Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Panel-Gene Transcriptomic Associations and Published-Variant Architecture in Parkinson Disease

Mark E. McCaulley

mark@siriusneuroscience.com

Sirius Neuroscience, LLC https://orcid.org/0000-0003-2059-622X

Research Article

Keywords: Parkinson disease, substantia nigra, transcriptomic meta-analysis, xenobiotic metabolism, glutathione, CYP2D6

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11097203/v2

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare potential competing interests as follows: The author is the sole inventor on a filed patent application relating to panel-based research scoring methods that use the locked 468-gene list reported here. The application did not determine gene membership, the high-confidence rule, or the reported results, and the constructs described in Supplementary Note S3 are stated to be unvalidated. Academic reuse of the deposited panel and code does not require a license from the author. The author declares no other competing interests.

Panel-Gene Transcriptomic Associations and Published-Variant Architecture in Parkinson Disease

Short title: PD panel transcriptomics and published-variant architecture

Mark E. McCaulley, M.D. · Sirius Neuroscience, LLC, Steamboat Springs, Colorado 80487, USA
ORCID: 0000-0003-2059-622X Correspondence: mark@siriusneuroscience.com

Keywords: Parkinson disease; substantia nigra; transcriptomic meta-analysis; xenobiotic metabolism; glutathione; CYP2D6

Companion note. Disease-restricted Parkinson disease analysis using the locked 468-gene panel and high-confidence rule of the methods-and-panel manuscript of this series (McCaulley 2026f). Not a pooled cross-disease analysis. Parallel companions apply the same panel and rule to one disease each; the Alzheimer disease companion is posted (McCaulley 2026a) and the others are in preparation (McCaulley 2026c, d, e). A separate methods-and-panel manuscript holds the panel definition, the excluded-symbol list, and the panel-level validation methods, including the negative list-level permutation on the locked 468-gene list. Those methods are not repeated here.

Abstract

Parkinson disease (PD) was examined in four primary adult-brain GEO cohorts (GSE49036, GSE20141, GSE8397, GSE7621) using a frozen 468-gene redox, metal, xenobiotic, and barrier panel. Inverse-variance random-effects meta-analysis (DerSimonian–Laird) produced 419 analyzable genes, 91 at Benjamini–Hochberg FDR < 0.05, and 56 high-confidence genes (FDR < 0.05, measured in ≥3 cohorts, direction-concordant in ≥70% of cohorts, and nominally significant in ≥2 cohorts). Median I^2 among high-confidence genes was 16.4%. In sensitivity analyses that varied probe-set selection, analysed GSE8397 at the donor level and added its second array, 28–41 genes met the same criteria, and 17 did so in every analysis (ALDH1A1, GCH1, GPX3, NDUFS1, GSS, CYP2S1, CES2, COMT, CBS, RELA, CYP2J2, FOXO4, CYP27A1, HMOX1, AKR1C3, CP, MAFF). In two independent substantia nigra cohorts that were not used in the meta-analysis and were examined post hoc, the high-confidence set was direction-concordant with the pooled estimates: in a bulk-nigra array cohort (GSE20292; 11 PD, 18 controls) 41 of 54 measured genes (76%, exact binomial p = 0.0002; Spearman ρ = 0.61), and in an RNA-seq cohort (GSE168496; 8 PD, 8 controls) 38 of 55 (69%, p = 0.006; ρ = 0.43). High-confidence calls included higher abundance of selected iron-handling and stress-response transcripts and lower abundance of selected glutathione- and mitochondrial-related transcripts. These are transcript-abundance findings; they do not themselves prove iron accumulation, glutathione depletion, or complex-I failure. A compiled variant inventory includes PD-relevant published alleles (including the CYP2D6 poor-metabolizer association and PON1 and ABCB1 pesticide-interaction literature). Expression and variant layers are not combined into a clinical score. GVI, DDI, and IDRS remain uncalibrated research formulae and are not used for inference in this paper. No statistical pooling with Alzheimer, ALS, multiple sclerosis, or autism cohorts was performed. This paper is a disease-restricted companion to the methods-and-panel manuscript of this series (McCaulley 2026f).

Introduction

Parkinson's disease is the fastest-growing neurological disorder worldwide (Su et al. 2025), with the Global Burden of Disease Study 2021 projecting 25.2 million cases by 2050. Cardinal motor features — bradykinesia, rigidity, resting tremor, and postural instability — reflect progressive loss of dopaminergic neurons in the substantia nigra pars compacta (SNpc), marked pathologically by α-

synuclein in Lewy bodies (Bloem et al. 2021). Prevailing paradigms emphasize α-synuclein aggregation and prion-like spreading (Spillantini et al. 1997), mitochondrial complex I deficiency (Kverneng et al. 2025; Schapira et al. 1990), oxidative stress (Jenner 2003), and neuroinflammation (Bloem et al. 2021). None fully explains selective SNpc vulnerability, variable age of onset, or the gene–environment interactions that modulate PD risk (Ahmed et al. 2017; Paul et al. 2016; Surmeier et al. 2017).

A prior Alzheimer paper using a smaller panel established the starting analytical framework (McCaulley 2026b). Disease-restricted companions apply the locked 468-gene panel and the same high-confidence rule to each disease. The posted Alzheimer disease companion reports 141 high-confidence genes (McCaulley 2026a). The locked workbook deposited with this series (Zenodo version 1.1, CONSOLIDATED_468_FINAL_7.6.26.xlsx) records 22 high-confidence genes in amyotrophic lateral sclerosis, 30 in multiple sclerosis, and none in the adult-brain meta-analysis of autism spectrum disorder; the companions reporting those three analyses are in preparation (McCaulley 2026c, d, e). The ALS companion has since re-analysed those seven cohorts at the donor level and reports 33 genes at FDR < 0.05 and 28 high-confidence, superseding the workbook figure of 22 (McCaulley 2026d). Those counts are not reused here as evidence for a pooled mechanism. Older working figures (for example an 81% ASD hit rate) are superseded by the locked workbook and are not claimed. Cross-disease gene-symbol overlap is descriptive only.

Several published features of PD make a pre-specified detoxification, metal, and barrier panel a reasonable test set rather than a discovery set. Glutathione (GSH) depletion in the SNpc is among the earliest biochemical changes reported in PD and precedes dopamine loss in the cited studies (Sian et al. 1994; Smeyne and Smeyne 2013; Sofic et al. 1992). The CYP2D6 poor-metabolizer phenotype carries a replicated PD association in two large variant meta-analyses (Lu et al. 2013; Lu et al. 2014). Gene–environment interactions between detoxification-gene variants and pesticide exposure produce some of the largest published odds ratios in PD epidemiology (Ahmed et al. 2017; Dutheil et al. 2010; Lee et al. 2013; Manthripragada et al. 2010; Narayan et al. 2015; Paul et al. 2016). Iron accumulates several-fold in neuromelanin-positive SNpc neurons, with published metallothionein responses (Carmona et al. 2024; Ebadi et al. 2005; Glaab and Schneider 2015; Michael et al. 2011; Miyazaki and Asanuma 2023). Nuclear NRF2 localization and HMOX1/NQO1 induction have been reported in surviving dopaminergic neurons (Ramsey et al. 2007). Those biochemical, imaging, and genetic findings are context. They are not outcomes measured in the present GEO contrasts.

This paper applies the locked panel and high-confidence rule to PD brain and places published PD-associated variants beside the transcript results. Primary inference is gene-level. A list-level permutation test of the locked 468-gene list is negative in the methods-and-panel manuscript (McCaulley 2026f) and is not re-interpreted here as enrichment. GVI, DDI, and IDRS are described only in the Supplementary Note and are not used to rank genes or assign patient-level risk.

Methods

Panel

The 468-gene panel is a curated, pathway-anchored gene set selected on mechanistic and pharmacogenomic grounds, not by data-driven feature selection within any PD dataset (McCaulley 2026f). This a priori definition avoids the circularity of selecting genes from the same contrasts used to test them. The panel evolved through documented intermediate builds from the 203-gene foundation in my prior Journal of Alzheimer's Disease Reports analysis (McCaulley 2026b) to the frozen 468-gene canonical panel (v468) applied uniformly across the five companion diseases. The

ten genes added in v468 relative to the immediately prior 458-gene panel are AIFM2, CUL3, FXN, GCH1, GLO1, HAGH, MFSD2A, RBX1, TF, and TFRC. The README of the frozen panel file lists nine of these ten and omits FXN; all ten form the file's CORE ferroptosis/iron/NRF2 tier. The panel's history was not purely additive: genes were both added and removed at the 389-gene reconciliation. The frozen v468 file is the only panel definition used in this paper, and the symbols excluded a priori are listed in the excluded-symbol file of the Zenodo deposit (version 1.0, https://doi.org/10.5281/zenodo.22178094). Because these ten genes entered the panel last, I repeated the locked PD meta-analysis without them, recomputing the Benjamini–Hochberg correction across the remaining 409 analyzable genes and leaving the high-confidence rule unchanged. Eighty-nine genes met FDR < 0.05 and 54 met the high-confidence rule; the only genes lost from either list were GCH1 and TF themselves, and no other gene changed FDR or high-confidence status.

