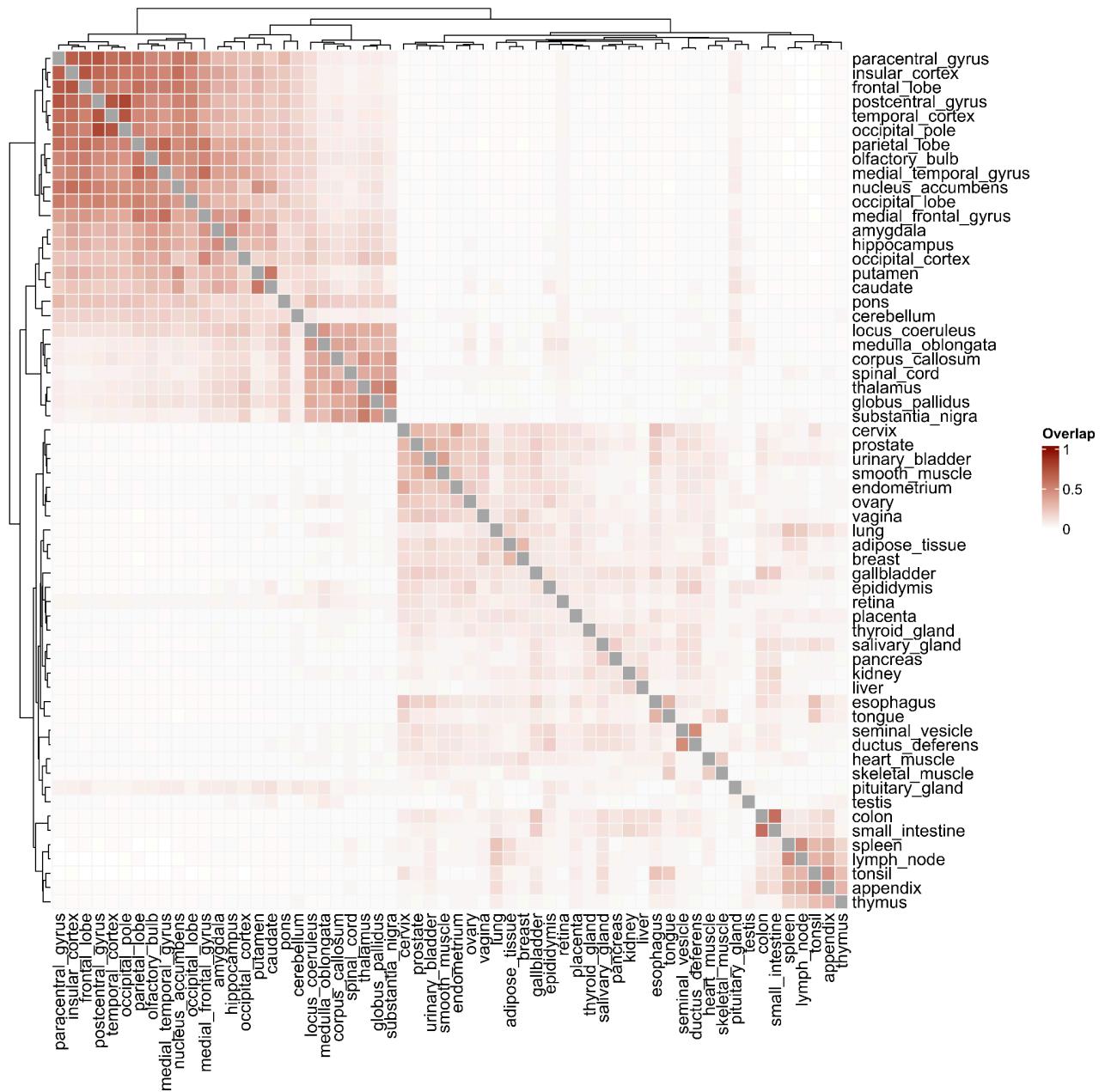


Supplementary Information

Supplementary Figures.....	2
Supplementary Tables.....	69

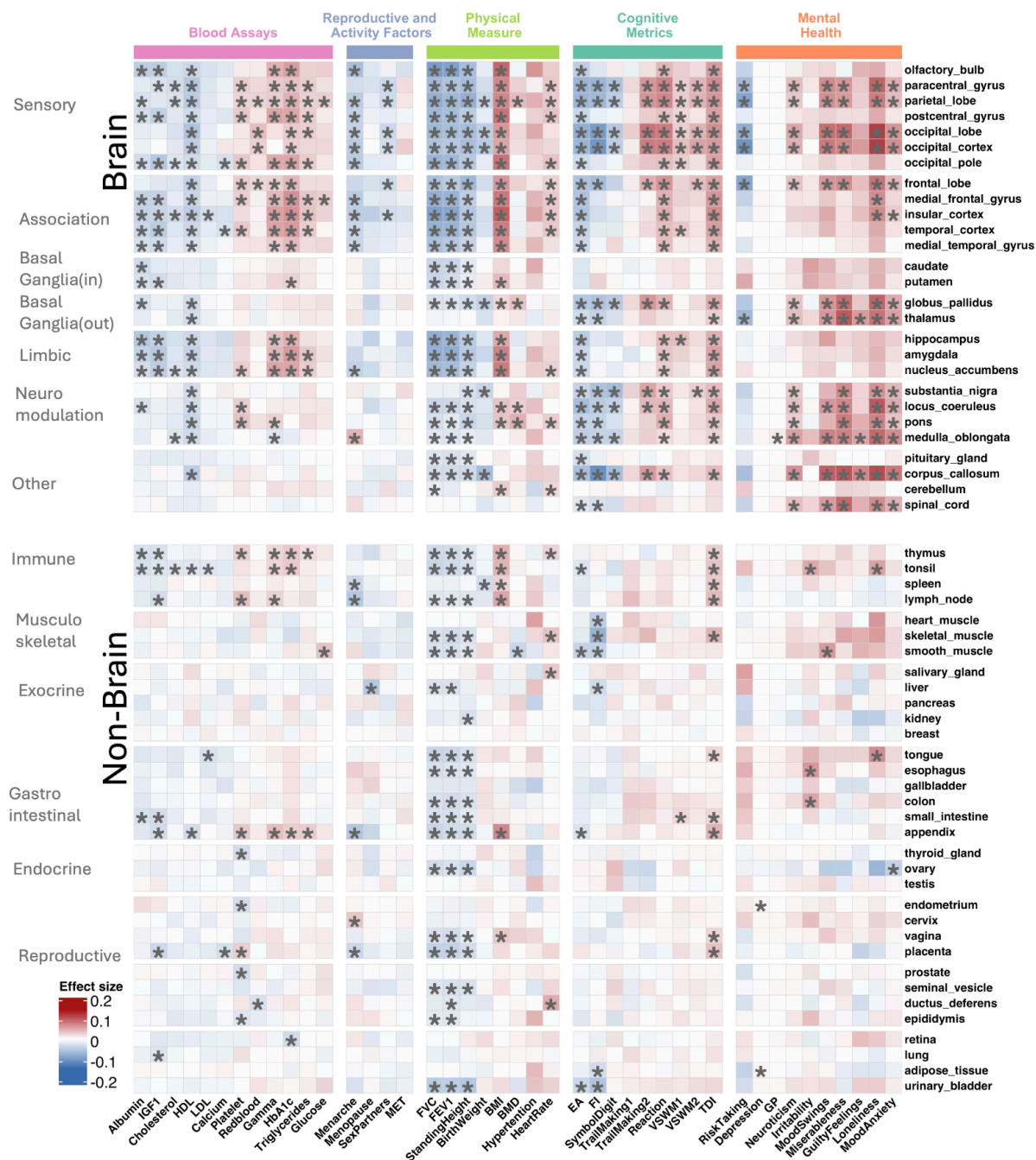
Supplementary Figures



Supplementary Fig. 1. Pairwise overlap matrix (Jaccard overlap) of tissue gene sets

The heatmap displays the pairwise overlap (1-Jaccard Distance) between tissue gene sets. Color encodes overlap magnitude. dark red = high overlap, pink = moderate, white = low (scale 0–1). The maximum, minimum, mean, and median overlap values are 77.2%, 0.07%, 7.9%, and 2.7%,

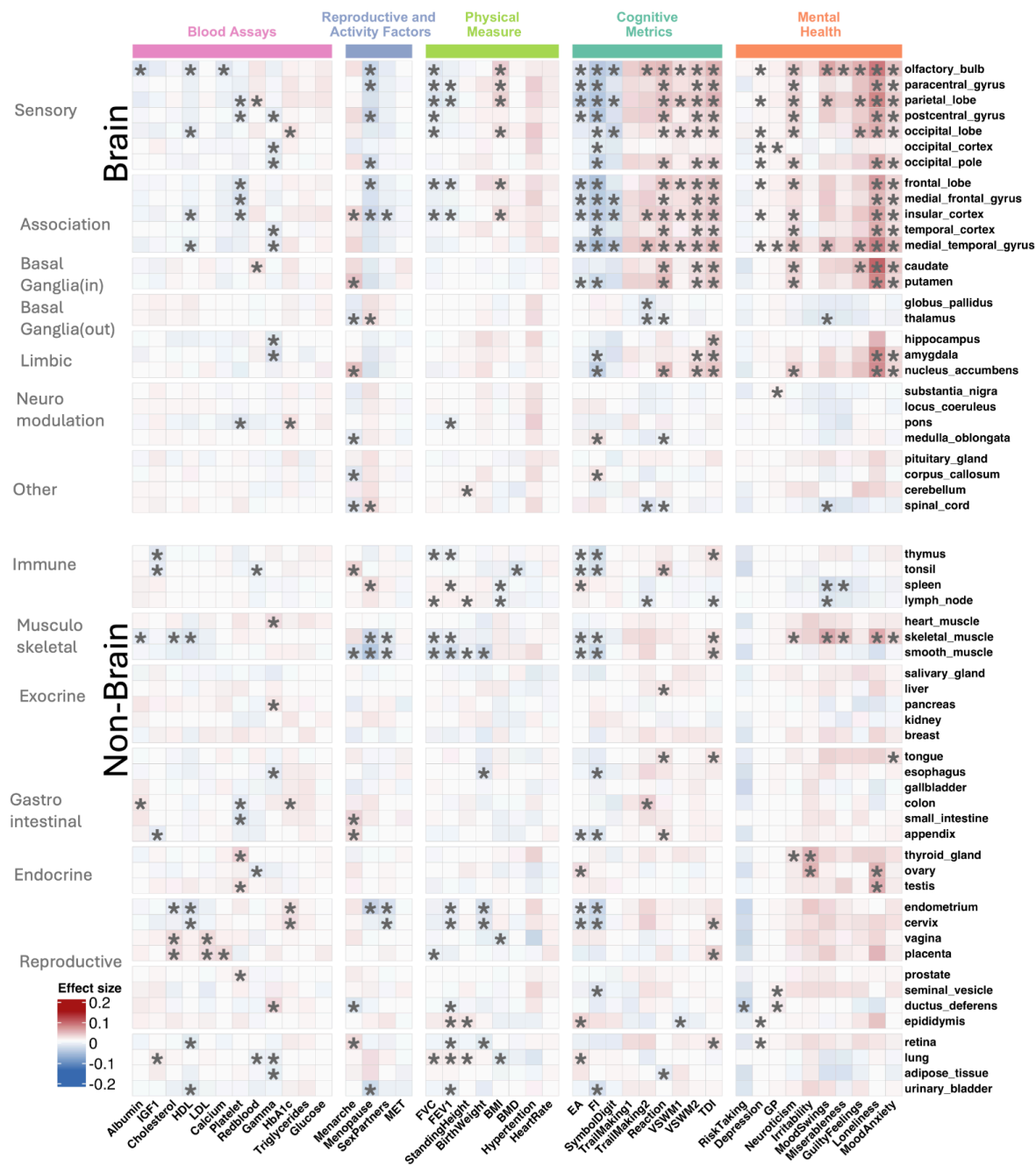
respectively. This is a descriptive similarity matrix; no statistical test is applied. Pairwise values are provided in Supplementary Data 3.



Supplementary Fig. 2. Heatmap of deletion effect sizes for whole-body tissue across traits

The heatmap displays deletion burden association effect sizes (β_1) between five trait categories

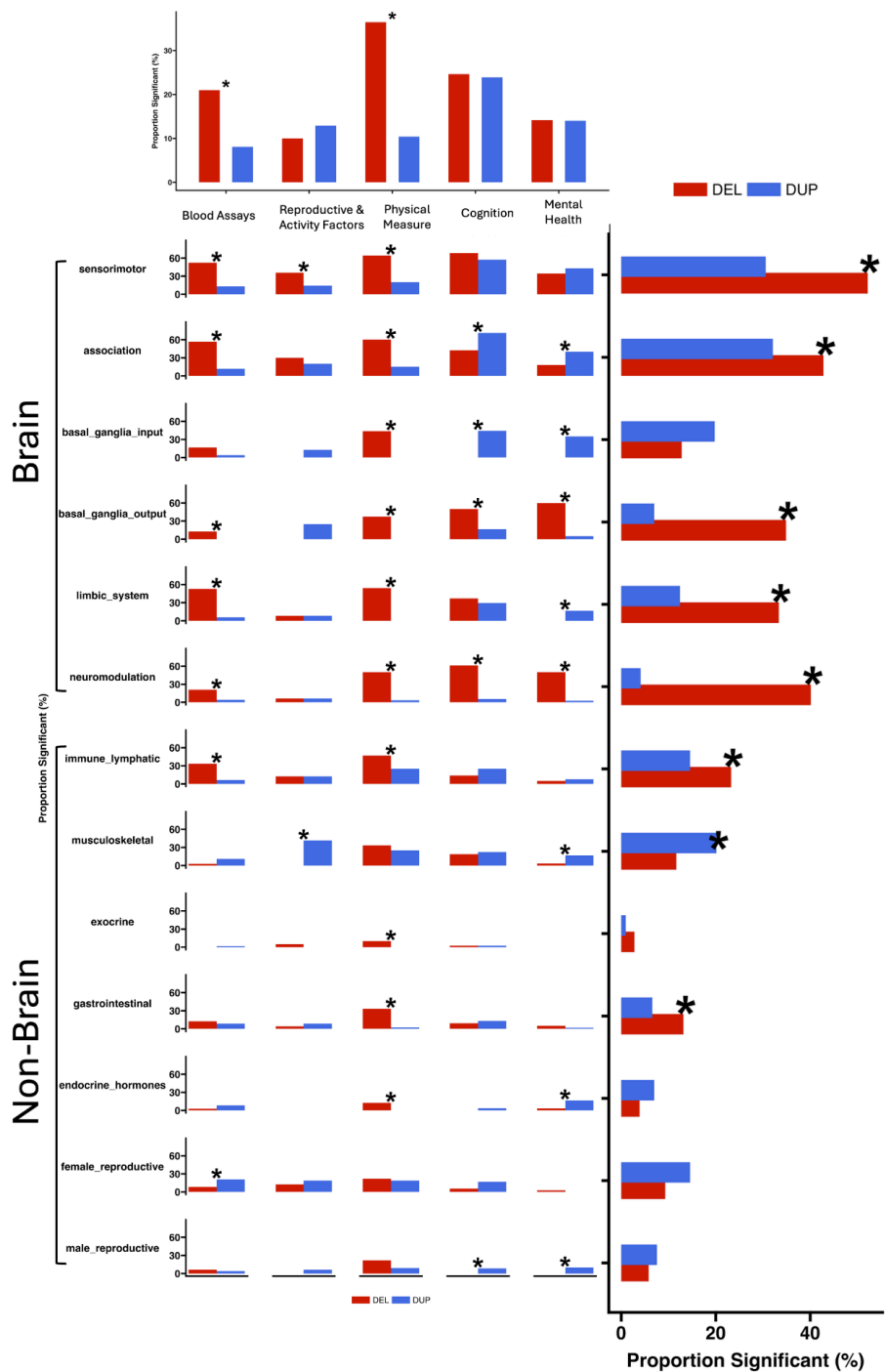
(x-axis) and tissue-specific gene sets (y-axis; categories/annotations at right). Color encodes the direction and magnitude of the standardized effect size (blue = negative; red = positive; scale centered at zero, spanning approximately -0.15 to $+0.16$). Black asterisks (*) indicate FDR-significant associations, determined using a two-sided t-test on the inside-set burden coefficient (β_1) from the FunBurd multiple linear regression model, with Benjamini–Hochberg FDR correction across 172 gene sets x 43 traits x 2 CNV types (14,792 tests; $q < 0.05$). All trait–gene-set combinations were tested to provide a complete picture, including biologically unexpected combinations (e.g. female-specific phenotypes such as menopause with male gene sets such as testis) that serve as negative controls. Full estimates, standard errors, and p-values are in Supplementary Data 4. Abbreviations. CNV, copy-number variant; FDR, false discovery rate.



Supplementary Fig. 3. Heatmap of duplication effect sizes for whole-body tissue across traits

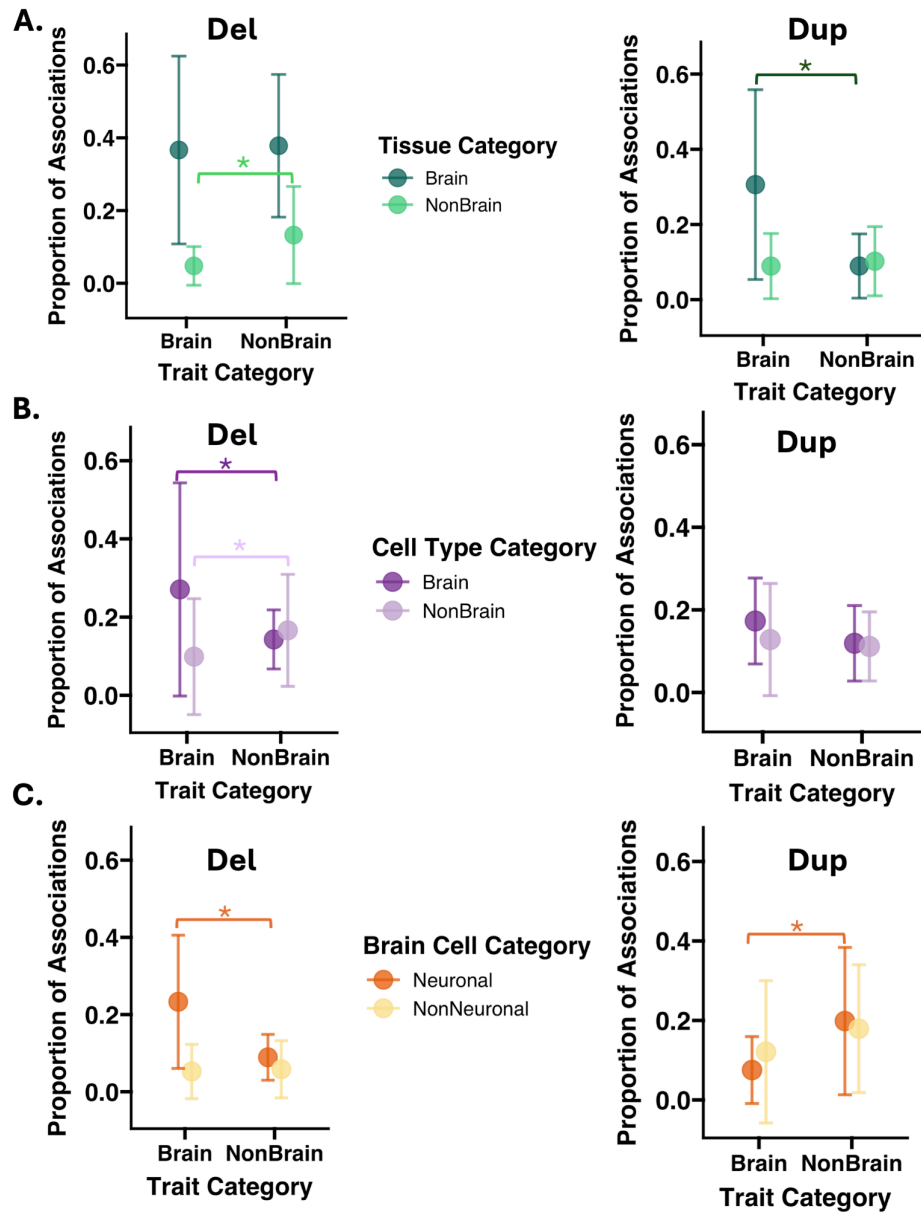
The heatmap displays duplication burden association effect sizes (β_1) between five trait categories (x-axis) and tissue-specific gene sets (y-axis; categories/annotations at right). Color encodes direction and magnitude of the standardized effect size (blue = negative; red = positive; scale centered at zero, ≈ -0.15 to $+0.16$). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the FunBurd multiple linear regression model);

Benjamini–Hochberg FDR across $172 \times 43 \times 2 = 14,792$ tests; $q < 0.05$). All trait–gene-set combinations were tested, including negative-control combinations. Full statistics in Supplementary Data 4.



Supplementary Fig. 4. Differences between deletion–duplication associations across trait and gene-set categories

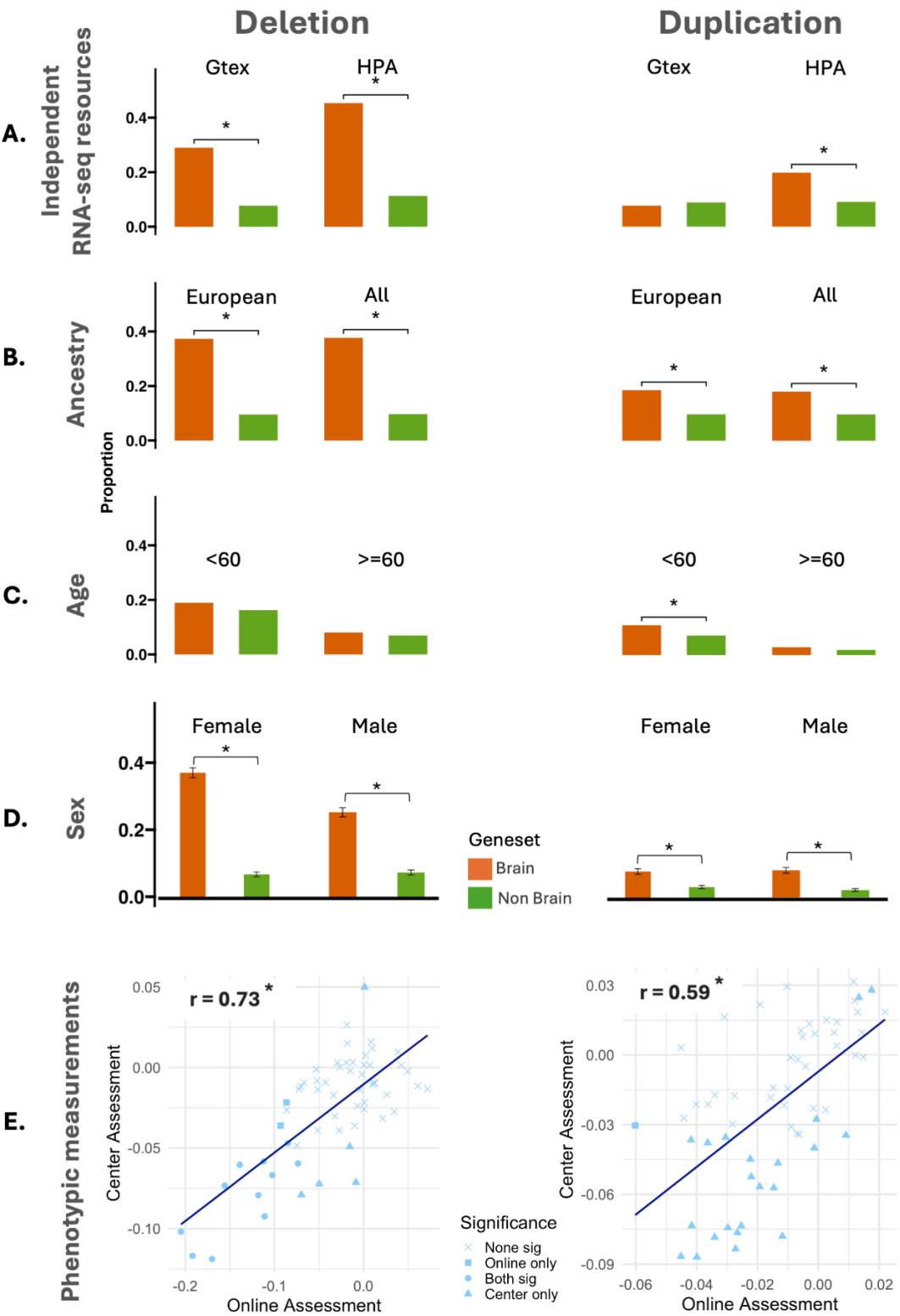
Bar plots summarize the difference in association level between deletions and duplications (proportion of FDR-significant associations) across five trait categories and functional gene-set categories/annotations. Bar height = the aggregate proportion of significant associations within each category; marginal plots show this proportion per trait category (top) and per gene-set category (right). Black asterisks (*) denote statistically significant deletion–duplication differences in the proportion of FDR-significant associations, determined using a two-sided test for equal proportions (Pearson's chi-squared test, no continuity correction) with Benjamini–Hochberg FDR correction applied within each panel ($q < 0.05$). Each proportion is computed over all deletion (or duplication) associations in the corresponding category, i.e. (number of traits in the category) $\times$ (number of tissue gene sets); associations span 43 traits $\times$ 60 whole-body tissue gene sets. Exact counts, test statistics, and FDR-adjusted p-values per panel are provided in Source Data (sheets SF4A–SF4C). Blue = duplication; red = deletion.



Supplementary Fig. 5. Summary of significant association proportions across tissue- and cell-type gene sets

Point range plots (A–C) show the proportion of $F_{\text{trait-gene-set}}$ associations (y-axis) for distinct gene sets (colors) across trait categories (x-axis); error bars represent the 95% confidence interval of the proportion. Gene sets are grouped by tissue (A; $n = 60$), whole-body cell types (B; $n = 81$), and whole-brain cell types (C; $n = 31$). An asterisk (*) denotes a statistically significant difference in association proportion between the two trait categories for a given gene set, determined using a two-sided two-proportion z-test after Benjamini–Hochberg FDR correction

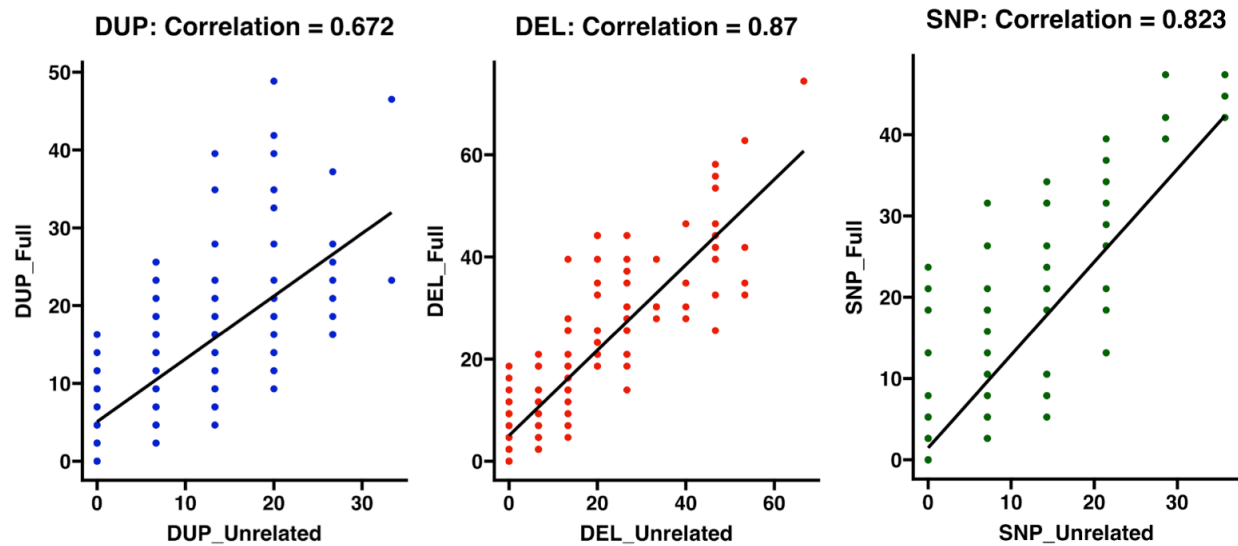
($q < 0.05$), indicating gene-set specificity to a trait category. Exact statistics in Supplementary Tables 2 and 4.



