



Polygenic Heterogeneity Across Obsessive-Compulsive Disorder Subgroups Defined by a Comorbid Diagnosis

Nora I. Strom^{1,2,3,4*}, Jakob Grove^{3,5,6}, Sandra M. Meier⁷, Marie Bækvad-Hansen^{5,8}, Judith Becker Nissen⁹, Thomas Damm Als^{3,5,6}, Matthew Halvorsen¹⁰, Merete Nordentoft^{5,11,12}, Preben B. Mortensen^{5,6,13,14}, David M. Hougaard^{5,8}, Thomas Werge^{5,11,15,16}, Ole Mors^{5,17}, Anders D. Børglum^{3,5,6}, James J. Crowley¹⁰, Jonas Bybjerg-Grauholm^{5,8} and Manuel Mattheisen^{2,3,7}

OPEN ACCESS

Edited by:

Lu Liu,
Peking University Sixth Hospital, China

Reviewed by:

Fotis Tsetsos,
Democritus University of Thrace,
Greece
Mohamed Abdulkadir,
Université de Genève, Switzerland

*Correspondence:

Nora I. Strom
stromnor@hu-berlin.de

Specialty section:

This article was submitted to
Behavioral and Psychiatric Genetics,
a section of the journal
Frontiers in Genetics

Received: 18 May 2021

Accepted: 27 July 2021

Published: 31 August 2021

Citation:

Strom NI, Grove J, Meier SM, Bækvad-Hansen M, Becker Nissen J, Damm Als T, Halvorsen M, Nordentoft M, Mortensen PB, Hougaard DM, Werge T, Mors O, Børglum AD, Crowley JJ, Bybjerg-Grauholm J and Mattheisen M (2021) Polygenic Heterogeneity Across Obsessive-Compulsive Disorder Subgroups Defined by a Comorbid Diagnosis. *Front. Genet.* 12:711624. doi: 10.3389/fgene.2021.711624

¹ Department of Psychology, Humboldt Universität zu Berlin, Berlin, Germany, ² Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden, ³ Department of Biomedicine and the iSEQ Center, Aarhus University, Aarhus, Denmark, ⁴ Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Munich, Germany, ⁵ The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Copenhagen, Denmark, ⁶ Center for Genomics and Personalized Medicine, Aarhus, Denmark, ⁷ Department of Psychiatry, Dalhousie University, Halifax, NS, Canada, ⁸ Center for Neonatal Screening, Department for Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark, ⁹ Center for Child and Adolescent Psychiatry, Aarhus University Hospital Risskov, Risskov, Denmark, ¹⁰ Department of Genetics, University of North Carolina, Chapel Hill, NC, United States, ¹¹ Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark, ¹² Copenhagen Research Centre for Mental Health (CORE), Mental Health Centre Copenhagen, Copenhagen University Hospital, Copenhagen, Denmark, ¹³ National Centre for Register-Based Research, Aarhus University, Aarhus, Denmark, ¹⁴ Centre for Integrated Register-based Research, Aarhus University, Aarhus, Denmark, ¹⁵ Institute of Biological Psychiatry, Mental Health Services, Copenhagen University Hospital, Copenhagen, Denmark, ¹⁶ Lundbeck Foundation Center for GeoGenetics, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark, ¹⁷ Psychosis Research Unit, Aarhus University Hospital, Aarhus, Denmark

Among patients with obsessive-compulsive disorder (OCD), 65–85% manifest another psychiatric disorder concomitantly or at some other time point during their life. OCD is highly heritable, as are many of its comorbidities. A possible genetic heterogeneity of OCD in relation to its comorbid conditions, however, has not yet been exhaustively explored. We used a framework of different approaches to study the genetic relationship of OCD with three commonly observed comorbidities, namely major depressive disorder (MDD), attention-deficit hyperactivity disorder (ADHD), and autism spectrum disorder (ASD). First, using publicly available summary statistics from large-scale genome-wide association studies, we compared genetic correlation patterns for OCD, MDD, ADHD, and ASD with 861 somatic and mental health phenotypes. Secondly, we examined how polygenic risk scores (PRS) of eight traits that showed heterogeneous correlation patterns with OCD, MDD, ADHD, and ASD partitioned across comorbid subgroups in OCD using independent unpublished data from the Lundbeck Foundation Initiative for Integrative Psychiatric Research (iPSYCH). The comorbid subgroups comprised of patients with only OCD ($N = 366$), OCD and MDD ($N = 1,052$), OCD and ADHD ($N = 443$), OCD and ASD ($N = 388$), and OCD with more than 1 comorbidity ($N = 429$). We found that PRS of all traits but BMI were significantly associated with OCD across all subgroups (neuroticism: $p = 1.19 \times 10^{-32}$, bipolar disorder: $p = 7.51 \times 10^{-8}$, anorexia nervosa:

$p = 3.52 \times 10^{-20}$, age at first birth: $p = 9.38 \times 10^{-5}$, educational attainment: $p = 1.56 \times 10^{-4}$, OCD: $p = 1.87 \times 10^{-6}$, insomnia: $p = 2.61 \times 10^{-5}$, BMI: $p = 0.15$). For age at first birth, educational attainment, and insomnia PRS estimates significantly differed across comorbid subgroups ($p = 2.29 \times 10^{-4}$, $p = 1.63 \times 10^{-4}$, and $p = 0.045$, respectively). Especially for anorexia nervosa, age at first birth, educational attainment, insomnia, and neuroticism the correlation patterns that emerged from genetic correlation analysis of OCD, MDD, ADHD, and ASD were mirrored in the PRS associations with the respective comorbid OCD groups. Dissecting the polygenic architecture, we found both quantitative and qualitative polygenic heterogeneity across OCD comorbid subgroups.

Keywords: obsessive-compulsive disorder, major depression, attention-deficit/hyperactivity disorder, autism, comorbidity, polygenic risk score, heterogeneity, genetic correlation

1. INTRODUCTION

Obsessive-Compulsive-Disorder (OCD) is a common, long-lasting and disabling neuropsychiatric disorder with an estimated lifetime prevalence of 1–3% (Weissman, 1998; U.S. International institutes of health (NIH), 2016). It is the fourth most common psychiatric disorder and has been ranked by the World Health Organization as being among the most disabling medical conditions world-wide as it can substantially impair the patient's social, occupational and academic functioning (Murray et al., 1996). OCD is considered a complex disorder with its risk likely being influenced by hundreds to thousands of genetic variants scattered across the genome, with small to modest additive effects (Craig, 2008; Taylor, 2013). Genome-wide association studies (GWAS) in OCD have found suggestive evidence for some single nucleotide polymorphisms (SNPs) and genes that are potentially involved in its pathogenesis (International Obsessive Compulsive Disorder Foundation Genetics, 2017). Yet, overall these findings remain rather inconclusive with no single genetic variant reliably replicating across individual studies (Sampaio et al., 2013; Bozorgmehr et al., 2017). These studies did, however, suggest that an increase in sample size will likely aid the identification of genome-wide significant loci, following the example of other psychiatric disorders like major depressive disorder (MDD) (Wray et al., 2018), attention-deficit hyperactivity-disorder (ADHD) (Demontis et al., 2019), or autism spectrum disorder (ASD) (Grove et al., 2019). Another reason for inconclusive findings may be that the majority of current studies of OCD do not account for or put enough emphasis on the heterogeneity of the disorder, though genetic findings may vary as a function of moderator variables (Mataix-Cols et al., 2005; Kulminski et al., 2016; Mattina and Steiner, 2016). One gene that is implicated in one subgroup of OCD patients may not be relevant for another, potentially making it more difficult to find true associations. As 65–85% of OCD patients manifest another psychiatric disorder concomitantly or at some other time point during their lifetime (Tükel et al., 2002; Nestadt et al., 2009; Gillan et al., 2017), often presenting very different symptoms (Ortiz et al., 2016), it raises the question whether comorbid patients form distinct (genetic) subgroups. Nestadt et al. (2009) proposed a sub-classification of OCD

based on comorbidity into three subgroups, with each group being associated with distinct clinical characteristics, prevalence rates, age-of-onsets, and sex-distributions. Dissecting OCD into more homogeneous and accurate sub-phenotypes based on comorbidity, may therefore lead to the successful identification of genetic risk variants for OCD (MacRae and Vasani, 2011; Kulminski et al., 2016).