A preliminary meta-analysis of the same four PD cohorts was run on an earlier 238-gene build of the panel before the panel was frozen (5 July 2026) and the analysis locked (6 July 2026). Twenty of the 56 high-confidence genes reported here entered the panel after that run: ADH1B, AKR1C3, ALDH6A1, BLVRA, CBS, CES2, CYP2J2, CYP2S1, CYP39A1, DDX41, GCH1, MAFF, NOTCH1, SLC22A3, SLC2A6, SLCO1A2, SLCO4A1, STS, TF, and TXNIP. Their PD values had not been examined before they were added. Restricted to the 222 genes common to both runs, the locked analysis yields 59 FDR-significant and 37 high-confidence genes: the 36 locked high-confidence genes in that set, plus SQSTM1. In the sensitivity analyses described below, the 20 later additions were no less robust than the 36 earlier genes (robust 7 versus 10; not robust 3 versus 5; Supplementary Table S5).

The panel is organized into five functional hubs on the frozen 468-gene list: Hub 1, xenobiotic and morphogen metabolism (230 genes); Hub 2, heavy-metal homeostasis (49 genes); Hub 3, neuro-vascular integrity (40 genes); Hub 4, antioxidant defense and redox homeostasis (92 genes); Hub 5, transcriptional regulation of detoxification (57 genes). Hub membership is an organizing device for display in this paper; it is not part of the locked panel definition in the methods-and-panel manuscript (McCaulley 2026f), and it is not a pathway-enrichment test.

Primary cohorts

I assembled publicly available transcriptomic datasets from the Gene Expression Omnibus (GEO). Primary inference is a four-cohort adult-brain random-effects meta-analysis of GSE49036, GSE20141, GSE8397, and GSE7621 (Table 1). Those four series are hybridization-array studies pooled at the per-gene effect-size level. They are not an RNA-seq pool. GSE68719 is a supporting single-cohort RNA-seq contrast. GSE20292 (array) and GSE168496 (RNA-seq) are substantia nigra cohorts that contributed nothing to the locked meta-analysis and are used only as independent checks of the high-confidence set. None of the three is a substitute for the locked set. I did not combine array and RNA-seq platforms into one index.

Table 1. Locked primary cohorts and supporting contrasts.

Accession	Role	Tissue	Platform	Design (as deposited)	Source
GSE7621	Locked	Substantia nigra	Affymetrix HG-U133 Plus 2.0 (GPL570)	16 PD / 9 control, bulk nigra	Lesnick et al. (2007); GEO
GSE49036	Locked	Substantia nigra	Affymetrix HG-U133 Plus 2.0 (GPL570)	28 arrays as deposited: 8 control (Braak 0), 5 iLBD (Braak 1–2), 7 PD (Braak 3–4), 8 PD (Braak 5–6); contrast used: PD, Braak 3–6 (n = 15) versus Braak 0 control (n = 8); iLBD	Dijkstra et al. (2015)

Accession	Role	Tissue	Platform	Design (as deposited)	Source
				excluded	
GSE8397	Locked	Medial and lateral substantia nigra; superior frontal gyrus	Affymetrix HG-U133A (GPL96); HG-U133B (GPL97) not used in the locked analysis	47 A-chip arrays as deposited: superior frontal gyrus 5 PD / 3 control, lateral SN 9 PD / 7 control, medial SN 15 PD / 8 control (SN controls include disease-control case MS155); contrast used: all PD versus all control A-chip arrays across the three regions, 29 PD versus 16 control arrays from 15 PD and 7 control donors, MS155 excluded, arrays treated as independent	Moran et al. (2006)
GSE20141	Locked	Laser-dissected SNpc neurons	Affymetrix HG-U133 Plus 2.0 (GPL570)	10 PD / 8 control neuronal isolates	Zheng et al. (2010)
GSE68719	Supporting	Prefrontal cortex (BA9)	RNA-seq	29 PD / 44 control; n = 73 in the covariate-adjusted screen	Dumitriu et al. (2016)
GSE20292	Independent check	Whole substantia nigra	Affymetrix HG-U133A (GPL96)	11 PD / 18 control as deposited; 3 control arrays re-scaled to log2 (see Methods)	Zhang et al. (2005)
GSE168496	Independent check	Substantia nigra	RNA-seq	Post-mortem SN, 16 samples as deposited	Tranchevent et al. (2023b)

The contrasts in Table 1 are those that reproduce the locked per-gene values deposited with this series (https://doi.org/10.5281/zenodo.22846772); they were confirmed by re-deriving all four cohort contrasts from the GEO series matrices (Supplementary Table S5). The earlier version of this paper described the GSE49036 contrast as Braak 5–6 versus Braak 0 and the GSE8397 contrast as lateral substantia nigra only, using both arrays; those descriptions did not match the analysis that produced the reported results and are corrected here. The meta-analysis run with the contrasts as previously described is reported in Supplementary Table S5 for comparison. GSE7621 standalone counts in Table 4 match the GEO record (16 PD / 9 control).

Meta-analysis and high-confidence rule

The primary inferential result is a four-cohort inverse-variance random-effects (DerSimonian–Laird) meta-analysis of author-processed PD-versus-control contrasts from the locked four series, pooled on the log2 fold-change scale. Per-cohort contrasts were computed with GEO2R, which fits limma linear models with empirical-Bayes moderated t-statistics (Barrett et al. 2013; Ritchie et al. 2015; Smyth 2004), without covariate adjustment, on the submitter-processed series matrices; the MAS5-processed series GSE20141 and GSE7621 were log2-transformed by GEO2R's automatic rule. Where a gene was represented by more than one probe set, the probe set with the lowest nominal p value in that cohort was retained; probe sets annotated to more than one gene symbol were excluded. Per-cohort standard errors were taken as |logFC / t|. DESeq2 lfcSE was not used for the locked array meta-analysis. A gene is high-confidence when Benjamini–Hochberg FDR < 0.05, it is measured in ≥3 of the four locked cohorts, direction is concordant in ≥70% of contributing cohorts,

and it is nominally significant ($p < 0.05$) in ≥2 cohorts; for that criterion, cohort-level p values were computed from the normal approximation, $2\Phi(-|\text{logFC} / \text{SE}|)$. That rule is locked across the series (McCaulley 2026f) and is not relaxed here.

Per-dataset descriptive counts restricted to panel genes at BH FDR < 0.05 and |log2FC| ≥ 0.5 are reported for transparency for GSE68719 and for GSE7621 as a standalone locked cohort. Those percentages are not pooled across diseases and are not used to declare list-level enrichment. Forest-plot confidence intervals in Figure 1 were computed as pooled estimate ± 1.96 × SE, with SE recovered as |pooled| / z(p) because the deposited per-gene file does not carry standard errors; GSS is drawn without an interval because its deposited p value underflows to zero.

Reproduction and sensitivity analyses

The locked result was first reproduced from the GEO series matrices with an independent implementation of the GEO2R two-group limma model (lmFit and eBayes, robust = FALSE) and the pooling and high-confidence rules above. The reproduction recovered 418 of the 419 analyzable genes (AFDN is carried under its former symbol, MLLT4, in the platform annotation used), pooled log2 fold-changes within 0.001 for 405 of them, 92 FDR-significant genes (91 in the lock; SULT4A1 differs by probe annotation), and the identical 56-gene high-confidence set.

Three features of the locked pipeline can inflate cohort-level effect sizes or precision: selection of the lowest-p probe set within each cohort, treatment of multiple GSE8397 arrays from the same donors as independent, and assessment of cohort-level significance by a normal approximation. The meta-analysis was therefore repeated under four alternative specifications, each using limma moderated-t p values for the cohort-level criterion: (A) the per-sample mean of all probe sets for each gene; (B) the single probe set with the highest average expression, a selection independent of the PD–control comparison (Miller et al. 2011); and (C, D) specifications A and B with GSE8397 restricted to substantia nigra, averaged within donor (15 PD, 7 control donors), and extended with its HG-U133B arrays. All other steps were unchanged. A gene was classed as robust if it met the high-confidence criteria under all four specifications, and as partly robust, fragile, or not robust if it did so under two or three, one, or none. These analyses were added after the panel and analysis plan were locked; they are reported as sensitivity analyses, and the locked analysis remains the primary result.

Published-variant inventory

Variants in Table 5 are a literature inventory. No variant was genotyped in the present GEO samples. Expression and variant layers are not combined into a score in the main analysis. Inclusion used a tiered minor-allele-frequency framework: common published PD-associated alleles (MAF ≥ 1% in any major reference population) and rare but published large-effect alleles. Formulae previously labelled GVI, DDI, and IDRS are confined to Supplementary Note S3 and are not calibrated against clinical outcomes.

Statistics and software

Random-effects pooling used the DerSimonian–Laird estimator on the log2 fold-change scale (DerSimonian and Laird 1986). Heterogeneity is reported as I^2 for high-confidence genes (median 16.4%). Multiple-testing control within the measurable panel genes used Benjamini–Hochberg FDR (Benjamini and Hochberg 1995). Probe-to-symbol mapping used the GEO platform annotation (GPL570 for GPL570 and GPL96 probe sets; GPL97 for HG-U133B probe sets in the sensitivity analyses). The reproduction and sensitivity analyses used a Python implementation of the two-group limma model; the scripts accompany Supplementary Table S5. Bulk tissue was not cell-type deconvolved. Pipeline files accompanying the locked panel are in the same Zenodo record.

