

# Supplementary Information: Overcoming false-positive gene-category enrichment in the analysis of spatially resolved transcriptomic brain atlas data.

Ben D. Fulcher<sup>1,2</sup>, Aurina Arnatkevičiūtė<sup>2</sup>, and Alex Fornito<sup>2</sup>

<sup>1</sup>School of Physics, Sydney University, Camperdown, 2006, NSW, Australia

<sup>2</sup>The Turner Institute for Brain and Mental Health, School of Psychological Sciences, and Monash Biomedical Imaging, Monash University, Victoria, Australia.

## S1 GO Category Case Study

To demonstrate the mechanisms underlying the major differences between randomized-gene nulls and randomized-phenotype nulls, especially in the case of spatially embedded (and spatially autocorrelated) transcriptional atlas data, we investigated the null distributions of three specific GO categories in detail. We selected two similarly-sized GO categories in mouse that have different levels of within-category coexpression: ‘regulation of dendritic spine morphogenesis’ (RDSM, 40 genes), and ‘zymogen activation’ (ZA, 42 genes).

### S1.1 Within-category coexpression drives false-positive bias

We first aimed to understand why, under the ensemble of random spatial brain phenotypes (SBPs), the ‘SBP-random’ ensemble, RDSM (CFPR = 13%) has a higher category false-positive rate (CFPR) than ZA (0.07%). Region  $\times$  gene expression matrices and gene  $\times$  gene coexpression matrices are plotted for RDSM and ZA in Figs S3A and B, respectively. Compared to a representative random set of 40 genes, shown in Fig. S3C, we find a much more consistent spatial patterning of gene expression in the RDSM category, resulting in increased within-category (gene–gene) coexpression. This spatial coherency of expression patterning is consistent with genes associated with regulating dendritic-spine morphogenesis (RDSM) having a coordinated brain function that varies characteristically across brain regions. By contrast, GO categories that play a minimal or less-specific role in brain function, such as zymogen activation (ZA), exhibit noisier expression patterns and lower within-category coexpression (that is visually similar to that of a random set of genes).

To understand how these differences in within-category coexpression affect category false-positive rates, we investigated the distribution of mean correlation coefficients between the expression profiles of genes in each category and the ensembles of randomized SBPs analyzed above. These category-score distributions are plotted for SBP-random and SBP-spatial ensembles in Figs S3D and E, respectively. A category with wider distributions than a random set of genes (the null comparison used in GSEA) will have a greater probability of obtaining a significant correlation to that ensemble of phenotypes. First we note that, consistent with both SBP-random and SBP-spatial ensembles containing no information about gene expression, all category-score distributions are symmetric about zero.

Under random phenotypes (Fig. S3D), RDSM has a wider distribution of category scores than ZA; it is more likely to exhibit a higher correlation to a random SBP. This widening is driven by the increased coexpression of RDSM genes, such that a chance correlation between a

random SBP and any single RDMS gene is amplified through a similar correlation with many other genes in the category. By contrast, categories like ZA contain genes with lower gene–gene coexpression, such that a chance correlation between an SBP and a given ZA gene is more likely to be cancelled out by a chance correlation in the opposite direction for another gene, driving category scores towards zero and resulting in a narrower category-score distribution.

In estimating category  $p$ -values, conventional GSEA compares the category score to that of a random set of the same number of genes, shown gray in Fig. S3D (where we have performed many randomizations of sets of 40 genes so as not to bias towards any particular random set). Due to the high coexpression of RDMS genes, this occurs for approximately 13% of SBP-random phenotypes, but for ZA, which has a similar coexpression structure as random genes, this occurs for just 0.07% of random SBPs.

## S1.2 The role of spatial autocorrelation

Relative to the SBP-random ensemble (of purely random phenotypes), the CFPR of RDMS increased under the SBP-spatial ensemble: 13%  $\rightarrow$  28%, whereas the CFPR for ZA decreased: 0.07%  $\rightarrow$  0.02%. Examining the category-score distributions in Fig. S3E, we first note that they are all much wider than for the SBP-random ensemble (Fig. S3D), including for random sets of genes. This is due to the predominance of spatial autocorrelation in the expression patterns of individual genes: a gene that exhibits a strongly autocorrelated expression map is more likely to be strongly correlated to a spatially autocorrelated SBP than a random spatial map. Relative to random genes, we see a wider category-score distribution for RDMS, and a narrower distribution for ZA, consistent with their CFPRs. Imposing the constraint of spatial autocorrelation (i.e., SBP-rand  $\rightarrow$  SBP-spatial) can thus either increase a GO category’s CFPR (if it exhibits a more similar spatial autocorrelation structure to the SBP-spatial ensemble) or decrease it (if it exhibits a less similar spatial autocorrelation structure).

## S2 Phenotype Enrichment

Full GO enrichment results for all phenotypes across all three null models are provided as data files in the data repository accompanying this article [1].