In recent years, a variety of genetic studies have shown that OCD shares some genetic background with the neuropsychiatric disorders it co-occurs with (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; O'Connell et al., 2018). The genetic correlation of OCD and tourette syndrome (TS) has been estimated at 0.41 ($SE = 0.15$) (Davis et al., 2013), with anorexia nervosa (AN) at 0.49 ($SE = 0.13$) (Yilmaz et al., 2020), with MDD at 0.21 ($SE = 0.05$) (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013), with ASD at 0.12 ($SE = 0.08$) (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013), and ranges between -0.17 ($SE = 0.07$) (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013) and 0.67 ($SE = 0.09$) (Hirschtritt et al., 2018; Goodman et al., 2020) for OCD and ADHD. With a quantitative genetic modeling approach Du Rietz et al. (2020) showed that the phenotypic association between ADHD and an externalizing factor, also loading onto OCD, was largely influenced by genetics and it was demonstrated that both ADHD factors (inattentive and hyperactive/impulsive symptoms) were genetically related to OCD (Hirschtritt et al., 2018). PRS derived from ASD genetic data predicted 0.11% of the phenotypic variance in OCD (Guo et al., 2017). More recently, evidence for disorder-specific genetic associations has also been demonstrated. Peyrot and Price (2021) identified two SNPs distinguishing OCD and ADHD, and one SNP distinguishing OCD and ASD, using a newly developed method to quantify the genetic differences between psychiatric disorders by testing for differences in allele frequencies between cases of two disorders. It has also been shown that the majority of genes that have been implicated in OCD, ASD, schizophrenia (SCZ), and bipolar disorder (BP) are disorder-specific (O'Connell et al., 2018) and that the phenotypic differences between ADHD and OCD are reflected in altered DNA methylation at specific sites, pointing toward heterogeneous regulatory changes in both disorders (Goodman et al., 2020). As OCD shows such a high and

specific genetic overlap with its comorbid neurodevelopmental and psychiatric disorders, while at the same time also presenting very unique genetic correlates, we explored whether OCD comorbid subgroups show a heterogeneous genetic architecture depending on the combination of co-occurring disorders.

In this paper we focused on the potential heterogeneity of OCD subgroups defined by comorbidity with MDD, ADHD, and/or ASD, as these disorders form the biggest comorbidity groups in the iPSYCH OCD sample. While MDD is the most commonly co-occurring diagnosis with OCD (~15–39.5%; Lochner et al., 2014), ADHD occurs in ~6–34% of OCD cases (Geller et al., 2004; Anholt et al., 2010) and OCD patients have a four-fold increased risk of developing ASD (Meier et al., 2015). Because specific markers associated with OCD have not yet been identified, we applied a variety of genome-wide analyses, neither looking for specific associated SNPs nor meta-analysing the iPSYCH samples with the current PGC OCD GWAS, as the sample-size increase would have only been marginal. Instead, in a first step we used publicly available summary statistics from the PGC to compare the genetic landscape of OCD patients to patients with either MDD, ADHD, or ASD. We dissected similarities and differences in correlation patterns of the four disorders with 861 other phenotypes. In a second step we used an independent and previously unpublished OCD dataset from iPSYCH and compared the polygenic architecture of comorbid samples of patients with an OCD diagnosis and a further diagnosis of either MDD, ADHD, ASD, or any combination thereof. We explored differences in polygenic risk score (PRS) load across the different OCD comorbid groups using a multivariate (multiple outcomes) multivariable (multiple covariates) regression, as introduced by Grove et al. (2019). As training datasets we used eight phenotypes from a variety of domains (psychiatric, personality/psychological, anthropomorphic/metabolic, education, and other) that exhibited a range of differing correlation patterns with OCD, MDD, ADHD, and ASD. As OCD, MDD, ADHD, and ASD showed heterogeneous genetic patterns in the analyses in step one, we hypothesized that (a) the comorbid OCD subgroups in the iPSYCH sample would show a heterogeneous association pattern with the PRSes, depending on the training dataset and the combination of comorbid disorders in the OCD subgroup, and (b) that this heterogeneity would be in line with the correlation patterns between OCD, MDD, ADHD, and ASD and the PRS training phenotypes. We expected that the heterogeneity across OCD co-morbid subgroups in the PRS analysis would vary depending on whether the correlations of MDD, ADHD, and ASD showed the same or opposing directions as OCD with the traits used as a training dataset in the PRS analyses (see **Figure 1** for an overview of performed analyses).

2. METHODS

2.1. Subjects

2.1.1. PGC Samples

Publicly available European ancestry GWAS summary statistics of OCD, MDD, ADHD, and ASD were downloaded from the Psychiatric Genomics Consortium (PGC) website (see here). A

description of sample sizes can be found in **Table 1**. Details about the cohorts and data processing have been described in the corresponding primary publications [OCD: International Obsessive Compulsive Disorder Foundation Genetics (2017), MDD: Wray et al. (2018), ADHD: Demontis et al. (2019), ASD: Grove et al. (2019)].

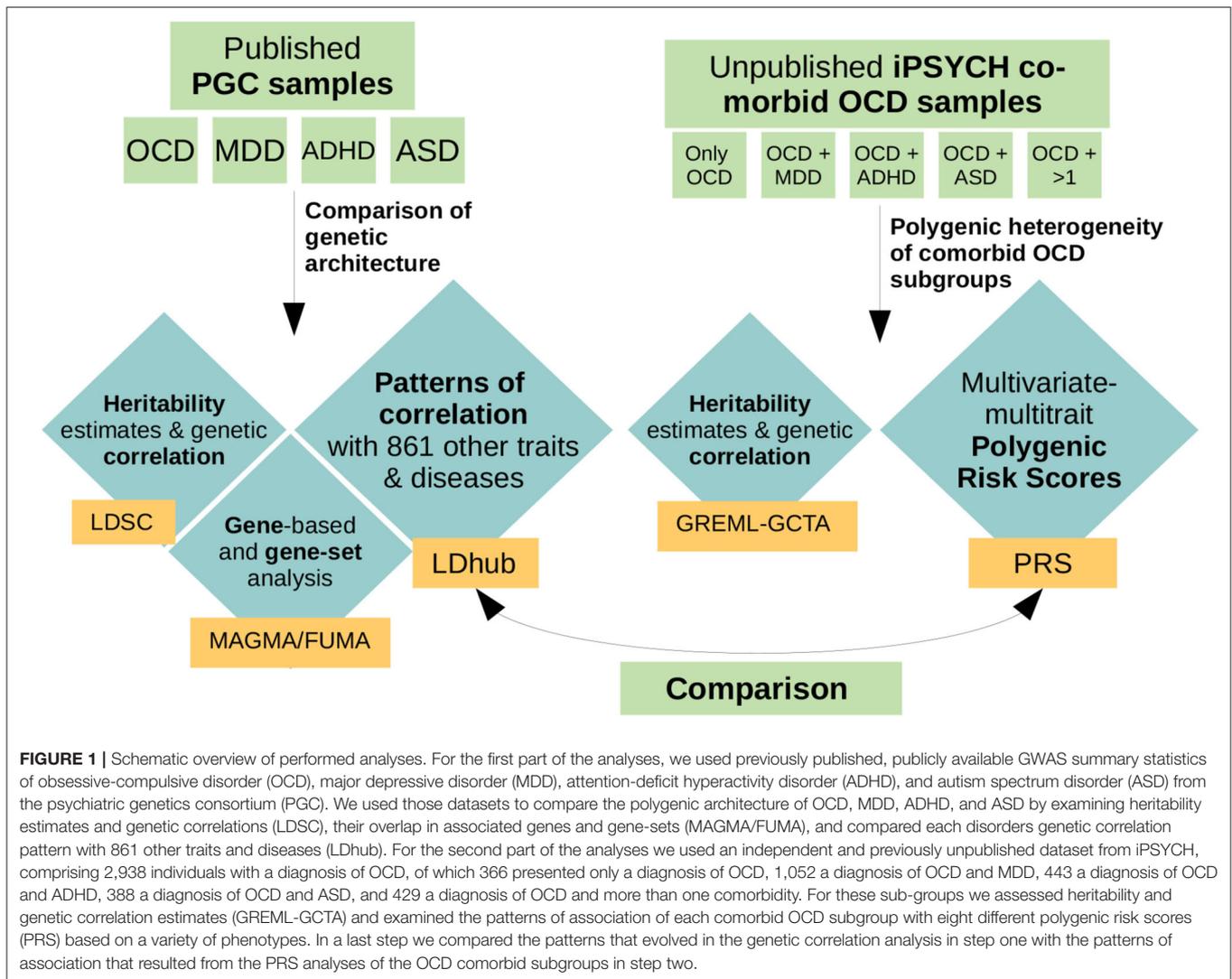
2.1.2. iPSYCH Comorbid OCD Sample

In the scope of the *Danish OCD and Tourette Study* (DOTS) within *The Lundbeck Foundation Initiative for Integrative Psychiatric Research* (iPSYCH), Danish nation-wide population-based case-cohort samples were collected and genotyped. The study was approved by the Regional Scientific Ethics Committee in Denmark and the Danish Data Protection Agency. All analyses of the samples were performed on the secured national GenomeDK high performance-computing cluster in Denmark (<https://genome.au.dk>). See Pedersen et al. (2018) for a detailed description of the overall cohort, array, genotyping, and quality control. Here we give a brief summary: The iPSYCH sample comprised 2,938 individuals with a diagnosis of OCD. All OCD patients that are included in the iPSYCH sample were either comorbid with one of the primary disorders in iPSYCH or were drawn from the population-based pool of controls. For each iPSYCH sample, DNA was obtained from the Danish Neonatal Screening Biobank (DNSB) at the Statens Serum Institut (SSI). Subsequent genotyping was performed in 23 batches on Illumina's PsychChip v 1.0 array (Illumina, San Diego, CA, USA) at the Broad Institute of MIT and Harvard (Cambridge, MA, USA). Cases were identified amongst all individuals in iPSYCH (cases and controls) as individuals that met ICD10 diagnostic criteria for OCD (F42). Controls were randomly selected (for a 4 to 1 matching with cases) from the same cohort, and excluded individuals with a diagnosis of F42. Genotypes were processed using the *Rapid Imputation and Computational Pipeline for Genome-Wide Association Studies* (ricopili) (Lam et al., 2020) performing stringent quality control of the data. Samples with call rates below 98% and individuals with a mismatch between sex obtained from genotyping and registered sex in the iPSYCH database were excluded. Related individuals were removed (randomly one individual per identified pair), principle component analyses were used to exclude ancestral outliers and the data was imputed using the 1,000 Genomes Project phase 3 reference panel (The 1000 Genomes Project Consortium, 2015). The final dataset included 10,411 controls and 2,678 cases of which 366 were diagnosed with only OCD (*onlyOCD*), 1,052 with OCD and MDD (*OCD+MDD*), 443 with OCD and ADHD (*OCD+ADHD*), 388 with OCD and ASD (*OCD+ASD*), and 429 with multiple comorbid disorders (*MC*) (see **Table 1**). Of the cases in the *MC* subgroup, 127 were diagnosed with OCD, MDD, and ASD; 140 with OCD, MDD, and ADHD; 129 with OCD, ASD, and ADHD; and 33 with OCD, ASD, ADHD, and MDD.

