Archival Report

Biological Psychiatry

Boosting Schizophrenia Genetics by Utilizing Genetic Overlap With Brain Morphology

Dennis van der Meer, Alexey A. Shadrin, Kevin O'Connell, Francesco Bettella, Srdjan Djurovic, Thomas Wolfers, Dag Alnæs, Ingrid Agartz, Olav B. Smeland, Ingrid Melle, Jennifer Monereo Sánchez, David E.J. Linden, Anders M. Dale, Lars T. Westlye, Ole A. Andreassen, Oleksandr Frei, and Tobias Kaufmann

ABSTRACT

BACKGROUND: Schizophrenia is a complex polygenic disorder with subtle, distributed abnormalities in brain morphology. There are indications of shared genetic architecture between schizophrenia and brain measures despite low genetic correlations. Through the use of analytical methods that allow for mixed directions of effects, this overlap may be leveraged to improve our understanding of underlying mechanisms of schizophrenia and enrich polygenic risk prediction outcome.

METHODS: We ran a multivariate genome-wide analysis of 175 brain morphology measures using data from 33,735 participants of the UK Biobank and analyzed the results in a conditional false discovery rate together with schizophrenia genome-wide association study summary statistics of the Psychiatric Genomics Consortium (PGC) Wave 3. We subsequently created a pleiotropy-enriched polygenic score based on the loci identified through the conditional false discovery rate approach and used this to predict schizophrenia in a nonoverlapping sample of 743 individuals with schizophrenia and 1074 healthy controls.

RESULTS: We found that 20% of the loci and 50% of the genes significantly associated with schizophrenia were also associated with brain morphology. The conditional false discovery rate analysis identified 428 loci, including 267 novel loci, significantly associated with brain-linked schizophrenia risk, with functional annotation indicating high relevance for brain tissue. The pleiotropy-enriched polygenic score explained more variance in liability than conventional polygenic scores across several scenarios.

CONCLUSIONS: Our results indicate strong genetic overlap between schizophrenia and brain morphology with mixed directions of effect. The results also illustrate the potential of exploiting polygenetic overlap between brain morphology and mental disorders to boost discovery of brain tissue–specific genetic variants and its use in polygenic risk frameworks.

https://doi.org/10.1016/j.biopsych.2021.12.007

Schizophrenia is a highly heritable complex brain disorder with a polygenic architecture, involving numerous common polymorphisms with small effects (1). While genome-wide association studies (GWASs) have identified hundreds of genetic variants associated with schizophrenia, together these explain only a fraction of its heritability (2). Identifying more variants will increase our understanding of the underlying disease mechanisms and improve genetic prediction, with potential for clinical utility. However, this is currently not feasible, as uncovering a significant portion of the heritability with conventional approaches will require sample sizes of more than a million individuals (3).

The complexity of the genetic architecture of schizophrenia has made it difficult to determine a genetic relationship with abnormal brain morphology phenotypes associated with the disorder (4). Patients with schizophrenia have on average a thinner cerebral cortex compared with healthy peers as well as reductions in cortical surface area (5). Subcortically, schizophrenia is associated with smaller hippocampal and amygdala volumes, larger lateral ventricles (6), and smaller cerebellum (7). In line with the clinical and genetic heterogeneity of schizophrenia, imaging studies have documented substantial brain structural heterogeneity (8), and specific regional deviations are shared by a only small percentage of individuals (9).

It is possible to boost the power of schizophrenia genetic studies by leveraging auxiliary genetic information contained in a related trait (10), such as brain morphology. Surprisingly, large-scale investigations have shown no or low genetic correlations between schizophrenia and brain structure (4,11). However, substantial genetic overlap may remain undetected by standard measures of genetic correlation owing to mixed directions of effects of shared genetic variants across the two traits, canceling each other out (12). Conditional false

SEE COMMENTARY ON PAGE 258

ISSN: 0006-3223

Biological Psychiatry August 15, 2022; 92:291–298 www.sobp.org/journal

discovery rate (cFDR) analysis uses overlapping variant associations, regardless of direction of effects, to re-rank the test statistics in a primary phenotype conditional on the associations in a secondary phenotype (10,13). The idea behind this is that in the presence of cross-trait enrichment, a variant with strong associations with both traits is more likely to represent a true association. This approach has been shown to allow for discovery of many novel variants (13), which have been found to replicate well (14,15). Further, owing to the distributed effects of genetic variants involved in brain morphology, the data can be analyzed more powerfully through a multivariate approach, capitalizing on the shared genetic signal across regional brain measures by considering the brain as an integrated unit (16), thereby providing greater statistical power and better alignment with the underlying biology than univariate approaches.

Here we characterized the polygenic overlap between schizophrenia and brain morphology, using a comprehensive set of brain morphology measures combined in a multivariate framework, and investigated whether this overlap can be leveraged to allow for discovery of common genetic variants contributing to brain-linked schizophrenia risk. We further sought to improve the polygenic prediction of schizophrenia by leveraging the newly identified genetic overlap.

METHODS AND MATERIALS

Participants

We used data from the UK Biobank (UKB) population cohort, under accession number 27412. The composition, set-up, and data-gathering protocols of UKB have been described elsewhere (17). After preprocessing, our sample size for the primary neuroimaging analyses, excluding individuals with schizophrenia, was 33,735, with a mean (SD) age of 64.33 (7.49) years, and 52.02% was female. For replication of the polygenic score (PGS) findings, we used 576 individuals with a schizophrenia diagnosis (mean [SD] age 55.76 [8.39] years, 35.60% female) and 317,139 individuals without any brain disorder diagnosis (mean age 57.43 [8.04] years, 54.90% female) not part of the neuroimaging sample.

