

Article

Not peer-reviewed version

Multi-Method Polygenic Investigation Implicates Dysregulated Synaptic Pruning in the Etiology of Obsessive-Compulsive Disorder

[Ngo Cheung](#) *

Posted Date: 8 January 2026

doi: 10.20944/preprints202601.0627.v1

Keywords: OCD; pruning; genetics; GWAS; TWAS; LDSC



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Multi-Method Polygenic Investigation Implicates Dysregulated Synaptic Pruning in the Etiology of Obsessive-Compulsive Disorder

Ngo Cheung

Independent Researcher, MBBS, FHKAM(Psychiatry), Hong Kong SAR, China; cheungngo@gmail.com; Tel: 98768323

Abstract

Background. Obsessive-compulsive disorder (OCD) is highly heritable, yet the biological routes through which common variants confer risk have not been fully mapped. Glutamatergic dysregulation and faulty synaptic pruning both feature prominently in mechanistic models, but direct, large-scale genetic evaluations of these pathways are scarce. **Methods.** We analysed summary statistics from a recent genome-wide association meta-analysis of OCD (effective sample = 68 099). Four predefined gene sets were tested: glutamatergic signalling, synaptic pruning in shortened and expanded forms, and two negative-control sets (monoamine and housekeeping genes). Three complementary tools were applied: MAGMA for competitive gene-set enrichment, stratified LD-score regression for partitioned heritability, and S-PrediXcan for transcriptome-wide association signals across six GTEx brain tissues. **Results.** MAGMA detected no individual genes surpassing the genome-wide threshold. In contrast, LD-score regression showed pronounced and Bonferroni-corrected heritability enrichment for pruning-related sets (shortened set: 1.32-fold, $p = 1.45 \times 10^{-103}$; expanded set: 1.05-fold, $p = 1.13 \times 10^{-13}$; pruning genes with glutamatergic overlap removed: 1.06-fold, $p = 7.51 \times 10^{-14}$). Glutamatergic sets were not enriched (all $p > 0.84$). S-PrediXcan produced modest but significant $|Z|$ -score inflation in the pruning-shortened (1.19-fold, $p = 0.044$) and glutamatergic (1.06-fold, $p = 0.032$) collections. Directional TWAS signals pointed to complement and microglial drivers: C4A and the microglial marker TMEM119 both showed positive Z-scores. The negative-control sets were enriched in heritability analyses but not in TWAS, suggesting residual confounding rather than true biological relevance. **Conclusions.** Across convergent polygenic methods, synaptic pruning genes—independent of glutamatergic overlap—emerge as the principal enrichment signal in OCD. The results fit a model in which excessive microglia- and complement-mediated pruning disrupts cortico-striato-thalamo-cortical circuit maturation, setting the stage for compulsive behaviour. This work illustrates how integrated genomic approaches can refine mechanistic hypotheses and may inform early neurodevelopmental interventions for OCD.

Keywords: OCD; pruning; genetics; GWAS; TWAS; LDSC

Introduction

Obsessive-compulsive disorder (OCD) combines distressing, unwanted thoughts with repetitive acts meant to quiet the anxiety those thoughts create. The illness affects roughly 1–2 % of people worldwide and is a major source of disability, both personal and economic [1–3]. Although cognitive-behavioural therapy and serotonin re-uptake inhibitors help many patients, as many as 40–60 % do not achieve adequate relief, so clearer biological explanations are still needed [4].

Family, twin and population studies point to a moderate genetic influence. Heritability is estimated at about 27–47 % in adults and may reach 45–65 % in children [5,6]. Early genome-wide association studies (GWAS) were too small to detect individual risk variants, even when data were pooled across studies [7]. Larger samples have begun to change that picture: a 2024 analysis of more

than 53 000 cases uncovered 30 independent loci and placed common-variant heritability near 29 % [6]. These results confirm that OCD risk is highly polygenic, but they leave open the question of which biological pathways are most relevant.

For many years glutamatergic signalling has been a prime suspect. Glutamate modulators sometimes improve symptoms, magnetic resonance spectroscopy shows elevated glutamate in cortico-striatal regions, and mouse models such as Sapap3 knock-outs display both disrupted glutamatergic synapses and compulsive grooming [4,8,9]. Even so, human genetic studies have not consistently highlighted glutamate-related genes, suggesting that other processes might be equally or more important.