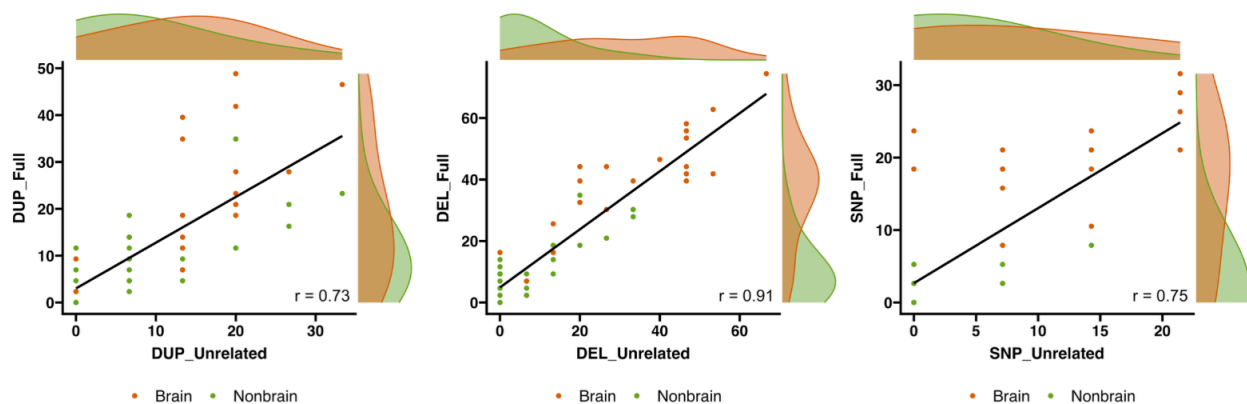
Supplementary Fig. 6. Sensitivity analysis of CNV associations across brain and non-brain gene sets

Each panel shows the proportion of significant CNV–trait associations (y-axis; functional burden pleiotropy) across brain and non-brain gene sets (x-axis) under one sensitivity condition: (A) independent RNA-seq resources (GTEx vs HPA); (B) ancestry (White British vs all UK Biobank ancestries); (C) age (< 60 vs ≥ 60 years); (D) sex; and (E) phenotype measurement method (Fluid Intelligence at the assessment centre, $n = 217,809$, vs online, $n = 92,106$). Where brain versus non-brain proportions are compared within a condition, significance was assessed using a two-sided two-proportion z-test with Benjamini–Hochberg FDR correction ($q < 0.05$). Per-comparison proportions and FDR-adjusted p-values are provided in Supplementary Data 5. Abbreviations: GTEx, Genotype-Tissue Expression; HPA, Human Protein Atlas.

A.

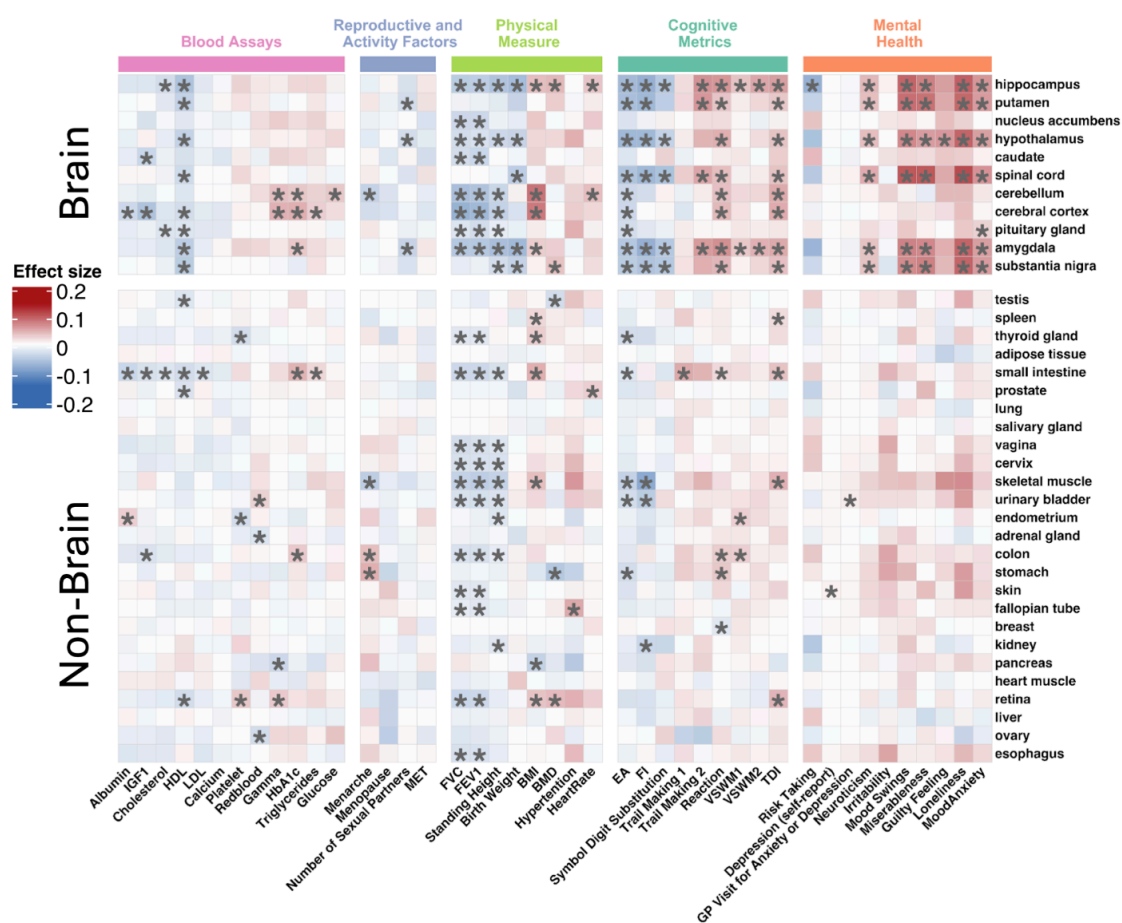


B.



Supplementary Fig. 7. Sensitivity analysis on pleiotropy computed by unrelated and all traits

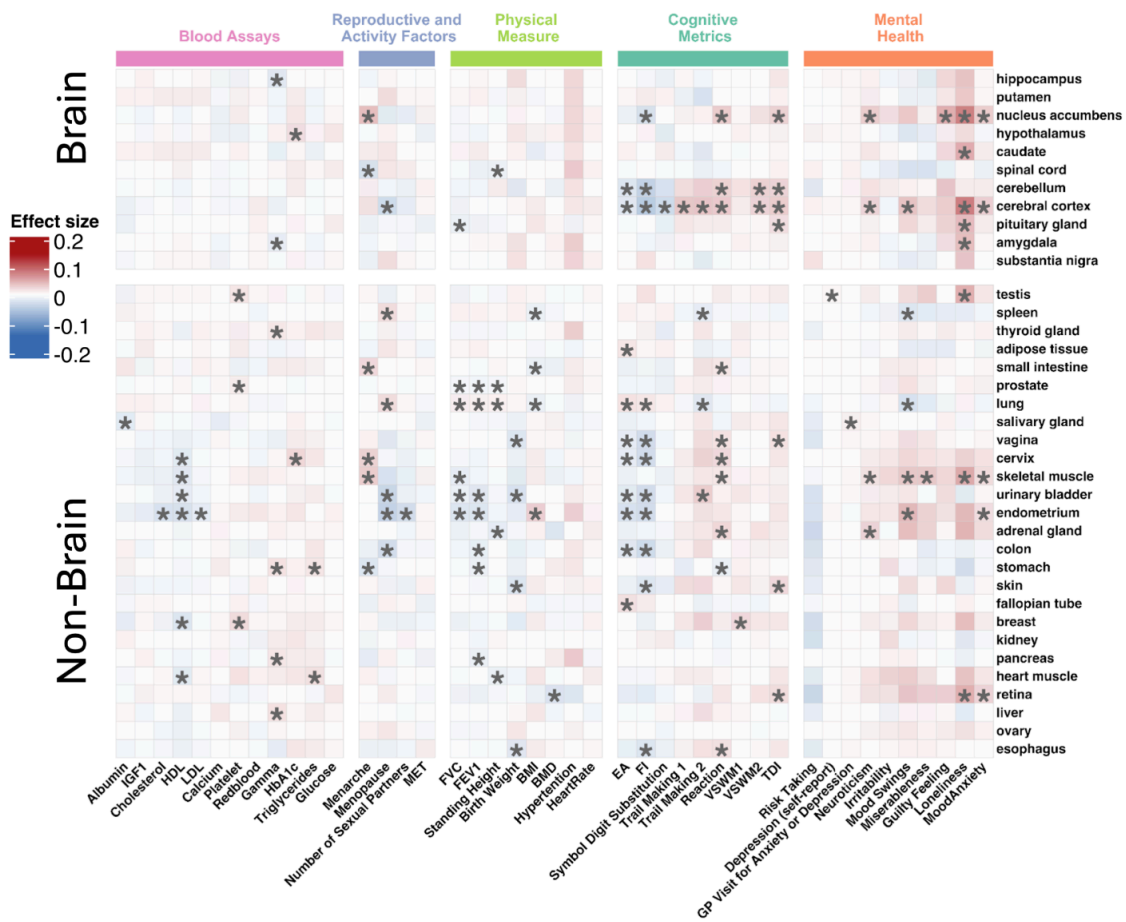
X-axis & Y-axis represent the pleiotropy calculated on unrelated and all traits, respectively, using A) all 172 gene sets and B) the 60 Tissue gene sets. Concordance is summarized using two-sided Pearson correlations. For pleiotropy on unrelated traits, we recalculated pleiotropy using a subset of 15 unrelated traits that have less than 0.3 correlation with other traits. The pleiotropy estimates from this subset remained highly correlated with those from the full set, with correlations of $r=0.68$ and $r=0.87$ for duplications and deletions, respectively. These correlations are higher for tissue gene sets, 0.73 and 0.91, with significant separation between brain and non-brain tissue gene sets.



Supplementary Fig. 8. Heatmap of deletion effect sizes for whole-body tissue (GTEx) across traits

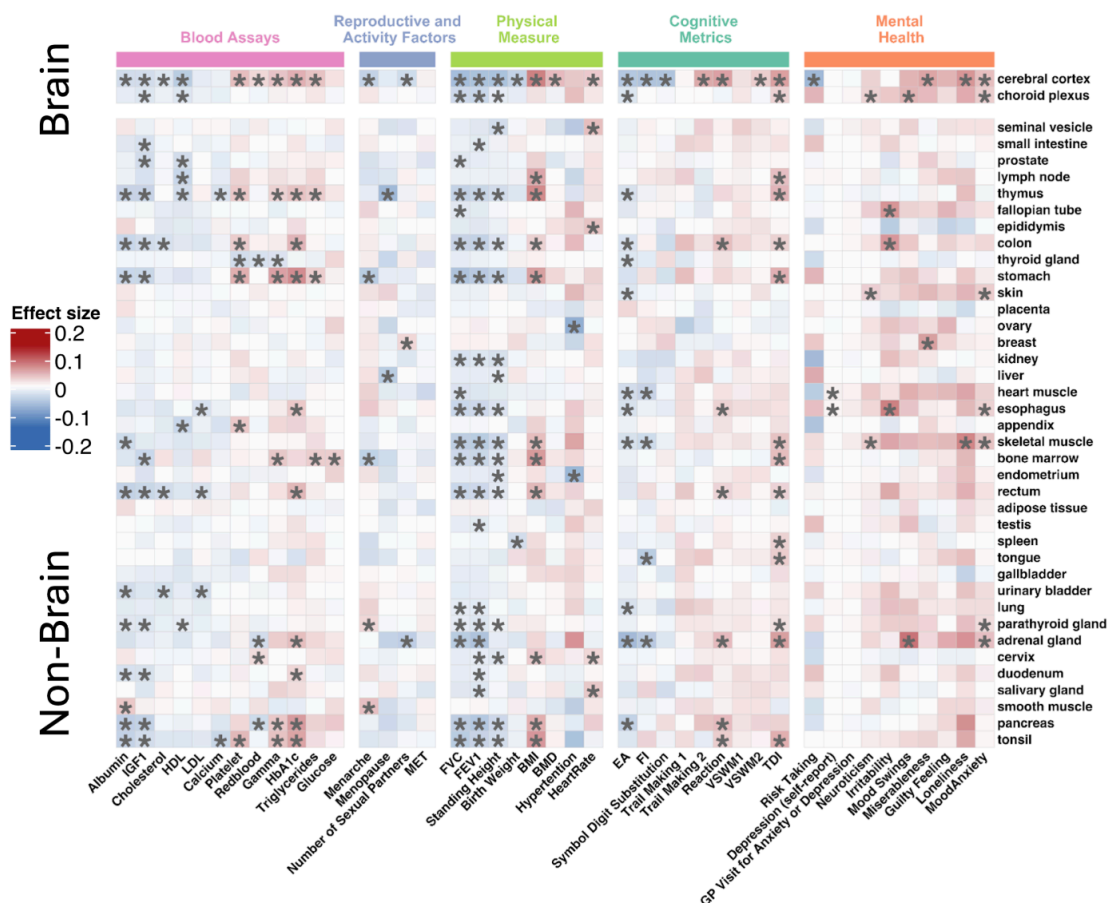
The heatmap displays deletion burden association effect sizes (β_1) between five trait categories (x-axis) and GTEx-based tissue-specific gene sets (y-axis). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations, determined using a two-sided t-test on β_1 from the

FunBurd multiple linear regression model, with Benjamini–Hochberg FDR correction across all GTEx gene sets $\times$ 43 traits $\times$ 2 CNV types ($q < 0.05$). Source data are provided in Source Data sheet SF8.



Supplementary Fig. 9. Heatmap of duplication effect sizes for whole-body tissue (GTEx) across traits

As in Supplementary Fig. 8, but for duplications. Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the FunBurd model; Benjamini–Hochberg FDR across all GTEx gene sets $\times$ 43 traits $\times$ 2 CNV types; $q < 0.05$). Source data are provided in Source Data sheet SF9.



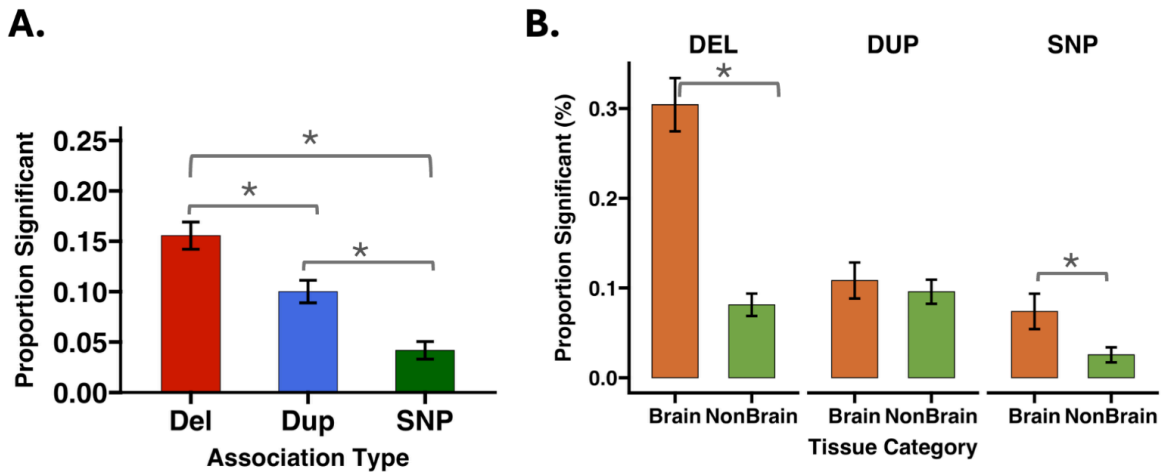
Supplementary Fig. 10. Heatmap of deletion effect sizes for whole-body tissue (HPA) across traits

The heatmap displays deletion burden association effect sizes (β_1) between five trait categories (x-axis) and Human Protein Atlas (HPA)-based tissue-specific gene sets (y-axis). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations, determined using a two-sided t-test on β_1 from the FunBurd multiple linear regression model, with Benjamini–Hochberg FDR correction across all HPA tissue gene sets $\times$ 43 traits $\times$ 2 CNV types ($q < 0.05$). Source data are provided in Source Data sheet SF10.



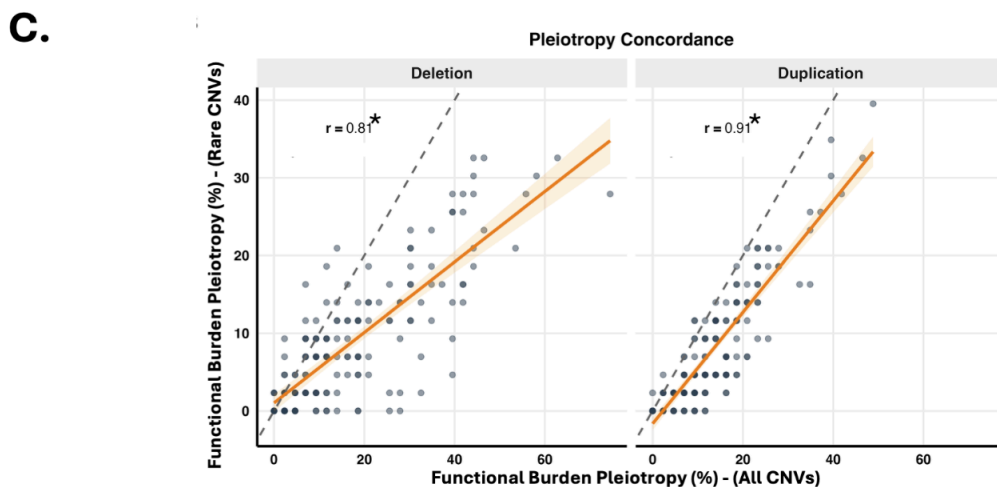
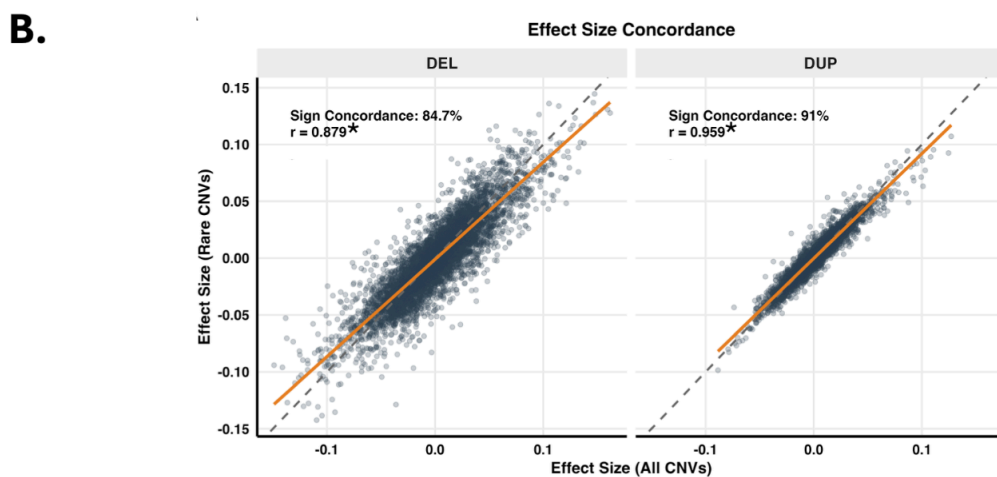
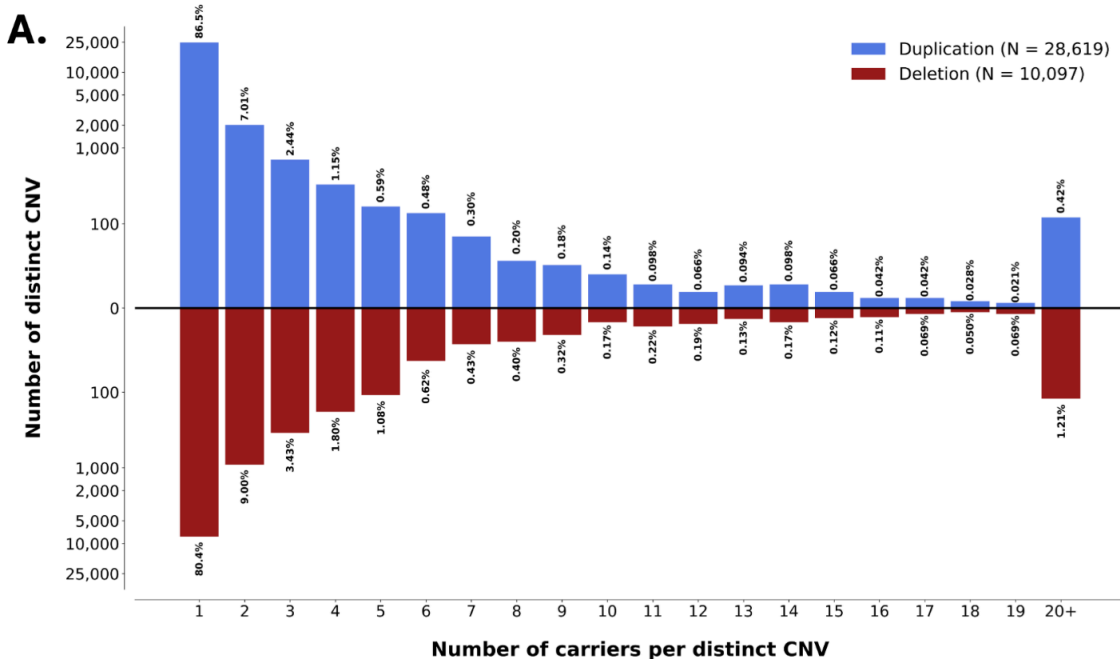
Supplementary Fig. 11. Heatmap of duplication effect sizes for whole-body tissue (HPA) across traits

As in Supplementary Fig. 10, but for duplications. Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the FunBurd model; Benjamini–Hochberg FDR across all HPA tissue gene sets $\times$ 43 traits $\times$ 2 CNV types; $q < 0.05$). Source data are provided in Source Data sheet SF11.



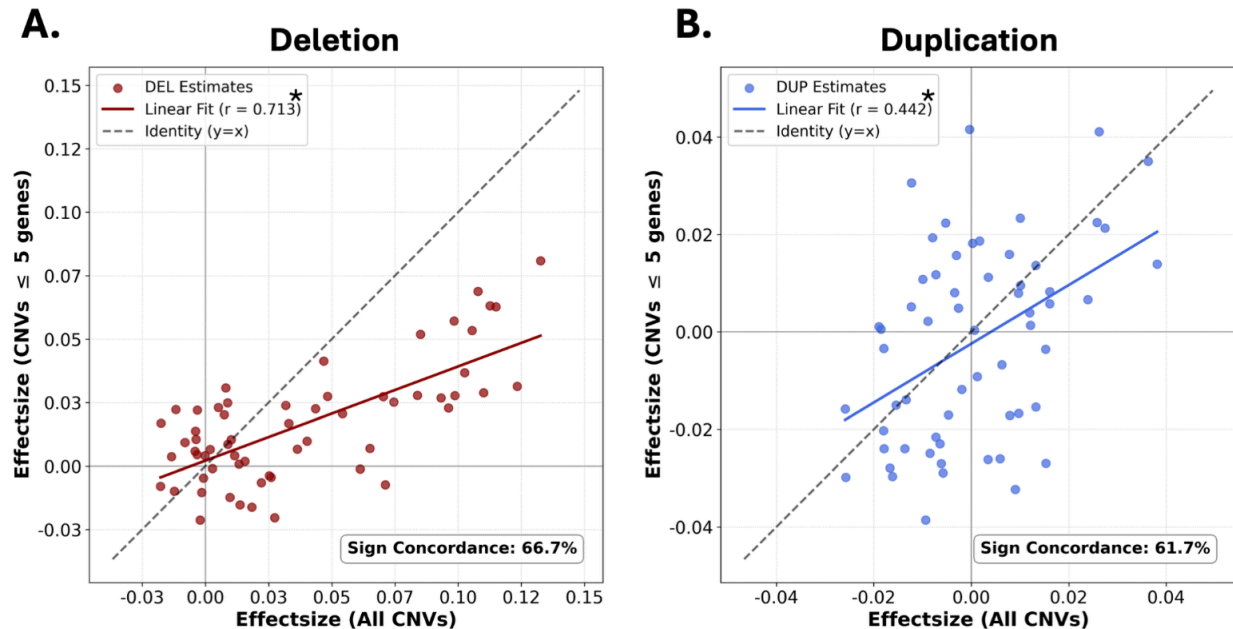
Supplementary Fig. 12. Proportion of significant associations across variant types and tissue categories on independent gene sets and traits

(A) Proportion of significant trait–gene-set associations by variant type: deletions (Del, red), duplications (Dup, blue) and single-nucleotide polymorphisms (SNP, green). (B) The same proportions stratified by tissue category (Brain, orange; Non-Brain, green) within each variant type. To account for non-independence, calculations used (i) gene sets with pairwise Jaccard overlap $\leq 20\%$ selected by LASSO regularization and (ii) 15 independent traits with mean pairwise phenotypic correlation < 0.3 . Error bars represent the standard error of the proportion. Asterisks (*) indicate significant differences between categories, determined using a two-sided two-proportion z-test with Benjamini–Hochberg FDR correction ($q < 0.05$). Source data are provided in Source Data sheets SF12A and SF12B.



Supplementary Fig. 13. Results are not related to a small number of recurrent CNVs

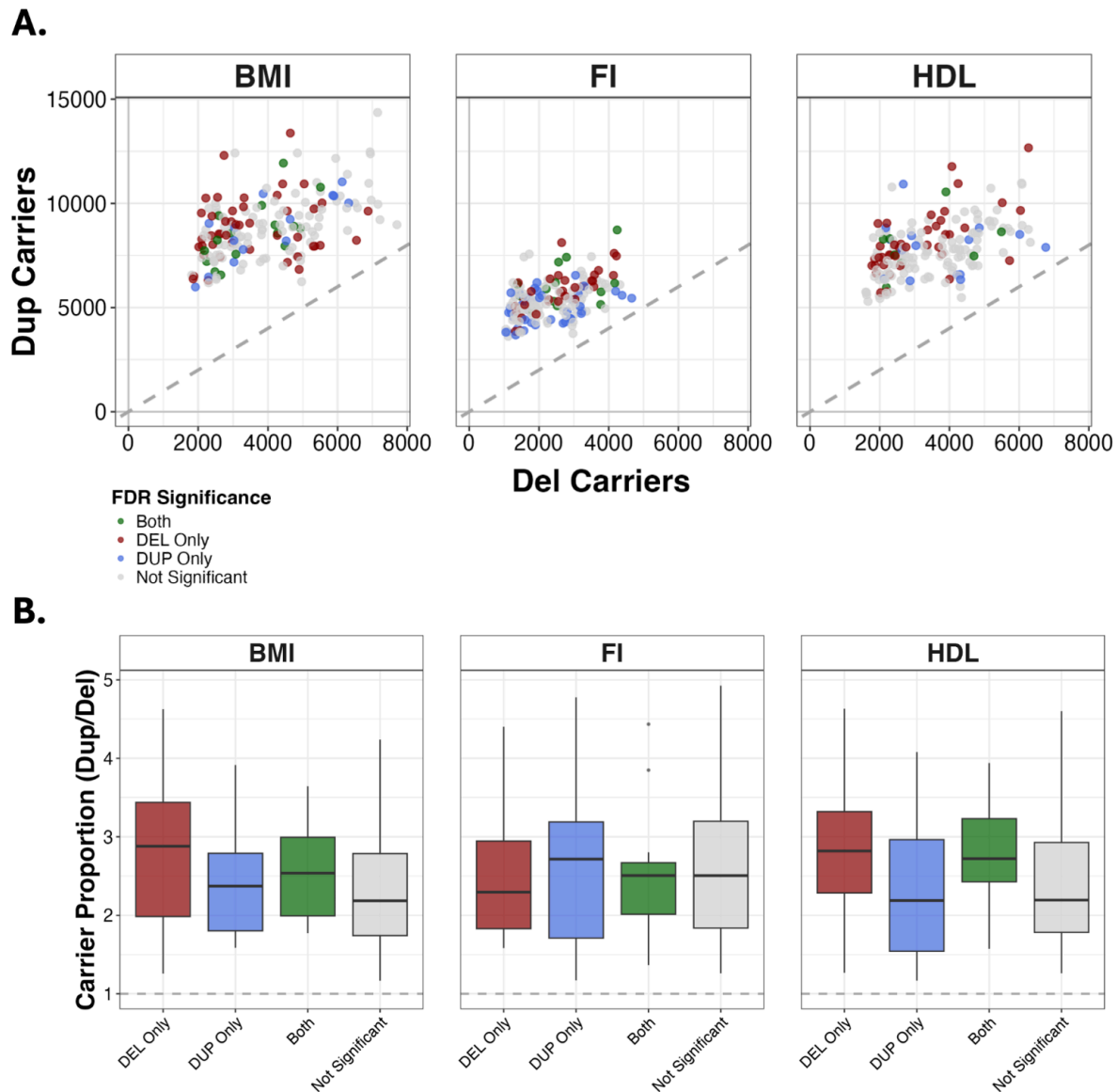
(A) Distribution of CNV Recurrence. Mirror bar chart showing the frequency of distinct duplications (N = 28,619, blue) and deletions (N = 10,097, red) based on carrier counts within the full biobank population analyzed for BMI; bar height = number of distinct CNVs at each carrier count. Most events are ultra-rare (≥ 20 carriers: 0.42% of duplications, 1.21% of deletions). (B) Effect-size concordance between estimates from all CNVs versus rare CNVs only (> 20 -carrier events excluded), across 43 traits and 172 gene sets; two-sided Pearson correlation $r = 0.88$; sign concordance 84.7% (deletions), 91% (duplications). (C) Concordance of functional burden pleiotropy rankings across the 172 gene sets (two-sided Pearson $r = 0.81$, p -Jaccard = $1.07e-41$ deletions, $r = 0.91$ duplications, p -Jaccard = $4.36e-65$). Orange lines = linear regression fits with 95% confidence intervals; dashed lines = identity ($y = x$). Source data are provided in Source Data sheets SF13B and SF13C.



