJ Open, 7*(9), e017172. doi:10.1136/bmjopen-2017-017172

Noh, H. J., Tang, R., Flannick, J., O'Dushlaine, C., Swofford, R., Howrigan, D., . . . Lindblad-Toh, K. (2017). Integrating evolutionary and regulatory information with a multispecies approach implicates genes and pathways in obsessive-compulsive disorder. *Nat Commun, 8*(1), 774. doi:10.1038/s41467-017-00831-x

Norgaard-Pedersen, B., & Hougaard, D. M. (2007). Storage policies and use of the Danish Newborn Screening Biobank. *Journal of inherited metabolic disease, 30*(4), 530-536. doi:10.1007/s10545-007-0631-x

Ogata, N., Gillis, T. E., Liu, X., Cunningham, S. M., Lowen, S. B., Adams, B. L., . . . Kaufman, M. J. (2013). Brain structural abnormalities in Doberman pinschers with canine compulsive disorder. *Prog Neuropsychopharmacol Biol Psychiatry, 45*, 1-6. doi:10.1016/j.pnpbp.2013.04.002

Ost, L. G., Havnen, A., Hansen, B., & Kvale, G. (2015). Cognitive behavioral treatments of obsessive-compulsive disorder. A systematic review and meta-analysis of studies published 1993-2014. *Clin Psychol Rev, 40*, 156-169. doi:10.1016/j.cpr.2015.06.003

Ost, L. G., Riise, E. N., Wergeland, G. J., Hansen, B., & Kvale, G. (2016). Cognitive behavioral and pharmacological treatments of OCD in children: A systematic review and meta-analysis. *J Anxiety Disord, 43*, 58-69. doi:10.1016/j.janxdis.2016.08.003

Overall, K. L. (2000). Natural animal models of human psychiatric conditions: assessment of mechanism and validity. *Prog Neuropsychopharmacol Biol Psychiatry, 24*(5), 727-776.

Overall, K. L., & Dunham, A. E. (2002). Clinical features and outcome in dogs and cats with obsessive-compulsive disorder: 126 cases (1989-2000). *J Am Vet Med Assoc, 221*(10), 1445-1452.

Pauls, D. L. (2010). The genetics of obsessive-compulsive disorder: a review. *Dialogues in clinical neuroscience, 12*(2), 149-163.

Pauls, D. L., Abramovitch, A., Rauch, S. L., & Geller, D. A. (2014). Obsessive-compulsive disorder: an integrative genetic and neurobiological perspective. *Nat Rev Neurosci, 15*(6), 410-424. doi:10.1038/nrn3746

Pe'er, I., Yelensky, R., Altshuler, D., & Daly, M. J. (2008). Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. *Genetic epidemiology, 32*(4), 381-385. doi:10.1002/gepi.20303

Petersson, F., Baadsgaard, M., & Thygesen, L. C. (2011). Danish registers on personal labour market affiliation. *Scandinavian journal of public health, 39*(7 Suppl), 95-98. doi:10.1177/1403494811408483

Rasmussen, S. A., & Tsuang, M. T. (1986). Clinical characteristics and family history in DSM-III obsessive-compulsive disorder. *The American journal of psychiatry, 143*(3), 317-322.

Rosenberg, C. M. (1967). Familial aspects of obsessional neurosis. *The British journal of psychiatry : the journal of mental science, 113*(497), 405-413.

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. doi:10.1007/s00439-011-1033-6

Ruscio, A. M., Stein, D. J., Chiu, W. T., & Kessler, R. C. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular psychiatry, 15*(1), 53-63. doi:10.1038/mp.2008.94

Samuels, J., Shugart, Y. Y., Grados, M. A., Willour, V. L., Bienvenu, O. J., Greenberg, B. D., . . . Nestadt, G. (2007). Significant linkage to compulsive hoarding on chromosome 14 in families with obsessive-compulsive disorder: results from the OCD Collaborative Genetics Study. *The American journal of psychiatry, 164*(3), 493-499. doi:10.1176/appi.ajp.164.3.493

Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. (2014). *Nature, 511*(7510), 421-427.