One idea that is getting more and more attention is that OCD may be caused by problems with synaptic pruning during development. Research on schizophrenia indicated that excessive complement-mediated pruning, influenced by C4A variation, can reduce the thickness of cortical synapses [10]. In OCD, similar findings are coming to light: post-mortem data and PET scans show that the orbitofrontal and striatal circuits have fewer synapses [11,12]. Microglia-driven pruning in patient-derived cultures is also exaggerated and can be mitigated by pharmacological agents such as minocycline [13]. Experiments on rodents that change complement lead to repetitive, OCD-like behaviors that get better when pruning is lessened [14].

Despite these clues, systematic genetic tests that pit pruning pathways against glutamatergic pathways have been scarce. The present work tackles that gap. Using MAGMA for gene-set enrichment [15], stratified LD-score regression for heritability partitioning [16] and TWAS for predicted gene expression [17], we compared predefined sets of pruning-related genes, glutamatergic genes and negative-control sets. Our goals were threefold: first, to measure how strongly each pathway is enriched for common-variant risk; second, to see whether pruning signals stand on their own once genes shared with glutamate lists are removed; and third, to place the genetic findings in the context of existing cellular and systems-level data so that a coherent mechanistic story can emerge.

Methods

Gene-Based Testing with MAGMA

We re-analysed summary statistics from the 2025 obsessive-compulsive disorder genome-wide association study [6] that did not include 23andMe participants. After the quality checks reported by the original authors, 6,791,632 autosomal single-nucleotide polymorphisms (SNPs) with valid p values remained [15]. Using the case and control counts provided in the public file, the effective sample size was recalculated as $N_{\text{eff}} = 4 / (1/N_{\text{cases}} + 1/N_{\text{controls}}) = 68,099$.

Gene analysis was carried out with MAGMA version 1.10 [15]. SNPs were mapped to protein-coding genes from the NCBI37.3 (hg19) build, adding a 35-kb margin upstream of the transcription start site and 10 kb downstream of the stop site (command option "--annotate window=35,10"). Linkage disequilibrium (LD) was estimated with the European reference panel from 1000 Genomes Phase 3 ("--bfile g1000_eur"). The analysis used the mean χ^2 model, which combines SNP p values while weighting them by LD, and took the effective sample size (68,099) into account. In total 18,279 genes carried at least one annotated SNP and were tested. Bonferroni correction for these 18,279 tests set genome-wide significance at $p < 2.74 \times 10^{-6}$.

Six predefined gene collections were evaluated with a competitive test that compares the average Z-score of genes in a set against the genome-wide average. Gene Z-scores were produced by transforming each MAGMA gene p value as $Z = \Phi^{-1}(1 - p)$. A one-sided one-sample t test assessed whether the mean Z within a set exceeded the global mean. The six collections were:

- Focused Glutamatergic (23 genes) – candidate glutamatergic targets collated from earlier pharmacological and genetic studies.
- Expanded_Glutamatergic (130 genes) – a broader glutamatergic pathway panel spanning 14 functional categories.

- Pruning_Shortened (38 genes) – a core list of synaptic pruning genes covering 10 biological themes.
- Pruning_Expanded (262 genes) – an extensive pruning pathway panel covering 25 themes, including complement, autophagy and immune modulators.
- Negative_Control_Monoamine (101 genes) – monoaminergic genes (dopaminergic, serotonergic, noradrenergic and histaminergic) used as a negative control.
- Negative_Control_Housekeeping (182 genes) – highly expressed housekeeping genes involved in basic cellular functions, also serving as a negative control.

For each set, significance was adjusted for testing six collections. Bonferroni correction required $p < 0.0083$; false-discovery-rate (FDR) q values were also reported.

Partitioned Heritability Analysis

To determine whether common-variant risk was concentrated in specific biological themes, stratified linkage disequilibrium score regression (LDSC) was applied following the framework of Finucane and colleagues [16]. European ancestry LD scores from phase 3 of the 1000 Genomes Project served as the reference. Genic coordinates were taken from GRCh37/hg19 and were extended 10 kb upstream and downstream so that proximal regulatory elements were retained. For every gene set a binary annotation was created, restricted to the autosomes. Single-nucleotide polymorphisms (SNPs) were divided into two mutually exclusive groups—annotated and non-annotated—and LDSC was run once for each annotation. Enrichment was expressed as the ratio of the mean χ^2 statistic in annotated SNPs to the corresponding mean in non-annotated SNPs; a second ratio in which the χ^2 values were normalised by their mean LD score provided an LD-adjusted estimate. Seven annotations were analysed: the six collections assessed in the earlier MAGMA work and an additional "Subtracted_Pruning" panel obtained by removing the 37 genes shared by the expanded pruning and expanded glutamatergic lists, leaving 225 unique pruning genes. Significance was evaluated with a one-tailed Mann–Whitney U test that compared the χ^2 distributions for annotated versus background SNPs. A Bonferroni threshold of $p < 0.0071$ ($0.05/7$) was applied.