We further used the TOP (Thematically Organized Psychosis) clinical cohort. For the polygenic scoring analyses, we had complete genetic data available for 743 individuals with schizophrenia (mean age 32.76 [13.29] years, 42.80% female) and 1074 healthy individuals (mean age 32.49 [10.00] years, 47.39% female). For association of univariate brain measures with diagnosis, we had available neuroimaging data of 457 individuals with schizophrenia (mean age 30.54 [9.47] years, 41.79% female) and 998 healthy individuals (mean age 33.33 [10.26] years, 45.49% female). Each sample was collected with the participants' written informed consent and with approval by local institutional review boards.

Genetic Data Preprocessing

We used UKB v3 imputed data, quality controlled as described by the UKB genetics team (18). We carried out standard quality check procedures, including removing individuals with more than 10% missingness, lead single nucleotide polymorphisms (SNPs) with more than 5% missingness, with an INFO score below 0.8, and failing the Hardy-Weinberg equilibrium test at $p = 1 \times 10^{-9}$. We set a minor allele frequency (MAF) threshold of 0.005, leaving 9,061,072 SNPs. This MAF was chosen in accordance with previous imaging genetics work with UKB data by our group (16). Extensive information on the TOP genetic processing is given in Supplement 1.

Image Acquisition

For the genetic analyses of neuroimaging data, we used UKB magnetic resonance imaging data released up to March 2020. T1-weighted scans were collected from four scanning sites throughout the United Kingdom, all on identically configured Siemens MAGNETOM Skyra 3T scanners (Siemens AG). The UKB core neuroimaging team has published extensive information on the applied procedures (19).

For the association of brain morphology with schizophrenia diagnosis, we used T1-weighted data from TOP, collected at the Oslo University Hospital on three scanners: a 1.5T Siemens MAGNETOM Sonata (Siemens AG), a 3.0T Signa HDxt (GE Healthcare), and a 3.0T GE 750 (GE Healthcare). For details on scanning protocols, see Brandt *et al.* (20) and Kaufman *et al.* (21).

Neuroimaging Data Preprocessing

We applied the standard recon-all-all processing pipeline of FreeSurfer v5.3, performing automated surface-based morphometry and subcortical segmentation (22,23). From the output, we extracted 175 global, subcortical, and cortical morphology measures. Table S1 in Supplement 1 contains all the measures included. We included both the left and the right hemisphere measures, if applicable.

For the imaging GWAS of UKB data, we first selected all individuals with White European ancestry, based on self-identification as White British and similar genetic ancestry. A cross-ethnic analysis including 5737 non-White participants and yielding highly similar results is provided in Supplement 1.

We excluded anyone with an ICD-10 schizophrenia diagnosis, indicated by an F2 code; individuals with bad structural scan quality, indicated by an age- and sex-adjusted Euler number (24) 3 standard deviations below the scanner site mean; or individuals with a global brain measure 5 standard deviations from the sample mean. Finally, we removed one of each genetically related pair of individuals, defined by a threshold of 0.0625 determined by genome-wide complex trait analysis (25).

We followed a similar preprocessing procedure for associating brain measures with schizophrenia in TOP. White European ethnicity was based on self-report. DSM-IV diagnosis of schizophrenia was determined based on the Structured Clinical Interview for DSM-IV Axis I disorders. Controls were individuals without brain damage or history of a severe psychiatric disorder themselves or first-degree relatives.

For both samples, separately, we regressed out age, sex, scanner site, Euler number, and the first 20 genetic principal components from each brain measure. We further regressed out estimated intracranial volume for subcortical volumes, mean thickness for regional thickness measures, and total surface area for regional surface area measures. We then

applied rank-based inverse normal transformation (26) to the residuals.

Genetic Analyses

A univariate GWAS on each of the 175 processed brain morphology measures was carried out using the additive model in PLINK2 (27). We ran a multivariate GWAS on all pre-residualized brain morphology measures using the Multivariate Omnibus Statistical Test (MOSTest). MOSTest yields one set of summary statistics that captures the significance of its association across all measures for each SNP. See van der Meer *et al.* (16) for an extensive description of MOSTest, its validation, and application to brain morphology data. The MOSTest software is openly available via https://github.com/precimed/mostest.

We further ran a GWAS on nonmelanoma skin cancer for use as a negative control, given that this is a heritable trait (28) with no significant association with psychiatric disorders (29). We made use of the UKB data, categorizing all individuals with an ICD-10 code C44 as cases, restricted to White Europeans without brain disorders. We ran a logistic regression with PLINK2, covarying for age, sex, and 20 genetic principal components.

We used summary statistics from the Psychiatric Genomics Consortium Wave 3 (PGC3) schizophrenia GWAS (30) to check for genetic overlap between brain morphology and schizophrenia. We selected meta-analyzed summary statistics excluding the TOP sample, preventing sample overlap, containing 13,025,668 SNPs for 50,965 individuals with schizophrenia and 68,049 controls.

We conducted cFDR analysis, conditioning the PGC3 schizophrenia GWAS on the brain morphology GWAS, through the pleioFDR tool using default settings (https://github.com/precimed/pleiofdr). PleioFDR builds on statistical pleiotropy, defined as the presence of associations with both traits (31). We set an FDR threshold of 0.01 as whole-genome significance, in accordance with recommendations.

