Research Article

A Claim-Governed Framework for Conservative Molecular Candidate Prioritization in Schizophrenia

Regio Marcos Pinto Abreu Filho[1]

1. Psychiatry Department, Military Police of the State of Rio de Janeiro (PMERJ), Brazil

Background. Schizophrenia genetics, postmortem transcriptomics, regulatory annotation, and biomedical literature resources produce large lists of plausible molecular candidates. These evidence layers are not interchangeable: genetic association does not prove differential expression, bulk expression does not establish cell-specific mechanism, literature co-mention does not validate disease specificity, and circuit annotation does not identify a therapeutic target.

Objective. To present a claim-governed candidate-prioritization framework that classifies evidence strength while explicitly blocking unsupported causal, biomarker, target, and validation claims.

Methods. The framework separates candidate intake, genetic-prior evidence, transcriptomic context, recurrence annotation, literature metadata resolution, specificity controls, negative controls, circuit-planning annotation, and final claim-firewall review. Candidate labels are assigned on a conservative ladder: L0 for unsupported or blocked candidates, L1 for hypothesis-generating single-route support, L2 for independent multi-route support, and L3 for experimental or clinical validation. Circuit annotation, novelty, low bibliometric saturation, and title/abstract co-mention are prohibited from increasing molecular claim level.

Illustrative audit result. In the frozen illustrative run reported here, 406 initial candidates were processed, 5 reached L1, and 0 reached L2. SNAP91 was retained only as a future-audit annotation, OLR1 as annotation-only mapping support, and GRM1 as a glutamatergic context-control gene. No gene was reported as a validated schizophrenia gene, biomarker, therapeutic target, or causal mechanism.

Conclusions. The framework is intended to reduce inferential inflation in computational neuropsychiatric candidate prioritization. The present demonstration is negative by design: it preserves auditable hypotheses without promoting weak signals into biological discovery claims.

Corresponding author: Regio Marcos Pinto Abreu Filho, regiomarcosabreu@gmail.com

Summary

This paper does not claim to have discovered a schizophrenia gene. It describes a conservative method for deciding when public genetic, expression, and literature signals are too weak to support strong biological claims. In the demonstration run, no candidate passed the threshold for independent multi-route support.

Introduction

Schizophrenia is a highly polygenic psychiatric disorder. Large genome-wide association studies have identified hundreds of associated loci and have shown enrichment of signals in genes expressed in excitatory and inhibitory neurons [1][2][3]. However, the translation of loci into causal genes, cell types, mechanisms, biomarkers, or treatment targets remains difficult. Most common-variant associations are non-coding, linkage disequilibrium spreads statistical signal across genomic intervals, and gene-mapping approaches can nominate different genes depending on the annotation layer used [4][5][6][7].

Transcriptomic resources add biological context but introduce a different set of limitations [8][9][10]. Bulk postmortem brain expression reflects mixtures of cell types and is affected by region, RNA quality, age, sex, medication exposure, agonal state, smoking, and other clinical or technical confounders. A differentially expressed gene is not necessarily causal, and a genetically prioritized gene is not necessarily differentially expressed. Literature retrieval creates another risk: a title or abstract hit can reflect background mention, broad psychiatric relevance, or unrelated biological function rather than schizophrenia-specific evidence.

The problem addressed here is therefore not simply candidate discovery. It is claim governance. Computational pipelines can produce plausible candidate narratives even when the evidence is indirect or non-independent. A conservative evidence framework should make the permitted claim explicit, define what evidence is insufficient, and prevent escalation from annotation to validation.

This manuscript presents a claim-governed framework for molecular candidate prioritization in schizophrenia. The framework is designed to retain weak candidates as auditable future-validation annotations while blocking unsupported claims about biomarkers, targets, mechanisms, validated genes, or causal disease processes.

Methods

Study design

This is a computational methods and evidence-governance study using public schizophrenia-relevant evidence layers. No new human participants, animal experiments, clinical interventions, identifiable data, primary biological samples, or patient-level datasets were collected. The methodological question was whether heterogeneous evidence can be organized into an auditable candidate framework without overpromoting unsupported biological claims.

Reporting status and reproducibility requirement

For a submission-ready version, the candidate matrix, evidence-layer matrix, code, query logs, and freeze manifest must be deposited in a public repository and referenced with a persistent identifier, consistent with FAIR principles and current expectations for transparent evidence reporting [11][12][13]. The illustrative run below should not be treated as independently reproducible unless those files accompany the manuscript. The minimum required files are specified in Table 5.

Candidate intake

Candidate intake begins with a defined gene universe derived from public schizophrenia-relevant genetic-prior and annotation sources. Each candidate must be represented by a stable HGNC-approved symbol and, where possible, an Ensembl identifier. Synonyms are retained for literature retrieval but do not replace the approved symbol. Ambiguous gene aliases are manually flagged before candidate scoring.