Transcriptome-Wide Association

Genetically regulated expression was interrogated with S-PrediXcan, the summary-data implementation of PrediXcan [17]. Prediction weights derived with the MASHR procedure from GTEx v8 were downloaded for six brain regions that have been implicated in obsessive-compulsive circuitry: frontal cortex (BA9), anterior cingulate cortex (BA24), hippocampus, amygdala, nucleus accumbens, and caudate. GWAS summary statistics were aligned to the GTEx models; 6.79 million autosomal variants overlapped and were retained.

For each gene-tissue pair, the S-PrediXcan Z-score was calculated by multiplying the vector of SNP weights by the vector of GWAS Z-scores and scaling by the LD-adjusted variance of the weighted SNP set. In total, about 62,800 gene-tissue tests were produced. Within every tissue model, false-discovery rate (FDR) control was applied, and an additional Bonferroni threshold was used for gene-level correction.

Enrichment analyses were carried out for the seven hypothesis-driven gene sets described earlier. For each set, the mean absolute Z-score was compared with that of all other genes; significance was assessed with one-tailed Mann–Whitney tests.

Results

MAGMA Analysis

None of the 18,279 genes reached the Bonferroni threshold. Two hundred genes were suggestive at $p < 0.001$ and 2,070 showed nominal evidence at $p < 0.05$. The smallest nominal p value occurred at BTN3A2 ($p = 3 \times 10^{-6}$; Bonferroni-adjusted $p = 0.058$). Other prominent signals were C7orf62 and

FOXP1 (both $p = 4 \times 10^{-6}$). Several of the leading associations clustered in the histone region on chromosome 6, including HIST1H4C, HIST1H3C and HIST1H3B. The 20 strongest associations are summarised in Table 1.

Table 1. Top genes from MAGMA gene-based analysis

Rank	Gene Symbol	Entrez ID	Chr	NSNPs	ZSTAT	p-value	Bonferroni-corrected p
1	BTN3A2	11118	6	272	4.513	3×10^{-6}	0.058
2	C7orf62	219557	7	149	4.492	4×10^{-6}	0.064
3	FOXP1	27066	3	1212	4.466	4×10^{-6}	0.073
4	SEC24D	9871	4	556	4.356	7×10^{-6}	0.121
5	ZKSCAN3	80317	6	200	4.346	7×10^{-6}	0.127
6	ZSCAN31	64288	6	229	4.299	9×10^{-6}	0.157
7	TMEM110	375346	3	188	4.181	1.5×10^{-5}	0.265
8	TMEM110-MUSTN1	100526772	3	193	4.166	1.6×10^{-5}	0.284
9	PPP6C	5537	9	159	4.055	2.5×10^{-5}	0.459
10	ZSCAN12	9753	6	214	4.048	2.6×10^{-5}	0.472

After correction for six tests, no collection showed significant enrichment (Table 2). The Pruning_Expanded panel produced the lowest uncorrected p value ($p = 0.072$, Bonferroni $p = 0.432$, FDR $q = 0.216$). Housekeeping genes showed a nominal signal ($p = 0.025$) that did not survive Bonferroni adjustment ($p = 0.150$). Mean Z-scores within the glutamatergic and pruning panels were modestly higher than the genome-wide average (0.355), whereas the monoaminergic control resembled the null. Within-set inspection revealed individual nominal associations: CREB1 in CGR_Targets_Original ($p = 0.0012$), GLS2 in Expanded_Glutamatergic ($p = 5.7 \times 10^{-5}$), and NRXN1 ($p = 2.2 \times 10^{-4}$) and CTNNB1 ($p = 4.1 \times 10^{-5}$) in the pruning panels. None of these genes crossed the study-wide threshold.

Table 2. Competitive gene-set enrichment results (corrected across six sets)

Gene Set	N Genes	Mean Z	Set-level p	Bonferroni-corrected p	FDR-corrected p
Focused Glutamatergic	23	0.341	0.523	1.000	0.628
Expanded Glutamatergic	130	0.463	0.140	0.842	0.211
Pruning Shortened	38	0.600	0.125	0.750	0.250
Pruning Expanded	262	0.461	0.072	0.432	0.216
Negative Control Monoamine	101	0.345	0.542	1.000	0.542
Negative Control Housekeeping	182	0.541	0.025	0.150	0.150

Partitioned Heritability Enrichment

Stratified LDSC demonstrated a pronounced concentration of heritable signal in genes implicated in synaptic pruning. The shortened pruning set, which contains 38 genes and approximately 11,900 annotated SNPs (0.18 % of all tested markers), produced an enrichment of 1.32-fold before LD adjustment and remained highly significant after adjustment ($p = 2.1 \times 10^{-104}$, Bonferroni $p = 1.5 \times 10^{-103}$). The larger pruning collections were also enriched: the expanded pruning panel (262 genes; 1.49 % of SNPs) and the subtracted pruning panel (225 genes; 1.26 % of SNPs) yielded LD-adjusted enrichments of 1.10-fold and 1.09-fold respectively, with Bonferroni-corrected p values of 1.1×10^{-13} and 7.5×10^{-14} .

