Online Resource 1. Supplementary notes

Transcriptomic associations of a pre-specified 468-gene redox and xenobiotic-metabolism panel across five neurological disorders, with a published-variant inventory

Mark E. McCaulley, M.D. · Sirius Neuroscience, LLC · Revised 3 October 2026

Machine-readable panel, excluded-symbol list, accession roles and analysis pipeline: Zenodo concept DOI https://doi.org/10.5281/zenodo.22178093 (the locked panel and rules are in version 1.0, https://doi.org/10.5281/zenodo.22178094). The disease-specific companion manuscripts are not part of this submission.

Contents

Table S0. Locked 468-gene panel — caption and tier arithmetic (full list in Online Resource 2).

Table S1. High-confidence overlap across the four adult neurodegenerative analyses (Online Resource 3).

Note S2. Locked primary GEO accessions and per-disease meta counts.

Note S3. Locked high-confidence rule.

Note S6. Definitions and intended use of GVI, DDI, and IDRS.

Table S0. Locked 468-gene panel

The immutable panel used in this article is the 468-gene list deposited at Zenodo (concept DOI https://doi.org/10.5281/zenodo.22178093; freeze date 2026-07-05). The full symbol-by-tier table is Online Resource 2. Genes were assembled by functional-family enumeration, not by data-driven feature selection. Genes were not added or removed on the basis of case-versus-control association. Named subsets reported in the main article are partitions of this same list and are not a revised invention list.

Tier arithmetic (locked)

CORE redox / metal / xenobiotic / barrier	n = 394
CORE cGAS–STING	n = 12
CORE ferroptosis / iron / NRF2	n = 10
SENSITIVITY (broader cellular-defense layer)	n = 40
Placental/sex-steroid module (labelled “ASD-maternal module” in the frozen panel file)	n = 12
Total	**n = 468**

Symbols added in the v458 → v468 freeze (n = 10): AIFM2, CUL3, FXN, GCH1, GLO1, HAGH, MFSD2A, RBX1, TF, TFRC. Twenty-two a priori excluded symbols (non-CNS structural or otherwise out of scope) are listed on sheet S0b of Online Resource 2 and are not panel members.

Governing rule for reuse: no manuscript, table, figure, or companion manuscript may use a gene symbol or count that is absent from the locked panel. Companions cite this Table S0 and the Zenodo DOI; they do not maintain a second panel.

Table S1. High-confidence directional-overlap genes

Table S1 is Online Resource 3. It lists the 53 genes that met the locked high-confidence rule (Note S3) in at least two of the four adult neurodegenerative within-disease meta-analyses (AD, PD, ALS, MS white

matter). Seven genes met the rule in three of those analyses: AKR1C3, CBS, CP, FTL, HMOX1, PINK1 and SLC40A1; none met it in all four. CBS was high-confidence up in AD and PD and high-confidence down in MS white matter; PINK1 was high-confidence down in AD, PD and ALS; CP and FTL were high-confidence up in PD, ALS and MS white matter. The ALS entries use the donor-resolved analysis (Note S2). No cross-disease p-value is attached. Sheet S1b of Online Resource 3 names MAFK as the sole FDR-significant gene in the adult ASD postmortem-brain analysis (1 of 392 analyzable panel genes); MAFK did not meet the high-confidence rule and is not an ASD claim of this article.

Note S2. Locked primary GEO accessions and per-disease meta counts

Locked primary within-disease meta-analyses used only the accessions below. Supporting and model-contrast series listed on the Zenodo GEO_manifest sheet are not added to these primary k counts.

AD brain (k = 6). GSE33000, GSE132903, GSE173955, GSE159699, GSE5281, GSE48350. Meta-analyzable panel genes 458; FDR-significant 167; high-confidence 141; median I^2 among high-confidence genes 79.3%.

PD substantia nigra (k = 4). GSE49036, GSE20141, GSE8397, GSE7621. Meta-analyzable 419; FDR-significant 91; high-confidence 56; median I^2 16.4%. Frontal-cortex series are supporting only.

ALS spinal cord / motor neuron (k = 7). GSE19332, GSE68605, GSE76220, GSE153960, GSE26927, GSE56500, GSE18920. Meta-analyzable 449; FDR-significant 33; high-confidence 28; median I^2 56.7%. Compartment stratification is retained. These figures come from a donor-resolved analysis (one value per donor; the exon-array cohorts GSE56500 and GSE18920 averaged within transcript cluster). The earlier locked figures for the same cohorts (456; 26; 22; median I^2 60.6%) pooled all tissues of GSE153960 and both tissue compartments of GSE18920 and selected exon-array probesets by significance, and are superseded.