2.2. Statistical Analyses

2.2.1. Gene-Based and Gene-Set Analysis

We performed gene-based- and gene-set association analysis of the PGC samples of OCD, MDD, ADHD, and ASD using



the web-based tool *Functional Mapping and Annotation of Genome-Wide Association Studies* (FUMA) v1.3.1 (Watanabe et al., 2017) and *Multi-marker Analysis of GenoMic Annotation* (MAGMA) v1.6 (de Leeuw et al., 2015), employing a multiple regression model while accounting for linkage disequilibrium (LD) between the markers. For both analyses, the default MAGMA settings (SNP-wise model for gene analysis and competitive model for gene-set analysis) were applied. First, FUMA defines genomic risk loci on the basis of independent lead SNPs (with $r^2 < 0.1$ between the independent lead SNPs), merging LD blocks that are physically closer than 250 kb or overlapping into a single locus. Only SNPs in LD with a lead SNP and a minimum association p -value of 0.05 were included for further analysis. Each risk locus is represented by the top lead SNP with the minimum p -value in the locus. For MDD, ASD, and ADHD the minimum p -value of included lead SNPs was set to 5×10^{-8} . Because the OCD GWAS had no SNPs exceeding the genome-wide threshold of 5×10^{-8} the threshold was arbitrarily lowered to 5×10^{-6} . The

minimum allele frequency (MAF) threshold was set to 0.01. One thousand genomes project phase 3 (The 1000 Genomes Project Consortium, 2015) was used as a reference panel to calculate LD across SNPs and genes and the MHC region was excluded. The gene-based p -values were computed by mapping SNPs to their corresponding gene(s) on the basis of their position in the genome. Positional mapping was based on ANNOVAR annotations and the maximum distance between SNPs and genes was set to 10 kb. To correct for multiple testing, Bonferroni correction and false-discovery rate (FDR) was applied for gene-analysis and gene-set analysis, respectively. For OCD, input SNPs were mapped to 18,709 protein-coding genes, genome-wide significance was defined at a Bonferroni corrected threshold of $p = 2.67 \times 10^{-6}$. FUMA tested curated gene-sets (c2.all) and gene ontology (GO) terms, using 10,894 gene-sets for FUMA \leq version 1.3.0 (ADHD) and 10,655 gene-sets for FUMA \geq version 1.3.1 (OCD, MDD, ASD). Gene-set p -values were computed using the gene-based p -values of all genes for each curated gene-set.

TABLE 1 | PGC and iPSYCH sample sizes, population prevalences and heritability estimates (h_g^2).

Phenotype	Ncases	Ncontrols	Ntotal	Ppoprev.	h_g^2 (SE)	P
PGC						
OCD	2,688	7,037	9,725	0.03	0.29 (0.05)	–
MDD	59,851	113,154	173,005	0.15	0.09 (0.01)	–
ADHD	19,099	34,194	53,293	0.05	0.21 (0.01)	–
ASD	18,382	27,969	46,351	0.01	0.11 (0.01)	–
iPSYCH						
onlyOCD	366	10,411	10,819	0.01	0.29 (0.09)	0.0003
OCD+MDD	1,052	10,411	11,543	0.005	0.08 (0.03)	0.0035
OCD+ADHD	443	10,411	10,901	0.002	0.04 (0.05)	0.2312
OCD+ASD	388	10,411	10,840	0.0003	0.03 (0.04)	0.2289
MC	429	10,411	10,890	0.0001	0.12 (0.03)	<0.0001

For PGC samples heritability (h_g^2) was estimated using LDSC, for the iPSYCH sub-samples univariate-GREML estimates of SNP-heritability are presented. All heritability estimates are reported on the liability scale (adjusted for population prevalence). Controls were the same for all iPSYCH subgroups. Abbreviations: Number of cases (Ncases), number of controls (Ncontrols), total number of individuals (Ntotal), population prevalence (Ppoprev.) heritability (h_g^2), standard error of the heritability estimate (SE), p-value of the heritability estimate (P), OCD subgroup with more than one comorbidity (MC). Within the MC group, 141 subjects are diagnosed with OCD, MDD, and ASD; 151 subjects are diagnosed with OCD, MDD, and ADHD, 153 subjects are diagnosed with OCD, ASD, and ADHD; and 34 subjects are diagnosed with OCD, MDD, ASD, and ADHD.

2.2.2. SNP-Heritability Estimates

SNP-heritability (h_g^2) was estimated using LDSC (Bulik-Sullivan et al., 2015a,b; Zheng et al., 2017) for the PGC samples and univariate *genetic-relationship restricted maximum likelihood* (GREML) as implemented in *Genome-wide Complex Trait Analysis* (GCTA) (Lee et al., 2011; Yang et al., 2011) for the iPSYCH OCD subgroups, as sample sizes of the subgroups were too small for LDSC and raw genotype data was available. For LDSC, freely available precomputed LD scores based on the European ancestry samples of the 1,000 G phase 3 (The 1000 Genomes Project Consortium, 2015), restricted to HapMap3 SNPs, were used. Before the analysis, standard LDSC filtering was applied. Poorly imputed SNPs with $INFO < 0.9$ were removed. For the conversion of observed-scale- to liability-scale estimates, previously reported disorder-specific prevalence rates were used (see Table 1).

For the comorbid iPSYCH samples the univariate GREML approach of GCTA was used. After removal of ancestry outliers, counts of each sub-phenotype were the following: *controls*: 10,411, *onlyOCD*: 366, *OCD+MDD*: 1,052, *OCD+ADHD*: 443, *OCD+ASD*: 388, *MC*: 429. A genetic relatedness matrix (GRM) was fitted, thereby providing relatedness estimates for all pairwise combinations of individuals. All indels were removed and the data was filtered on *genotype probability* > 0.8 , *missing rate* < 0.01 and *MAF* > 0.05 . GRM was estimated for each individual autosome and subsequently merged into a single GRM based on all autosomes. h_g^2 estimation for each OCD sub-phenotype was performed including the first four principle components (PCs) as continuous covariates together with any other PC that was nominally significantly associated to the phenotype. Waves were included as categorical indicator covariates. Lacking proper population prevalence estimates for subgroups, prevalence rates for comorbid conditions were estimated by multiplying the prevalence for each comorbid disorder with the OCD prevalence (3%). The prevalence for the OCD subgroup with more than

one comorbid disorder was estimated to be lower than any of the other prevalence rates at an arbitrary value of 0.01%, as the multiplication of more than two prevalence rates would strongly underestimate the true prevalence. Because at least one other psychiatric disorder is present in approximately two thirds of OCD patients (Tükel et al., 2002; Gillan et al., 2017), the prevalence for *only OCD*, without any comorbid diagnosis, was set to 1% (one third of the general OCD prevalence). See Table 1 for a list of all population prevalence estimates.

2.2.3. Genetic Correlation Estimates

Using LDSC (Bulik-Sullivan et al., 2015a,b) we estimated the genetic correlation (r_G) of OCD with MDD, ADHD, and ASD. We further estimated each disorder's genetic correlation with 861 other phenotypes using LDSC as implemented in LDhub (Zheng et al., 2017) (for 855 traits) and LDSC (for six additional datasets/traits not contained in the LDhub database). We corrected for multiple testing by setting the significance threshold to a Bonferroni-corrected p-value (dividing 0.05 by the number of valid tests per disorder). We then compared the correlation patterns that emerged for OCD to those of MDD, ADHD, and ASD.

Bi-variate GREML as implemented in GCTA was used to estimate the genetic correlation between the iPSYCH OCD subgroup samples. The controls were split proportionally in order to guarantee an independent control group for each comorbid subgroup in every pairwise comparison.