Independent significant SNPs and genomic loci were identified from the resulting GWAS summary statistics in accordance with the Functional Mapping and Annotation (FUMA) SNP2GENE definition (32). We also used FUMA (https://fuma.ctglab.nl/) to map SNPs to genes, using default settings. Given that cFDR output is incompatible with FUMA, we first identified lead SNPs according to FUMA definitions by clumping these summary statistics at an FDR threshold of 0.01 (https://github.com/precimed/python_convert) before using the FUMAs predefined lead SNPs option.

We carried out gene-based analyses through MAGMA v1.08 for the GWASs of brain morphology and schizophrenia (not for cFDR, as this is not valid), with a SNP-wide mean model and the 1000 Genomes Phase 3 EUR reference panel.

Hypergeometric tests were applied through FUMA GENE2FUNC function to calculate overlap between the sets of mapped genes and differentially expressed genes in each of the 54 tissues available in the Genotype-Tissue Expression (GTEx) v8 database. Differentially expressed genes were defined as genes with log2 transformed, normalized expression values (reads per kilobase million, zero mean) with p value $\leq$ .05 after Bonferroni correction and absolute log fold change $\geq$0.58 in a given tissue, the default set by FUMA.

Polygenic Scoring

We calculated PGSs through PRSice v2 (33). We first clumped each of the summary statistics (skin cancer, brain morphology, schizophrenia, and cFDR) through PLINK, using an MAF threshold of 0.05, p-value threshold of 1, linkage disequilibrium threshold of 0.1, and distance threshold of 10 mb. We then ran PRSice, calculating PGSs for the individuals in TOP (and UKB replication sample) based on the PGC3 summary statistics, using the most significant lead SNPs from each of the four clumped summary statistics as determined by the indicated PGS threshold.

We used logistic regression to calculate the log odds and Nagelkerke R^2 of the PGSs at each threshold, with schizophrenia diagnosis as outcome in TOP and UKB separately. We included age, sex, batch, and the first 20 genetic principal components as covariates. See Supplement 1 for robustness analyses, whereby we created PGSs based on different GWASs to predict schizophrenia and intelligence.

Statistical Analyses

All downstream analyses were run in R v3.6.1 (R Foundation for Statistical Computing). Graphs were created using ggplot2 (34), and brain maps were created through pysurfer https://pysurfer.github.io/. Coding applied for this study and an overview of the steps involved in the novel polygenic scoring approach, pleiotropy-enriched PGS (pleioPGS), are available at https://github.com/norment/open-science.

RESULTS

There was considerable overlap between the loci and genes identified in the PGC3 schizophrenia GWAS and those discovered through the multivariate analysis of 175 brain regions. There were 571 independent unique genetic loci associated with brain morphology at $p < 5 \times 10^{-8}$. Of these, 32 loci were also among the 164 genome-wide significant loci (i.e., 19.5%) associated with schizophrenia (Figure 1A). Through MAGMA (32,35), we found 1586 genes significantly associated with brain morphology after multiple comparisons correction, of which 252 were also among the 508 genes (i.e., 49.6%) identified through the schizophrenia GWAS. Supplements 2 and 3 list all discovered loci and genes.

Conditional FDR

The cFDR analysis, conditioning the schizophrenia GWAS on the multivariate brain morphology GWAS, further indicated genetic overlap (Figure 1B). Through cFDR, we identified 428 independent significant loci, of which 267 were novel compared with the schizophrenia GWAS (Figure S1 in Supplement 1). Supplement 2 lists information on all identified loci.

As follow-up of the main analyses using multivariate data, we checked for significant genetic overlap between schizophrenia and the individual univariate brain measures. The number of cFDR-discovered loci were between 235 and 315 per region (median = 270), while genetic correlations were between −0.11 and 0.11 (median = −0.01, maximum

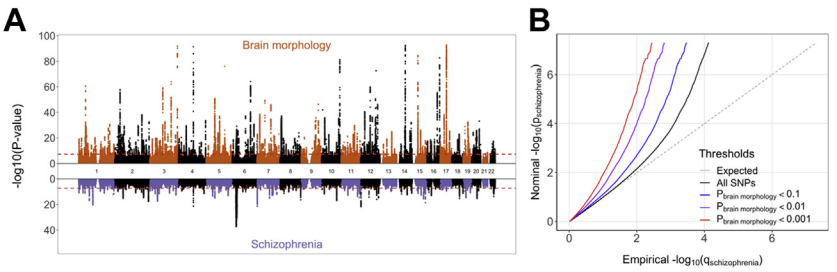

Figure 1. Identification of genetic overlap between schizophrenia and brain morphology. **(A)** Miami plot, depicting the $-\log_{10}$ (*p* values) on the y-axis, of each genetic variant by their genomic position per chromosome on the x-axis, as identified through the multivariate brain morphology GWAS (top half) and schizophrenia GWAS (bottom half). The red dashed lines indicate the genome-wide significance threshold of 5×10^{-8}. Note that the y-axis is clipped at $-\log_{10}$ (*p* value) = 100. **(B)** Quantile-quantile plot conditioning the schizophrenia GWAS on the multivariate brain morphology GWAS. Successive leftward deflection from the null distribution (dotted line) of the observed *p*-value distributions at more stringent significance thresholds (color-coded lines) indicates genetic overlap between the two traits. GWAS, genome-wide association study; SNPs, single nucleotide polymorphisms.

p value 1.2×10^{-4}). The full results are summarized in Supplement 5.