Supplementary Fig. 14. Excluding large multigenic CNVs (≥ 5 Genes) shows concordant effect size profiles

Scatter plots compare burden effect sizes (β_1) from the primary model (all CNVs; x-axis) versus a model excluding CNVs encompassing ≥ 5 protein-coding genes (y-axis), across FANTOM tissue gene sets for BMI, for (A) deletions and (B) duplications. Dashed line = identity ($y = x$); solid line = linear regression fit. Concordance was assessed using a two-sided Pearson correlation; the asterisk denotes a two-sided P-Jaccard permutation test (1,000 nulls; $p < 0.05$) (A: $r = 0.713$, p -Jaccard = $1.65e-10$; B: $r = 0.442$, p -Jaccard = $4.13e-04$). Text boxes report sign concordance (percentage of gene sets retaining effect direction across models). The positive correlations and

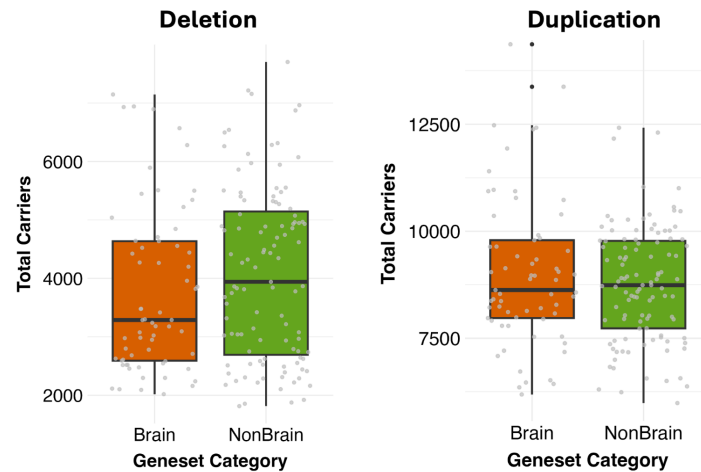
high sign concordance indicate that associations are not driven by large multigenic CNVs. Source data are provided in Source Data sheet SF14.



Supplementary Fig. 15. Specific effects of deletions and duplications are not explained by power imbalance

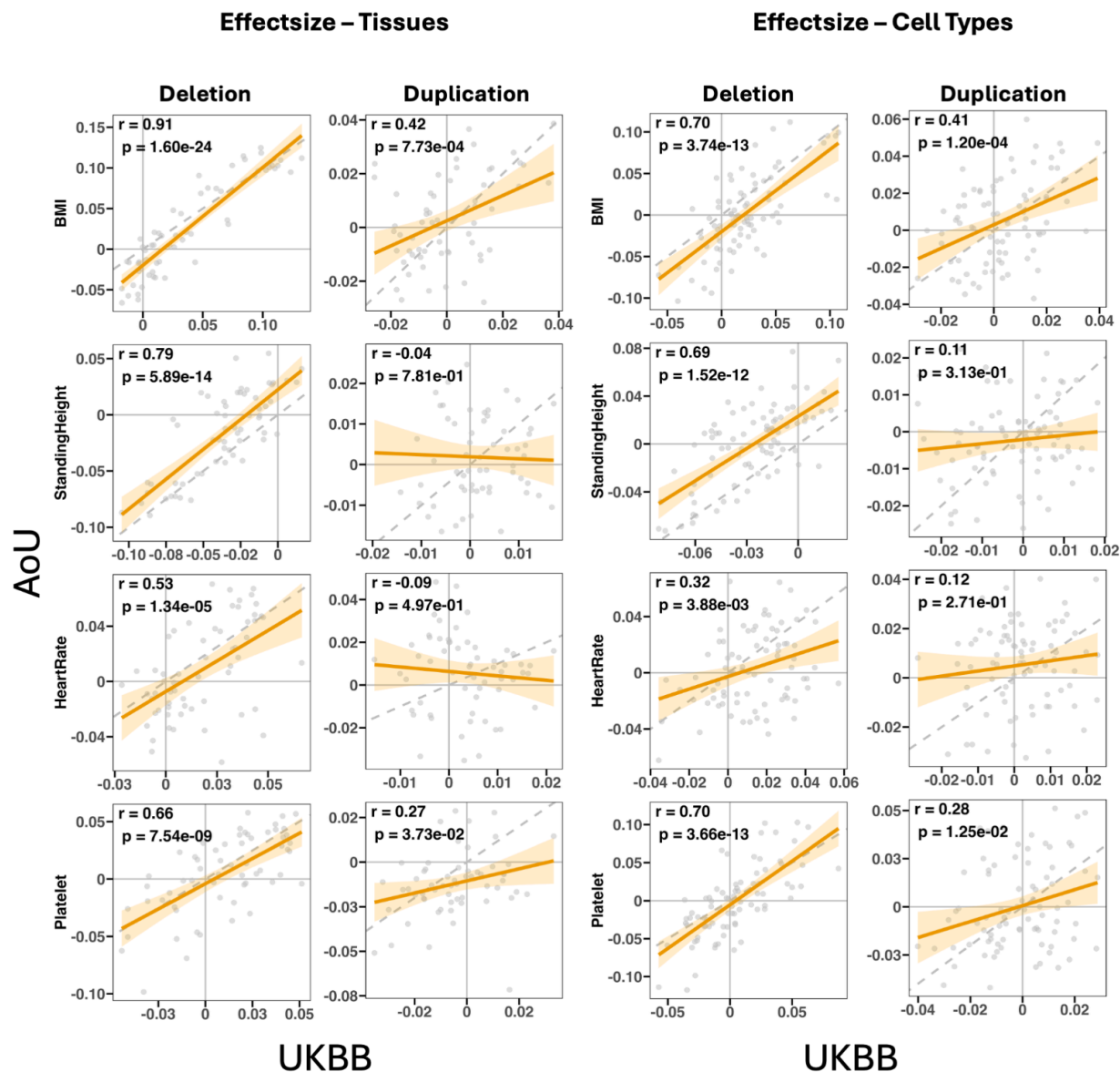
(A) Duplication versus deletion carrier counts across 172 gene sets, for three traits (BMI, Fluid Intelligence, HDL cholesterol). Points are colored by FDR significance category: both significant (green), deletion-only (red), duplication-only (blue), neither (grey); dashed line = identity ($y = x$). Across categories, duplication carriers were 2–3 $\times$ more frequent than deletion carriers, so reaching significance for one CNV type is not predicted by relative carrier number. (B) Box plots

of the duplication-to-deletion carrier ratio across significant trait–gene-set associations, by the same significance categories. Center line = median; box bounds = 25th and 75th percentiles; whiskers = $1.5 \times$ interquartile range; points beyond whiskers plotted as outliers; each point = one gene set. Pairwise differences were tested using a two-sided Wilcoxon rank-sum test with FDR correction; no comparison was significant (all $q > 0.05$). The median ratio remained between 2 and 3 across all categories. Source data are provided in Source Data sheet SF15B.



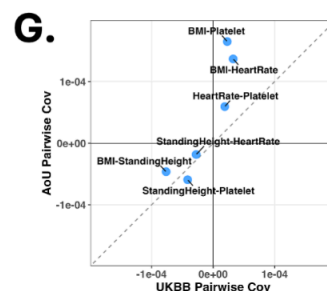
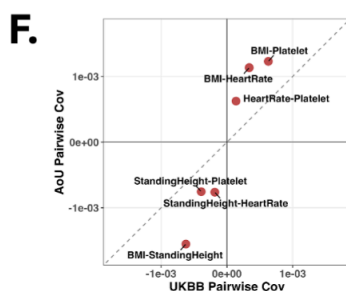
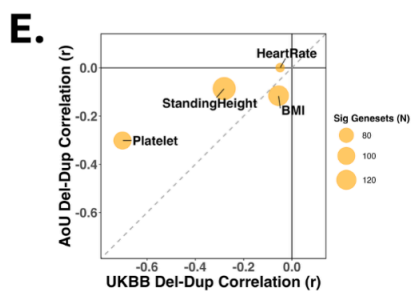
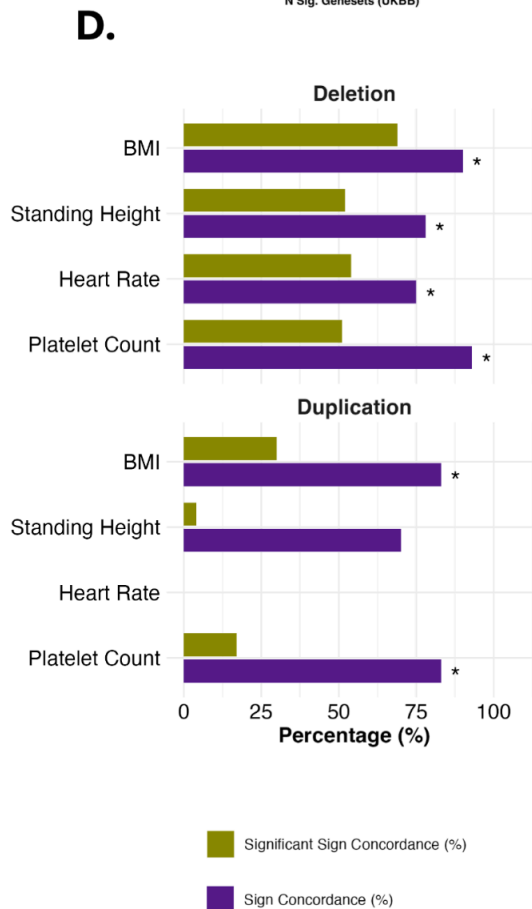
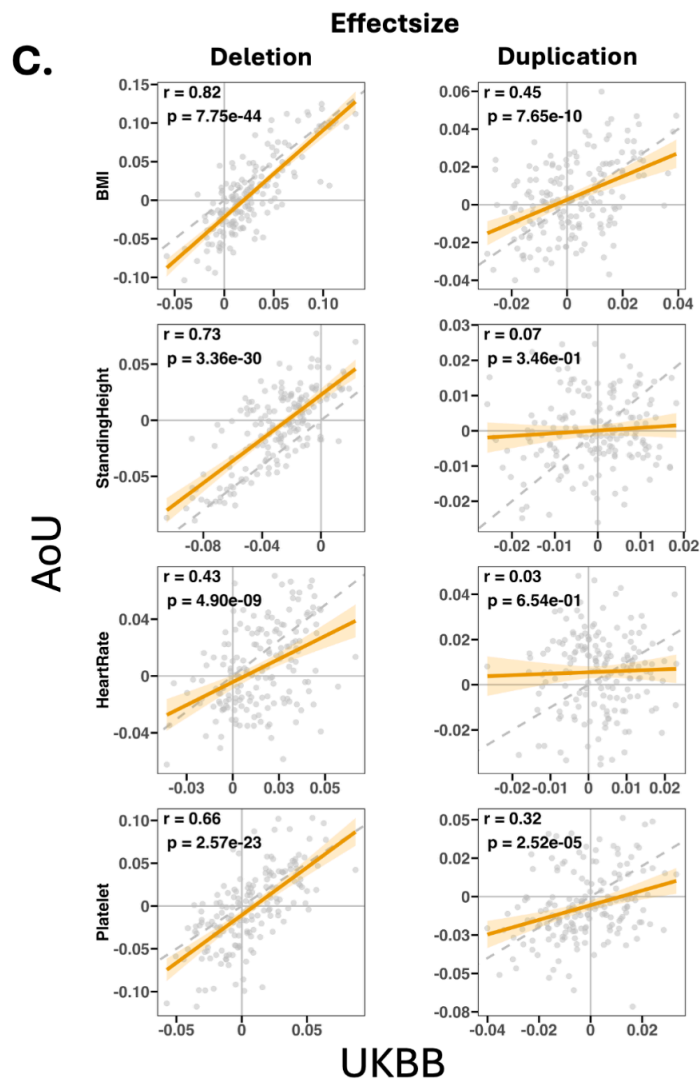
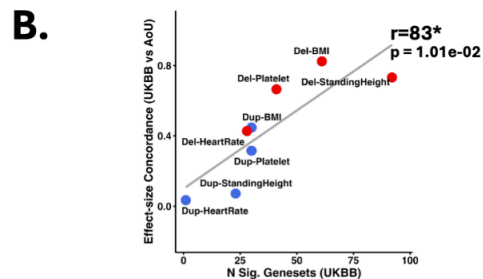
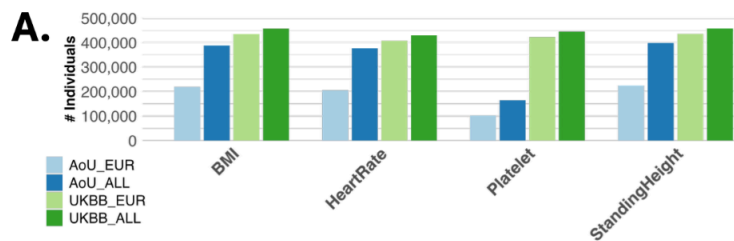
Supplementary Fig. 16. No differences in the number of carriers between the brain and non-brain gene sets

Box plots show the total number of CNV carriers for deletions (left) and duplications (right) across brain-specific and non-brain gene sets, using BMI as a representative trait; each point is one gene set ($n = 65$ brain, $n = 107$ non-brain gene sets). Center line = median; box bounds = 25th and 75th percentiles; whiskers = $1.5 \times$ interquartile range; points beyond whiskers plotted as outliers. Brain versus non-brain carrier counts were compared using a two-sided Wilcoxon rank-sum test: deletions $p = 0.15$, duplications $p = 0.69$. The absence of a difference indicates that the higher functional pleiotropy of brain gene sets is not driven by CNV frequency or statistical power. Source data are provided in Source Data sheet SF16.



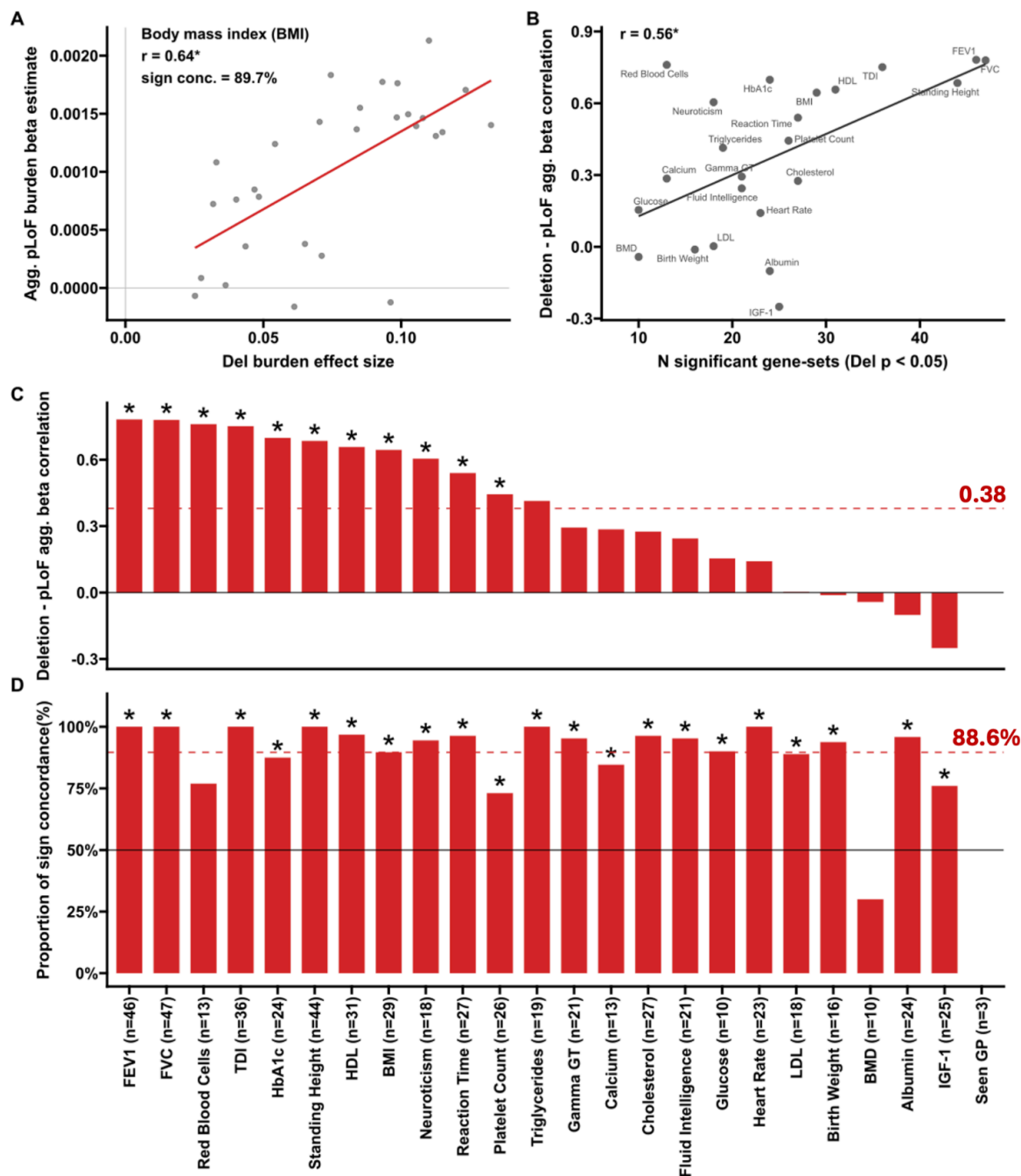
Supplementary Fig. 17. Replication of tissue-specific CNV burden associations in All of Us (AoU)

Scatter plots compare effect size estimates (β_1) between UK Biobank (x-axis) and All of Us (y-axis) across 60 whole-body tissue gene sets (left) and 81 whole-body cell-type gene sets (right), for four replication traits (BMI, Heart Rate, Platelet Count, Standing Height). Two-sided Pearson correlations (r) denote between-biobank concordance for deletions (left) and duplications (right). Orange lines and shaded bands = linear regression fits with 95% confidence intervals; dashed lines = identity ($y = x$). Full AoU estimates in Supplementary Data 6; replication statistics in Supplementary Data 7.



Supplementary Fig. 18. Replication of CNV burden associations in the All of Us (AoU) across 172 gene sets

(A) Sample sizes: participants with genotype and phenotype data in AoU and UK Biobank for the four replication traits, by European-specific (EUR) and all-ancestry (ALL) groupings; replication used European-ancestry participants to match the main analyses. (B) Replication strength vs discovery power: per-trait cross-cohort concordance (two-sided Pearson r) versus the number of FDR-significant gene sets in UKBB (Pearson correlation test, $r = 0.83$, $p = 1.01e-02$); bubble size = number of FDR-significant gene sets in UKBB. (C) Gene-set effect-size concordance: 172 gene-set effect sizes (β_1) compared between UKBB (x-axis) and AoU (y-axis) for deletions and duplications; two-sided Pearson r per comparison and FDR significance. The solid line denotes the linear regression fit, with shaded regions representing the 95% confidence intervals, alongside a dashed line representing perfect concordance ($y = x$). (D) Directional replication: sign concordance (purple) and significant sign concordance (olive; nominal same-direction $p < 0.05$) for deletions (top) and duplications (bottom). Asterisks denote one-sided exact binomial test significance against the 50% null, FDR-corrected. Across all 172 gene sets, deletion associations showed 84% sign concordance (186/222; binomial $p = 7e-26$) with 57% reaching nominal significance in the same direction; duplication associations showed 79% sign concordance (66/84; $p = 7e-8$) with 18% reaching nominal significance. (E) Per-trait deletion–duplication correlation concordance, AoU (y-axis) vs UKBB (x-axis); bubble size = number of FDR-significant gene sets. (F–G) Between-trait burden covariance concordance for (F) deletions (red) and (G) duplications (blue), UKBB (x-axis) vs AoU (y-axis), across 172 gene sets. Orange lines = linear fit with 95% CI; dashed lines = identity ($y = x$). Full statistics in Supplementary Data 6 and 7.

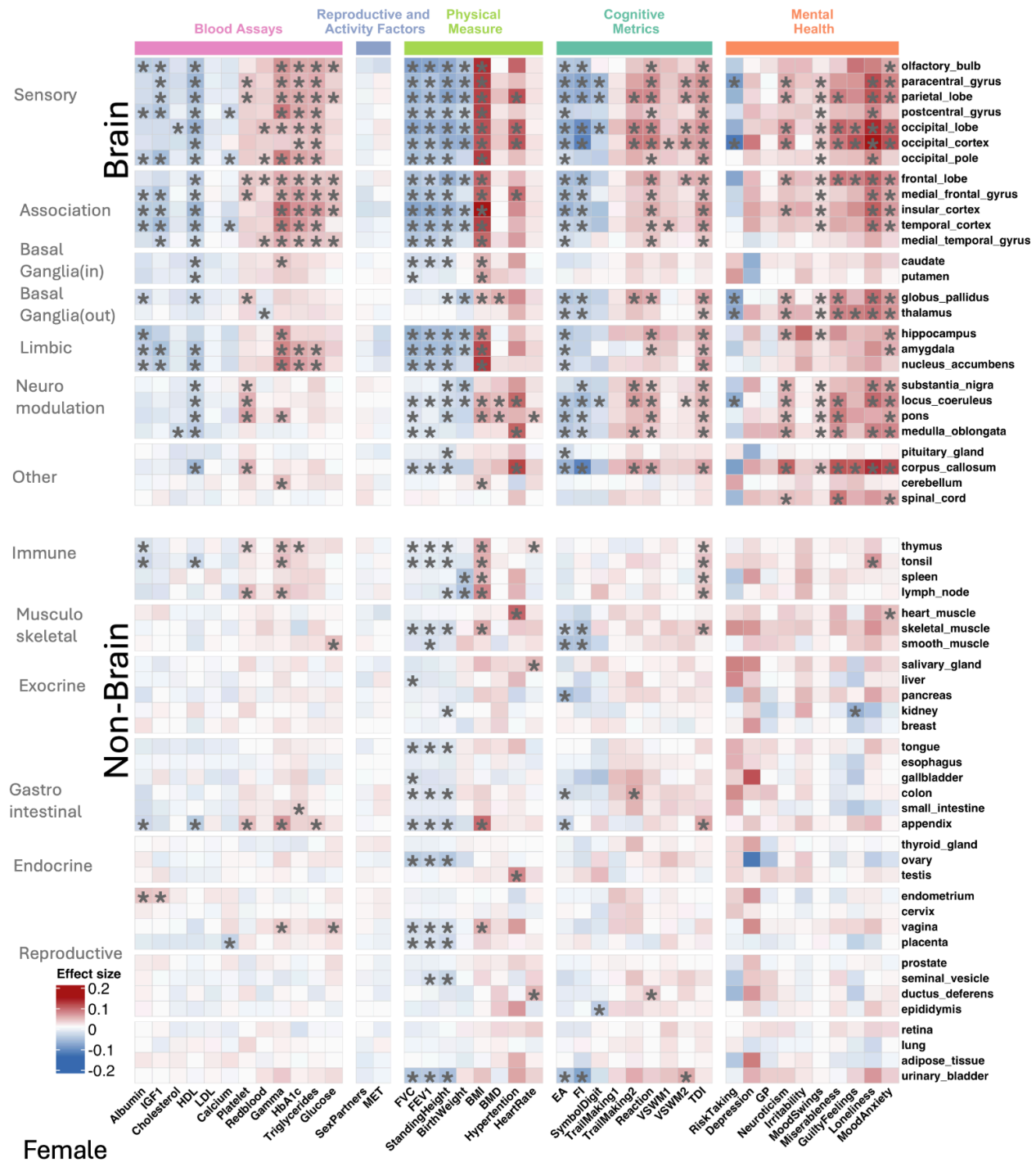


Supplementary Fig. 19. Concordance between CNV deletion burden and aggregated pLoF burden effect sizes (tissue gene-sets)

(A) CNV burden (deletion) vs aggregated pLoF burden beta estimates for an example trait (Body mass index, BMI) across nominally significant gene-sets (Deletion $p < 0.05$). Red line. linear fit; * $p\text{-Jaccard} < 0.05$. Two-sided Pearson correlation and sign concordance are reported in the panel.

(B) Per-trait Del-pLoF Pearson correlation versus the number of nominally significant gene sets

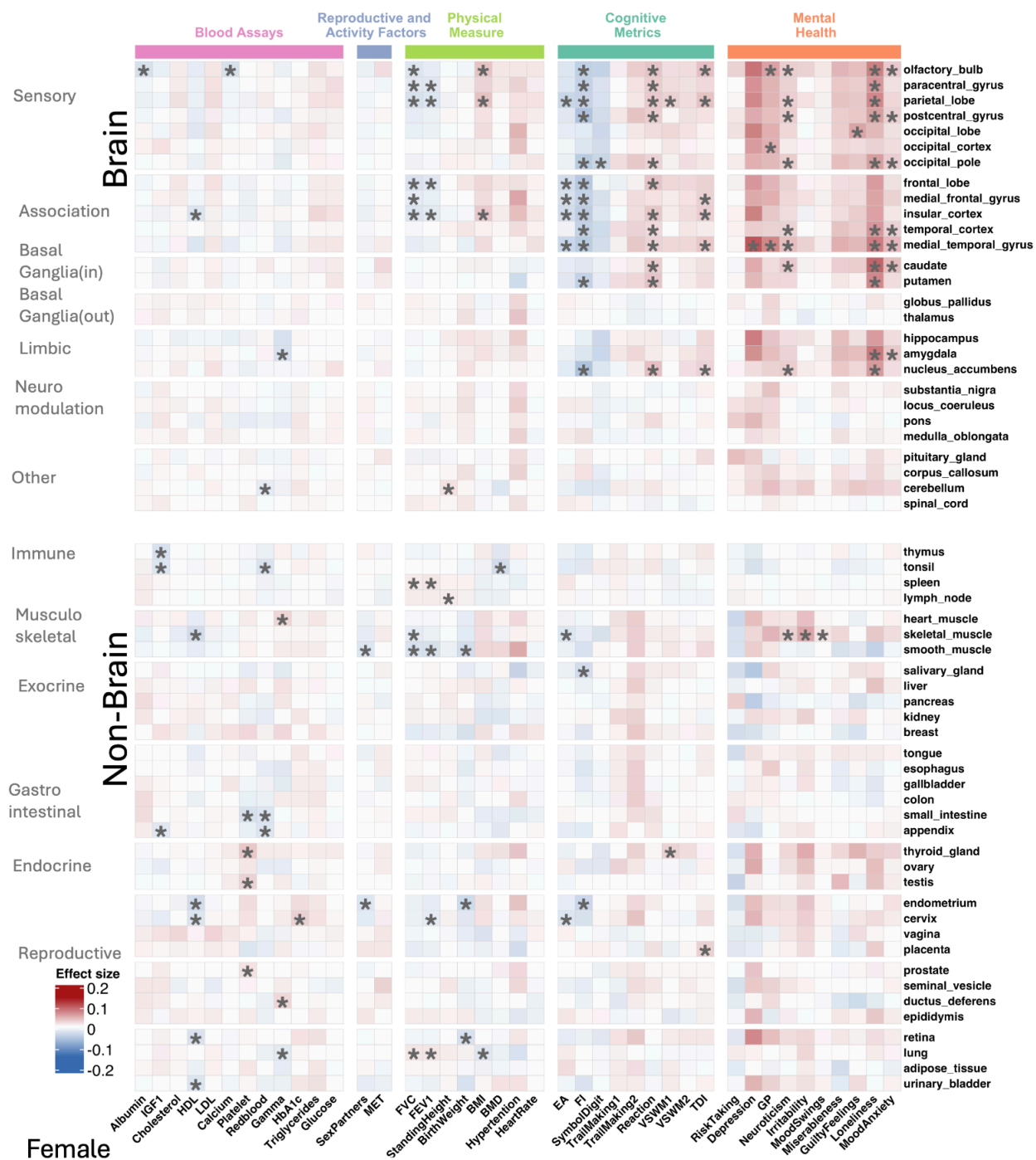
per trait (24 traits with both deletion and pLoF burden estimates). (C) Deletion–pLoF two-sided Pearson correlations across the 24 traits, ordered by decreasing r ; values in parentheses = number of significant gene sets per trait (minimum 6, i.e. 10% of the 60 tissue gene sets); traits below this threshold are shown as empty bars; dashed red line = mean r ; asterisks = Benjamini–Hochberg FDR-corrected Pearson correlation $p < 0.05$ (24 tests). (D) Sign concordance (proportion of gene sets with same-sign deletion and pLoF effects) per trait; black line = 50% chance level; dashed red line = mean; asterisks = FDR < 0.05 (one-sided exact binomial test vs. 0.5, Benjamini–Hochberg across 24 tests). Per-gene-set pLoF estimates and per-panel plotted values are in Source Data (sheets SF19A–SF19D) and Supplementary Data 8.