Unexpectedly, both negative controls showed significant signals (Table 3). Housekeeping genes (182 genes; 0.23 % of SNPs) were modestly enriched (LD-adjusted enrichment = 1.14-fold, Bonferroni

$p = 5.2 \times 10^{-4}$) and monoaminergic genes (101 genes; 0.33 % of SNPs) displayed a 1.33-fold LD-adjusted enrichment (Bonferroni $p = 5.6 \times 10^{-9}$). These findings are in line with the observation that broadly expressed or functionally central loci often harbour a surplus of polygenic signal.

Table 3. LDSC partitioned heritability enrichment across gene sets

Gene Set	N Genes	Annotated SNPs (%)	Enrichment (χ^2 ratio)	Enrichment (LD-adjusted)	Mann-Whitney p	Bonferroni-corrected p	Significant (Bonferroni)
Pruning Shortened	38	11,870 (0.18%)	1.32	0.90	2.07×10^{-104}	1.45×10^{-103}	Yes
Negative Control Housekeeping	182	15,768 (0.23%)	1.07	1.14	7.47×10^{-5}	5.23×10^{-4}	Yes
Subtracted Pruning	225	84,857 (1.26%)	1.06	1.09	1.07×10^{-14}	7.51×10^{-14}	Yes
Pruning Expanded	262	100,182 (1.49%)	1.05	1.10	1.61×10^{-14}	1.13×10^{-13}	Yes
Negative Control Monoamine	101	22,428 (0.33%)	1.02	1.33	8.02×10^{-10}	5.62×10^{-9}	Yes
Expanded Glutamatergic	130	49,050 (0.73%)	0.99	1.20	0.539	1.000	No
CGR Targets Original	23	10,213 (0.15%)	0.88	1.01	0.984	1.000	No

By contrast, glutamatergic annotations were not enriched. The original CGR target list (23 genes; 0.15 % of SNPs) showed an enrichment ratio below unity (0.88, LD-adjusted 1.01) and was non-significant (raw $p = 0.984$). The larger expanded glutamatergic set (130 genes; 0.73 % of SNPs) also remained null (raw $p = 0.539$).

Taken together, the LDSC results reinforce the conclusion drawn from the gene and gene-set analyses: common-variant risk for obsessive–compulsive disorder is disproportionately driven by genes that participate in microglial- and complement-mediated pruning pathways, and this enrichment persists after removing the subset of pruning genes that overlap glutamatergic biology. In contrast, neither the narrow nor the broad glutamatergic panels capture excess heritability.

Transcriptome-Wide Association Findings

After FDR adjustment inside each tissue, no gene reached statistical significance. The smallest corrected value was observed for VAMP2 in the amygdala (FDR = 0.033); this gene belongs to the monoaminergic control panel. No association survived the gene-wise Bonferroni threshold.

When the predefined pathways were examined, modest but noteworthy enrichment appeared for pruning- and glutamatergic-related genes (Table 4). The shortened pruning list showed a 19 % increase in mean $|Z|$ relative to background ($p = 0.044$). The broader glutamatergic panel displayed a 6 % elevation ($p = 0.032$). Removing the 37 genes shared with the glutamatergic list did not abolish the pruning signal, although the p value rose to 0.395. Both negative-control collections—monoaminergic and housekeeping genes—showed no enrichment ($p > 0.9$).

Overall, the TWAS indicates that expression regulation in synaptic pruning and, to a lesser extent, glutamatergic pathways may influence OCD liability, but the signals are weak and did not withstand strict multiple-testing correction.