Independent-cohort checks of the high-confidence set

Two substantia nigra cohorts that contributed nothing to the locked meta-analysis were used to check the 56 high-confidence genes. GSE20292 (Zhang et al. 2005) is a whole-substantia-nigra Affymetrix HG-U133A series of 11 PD donors and 18 controls. Its series matrix holds 26 RMA-normalized log2 arrays and three control arrays (GSM606624–GSM606626, added to the series in 2010 and not part of the 2005 report) deposited on a linear scale; the three were log2-transformed, all 29 arrays were quantile-normalized together, and PD-versus-control log2 fold-changes were computed per probe with a Welch t-test, taking the probe with the smallest p value for each gene. Because PD donors were older (mean 75.5 versus 66.8 years), each probe was also fitted by ordinary least squares with age and sex as covariates, and the adjusted estimates are reported alongside. As a further sensitivity analysis, the comparison was repeated on the 26 original arrays alone, excluding the three late-added arrays and without renormalization. GSE168496 (Tranchevent et al. 2023b) comprises bulk substantia nigra RNA-seq from 8 PD donors and 8 non-demented controls; group assignment was taken from the GEO series matrix. The primary comparison used the PD-versus-control DESeq2 results generated by NCBI GEO2R from NCBI-produced gene counts, so that no author-side processing choice enters the check; CYP2E1 is absent from that table and is reported as not measured. As a sensitivity analysis, gene-level abundance was also obtained by summing the transcript-level values of the deposited preprocessed matrix over Ensembl transcripts of each gene, with log2 fold-change computed as log2 of the ratio of group means (+1 offset) and a Welch t-test on log2(x + 1). In each cohort, direction concordance of the high-confidence genes with the pooled estimate was tested against 50% by exact binomial test, rank agreement by Spearman correlation, and concordance was examined separately for genes lower and higher in PD. Per-gene values for both cohorts and all routes are in Supplementary Table S4. These checks test direction and rank agreement with the locked signature; they did not alter any locked count.

Results

Primary four-cohort meta-analysis

Of 468 panel genes, 419 were analyzable, that is, measured in at least two of the four locked cohorts so that a pooled estimate could be formed; the high-confidence rule additionally requires measurement in at least three. Ninety-one genes met FDR < 0.05. Fifty-six met the locked high-confidence rule (Table 2; Figure 1). Thirty-six of the fifty-six are higher in PD and twenty are lower. Median I^2 among the high-confidence genes is 16.4%. The full FDR-significant and high-confidence lists are the per-gene tables deposited at https://doi.org/10.5281/zenodo.22846772 and summarized in Supplementary Table S2 for every high-confidence gene named with a point estimate in this text.

The 56 genes organize across the five hubs into a two-directional transcript pattern. Iron-handling and NRF2-axis transcripts are higher: ceruloplasmin (CP +0.98), metallothioneins MT1G +0.86, MT1F +0.85, and MT1X +0.74, transferrin (TF +0.57), ferroportin (SLC40A1 +0.52), HMOX1 +0.66, and NFE2L2 +0.50. Glutathione-synthesis and selected mitochondrial transcripts are lower: GSS −0.62, GPX3 −0.66, GCH1 −0.96, NDUFS1 −0.62, and PINK1 −0.39. The largest single pooled change among named high-confidence genes is ALDH1A1 at −1.66, the aldehyde dehydrogenase that oxidizes the dopamine metabolite DOPAL. Additional named high-confidence transcripts include AKR1C3 +0.95, ADH1B +0.73, COMT +0.40, TXNIP +0.81, DNAJB1 +0.78, MAFF +1.39, FOXO1 +0.50, FOXO3 +0.41, PPARA +0.50, RELA +0.41, SLCO4A1 +1.39, TJP2 +0.49, SLCO1A2 +0.44, NOTCH1 +0.44, STS −0.68, GSTA4 −0.46, ATP7A −0.23, SLC2A6 −0.52, DDX41 −0.36, and ZNF580 −0.33. Four of the genes named in this paragraph (TF, SLC40A1, SLCO1A2, and ATP7A) did not meet the high-confidence criteria in any sensitivity analysis and are not interpreted further (see Sensitivity of the high-confidence set).

These are bulk-tissue transcript-abundance findings. They do not measure iron content, glutathione concentration, or respiratory-chain activity in these cohorts. Higher metallothionein or HMOX1 transcript does not establish ferroptosis. CLDN5 is a panel member with a pooled estimate of +0.35 (FDR 0.04) in this PD meta-analysis; that is not a measurement of barrier function. CYP26A1 (+0.06, FDR 0.64), CYP26B1 (−0.14, FDR 0.80), and NOS2 (+0.07, FDR 0.62) do not meet significance in the locked PD pool.

Table 2. High-confidence genes of the four-cohort PD random-effects meta-analysis, by functional hub (56 genes).

Functional hub	HC (n)	Up	Down	Representative high-confidence genes (pooled log2FC)
Hub 1 — Xenobiotic and morphogen metabolism	15	9	6	AKR1C3† +0.95, ADH1B +0.73, COMT† +0.40; ALDH1A1† −1.66, STS −0.68, GSTA4 −0.46
Hub 2 — Heavy-metal homeostasis	9	8	1	CP† +0.98, MT1G +0.86, MT1F +0.85, MT1X +0.74, TF‡ +0.57, SLC40A1‡ +0.52; ATP7A‡ −0.23
Hub 3 — Neuro-vascular integrity	7	6	1	SLCO4A1 +1.39, TJP2 +0.49, SLCO1A2‡ +0.44, NOTCH1 +0.44; SLC2A6 −0.52
Hub 4 — Antioxidant defense and redox	13	5	8	TXNIP +0.81, HMOX1† +0.66, NFE2L2 +0.50, DNAJB1 +0.78; GCH1† −0.96, GPX3† −0.66, GSS† −0.62, NDUFS1† −0.62
Hub 5 — Transcriptional regulation	12	8	4	MAFF† +1.39, FOXO1 +0.50, PPARA +0.50, RELA† +0.41, FOXO3 +0.41; DDX41 −0.36, ZNF580 −0.33

BH-adjusted p < 0.05 for all listed genes. Positive log2FC = higher in PD. The complete 56-gene and 91-gene lists are in the Zenodo per-gene file. † High-confidence in all four sensitivity analyses; ‡ high-confidence in none (Supplementary Table S5).

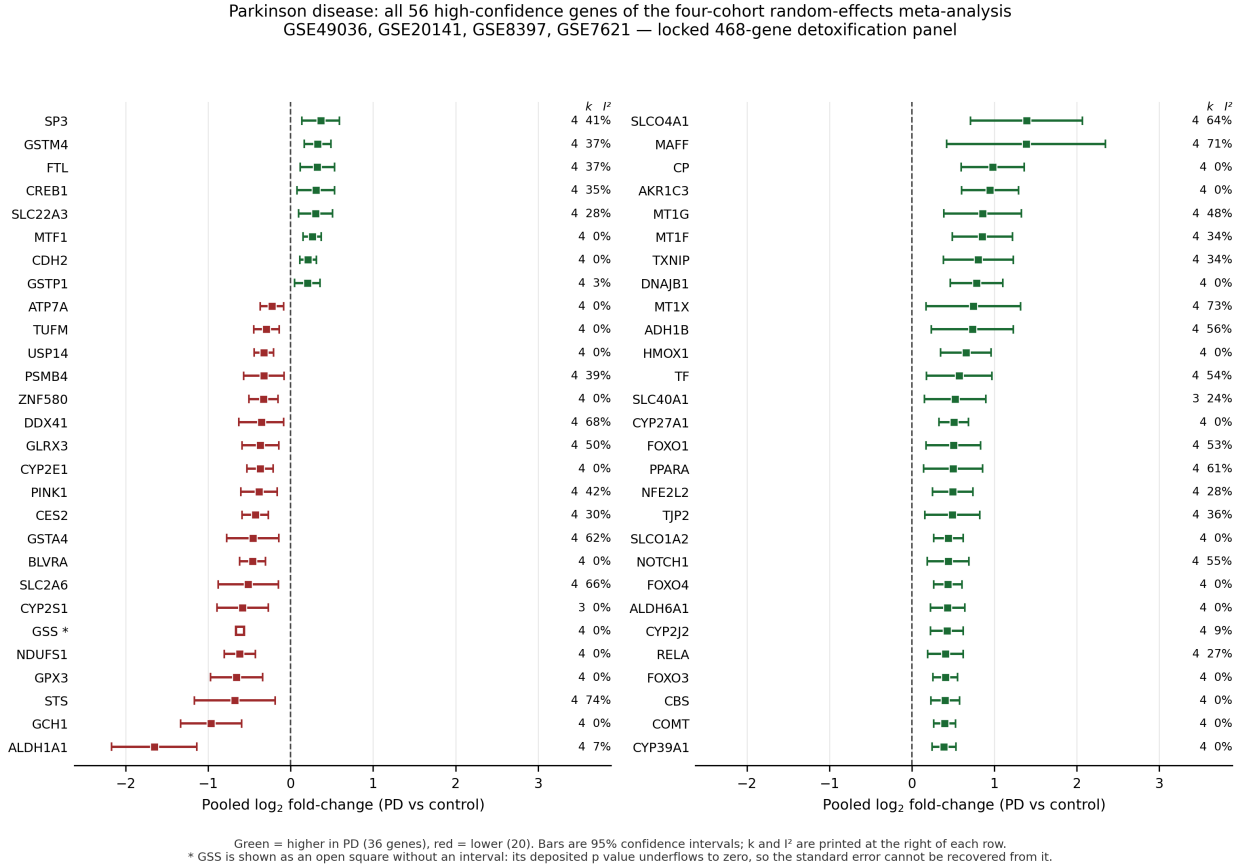


Figure 1. All 56 high-confidence genes of the four-cohort Parkinson disease random-effects meta-analysis (DerSimonian–Laird) of the locked 468-gene panel across GSE49036, GSE20141, GSE8397, and GSE7621, ordered by pooled effect. A gene is high-confidence when it is FDR-significant, measured in at least three cohorts, direction-concordant in at least 70% of contributing cohorts, and nominally significant in at least two. Points are pooled log2 fold-changes; bars are 95% confidence intervals computed as the pooled estimate ± 1.96 × SE, with SE recovered as |pooled| / z(p) because the deposited per-gene file does not carry standard errors; k and I^2 are printed at the right of each row. Thirty-six genes are higher in PD and twenty lower. GSS is drawn as an open square without an interval because its deposited p value underflows to zero. Median I^2 across the 56 is 16.4%. Values are those deposited at

https://doi.org/10.5281/zenodo.22846772. The robustness of each gene to the sensitivity analyses is given in Supplementary Table S5.