2.2.4. Multivariate-Multitrait PRS Analyses (PRS)

By applying multivariate (multiple outcomes) multivariable (multiple covariates) regression (Grove et al., 2019) we examined the distribution of PRSs based on OCD (International Obsessive Compulsive Disorder Foundation Genetics, 2017), neuroticism (Nagel et al., 2018), anorexia nervosa (AN) (Watson et al., 2019), bipolar disorder (BP) (Stahl et al., 2019), Educational Attainment

(EA) (Lee et al., 2018), body mass index (BMI) (Yengo et al., 2018), age at first birth (AFB) (Barban et al., 2016), and insomnia (Jansen et al., 2019), over the OCD comorbid subgroups. For the calculation of PRSs, the summary statistics of interest were clumped by applying standard *ricopili* parameters. Prior to clumping overlapping SNPs between the iPSYCH data and the external summary statistics were extracted and strand ambiguous A/T and C/G SNPs with a frequency between 0.4 and 0.6 were removed to avoid potential strand conflicts. PRS were generated at the default p -value thresholds of 5×10^{-8} , 1×10^{-6} , 1×10^{-4} , 0.001, 0.01, 0.05, 0.1, 0.2, 0.5, and 1 as a weighted sum of the risk allele dosages. Prior to analysis scores were normalized. After the PRS were calculated, the scores were regressed onto the OCD subgroups to evaluate the genetic overlap between the phenotypes and the OCD subgroups. Batch effects from genotyping waves and PCs in the comorbid OCD data were adjusted for in the multivariate multivariable regression. The advantage of a multivariate regression is that it can handle a possible correlation among the PRSes, making it possible to test a great number of hypotheses across PRSes and subtypes. The approach is statistically very powerful which enables us to conduct these analyses even with sample sizes too small to conduct a GWAS or LDSC analysis.

3. RESULTS

3.1. Comparing the Genetic Architecture of OCD, MDD, ADHD, and ASD

3.1.1. Gene and Gene-Set Analysis

First, we performed gene-based- and gene-set association analysis of the PGC samples of OCD, MDD, ADHD, and ASD using MAGMA/FUMA, thereby looking for potential overlaps in associated genes and gene-sets between the four disorders. When looking at 13 genes that showed suggestive association for OCD ($p < 1 \times 10^{-4}$; strongest association for *KIT Proto-Oncogene Receptor Tyrosine Kinase* on chromosome 4, $p = 2.46 \times 10^{-7}$) there was no evident overlap with significant genes of the other disorders (see **Supplementary Table 1**). Furthermore, no gene-set ($p \leq 9.7 \times 10^{-5}$) overlapped between OCD, MDD, ADHD, and ASD (see **Supplementary Table 2** for gene-set results of OCD).

3.1.2. Heritability and Genetic Correlations

Next, we computed SNP heritabilities (h_g^2) of OCD, MDD, ADHD, and ASD (see **Table 1**) and calculated cross-trait genetic correlations (r_G) between each pair of disorders using LDSC (Bulik-Sullivan et al., 2015a,b). OCD was significantly positively correlated with MDD ($r_G = 0.23$, $SE = 0.07$, $p = 0.0005$) and nominally significantly negatively correlated with ADHD ($r_G = -0.17$, $SE = 0.07$, $p = 0.02$), while the correlation between OCD and ASD did not reach significance ($r_G = 0.12$, $SE = 0.08$, $p = 0.15$).

To investigate the extent of genetic overlap between OCD and an array of other phenotypes, we estimated its genetic correlations with 861 psychiatric and other medical diseases, disorders, and traits using bivariate LD score regression (Bulik-Sullivan et al., 2015a,b; Zheng et al., 2017). The same analysis was

also performed for MDD, ADHD, and ASD as we were interested in similarities and differences in patterns of correlations between the four disorders. 777 (for OCD and ADHD), 778 (for ADHD), and 779 (for MDD) genetic correlations yielded interpretable results, the remaining estimations resulted in “NA,” due to small sample size and non-significant heritability. We therefore set the significance threshold to a Bonferroni-corrected p -value of $0.05/779 = 6.42 \times 10^{-5}$. Of the tested diseases and traits, 45 were significantly correlated with OCD, 249 with MDD, 285 with ADHD, and 52 with ASD (disregarding a phenotype if has been represented by a different dataset already). Forty traits overlapped between OCD and MDD, 37 between OCD and ADHD, and 12 between OCD and ASD. Nine traits were significantly associated with all four disorders, of which five demonstrated the same direction of effect (see **Supplementary Table 3**).

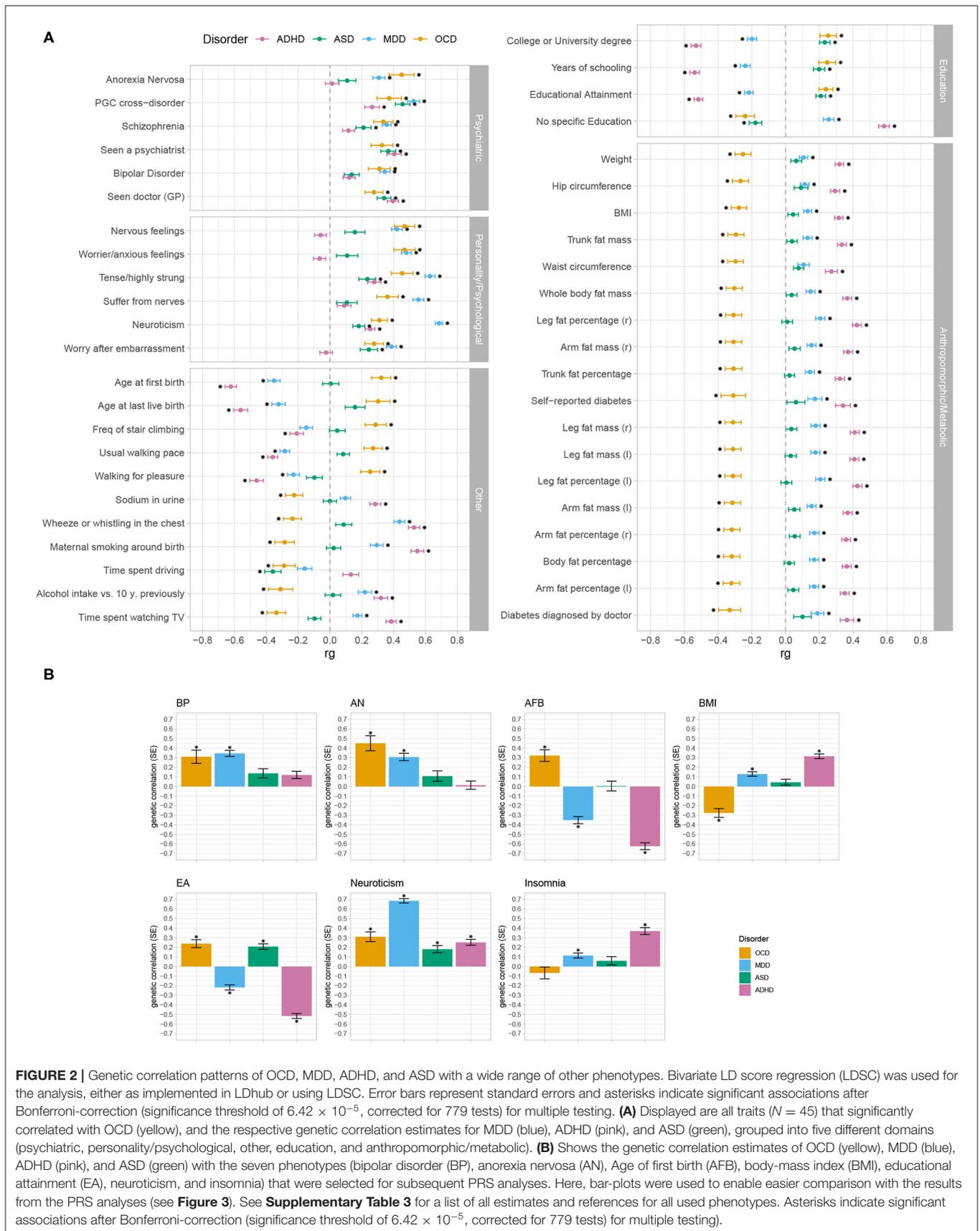
All phenotypes that significantly correlated with OCD were grouped into five categories: psychiatric, personality/psychological, anthropomorphic/metabolic, education, and other (see **Figure 2A**). Across the four disorders (OCD, MDD, ADHD, and ASD), differences in their patterns of correlations emerged. While all four disorders generally showed positive associations with traits in *psychiatric disorders* and *personality/psychological* traits, ASD and ADHD exhibited fewer significant and in several cases less strong associations compared to OCD and MDD. In the category of *other*, OCD was positively correlated with fertility parameters (*AFB* and *Age at last life birth*) and movement parameters, and negatively correlated with all other parameters, while ADHD and MDD generally showed the opposite pattern of correlation. While OCD and ASD positively correlated with *education* traits, ADHD and MDD negatively correlated with all *education* parameters (for *no specific qualifications* the pattern was reversed). In the category of *anthropomorphic and metabolic* traits, OCD significantly correlated negatively with all reported parameters, while MDD correlated moderately positively and ADHD strongly positively with the same phenotypes. In this category, ASD did not significantly correlate with any of the traits.

We further selected an array of traits across all five domains (see **Figure 2B**) to evaluate how PRS based on a broad spectrum of phenotypes with varying patterns of correlations with OCD, MDD, ADHD, and ASD, partition across comorbid OCD subgroups. See section 3.2.2 for details on which traits were selected for analysis.