Supplementary Fig. 20. Heatmap of deletion effect sizes for whole-body tissue across traits for Female

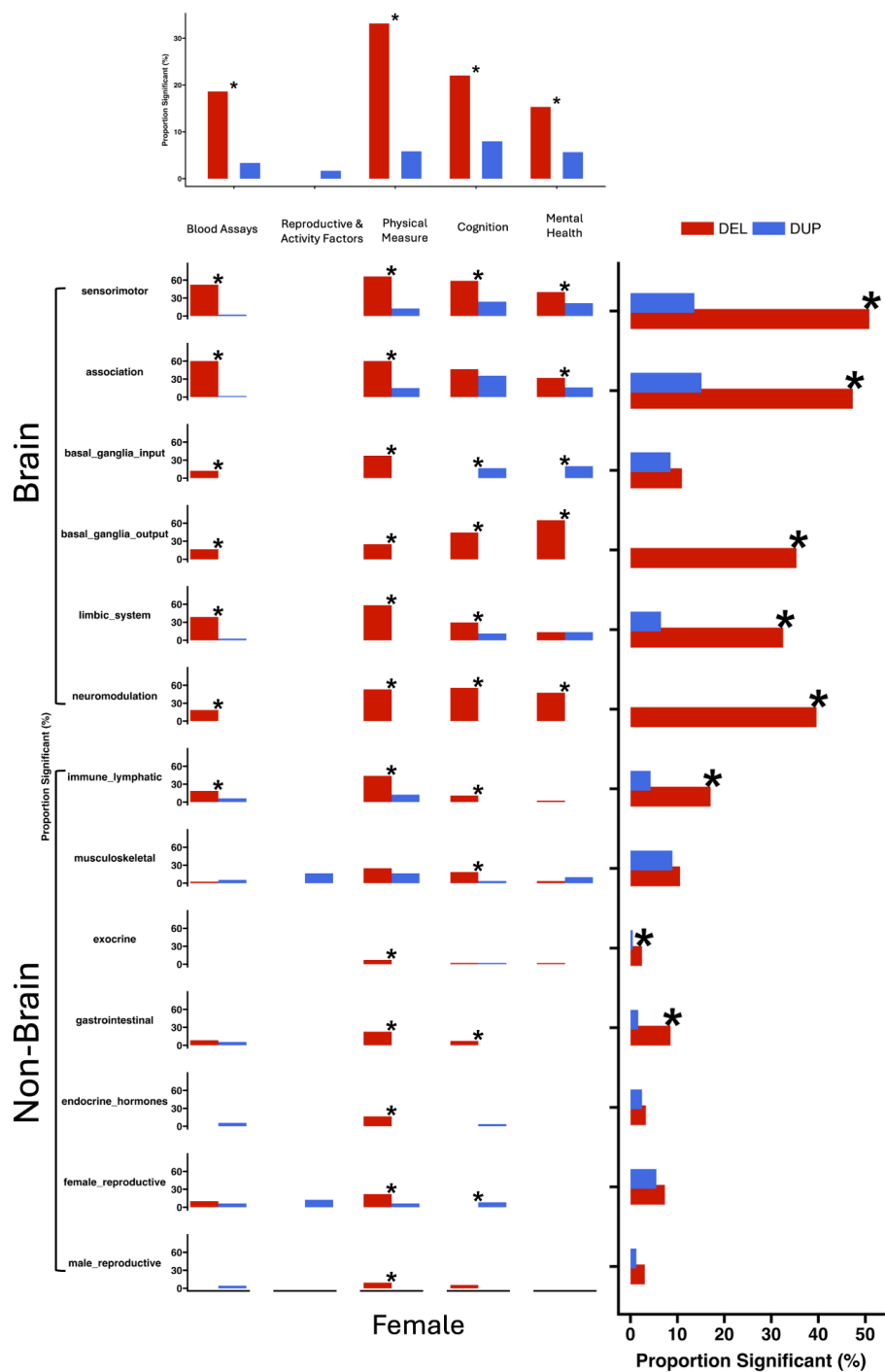
The heatmap displays female-specific deletion burden association effect sizes (β_1) between five trait categories (x-axis) and tissue-specific gene sets (y-axis). Color encodes the standardized

effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations, determined using a two-sided t-test on β_1 from the sex-stratified FunBurd model, with Benjamini–Hochberg FDR correction across 172 gene sets $\times$ 43 traits $\times$ 2 CNV types within the female stratum ($q < 0.05$). Full sex-stratified estimates in Supplementary Data 9.



Supplementary Fig. 21. Heatmap of duplication effect sizes for whole-body tissue across traits for Female

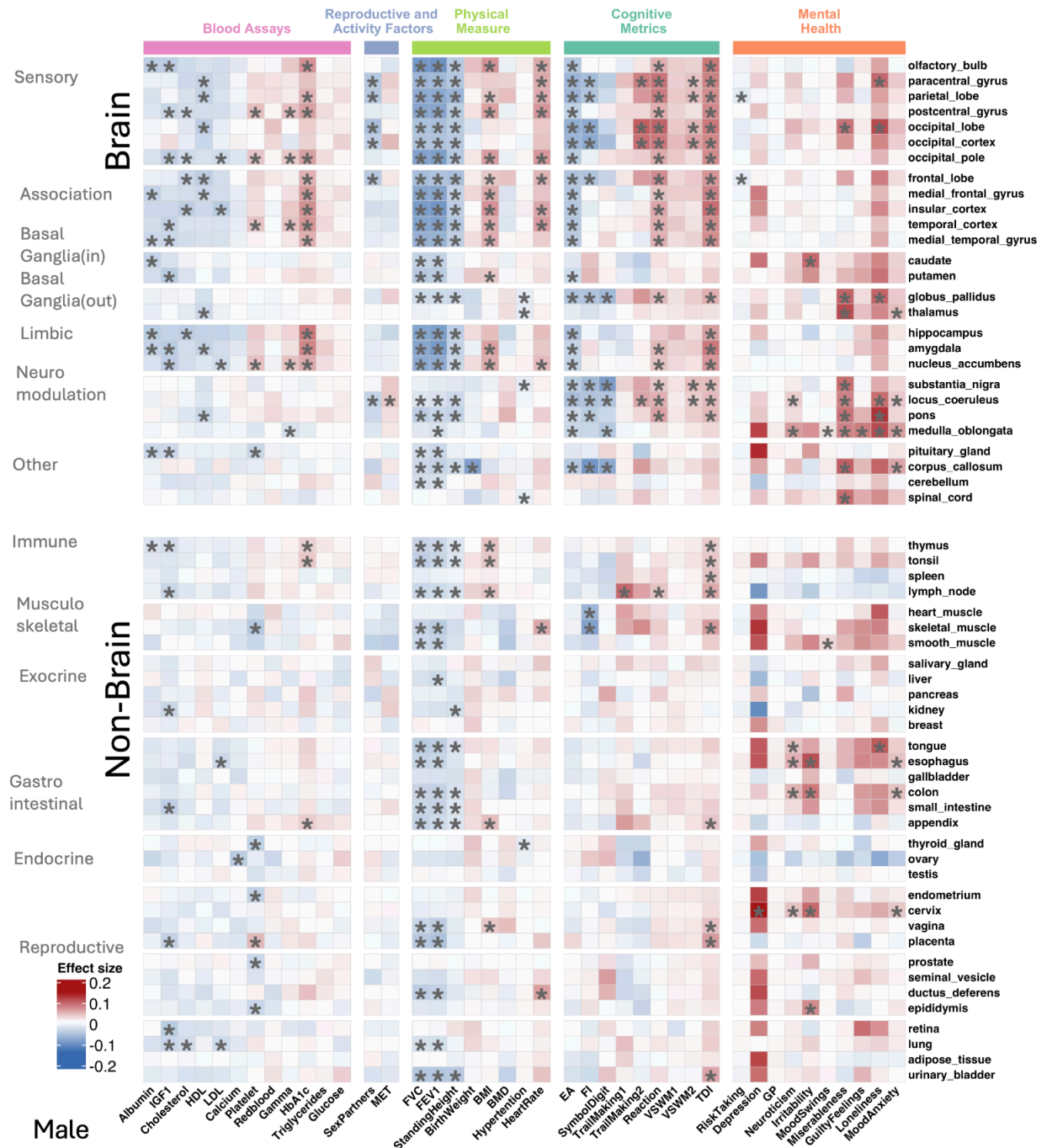
As in Supplementary Fig. 20, but for duplications (female stratum). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the sex-stratified FunBurd model; Benjamini–Hochberg FDR across $172 \times 43 \times 2$ within the female stratum; $q < 0.05$). Full statistics in Supplementary Data 9.



Supplementary Fig. 22. Differences between del-dup associations across trait and gene-set categories for Female

Bar plots summarize the difference in association level between deletions and duplications (proportion of FDR-significant associations) across five trait categories and functional gene-set

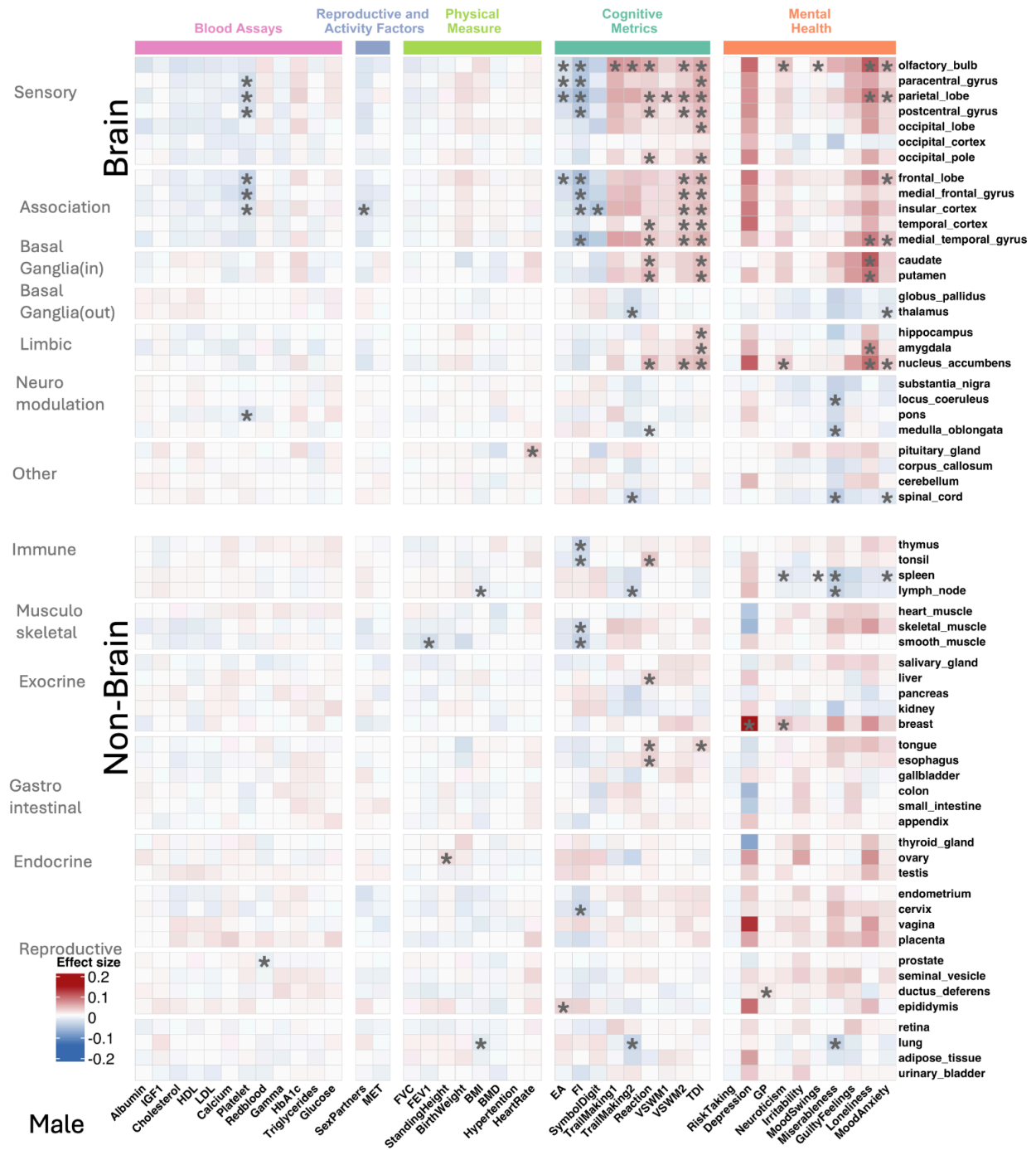
categories/annotations, for the female stratum. Bar height = aggregate proportion of significant associations; marginal plots show this per trait category (top) and gene-set category (right). Black asterisks (*) denote statistically significant proportion differences, determined using a two-sided test for equal proportions (Pearson's chi-squared test) with Benjamini–Hochberg FDR correction ($q < 0.05$). For mental-health traits, female-versus-male proportion differences were additionally tested (two-sided two-proportion z-test, FDR-corrected; e.g., deletions, Mood Swings: female 30% vs male 3.3%, FDR $p = 1.8e-3$; Supplementary Table 3). Source data are provided in Source Data sheets SF22A, SF22B, SF22C.



Supplementary Fig. 23. Heatmap of deletion effect sizes for whole-body tissue across traits for Male

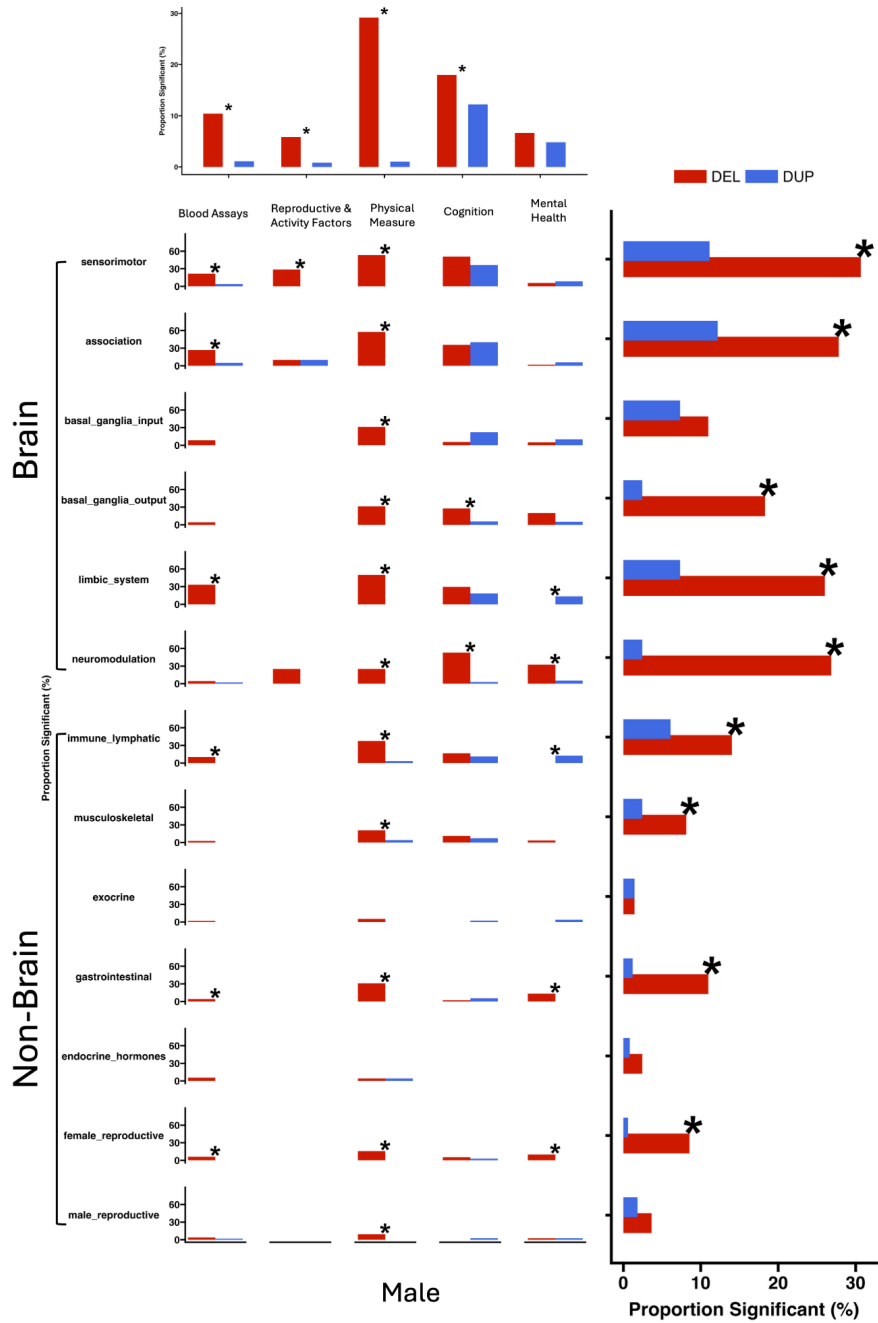
The heatmap displays male-specific deletion burden association effect sizes (β_1) between five trait categories (x-axis) and tissue-specific gene sets (y-axis). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the sex-stratified FunBurd model);

Benjamini–Hochberg FDR across $172 \times 43 \times 2$ within the male stratum; $q < 0.05$). Full statistics in Supplementary Data 9.



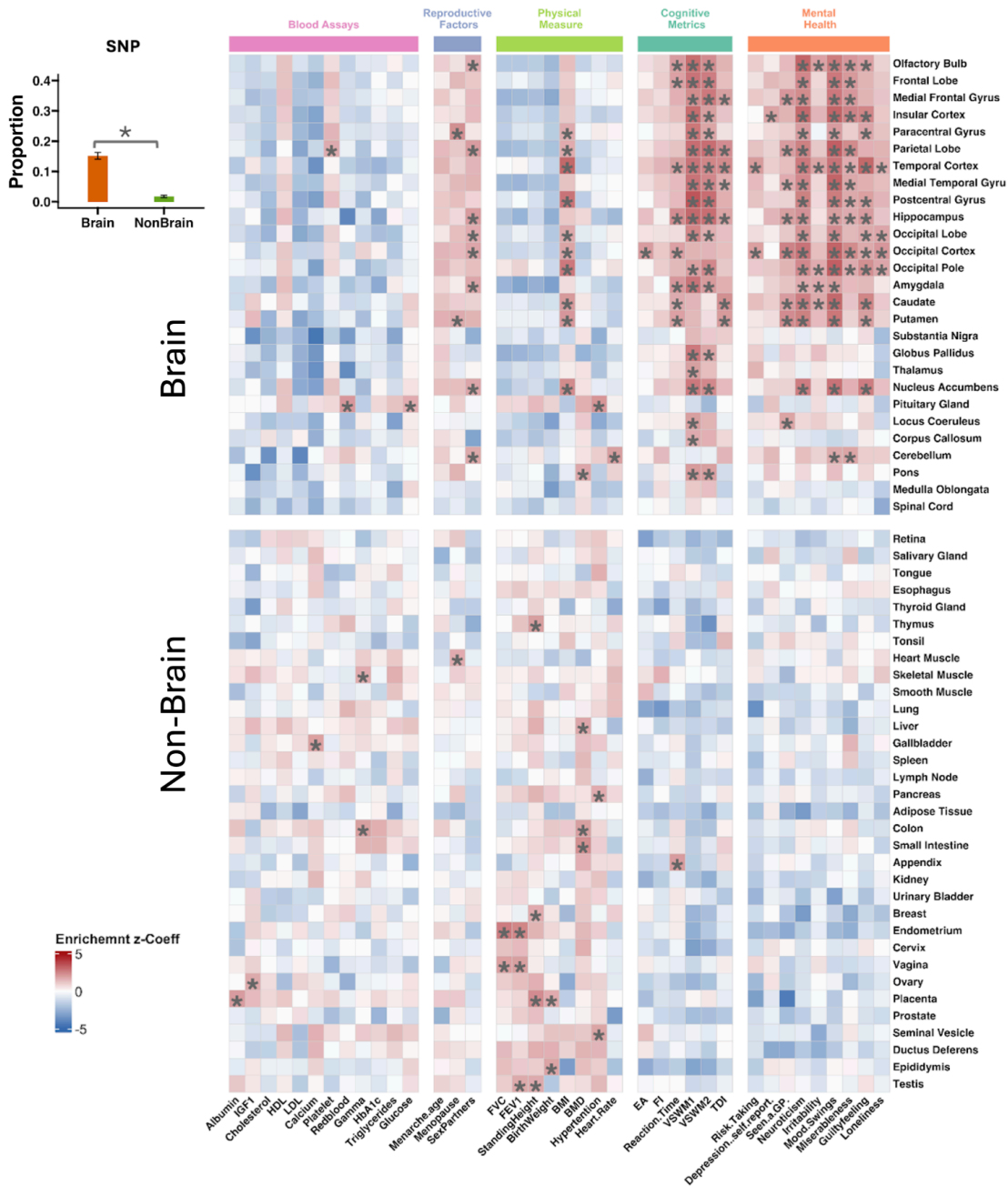
Supplementary Fig. 24. Heatmap of duplication effect sizes for whole-body tissue across traits for Male

As in Supplementary Fig. 23, but for duplications (male stratum). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the sex-stratified FunBurd model; Benjamini–Hochberg FDR across $172 \times 43 \times 2$ within the male stratum; $q < 0.05$). Full statistics in Supplementary Data 9.



Supplementary Fig. 25. Differences between del-dup associations across trait and gene-set categories for Male

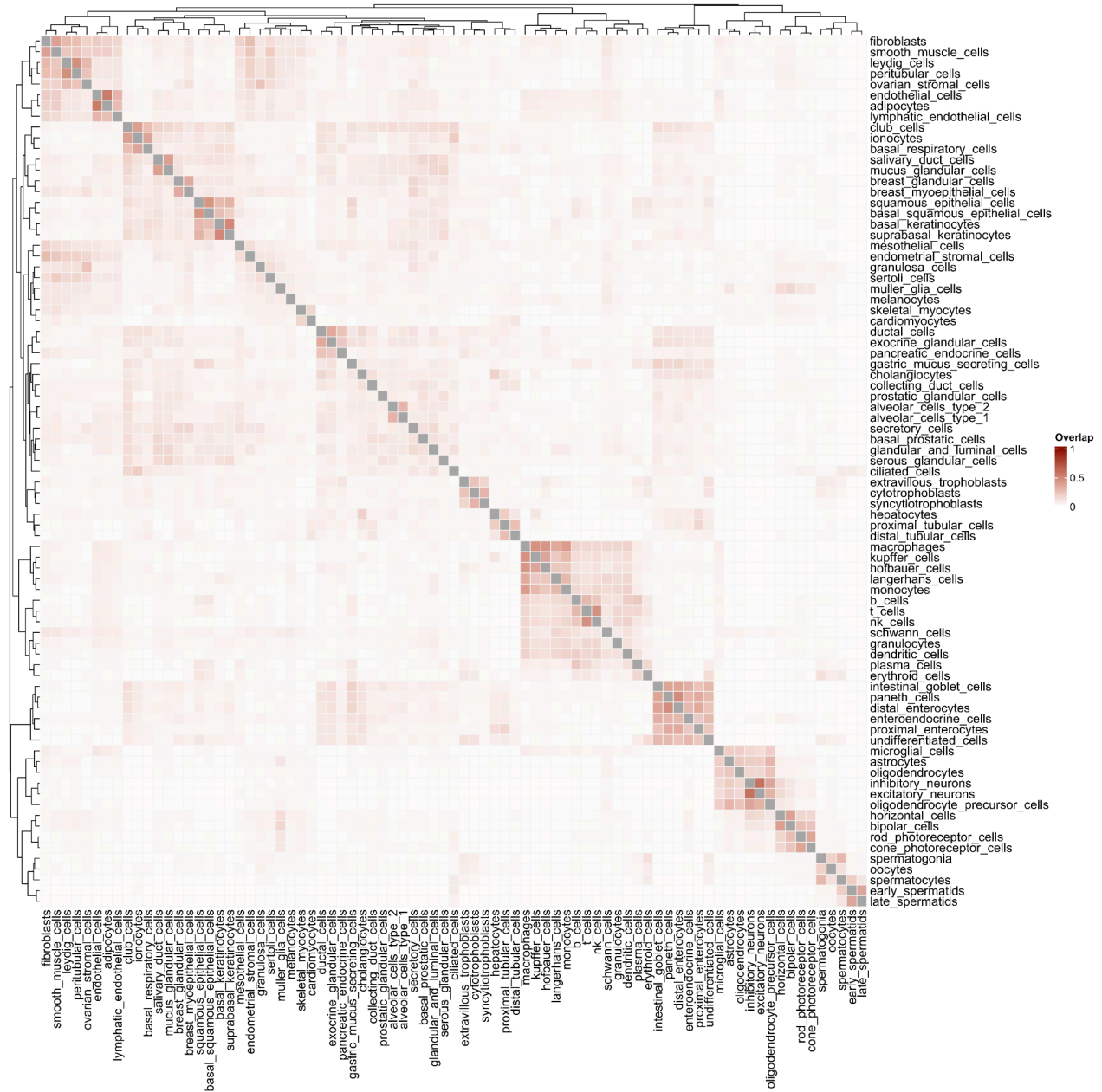
Bar plots summarize the difference in association level between deletions and duplications (proportion of FDR-significant associations) across five trait categories and functional gene-set categories/annotations, for the male stratum. Bar height = aggregate proportion of significant associations; marginal plots show this per trait category (top) and gene-set category (right). Black asterisks (*) denote statistically significant proportion differences (two-sided Pearson's chi-squared test; Benjamini–Hochberg FDR, $q < 0.05$). Source data are provided in Source Data sheets SF25A, SF25B, SF25C.



Supplementary Fig. 26. Heatmap of GWAS enrichments for whole-body tissues across traits using S-LDSC

The heatmap displays stratified LD-score regression (S-LDSC) GWAS enrichments for five trait categories (x-axis) across tissue-specific gene sets (y-axis). Color encodes the S-LDSC enrichment regression coefficient Z-score (blue = negative; red = positive; centered at zero, spanning approximately -6.2 to $+5.7$). Black asterisks (*) indicate nominal significance of the coefficient Z-score, determined using a one-sided test ($p = 1 - \Phi(Z)$, reflecting the directional

enrichment hypothesis; nominal $\alpha = 0.05$, no multiple-comparison adjustment). Per-trait/gene-set enrichment, standard error, and p-values are in Supplementary Data 10. Abbreviation: S-LDSC, stratified LD-score regression.