Sensitivity of the high-confidence set

The locked count depends on analytic choices (Supplementary Table S5). Under the four sensitivity specifications, 79–89 genes were FDR-significant and 28–41 met the high-confidence criteria; most of the attrition came from probe-set selection, and replacing the lowest-p probe set with the highest-expressed probe set alone reduced the high-confidence count to 41. With the GSE49036 and GSE8397 contrasts as described in the earlier version of this paper, 73 genes were FDR-significant and 45 high-confidence (44 when cohort P values are taken from the limma moderated t, as in the four specifications; Supplementary Table S5).

Seventeen genes met the high-confidence criteria under every specification. Seven were lower in PD (ALDH1A1, GCH1, GPX3, NDUFS1, GSS, CYP2S1, and CES2) and ten higher (COMT, CBS, RELA, CYP2J2, FOXO4, CYP27A1, HMOX1, AKR1C3, CP, and MAFF), and their pooled effects changed little across specifications (for example, ALDH1A1 −1.66 locked and −1.71 under specification D). A further 24 locked high-confidence genes, including NFE2L2, PINK1, MT1F, MT1G, and DNAJB1, met the criteria under two or three specifications, and seven under one. Eight met them under none: TF, SLCO1A2, CYP39A1, SLC40A1, and TUFM were not FDR-significant under the corrected specifications, and SP3, CREB1, and ATP7A remained FDR-significant but were not nominally significant in two or more cohorts. NFE2L2 was robust to single-probe selection (pooled log2FC 0.46–0.48, FDR ≤ 0.003) but attenuated when all probe sets were averaged (0.28–0.29, FDR 0.21).

Supporting single-cohort contrasts

In the covariate-adjusted screen of GSE68719 (prefrontal cortex, n = 73; supporting, not a locked primary cohort), 20 panel genes were significant under the descriptive rule used for single-cohort counts (BH FDR < 0.05 and |log2FC| ≥ 0.5), including seven metallothioneins plus HSPA1A and DNAJB1 (Table 3). That modest single-cohort signal is below an earlier covariate-unadjusted estimate. GSE168496 yielded one panel gene on the same descriptive rule (ALDH1A1). GSE7621 is one of the four locked meta cohorts; its standalone descriptive count was 0 (Table 4) and is shown only for transparency. A single-cohort null cannot distinguish limited array sensitivity from absent signal. GSE7621 still contributes to the four-cohort pool.

Table 3. Supporting single-cohort descriptive count for GSE68719 (not a locked primary meta cohort).

Dataset	Tissue	n	Measurable	DEGs	Up	Down	Notes
GSE68719	PFC BA9	73	399	20	18	2	Covariate-adjusted (PMI, age, RIN); supporting RNA-seq

GSE168496 (RNA-seq SN) had one panel gene meeting the same descriptive rule (ALDH1A1, log2FC −2.26) and is not tabulated as a single-cohort count; its use, with GSE20292, as an independent check of the high-confidence set is reported below. Primary inference remains the four-cohort meta-analysis (91 FDR / 56 high-confidence).

Table 4. Standalone descriptive count for GSE7621, one locked meta cohort, shown on its own.

Dataset	Tissue	Platform	Samples	Measurable	DEGs
GSE7621	SN	Affymetrix array	16 PD / 9 control	422	0

Standalone counts are descriptive. GSE7621 measures 422 of the 468 panel genes, so its null count is not a coverage artefact; with 25 arrays it cannot distinguish limited sensitivity from absent signal.

Independent-cohort checks of the high-confidence set

In GSE20292, 54 of the 56 high-confidence genes were measurable on the HG-U133A array (CYP2S1 and SLC40A1 are absent). Forty-one of the 54 (76%) changed in the same direction as the pooled estimate (exact binomial p = 0.0002), and the GSE20292 log2 fold-changes correlated with the pooled estimates in rank (Spearman ρ = 0.61, p = 1 × 10^{-6}; Pearson r = 0.73; Figure 2A). Twenty-two genes were nominally significant in the concordant direction, including ALDH1A1 (−1.59, p = 0.003),

NDUFS1 (−1.09, p = 0.0004), PSMB4, BLVRA, USP14, PINK1 and GSS among the lower genes and MAFF (+1.22), SLCO4A1 (+1.10), TF (+0.92), CYP27A1 (+0.70), FOXO1 (+0.69), FOXO3 and RELA among the higher genes. Concordance was 17 of 19 for genes lower in PD and 24 of 35 for genes higher in PD (p = 0.041). Adjustment for age and sex attenuated but did not remove the result (37 of 54 concordant, 69%, p = 0.009; ρ = 0.39, p = 0.004; r = 0.54), and adjusted and unadjusted estimates agreed closely (ρ = 0.83 between them). Direction concordance was unchanged when the three late-added arrays were excluded entirely (26 original arrays, no renormalization: 41 of 54 concordant, p = 0.0002), although the rank correlation was lower (ρ = 0.46, p = 0.0004; r = 0.63). The dopaminergic markers TH, SLC6A3 and DDC were reduced as expected (log2FC −1.10 to −1.61), so the cohort shows the neuronal loss characteristic of PD nigra; the panel result is not an artefact of its absence.

In GSE168496, 55 of the 56 high-confidence genes were measured. Thirty-eight of the 55 (69%) changed in the same direction as the pooled estimate (exact binomial p = 0.006), and the log2 fold-changes correlated with the pooled estimates in rank (Spearman ρ = 0.43, p = 0.001; Figure 2B). ALDH1A1 was the only gene individually significant after correction (log2FC −2.26, padj < 0.001), and GCH1, GSS, CYP2S1 and TUFM were nominally significant in the concordant direction; at n = 16, per-gene significance is not expected and is not claimed. Concordance was 17 of 19 for genes lower in PD (89%, p = 0.0007) but 21 of 36 for genes higher in PD (58%, p = 0.41); the discordant higher-in-PD genes were concentrated among the iron-handling and metallothionein transcripts (MT1G, MT1X, TF, SLC40A1) and also included MAFF and TXNIP; several of these replicated in GSE20292. The transcript-sum sensitivity route gave similar rank agreement (ρ = 0.41, p = 0.002; ρ = 0.93 between the two routes) but weaker direction concordance (35 of 56, 63%, exact binomial p = 0.08). Across the two cohorts, the lower-in-PD arm of the signature replicated almost completely in both, and the higher-in-PD arm replicated in the array cohort but only partly in the RNA-seq cohort (Supplementary Table S4).

Figure 2. Independent-cohort checks of the 56 high-confidence genes in two substantia nigra cohorts that contributed nothing to the locked meta-analysis. Each point is one gene, labelled by symbol: pooled log2 fold-change from the four locked cohorts on the horizontal axis, external-cohort log2 fold-change (PD versus control) on the vertical axis. Orange, higher in PD in the pooled analysis; blue, lower. Filled symbols are direction-concordant, open symbols discordant; the dashed line is identity. (A) GSE20292, whole substantia nigra HG-U133A array, 11 PD and 18 controls, all arrays on a common log2 scale and quantile-normalized (unadjusted estimates; age- and sex-adjusted estimates are in Supplementary Table S4); 54 genes shown, 41 concordant (exact binomial p = 0.0002), Spearman ρ = 0.61. (B) GSE168496, substantia nigra RNA-seq, 8 PD and 8 controls, NCBI GEO2R DESeq2 estimates; 55 genes shown (CYP2E1 absent), 38 concordant (p = 0.006), ρ = 0.43. The vertical axis of panel B is limited to −1.2 to 0.9 so that the clustered genes remain legible; ALDH1A1 (−2.26) lies below that range and is drawn as a triangle at the axis floor with its value stated. Per-gene values for both cohorts are in Supplementary Table S4.

Published-variant architecture

Table 5 lists published PD-associated alleles that map onto panel genes or onto closely related xenobiotic, conjugation, transport, antioxidant, or iron-handling loci. Individual main-effect odds ratios are generally modest (approximately 1.1–1.4). Larger published effects appear in gene–environment analyses: PON1 genotype with organophosphate exposure (Lee et al. 2013; Manthripragada et al. 2010), NOS1 variants with organophosphate exposure (Paul et al. 2016), and ABCB1 variants with organophosphate or organochlorine exposure (Dutheil et al. 2010; Narayan et al. 2015). Those interaction estimates come from retrospective case-control studies and are susceptible to recall bias and exposure misclassification. They are not estimated from the present transcript data.