3.2. Dissection of the Polygenic Architecture of Comorbid OCD Subgroups

3.2.1. Heritability and Genetic Correlations Among the Subgroups

Next, we explored the polygenic heterogeneity across OCD comorbid subgroups. We examined how h_g^2 partitioned across the comorbid OCD subgroups and estimated the genetic correlation among these groups using GCTA (Yang et al., 2011). Univariate GREML analysis revealed significant h_g^2 for the *onlyOCD*, *OCD+MDD*, and *MC* subgroups (see **Table 1** for all h_g^2 estimates). Pairwise comparisons of genetic correlations (r_G) of



the sub-phenotypes were estimated with bivariate GREML. Each subgroup demonstrated a high genetic correlation with all other subgroups (between 0.2 and 1; see **Supplementary Table 4** for a list of the results). Standard errors were generally very high for all pairwise correlations, making it difficult to interpret the results.

3.2.2. Cross-Trait PRS Analyses

To examine a possible polygenic heterogeneity of OCD, we further investigated how PRS trained on different phenotypes (OCD, neuroticism, EA, AN, BP, BMI, AFB, and insomnia) distribute across the iPSYCH OCD subgroups defined by a comorbid diagnosis of either MDD, ADHD, and/or ASD. Traits for the PRS analysis were selected from across all tested domains (psychiatric, personality/psychological, anthropomorphic/metabolic, education, and other) and were chosen in view of their different correlation patterns with OCD, MDD, ADHD, and ASD. Thereby, we wanted to explore whether different correlation patterns with OCD, MDD, ADHD, and ASD would translate into differing patterns in the PRS analysis across the OCD comorbid subgroups. The traits used as training datasets in the PRS analysis either showed (a) a significant correlation with OCD, MDD, ADHD, and ASD in either the same direction (BMI) or differing directions (EA); or (b) a significant correlation with OCD and either one (BP) or two (AFB, BMI, AN) other disorders; or (c) no significant correlation with OCD but a significant correlation with two other tested disorders (insomnia). Further, OCD itself was included as a training dataset for the PRS analysis. With this selection of phenotypes we aimed to explore whether a heterogeneous genetic correlation pattern between a phenotype and OCD, MDD, ADHD, and ASD translates into heterogeneous PRS loadings in the OCD comorbid subgroups.

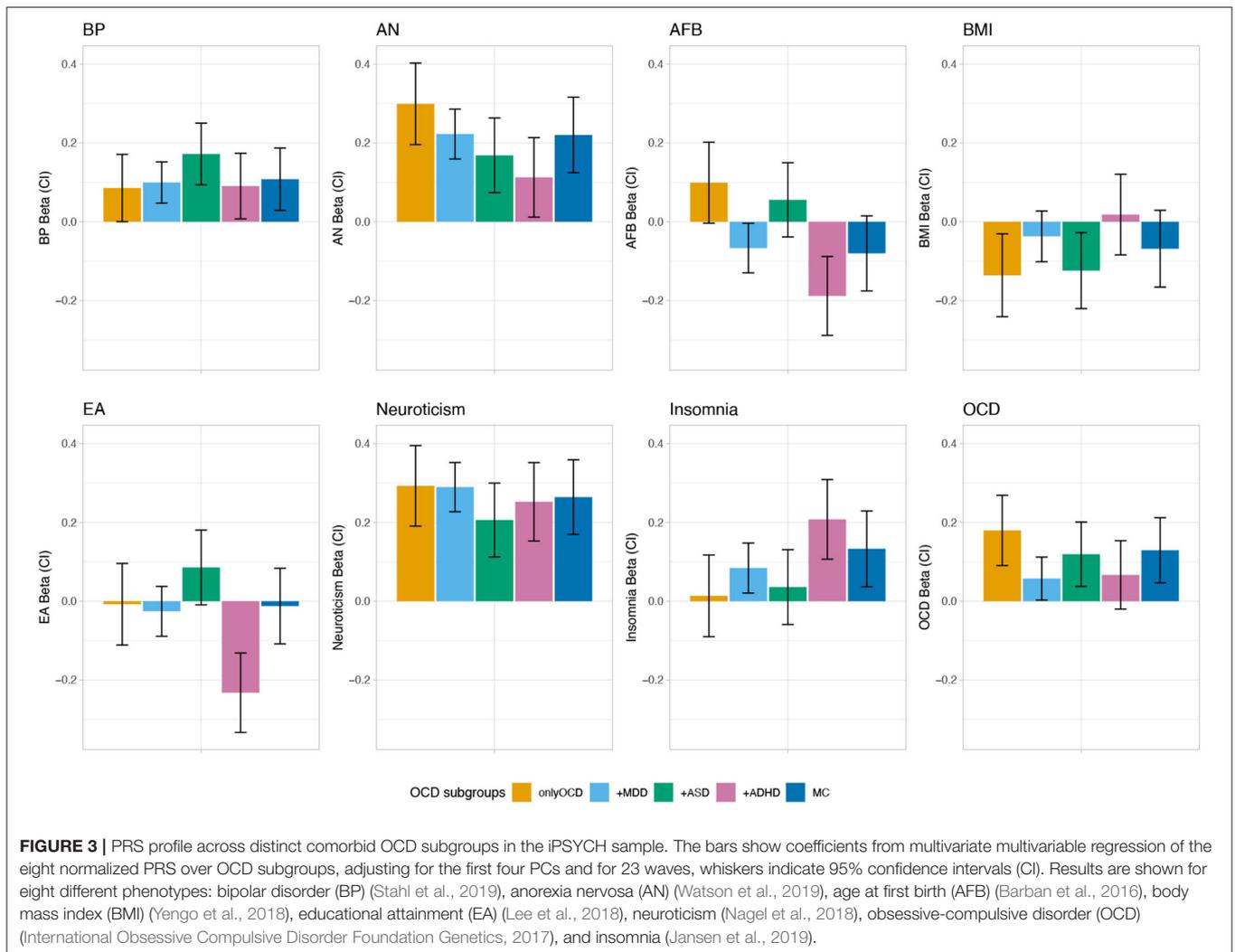
The PRS analysis can be read as a linear regression with the beta value indicating the mean level of PRS relative to the controls, adjusted for the other variables and covariates (first four principle components and batches). First, for each phenotype, it was tested whether the betas of the PRS analyses were significantly different from zero across all OCD comorbid subgroups. Neuroticism, BP, AN, AFB, EA, OCD, and insomnia showed significant associations with the iPSYCH OCD samples ($p = 1.19 \times 10^{-32}$, $p = 7.51 \times 10^{-8}$, $p = 3.52 \times 10^{-20}$, $p = 9.38 \times 10^{-5}$, $p = 1.56 \times 10^{-4}$, $p = 1.87 \times 10^{-6}$, $p = 2.61 \times 10^{-5}$, respectively; see **Supplementary Table 5**). Of the eight phenotypes tested (neuroticism, BP, AN, AFB, EA, OCD, BMI, and insomnia) for association with the OCD comorbid subgroups, AFB ($p = 2.29 \times 10^{-4}$), EA ($p = 1.63 \times 10^{-4}$), and insomnia ($p = 0.045$) showed a significant heterogeneity across OCD subgroups. BP and AN were positively associated with all OCD subgroups, while the other traits showed significant associations with some of the OCD comorbid subgroups, but not with all (see **Figure 3** and **Supplementary Table 6**). For AFB the strongest, though non-significant, positive associations were with the *onlyOCD* group ($Beta = 0.099$, $CII = -0.004$, $CIU = 0.202$, $p = 0.059$), followed by the *OCD+ASD* ($Beta = 0.056$, $CII = -0.039$, $CIU = 0.15$, $p = 0.247$) group. The strongest negative association was with the *OCD+ADHD* group ($Beta = -0.188$,

$CII = -0.288$, $CIU = -0.088$, $p = 2.29 \times 10^{-4}$), followed by the, though non-significant, *MC* group ($Beta = -0.08$, $CII = -0.176$, $CIU = 0.015$, $p = 0.098$), and *OCD+MDD* group ($Beta = -0.067$, $CII = -0.13$, $CIU = -0.004$, $p = 0.037$) (see **Figure 3** and **Supplementary Tables 6, 7** for results of all tested phenotypes). For EA, there was a strong negative association with *OCD+ADHD* ($Beta = -0.232$, $CII = -0.333$, $CIU = -0.131$, $p = 6.36 \times 10^{-6}$) and a trend for a positive association with *OCD+ASD* ($Beta = 0.086$, $CII = -0.009$, $CIU = 0.180$, $p = 0.077$), while the other subgroups demonstrated scores around zero. For the PRS based on insomnia the strongest positive association was with the *OCD+ADHD* ($Beta = 0.208$, $CII = 0.107$, $CIU = 0.309$, $p = 5.62 \times 10^{-5}$) group, followed by the *MC* ($Beta = 0.133$, $CII = 0.037$, $CIU = 0.229$, $p = 5.62 \times 10^{-3}$) and the *OCD+MDD* group ($Beta = 0.084$, $CII = 0.021$, $CIU = 0.148$, $p = 9.39 \times 10^{-3}$).