The frozen illustrative run contained 406 initial candidates. The manuscript does not interpret the number 406 as a discovery result; it is only the size of the candidate universe passed into the audit pipeline.

Evidence layers

Layer	Permitted role	Prohibited claim
Genetic-prior evidence	Association or prioritization context from schizophrenia-relevant genetic studies.	Differential expression, mechanism, target status, or causal gene status.
Transcriptomic context	Expression or dysregulation context, with dataset-specific limitations.	Cell-specific mechanism or disease causality from bulk data alone.
Regulatory or chromatin annotation	Hypothesis support when independently linked to a candidate gene.	Validation unless experimentally tested in an appropriate model.
Secondary recurrence	Recurrence annotation across independent resources.	Replication unless source independence and methods are shown.
Literature metadata resolver	Title/abstract-level context and future-audit routing.	Full-text validation, target discovery, biomarker discovery, or mechanism.
Bibliometric saturation	Low or high literature saturation flag.	Biological importance, novelty proof, or disease specificity.
Negative controls	Detection of nonspecific or overbroad behavior.	Disease validation.
Circuit annotation	Planning of future functional validation tasks.	Molecular mechanism or clinical circuit biomarker.
Claim firewall	Final textual audit against prohibited claims.	Evidence generation by itself.

Claim ladder

Level	Definition	Permitted manuscript language	Blocked language
L0	Unsupported, failed evidence floor, negative-control caution, nonspecific, or exploratory only.	Exploratory, blocked, unsupported, context only.	Candidate, target, biomarker, mechanism, validation.
L1	Hypothesis-generating support from one admissible evidence route, without independent multi-route confirmation.	Future-audit candidate, annotation-level support, hypothesis-generating.	Validated gene, causal gene, therapeutic target, biomarker.
L2	Independent multi-route support with documented source independence and specificity controls.	Evidence-qualified candidate, still requiring experimental validation.	Validated target or mechanism unless functional evidence exists.
L3	Experimental or clinical validation in an appropriate biological or clinical system.	Validated only for the exact assay, context, and endpoint tested.	Generalized clinical utility unless prospectively evaluated.

The illustrative run allowed L0–L2 classification. No candidate reached L2. No L3 evidence was generated or claimed.

Evidence-floor rule

The evidence floor is applied before novelty assessment, literature saturation, circuit annotation, or simulation. A candidate cannot be promoted by being under-studied, familiar, biologically plausible, or co-mentioned with schizophrenia. Promotion requires admissible evidence and documented independence between supporting routes.

For full reproducibility, each route-specific threshold must be supplied in a machine-readable table. A route may be binary, ordinal, or quantitative, but it must be prespecified, source-specific, and auditable.

Literature metadata resolver

The resolver is a metadata triage tool, not a systematic review and not a validation method. If the literature component is later framed as evidence mapping or a scoping review, the search and screening report should follow PRISMA-ScR expectations [11]. In this framework, the resolver records whether the approved gene symbol or approved synonym is present in the title or abstract, whether schizophrenia or psychosis context is present, whether brain, expression, genetic, regulatory, or cell-type language is present, and whether overclaim-sensitive terms are detected. When a metadata hit is ambiguous, the output is downgraded to future-audit or specificity-warning status until full-text review resolves the context.

A submission-ready implementation must provide query strings, database names, API versions, query dates, synonym dictionaries, manual disambiguation rules, and PMID or Europe PMC identifiers for all retained literature hits.

Specificity and negative controls

Specificity controls are required because psychiatric, neurodevelopmental, inflammatory, metabolic, and synaptic genes may appear across multiple disorders. Cross-disorder recurrence can be biologically important but is not schizophrenia-specific evidence unless a schizophrenia-specific route is independently demonstrated. Negative controls are used to detect pipelines that indiscriminately promote broad genes, highly cited genes, or generic brain genes.

Circuit and simulation annotation

Circuit-level annotations may help plan future experiments, for example by linking candidate hypotheses to salience processing, auditory prediction, source monitoring, or excitatory/inhibitory balance. These annotations are planning-only. They do not increase molecular claim level and do not validate a candidate gene.

Claim firewall

A final claim firewall is applied to the manuscript and candidate outputs. Forbidden positive claims include therapeutic target, validated target, biomarker, diagnostic marker, causal mechanism, validated schizophrenia gene, discovered target, brain degeneration, and L2 candidate unless the relevant claim level is actually met. Negated uses are permitted only when used to state limitations or blocked claims.