CYP2D6*4 (rs3892097) and the poor-metabolizer phenotype have replicated PD associations in two variant meta-analyses (Lu et al. 2013; Lu et al. 2014). GSTM1 null has a modest Caucasian-specific association in published GST meta-analyses (Liu et al. 2015; Wang and Wang 2014; Weikang et al. 2016). SOD2 p.V16A (rs4880) shows a recessive association in a Han Chinese meta-analysis (OR 1.77, 95% CI 1.15–2.71) (Wu et al. 2021); that estimate is population-restricted and is not treated here as a universal main effect. HFE and NQO1 rows are carried as inventory items without a pooled OR. No row in Table 5 was genotyped in these GEO cohorts.

Table 5. Published PD-relevant variants on or adjacent to the locked panel (inventory only).

Gene / variant	Category	Published PD association	Reference
CYP2D6*4 (rs3892097)	Phase I	OR 1.28 (1.14–1.43) for A allele; 22 studies	Lu et al. 2013
CYP2D6 poor-metabolizer	Phase I	OR 1.33 (1.01–1.74) vs extensive	Lu et al. 2014

Gene / variant	Category	Published PD association	Reference
phenotype		metabolizers; 29 studies	
GSTM1 null	Phase II	OR 1.11–1.16 in Caucasian meta-analyses	Liu et al. 2015; Wang and Wang 2014; Weikang et al. 2016
GSTT1 null	Phase II	Modest; reported mainly in combination analyses	Liu et al. 2015; Wang and Wang 2014
GSTP1 Ile105Val	Phase II	Variable; context-dependent; no meta-analysis cited here	—
PON1 Q192R / MM	Organophosphate hydrolysis	Modifier with OP exposure (OR ~2.5–3.3 in cited studies)	Lee et al. 2013; Manthripragada et al. 2010
ABCB1 C3435T (rs1045642)	BBB efflux	Interaction with pesticides (OR 3.0–7.9 in cited studies)	Dutheil et al. 2010; Narayan et al. 2015
SOD2 p.V16A (rs4880)	MnSOD targeting	Recessive CC OR 1.77 (1.15–2.71) in Han Chinese meta-analysis	Wu et al. 2021
NQO1 C609T	Antioxidant	No significant main-effect meta-analysis cited here	—
HFE C282Y / H63D	Iron homeostasis	Variable; possible age-of-onset modifier; no pooled OR assigned	—

This is a proposed inventory. Protective combination genotypes reported in a single GST multi-locus subgroup analysis (Fan et al. 2022) require replication and are not promoted to a cross-disease "vulnerability core."

Discussion

What the transcripts do and do not show

The high-confidence set recovers, as transcript abundance, gene systems that the PD literature has measured by other means: iron handling (Carmona et al. 2024; Ebadi et al. 2005; Glaab and Schneider 2015; Michael et al. 2011; Miyazaki and Asanuma 2023), NRF2-axis engagement (Cook et al. 2011; Kaidery et al. 2013; Ramsey et al. 2007), glutathione synthesis (Garrido et al. 2011; Shukla et al. 2023; Sian et al. 1994; Smeyne and Smeyne 2013; Sofic et al. 1992), and complex I–related genes (Chinta and Andersen 2006; Flønes et al. 2024; Haas et al. 1995; Keeney et al. 2006; Kverneng et al. 2025; Schapira et al. 1990). Alignment with prior systems-level analyses of parkinsonian brain (Dumitriu et al. 2016; Lesnick et al. 2007; Tranchevent et al. 2023a) is descriptive. The present design cannot convert a higher CP or MT1G transcript into a claim of nigral iron accumulation, or a lower GSS transcript into a claim of GSH depletion, in these samples. Sian et al. (1994) and Sofic et al. (1992) measured glutathione. Schapira et al. (1990) and later complex-I stratification studies (Flønes et al. 2024) measured respiratory-chain function. Those papers are cited as context, not as results of this meta-analysis.

ALDH1A1 at −1.66 is the largest named high-confidence change. ALDH1A1 oxidizes DOPAL. That biochemical role is established in the literature; DOPAL was not measured here. Lower PINK1 and NDUFS1 transcripts are compatible with published mitochondrial narratives and remain transcript calls.

Interpretation in this paper rests mainly on the 17 genes that were robust to every sensitivity specification. The iron-handling component rests on CP and HMOX1, which were robust; transferrin and ferroportin were not. The glutathione and mitochondrial calls rest on GSS, GPX3, GCH1, and NDUFS1, which were robust, and on PINK1, which was partly robust. ALDH1A1 was robust and

replicated in both independent cohorts. Several features of the locked pipeline favour detection over precision; they are retained in the primary analysis because it was fixed before they were examined, and they are reported side by side with the alternatives so that readers can judge which results depend on them.

The two independent checks support the locked signature in direction and rank. In GSE20292 three-quarters of the high-confidence genes moved with the pooled estimate and the rank correlation (ρ = 0.61) is of the same order as the external replications reported for the other diseases in the methods-and-panel manuscript (McCaulley 2026f); in GSE168496 the agreement was weaker (ρ = 0.43) and confined largely to the genes lower in PD. The genes lower in PD, the glutathione-synthesis and mitochondrial calls among them, replicated almost completely in both cohorts, and ALDH1A1 replicated in both, at nominal significance in GSE20292 and after FDR correction in GSE168496. The genes higher in PD replicated in the array cohort but at little better than chance in the RNA-seq cohort, where the misses were the iron-handling and metallothionein transcripts. One reading is that the higher-in-PD component of the signature is partly a signal of cellular composition, since metallothionein and iron-handling transcripts are enriched in astrocytes and microglia and rise with gliosis, and the two cohorts may differ in the balance of neuronal loss and glial response; that is a hypothesis about bulk tissue, not a demonstration, and it is the reason this paper does not convert higher metallothionein or CP transcript into a claim about iron. The methods-and-panel manuscript as submitted reported that bulk PD substantia nigra did not replicate externally; that statement rested on an incomplete GEO2R export of GSE20292 (a download truncated at 1 MiB, which retained only the 41 most significant of the 54 measurable high-confidence genes) in which three control arrays were also on a linear rather than log2 scale; the complete, rescaled data give the result reported here, and the statement is corrected in the revision of that manuscript.

Published biochemistry and pharmacology stay in their lane

GSH depletion via GCLC RNAi causes dopaminergic loss in rat SNpc in a cited model (Garrido et al. 2011). NRF2 activators restore GSH and metabolic activity in olfactory neurosphere cells from sporadic PD and attenuate MPTP toxicity in mice (Cook et al. 2011; Kaidery et al. 2013). Soluble epoxide hydrolase (EPHX2) protein is up-regulated in dementia-with-Lewy-body striatum, and sEH inhibition or knockout is protective in MPTP and related models (Lakkappa et al. 2019; Ren et al. 2018; Sun et al. 2022); EC5026 has entered a Phase 1 trial in PD (NCT07142044; ClinicalTrials.gov, recruiting). P-glycoprotein function measured by [11C]-verapamil PET declines with advancing PD (Bartels et al. 2008); that is a transporter-function finding and is not interchangeable with ABCB1 transcript in the present meta-analysis. None of those reports assigns N-acetylcysteine, zinc, iron chelation, sulforaphane, or an sEH inhibitor to a patient on the basis of this GEO panel.

A timing contrast with the ASD companion has been drawn for CYP26 family members and for NOS2. In the locked adult PD pool, CYP26A1/B1 and NOS2 are not significant. The locked adult-brain ASD meta-analysis is likewise null for those symbols (McCaulley 2026c). Developmental-model or CSF/PET findings cited in the companions are therefore not re-tested here and are not treated as opposite locked-meta directions.

Descriptive comparison with the AD companion paper

Table 6 is a side-by-side with the AD companion paper (McCaulley 2026a). It is not a pooled analysis and does not propose a lifespan mechanism.

Table 6. Descriptive comparison with the AD companion paper (McCaulley 2026a). Not a pooled analysis.

Feature	AD companion (McCaulley 2026a)	This PD paper
Locked adult-brain result	141 high-confidence genes	91 FDR / 56 high-confidence (4 locked cohorts); 17 robust to all sensitivity analyses
Primary tissue in locked pool	Cortex / hippocampus (companion)	Substantia nigra, laser-captured SNpc neurons and superior frontal gyrus (GSE8397 A-chip arrays); supporting PFC
Named Phase I emphasis	No single dominant CYP in the locked claim	Published CYP2D6 poor-metabolizer association; not a transcript HC claim for CYP2D6

Feature	AD companion (McCaulley 2026a)	This PD paper
Metallothioneins	Eight metallothioneins HC higher (MT1A, MT1E, MT1F, MT1G, MT1H, MT1M, MT1X, MT2A)	MT1G, MT1F, MT1X among the 56 HC genes
GSH system	GSS/GSR among AD transcript calls	GSS and GPX3 HC lower; tissue GSH not measured
Gene–environment literature	APOE, metals as published context	PON1 × organophosphates; ABCB1 × pesticides (Table 5); CYP2D6 poor-metabolizer main effect, as published context

Shared gene symbols across companions are descriptive.