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), 763-770. doi:10.1038/sj.mp.4001847

Skoog, G., & Skoog, I. (1999). A 40-year follow-up of patients with obsessive-compulsive disorder [see commetns]. *Archives of general psychiatry, 56*(2), 121-127.

Steinhausen, H. C., Bisgaard, C., Munk-Jorgensen, P., & Helenius, D. (2013). Family aggregation and risk factors of obsessive-compulsive disorders in a nationwide three-generation study. *Depression and anxiety, 30*(12), 1177-1184. doi:10.1002/da.22163

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), 788-798. doi:10.1038/mp.2012.85

Sullivan, P. F., Daly, M. J., & O'Donovan, M. (2012). Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature reviews. Genetics, 13*(8), 537-551. doi:10.1038/nrg3240

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), 264-271.

Tang, R., Noh, H. J., Wang, D., Sigurdsson, S., Swofford, R., Perloski, M., . . . Karlsson, E. K. (2014). Candidate genes and functional noncoding variants identified in a canine model of obsessive-compulsive disorder. *Genome biology, 15*(3), R25. doi:10.1186/gb-2014-15-3-r25

Taylor, S. (2013). Molecular genetics of obsessive-compulsive disorder: a comprehensive meta-analysis of genetic association studies. *Molecular psychiatry, 18*(7), 799-805. doi:10.1038/mp.2012.76

Tengvall, K., Kierczak, M., Bergvall, K., Olsson, M., Frankowiack, M., Farias, F. H., . . . Lindblad-Toh, K. (2013). Genome-wide analysis in German shepherd dogs reveals association of a locus on CFA 27 with atopic dermatitis. *PLoS genetics, 9*(5), e1003475. doi:10.1371/journal.pgen.1003475

Thomas, D. (2010a). Gene--environment-wide association studies: emerging approaches. *Nature reviews. Genetics, 11*(4), 259-272. doi:10.1038/nrg2764

Thomas, D. (2010b). Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. *Annu Rev Public Health, 31*, 21-36. doi:10.1146/annurev.publhealth.012809.103619

van der Sluis, S., Posthuma, D., Nivard, M. G., Verhage, M., & Dolan, C. V. (2013). Power in GWAS: lifting the curse of the clinical cut-off. *Molecular psychiatry, 18*(1), 2-3. doi:10.1038/mp.2012.65

van der Sluis, S., Verhage, M., Posthuma, D., & Dolan, C. V. (2010). Phenotypic complexity, measurement bias, and poor phenotypic resolution contribute to the missing heritability problem in genetic association studies. *PLoS One, 5*(11), e13929. doi:10.1371/journal.pone.0013929

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. doi:10.1375/183242705774310060

Welch, J. M., Lu, J., Rodriguiz, R. M., Trotta, N. C., Peca, J., Ding, J. D., . . . Feng, G. (2007). Cortico-striatal synaptic defects and OCD-like behaviours in Sapap3-mutant mice. *Nature, 448*(7156), 894-900. doi:10.1038/nature06104

Wettermark, B., Hammar, N., Fored, C. M., Leimanis, A., Otterblad Olausson, P., Bergman, U., . . . Rosen, M. (2007). The new Swedish Prescribed Drug Register--opportunities for pharmacoepidemiological research and experience from the first six months. *Pharmacoepidemiol Drug Saf, 16*(7), 726-735. doi:10.1002/pds.1294

Willour, V. L., Yao Shugart, Y., Samuels, J., Grados, M., Cullen, B., Bienvenu, O. J., 3rd, . . . Nestadt, G. (2004). Replication study supports evidence for linkage to 9p24 in obsessive-compulsive disorder. *American journal of human genetics, 75*(3), 508-513. doi:10.1086/423899