4. DISCUSSION

In the present study we first looked at genetic similarities and differences between OCD and the three psychiatric disorders MDD, ADHD, and ASD, with a specific emphasis on the genetic correlation patterns of each of the four disorders with 861 somatic and mental health phenotypes. In a second step we used genome-wide data of an independent set of OCD patients from iPSYCH for which we defined five OCD subgroups based on the patients' comorbidity with MDD, ADHD and/or ASD (*onlyOCD*, *OCD+MDD*, *OCD+ADHD*, *OCD+ASD*, and *MC*). Using eight different traits (BP, AN, AFB, BMI, EA, neuroticism, insomnia, and OCD) as training data sets, we applied PRS analysis across the comorbid OCD subgroups. Traits were selected from a variety of domains (psychiatric, personality/psychological, anthropomorphic/metabolic, education, and other), on the basis of their differential correlation patterns with OCD, MDD, ADHD, and ASD. We hypothesized that (a) the comorbid OCD subgroups show a heterogeneous association pattern with the PRSes, depending on the training dataset and the combination of comorbid disorders in the OCD subgroup, and (b) that the heterogeneous association patterns of the comorbid OCD subgroups are in accordance with the correlation patterns between OCD, MDD, ADHD, and ASD and the PRS training phenotypes that were reported in the first part of the manuscript. For example, if OCD showed a positive correlation with trait A and MDD a negative correlation with trait A, we expected that the *onlyOCD* group would show a higher association with the PRS based on trait A than the comorbid subgroup of *OCD+MDD*, while we hypothesized that a positive correlation of both, MDD and OCD, with trait A would translate into either an increased or similar association of the PRS based on trait A with the *OCD+MDD* comorbid subgroup compared to the *onlyOCD* group.

The genetic correlation patterns that emerged in the first part of the analysis are generally in accordance with symptomatic and clinical observations of OCD, MDD, ADHD, and ASD patients. As the sample size of a GWAS has an influence on the standard error and p -value of genetic correlation point estimates, it was expected that the OCD GWAS showed a



lower number of significant genetic correlations with the 861 tested traits, as compared to the larger MDD, ADHD, and ASD GWASs. We cannot exclude the possibility that with an increase in sample sizes more genetic correlations will become significant. However, as we only consider significant genetic correlations for interpretation, this should not have an influence on the results discussed here. All four disorders displayed positive associations with most of the other psychiatric disorders and with personality/psychological parameters, such as BMI, worry, and tense feelings. The genetic correlation of OCD with anthropomorphic and metabolic traits was negative, while MDD and ADHD showed a positive correlation. This is in line with the observation that OCD is genetically positively correlated with AN (The Brainstorm Consortium et al., 2018), as AN correlates negatively with weight parameters on a symptomatic and genetic level (Speranza et al., 2001; Duncan et al., 2017). OCD and ASD showed a positive correlation with education parameters and OCD correlated negatively with age at first birth (there was no significant correlation between ASD and age at first birth), while the pattern was reversed for ADHD

and MDD. Dalsgaard et al. (2020) recently demonstrated that males with OCD achieved significantly higher school grades than individuals without a psychiatric disorder, while people with other psychiatric disorders (except AN) had significantly lower grades. It was also shown that higher education and socio-economic status are associated with higher maternal age at first birth (Van Roode et al., 2017) and that children of young mothers were disadvantaged in schooling (Fall et al., 2015).

Because the four disorders showed differential genetic correlation patterns, we presumed that the polygenic architecture of comorbid OCD subgroups would vary depending on their comorbid diagnosis. We first looked at heritability estimates and genetic correlations between the comorbid OCD subgroups. The *onlyOCD* and the *MC* group demonstrated the highest heritability estimates, while the *OCD+ASD* group displayed the lowest heritability estimates compared to all other subgroups. As sample sizes in each comorbidity group were quite low, SE were generally high and not all of the heritability estimates and none of the genetic correlation estimates between the comorbid subgroups reached significance.

In a last step we then applied PRS analysis across the iPSYCH OCD comorbid subgroups. Rather than selecting traits used as training datasets on a theoretical or clinical background, they were selected in view of their different correlation patterns with OCD, MDD, ADHD, and ASD across a wide range of psychiatric and somatic phenotypes, as we wanted to explore whether the different directions of correlations would be mirrored in the PRS analysis of the OCD comorbid subgroups. For traits for which OCD, MDD, ADHD, and ASD showed a heterogeneous genetic correlation pattern (EA, AFB, BMI) we hypothesized that PRSes based on those traits would also exhibit a heterogeneous pattern of association with the comorbid OCD subgroups. For EA and AFB the pattern of PRS loadings that emerged across the OCD comorbid subgroups closely mirrored the concordance structure of the genetic correlations between OCD and MDD, ADHD, and ASD. OCD and ASD correlated positively with Years of schooling and College or university degree, while it was the opposite for ADHD and MDD. Accordingly, in the PRS analysis the *OCD+ADHD* group had the highest negative loading for EA, while the EA PRS estimate was positive in the *OCD+ASD* group. Similarities between the correlation analysis and PRS analysis could also be shown for AFB. OCD correlated positively, MDD and ADHD negatively with AFB. ASD did not show a significant correlation. Similarly, in the PRS analysis, AFB was positively associated with disease status in the *onlyOCD* group and to a lower degree also in the *OCD+ASD* group, while it was negatively associated with the *OCD+MDD*, *OCD+ADHD*, and *MC* group. The PRS loadings for BMI was the most negative for the *onlyOCD* group, but also showed a negative association with *OCD+ASD* and *MC*, while, somewhat surprisingly, the other OCD subgroups were not significantly associated with the BMI PRS. One possible explanation for this pattern may be that the negative correlation between OCD and BMI and the positive correlations between ADHD and BMI, as well as between MDD and BMI translate into a null-finding in the PRS finding for BMI because the opposing correlations may evoke counteracting effects in the comorbid subgroups. As neuroticism showed a fairly homogeneous correlation with OCD, MDD, ADHD, and ASD, we expected no polygenic heterogeneity across comorbid OCD subgroups. Similarly, for AN and BP, which correlated significantly positively with OCD and MDD, and positively but non-significantly with ASD and ADHD, we expected a rather homogeneous pattern of association with PRSs across the subgroups, with stronger associations for *onlyOCD* and *OCD+MDD*. This was indeed the case, as PRSes of neuroticism, and BP were associated with OCD across all comorbid subgroups with no significant differences in estimates between the OCD comorbid subgroups. For AN, the pattern of correlations was mirrored closely in the PRS analysis—*onlyOCD* and *OCD+MDD* demonstrated the highest PRS estimates, followed by *OCD+ASD* and *OCD+ADHD*, with a significant difference between the highest estimate for *onlyOCD* and the lowest estimate for *OCD+ADHD*. Because we were also interested how PRS estimates change for traits which showed no correlation with OCD but with some of the other three disorders, we also included insomnia in the PRS analysis. While the insomnia PRS was not significantly associated with the

onlyOCD subgroup, it showed significant associations with the *OCD+MDD*, *OCD+ADHD*, and *MC* subgroups, indicating that a comorbid diagnosis might change the association of OCD and insomnia.

To conclude, the different PRS estimates across OCD subsets provide the first evidence for a heterogeneous and qualitatively different genetic architecture of OCD subgroups defined by a comorbid diagnosis of MDD, ADHD, and/or ASD. Traits that show a heterogeneous genetic correlation pattern with OCD, MDD, ADHD, and ASD generally also exhibit a heterogeneous pattern of estimations in PRS analysis across OCD comorbid subgroups. This was especially shown for AFB, and EA. While being unique in its approach, results of the present study are in accordance with previous research by Hirschtritt et al. (2018) who examined OCD- and ADHD-symptom dimensions in TS cases and identified unique OCD symptom subgroups that were differentially associated with other comorbid psychiatric disorders. Both, OCD symptom subgroups and comorbid subgroups, may be markers of distinct underlying patterns of psychopathology and genetic architecture.

Because heterogeneous genetic architectures could potentially point toward heterogeneous disease mechanisms, the context in which OCD occurs may have implications for diagnostic criteria and treatment that might not have been considered sufficiently in past and present research and clinical practice. Pallanti et al. (2011), for example, showed that OCD in the presence of comorbid conditions is often associated with non-response to treatment, indicating differential clinical characteristics. Also, for the success of GWAS analyses, it may be beneficial to focus on (sub)phenotype definitions rather than solely relying on increasing sample size. As MacRae and Vasan (2011) and Kulminski et al. (2016) have discussed, increasing the size of many human disease cohorts is likely only to upscale the heterogeneity in parallel. Especially for cross-disorder GWAS analyses, which have gained a lot of attention recently (Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Grotzinger et al., 2019; Lee et al., 2019; Abdellaoui et al., 2020), it may be crucial to account for comorbidities to avoid confounding of genetic similarities and differences between psychiatric disorders. One limitation of this study is the right censoring of comorbidities. While ADHD and ASD are neurodevelopmental disorders with mostly a childhood onset (and some persistency into adulthood), MDD usually occurs with an onset in late adolescence and adulthood. Therefore, the possibility that an individual develops a comorbidity, or another comorbidity on top of an already existing one, cannot be ruled out and may be higher for disorders with a later age of onset. Inherently, iPSYCH is a longitudinal study. As with other studies, however, it may be the case that some study participants (e.g., those originally ascertained for their ADHD and/or ASD diagnosis) were included at a time point at which the follow-up time was not sufficient to capture a later diagnosis of one of the comorbidities under study (e.g., MDD). While right censoring may dampen some of the observed effects, it is unlikely to alter the overall observations of this study and its main finding of a heterogeneous genetic architecture of comorbid subgroups.