Supplementary Fig. 27. Pairwise overlap matrix of whole-body gene sets

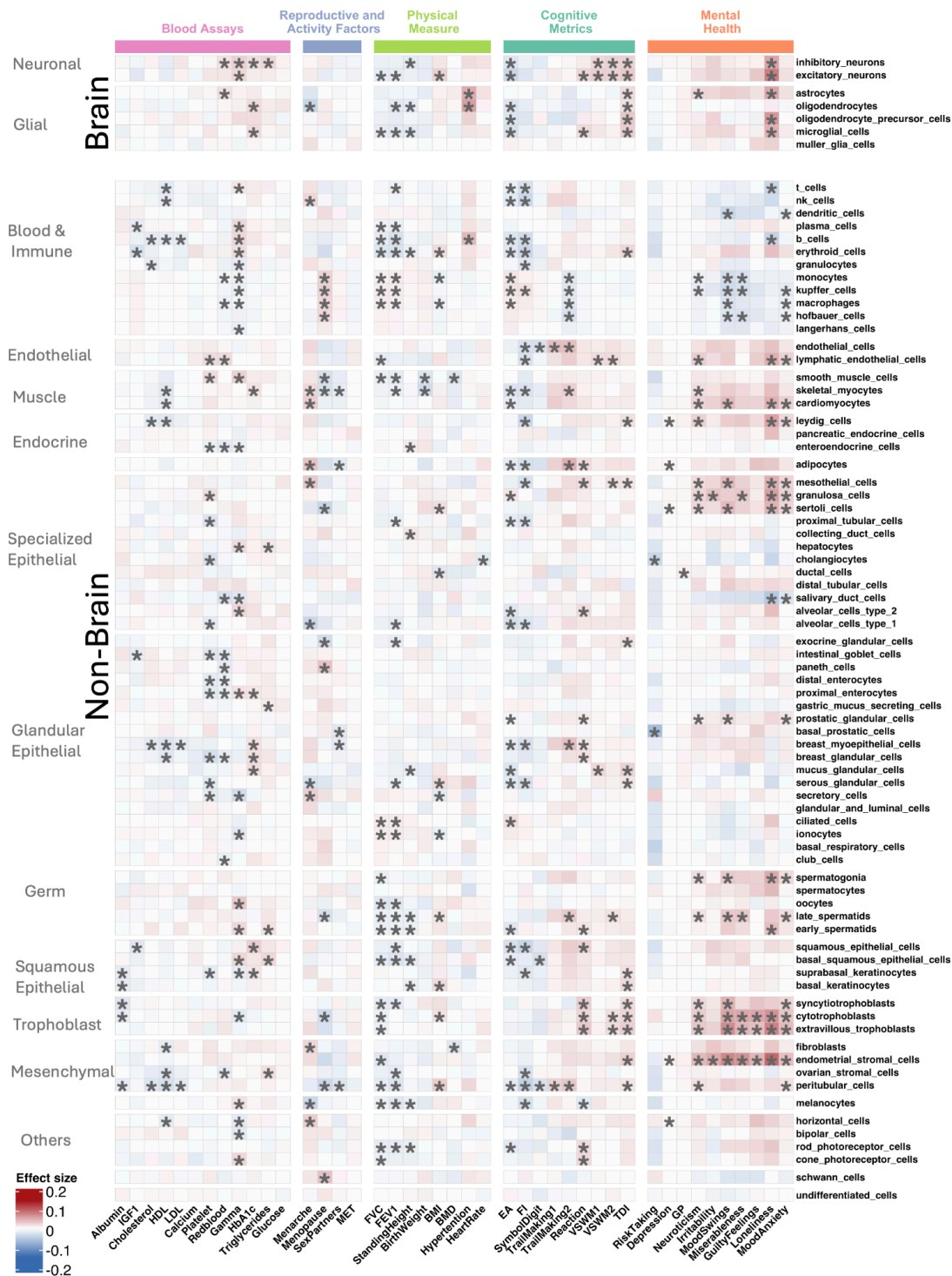
The heatmap displays the pairwise overlap (1 – Jaccard distance) between the 81 whole-body cell-type gene sets. Color encodes overlap magnitude: dark red = high overlap, white = low (scale 0–1). The maximum, minimum, mean, and median overlap values are 58.8%, 0.3%, 4.6%,

and 3.0%, respectively. Descriptive similarity matrix; no statistical test applied. Pairwise values in Supplementary Data 3.



Supplementary Fig. 28. Heatmap of deletion effect sizes for whole-body cell types across traits

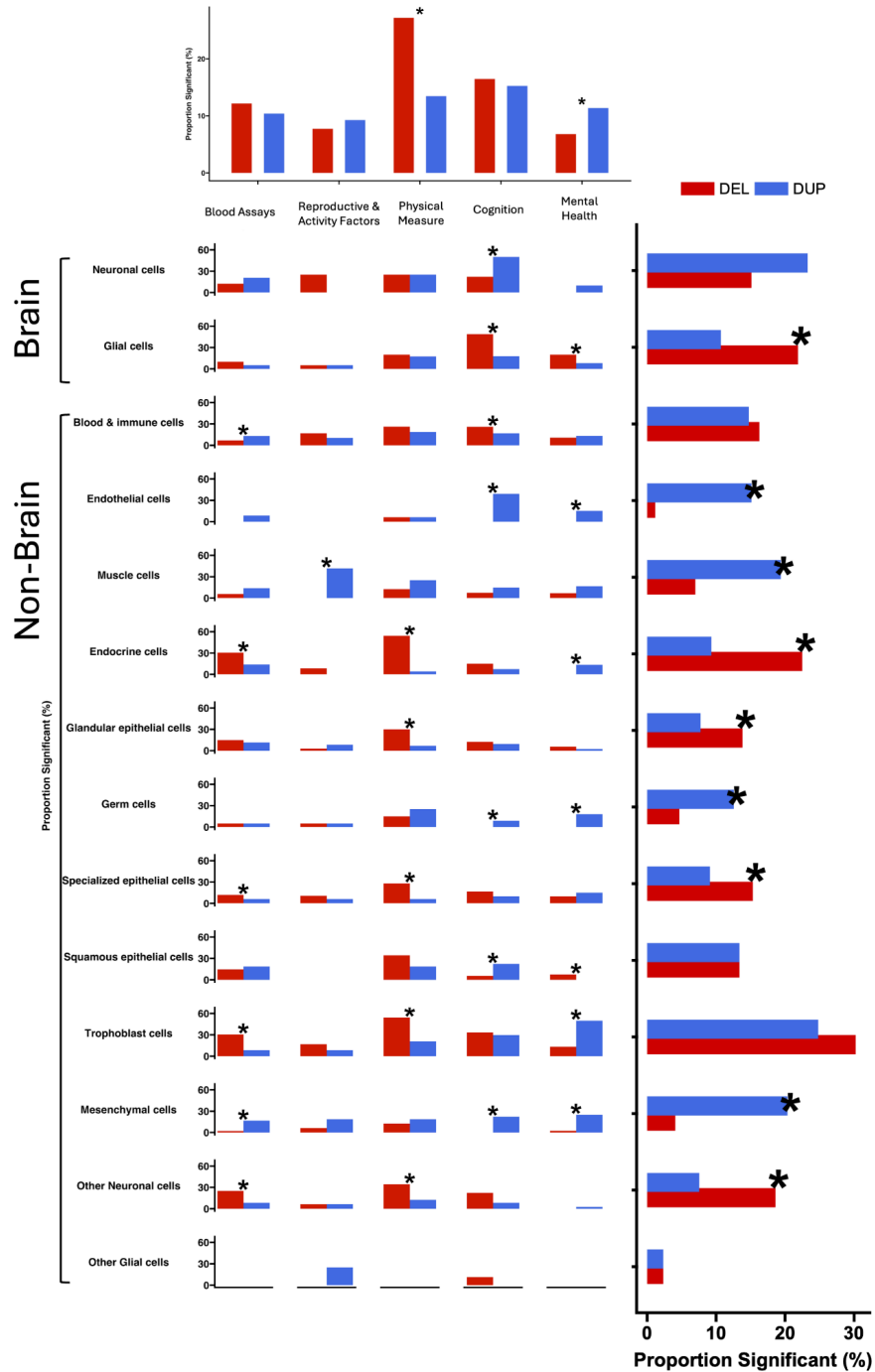
The heatmap displays deletion burden association effect sizes (β_1) between five trait categories (x-axis) and whole-body cell-type-specific gene sets (y-axis; categories/annotations at right). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations, determined using a two-sided t-test on β_1 from the FunBurd multiple linear regression model, with Benjamini–Hochberg FDR correction across 172 gene sets $\times$ 43 traits $\times$ 2 CNV types ($q < 0.05$). Full statistics in Supplementary Data 4.



Supplementary Fig. 29. Heatmap of duplication effect sizes for whole-body cell types across traits

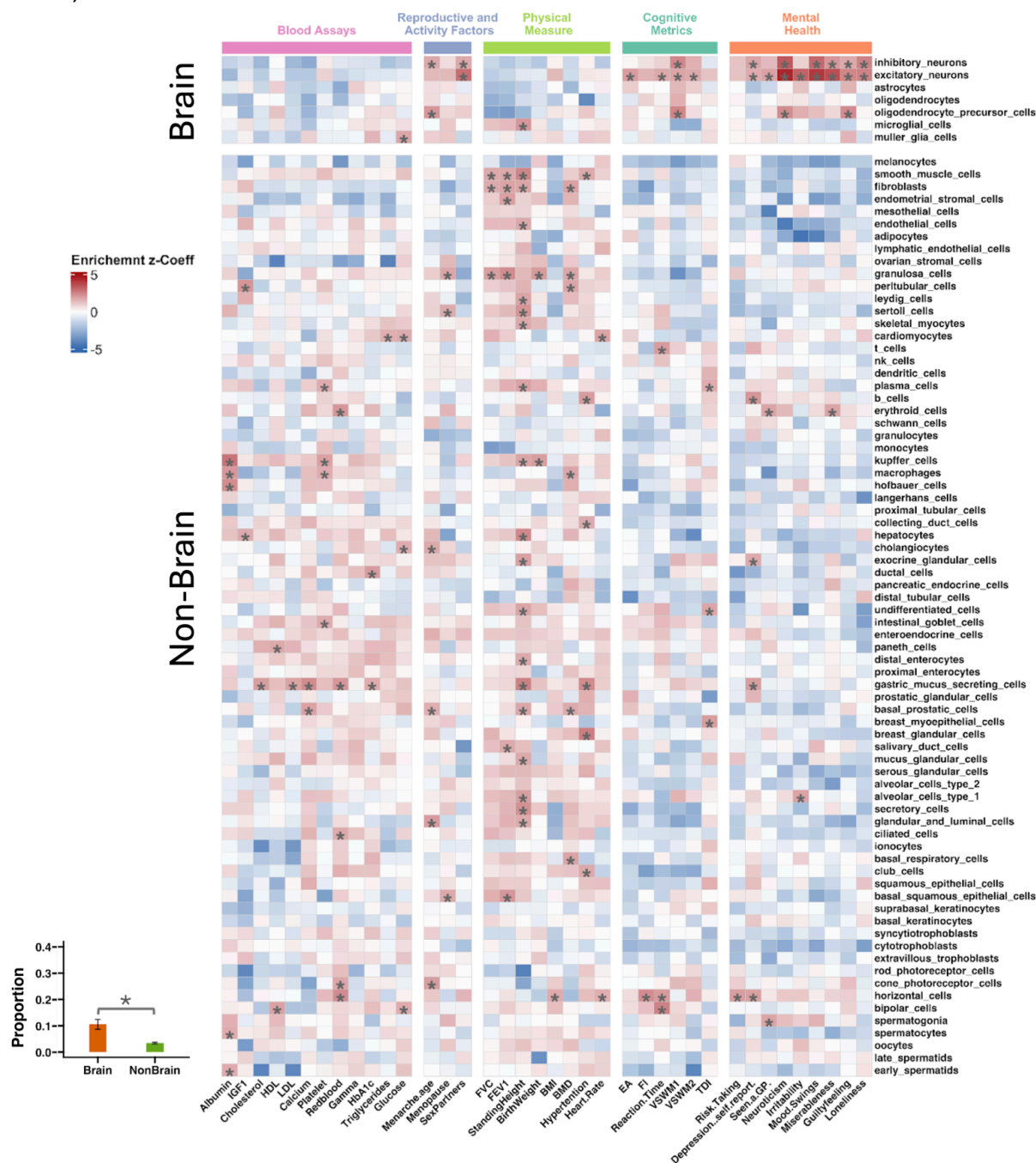
As in Supplementary Fig. 28 but for duplications. The heatmap displays duplication burden association effect sizes (β_1) between five trait categories (x-axis) and whole-body cell-type-specific gene sets (y-axis; categories/annotations at right). Color encodes the

standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the FunBurd model; Benjamini–Hochberg FDR across $172 \times 43 \times 2$; $q < 0.05$). Full statistics in Supplementary Data 4.



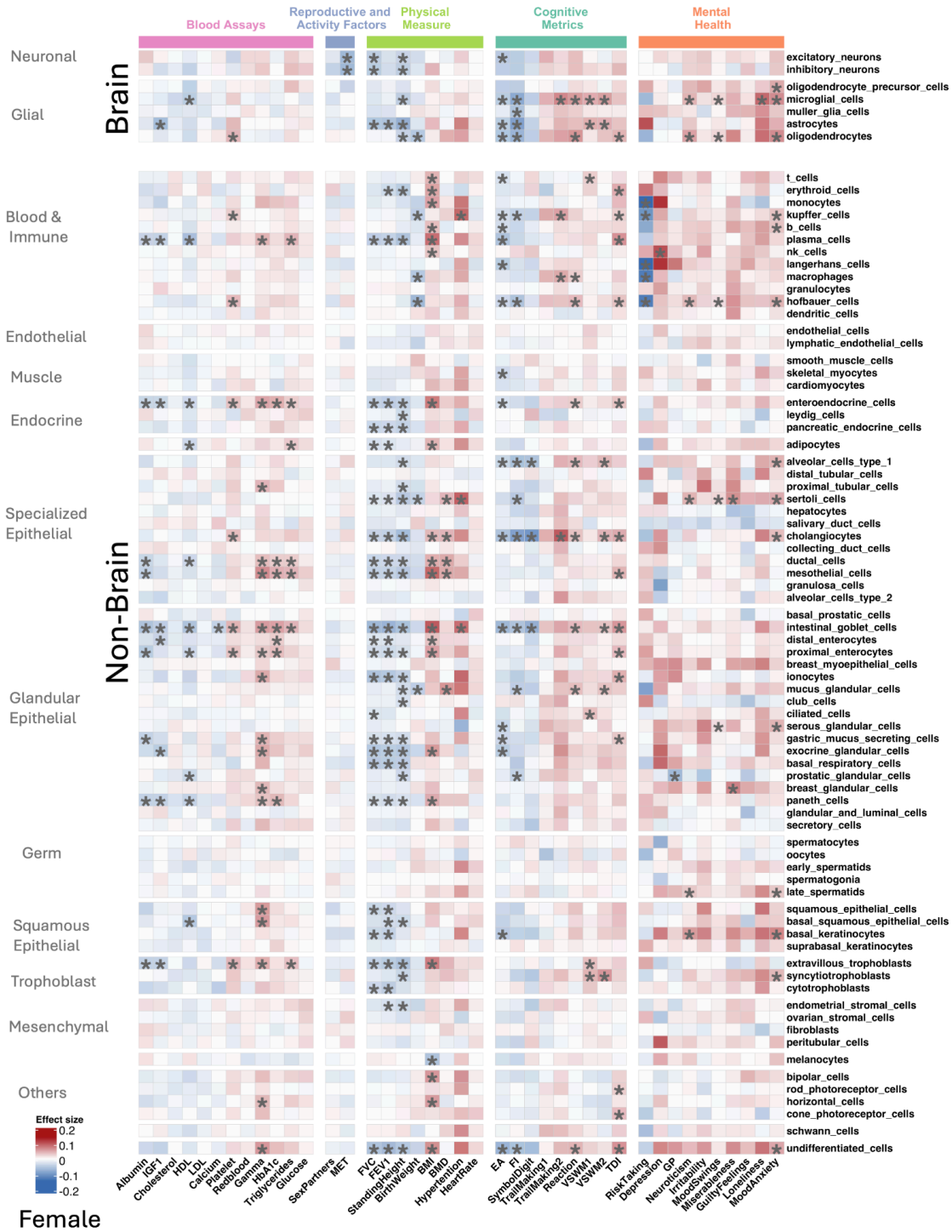
Supplementary Fig. 30. Differences between del-dup associations across trait and cell-type gene-set categories

Bar plots summarize the difference in association level between deletions and duplications (proportion of FDR-significant associations) across five trait categories and whole-body cell-type gene-set categories/annotations. Bar height = aggregate proportion of significant associations; marginal plots show this per trait category (top) and gene-set category (right). Black asterisks (*) denote statistically significant proportion differences (two-sided Pearson's chi-squared test; Benjamini–Hochberg FDR, $q < 0.05$). Source data are provided in Source Data sheets SF30A, SF30B, SF30C.



Supplementary Fig. 31. Heatmap of GWAS enrichments for whole-body cell types across traits using S-LDSC

The heatmap displays stratified LD-score regression (S-LDSC) GWAS enrichments for five trait categories (x-axis) across whole-body cell-type gene sets (y-axis). Color encodes the S-LDSC enrichment regression coefficient Z-score (blue = negative; red = positive; centered at zero, ≈ -6.2 to $+5.7$). Black asterisks (*) indicate nominal significance of the coefficient Z-score (one-sided test, $p = 1 - \Phi(Z)$; nominal $\alpha = 0.05$, no multiple-comparison adjustment). Statistics in Supplementary Data 10.



Supplementary Fig. 32. Heatmap of deletion effect sizes for whole-body cell-type across traits for Female

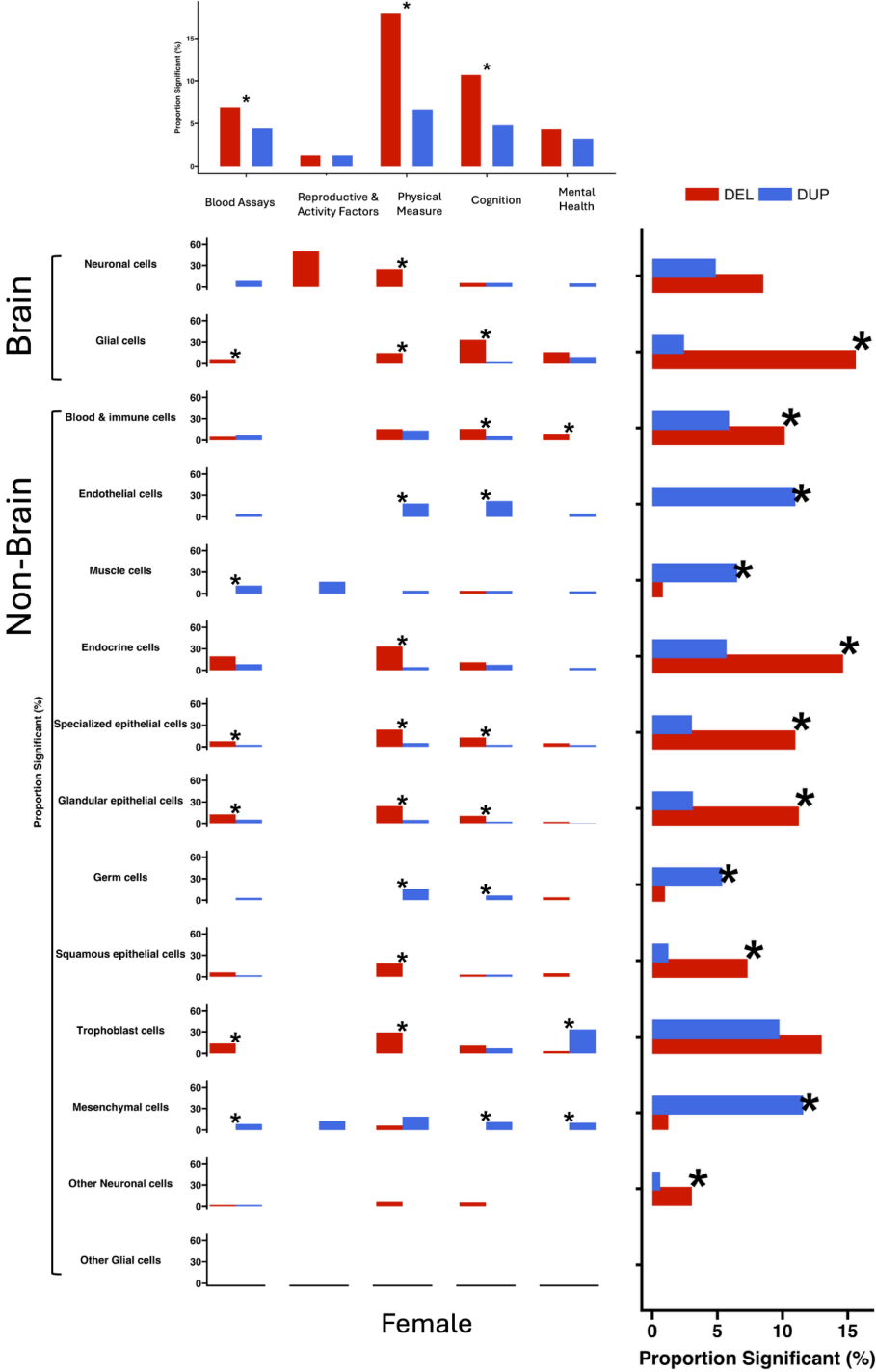
The heatmap displays female-specific deletion burden association effect sizes (β_1) between five trait categories (x-axis) and whole-body cell-type-specific gene sets (y-axis; categories/annotations at right). Color encodes the standardized effect size (blue = negative;

red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the sex-stratified FunBurd model; Benjamini–Hochberg FDR within the female stratum; $q < 0.05$). Full statistics in Supplementary Data 9.



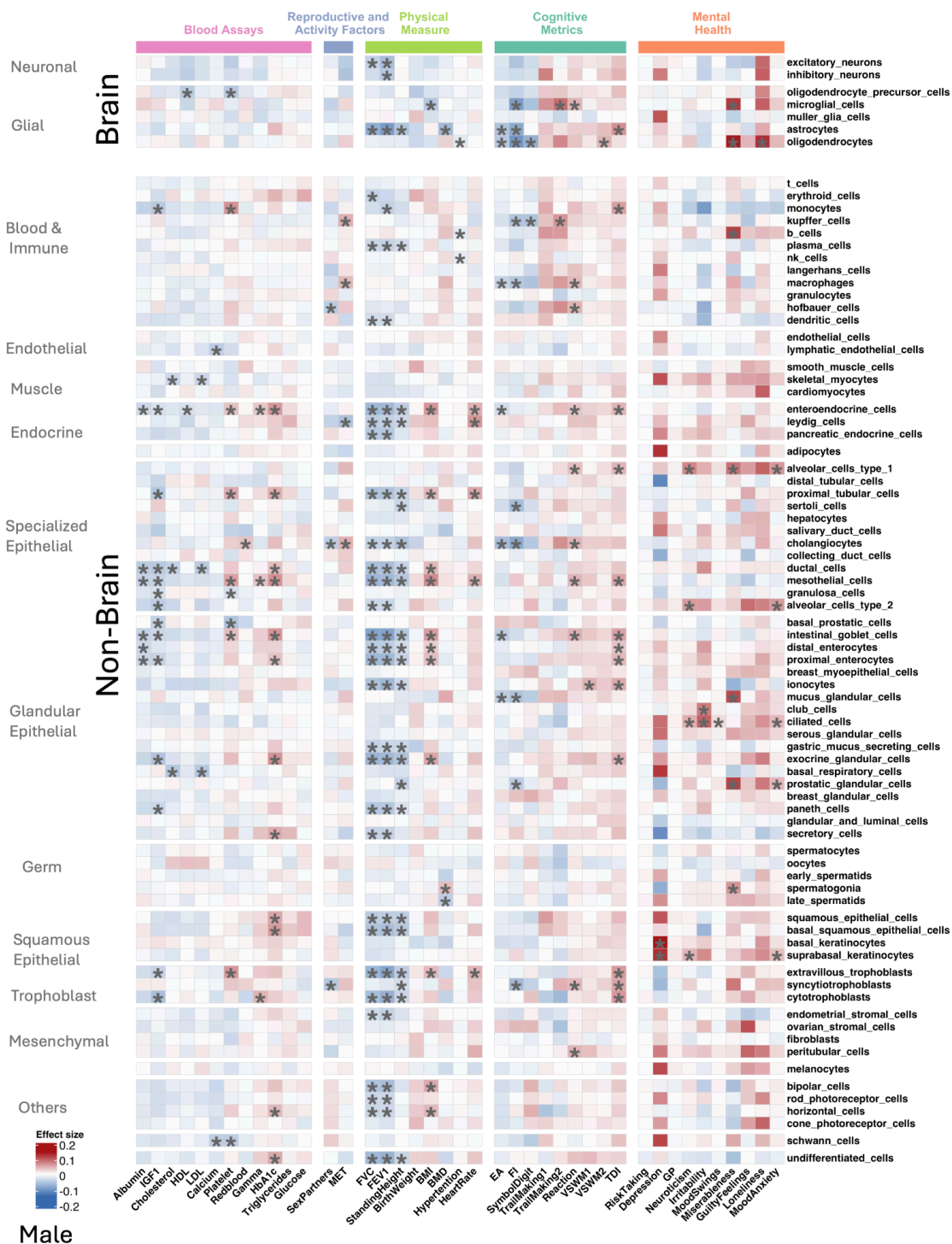
Supplementary Fig. 33. Heatmap of duplication effect sizes for whole-body cells across traits for Female

As in Supplementary Fig. 32 but for duplications (female stratum). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the sex-stratified FunBurd model; Benjamini–Hochberg FDR within the female stratum; $q < 0.05$). Full statistics in Supplementary Data 9.



Supplementary Fig. 34. Differences between del-dup associations across trait and cell-type gene-set categories for Female

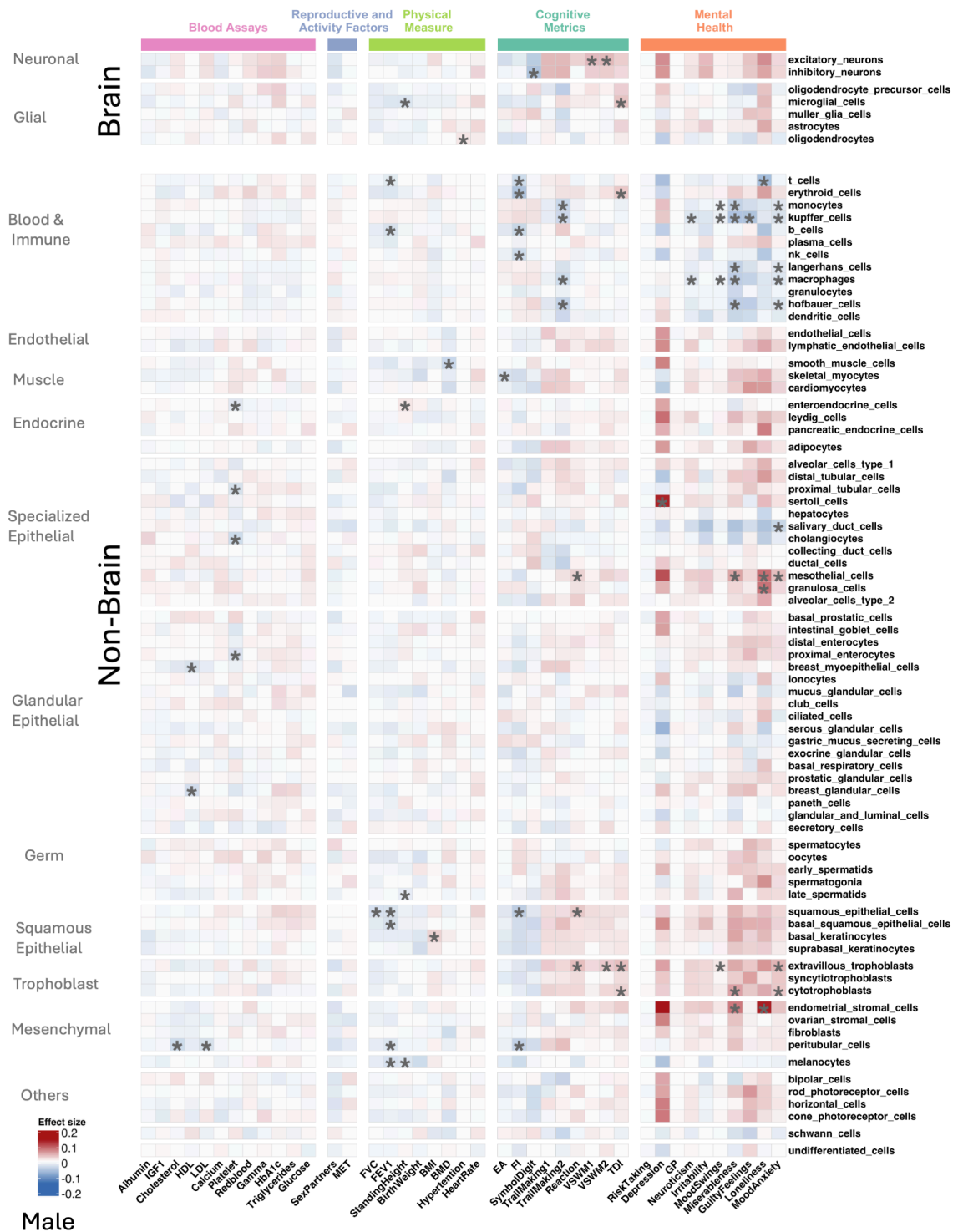
Bar plots summarize the difference in association level between deletions and duplications (proportion of FDR-significant associations) across five trait categories and whole-body cell-type gene-set categories/annotations, for the female stratum. Bar height = aggregate proportion of significant associations; marginal plots show this per trait category (top) and gene-set category (right). Black asterisks (*) denote statistically significant proportion differences (two-sided Pearson's chi-squared test; Benjamini–Hochberg FDR, $q < 0.05$). Source data are provided in Source Data sheets SF34A, SF34B, SF34C.