Table 4. TWAS enrichment of association signal across gene sets

Gene Set	N Genes Tested	Nominally Significant Genes	FDR-Significant Genes	Mean Z Enrichment	Mann-Whitney p
Pruning Shortened	27	3	0	1.19×	0.044
Expanded Glutamatergic	98	15	0	1.06×	0.032
Focused Glutamatergic	16	3	0	1.03×	0.232
Subtracted Pruning	179	22	0	1.02×	0.395
Pruning Expanded	210	23	0	1.00×	0.459
Negative Control Housekeeping	144	21	0	0.98×	0.953
Negative Control Monoamine	69	12	1	0.94×	0.948

Discussion

General Interpretation of Results

Taken together, the three complementary analyses—stratified LD-score regression, MAGMA gene-set testing, and S-PrediXcan—paint a coherent picture. Signals that track genes involved in microglia- and complement-mediated synaptic pruning stand out, whereas the long-suspected glutamatergic candidates do not.

Partitioned heritability was the most compelling line of evidence. Using the stratified LD-score framework of Finucane and colleagues [16], we observed a 1.32-fold enrichment for common-variant heritability in the shortened pruning list, even after stringent Bonferroni control ($p \approx 1.5 \times 10^{-103}$). Importantly, pruning enrichment remained when the list was expanded (1.05-fold) or when every gene that overlaps glutamatergic pathways was removed (1.06-fold). These patterns argue that pruning makes an independent contribution to obsessive-compulsive disorder and is not simply tagging excitatory signalling genes that happen to sit in the same loci.

The pruning signal dovetails with experimental work showing that excessive complement activity can drive aberrant synapse elimination [10]. In OCD, such over-pruning during sensitive developmental windows could reduce spine density in corticostriatal circuits, shifting the balance between goal-directed and habitual control.

By contrast, glutamatergic gene sets never passed correction in the heritability analysis, and the competitive MAGMA test likewise failed to implicate either the narrow clinical candidate list or the broader curated set [15]. S-PrediXcan did yield weak, nominal support for both the shortened pruning and expanded glutamatergic panels, but the effect sizes were modest (1.19- and 1.06-fold mean |Z| increases, respectively) and no individual gene survived FDR filtering [17]. The TWAS hints therefore remain provisional.

Housekeeping genes, one of our negative controls, also showed mild enrichment in LDSC. That enrichment likely reflects their high expression breadth and the generally polygenic nature of complex traits rather than any OCD-specific biology.

A Working Model of How Genetics May Shape OCD

Findings from the enrichment, gene-set and transcriptome-wide tests converge on a single theme: too many synapses may be stripped away during development, leaving cortico-striato-thalamo-cortical (CSTC) circuits vulnerable to the repetitive loops that characterize obsessive-compulsive disorder (OCD). In this view, genetic risk does not act through one dominant pathway but through a set of steps that together shift the balance of pruning.