Functional Annotation

There were 50 lead SNPs with a Combined Annotation-Dependent Depletion score above 12.37, categorizing them as deleterious (36), versus 20 of the schizophrenia GWAS lead SNPs. There were no significant differences in distributions of minimum chromatin state scores or RegulomeDB scores (37), indicating that the SNPs identified through either cFDR or the schizophrenia GWAS were equally likely to be located in regulatory regions.

Gene Mapping

The lead SNPs identified by the cFDR analysis were mapped onto 1114 genes, of which 645 genes were not mapped through the schizophrenia GWAS. Among the top novel genes identified were *RORA*, encoding a nuclear receptor essential for development of the cerebellum (38) and for regulating circadian rhythm and immune function (39), hailed as a strong potential drug target (40); *BCL11A*, a transcription factor important for cortical development by regulating axon outgrowth (41), coupled to intellectual disability (42); and *ALG14*, encoding a catalyst of glycosylation, mutation of which is known to cause early, severe neurodegeneration (43) and behavioral problems (44). Information on each mapped gene is listed in Supplement 4.

Further, the set of genes mapped through the cFDR loci was significantly enriched for genes with known differential expression in the brain, particularly in the cerebellum and frontal cortex, but not in nonbrain tissues. This contrasts with the genes identified through the brain morphology GWAS ($n = 2245$) and schizophrenia GWAS ($n = 660$), which showed less specificity of expression in brain tissue (Figure 2). As we have previously shown that the brain morphology GWAS does produce highly significant associations with genetic pathways central for brain development (16), we speculate that these findings speak to the nature of this test, capturing specificity, which is lost for the multivariate brain morphology data owing to the large number of genes identified.

The cFDR-identified genes were significantly overrepresented in Gene Ontology (GO) nervous system

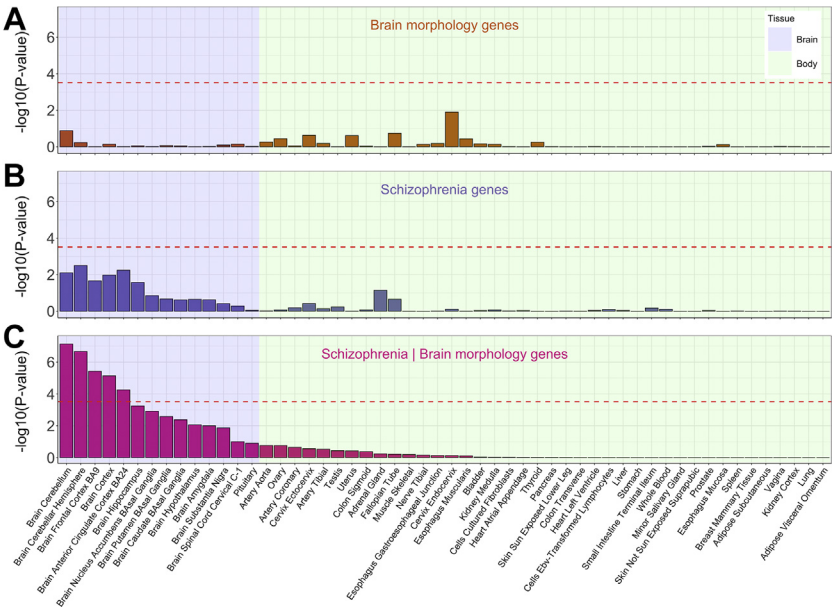

Figure 2. Brain tissue–specific expression of the mapped genes. **(A)** Results for the genes mapped through the multivariate brain morphology GWAS. **(B)** The schizophrenia GWAS and **(C)** conditioning the schizophrenia GWAS on the multivariate brain morphology GWAS. The y-axis shows the $-\log_{10}$ (*p* values) of the hypergeometric tests, testing for enrichment of the identified genes among gene sets with tissue-specific gene expression based on the GTEx v8 database. The x-axis shows each of the 54 tissues, split by brain and body tissues, and sorted within these categories by the significance of the pleiotropy-informed gene set enrichment. The horizontal dashed red line indicates the multiple comparisons–corrected significance threshold, at $p = .05/(54 \text{ tissues} \times 3 \text{ analyses})$. BA, Brodmann area; GWAS, genome-wide association study.

development gene sets and previously identified through GWASs of bipolar disorder, neuroticism, cognitive ability, and intelligence. This was also the case when rerunning these analyses on only the 645 novel genes (Figures S3 and S4 in Supplement 1).

We additionally compared our gene list with two studies coupling transcriptomics data to GWAS summary statistics to boost identification of schizophrenia genes not linked to the PGC set of significant loci. Of the non-PGC identified genes, 16 were among the 62 additional genes reported by Gusev *et al.* (45), 43 were overlapping with the 162 additional genes identified by Walker *et al.* (46), and 7 were overlapping between all three.