Supplementary Fig. 35. Heatmap of deletion effect sizes for whole-body cell type across traits for Male

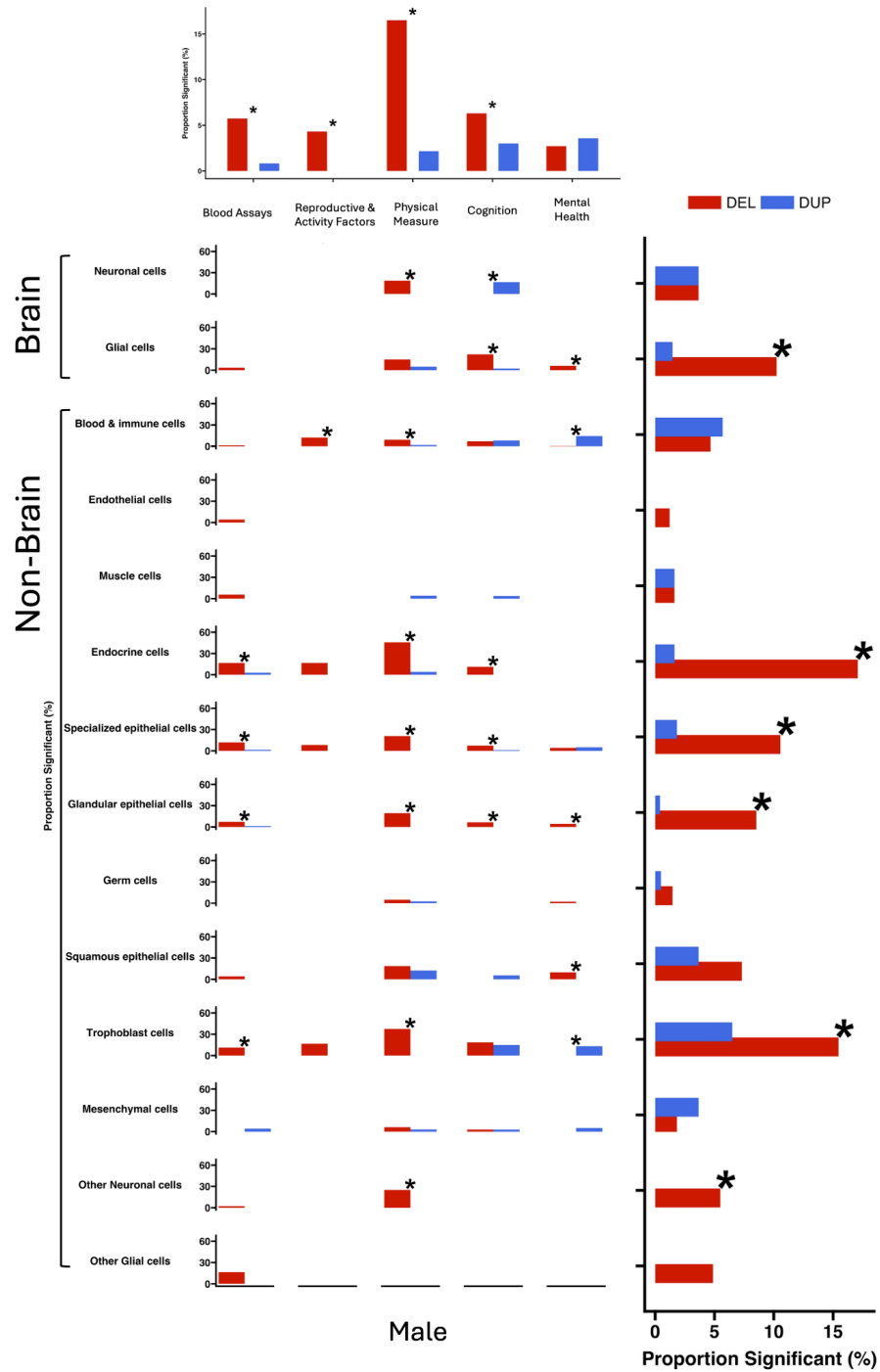
The heatmap displays male-specific deletion burden association effect sizes (β_1) between five trait categories (x-axis) and whole-body cell-type-specific gene sets (y-axis; categories/annotations at right). Color encodes the standardized effect size (blue = negative;

red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the sex-stratified FunBurd model; Benjamini–Hochberg FDR within the male stratum; $q < 0.05$). Full statistics in Supplementary Data 9.



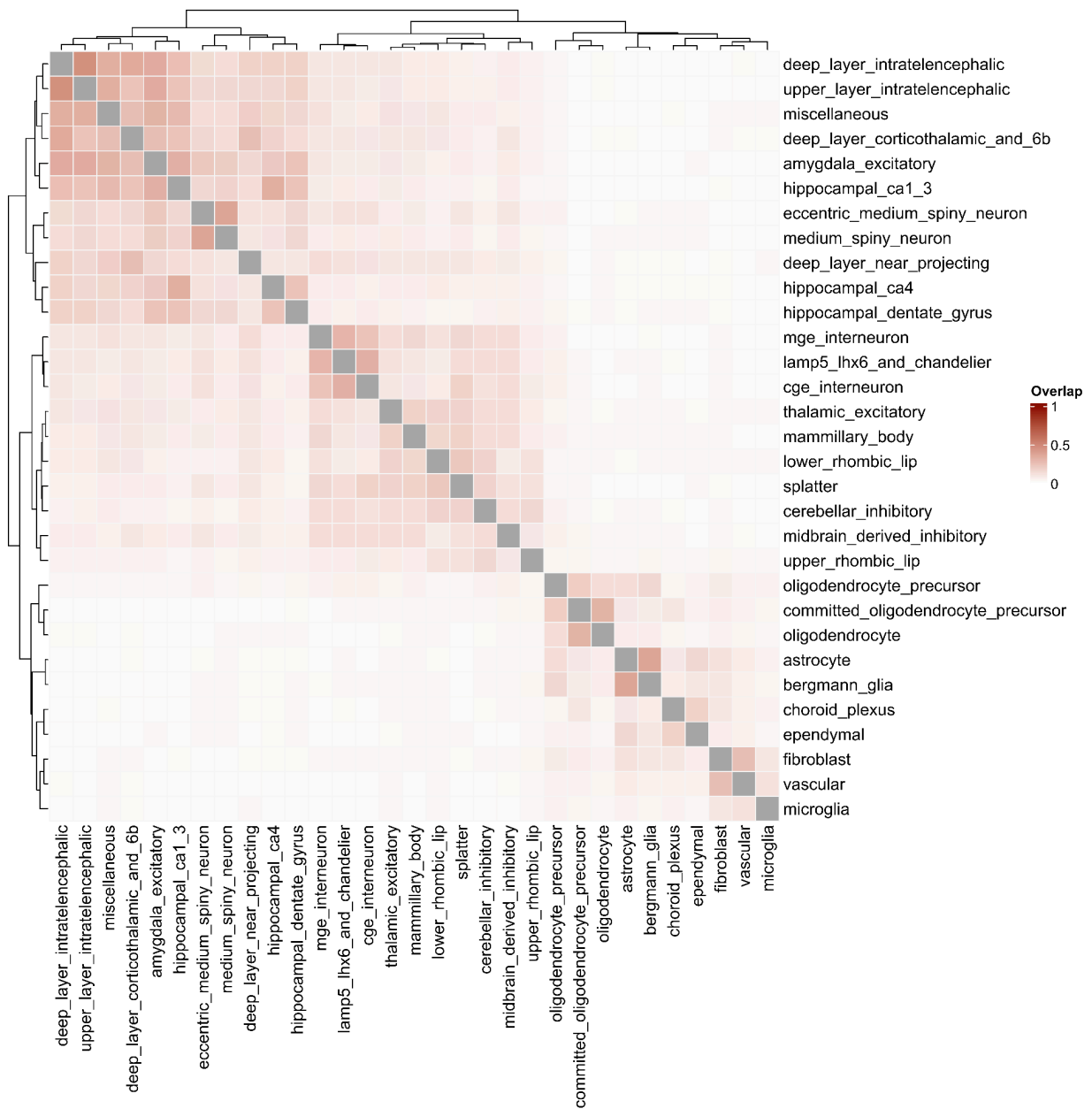
Supplementary Fig. 36. Heatmap of duplication effect sizes for whole-body cells across traits for Male

As in Supplementary Fig. 35, but for duplications (male stratum). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the sex-stratified FunBurd model; Benjamini–Hochberg FDR within the male stratum; $q < 0.05$). Full statistics in Supplementary Data 9.



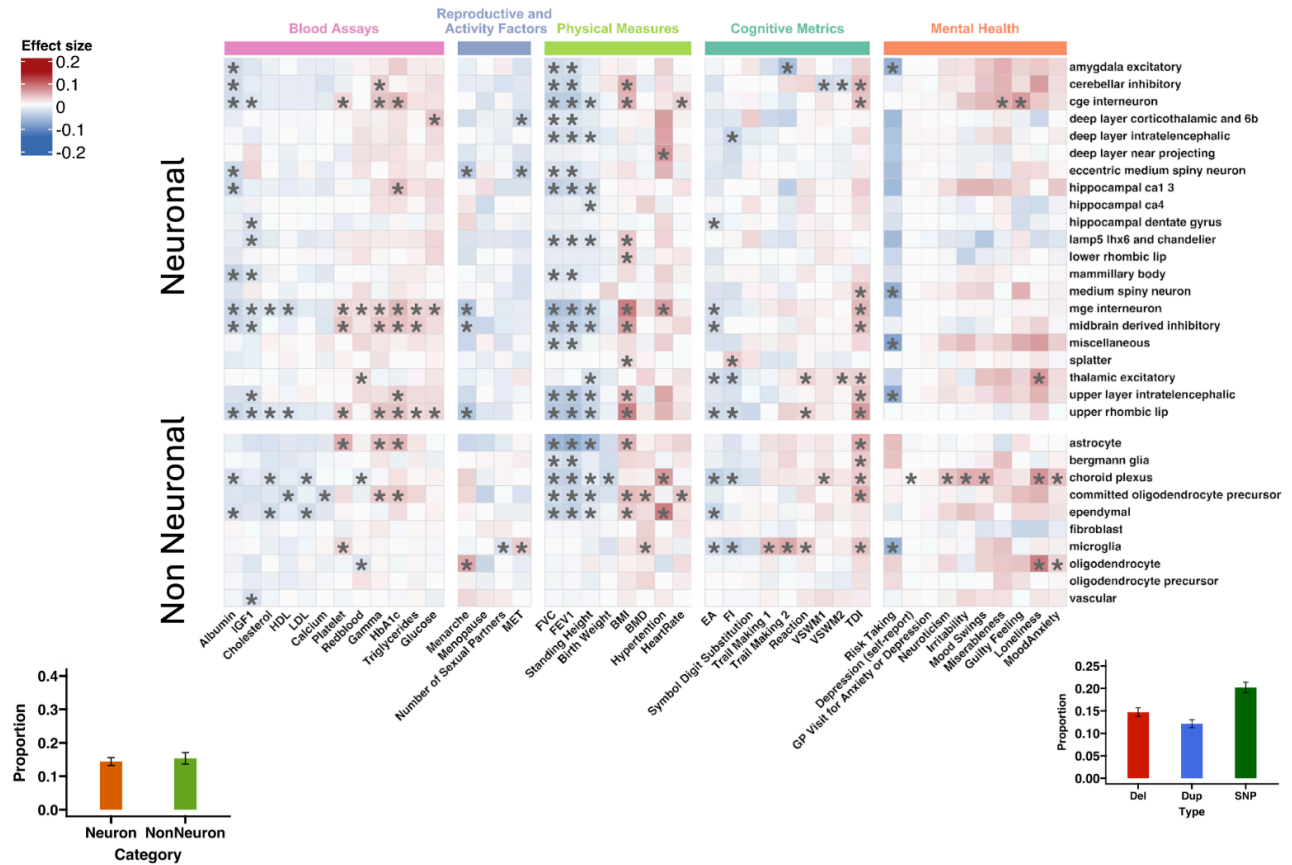
Supplementary Fig. 37. Differences between del-dup associations across trait and gene-set categories for Male

Bar plots summarize the difference in association level between deletions and duplications (proportion of FDR-significant associations) across five trait categories and whole-body cell-type gene-set categories/annotations, for the male stratum. Bar height = aggregate proportion of significant associations; marginal plots show this per trait category (top) and gene-set category (right). Black asterisks (*) denote statistically significant proportion differences (two-sided Pearson's chi-squared test; Benjamini–Hochberg FDR, $q < 0.05$). Source data are provided in Source Data sheets SF37A, SF37B, SF37C.



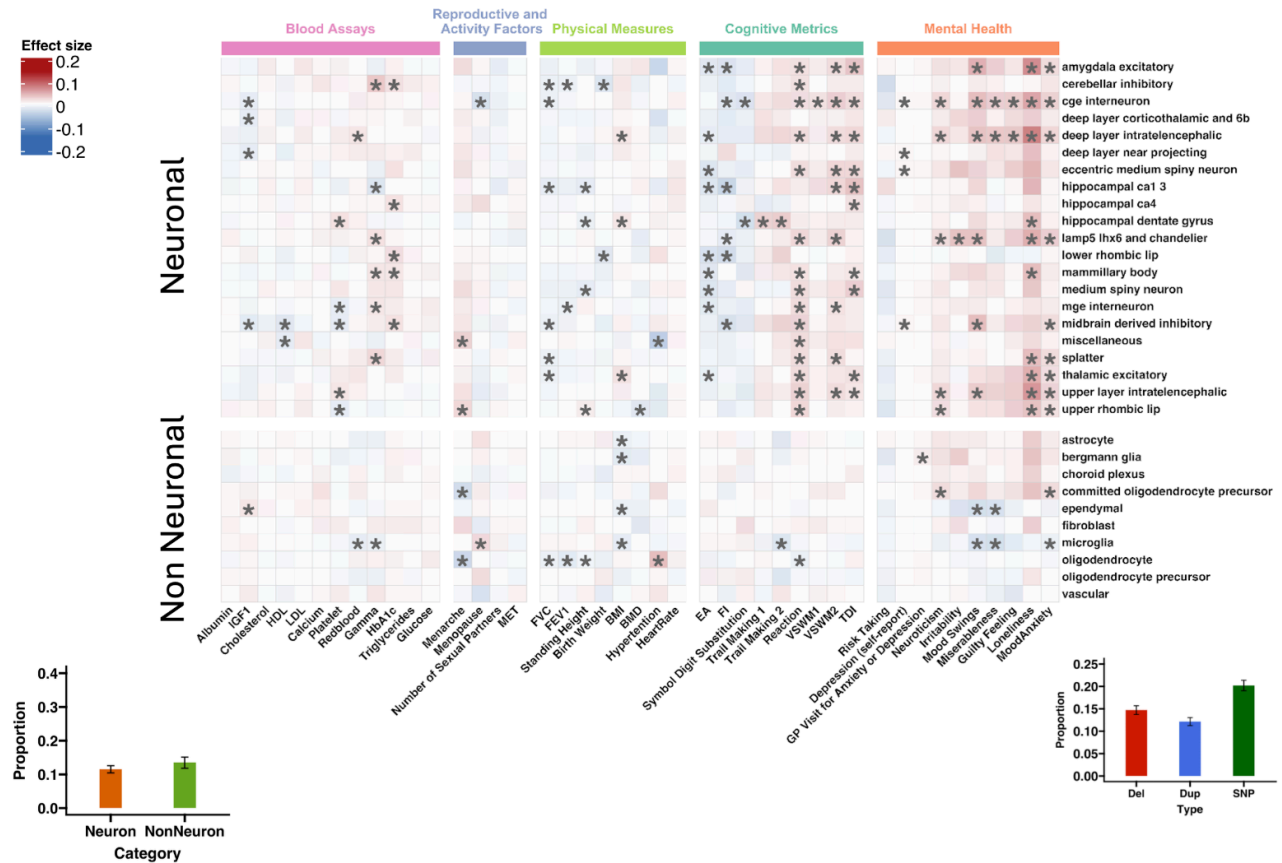
Supplementary Fig. 38. Pairwise overlap matrix of whole brain cell type gene sets

The heatmap displays the pairwise overlap (1 – Jaccard distance) between the 31 whole-brain cell-type gene sets. Color encodes overlap magnitude: dark red = high overlap, white = low (scale 0–1). The maximum, minimum, mean, and median overlap values are 46.7%, 0.98%, 7.5%, and 4.9%, respectively. Descriptive similarity matrix; no statistical test applied. Pairwise values in Supplementary Data 3.



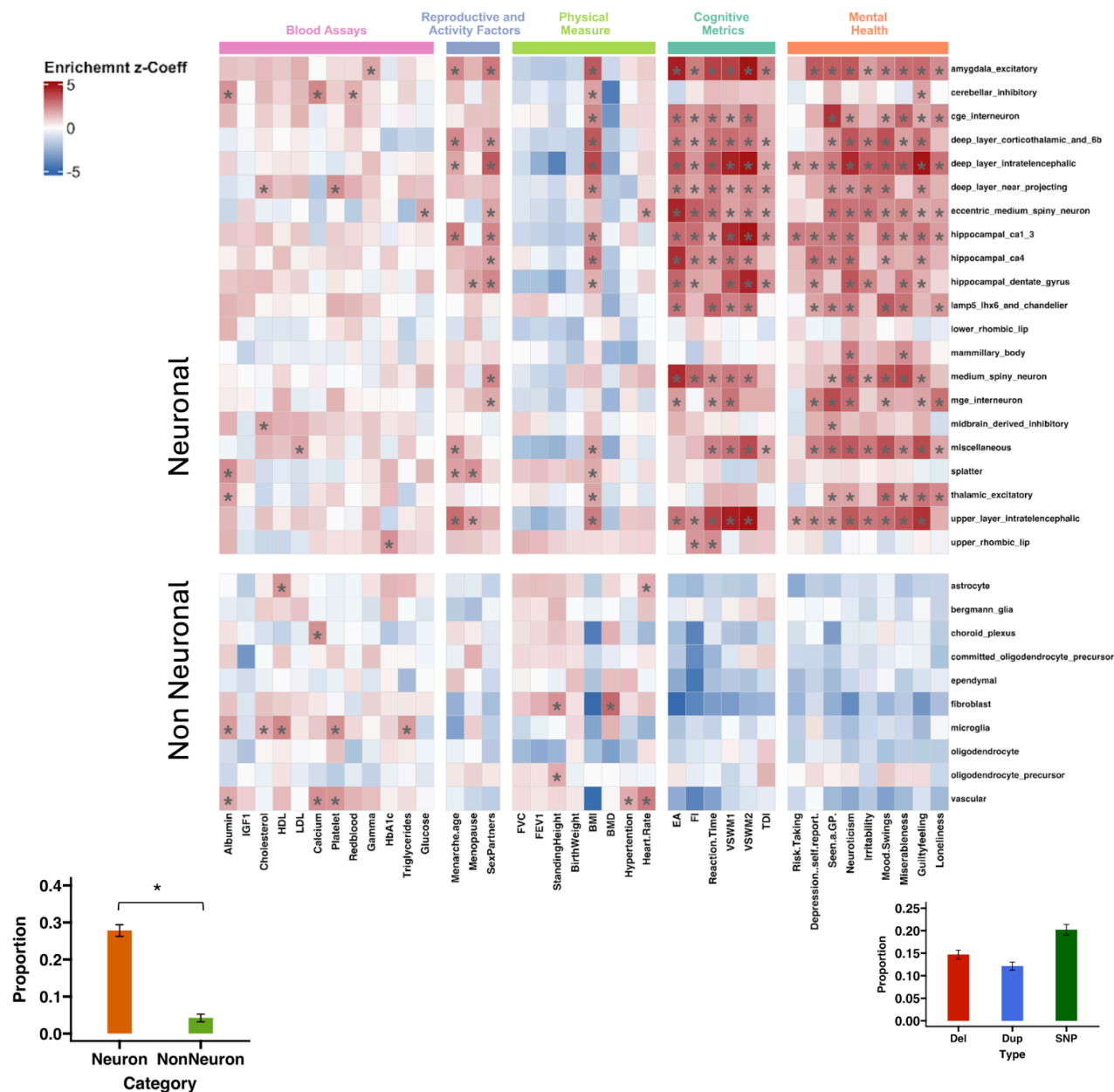
Supplementary Fig. 39. Heatmap of deletion effect sizes for whole-brain cell-type across traits

The heatmap displays deletion burden association effect sizes (β_1) between five trait categories (x-axis) and whole-brain cell-type-specific gene sets (y-axis). Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations, determined using a two-sided t-test on β_1 from the FunBurd multiple linear regression model, with Benjamini–Hochberg FDR correction across 172 gene sets $\times$ 43 traits $\times$ 2 CNV types (14,792 tests; $q < 0.05$). Full statistics in Supplementary Data 4.



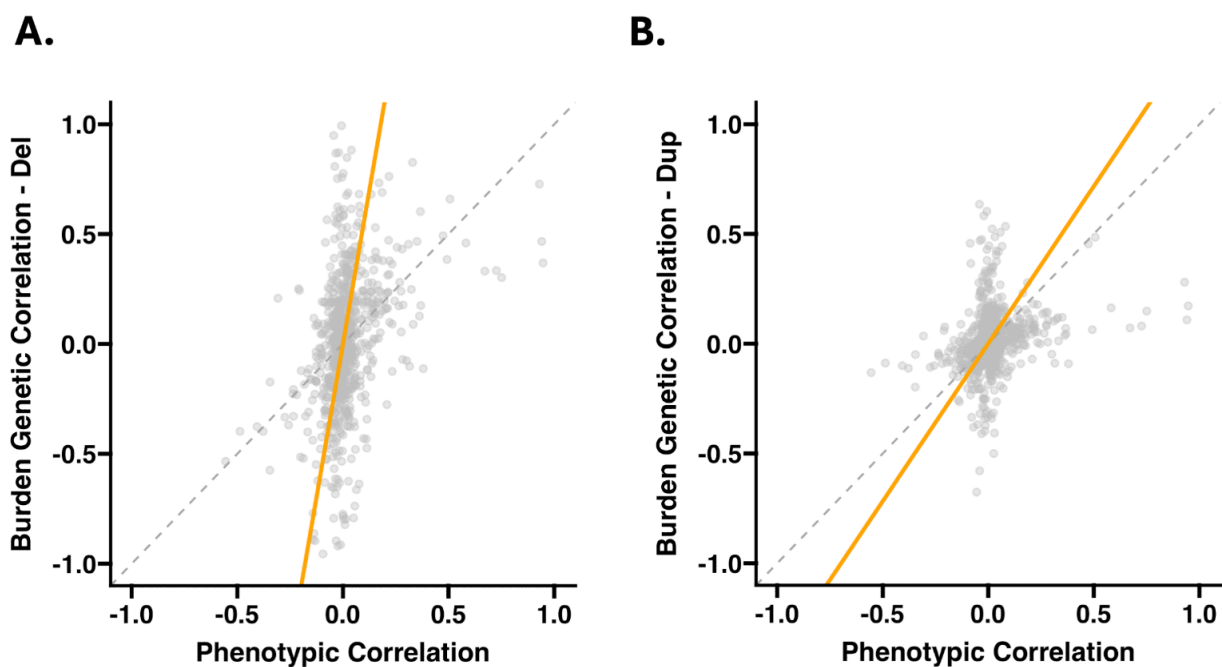
Supplementary Fig. 40. Heatmap of duplication effect sizes for whole-brain cell-type across traits

As in Supplementary Fig. 39, but for duplications. Color encodes the standardized effect size (blue = negative; red = positive; centered at zero). Black asterisks (*) indicate FDR-significant associations (two-sided t-test on β_1 from the FunBurd model; Benjamini–Hochberg FDR across $172 \times 43 \times 2 = 14,792$ tests; $q < 0.05$). Full statistics in Supplementary Data 4.



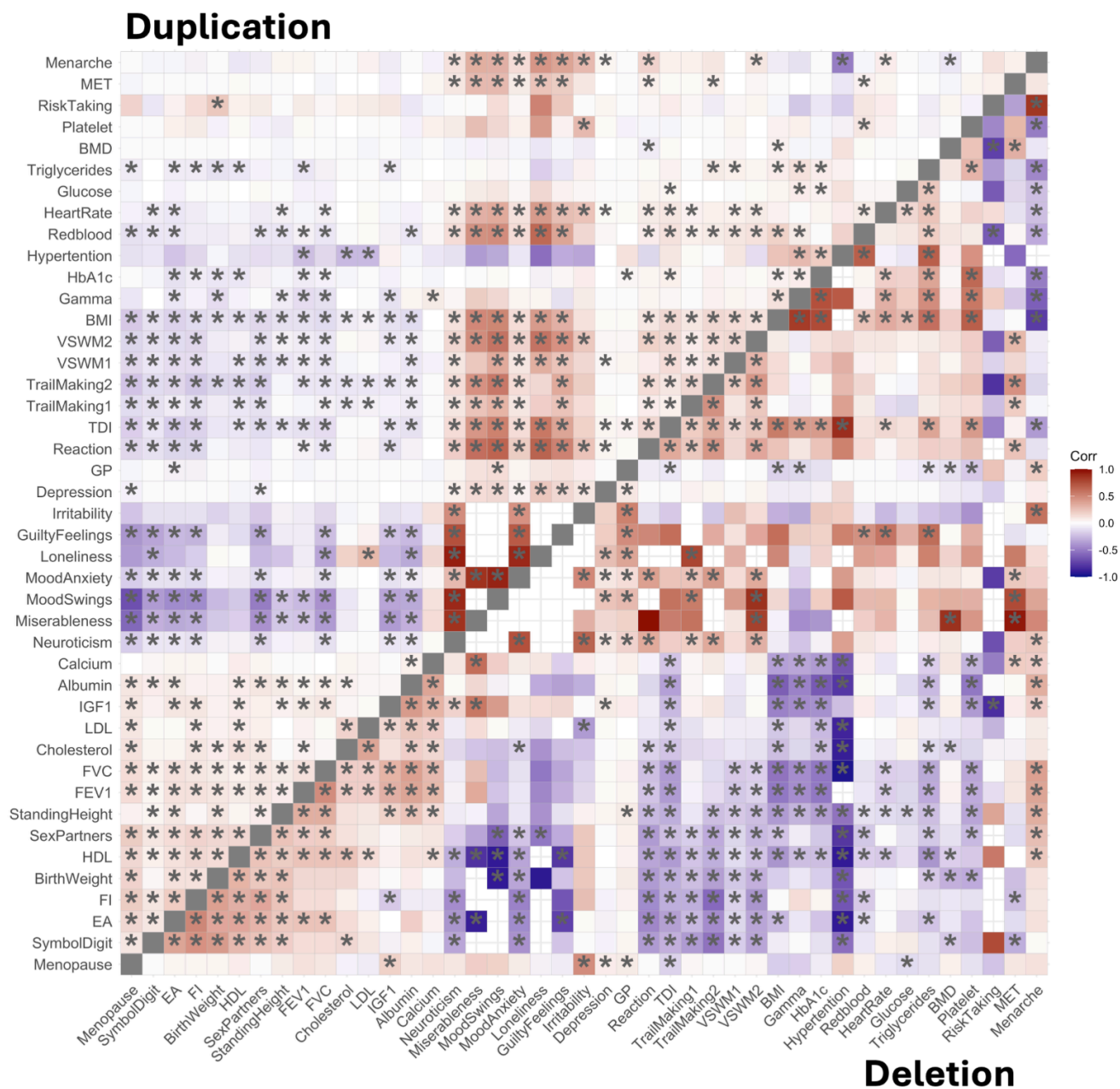
Supplementary Fig. 41. Heatmap of GWAS enrichments for whole-brain cell types across traits using S-LDSC

The heatmap displays stratified LD-score regression (S-LDSC) GWAS enrichments for five trait categories (x-axis) across whole-brain cell-type gene sets (y-axis). Color encodes the S-LDSC enrichment regression coefficient Z-score (blue = negative; red = positive; centered at zero, ≈ -6.2 to $+5.7$). Black asterisks (*) indicate nominal significance of the coefficient Z-score (one-sided test, $p = 1 - \Phi(Z)$; nominal $\alpha = 0.05$, no multiple-comparison adjustment). Statistics in Supplementary Data 10.



Supplementary Fig. 42. Comparison of between-trait CNV burden correlations with phenotypic correlations

Concordance scatter plots compare between-trait (A) deletion and (B) duplication burden correlations (y-axis) with between-trait phenotypic correlations (x-axis). Concordance is summarized using the Mantel statistic r (A: $r=0.33$, $p=0.001$; B: $r=0.23$, $p=0.001$). Full between-trait correlations (burden, SNP, LoF SNV, phenotypic) are in Supplementary Data 12; concordance statistics (Total Least Squares, Mantel, concordance correlation coefficient) are in Supplementary Data 13.