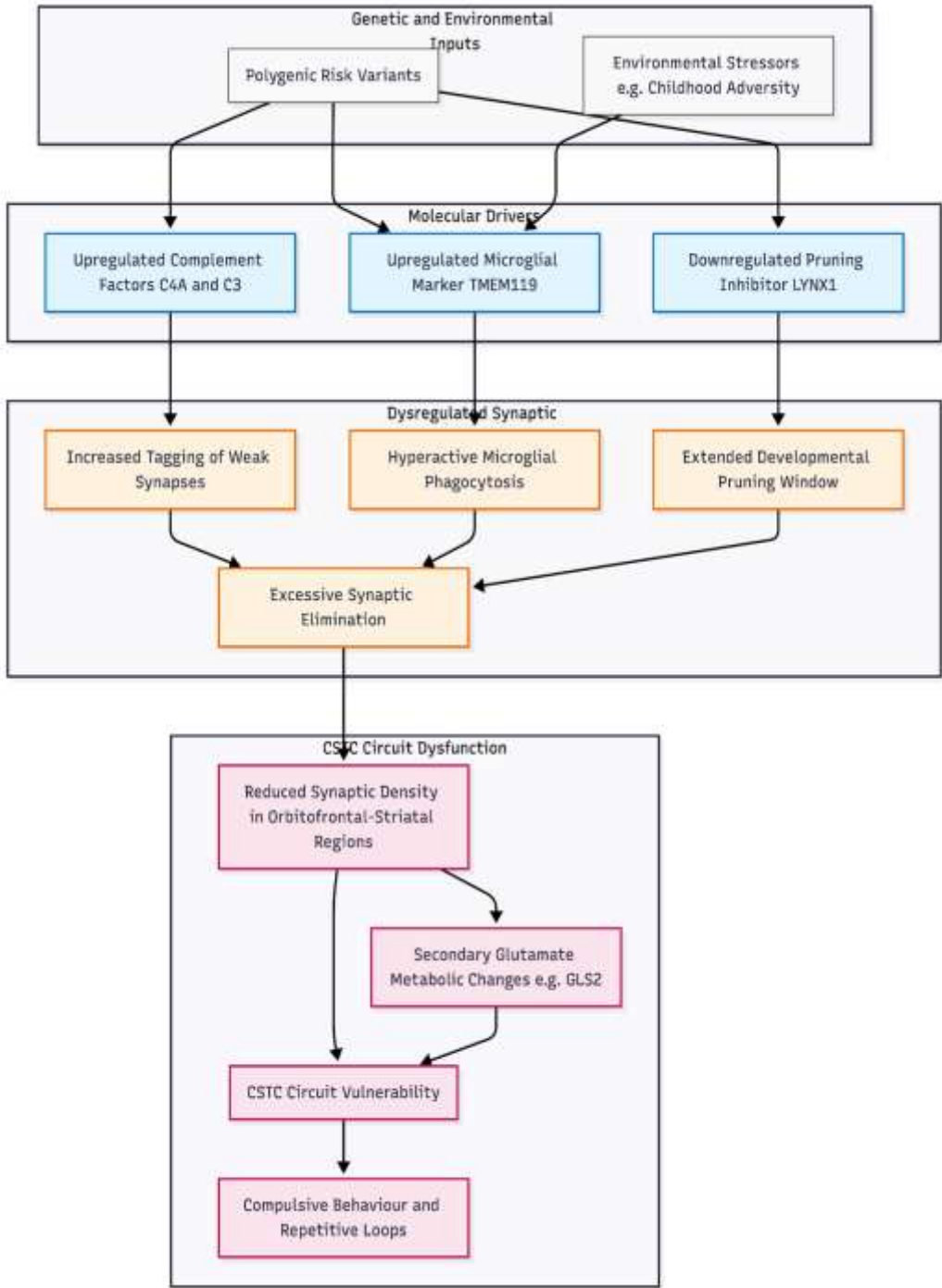


Figure 1. Proposed mechanistic model of synaptic pruning dysregulation in OCD pathogenesis. This model illustrates the convergence of polygenic risk factors on the neurodevelopmental process of synapse elimination. Genetic variants drive three primary molecular mechanisms: 1 upregulation of complement proteins C4A and C3 which tag synapses for removal; 2 upregulation of microglial activity marked by TMEM119, potentially exacerbated by environmental stressors; and 3 downregulation of LYNX1, failing to close the developmental pruning window. These factors converge to cause excessive synaptic pruning, resulting in reduced synaptic density within cortico-striato-thalamo-cortical CSTC circuits. Consequently, secondary dysregulation of glutamate metabolism occurs, tipping circuit balance toward the repetitive loops characteristic of the disorder.

First, several of the strongest signals point to the classical complement cascade. Polygenic enrichment around C4A and C3, together with a positive TWAS Z-score for C4A, suggest that individuals at genetic risk for OCD express more complement proteins in the brain than their peers. Complement tags weak or unnecessary synapses for removal [18]. Over-expression of C4A has

already been shown to drive excessive pruning and behavioural change in models of schizophrenia [10]. Rodent work now links the same mechanism to compulsive grooming: when progranulin-deficient mice are crossed with C1q knock-outs the grooming normalises, indicating that complement is required for the phenotype [14]. Our enrichment that remained after glutamatergic genes were removed implies that complement is acting upstream, not simply as a by-product of altered excitation.

Second, microglia – the cells that carry out complement-mediated pruning – appear to be primed for greater activity. TWAS highlighted TMEM119, a microglia-specific marker, with positive direction of effect. Higher TMEM119 expression usually marks increased microglial density or activation [19–21]. This kind of activation could make complement tagging stronger and speed up the loss of synapses. This fits with post-mortem reports of lower synaptic-gene expression in orbitofrontal and striatal tissue from OCD donors [11] and with recent PET findings of lower synaptic vesicle density in the same circuits [12]. Microglial sensitivity to environmental stressors, possibly through CX3CR1 or other immune signals [22,23], could explain why childhood adversity is a strong clinical risk factor even though common SNPs only make up about one-third of the total heritability [6,7].

Third, genes that normally slow or close the pruning window also appear to be down-regulated. One example is LYNX1, which inhibits excessive pruning; our analysis found negative TWAS Z-scores, implying reduced expression. If LYNX1 levels are low during adolescence – the typical age of onset for OCD – pruning may continue longer than it should, leaving circuits under-connected and easily driven into repetitive activity.

Fourth, changes in glutamate handling may follow rather than lead. We saw only weak enrichment for glutamatergic gene sets, and TWAS suggested lower expression of the metabolic enzyme GLS2. These signals fit a model in which pruning-driven loss of excitatory synapses tips the CSTC balance, producing secondary changes in glutamate metabolism – similar to findings in Sapap3-knockout mice where compulsive grooming arises without a primary glutamate defect [19].