Results

Candidate attrition

Stage	Number of candidates	Interpretation
Initial candidate universe	406	Genes entering the audit pipeline; not a discovery result.
L1 retained	5	Hypothesis-generating single-route or limited annotation support.
L2 retained	0	No candidate met independent multi-route evidence requirements.
Final discussed rows	3	SNAP91, OLR1, and GRM1 retained only for limited interpretation.

The main result is negative. The pipeline did not identify a validated schizophrenia gene, biomarker, therapeutic target, causal mechanism, or L2 molecular candidate.

Final candidate status

Gene	Final status	Permitted interpretation	Required next evidence
SNAP91	Future-audit annotation	Most informative retained audit case in the illustrative run; not validated and not L2.	Specificity controls, independent transcriptomic or regulatory recurrence, full source audit, and functional follow-up if stronger evidence emerges.
OLR1	Annotation-only support	Weak future-audit annotation; no target, biomarker, or mechanism claim.	Independent recurrence, full-text confirmation, alias disambiguation, and schizophrenia-specific support.
GRM1	Context-control gene	Glutamatergic excitatory/inhibitory context and calibration control; not a new pipeline discovery.	Use as comparator unless it meets the same evidence floor as other candidates.

Resolver and specificity audit

The literature metadata resolver refined interpretation but did not change molecular claim levels. Specificity reconciliation retained caution where evidence was broad, cross-disorder, or not clearly schizophrenia-specific. Candidate-specific evidence was therefore preserved as future-audit information rather than converted into biological validation.

Claim-firewall result

The release-state manuscript contained no affirmative claim that SNAP91, OLR1, or GRM1 is a validated schizophrenia gene, biomarker, therapeutic target, causal mechanism, or L2 candidate. This negative firewall result is a reporting safeguard rather than a biological result.

Minimum reproducibility package for peer review

File	Minimum contents	Required status before submission
Supplementary Table S1	All 406 candidates, approved symbol, identifiers, source membership, evidence columns, pass/fail flags.	Required.
Supplementary Table S2	All L1 candidates, scoring routes, reasons for non-promotion, specificity flags.	Required.
Supplementary Table S3	Negative controls and specificity controls with expected and observed behavior.	Required.
Supplementary Table S4	Literature resolver output with PMID/PMCID/Europe PMC IDs, query date, field hit, and disambiguation status.	Required.
Freeze manifest	Data versions, download dates, code commit, package versions, checksums, output hashes.	Required.
Code archive	Executable code or deterministic scripts plus README.	Required.

Discussion

The principal contribution of this work is not the promotion of specific molecular candidates. Its contribution is a conservative audit structure for candidate handling in schizophrenia bioinformatics. In the illustrative frozen run, no candidate reached L2. This outcome is compatible with the central premise: public evidence layers can generate many plausible molecular narratives, but most such narratives should remain below the level of validated biological claims.

The framework addresses a recurrent problem in post-GWAS psychiatric biology: evidence-layer conflation [4][5][6][7]. A locus association may be treated as a gene association; a bulk expression contrast may be treated as a cell-specific mechanism; a literature hit may be treated as validation; and biological familiarity may be treated as discovery. The claim ladder makes these transitions explicit and blocks them unless evidence requirements are met.

SNAP91 illustrates a candidate that may be worth future audit but should not be promoted without independent recurrence and specificity controls. OLR1 illustrates the weakness of annotation-only

support when disease specificity and full-text evidence are not established. GRM1 illustrates the correct handling of a biologically plausible, literature-rich glutamatergic gene: it can serve as a context-control comparator without being reported as a novel discovery.

The method is deliberately conservative. This conservatism reduces false-positive interpretive escalation but may also withhold promotion from true biological candidates that have sparse evidence. Therefore, the framework should be evaluated against positive controls, negative controls, and independent datasets before being used for prioritization decisions beyond audit planning.

Limitations

- The illustrative run does not constitute experimental validation, functional perturbation, animal-model evidence, organoid validation, clinical prediction, or therapeutic target discovery.
- Bulk postmortem transcriptomic evidence cannot establish cell-type-specific mechanism without appropriate deconvolution or single-cell validation.
- Literature metadata resolution cannot replace systematic full-text review.
- The negative result of 0 L2 candidates is only scientifically interpretable if thresholds were prespecified and the complete candidate matrix is deposited.
- Candidate labels for SNAP91, OLR1, and GRM1 should be interpreted as audit labels, not biological conclusions.
- The current rewritten version requires repository deposition of code, candidate matrices, and freeze manifest before peer-review submission.

Conclusions

A claim-governed framework for schizophrenia molecular candidate prioritization can reduce overinterpretation of heterogeneous public evidence. In the illustrative audit, 406 initial candidates were processed, 5 reached L1, and 0 reached L2. SNAP91, OLR1, and GRM1 were retained only for limited audit or control roles. The framework should be viewed as an evidence-governance method that preserves future-validation hypotheses while blocking unsupported claims about validated genes, biomarkers, therapeutic targets, and causal mechanisms.