Limitations

Primary inference rests on four public array series of post-mortem brain. Sample sizes per series are small by contemporary RNA-seq standards. Bulk tissue mixes neuronal loss with gliosis; compositional change can mimic transcript change. GSE20141 is laser-captured SNpc neurons and is therefore closer to a neuronal compartment than GSE7621 or GSE49036, but four-cohort pooling still mixes compartments. Array platforms incompletely cover the newest panel members. The locked pipeline selected the lowest-p probe set per gene within each cohort, which inflates apparent effect sizes; treated GSE8397 arrays from the same donors, including frontal-cortex arrays, as independent; used only the HG-U133A arrays of GSE8397; and assessed cohort-level significance with a normal approximation that ignores the small residual degrees of freedom of these series. Sensitivity analyses that correct these features leave 17 of the 56 high-confidence genes robust to every correction (Supplementary Table S5). Standard errors for Figure 1 were recovered from deposited p values rather than from stored variances. Covariate adjustment (PMI, age, RIN) is explicit for the supporting GSE68719 screen and is not uniformly reconstructable from every older array series. Gene–environment odds ratios are from retrospective studies. Variants in Table 5 were not typed in these samples. GVI, DDI, and IDRS are uncalibrated and are not offered for patient-level assessment (Supplementary Note S3). Cross-companion comparisons use different tissues and platforms. The two independent checks support the high-confidence set in direction and rank, not in per-gene significance; the genes higher in PD replicated in the array cohort but at little better than chance in the RNA-seq cohort, an asymmetry consistent with a composition-sensitive component of the bulk-tissue signal that neuronal loss and gliosis can produce. Both external cohorts are small (29 and 16 samples), GSE20292 donors differ in age between groups, GSE20292 and the locked cohort GSE20141 were both re-deposited under the GSE20295 umbrella of Zheng et al. (2010) and donor-level independence of the whole-nigra arrays from the laser-captured neuronal isolates cannot be verified from the records, and the checks were made after the meta-analysis was locked, not pre-specified.

Acknowledgments

The author thanks the investigators who generated and deposited the GEO datasets analyzed in this work, and acknowledges the participants in the original studies.

Declarations

Competing interests

The author is the sole inventor on a filed patent application relating to panel-based research scoring methods that use the locked 468-gene list reported here. The application did not determine gene membership, the high-confidence rule, or the reported results, and the constructs described in

Supplementary Note S3 are stated to be unvalidated. Academic reuse of the deposited panel and code does not require a license from the author. The author declares no other competing interests.

Funding

This Parkinson disease analysis received no external funding and was conducted independently through Sirius Neuroscience, LLC. Sequencing and liver-tissue purchase that supported the author's prior Alzheimer disease study (McCaulley 2026b) were funded by Daniel Foley; that support did not fund the present Parkinson disease analysis or the GEO meta-analysis reported here.

Ethics approval

The archived transcriptomic analyses used only publicly available, de-identified gene-expression data from the NCBI Gene Expression Omnibus. Those datasets carry the ethical oversight of their originating studies. No identifiable private information was accessed, and no living human subjects were enrolled, recruited, or intervened upon by the author.

Consent to participate

Not applicable; no human participants were enrolled.

Consent to publish

Not applicable.

Data availability

Archived analyses use the public GEO accessions in Table 1. The locked 468-gene panel, excluded-symbol list, locked per-disease meta counts, methods note, per-gene PD meta-analysis table, and analysis pipeline are deposited at Zenodo. Version 1.0 (https://doi.org/10.5281/zenodo.22178094) holds the panel, excluded-symbol list, analysis pipeline, and locked per-disease meta counts; version 1.1 (https://doi.org/10.5281/zenodo.22181129) adds the per-gene meta-analysis tables, including the PD table with the high-confidence flag; version 1.2 (https://doi.org/10.5281/zenodo.22846772) adds the same PD table as a standalone workbook and CSV with the locked cohort contrast definitions, and is the version cited for the per-gene values in this paper. The all-versions DOI https://doi.org/10.5281/zenodo.22178093 resolves to the latest version. Supplementary Tables S1–S2 and Note S3 accompany this paper; Supplementary Table S4 (GSE20292 and GSE168496 per-gene values by all routes) is a separate workbook, as is Supplementary Table S5 (reproduction of the locked analysis and sensitivity analyses of the high-confidence set, with scripts). Figures 1 and 2 are provided as separate high-resolution files.

Author contributions

M.E.M. is the sole author and performed all aspects of this work: conceptualization, panel application, dataset analysis, interpretation, and manuscript writing.

Use of generative AI

During the preparation of this work the author used Grok (xAI), Claude (Anthropic), and Perplexity to assist with literature synthesis, bioinformatic pipeline construction, editorial refinement, and

critical review of claims and citations. After using these tools, the author reviewed and edited the content and takes full responsibility for the content of the article. No generative tool is an author.

References

Ahmed H, Abushouk AI, Gabr M, Negida A, Abdel-Daim MM (2017) Parkinson's disease and pesticides: a meta-analysis of disease connection and genetic alterations. Biomed Pharmacother 90:638–649. https://doi.org/10.1016/j.biopha.2017.03.100

Barrett T et al (2013) NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res 41:D991–D995. https://doi.org/10.1093/nar/gks1193

Bartels AL et al (2008) Decreased blood-brain barrier P-glycoprotein function in the progression of Parkinson's disease, PSP and MSA. J Neural Transm 115:1001–1009. https://doi.org/10.1007/s00702-008-0030-y

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57:289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bloem BR, Okun MS, Klein C (2021) Parkinson's disease. Lancet 397:2284–2303. https://doi.org/10.1016/S0140-6736(21)00218-X

Carmona A et al (2024) Metal dyshomeostasis in the substantia nigra of patients with Parkinson's disease or multiple sclerosis. J Neurochem 168:128–141. https://doi.org/10.1111/jnc.16040

Chinta SJ, Andersen JK (2006) Reversible inhibition of mitochondrial complex I activity following chronic dopaminergic glutathione depletion in vitro: implications for Parkinson's disease. Free Radic Biol Med 41:1442–1448. https://doi.org/10.1016/j.freeradbiomed.2006.08.002

Cook AL et al (2011) NRF2 activation restores disease related metabolic deficiencies in olfactory neurosphere-derived cells from patients with sporadic Parkinson's disease. PLoS One 6:e21907. https://doi.org/10.1371/journal.pone.0021907

DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Trials 7:177–188. https://doi.org/10.1016/0197-2456(86)90046-2

Dijkstra AA et al (2015) Evidence for immune response, axonal dysfunction and reduced endocytosis in the substantia nigra in early stage Parkinson's disease. PLoS One 10:e0128651. https://doi.org/10.1371/journal.pone.0128651

Dumitriu A et al (2016) Integrative analyses of proteomics and RNA transcriptomics implicate mitochondrial processes, protein folding pathways and GWAS loci in Parkinson disease. BMC Med Genomics 9:5. https://doi.org/10.1186/s12920-016-0164-y

Dutheil F, Beaune P, Tzourio C, Loriot MA, Elbaz A (2010) Interaction between ABCB1 and professional exposure to organochlorine insecticides in Parkinson disease. Arch Neurol 67:739–745. https://doi.org/10.1001/archneurol.2010.101

Ebadi M et al (2005) Metallothionein-mediated neuroprotection in genetically engineered mouse models of Parkinson's disease. Brain Res Mol Brain Res 134:67–75. https://doi.org/10.1016/j.molbrainres.2004.09.011

Fan HH et al (2022) Polymorphisms of cytochromes P450 and glutathione S-transferases synergistically modulate risk for Parkinson's disease. Front Aging Neurosci 14:888942. https://doi.org/10.3389/fnagi.2022.888942

Fløynes IH et al (2024) Mitochondrial complex I deficiency stratifies idiopathic Parkinson's disease. Nat Commun 15:3631. https://doi.org/10.1038/s41467-024-47867-4

Garrido M et al (2011) Glutathione depletion and overproduction both initiate degeneration of nigral dopaminergic neurons. Acta Neuropathol 121:475–485. https://doi.org/10.1007/s00401-010-0791-x

Glaab E, Schneider R (2015) Comparative pathway and network analysis of brain transcriptome changes during adult aging and in Parkinson's disease. Neurobiol Dis 74:1–13. https://doi.org/10.1016/j.nbd.2014.11.002

Haas RH et al (1995) Low platelet mitochondrial complex I and complex II/III activity in early untreated Parkinson's disease. Ann Neurol 37:714–722. https://doi.org/10.1002/ana.410370604

Jenner P (2003) Oxidative stress in Parkinson's disease. Ann Neurol 53 Suppl 3:S26–S36. https://doi.org/10.1002/ana.10483

Kaidery NA et al (2013) Targeting Nrf2-mediated gene transcription by extremely potent synthetic triterpenoids attenuate dopaminergic neurotoxicity in the MPTP mouse model of Parkinson's disease. Antioxid Redox Signal 18:139–157. https://doi.org/10.1089/ars.2011.4491

Keeney PM, Xie J, Capaldi RA, Bennett JP Jr (2006) Parkinson's disease brain mitochondrial complex I has oxidatively damaged subunits and is functionally impaired and misassembled. J Neurosci 26:5256–5264. https://doi.org/10.1523/JNEUROSCI.0984-06.2006

Kverneng SU et al (2025) Mitochondrial complex I deficiency occurs in skeletal muscle of a subgroup of individuals with Parkinson's disease. Commun Med (Lond) 5:141. https://doi.org/10.1038/s43856-025-00817-7