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

Wu, Y., Liu, X., Luo, H., Deng, W., Zhao, G., Wang, Q., . . . Li, T. (2012). Advanced paternal age increases the risk of schizophrenia and obsessive-compulsive disorder in a Chinese Han population. *Psychiatry research, 198*(3), 353-359. doi:10.1016/j.psychres.2012.01.020

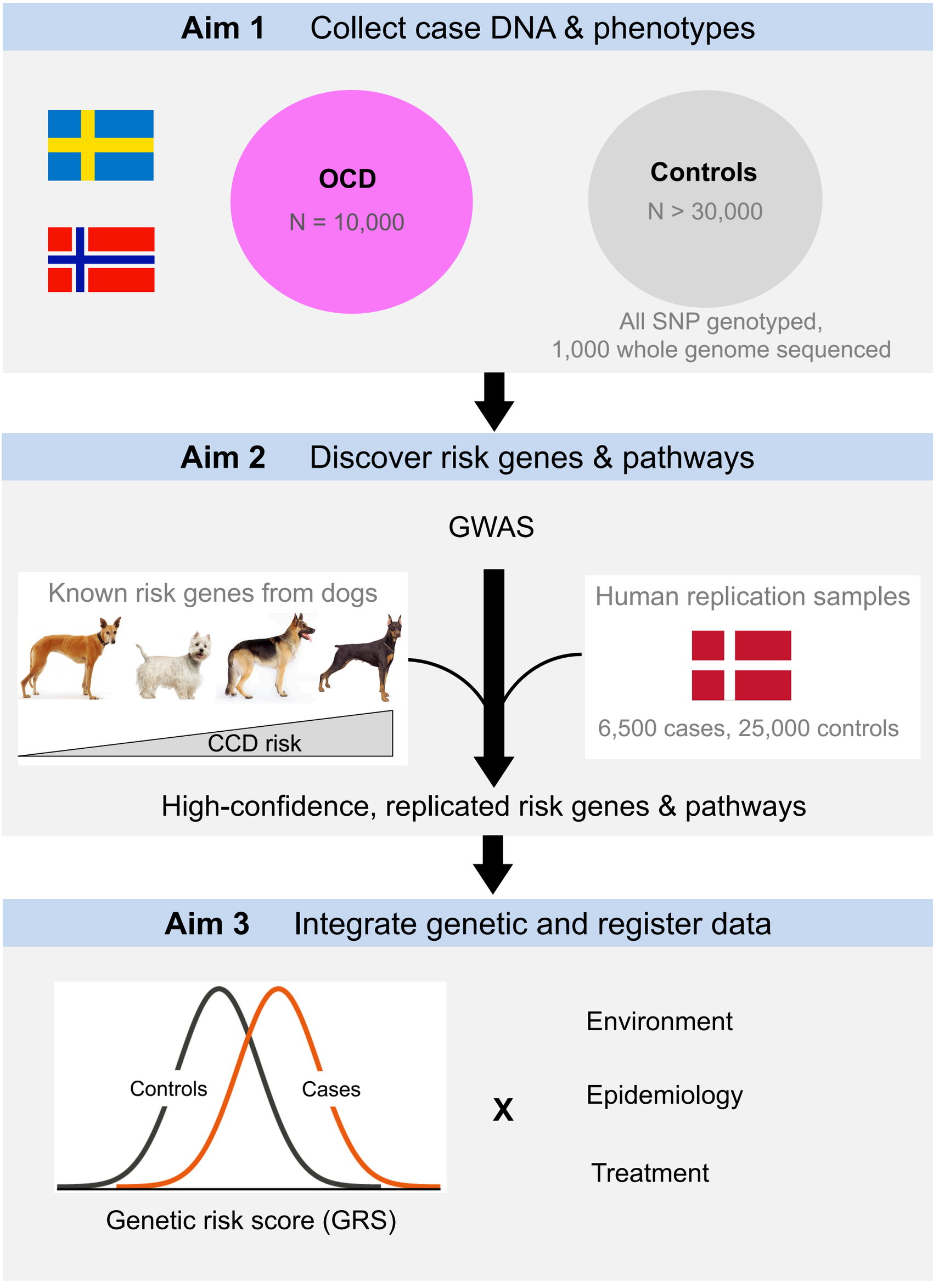
Zuchner, S., Wendland, J. R., Ashley-Koch, A. E., Collins, A. L., Tran-Viet, K. N., Quinn, K., . . . Murphy, D. L. (2009). Multiple rare SAPAP3 missense variants in trichotillomania and OCD. *Molecular psychiatry, 14*(1), 6-9. doi:10.1038/mp.2008.83