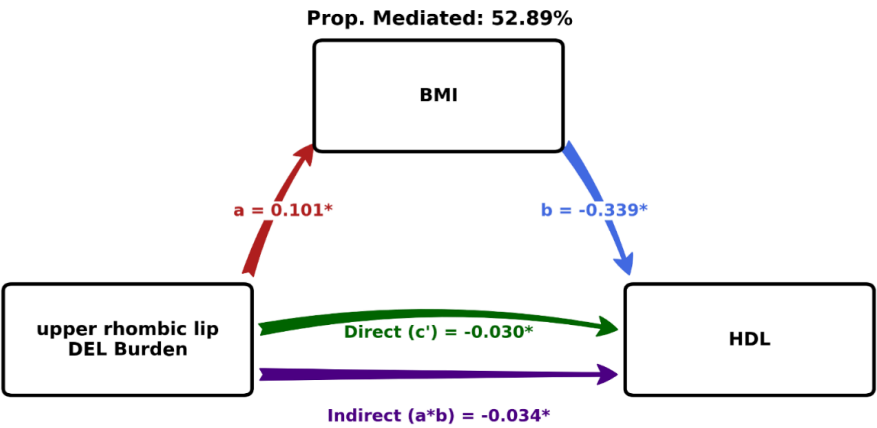


Supplementary Fig. 43. Pairwise CNV burden genetic correlations across 43 traits

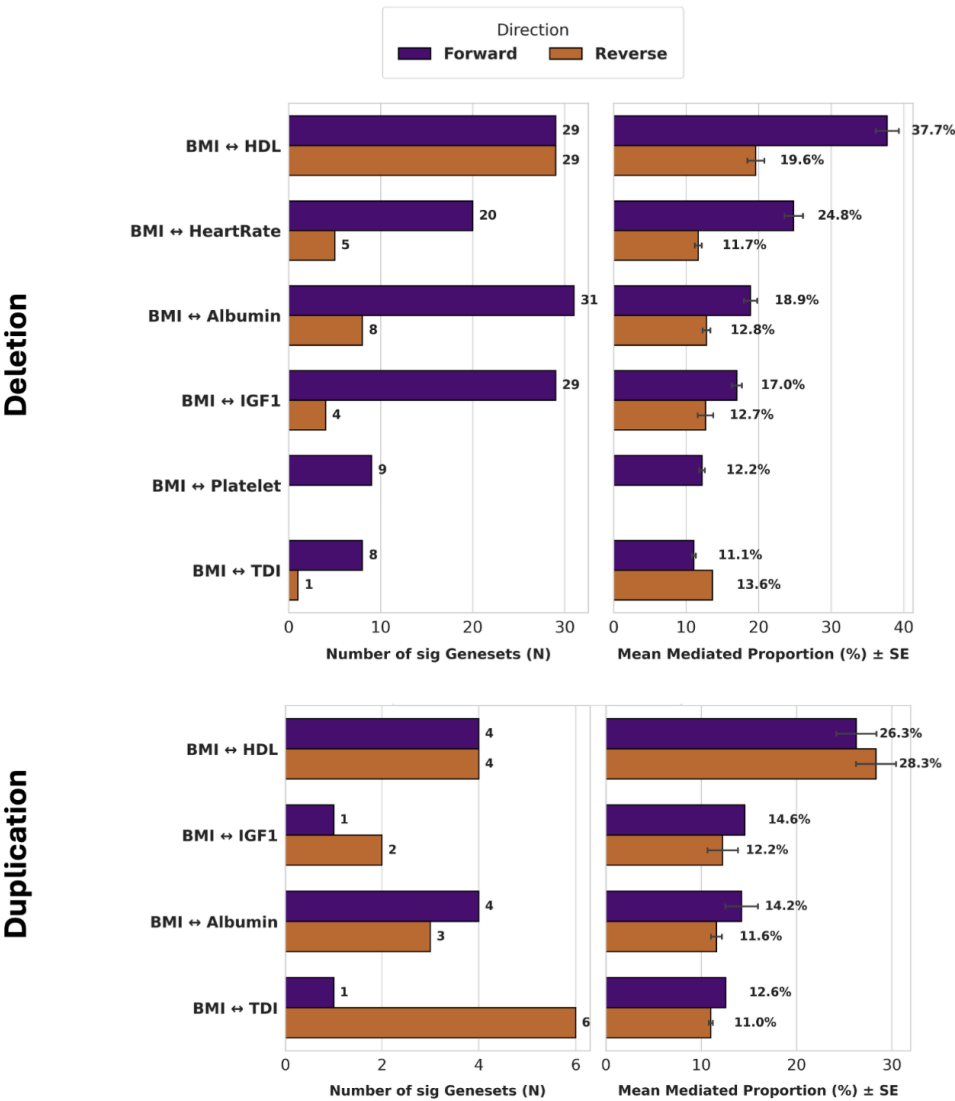
Heatmap of pairwise between-trait CNV burden correlations (Pearson): deletions in the lower triangle, duplications in the upper triangle. Color encodes correlation sign and magnitude: purple = negative, red = positive. Correlations with FDR-corrected significance are marked with an asterisk (*), determined using a two-sided non-parametric permutation test preserving Jaccard distance (P-Jaccard; 1,000 nulls) with Benjamini–Hochberg FDR correction ($q < 0.05$).

White cells indicate burden correlations that exceeded 1 after adjustment for CNV heritability (R^2 , variance explained; Methods). Full correlations and p-values in Supplementary Data 12.

A.



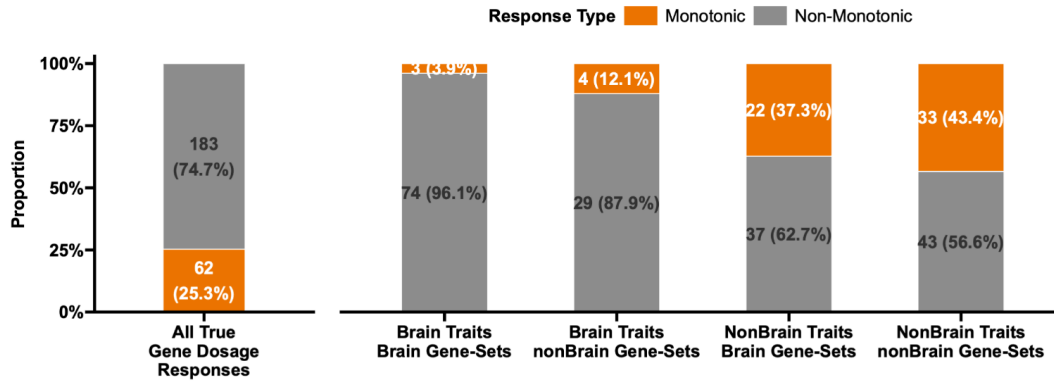
B.



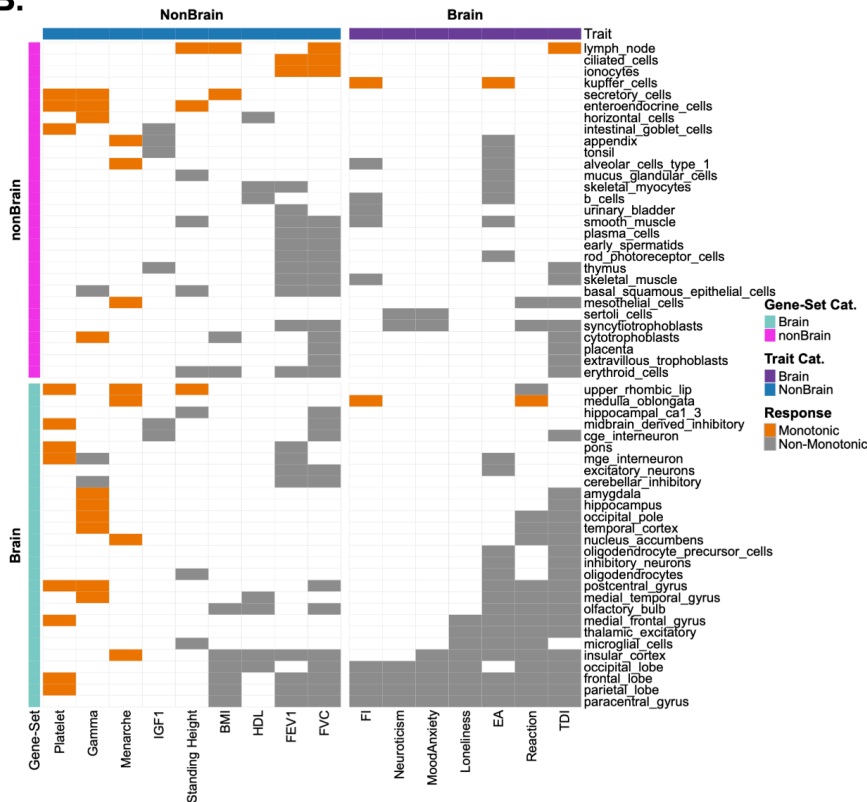
Supplementary Fig. 44. Mediation analysis to distinguish biological and mediated pleiotropy

(A) Representative deletion mediation model: decomposition of the deletion burden effect of the "upper rhombic lip" gene set on HDL cholesterol, with BMI as a mediator. The total effect is partitioned into a direct genetic effect ($c' = -0.030$, $q=2.4e-2$, representing biological pleiotropy) and an indirect effect ($a * b = -0.034$ $q=1.5e-12$, representing mediated pleiotropy). In this specific model, 52.89% of the association is mediated through BMI (FDR $q=5.3e-10$). Path a (0.101, FDR $q=3e-19$) represents the burden effect on the mediator; path b ($b = -0.339$, $q<0.05$) represents the effect of the mediator on the outcome. Asterisks * denote significant path coefficients ($p < 0.05$). Asterisks (*) denote significant path coefficients ($p < 0.05$); the Sobel test assessed the indirect-path significance. (B) Summary of Bidirectional Mediation Across Correlated Trait Pairs. Results of mediation analysis for six trait pairs with CNV burden correlations > 0.3 , analyzed across all 172 gene sets for deletions and duplications. Left panels indicate the number of significant gene sets (N) identifying a mediation path ($a * b$), where both the exposure-to-mediator (path a) and the direct effect (path c') were also statistically significant. Right panels display the mean mediated proportion (%) and standard error (SE) across significant gene sets. Analysis was performed in two directions. forward (BMI $\rightarrow$ Trait, purple) and reverse (BMI $\leftarrow$ Trait, brown). While the direct genetic effect explained the majority of the shared genetic architecture (mean mediation 16%), deletions for BMI–HDL showed a higher mediated proportion (37.7%), suggesting specific metabolic dependencies for deletion-driven associations. Full FDR-corrected path coefficients and p-values in Supplementary Data 15.

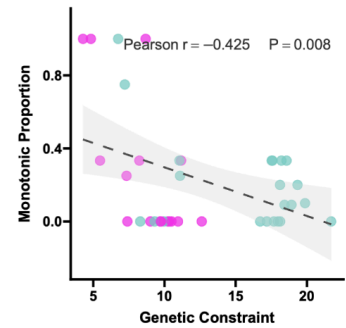
A.



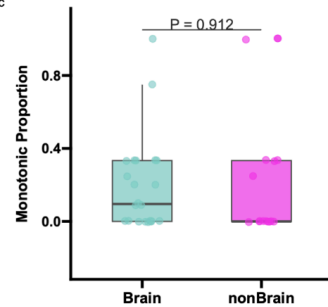
B.



C.



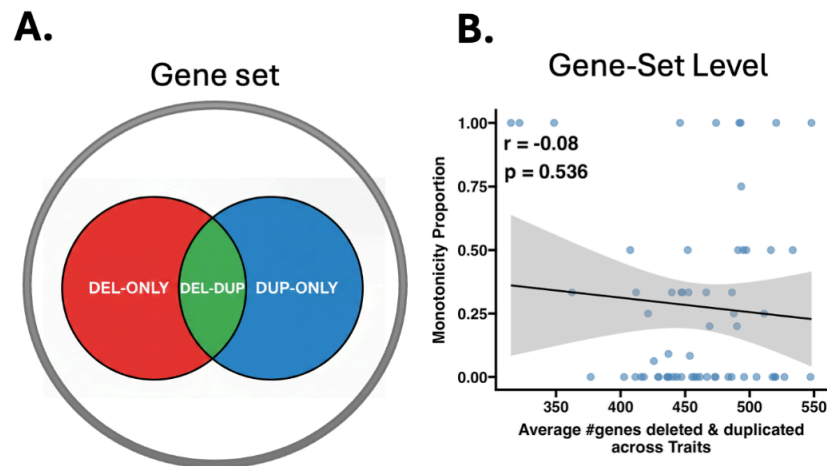
D.



Supplementary Fig. 45. Characteristics of true gene dosage responses across associated traits and gene-sets

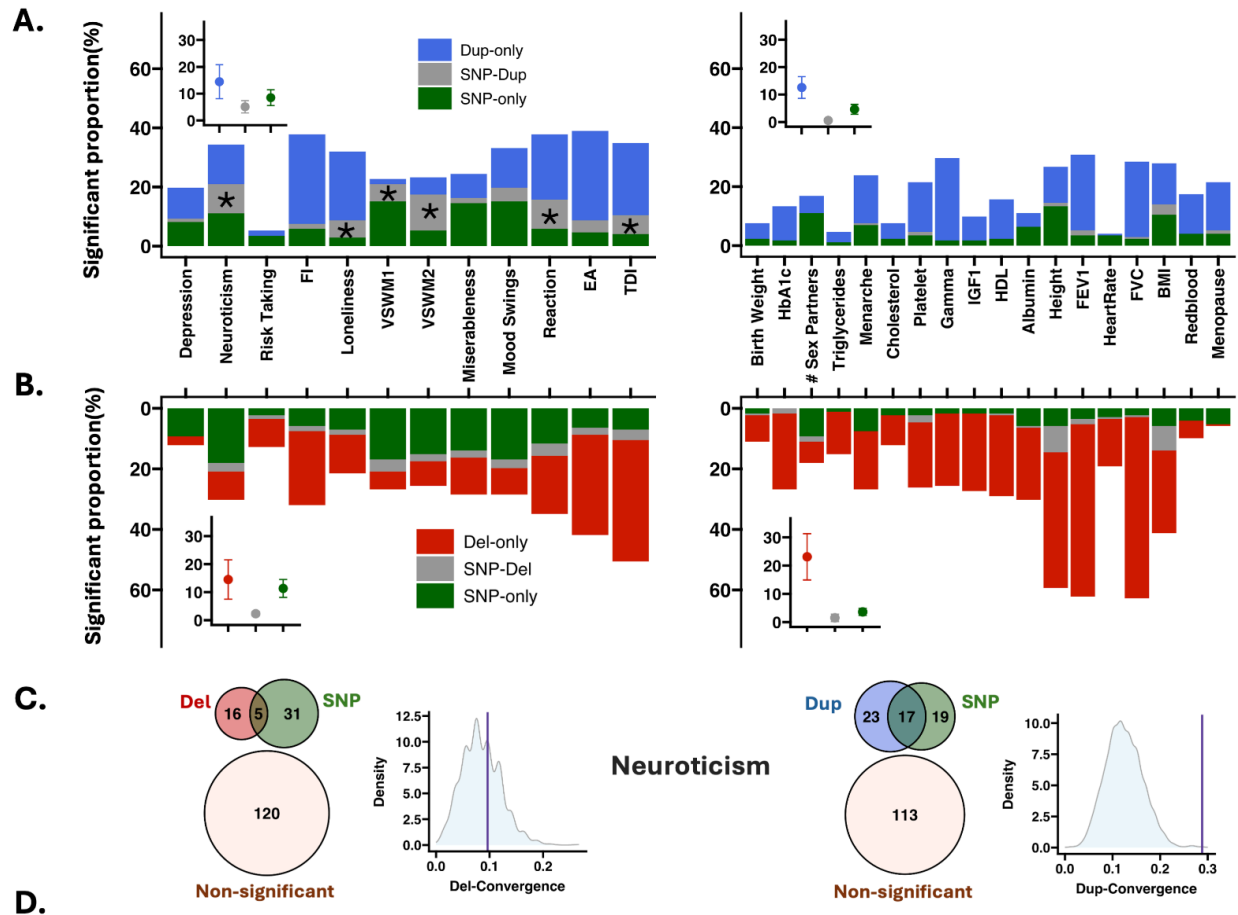
Analysis is restricted to gene-set–trait pairs FDR-significant for both deletions and duplications, further restricted to traits associated with ≥ 5 gene sets and gene sets associated with ≥ 2 traits ($n = 245$ pairs; 16 traits, 88 gene sets). (A) Stacked bar plots of the proportion of monotonic (opposite-direction, orange) versus non-monotonic (same-direction, gray) responses: 62 pairs (25.3%) monotonic, 183 (74.7%) non-monotonic (two-sided exact binomial test against a null proportion of 0.50, $p = 4.8e-15$). Brain traits paired with brain gene sets are 96.1% non-monotonic; non-brain traits show higher monotonic proportions (37.3–43.4%). (B)

Heatmap of the monotonic/non-monotonic distribution across 16 traits and 38 of 88 gene sets (those with > 2 associations). (C) Two-sided Pearson correlation between genetic constraint and the gene-set proportion of monotonic responses ($r = -0.425$, $p = 0.008$; $n = 38$ gene sets). (D) Box plot comparing the proportion of monotonic responses between brain and non-brain gene sets (two-sided Wilcoxon rank-sum test, $p = 0.912$; $n = 22$ brain, 16 non-brain gene sets). Center line = median; box bounds = 25th and 75th percentiles; whiskers = $1.5 \times \text{IQR}$; points beyond whiskers plotted as outliers; each point = one gene set. Per-gene-set and Per-trait counts in Supplementary Data 11 and 14. Source Data in sheets SF45A-SF45D.



Supplementary Fig. 46. Ability to detect monotonicity is not driven by different genes in a gene set affected by deletions and duplications

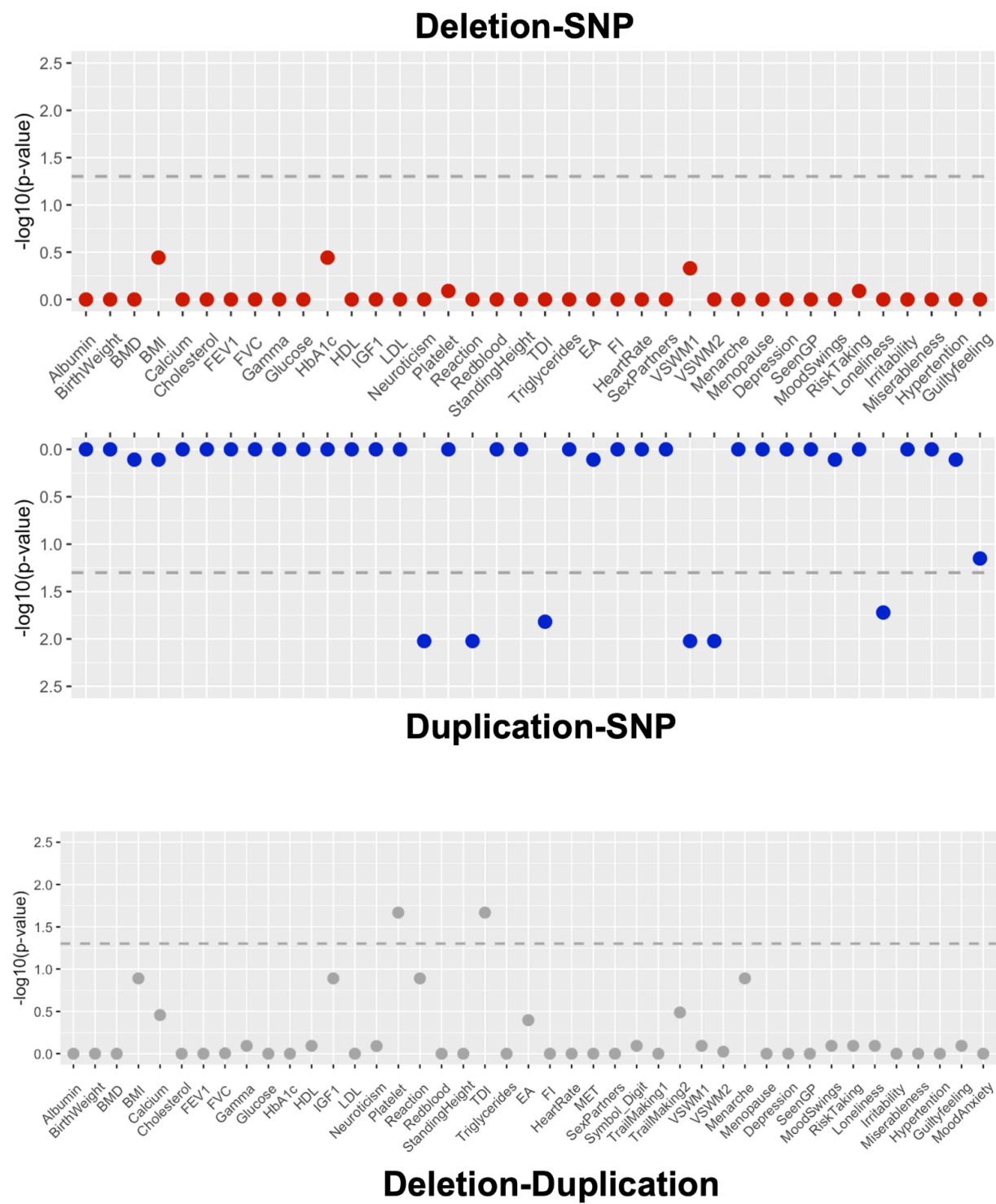
(A) Partitioning of CNV-Impacted Genes Within Gene Sets. Venn diagram illustrating the schematic proportion of genes within a gene set hit exclusively by deletions (DEL-ONLY, red), exclusively by duplications (DUP-ONLY, blue), or by both CNV types (DEL-DUP, green). On average, 35% of genes per gene set are disrupted by both deletions and duplications. (B) Scatter plot showing the number of genes encompassed in both deletions and duplications (x-axis) versus the proportion of monotonic responses (y-axis). Each point is a gene set trait association. No significant correlation was observed (two-sided Pearson correlation, $r = -0.08$, $p = 0.54$; $n = 57$). Monotonicity is computed exclusively among true gene dosage responses. gene-set-trait pairs where both deletion and duplication associations reach FDR significance (57 out of 172 gene sets and 16 out of 43 traits remained). Black lines represent linear regression fits with shaded 95% confidence intervals. Source data are provided in Source Data sheet SF46B.



Supplementary Fig. 47. Functional overlap between common variants and CNVs

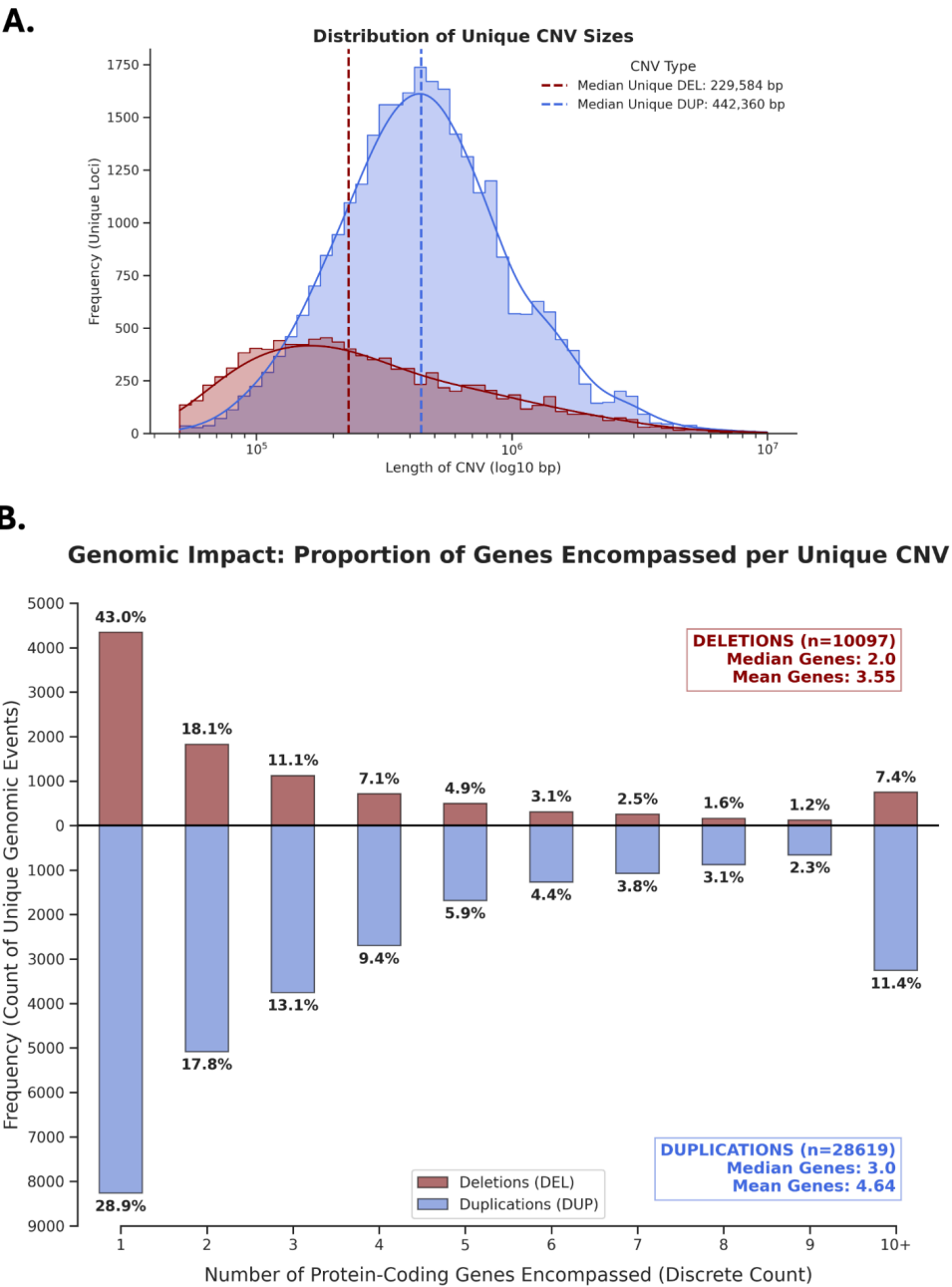
A) The stacked bar plot represents the proportion of the functional gene sets overlapping between common variants and duplications for each trait. Overlapping functional gene sets (enriched in common variant and CNV signal) are gray, and non-overlapping functional gene sets are represented in blue (associations only significant for duplications or deletions). Stars denote significant functional overlap, determined using a two-sided non-parametric permutation test (1,000 nulls) with Benjamini–Hochberg FDR correction ($q < 0.05$). Y-axis: proportion of associated gene sets; X-axis: Brain and non-brain traits. The inserts represent the average convergence (between common variants and duplications) and average duplication-specific associations across traits. The same information on functional convergence is shown in (B) between deletions and common variants (SNPs). C) illustrates the example of the functional overlap between SNPs and CNVs for Neuroticism using a Venn diagram. The density plot represents the null distribution of overlap significance (1000

permutations), with the observed overlap shown by the purple vertical line. Source data are provided in Source Data sheets SF47A-SF47B.