Lakkappa N, Krishnamurthy PT, Pandareesh MD, Hammock BD, Hwang SH (2019) Soluble epoxide hydrolase inhibitor, APAU, protects dopaminergic neurons against rotenone induced neurotoxicity: implications for Parkinson's disease. Neurotoxicology 70:135–145. https://doi.org/10.1016/j.neuro.2018.11.010

Lee PC, Rhodes SL, Sinsheimer JS, Bronstein J, Ritz B (2013) Functional paraoxonase 1 variants modify the risk of Parkinson's disease due to organophosphate exposure. Environ Int 56:42–47. https://doi.org/10.1016/j.envint.2013.03.004

Lesnick TG et al (2007) A genomic pathway approach to a complex disease: axon guidance and Parkinson disease. PLoS Genet 3:e98. https://doi.org/10.1371/journal.pgen.0030098

Liu H, Peng J, Gao J, Zheng F, Tie C (2015) Glutathione S-transferase T1 and M1 null genotypes and Parkinson's disease risk: evidence from an updated meta-analysis. Neurol Sci 36:1559–1565. https://doi.org/10.1007/s10072-015-2159-4

Lu Y et al (2013) CYP2D6*4 allele polymorphism increases the risk of Parkinson's disease: evidence from meta-analysis. PLoS One 8:e84413. https://doi.org/10.1371/journal.pone.0084413

Lu Y et al (2014) CYP2D6 phenotypes and Parkinson's disease risk: a meta-analysis. J Neurol Sci 336:161–168. https://doi.org/10.1016/j.jns.2013.10.030

Manthripragada AD, Costello S, Cockburn MG, Bronstein JM, Ritz B (2010) Paraoxonase 1, agricultural organophosphate exposure, and Parkinson disease. Epidemiology 21:87–94. https://doi.org/10.1097/EDE.0b013e3181c15ec6

McCaulley ME (2026a) A locked 468-gene detoxification panel in Alzheimer's disease brain, liver, and blood: 141 high-confidence cerebral associations including lower CYP46A1 transcript abundance, and a 12-sample hepatic CYP7A1 signal. Research Square, version 1. https://doi.org/10.21203/rs.3.rs-10999456/v1

McCaulley ME (2026b) Impaired pathways of detoxification and risk of Alzheimer's disease: differential gene expression studies via gene set expression analysis. J Alzheimers Dis Rep 10:25424823261419382. https://doi.org/10.1177/25424823261419382

McCaulley ME (2026c) A null adult postmortem-brain test of a locked 468-gene panel in autism spectrum disorder, with published maternal-axis architecture as context. Companion manuscript of this series, in preparation

McCaulley ME (2026d) Panel-gene transcriptomic associations in amyotrophic lateral sclerosis: a donor-resolved seven-cohort analysis of a locked 468-gene detoxification panel. Companion manuscript of this series, in preparation

McCaulley ME (2026e) Panel-gene transcriptomic associations in multiple sclerosis: a three-cohort white-matter analysis of a locked 468-gene detoxification panel. Companion manuscript of this series, in preparation

McCaulley ME (2026f) Transcriptomic associations of a pre-specified 468-gene redox and xenobiotic-metabolism panel across five neurological disorders, with a published-variant inventory. Methods-and-panel manuscript of this series, under peer review at Molecular Neurobiology

Michael GJ et al (2011) Up-regulation of metallothionein gene expression in parkinsonian astrocytes. Neurogenetics 12:295–305. https://doi.org/10.1007/s10048-011-0294-5

Miller JA et al (2011) Strategies for aggregating gene expression data: the collapseRows R function. BMC Bioinformatics 12:322. https://doi.org/10.1186/1471-2105-12-322

Miyazaki I, Asanuma M (2023) Multifunctional metallothioneins as a target for neuroprotection in Parkinson's disease. Antioxidants (Basel) 12:894. https://doi.org/10.3390/antiox12040894

Moran LB et al (2006) Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. Neurogenetics 7:1–11. https://doi.org/10.1007/s10048-005-0020-2

Narayan S et al (2015) Genetic variability in ABCB1, occupational pesticide exposure, and Parkinson's disease. Environ Res 143:98–106. https://doi.org/10.1016/j.envres.2015.08.022

Paul KC et al (2016) Organophosphate pesticide exposures, nitric oxide synthase gene variants, and gene–pesticide interactions in a case-control study of Parkinson's disease, California (USA). Environ Health Perspect 124:570–577. https://doi.org/10.1289/ehp.1408976

Ramsey CP et al (2007) Expression of Nrf2 in neurodegenerative diseases. J Neuropathol Exp Neurol 66:75–85. https://doi.org/10.1097/nen.0b013e31802d6da9

Ren Q et al (2018) Soluble epoxide hydrolase plays a key role in the pathogenesis of Parkinson's disease. Proc Natl Acad Sci USA 115:E5815–E5823. https://doi.org/10.1073/pnas.1802179115

Ritchie ME et al (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47. https://doi.org/10.1093/nar/gkv007

Schapira AH et al (1990) Mitochondrial complex I deficiency in Parkinson's disease. J Neurochem 54:823–827. https://doi.org/10.1111/j.1471-4159.1990.tb02325.x

Shukla D et al (2023) Glutathione depletion and concomitant elevation of susceptibility in patients with Parkinson's disease: state-of-the-art MR spectroscopy and neuropsychological study. ACS Chem Neurosci 14:4383–4394. https://doi.org/10.1021/acschemneuro.3c00717

Sian J et al (1994) Alterations in glutathione levels in Parkinson's disease and other neurodegenerative disorders affecting basal ganglia. Ann Neurol 36:348–355. https://doi.org/10.1002/ana.410360305

Smeyne M, Smeyne RJ (2013) Glutathione metabolism and Parkinson's disease. Free Radic Biol Med 62:13–25. https://doi.org/10.1016/j.freeradbiomed.2013.05.001

Smyth GK (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3:Article3. https://doi.org/10.2202/1544-6115.1027

Sofic E, Lange KW, Jellinger K, Riederer P (1992) Reduced and oxidized glutathione in the substantia nigra of patients with Parkinson's disease. Neurosci Lett 142:128–130. https://doi.org/10.1016/0304-3940(92)90355-b

Spillantini MG et al (1997) α-Synuclein in Lewy bodies. Nature 388:839–840. https://doi.org/10.1038/42166

Su D et al (2025) Projections for prevalence of Parkinson's disease and its driving factors in 195 countries and territories to 2050: modelling study of Global Burden of Disease Study 2021. BMJ 388:e080952. https://doi.org/10.1136/bmj-2024-080952

Sun CP et al (2022) Kurarinone alleviated Parkinson's disease via stabilization of epoxyeicosatrienoic acids in animal model. Proc Natl Acad Sci USA 119:e2118818119. https://doi.org/10.1073/pnas.2118818119

Surmeier DJ, Obeso JA, Halliday GM (2017) Selective neuronal vulnerability in Parkinson disease. Nat Rev Neurosci 18:101–113. https://doi.org/10.1038/nrn.2016.178

Tranchevent LC, Halder R, Glaab E (2023a) Systems level analysis of sex-dependent gene expression changes in Parkinson's disease. NPJ Parkinsons Dis 9:8. https://doi.org/10.1038/s41531-023-00446-8

Tranchevent LC, Halder R, Glaab E (2023b) Transcriptomics analysis of post-mortem substantia nigra samples from Parkinson's disease patients and controls. NCBI Gene Expression Omnibus, GSE168496. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE168496

Wang T, Wang B (2014) Association between glutathione S-transferase M1/glutathione S-transferase T1 polymorphisms and Parkinson's disease: a meta-analysis. J Neurol Sci 338:65–70. https://doi.org/10.1016/j.jns.2013.12.018

Weikang C, Jie L, Likang L, Weiwen Q, Liping L (2016) A meta-analysis of association between glutathione S-transferase M1 gene polymorphism and Parkinson's disease susceptibility. Open Med (Wars) 11:578–583. https://doi.org/10.1515/med-2016-0094

Wu YR et al (2021) Association of SOD2 p.V16A polymorphism with Parkinson's disease: a meta-analysis in Han Chinese. J Formos Med Assoc 120:501–507. https://doi.org/10.1016/j.jfma.2020.06.023

Zhang Y, James M, Middleton FA, Davis RL (2005) Transcriptional analysis of multiple brain regions in Parkinson's disease supports the involvement of specific protein processing, energy metabolism, and signaling pathways, and suggests novel disease mechanisms. Am J Med Genet B Neuropsychiatr Genet 137B:5–16. https://doi.org/10.1002/ajmg.b.30195

Zheng B et al (2010) PGC-1α, a potential therapeutic target for early intervention in Parkinson's disease. Sci Transl Med 2:52ra73. https://doi.org/10.1126/scitranslmed.3001059