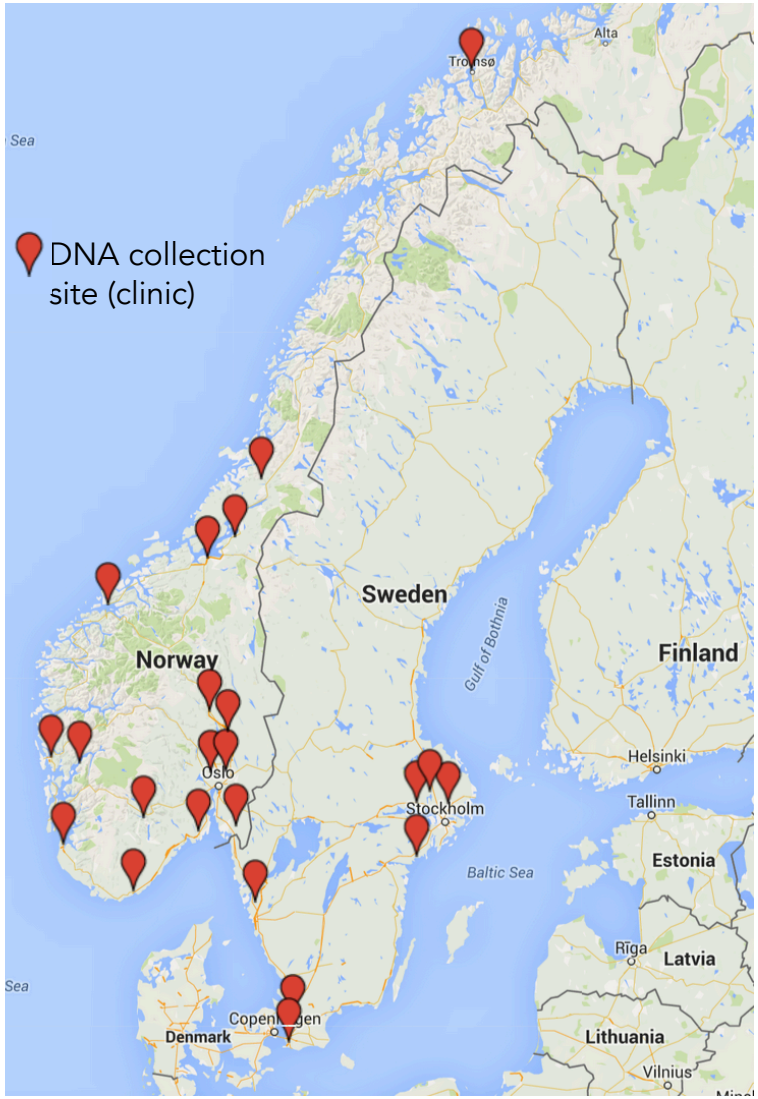
FIGURE LEGENDS

Figure 1. NORDiC study overview.

Figure 2. NORDiC sample collection sites.