The present study should be viewed as a pilot study and exploratory in nature. In the future, it would be of interest to conduct similar analyses with a broader range of correlated phenotypes and to include other related and comorbid disorders, such as schizophrenia, BP, AN, Tourette's syndrome and anxiety disorders. It has also been suggested that the onset of OCD (early vs. late) (Hemmings et al., 2004; Walitza et al., 2010; Taylor, 2011), sex (male vs. female) (Khramtsova et al., 2019), or different symptom dimensions of OCD (Hasler et al., 2005) present differing underlying genetic architectures.

DATA AVAILABILITY STATEMENT

The data used in this study can be made available upon request to the authors. Requests to access these datasets should be directed to Anders D. Børghlum, anders@biomed.au.dk.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Regional Scientific Ethics Committee in Denmark and the Danish Data Protection Agency. Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

NS and MM contributed to the conception and design of the study. AB, OM, MN, TW, DH, and PM contributed to the conception and design of the original iPSYCH study. SM, MB-H, JB, TD, MH, JC, and JB-G contributed to the data collection

REFERENCES

- Abdellaoui, A., Smit, D. J. A., van den Brink, W., Denys, D., and Verweij, K. J. H. (2020). Genomic relationships across psychiatric disorders including substance use disorders. *Drug and Alcohol Dependence*. 220:108535. doi: 10.1016/j.drugalcdep.2021.108535
- Anholt, G. E., Cath, D. C., Van Oppen, P., Eikelenboom, M., Smit, J. H., Van Megen, H., et al. (2010). Autism and ADHD symptoms in patients with OCD: are they associated with specific OC symptom dimensions or OC symptom severity. *J. Autism Dev. Disord.* 40, 580–589. doi: 10.1007/s10803-009-0922-1
- Barban, N., Jansen, R., De Vlaming, R., Vaez, A., Mandemakers, J. J., Tropf, F. C., et al. (2016). Genome-wide analysis identifies 12 loci influencing human reproductive behavior. *Nat. Genet.* 48, 1–7. doi: 10.1038/ng.3698
- Bozorgmehr, A., Ghadirivassfi, M., and Shahsavand Ananloo, E. (2017). Obsessive-compulsive disorder, which genes? Which functions? Which pathways? An integrated holistic view regarding OCD and its complex genetic etiology. *J. Neurogenet.* 31, 153–160. doi: 10.1080/01677063.2017.1336236
- Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P. R., et al. (2015a). An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* 47, 1236–1241. doi: 10.1038/ng.3406
- Bulik-Sullivan, B. K., Loh, P.-R., Finucane, H. K., Ripke, S., Yang, J., Patterson, N., et al. (2015b). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* 47, 291–295. doi: 10.1038/ng.3211

and organization of the database. NS, JG, and TD performed the statistical analysis. NS wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

The iPSYCH team was funded by the Lundbeck Foundation (R102-A9118, R155-2014-1724, and R248-2017-2003), the EU H2020 Program (Grant No. 667302, CoCA), NIMH (1U01MH109514-01) and the universities and university hospitals of Aarhus and Copenhagen. This study was supported by NIH grants R01MH105500 and R01MH110427.

ACKNOWLEDGMENTS

This research has been conducted using the Danish National Biobank resource, supported by the Novo Nordisk Foundation. High-performance computer capacity for handling and statistical analysis of iPSYCH data on the GenomeDK HPC facility was provided by the Center for Genomics and Personalized Medicine and the Centre for Integrative Sequencing, iSEQ, Aarhus University, Denmark. We acknowledge support by the German Research Foundation (DFG) and the Open Access Publication Fund of Humboldt-Universität zu Berlin.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.711624/full#supplementary-material>

- Craig, J. (2008). Complex diseases: research and applications. *Nature Education*. 1:184. Available online at: <https://www.nature.com/scitable/topicpage/complex-diseases-research-and-applications-748/>.
- Cross-Disorder Group of the Psychiatric Genomics Consortium (2013). Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 381, 1371–1379. doi: 10.1016/S0140-6736(12)62129-1
- Dalsgaard, S., McGrath, J., Østergaard, S. D., Wray, N. R., Pedersen, C. B., Mortensen, P. B., et al. (2020). Association of mental disorder in childhood and adolescence with subsequent educational achievement. *JAMA Psychiatry* 77, 797–805. doi: 10.1001/jamapsychiatry.2020.0217
- Davis, L. K., Yu, D., Keenan, C. L., Gamazon, E. R., Konkashbaev, A. I., Derks, E. M., et al. (2013). Partitioning the heritability of tourette syndrome and obsessive compulsive disorder reveals differences in genetic architecture. *PLoS Genet.* 9:e1003864. doi: 10.1371/journal.pgen.1003864
- de Leeuw, C. A., Mooij, J. M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* 11:e1004219. doi: 10.1371/journal.pcbi.1004219
- Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., et al. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* 51, 63–75. doi: 10.1038/s41588-018-0269-7
- Du Rietz, E., Petteesson, E., Brikell, I., Ghirardi, L., Chen, Q., Hartman, C., et al. (2020). Overlap between attention-deficit hyperactivity disorder and neurodevelopmental, externalising and internalising disorders: separating