Abbreviations

Abbreviation	Meaning
E/I	Excitatory/inhibitory.
GWAS	Genome-wide association study.
HGNC	HUGO Gene Nomenclature Committee.
L0–L3	Claim ladder levels used by the audit framework.
PMID	PubMed identifier.

Notes

Short title: Claim-governed schizophrenia candidate prioritization

Statements and Declarations

Funding

No external funding was received for this work.

Potential Competing Interests

The author declares no competing interests.

Ethics

Not applicable. This study used public or derived computational evidence layers and did not involve new human participants, animal experiments, identifiable clinical data, or biological samples.

Data Availability

For peer-review submission, the complete candidate matrix, evidence-layer matrix, resolver outputs, negative-control results, code, README, and freeze manifest must be deposited in a persistent repository. Until those files are available, the results should be interpreted as a non-reproducible illustrative audit rather than an independently verifiable computational result.

Authors Contributions

RMPAF conceived the study, designed the audit framework, performed the original analysis, interpreted the outputs, and is responsible for the scientific content of the manuscript.

Use of Generative AI

AI-assisted editorial revision and critical restructuring were used after the original draft. The author remains responsible for verifying all scientific content, references, data, code, and claims before submission.

References

1. ^*Trubetskoy V, Pardinas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. (2022). "Mapping Genomic Loci Implicates Genes and Synaptic Biology in Schizophrenia." Nature.* ***604****:502–508. doi:10.1038/s41586-022-04434-5.*
2. ^*Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). "Biological Insights from 108 Schizophrenia-Associated Genetic Loci." Nature.* ***511****:421–427. doi:10.1038/nature13595.*
3. ^*Pardinas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. (2018). "Common Schizophrenia Alleles Are Enriched in Mutation-Intolerant Genes and in Regions Under Strong Background Selection." Nat Genet.* ***50****:381–389. doi:10.1038/s41588-018-0059-2.*
4. a, b*Watanabe K, Taskesen E, van Bochoven A, Posthuma D (2017). "Functional Mapping and Annotation of Genetic Associations with FUMA." Nat Commun.* ***8****:1826. doi:10.1038/s41467-017-01261-5.*
5. a, b*Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR, et al. (2015). "Partitioning Heritability by Functional Annotation Using Genome-Wide Association Summary Statistics." Nat Genet.* ***47****:1228–1235. doi:10.1038/ng.3404.*
6. a, b*Nasser J, Bergman DT, Fulco CP, Guckelberger P, Doughty BR, Patwardhan TA, et al. (2021). "Genome-Wide Enhancer Maps Link Risk Variants to Disease Genes." Nature.* ***593****:238–243. doi:10.1038/s41586-021-03446-x.*
7. a, b*Powell SK, Liao W, Ghorbani S, et al. (2026). "3D Genetic Architecture of Schizophrenia Risk Across Three Neuronal Subtypes." Mol Psychiatry.* ***31****:2216–2231. doi:10.1038/s41380-025-03352-y.*
8. ^*Fromer M, Roussos P, Sieberts SK, Johnson JS, Kavanagh DH, Perumal TM, et al. (2016). "Gene Expression Elucidates Functional Impact of Polygenic Risk for Schizophrenia." Nat Neurosci.* ***19****:1442–1453. doi:10.1038/n*

n.4399.

9. ^*Jaffe AE, Straub RE, Shin JH, Tao R, Gao Y, Collado-Torres L, et al. (2018). "Developmental and Genetic Regulation of the Human Cortex Transcriptome Illuminate Schizophrenia Pathogenesis." Nat Neurosci. **21**:1117–1125. doi:10.1038/s41593-018-0197-y.*
10. ^*Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, et al. (2018). "Transcriptome-Wide Isoform-Level Dysregulation in Autism, Schizophrenia, and Bipolar Disorder." Science. **362**:eaat8127. doi:10.1126/science.aat8127.*
11. ^a, ^b*Tricco AC, Lillie E, Zarin W, OBrien KK, Colquhoun H, Levac D, et al. (2018). "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation." Ann Intern Med. **169**:467–473. doi:10.7326/M18-0850.*
12. ^*Little J, Higgins JPT, Ioannidis JPA, Moher D, Gagnon F, von Elm E, et al. (2009). "STrengthening the REporting of Genetic Association Studies (STREGA): An Extension of the STROBE Statement." PLoS Med. **6**:e22. doi:10.1371/journal.pmed.1000022.*
13. ^*Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. (2016). "The FAIR Guiding Principles for Scientific Data Management and Stewardship." Sci Data. **3**:160018. doi:10.1038/sdata.2016.18.*

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.