Figure 3. Statistical power for OCD GWAS across the allelic spectrum. Curves indicate minimal detectable relative risk with 80% power. For reference, schizophrenia genome-wide significant SNPs are shown in green.("Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci," 2014)


Aim 1 Collect case DNA & phenotypes
OCD
N = 10,000
Controls
N > 30,000
All SNP genotyped,
1,000 whole genome sequenced
Aim 2 Discover risk genes & pathways
GWAS
Known risk genes from dogs
CCD risk
Human replication samples
6,500 cases, 25,000 controls
High-confidence, replicated risk genes & pathways
Aim 3 Integrate genetic and register data
Controls
Cases
Genetic risk score (GRS)
x
Environment
Epidemiology
Treatment



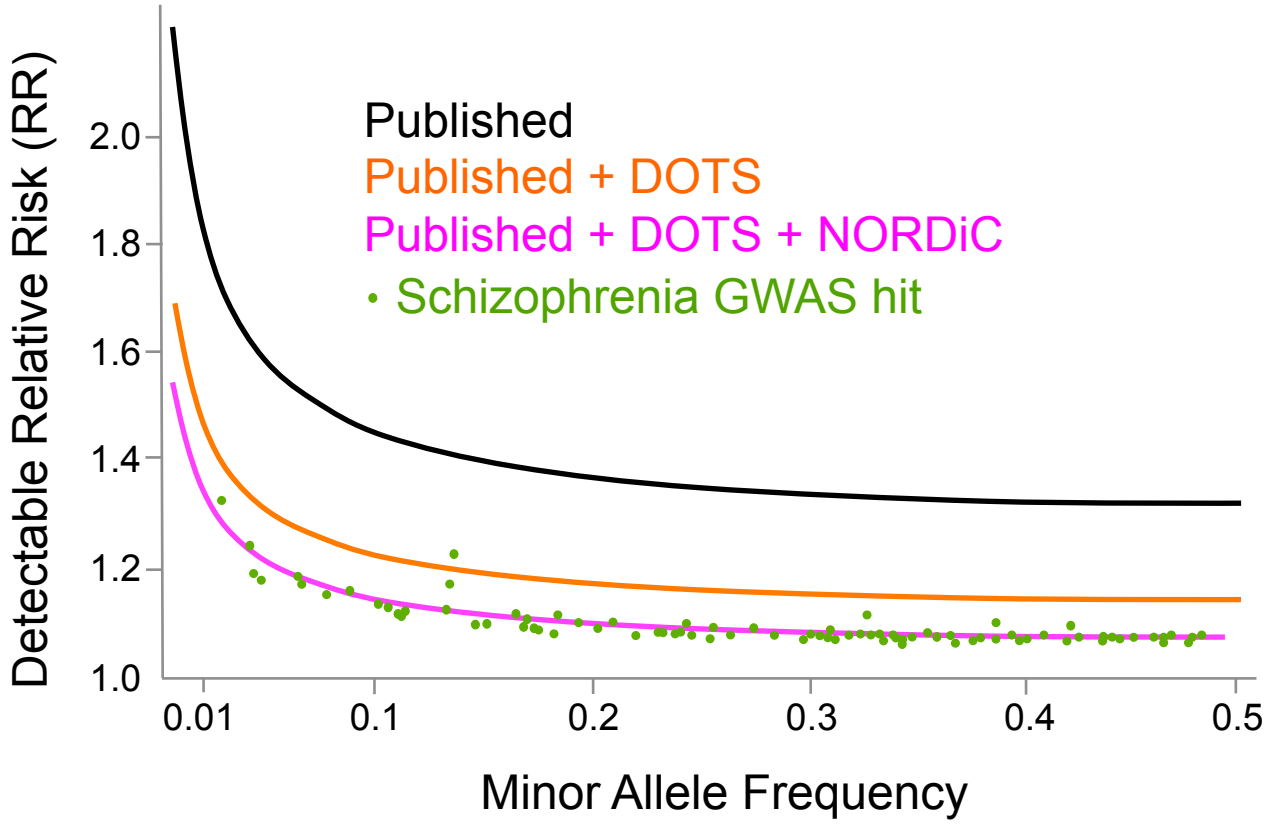

Detectable Relative Risk (RR)
2.0
1.8
1.6
1.4
1.2
1.0
0.01
0.1
0.2
0.3
0.4
0.5
Minor Allele Frequency
Published
Published + DOTS
Published + DOTS + NORDiC
Schizophrenia GWAS hit