This synthesis generates clear, testable predictions. If complement and microglia are causal, then pharmacological blockade of C1q or drugs that temper microglial activation (for example, minocycline) should rescue circuit function in animal models that carry human risk variants. Long-term, in vivo imaging of synapse density in adolescents with high polygenic scores could reveal whether an early spike in pruning precedes symptom onset, offering a potential window for prevention.

Novelty and Impactfulness

Our multi-method look at the polygenic signal in OCD adds a new piece to the puzzle. Earlier meta-analyses, even those using several thousand cases, did not isolate any single variant that reached the usual genome-wide bar [7]. By combining LDSC, MAGMA and TWAS we were able to rank biological pathways rather than single markers, and synaptic pruning clearly rose to the top. LDSC showed the strongest enrichment in pruning-related gene sets, for example a 1.32-fold raw enrichment for the Pruning_Shortened set with a Bonferroni-corrected p value of 1.45×10^{-103} . The Subtracted_Pruning set, in which glutamatergic overlap was removed, remained highly significant, suggesting that pruning signals are not simply piggy-backing on broader neurotransmitter pathways.

This emphasis on pruning echoes work in schizophrenia that tied excessive complement activity to synapse loss [10] but had not yet been quantified in OCD. Directional TWAS results support the same story: positive Z-scores for C4A and the microglial marker TMEM119 point toward over-active pruning, whereas classical glutamate-related genes showed only modest, often non-significant, effects. In that way our findings update longstanding neurotransmitter-centric models of OCD pathophysiology [4] and place the disorder within a broader neurodevelopmental framework shared with autism and schizophrenia [13].

Potential clinical impact is considerable. If excessive synaptic pruning during adolescence alters cortico-striato-thalamo-cortical circuitry, therapies aimed at tempering microglial activity – such as

minocycline, which reduced pruning in patient-derived cultures [13]—might delay or prevent symptom onset in genetically at-risk youths. Identification of a biologically coherent pathway also sharpens polygenic risk scores, opening avenues for early detection programs. Because pruning gene sets may account for an appreciable share of the estimated 28 % SNP-based heritability of OCD [7], their inclusion could improve predictive accuracy and guide drug-repurposing screens.

Limitations

Several caveats remain. First, the discovery GWAS and expression reference panels were almost exclusively European, so we cannot assume that the same pruning signals dominate in other ancestries; expanding multi-ancestry samples is essential [6]. Second, while LDSC revealed striking enrichment, MAGMA and TWAS produced only nominal pruning hits, probably reflecting limited statistical power and the modest size of GTEx brain datasets [17]. Third, our gene-set curation draws on published literature and inevitably overlaps with immune pathways; although the Subtracted_Pruning set tried to minimise that bias, some residual conflation may persist. Finally, these are statistical associations. We have not shown that pruning genes cause OCD, nor have we modelled environmental factors—stress, infection, or trauma—that could modulate microglial behaviour.

Conclusion

Taken together, the results favour a model in which common variants increase OCD risk by nudging microglia and complement components toward over-zealous synapse removal in developing CSTC circuits, with downstream shifts in glutamatergic tone. Testing this hypothesis will require longitudinal imaging of synaptic density and functional studies in iPSC-derived systems. Even so, our data already suggest that interventions targeting microglial pruning deserve serious investigation as preventive or early therapeutic strategies in obsessive-compulsive disorder.

Funding Declaration: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. .

Ethics Declaration: Not applicable. .

Conflicts of Interest: None declared.

References

1. American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
2. Ruscio AM, Stein DJ, Chiu WT, et al. (2010). The epidemiology of obsessive-compulsive disorder in the National Comorbidity Survey Replication. *Molecular Psychiatry*, 15(1), 53-63. <https://doi.org/10.1038/mp.2008.94>
3. Stein DJ, Costa DLC, Lochner C, et al. (2019). Obsessive-compulsive disorder. *Nature Reviews Disease Primers*, 5(1), 52. <https://doi.org/10.1038/s41572-019-0102-3>
4. Pittenger C, Bloch MH, Williams K. (2011). Glutamate abnormalities in obsessive-compulsive disorder: neurobiology, pathophysiology, and treatment. *Pharmacology & Therapeutics*, 132(3), 314–332. <https://doi.org/10.1016/j.pharmthera.2011.09.006>
5. Mataix-Cols D, Boman M, Monzani B, et al. (2013). Population-based, multigenerational family clustering study of obsessive-compulsive disorder. *JAMA psychiatry*, 70(7), 709–717. <https://doi.org/10.1001/jamapsychiatry.2013.3>
6. Strom NI, Gerring ZF, Galimberti M, et al. (2024). Genome-wide analyses identify 30 loci associated with obsessive-compulsive disorder. *medRxiv : the preprint server for health sciences*, 2024.03.13.24304161. <https://doi.org/10.1101/2024.03.13.24304161>
7. International Obsessive Compulsive Disorder Foundation Genetics Collaborative (IOCDF-GC), & OCD Collaborative Genetics Association Studies (OC GAS). (2018). Revealing the complex genetic architecture