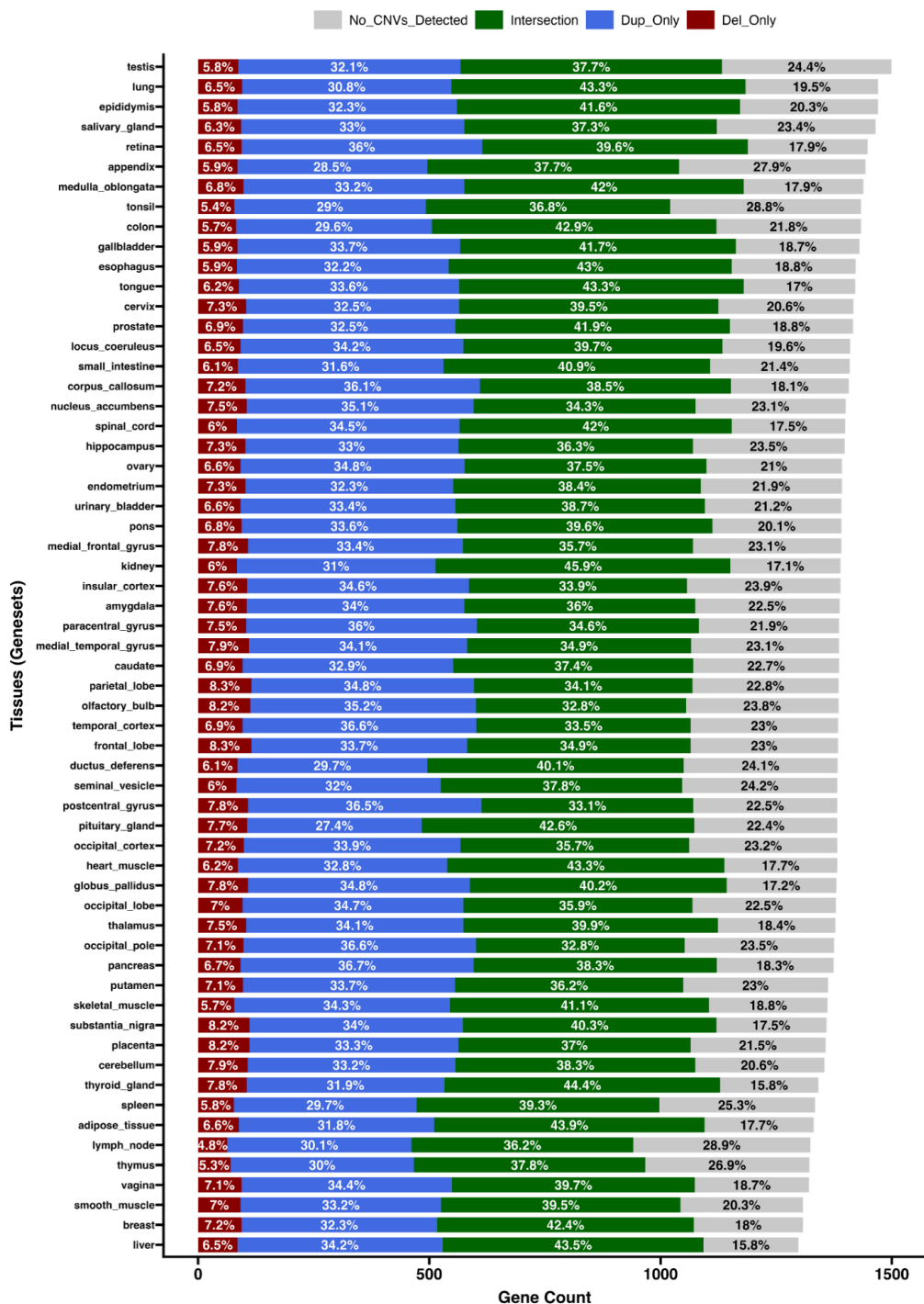
Supplementary Fig. 48. Permutation test P-Values (1000 nulls) for functional overlap between CNV associations and GWAS enrichments, and between deletion and duplication

Matrix of p-values from the two-sided non-parametric permutation test (1,000 nulls) assessing functional gene-set overlap between (i) CNV associations and GWAS S-LDSC enrichments and (ii) deletions and duplications, per trait. Companion to Supplementary Fig. 47. Source data are provided in Source Data sheets SF48.



Supplementary Fig. 49. Genomic characteristics and gene content of unique deletions and duplications

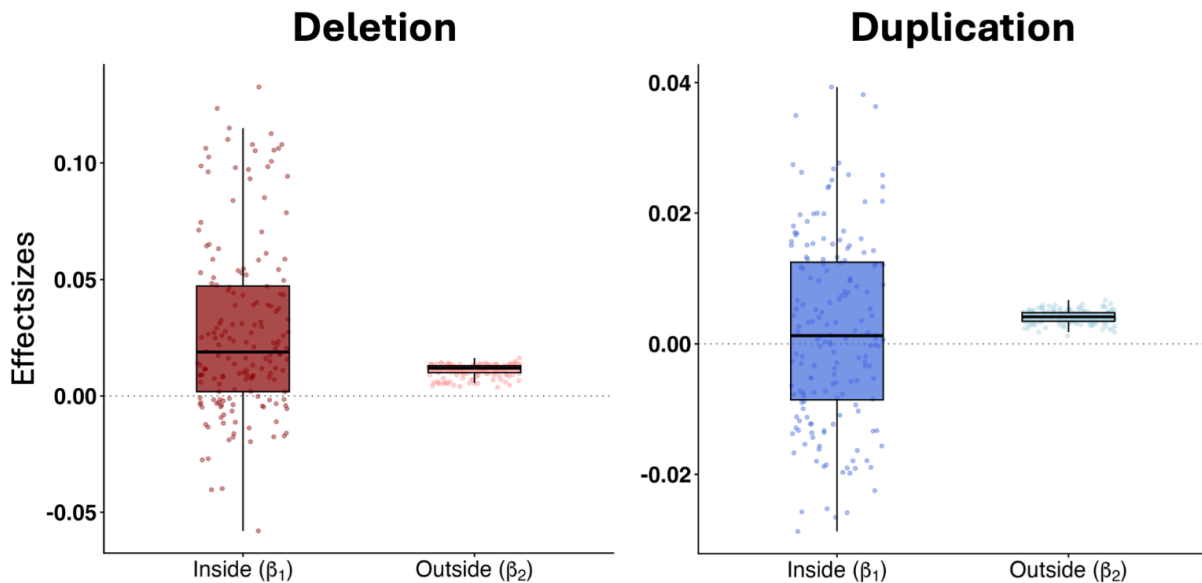
(A) Distribution of Unique CNV Sizes. Histogram and density plot showing the length distribution (log10 base pairs) for unique deletions (DEL, red) and duplications (DUP, blue). Vertical dashed lines indicate the median size for each CNV type. 229,584 bp for deletions and 442,360 bp for duplications. (B) Proportion of Genes Encompassed per Unique CNV. Mirror bar chart displaying the frequency and percentage of unique genomic events categorized by the number of genes (protein-coding) they encompass (ranging from 1 to 10+). Deletions (n = 10,097) show a median of 2.0 genes and a mean of 3.55 genes per event. Duplications (n = 28,619) show a higher gene density with a median of 3.0 genes and a mean of 4.64 genes per event. Descriptive distributions; no statistical test applied. Summary statistics (mean, median, IQR, range) in Supplementary Table 6.



Supplementary Fig. 50. Proportion of genes in tissue-specific gene sets hit by observed CNVs (for BMI)

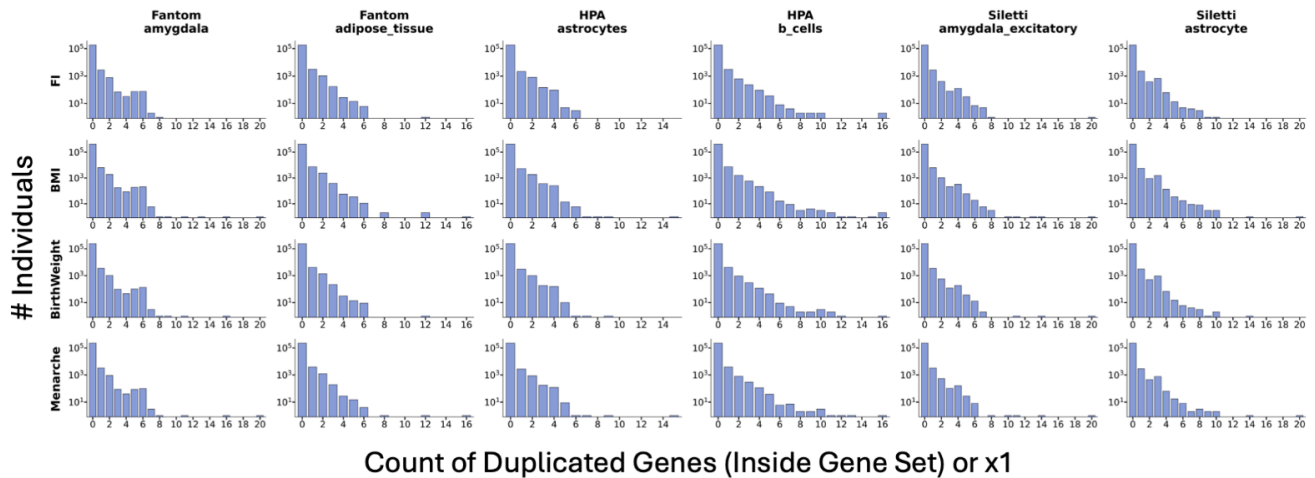
Horizontal stacked bar chart depicting the intersection between 60 FANTOM5 tissue-specific gene sets and the CNVs observed in our dataset. Each bar represents the total gene count for a defined gene set, partitioned by the specific type of CNV hit. Genes hit exclusively by observed

deletions (Del_Only, red), genes hit exclusively by observed duplications (Dup_Only, blue), and genes hit by both CNV types (Intersection, green). The grey segment (No_CNVs_Detected) accounts for genes within each defined set that were not hit by any CNVs in our dataset. Internal labels indicate the percentage of the total gene count within each defined gene set that is hit by observed CNVs in each respective category. Descriptive (compositional) plot; no statistical test applied. Source data are provided in Source Data sheet SF50.



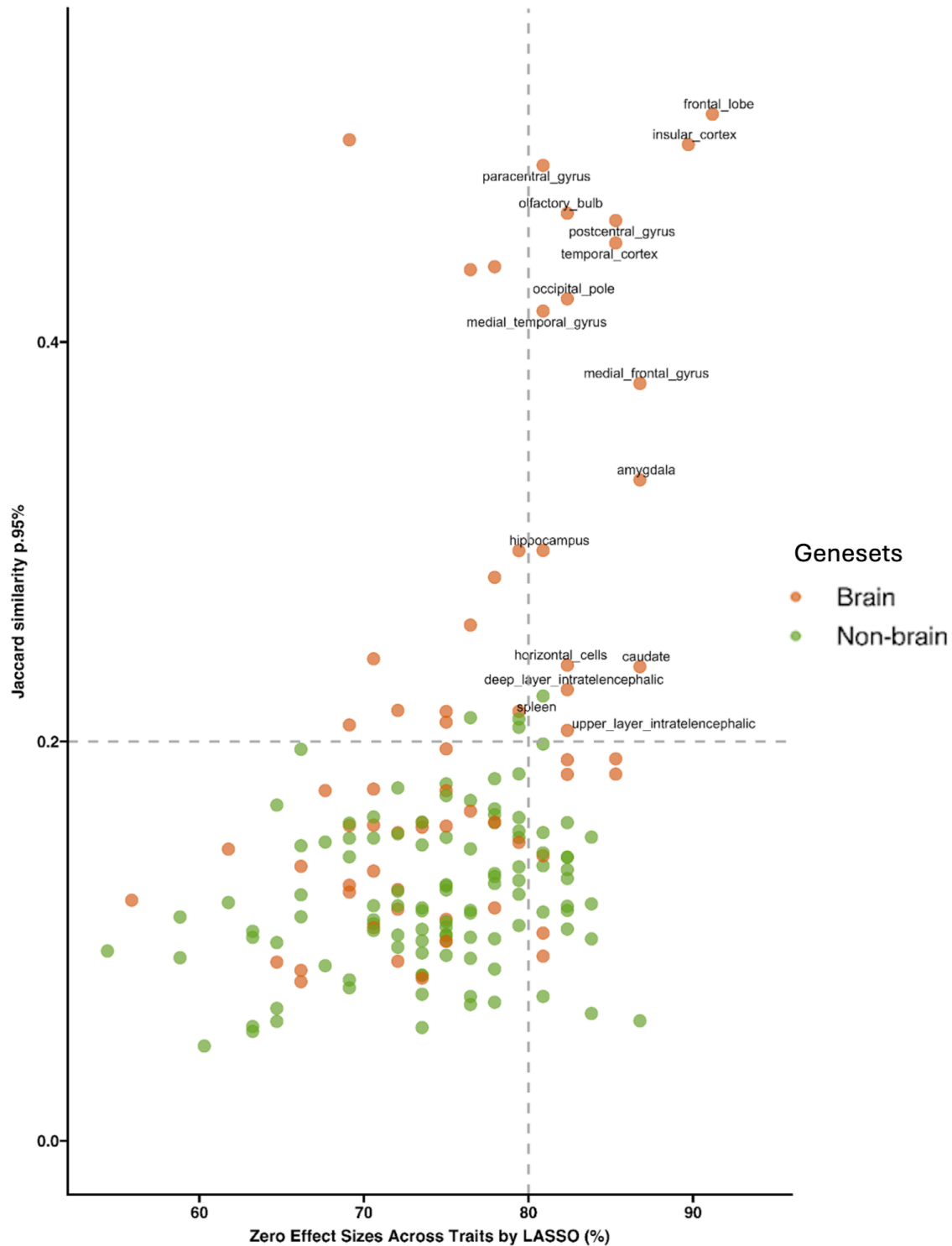
Supplementary Fig. 51. Comparison of burden effect sizes for gene set membership inside (β_1) and outside (β_2) for BMI

Box plots comparing the effect size estimates for genes inside the target gene set (β_1) versus genes outside the set (β_2) across all 172 gene sets, shown separately for deletions (left) and duplications (right). Center line = median; box bounds = 25th and 75th percentiles; whiskers = $1.5 \times$ interquartile range; data points shown as a jittered overlay (each point = one gene set). Inside effects show greater variance than outside-set effects for both CNV types, consistent with gene-set-specific signal enrichment. The center line denotes the median, the box represents the interquartile range (IQR), and data points are shown as a jittered overlay to illustrate the full distribution of the gene sets' effect sizes. Full per-trait β_1 and β_2 statistics in Supplementary Data 4.

A.**Deletion****B.****Duplication**

Supplementary Fig. 52. Distribution of the number of genes hit by CNVs per individual across representative gene sets and traits

(A) Deletion Count Distributions. Grid of histograms showing the frequency of individuals (y-axis) categorized by the discrete count of deleted genes (x-axis; x1) within six representative gene sets across four traits. (B) Duplication Count Distributions. The observed distributions follow a Poisson-like decay, where the majority of individuals carry zero hits within a specific gene set, and a small fraction of the population carries multiple disrupted genes. Descriptive distributions; no statistical test applied. Source data are provided in Source Data sheet SF52.



Supplementary Fig. 53. Scatterplot of geneset Jaccard similarity (95th percentile) vs geneset zero effect sizes across traits excluded by LASSO (%)

Scatter plot of gene-set Jaccard similarity (95th percentile across pairs; x-axis) versus the proportion of near-zero effect sizes across traits (y-axis). Gene sets with Jaccard similarity > 0.2

and > 80% zero effect sizes were excluded from the genetic-correlation analysis. Descriptive selection plot; no statistical test applied.



Supplementary Fig. 54. Performance of FunBurd running in parallel for constraint nulls genesets and traits (43,000 Jobs)

Computational performance of the parallelized FunBurd normative-null pipeline, executed across 1,000 constraint-based null gene sets $\times$ 43 traits, with multithreaded evaluation of 20 constraint levels (5%–100%) per run, in total 43 traits $\times$ 1,000 nulls $\times$ 20 levels $\times$ 2 CNV types =

1,720,000 FunBurd associations. Descriptive; no statistical test applied. Statistics of the performance of hybrid jobs derived from Compute Canada.

Supplementary Tables

Supplementary Table 1

Statistics for comparing proportion of significant associations/enrichments between deletions, duplications, and SNPs.

Geneset_Type	Type1	Type2	p_value	statistic	estimate_prop_Type1	estimate_prop_Type2	p_value_fdr
Fantom	SNP	Del	2.90e-42	185.603	0.0785	0.2202	1.30e-41
Fantom	SNP	Dup	1.42e-10	41.1367	0.0785	0.1364	2.56e-10
Fantom	Del	Dup	5.30e-15	61.1467	0.2202	0.1364	1.19e-14
HPA	SNP	Del	6.18e-45	197.84	0.04	0.1418	5.56e-44
HPA	SNP	Dup	3.22e-32	139.6194	0.04	0.1209	9.67e-32
HPA	Del	Dup	0.0107	6.5223	0.1418	0.1209	0.012
Siletti	SNP	Del	3.37e-04	12.8502	0.202	0.147	4.34e-04
Siletti	SNP	Dup	5.13e-08	29.6669	0.202	0.1215	7.70e-08
Siletti	Del	Dup	0.0609	3.5137	0.147	0.1215	0.0609

Column descriptions

Column	Description
Geneset_Type	Geneset type, with either of the three values: i) Tissue_Fantom60 (60 gene sets derived from whole body tissue RNAseq data from Fantom); ii) SC_HPA81 (81 gene sets derived from the whole body single cell RNAseq data from Human Protein Atlas, HPA); and iii) SC_Siletti31 (31 gene sets derived from Siletti et al. Science 2023, from whole brain single cell RNA seq data)
Type1	First variant type being compared (e.g., SNP, Del, Dup)
Type2	Second variant type being compared (e.g., SNP, Del, Dup)
p_value	Raw p-value from the proportion test comparing proportion of significant associations between Type1 and Type2
statistic	Test statistic (chi-squared) from the proportion test
estimate_prop_Type1	Proportion of significant associations for variant Type1
estimate_prop_Type2	Proportion of significant associations for variant Type2
p_value_fdr	FDR-adjusted p-value from the proportion test

Supplementary Table 2

Statistics for comparing proportion of significant associations/enrichments between brain and non-brain gene sets, across variant types (deletions, duplications, SNPs) and gene-set categories.

Type	Geneset_Type	geneset_cat1	geneset_cat2	p_value	statistic	estimate_prop_cat1	estimate_prop_cat2	p_value_fdr
Del	Fantom	Brain	NonBrain	4.89e-64	285.4459	0.373	0.0951	4.40e-63
Del	HPA	Brain	NonBrain	0.0037	8.4412	0.1993	0.1364	0.0055
Del	Siletti	Neuron	NonNeuronal	0.7067	0.1416	0.144	0.1535	0.7067
Dup	Fantom	Brain	NonBrain	9.94e-11	41.8324	0.1852	0.0965	2.24e-10
Dup	HPA	Brain	NonBrain	0.2578	1.2806	0.1429	0.1188	0.3314
Dup	Siletti	Neuron	NonNeuronal	0.3472	0.8836	0.1152	0.1349	0.3906
SNP	Fantom	Brain	NonBrain	8.91e-32	137.6017	0.152	0.0183	4.01e-31
SNP	HPA	Brain	NonBrain	3.29e-08	30.5278	0.1053	0.0338	5.92e-08
SNP	Siletti	Neuron	NonNeuronal	8.26e-21	87.5402	0.2782	0.0421	2.48e-20

Column descriptions

Column	Description
Type	Variant type (Del, Dup, or SNP) for which the brain vs non-brain comparison was performed (Neuron vs Non-Neuron for Siletti single cell data)
Geneset_Type	Geneset type, with either of the three values: i) Tissue_Fantom60 (60 gene sets derived from whole body tissue RNAseq data from Fantom); ii) SC_HPA81 (81 gene sets derived from the whole body single cell RNAseq data from Human Protein Atlas, HPA); and iii) SC_Siletti31 (31 gene sets derived from Siletti et al. Science 2023, from whole brain single cell RNA seq data)
geneset_cat1	Gene set category 1, Brain or Neuronal (for Siletti) gene sets
geneset_cat2	Gene set category 2, Non-Brain or Non-Neuronal (for Siletti) gene sets
p_value	Raw p-value from the proportion test comparing brain vs non-brain gene-set proportions
statistic	Test statistic (chi-squared) from the proportion test
estimate_prop_cat1	Proportion of significant associations in brain-related tissues/cell types (Neuron cell types for Siletti)
estimate_prop_cat2	Proportion of significant associations in non-brain-related tissues/cell types (Non-Neuron cell types for Siletti)
p_value_fdr	FDR-adjusted p-value from the proportion test

Supplementary Table 3

Sex-difference statistics for Mental Health traits: two-proportion tests comparing Female vs Male significance proportions for deletions and duplications.

CNV_Type	Trait	Sig_F	Total_F	Prop_F	Sig_M	Total_M	Prop_M	Raw_P_Value	FDR_P_Value	Significant_Sex_Diff
DEL	MoodSwings	18	60	0.3	2	60	0.0333	8.88e-05	0.0018	Yes
DEL	MoodAnxiety	20	60	0.3333	7	60	0.1167	0.0045	0.0448	Yes
DEL	Loneliness	17	60	0.2833	7	60	0.1167	0.0225	0.1224	No
DEL	Neuroticism	15	60	0.25	6	60	0.1	0.0306	0.1224	No
DEL	Irritability	0	60	0	5	60	0.0833	0.0573	0.1911	No
DEL	GuiltyFeelings	6	60	0.1	1	60	0.0167	0.114	0.3256	No
DEL	RiskTaking	5	60	0.0833	2	60	0.0333	0.439	0.878	No
DEL	Depression	0	60	0	1	60	0.0167	1	1	No
DEL	GP	0	60	0	0	60	0	1	1	No
DEL	Miserableness	11	60	0.1833	9	60	0.15	0.6242	1	No
DUP	Miserableness	0	60	0	6	60	0.1	0.0274	0.1224	No
DUP	Neuroticism	9	60	0.15	4	60	0.0667	0.1419	0.3549	No
DUP	Loneliness	11	60	0.1833	7	60	0.1167	0.3065	0.6811	No
DUP	Depression	1	60	0.0167	1	60	0.0167	1	1	No
DUP	GP	3	60	0.05	1	60	0.0167	0.6186	1	No
DUP	GuiltyFeelings	1	60	0.0167	0	60	0	1	1	No
DUP	Irritability	1	60	0.0167	0	60	0	1	1	No
DUP	MoodAnxiety	7	60	0.1167	8	60	0.1333	0.7825	1	No
DUP	MoodSwings	1	60	0.0167	2	60	0.0333	1	1	No
DUP	RiskTaking	0	60	0	0	60	0	1	1	No

Column descriptions

Column	Description
CNV_Type	CNV type (deletion or duplication)
Trait	Mental health trait / phenotype name
Sig_F	Number of FDR-significant gene-set associations in females
Total_F	Total number of gene-set associations tested in females
Prop_F	Proportion of gene-set associations FDR-significant in females (Sig_F / Total_F)
Sig_M	Number of FDR-significant gene-set associations in males
Total_M	Total number of gene-set associations tested in males
Prop_M	Proportion of gene-set associations FDR-significant in males (Sig_M / Total_M)
Raw_P_Value	Raw p-value from the two-proportion test comparing female vs male significance proportions
FDR_P_Value	FDR-adjusted p-value from the two-proportion test
Significant_Sex_Diff	Indicator (Yes/No) of whether the female vs male proportion difference is FDR-significant

Supplementary Table 4

Statistics for comparing proportion of significant associations for a given gene set category (Brain vs NonBrain) across trait categories.

Test_Name	Chi_Squared_Statistic	P_Value	Prop_1_Estimate	Prop_2_Estimate	FDR_pval	DF
Fantom_DUP_BrainTissue	87.5437	8.24e-21	0.306	0.0895	9.89e-20	1
Fantom_DUP_NBrainTissue	0.5333	0.4652	0.0893	0.1023	0.5582	1
Fantom_DEL_BrainTissue	0.1192	0.7299	0.3665	0.3781	0.7962	1
Fantom_DEL_NBrainTissue	28.2077	1.09e-07	0.0478	0.1326	3.27e-07	1
HPA_DUP_BrainCell	1.3477	0.2457	0.1729	0.119	0.3276	1
HPA_DUP_NBrainCell	1.895	0.1686	0.128	0.1115	0.253	1
HPA_DEL_BrainCell	6.8192	0.009	0.2707	0.1429	0.018	1
HPA_DEL_NBrainCell	29.5548	5.44e-08	0.0989	0.1661	2.17e-07	1
Siletti_DUP_NeuronCell	34.4652	4.34e-09	0.2331	0.0893	2.60e-08	1
Siletti_DUP_NonNeuronCell	0.002	0.9647	0.0526	0.0583	0.9647	1
Siletti_DEL_NeuronCell	26.448	2.71e-07	0.0752	0.1984	6.50e-07	1
Siletti_DEL_NonNeuronCell	2.3273	0.1271	0.1211	0.1792	0.2179	1

Column descriptions

Column	Description
Test_Name	Label identifying the comparison (e.g., GenesetType_CNVtype_GenesetCategory) for proportion test across trait categories
Chi_Squared_Statistic	Test statistic (chi-squared) from the proportion test across trait categories
P_Value	Raw p-value from the proportion test across trait categories
Prop_1_Estimate	Proportion of significant associations in trait category 1 for the given gene set/variant combination
Prop_2_Estimate	Proportion of significant associations in trait category 2 for the given gene set/variant combination
FDR_pval	FDR-adjusted p-value from the proportion test across trait categories
DF	Degrees of freedom of the chi-squared test

Supplementary Table 5

Statistics for correlations between genetic constraint measures and functional burden pleiotropy (deletions, duplications, GWAS S-LDSC enrichments), with Jaccard-permutation-based p-values .

stat	constraint	cor	pvalue_Jaccard	adj_pvalue_Jaccard
functional_burden_pleiotropy_deletion	measure_LEOUF_topdecile	0.3899	0.012	0.0229
functional_burden_pleiotropy_deletion	constraint_measure_missense_Z_topdecile	0.502	0.001	0.0026
functional_burden_pleiotropy_deletion	constraint_measure_s_het_topdecile	0.2637	0.112	0.147
functional_burden_pleiotropy_deletion	constraint_measure_CDS_topdecile	0.4633	0.001	0.0026
functional_burden_pleiotropy_deletion	constraint_measure_pHaplo_topdecile	0.11	0.321	0.321
functional_burden_pleiotropy_deletion	constraint_measure_pTriplo_topdecile	0.4063	0.008	0.0168
functional_burden_pleiotropy_deletion	constraint_measure_GeneLenBP_topdecile	0.3221	0.041	0.0615
functional_burden_pleiotropy_duplication	measure_LEOUF_topdecile	0.4385	0.001	0.0026
functional_burden_pleiotropy_duplication	constraint_measure_missense_Z_topdecile	0.4963	0.001	0.0026
functional_burden_pleiotropy_duplication	constraint_measure_s_het_topdecile	0.392	0.001	0.0026
functional_burden_pleiotropy_duplication	constraint_measure_CDS_topdecile	0.4903	0.001	0.0026
functional_burden_pleiotropy_duplication	constraint_measure_pHaplo_topdecile	0.213	0.156	0.182
functional_burden_pleiotropy_duplication	constraint_measure_pTriplo_topdecile	0.4108	0.001	0.0026
functional_burden_pleiotropy_duplication	constraint_measure_GeneLenBP_topdecile	0.246	0.093	0.1302
functional_pleiotropy_GWASenrichment	measure_LEOUF_topdecile	0.5182	0.026	0.042
functional_pleiotropy_GWASenrichment	constraint_measure_missense_Z_topdecile	0.6214	0.001	0.0026
functional_pleiotropy_GWASenrichment	constraint_measure_s_het_topdecile	0.2652	0.241	0.253
functional_pleiotropy_GWASenrichment	constraint_measure_CDS_topdecile	0.595	0.002	0.0047
functional_pleiotropy_GWASenrichment	constraint_measure_pHaplo_topdecile	0.2872	0.227	0.2509
functional_pleiotropy_GWASenrichment	constraint_measure_pTriplo_topdecile	0.5111	0.023	0.0403
functional_pleiotropy_GWASenrichment	constraint_measure_GeneLenBP_topdecile	0.4003	0.127	0.1569

Column descriptions

Column	Description
stat	Functional pleiotropy measure (deletion, duplication, or GWAS enrichment) used in the correlation test
constraint	Specific gene constraint measure being tested (e.g., LOEUF top decile, pLI top decile, S-het top decile)
cor	Pearson correlation coefficient between functional pleiotropy and the constraint measure
pvalue_Jaccard	Raw p-value (P-Jaccard permutation test)
adj_pvalue_Jaccard	FDR-adjusted p-value (P-Jaccard permutation test)

Supplementary Table 6

CNV properties: length distribution and number of protein-coding genes encompassed per unique CNV (deletions and duplications).

Metric	DEL	DUP	ALL_UNIQUE_CNVs
Unique_Count	10097	28619	38716
CNV_Size_Mean_bp	510107.46	632551.85	600618.77
CNV_Size_Median_bp	229584	442360	397162
CNV_Size_Min_bp	50065	50105	50065
CNV_Size_Max_bp	8978549	9975356	9975356
CNV_Size_IQR_bp	406293	459917.5	480125.5
Genes_Encompassed_Mean	3.55	4.64	4.35
Genes_Encompassed_Median	2	3	2
Genes_Encompassed_Min	1	1	1
Genes_Encompassed_Max	70	186	186
Genes_Encompassed_IQR	3	4	4

Row descriptions

Row	Description
CNV_Type	CNV type (DEL for deletions, DUP for duplications)
Unique_Count	Number of unique CNVs (distinct loci) included in the analysis
CNV_Size_Mean_bp	Mean size of unique CNVs, in base pairs (bp)
CNV_Size_Median_bp	Median size of unique CNVs, in base pairs (bp)
CNV_Size_Min_bp	Minimum size of unique CNVs, in base pairs (bp)
CNV_Size_Max_bp	Maximum size of unique CNVs, in base pairs (bp)
CNV_Size_IQR_bp	Interquartile range (IQR) of unique CNV sizes, in base pairs (bp)
Genes_Encompassed_Mean	Mean number of protein-coding genes fully encompassed per unique CNV
Genes_Encompassed_Median	Median number of protein-coding genes fully encompassed per unique CNV
Genes_Encompassed_Min	Minimum number of protein-coding genes fully encompassed per unique CNV
Genes_Encompassed_Max	Maximum number of protein-coding genes fully encompassed per unique CNV
Genes_Encompassed_IQR	Interquartile range (IQR) of the number of protein-coding genes fully encompassed per unique CNV