MS white matter (k = 3). GSE138614, GSE32915, GSE108000. Meta-analyzable 442; FDR-significant 78; high-confidence 30; median I^2 0%. Gray-matter series are reported separately and are not in this primary k.

ASD adult postmortem brain (k = 4). GSE102741, GSE28521, GSE269105, GSE64018. Meta-analyzable 392; FDR-significant 1 (MAFK); high-confidence 0.

RNA-seq cohorts were analyzed with PyDESeq2 and microarray cohorts with limma. P values were Benjamini–Hochberg FDR-adjusted within the measurable panel of each meta-analysis.

Note S3. Locked high-confidence rule

A gene is high-confidence in a given disease if and only if all four conditions hold in that disease's locked primary meta-analysis:

1. Benjamini–Hochberg FDR < 0.05.

2. Measured in at least three primary cohorts ($k \geq 3$).

3. Direction-concordant with the pooled effect in at least 70% of those cohorts.

4. Independently nominally significant ($p < 0.05$) in at least two cohorts.

The rule has no fold-change floor and no I^2 cap. The $|\log2FC| \geq 0.5$ floor belongs only to the per-dataset DDI (Note S6). The rule is therefore applicable only where primary $k \geq 3$ (AD, PD substantia nigra, ALS, MS white matter). Genes measured in only two cohorts cannot meet it. Primary inference is self-contained at the gene. List-level competitive inference used gene-label permutation of power-matched random 468-gene sets from the measurable background; that test did not show that the full panel out-enriched those random sets.

CAMERA is the parametric competitive counterpart and was not used to define, revise, or interpret the locked panel.

Note S6. Definitions and intended use of the GVI, DDI, and IDRS

The Genetic Vulnerability Index (GVI), Detoxification Dysregulation Index (DDI), and Integrated Detoxification Risk Score (IDRS) are proposed research quantities intended to organize genetic and expression-level evidence. They are not diagnostic scores, calibrated disease-risk estimates, prognostic instruments, treatment-selection tools, or measures of treatment response.

Genetic Vulnerability Index. For an individual carrying a set of qualifying variants V, GVI = Σ ln(OR_v), where OR_v is the published disease-associated odds ratio assigned to qualifying variant v. The corresponding multiplicative quantity is exp(GVI) = Π OR_v. A variant with OR < 1 contributes a negative term. This formulation assumes additivity on the log-odds-ratio scale and does not model linkage disequilibrium, epistasis, ancestry-specific effects, exposure interactions, penetrance, age dependence, or clinical covariates. Only variants with an established odds ratio enter a numerical GVI. Odds ratios are taken from cited source studies and are not re-estimated here.

Detoxification Dysregulation Index. For a given transcriptomic dataset, DDI is expressed as a percent (0–100): DDI = 100 × (N with BH-adjusted $P < 0.05$ and $|\log2FC| \geq 0.5$) / (N measurable panel genes in that dataset). The $|\log2FC| \geq 0.5$ floor applies to this descriptive, per-dataset DDI only; it does not define significance or high confidence in the random-effects meta-analyses. DDI is a tissue- and dataset-level descriptor, not a patient-level measure unless independently validated.

Integrated Detoxification Risk Score. The composite used in this article is IDRS = GVI + DDI/10, where DDI is the percent form defined above. Division of DDI by 10 is a heuristic scaling convention. It is not a clinically calibrated weighting, and IDRS is not an estimated probability, odds ratio, absolute risk, or actionable threshold. A z-score composite of DDI and GVI, if shown elsewhere as an illustration, is a different construct and is not the IDRS of this article.

Worked illustration. If GVI = 1.55 and DDI = 20 (20% of measurable panel genes meet the DDI rule), then IDRS = 1.55 + 20/10 = 3.55. This arithmetic demonstrates only the behavior of the proposed construction. It does not estimate an individual's disease risk or expected benefit from any intervention.

The GVI and DDI arise from different measurement contexts (germline associations versus acquired tissue-level expression). Their combination is a hypothesis-generating organizational model. Prospective, multi-ancestry validation against prespecified clinical outcomes is required before any translational use.

Companion manuscripts

Five disease-resolved companions (AD, ASD, PD, MS, ALS) use this Table S0 and this Note S3 verbatim and cite the Zenodo concept DOI. The AD and PD companions are posted at Research Square (https://doi.org/10.21203/rs.3.rs-10999456/v1 and https://doi.org/10.21203/rs.3.rs-11097203/v1); the ASD, MS and ALS companions are also posted there (https://doi.org/10.21203/rs.3.rs-11553062/v1, https://doi.org/10.21203/rs.3.rs-11530178/v1 and https://doi.org/10.21203/rs.3.rs-11552529/v1). None is part of this submission.