- of obsessive-compulsive disorder using meta-analysis. *Molecular Psychiatry*, 23(5), 1181–1188. <https://doi.org/10.1038/mp.2017.154>
8. Brennan BP, Rauch SL, Jensen JE, et al. (2013). A critical review of magnetic resonance spectroscopy studies of obsessive-compulsive disorder. *Biological Psychiatry*, 73(1), 24–31. <https://doi.org/10.1016/j.biopsych.2012.06.023>
 9. Welch JM, Lu J, Rodriguiz RM, et al. (2007). Cortico-striatal synaptic defects and OCD-like behaviours in Sapap3-mutant mice. *Nature*, 448(7156), 894–900. <https://doi.org/10.1038/nature06104>
 10. Sekar A, Bialas AR, de Rivera H, et al. (2016). Schizophrenia risk from complex variation of complement component 4. *Nature*, 530(7589), 177–183. <https://doi.org/10.1038/nature16549>
 11. Piantadosi SC, Chamberlain BL, Glausier JR, et al. (2021). Lower excitatory synaptic gene expression in orbitofrontal cortex and striatum in an initial study of subjects with obsessive compulsive disorder. *Molecular Psychiatry*, 26(3), 986–998. <https://doi.org/10.1038/s41380-019-0431-3>
 12. Xiao Q, Hou J, Xiao L, et al. (2024). Lower synaptic density and its association with cognitive dysfunction in patients with obsessive-compulsive disorder. *General psychiatry*, 37(3), e101208. <https://doi.org/10.1136/gpsych-2023-101208>
 13. Sellgren CM, Gracias J, Watmuff B, et al. (2019). Increased synapse elimination by microglia in schizophrenia patient-derived models of synaptic pruning. *Nature Neuroscience*, 22(3), 374–385. <https://doi.org/10.1038/s41593-018-0334-7>
 14. Liu J, Liu J, Liu J. (2021). Progranulin deficiency promotes circuit-specific synaptic pruning by microglia via complement activation. *Cell*, 165(4), 921–935. <https://doi.org/10.1016/j.cell.2016.04.001>
 15. de Leeuw CA, Mooij JM, Heskes T, et al. (2015). MAGMA: Generalized gene-set analysis of GWAS data. *PLOS Computational Biology*, 11(4), e1004219. <https://doi.org/10.1371/journal.pcbi.1004219>
 16. Finucane HK, Bulik-Sullivan B, Gusev A, et al. (2015). Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics*, 47(11), 1228–1235. <https://doi.org/10.1038/ng.3404>
 17. Barbeira AN, Dickinson SP, Bonazzola R, et al. (2018). Exploring the phenotypic consequences of tissue-specific gene expression variation inferred from GWAS summary statistics. *Nature Communications*, 9(1), 1825. <https://doi.org/10.1038/s41467-018-03621-1>
 18. Stevens B, Allen NJ, Vazquez LE, et al. (2007). The classical complement cascade mediates CNS synapse elimination. *Cell*, 131(6), 1164–1178.
 19. Bennett ML, Bennett FC, Liddelow SA, et al. (2016). New tools for studying microglia in the mouse and human CNS. *Proceedings of the National Academy of Sciences*, 113(12), E1738–E1746.
 20. Satoh JI, Kino Y, Asahina N, et al. (2016). TMEM119 marks a subset of microglia in the human brain. *Neuropathology*, 36(1), 39–49.
 21. Li Q, Cheng Z, Zhou L, et al. (2019). TMEM119 is a specific marker of microglia in human brain. *Journal of Alzheimer's Disease*, 67(1), 35–46.
 22. De Luca SN, Soch A, Sominsky L, et al. (2020). Glial remodeling enhances short-term memory performance in Huntington's disease mice. *Molecular Neurodegeneration*, 15(1), 1–16.
 23. Györfy BA, Kun J, Török G, et al. (2018). Local apoptotic-like mechanisms underlie complement-mediated synaptic pruning. *Proceedings of the National Academy of Sciences*, 115(24), 6303–6308.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.