Polygenic Scoring

We hypothesized that a pleiotropy-informed approach, identifying SNPs that influence schizophrenia conditioned on brain morphology in independent samples, can reduce noise in the creation of PGSs owing to locus selection being based on two traits, lowering the likelihood of chance findings. To test this, we created PGSs to predict schizophrenia diagnosis for participants of the TOP clinical cohort. We first selected the most significant lead SNPs identified by 1) the multivariate brain morphology GWAS; 2) the schizophrenia GWAS; 3) the cFDR (schizophrenia/multivariate brain morphology) analysis; or 4) nonmelanoma skin cancer, as a control measure. We then calculated the PGSs for each of these sets of lead SNPs, based on the regression coefficients from the schizophrenia GWAS summary statistics (excluding the TOP cohort). We calculated the scores for a number of lead SNPs rather than for significance thresholds to allow for direct comparison between the approaches, at equal numbers of SNPs in each set. Here, the 200 SNPs threshold corresponds approximately to whole-genome significance, and the 30,000 SNPs threshold corresponds approximately to the number of lead SNPs identified in the schizophrenia GWAS at nominal p = .05, the threshold commonly applied for schizophrenia PGSs (30). As shown in Figure 3, at these thresholds, the cFDR-based pleioPGSs showed best prediction. Expressed as variance in liability, pleioPGSs explained 0.013 with the top 200 SNPs versus 0.008 for regular PGSs. With the top 30,000 SNPs, this was 0.077 versus 0.069. Figure S5 in Supplement 1 shows the explained variance in schizophrenia diagnosis per 1000 lead SNPs for each locus selection strategy through a sliding window approach. This graph illustrates that the variance explained by each subsequent window goes down strongest for the pleioPGS and the standard schizophrenia-based PGS, while there is no difference between the first and the last window of SNPs based on the skin cancer GWAS, indicating the lack of information captured by selecting loci based on this trait.

We subsequently sought to determine the robustness of these findings by rerunning the PGS analyses with different settings. In each case, the pleioPGS approach performed as good as or better than the conventional PGS approach. We first tested the use of a different replication sample: when calculating scores on 576 UKB participants with a schizophrenia diagnosis versus 317,139 individuals without brain disorders, the pleioPGS explained 0.022 (top 200 SNPs) and 0.063 (top 30,000 SNPs) variance in liability compared with 0.012 and 0.059 for regular PGSs. Second, we used an earlier subset of the schizophrenia GWAS to generate the PGSs, namely, the smaller (and thus likely noisier) PGC2 schizophrenia GWAS (2), with the TOP cohort excluded. With this, the explained liability in TOP was 0.017 and 0.057 for pleioPGS versus 0.009 and 0.050 for regular PGSs. Third, we checked how the approach generalizes to a different outcome measure: we first ran a GWAS of fluid intelligence on 100,000 UKB participants, followed by a cFDR analysis whereby we conditioned this GWAS on the multivariate brain morphology GWAS. We then created PGSs based on the resulting summary statistics and predicted intelligence in a random holdout set of 18,515 participants. Here, the pleioPGS approach explained 0.023 and 0.059 variance at the two thresholds versus 0.020 and 0.055 for regular PGSs. Therefore, with each of these three analyses, we saw that pleioPGS performed better at lower thresholds and that it reached a plateau earlier, indicating that it may allow for improved prioritization of genetic variants to study and predict brain traits (see Supplement 1 for more details and the full results).

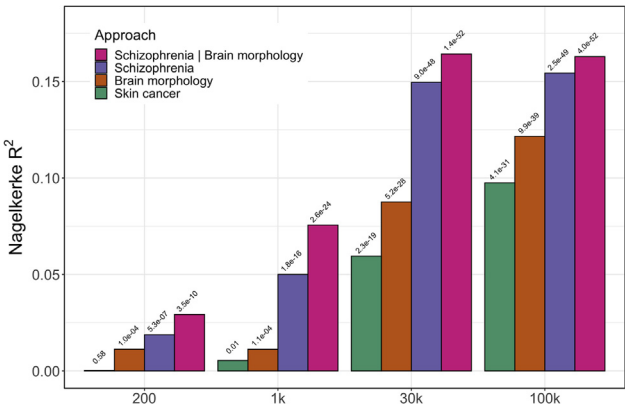

Figure 3. Prediction of schizophrenia diagnosis by polygenic scores in the TOP sample. The x-axis indicates the number of most significant single nucleotide polymorphisms that went into the polygenic scores, based on the results from the skin cancer GWAS (green bars), brain morphology GWAS (orange bars), schizophrenia GWAS (purple bars), or conditional false discovery rate analysis conditioning the schizophrenia GWAS on the multivariate brain morphology GWAS (pink bars). The y-axis indicates Nagelkerke R^2. The p values of the association between the polygenic scores and schizophrenia diagnosis are listed above the bars. GWAS, genome-wide association study; TOP, Thematically Organized Psychosis.

DISCUSSION

With the present study, we provide evidence of polygenic overlap between schizophrenia and brain morphology that has not been detected with standard genetic approaches (4,11). We showed that this shared genetic architecture may be used to discover novel genetic variants that may play a role in schizophrenia, identifying 267 loci not discovered by the PGC and implicating schizophrenia-associated genes with high relevance for brain morphology. Lastly, we found that PGSs based on the shared genetic architecture between schizophrenia and brain magnetic resonance imaging morphology can be used for predicting schizophrenia.

The present identification of genetic overlap between schizophrenia and brain morphology probably results from the application of analytical approaches more in line with the complex nature of the polygenic architecture of the brain and brain-related disorders such as schizophrenia. In particular, the distributed nature of genetic signal relevant to schizophrenia across the brain supports the use of multivariate methods (16), and the mixed directions of effects involved in the genetic overlap support the use of methods that do not rely on globally consistent directionality of effects (13,16). The high polygenicity of nearly all clinically relevant traits suggest that there are a large number of neurobiological mechanisms involved (3), and there is often large genetic overlap between many pairs of traits, while there is a relatively low genetic correlation (12). This indicates that some of these mechanisms strengthen both traits, while others inhibit the expression of one trait and strengthen the other. Given this, the approaches put forth here are able to better characterize the complex genetic architecture beyond genetic correlation, providing a more accurate model of the genetic relationship between brain-related traits and disorders (1), essential for our understanding of the etiology of schizophrenia and other mental disorders.