- unique from general psychopathology effects. *Br. J. Psychiatry* 218, 35–42. doi: 10.1192/bjp.2020.152
- Duncan, L., Yilmaz, Z., Walters, R., Goldstein, J., Anttila, V., and Bulik-Sullivan, B. K. (2017). Genome-wide association study reveals first locus for anorexia nervosa and metabolic correlations. *Am. J. Psychiatry* 174, 850–858. doi: 10.1176/appi.ajp.2017.16121402
- Fall, C. H., Sachdev, H. S., Osmond, C., Restrepo-Mendez, M. C., Victora, C., Martorell, R., et al. (2015). Association between maternal age at childbirth and child and adult outcomes in the offspring: a prospective study in five low-income and middle-income countries (COHORTS collaboration). *Lancet Global Health* 3, e366–e377. doi: 10.1016/S2214-109X(15)00038-8
- Geller, D. A., Biederman, J., Faraone, S., Spencer, T., Doyle, R., Mullin, B., et al. (2004). Re-examining comorbidity of obsessive compulsive and attention-deficit hyperactivity disorder using an empirically derived taxonomy. *Eur. Child Adolesc. Psychiatry* 13, 83–91. doi: 10.1007/s00787-004-0379-x
- Gillan, C. M., Fineberg, N. A., and Robbins, T. W. (2017). A trans-diagnostic perspective on obsessive-compulsive disorder. *Psychol. Med.* 47, 1528–1548. doi: 10.1017/S0033291716002786
- Goodman, S. J., Burton, C. L., Butcher, D. T., Siu, M. T., Lemire, M., Chater-Diehl, E., et al. (2020). Obsessive-compulsive disorder and attention-deficit/hyperactivity disorder: distinct associations with DNA methylation and genetic variation. *J. Neurodev. Disord.* 12:23. doi: 10.1186/s11689-020-09324-3
- Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., et al. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat. Hum. Behav.* 3, 513–525. doi: 10.1038/s41562-019-0566-x
- Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., et al. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* 51, 431–444. doi: 10.1038/s41588-019-0344-8
- Guo, W., Samuels, J. F., Wang, Y., Cao, H., Ritter, M., Nestadt, P. S., et al. (2017). Polygenic risk score and heritability estimates reveals a genetic relationship between ASD and OCD. *Eur. Neuropsychopharmacol.* 27, 657–666. doi: 10.1016/j.euroneuro.2017.03.011
- Hasler, G., LaSalle-Ricci, V. H., Ronquillo, J. G., Crawley, S. A., Cochran, L. W., Kazuba, D., et al. (2005). Obsessive-compulsive disorder symptom dimensions show specific relationships to psychiatric comorbidity. *Psychiatry Res.* 135, 121–132. doi: 10.1016/j.psychres.2005.03.003
- Hemmings, S. M., Kinnear, C. J., Lochner, C., Niehaus, D. J., Knowles, J. A., Moolman-Smook, J. C., et al. (2004). Early- versus late-onset obsessive-compulsive disorder: investigating genetic and clinical correlates. *Psychiatry Res.* 128, 175–182. doi: 10.1016/j.psychres.2004.05.007
- Hirschtritt, M. E., Darrow, S. M., Illmann, C., Osiecki, L., Grados, M., Sandor, P., et al. (2018). Genetic and phenotypic overlap of specific obsessive-compulsive and attention-deficit/hyperactive subtypes with Tourette syndrome. *Psychol. Med.* 48, 279–293. doi: 10.1017/S0033291717001672
- International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC) and OCD Collaborative Genetics Association Studies and OCGAS (2017). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Mol. Psychiatry* 23, 1181–1188. doi: 10.1038/mp.2017.154
- Jansen, P. R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A. R., et al. (2019). Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat. Genet.* 51, 394–403. doi: 10.1038/s41588-018-0333-3
- Khrantsova, E. A., Heldman, R., Derks, E. M., Yu, D., Davis, L. K., and Stranger, B. E. (2019). Sex differences in the genetic architecture of obsessive-compulsive disorder. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 180, 351–364. doi: 10.1002/ajmg.b.32687
- Kulminski, A. M., Loika, Y., Culminskaya, I., Arbeeve, K. G., Ukraintseva, S. V., Stallard, E., et al. (2016). Explicating heterogeneity of complex traits has strong potential for improving GWAS efficiency. *Sci. Rep.* 6, 1–8. doi: 10.1038/srep35390
- Lam, M., Awasthi, S., Watson, H. J., Goldstein, J., Panagiotaropoulou, G., Trubetskoy, V., et al. (2020). RICOPIIL: Rapid Imputation for COnsortias PIpeLine. *Bioinformatics* 36, 930–933. doi: 10.1093/bioinformatics/btz633
- Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., et al. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* 50, 1112–1121. doi: 10.1038/s41588-018-0147-3
- Lee, P. H., Anttila, V., Won, H., Feng, Y. C. A., Rosenthal, J., Zhu, Z., et al. (2019). Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell* 179, 1469–1482. doi: 10.1016/j.cell.2019.11.020
- Lee, S. H., Wray, N. R., Goddard, M. E., and Visscher, P. M. (2011). Estimating missing heritability for disease from genome-wide association studies. *Am. J. Hum. Genet.* 88, 294–305. doi: 10.1016/j.ajhg.2011.02.002
- Lochner, C., Fineberg, N. A., Zohar, J., Van Ameringen, M., Juven-Wetzler, A., Altamura, A. C., et al. (2014). Comorbidity in obsessive-compulsive disorder (OCD): a report from the International College of Obsessive-Compulsive Spectrum Disorders (ICOCS). *Compr. Psychiatry* 55, 1513–1519. doi: 10.1016/j.comppsy.2014.05.020
- MacRae, C. A., and Vasan, R. S. (2011). Next-generation genome-wide association studies time to focus on phenotype? *Circ. Cardiovasc. Genet.* 4, 334–336. doi: 10.1161/CIRCGENETICS.111.960765
- Mataix-Cols, D., Conceição do Rosario-Campos, M. D., and Leckman, J. F. (2005). A multidimensional model of obsessive-compulsive disorder. *Am. J. Psychiatry* 162, 228–238. doi: 10.1176/appi.ajp.162.2.228
- Mattina, G. F., and Steiner, M. (2016). The need for inclusion of sex and age of onset variables in genetic association studies of obsessive-compulsive disorder: overview. *Prog. Neuropsychopharmacol. Biol. Psychiatry* 67, 107–116. doi: 10.1016/j.pnpbp.2016.01.012
- Meier, S. M., Petersen, L., Schendel, D. E., Mattheisen, M., Mortensen, P. B., and Mors, O. (2015). Obsessive-compulsive disorder and autism spectrum disorders: longitudinal and offspring risk. *PLoS ONE* 10:e141703. doi: 10.1371/journal.pone.0141703
- Murray, C. J. L., Lopez, A. D., World Health Organization, World Bank, and Harvard School of Public Health (1996). *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries and Risk Factors in 1990 and Projected to 2020*, eds C. J. L. Murray and A. D. Lopez. World Health Organization. Available online at: <https://apps.who.int/iris/handle/10665/41864>
- Nagel, M., Watanabe, K., Stringer, S., Posthuma, D., and Van Der Sluis, S. (2018). Item-level analyses reveal genetic heterogeneity in neuroticism. *Nat. Commun.* 9:905. doi: 10.1038/s41467-018-03242-8
- Nestadt, G., Di, C. Z., Riddle, M. A., Grados, M. A., Greenberg, B. D., Fyer, A. J., et al. (2009). Obsessive-compulsive disorder: subclassification based on co-morbidity. *Psychol. Med.* 39, 1491–1501. doi: 10.1017/S0033291708004753
- O'Connell, K. S., McGregor, N. W., Lochner, C., Emsley, R., and Warnich, L. (2018). The genetic architecture of schizophrenia, bipolar disorder, obsessive-compulsive disorder and autism spectrum disorder. *Mol. Cell. Neurosci.* 88, 300–307. doi: 10.1016/j.mcn.2018.02.010
- Ortiz, A. E., Morer, A., Moreno, E., Plana, M. T., Cordovilla, C., and Lázaro, L. (2016). Clinical significance of psychiatric comorbidity in children and adolescents with obsessive-compulsive disorder: subtyping a complex disorder. *Eur. Arch. Psychiatry Clin. Neurosci.* 266, 199–208. doi: 10.1007/s00406-015-0642-9
- Pallanti, S., Grassi, G., Sarrecchia, E. D., Cantisani, A., and Pellegrini, M. (2011). Obsessive-compulsive disorder comorbidity: clinical assessment and therapeutic implications. *Front. Psychiatry* 2:70. doi: 10.3389/fpsy.2011.00070
- Pedersen, C. B., Bybjerg-Grauholm, J., Pedersen, M. G., Grove, J., Agerbo, E., Bækvad-Hansen, M., et al. (2018). The iPSYCH2012 case-cohort sample: new directions for unravelling genetic and environmental architectures of severe mental disorders. *Mol. Psychiatry* 23, 6–14. doi: 10.1038/mp.2017.196
- Peyrot, W. J., and Price, A. L. (2021). Identifying loci with different allele frequencies among cases of eight psychiatric disorders using CC-GWAS. *Nat. Genet.* 53, 445–454. doi: 10.1038/s41588-021-00787-1
- Sampaio, A., Pacheco Lins, R., Daltro-Oliveira, R., De Castro Quarantini, L., Conceicao Do Rosario, M., Euripedes, C., et al. (2013). Genetic association studies in obsessive-compulsive disorder. *Rev. Psiq. Clin. Brazil* 40, 177–190. doi: 10.1590/S0101-60832013000500003
- Speranza, M., Corcos, M., Godart, N., Loas, G., Guilbaud, O., Jeammet, P., et al. (2001). Obsessive compulsive disorders in eating disorders. *Eat Behav.* 2, 193–207. doi: 10.1016/S1471-0153(01)00035-6
- Stahl, E. A., Breen, G., Forstner, A. J., McQuillin, A., Ripke, S., Trubetskoy, V., et al. (2019). Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nat. Genet.* 51, 793–803. doi: 10.1038/s41588-019-0397-8

- Taylor, S. (2011). Early versus late onset obsessive-compulsive disorder: evidence for distinct subtypes. *Clin. Psychol. Rev.* 31, 1083–1100. doi: 10.1016/j.cpr.2011.06.007
- Taylor, S. (2013). Molecular genetics of obsessive-compulsive disorder: A comprehensive meta-analysis of genetic association studies. *Mol. Psychiatry* 18, 799–805. doi: 10.1038/mp.2012.76
- The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature* 526, 68–74. doi: 10.1038/nature15393
- The Brainstorm Consortium, Anttila, V., Bulik-Sullivan, B., Finucane, H. K., Walters, R. K., Bras, J., et al. (2018). Analysis of shared heritability in common disorders of the brain. *Science* 360:1313. doi: 10.1126/science.aap8757
- Tükel, R., Polat, A., Özdemir, z., Aksüt, D., and Türksöy, N. (2002). Comorbid conditions in obsessive-compulsive disorder. *Compr. Psychiatry* 43, 204–209. doi: 10.1053/comp.2002.32355
- U.S. International Institutes of Health (NIH) (2016). *What is Obsessive-Compulsive Disorder (OCD)?* U.S. International Institutes of Health (NIH).
- Van Roode, T., Sharples, K., Dickson, N., and Paul, C. (2017). Life-Course relationship between socioeconomic circumstances and timing of first birth in a birth cohort. *PLoS ONE* 12:e170170. doi: 10.1371/journal.pone.0170170
- Walitza, S., Wendland, J. R., Gruenblatt, E., Warnke, A., Sontag, T. A., Tucha, O., et al. (2010). Genetics of early-onset obsessive-compulsive disorder. *Eur. Child Adolesc. Psychiatry* 19, 227–235. doi: 10.1007/s00787-010-0087-7
- Watanabe, K., Taskesen, E., Van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* 8, 1–10. doi: 10.1038/s41467-017-01261-5
- Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R., Gaspar, H. A., et al. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat. Genet.* 51, 1207–1214. doi: 10.1038/s41588-019-0439-2
- Weissman, M. M. (1998). Cross-national epidemiology of obsessive-compulsive disorder. *CNS Spectr.* 3, 18–21. doi: 10.1017/S1092852900007136
- Wray, N., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., et al. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* 50, 668–681.
- Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi: 10.1016/j.ajhg.2010.11.011
- Yengo, L., Sidorenko, J., Kemper, K. E., Zheng, Z., Wood, A. R., Weedon, M. N., et al. (2018). Meta-analysis of genome-wide association studies for height and body mass index in 700 000 individuals of European ancestry. *Hum. Mol. Genet.* 27, 3641–3649. doi: 10.1093/hmg/ddy271
- Yilmaz, Z., Halvorsen, M., Bryois, J., Yu, D., Thornton, L. M., Zerwas, S., et al. (2020). Examination of the shared genetic basis of anorexia nervosa and obsessive-compulsive disorder. *Mol. Psychiatry* 25, 2036–2046. doi: 10.1038/s41380-018-0115-4
- Zheng, J., Erzurumluoglu, A. M., Elsworth, B. L., Kemp, J. P., Howe, L., Haycock, P. C., et al. (2017). LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics* 33, 272–279. doi: 10.1093/bioinformatics/btw613

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer FT declared a past co-authorship with the authors AB, JC, MM to the handling editor.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Strom, Grove, Meier, Bækvad-Hansen, Becker Nissen, Damm Als, Halvorsen, Nordentoft, Mortensen, Hougaard, Werge, Mors, Børglum, Crowley, Bybjerg-Grauholm and Mattheisen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.