Table 1. Harmonized OCD diagnostic, treatment and outcome measures across Sweden and Norway

Age	Diagnostic	Clinician-administered	Self-report	Family-report
Pediatric	MINI-KID, OCD-RD modules	CY-BOCS, C-GAS, MADRS, CGI-Severity, Improvement	OCI-CV, PHQ-9	FAS
Adult	MINI, OCD-RD modules	Y-BOCS, GAF, MADRS, CGI-Severity, Improvement	OCI-R, PHQ-9	FAS

C-GAS: Children's Global Assessment Scale
CGI: Clinical Global Impression
CY-BOCS: Children's Yale-Brown Obsessive-Compulsive
FAS: Family Accommodation Scale
GAF: Global Assessment of Functioning
MADRS: Montgomery–Åsberg Depression Rating Scale
PHQ-9: Patient Health Questionnaire-9
MINI: Mini-International Neuropsychiatric Interview
MINI-KID: Mini-Intl. Neuropsychiatric Interview -
OCD-RD: diagnosis of BDD, hoarding, skin-pick, hair-
OCI-CV: Obsessive-Compulsive Inventory – Child
OCI-R: Obsessive-Compulsive Inventory – Revised
Y-BOCS: Yale-Brown Obsessive-Compulsive Scale

Table 2. Heterogeneity measures		
Possible subtype	Clinical information	Register information
Tic-related OCD	Structured diagnostic interviews, including questions about lifetime occurrence of tics; medical records	National Patient Register: Lifetime diagnoses of tic disorders
Early onset OCD	Approximately half of our cohort will consist of pediatric OCD cases	
OCD symptom dimensions	Clinician administered: YBOCS symptom checklist at baseline Self-report: OCI-R (adult samples) and OCI-CV (pediatric samples)	
Familial vs sporadic OCD	Self-reported family history of OCD and related disorders; medical records	Multigenerational and National Patient Register: Lifetime diagnoses of OCD in 1st, 2nd and 3rd degree relatives
Family history of severe mental illness (e.g. psychosis, autism)	Self-reported family history; medical records	Multigenerational and National Patient Register: Lifetime diagnoses in first, second and third-degree relatives
Treatment refractoriness	Incomplete or no treatment response with SRIs and/or CBT, measured with the YBOCS at post-treatment, medical records	

Table 3. Examples of Scandinavian registers		Reference		
Type	Information	Sweden	Norway	Denmark
Medical Birth Register (MBR)	Data on place of birth, birth weight, mother's parity, mother's age, type of delivery, Apgar scores, obstetrical complications (e.g., eclampsia).	(O. Axelsson, 2003)	(Langhoff-Roos et al., 2014)	(Knudsen & Olsen, 1998)
National Patient Register (NPR)	Information on treatment at all hospitals, including both in- and outpatient psychiatric contacts (e.g., diagnostic and treatment codes).	(Ludvigsson et al., 2011)	www.npr.no	(Lynge, Sandegaard, & Rebolj, 2011), (Andersen, Olivarius Nde, & Krasnik, 2011)
Comprehensive statistics (CS)	Exhaustive information on sociodemographic characteristics (e.g., living situation, marital status, education levels, employment, income).	www.scb.se	www.ssb.no	(Baadsgaard & Quitzau, 2011; Petersson, Baadsgaard, & Thygesen, 2011)
Prescription drug register (PDR)	Individual-level data for all prescriptions dispensed for in- and out-patients (e.g. variables at level of drug user, prescriber, pharmacy).	(Wettermark et al., 2007)	www.norpd.no	(Kildemoes, Sorensen, & Hallas, 2011)
Cause of Death Register (CDR)	Information on mortality from death certificates (e.g., cause of death according to ICD codes).	(Johansson & Westerling, 2000)	www.fhi.no	(Helweg-Larsen, 2011; Juel & Helweg-Larsen, 1999)