The high yield in terms of genome-wide significant loci from the cFDR analyses further support the value of considering genetic overlap between schizophrenia and brain morphology, providing important leads for understanding the neurobiological underpinnings of this disorder. Through the cFDR approach, we more than doubled the number of loci identified compared with the largest schizophrenia GWAS available to date (30). Conditioning a schizophrenia GWAS on a multivariate brain morphology GWAS allows for the identification of genetic determinants, particularly of brain-linked schizophrenia risk. This is further supported by functional annotation of these cFDR loci, with the mapped genes being linked to brain tissue–specific gene expression. A subset of the genes not reported by the PGC were also found in studies identifying schizophrenia risk genes on the basis of brain transcriptome data (45,46), thereby increasing the confidence in these findings through converging evidence. Further experimental interrogation of the identified loci may uncover biological insights that can lead to better understanding of the disease mechanisms and inform the development of novel treatment regimens. In addition, the specificity attained by our current approach may be valuable to parse schizophrenia genetics. Here we applied a brain morphology GWAS to boost the identification of nervous system genes associated with schizophrenia; we expect that other novel genes may be found when conditioning the schizophrenia GWAS on other system-specific GWAS, such as immune-related disorders or traits.

We further found that a pleioPGS has the potential to improve PGSs. We note that this was specifically the case for more stringent significance thresholds, in line with the reasoning that this approach brings about greater specificity, being particularly relevant for identifying the subset of loci that are most relevant for brain-linked schizophrenia risk. This approach appears robust, as the improved performance is also seen when using a subset of the main schizophrenia GWAS, i.e., less statistical power, as well as when using another independent schizophrenia test sample. We also found a similar pattern of results when analyzing another brain-related trait, fluid intelligence. This suggests that the concept underlying the pleioPGS may be widely applicable. This approach is based on the selection of loci identified by conditioning the genetics of one trait on that of another trait, leading to better out-of-sample prediction, possibly by reducing noise in GWAS signal and identifying more biologically relevant genetic variants. This is conceptually similar to a previous report that PGSs based on genetic variants specific to placental tissue are more predictive of schizophrenia than PGSs based on nonspecific variants (47). We are currently following up on this approach with the investigation of more pairs of traits.

It should be noted that the presence of statistical pleiotropy leveraged in this study to improve gene discovery and prediction can be partly driven by patterns of linkage disequilibrium and MAF (31). This warrants replication of the findings in new samples in the future. We also note the discrepancy in demographics between the elderly population–based UKB and younger clinical TOP sample, which may have contributed to larger error in the polygenic scoring analyses. Further, experimental studies are needed to clarify the molecular mechanisms underlying the current statistical associations. Last, we note that our aim was to illustrate the broad concept of leveraging genetic overlap to aid discovery and prediction. We thereby recognize that we did not explore other tools to achieve this, as that falls outside the scope of the current work, leaving that for future studies. Regardless of these limitations, our approach may be of particular clinical interest in the context of comorbidity; by conditioning a disease on comorbid traits or disorders, we may be able to improve the prediction of clinical subgroups.

To conclude, we provide evidence in this study for polygenic overlap between schizophrenia and brain morphology. This enabled us to identify 267 novel loci and 645 genes with specific relevance for the nervous system in schizophrenia. It also made it possible to develop enriched PGSs based on overlapping genetic architecture between brain morphology and schizophrenia. This suggests that the current approach can be used to attain greater specificity in identifying the underlying biological pathways and improve our understanding of the disease mechanisms of schizophrenia.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the Research Council of Norway (Grant Nos. 276082 [to TK], 213837 [to OAA], 223273 [to OAA], 204966/F20 [to LTW], 229129 [to OAA], 249795/F20 [to LTW], 225989 [to OAA], 248778 [to OAA], 249795 [to LTW], 298646 [to LTW], 300767 [to LTW], and 323961 [to TK]), South-Eastern Norway Regional Health Authority (Grant Nos. 2013-123 [to OAA], 2014-097 [to LTW], 2015-073 [to LTW], 2016-064 [to OAA], 2017-004 [to IA], and 2019-101 [to LTW]), Stiftelsen Kristian Gerhard Jebsen (Grant No. SKGJ-Med-008 [to OAA and LTW]), European Research Council Starting Grant under the European Union Horizon 2020 Research and Innovation Programme (Grant No. 802998 [to LTW]), and European Union Horizon 2020 Research and Innovation Action Grant (Grant No. 847776 CoMorMent [to OAA])

This work was partly performed on the TSD (Tjeneste for Sensitive Data) facilities, owned by the University of Oslo, operated and developed by the TSD service group at the University of Oslo, IT Department (USIT) (tsd-drift@usit.uio.no). Computations were also performed on resources provided by UNINETT Sigma2, the national infrastructure for high performance computing and data storage in Norway.

DvdM conceived the study. DvdM and TK pre-processed the data. DvdM performed all analyses, with conceptual input from TK, AAS, OF, JMS, OAA,

and LTW. All authors contributed to interpretation of results. DvdM drafted the manuscript, and all authors contributed to and approved the final manuscript.

A previous version of this article was published as a preprint on medRxiv: https://doi.org/10.1101/2020.08.03.20167510.