Figures

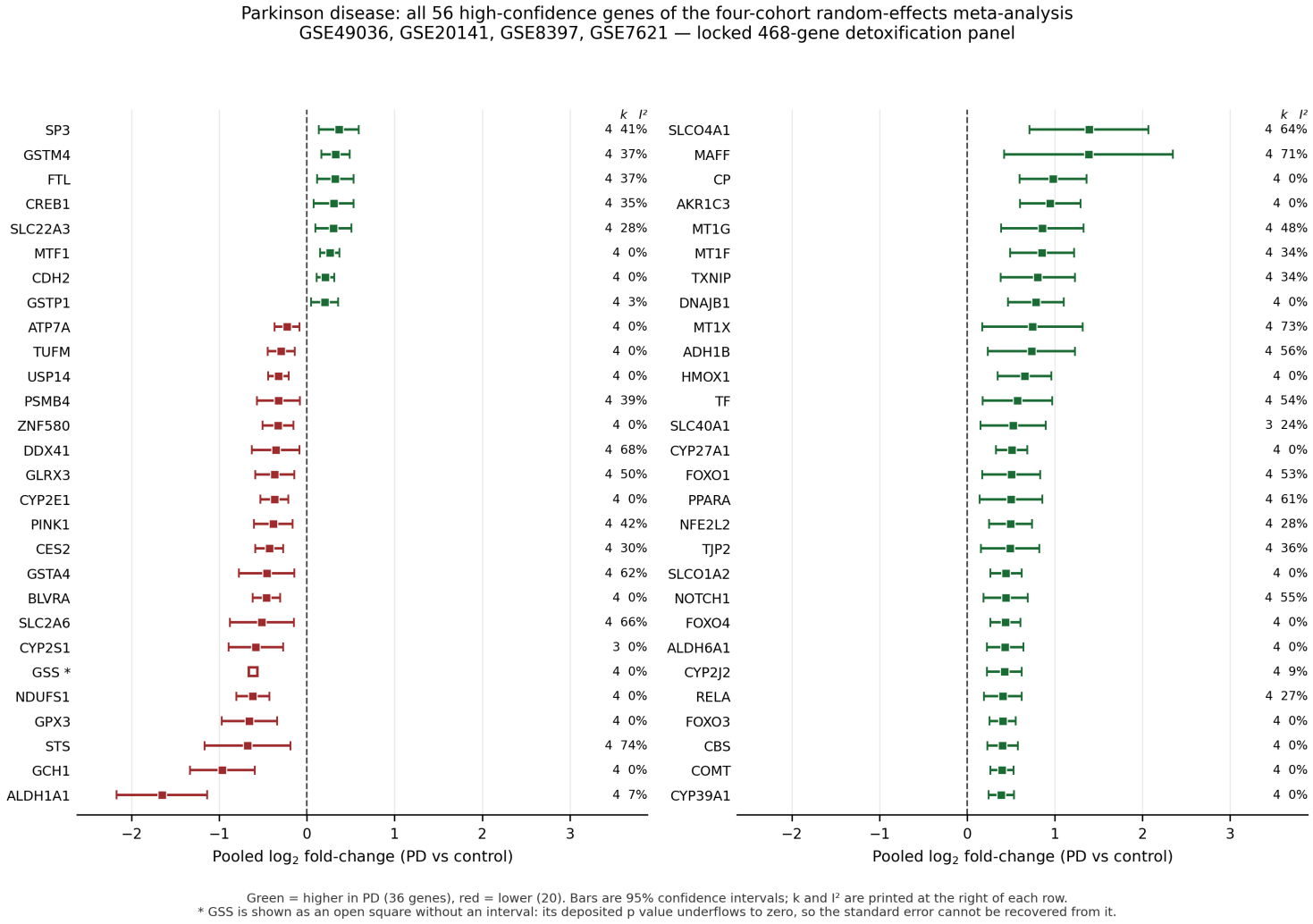


Figure 1

All 56 high-confidence genes of the four-cohort Parkinson disease random-effects meta-analysis (DerSimonian–

Laird) of the locked 468-gene panel across GSE49036, GSE20141, GSE8397, and GSE7621, ordered by pooled effect. A

gene is high-confidence when it is FDR-significant, measured in at least three cohorts, direction-concordant in at least

70% of contributing cohorts, and nominally significant in at least two. Points are pooled log2 fold-changes; bars are 95%

confidence intervals computed as the pooled estimate ± 1.96 × SE, with SE recovered as |pooled| / z(p) because the

deposited per-gene file does not carry standard errors; k and I² are printed at the right of each row. Thirty-six genes are

higher in PD and twenty lower. GSS is drawn as an open square without an interval because its deposited p value

underflows to zero. Median I² across the 56 is 16.4%. Values are those deposited athttps://doi.org/10.5281/zenodo.22846772. The robustness of each gene to the sensitivity analyses is given in

Supplementary Table S5.

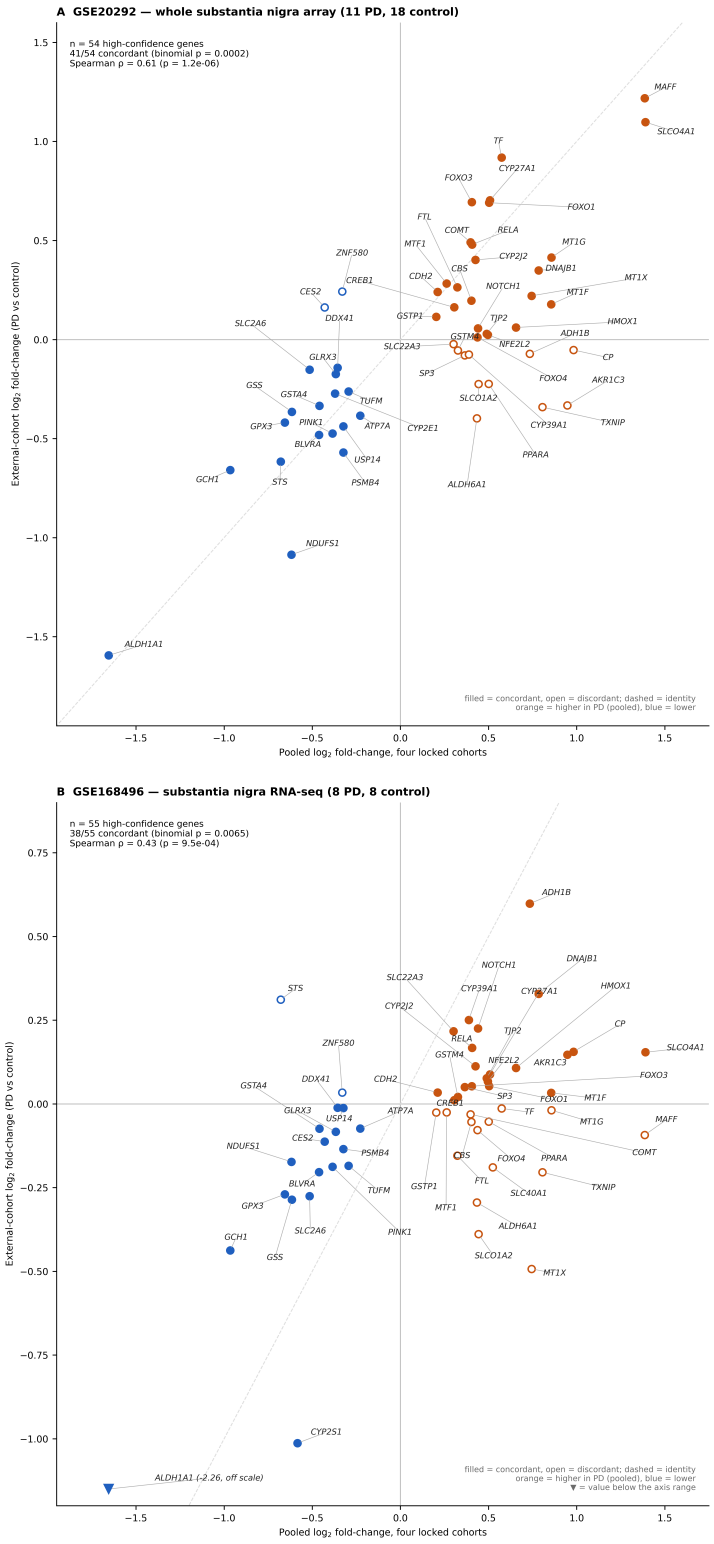


Figure 2

Independent-cohort checks of the 56 high-confidence genes in two substantia nigra cohorts that contributed

nothing to the locked meta-analysis. Each point is one gene, labelled by symbol: pooled log2 fold-change from the four locked

cohorts on the horizontal axis, external-cohort log2 fold-change (PD versus control) on the vertical axis. Orange, higher in PD

in the pooled analysis; blue, lower. Filled symbols are direction-concordant, open symbols discordant; the dashed line is

identity. (A) GSE20292, whole substantia nigra HG-U133A array, 11 PD and 18 controls, all arrays on a common log2 scale

and quantile-normalized (unadjusted estimates; age- and sex-adjusted estimates are in Supplementary Table S4); 54 genesshown, 41 concordant (exact binomial p = 0.0002), Spearman ρ = 0.61. (B) GSE168496, substantia nigra RNA-seq, 8 PD and 8

controls, NCBI GEO2R DESeq2 estimates; 55 genes shown (CYP2E1 absent), 38 concordant (p = 0.006), ρ = 0.43. The vertical

axis of panel B is limited to −1.2 to 0.9 so that the clustered genes remain legible; ALDH1A1 (−2.26) lies below that range andis drawn as a triangle at the axis floor with its value stated. Per-gene values for both cohorts are in Supplementary Table S4.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- McCaulleyPD468panelSupplementary20261014.docx
- PDSupplementaryTableS4externalchecks.xlsx
- SupplTableS5PDHCreproductionsensitivity20260930.xlsx
- McCaulleyPDTableS5scripts20261015.zip