The data incorporated in this work were gathered from public resources and the Thematically Organized Psychosis study. The code is available via https://github.com/precimed (GPLv3 license). Correspondence and requests for materials should be addressed to d.v.d.meer@medisin.uio.no.

OAA has received speaker's honorarium from Lundbeck and Sunovion and is a consultant to HealthLytix. AMD is a founder of and holds equity in CorTechs Labs, Inc., and serves on its Scientific Advisory Board. He is a member of the Scientific Advisory Board of Human Longevity, Inc., and receives funding through research agreements with GE Healthcare and Medtronic. The terms of these arrangements have been reviewed and approved by the University of California San Diego in accordance with its conflict of interest policies. All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Norwegian Centre for Mental Disorders Research (DvdM, AAS, KO, FB, TW, DA, IA, OBS, IM, LTW, OAA, OF, TK), KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital, and Institute of Clinical Medicine, University of Oslo; Department of Psychology (LTW) and Centre for Bioinformatics (OF), Department of Informatics, University of Oslo; Department of Medical Genetics (SD), Oslo University Hospital, Oslo; Norwegian Centre for Mental Disorders Research (SD), Department of Clinical Science, University of Bergen, Bergen, Norway; School of Mental Health and Neuroscience (DvdM, JMS, DEJL), Faculty of Health, Medicine and Life Sciences, Maastricht University; Department of Radiology and Nuclear Medicine (JMS), Maastricht University Medical Center, Maastricht, the Netherlands; Center for Multimodal Imaging and Genetics (AMD), University of California San Diego, La Jolla, California; and Department of Psychiatry and Psychotherapy (TK), University of Tübingen, Tübingen, Germany.

Address correspondence to Dennis van der Meer, Ph.D., at d.v.d.meer@medisin.uio.no.

Received Jul 21, 2021; revised Dec 6, 2021; accepted Dec 9, 2021.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2021.12.007.

REFERENCES

1. Smeland O, Frei O, Dale A, Andreassen OA (2020): The polygenic architecture of schizophrenia—rethinking pathogenesis and nosology. Nat Rev Neurol 16:366–379.
2. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014): Biological insights from 108 schizophrenia-associated genetic loci. Nature 511:421–427.
3. Holland D, Frei O, Desikan R, Fan CC, Shadrin AA, Smeland OB, *et al.* (2020): Beyond SNP heritability: Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. PLoS Genet 16:e1008612.
4. Franke B, Stein JL, Ripke S, Anttila V, Hibar DP, van Hulzen KJE, *et al.* (2016): Genetic influences on schizophrenia and subcortical brain volumes: Large-scale proof of concept. Nat Neurosci 19:420–431.
5. van Erp TGM, Walton E, Hibar DP, Schmaal L, Jiang W, Glahn DC, *et al.* (2018): Cortical brain abnormalities in 4474 individuals with schizophrenia and 5098 control subjects via the Enhancing Neuro Imaging Genetics Through Meta Analysis (ENIGMA) Consortium. Biol Psychiatry 84:644–654.
6. van Erp TGM, Hibar DP, Rasmussen JM, Glahn DC, Pearlson GD, Andreassen OA, *et al.* (2016): Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. Mol Psychiatry 21:547–553.
7. Moberget T, Doan NT, Alnæs D, Kaufmann T, Córdova-Palomera A, Lagerberg TV, *et al.* (2018): Cerebellar volume and cerebellocerebral structural covariance in schizophrenia: A multisite mega-analysis of 983 patients and 1349 healthy controls. Mol Psychiatry 23:1512–1520.
8. Alnæs D, Kaufmann T, van der Meer D, Córdova-Palomera A, Rokicki J, Moberget T, *et al.* (2019): Brain heterogeneity in schizophrenia and its association with polygenic risk. JAMA Psychiatry 76:739–748.
9. Wolfers T, Doan NT, Kaufmann T, Alnæs D, Moberget T, Agartz I, *et al.* (2018): Mapping the heterogeneous phenotype of schizophrenia and bipolar disorder using normative models. JAMA Psychiatry 75:1146–1155.
10. Andreassen OA, Thompson WK, Dale AM (2013): Boosting the power of schizophrenia genetics by leveraging new statistical tools. Schizophr Bull 40:13–17.
11. Ohi K, Shimada T, Kataoka Y, Yasuyama T, Kawasaki Y, Shioiri T, Thompson PM (2020): Genetic correlations between subcortical brain volumes and psychiatric disorders. Br J Psychiatry 216:280–283.
12. Frei O, Holland D, Smeland OB, Shadrin AA, Fan CC, Maeland S, *et al.* (2019): Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. Nat Commun 10:2417.
13. Smeland OB, Frei O, Shadrin A, O'Connell K, Fan CC, Bahrami S, *et al.* (2020): Discovery of shared genomic loci using the conditional false discovery rate approach. Hum Genet 139:85–94.
14. Wiström ED, O'Connell KS, Karadag N, Bahrami S, Hindley GFL, Lin A, *et al.* (2022): Genome-wide analysis reveals genetic overlap between alcohol use behaviours, schizophrenia and bipolar disorder and identifies novel shared risk loci. Addiction 117:600–610.
15. Cheng W, Frei O, van der Meer D, Wang Y, O'Connell KS, Chu Y, *et al.* (2021): Genetic association between schizophrenia and cortical brain surface area and thickness. JAMA Psychiatry 78:1020–1030.
16. van der Meer D, Frei O, Kaufmann T, Shadrin AA, Devor A, Smeland OB, *et al.* (2020): Understanding the genetic determinants of the brain with MOSTest [published correction appears in Nat Commun 2020; 14:4700]. Nat Commun 11:3512.
17. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, *et al.* (2015): UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 12:e1001779.
18. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, *et al.* (2018): The UK Biobank resource with deep phenotyping and genomic data. Nature 562:203–209.
19. Miller KL, Alfaro-Almagro F, Bangerter NK, Thomas DL, Yacoub E, Xu J, *et al.* (2016): Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci 19:1523–1536.
20. Brandt CL, Kaufmann T, Agartz I, Hugdahl K, Jensen J, Ueland T, *et al.* (2015): Cognitive effort and schizophrenia modulate large-scale functional brain connectivity. Schizophr Bull 41:1360–1369.
21. Kaufmann T, Skatun KC, Alnaes D, Doan NT, Duff EP, Tonnesen S, *et al.* (2015): Disintegration of sensorimotor brain networks in schizophrenia. Schizophr Bull 41:1326–1335.
22. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, *et al.* (2002): Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. Neuron 33:341–355.
23. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, *et al.* (2006): An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980.
24. Rosen AFG, Roalf DR, Ruparel K, Blake J, Seelaus K, Villa LP, *et al.* (2018): Quantitative assessment of structural image quality. Neuroimage 169:407–418.
25. Yang J, Lee SH, Goddard ME, Visscher PM (2011): GCTA: A tool for genome-wide complex trait analysis. Am J Hum Genet 88:76–82.
26. Beasley TM, Erickson S, Allison DB (2009): Rank-based inverse normal transformations are increasingly used, but are they merited? Behav Genet 39:580–595.
27. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015): Second-generation PLINK: Rising to the challenge of larger and richer datasets. Gigascience 4:7.

28. Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, *et al.* (2016): Familial risk and heritability of cancer among twins in Nordic countries. JAMA 315:68–76.
29. Lu D, Andersson TM, Fall K, Hultman CM, Czene K, Valdimarsdóttir U, Fang F (2016): Clinical diagnosis of mental disorders immediately before and after cancer diagnosis: A nationwide matched cohort study in Sweden. JAMA Oncol 2:1188–1196.
30. Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke S, Walters JTR, O'Donovan MC (2020): Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. medRxiv. https://doi.org/10.1101/2020.09.12.20192922.
31. Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC, *et al.* (2019): A global overview of pleiotropy and genetic architecture in complex traits [published correction appears in Nat Genet 2020; 52:535]. Nat Genet 51:1339–1348.
32. Watanabe K, Taskesen E, Bochoven A, Posthuma D (2017): Functional mapping and annotation of genetic associations with FUMA. Nat Commun 8:1826.
33. Choi SW, O'Reilly PF (2019): PRSice-2: Polygenic Risk Score software for biobank-scale data. Gigascience 8:giz082.
34. Wickam H (2009): ggplot2: Elegant Graphics for Data Analysis. New York: Springer.
35. de Leeuw CA, Mooij JM, Heskes T, Posthuma D (2015): MAGMA: Generalized gene-set analysis of GWAS data. PLoS Comput Biol 11: e1004219.
36. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J (2014): A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 46:310–315.
37. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, *et al.* (2012): Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 22:1790–1797.
38. Hamilton BA, Frankel WN, Kerrebrock AW, Hawkins TL, FitzHugh W, Kusumi K, *et al.* (1996): Disruption of the nuclear hormone receptor RORα in staggerer mice. Nature 379:736–739.
39. Jetten AM (2009): Retinoid-related orphan receptors (RORs): Critical roles in development, immunity, circadian rhythm, and cellular metabolism. Nucl Recept Signal 7:e003.
40. Kojetin DJ, Burris TP (2014): REV-ERB and ROR nuclear receptors as drug targets. Nat Rev Drug Discov 13:197–216.
41. Kuo TY, Hong CJ, Hsueh YP (2009): Bcl11A/CTIP1 regulates expression of DCC and MAP1b in control of axon branching and dendrite outgrowth. Mol Cell Neurosci 42:195–207.
42. Dias C, Estruch SB, Graham SA, McRae J, Sawiak SJ, Hurst JA, *et al.* (2016): BCL11A haploinsufficiency causes an intellectual disability syndrome and dysregulates transcription. Am J Hum Genet 99:253–274.
43. Schorling DC, Rost S, Lefeber DJ, Brady L, Müller CR, Korinthenberg R, *et al.* (2017): Early and lethal neurodegeneration with myasthenic and myopathic features: A new ALG14-CDG. Neurology 89:657–664.
44. Kvarnung M, Taylan F, Nilsson D, Anderlid B, Malmgren H, Lagerstedt-Robinson K, *et al.* (2018): Genomic screening in rare disorders: New mutations and phenotypes, highlighting ALG14 as a novel cause of severe intellectual disability. Clin Genet 94:528–537.
45. Gusev A, Mancuso N, Won H, Kousi M, Finucane HK, Reshef Y, *et al.* (2018): Transcriptome-wide association study of schizophrenia and chromatin activity yields mechanistic disease insights. Nat Genet 50:538–548.
46. Walker RL, Ramaswami G, Hartl C, Mancuso N, Gandal MJ, de la Torre-Ubieta L, *et al.* (2019): Genetic control of expression and splicing in developing human brain informs disease mechanisms. Cell 179:750–771.e22.
47. Ursini G, Punzi G, Chen Q, Marenco S, Robinson JF, Porcelli A, *et al.* (2018): Convergence of placenta biology and genetic risk for schizophrenia. Nat Med 24:792–801.