For human cortex:

- Node structural connectivity betweenness,  $B$ , `EnrichmentThreeWays_human_cortex_betweenness.csv`
- Node structural connectivity degree,  $k$ , `EnrichmentThreeWays_human_cortex_degree.csv`

And for mouse (whole-brain and cortex), where all files share the `EnrichmentThreeWays_mouse` prefix; suffixes are given below:

- Node structural connectivity betweenness: `_all_betweenness.csv`, `_cortex_betweenness.csv`
- Node structural connectivity degree: `_all_degree.csv`, `_cortex_degree.csv`
- VIP+ cell density: `_all_VIP.csv`, `_cortex_VIP.csv`
- SST+ cell density: `_all_SST.csv`, `_cortex_SST.csv`
- PV+ cell density: `_all_PV.csv`, `_cortex_PV.csv`
- oligodendrocyte density: `_all_oligodendrocytes.csv`, `_cortex_oligodendrocytes.csv`
- neural density: `_all_neurons.csv`, `_cortex_neurons.csv`
- glia density: `_all_glia.csv`, `_cortex_glia.csv`

- microglia density: `_all_microglia.csv`, `_cortex_microglia.csv`
- excitatory density: `_all_excitatory.csv`, `_cortex_excitatory.csv`
- inhibitory density: `_all_inhibitory.csv`, `_cortex_inhibitory.csv`

### S3 Literature Survey

Enrichment on the basis of spatial patterns of expression has not been performed consistently in the existing literature. Studies have processed the data differently (including substantially different methods for normalizing and filtering genes), and performed the enrichment differently (using different software packages, annotation systems, and including different sets of categories for enrichment). Our survey focuses on studies that have reported gene-set enrichment for Biological Processes in the Gene Ontology (GO) [2].

#### S3.1 Allen Mouse Brain Atlas

The following studies report results of GSEA using transcriptional data from the Allen Mouse Brain Atlas (AMBA) [3]:

1. French et al. [4] (mouse expression and rat connectivity) 3976 coronal genes using ORA (*ermineJ* [5]) for the two anti-correlated expression patterns: NE (neuron-enriched pattern) and OE (oligodendrocyte-enriched pattern). Data taken from Supplementary Data Sheet 3.
2. French and Pavlidis [6] (mouse expression and rat connectivity) 17 530 genes using ORA (*ermineJ* [5]). Data taken from proximity-corrected enrichment for (i) outgoing connectivity, and (ii) incoming connectivity (Table S5).
3. Ji et al. [7]: Tabulated enrichment categories that appeared often across many analyses using 4084 coronal section genes: all brain structures and different ways of measuring connectivity. Data taken from Table 1.
4. Fakhry and Ji [8]: 4084 (coronal-section) genes that are predictive of voxel-level brain connectivity using ORA. Data taken from Fig. 5, summarizing each GO category as the number of injection structures for which it was significant.
5. Rubinov et al. [9]: 3380 genes (assayed multiple times) scored by partial least squares for nodal participation metrics, doing enrichment on the genes in the top 25% of positive weights. Data taken from Table S3.
6. Fulcher and Fornito [10]: 17 642 genes scored by differences in gene coexpression contribution (GCC) scores using GSR using *ermineJ* [5]. Two GSEA were performed, for: (i) connected versus unconnected pairs of brain areas (data taken from Table S1), (ii) rich and feeder connections versus peripheral connections (data taken from Table S5).
7. Mills et al. [11]: 3079 (coronal-section) genes, enrichment for processes showing a strong relationship between CGE and functional connectivity (FC) using ORA (*ermineJ* [5]). Data taken from Table 3.
8. Ko et al. [12]: 170 neuron-, 44 oligodendrocyte-, and 50 astrocyte-specific genes for the coronal plane (gene sets taken from [13] using a 10-fold threshold, and including only genes that could be matched to AMBA data). Enrichment performed on each gene set, with results summarized in the text (Page 2).

## S3.2 Allen Human Brain Atlas

The following studies report the results of GSEA using data from the Allen Human Brain Atlas (AHBA) [14]:

1. Tan et al. [15]: 20 444 genes, ORA enrichment (using *DAVID* [16]) on top 100 genes for the most positive and negatively correlated expression patterns (one marks neurons, the other for oligodendrocytes). Data taken from Table S6 ( $p$ -values for ‘positive 100’) and Page 8 text (list of significant categories for ‘negative 100’).
2. Richiardi et al. [17]: 16 906 genes: enrichment on strength of correlated gene expression within functional networks (relative to between functional networks) performed with *DAVID* [16] and Panther [18]. Data taken from Table S6.
3. French and Paus [19]: 20 737 genes, ROC enrichment (ermineJ [5]) for measures of inter-subject consistency and inconsistency. Data taken from supplementary files from [https://figshare.com/articles/A\\_FreeSurfer\\_view\\_of\\_the\\_cortical\\_transcriptome\\_generated\\_from\\_the\\_Allen\\_Human\\_Brain\\_Atlas/1439749](https://figshare.com/articles/A_FreeSurfer_view_of_the_cortical_transcriptome_generated_from_the_Allen_Human_Brain_Atlas/1439749): InconsistentGOGroups.tsv and ConsistentGOGroups
4. Vértes et al. [20]: 20 737 genes, enrichment on topologically integrative hubs (of functional connectivity), for positive and negative contributions to partial least squares components 1, 2, and 3. Data from supplementary file: *Vertes-rstb20150362suppl.xlsx*.
5. Parkes et al. [21]: 19 343 genes, enrichment performed using GSR (ermineJ [5]) on coefficients of principal components (PCs) of gene expression in the striatum, for PCs 1, 2, 5, and 9. Data obtained directly from the author.
6. Forest et al. [22]: 20 783 genes, enrichment performed on gene clusters formed using WGCNA [23] using *Cytoscape* [24]. Data taken from Tables S3 and S8 for full and reduced models, respectively.
7. Whitaker et al. [25]: 20 737 genes, enrichment performed on partial least squares component 2 using GORILLA [26]. Data taken from supplementary data table: *WhitakerVertes\_PLSEnrichmen*
8. Romme et al. [27]: 20 737 genes, enrichment on top 100 genes with strongest correlation to SCZ connectome disconnectivity using ORA (*Panther* [18]). Data taken from supplementary tables.
9. Liu et al. [28]: 20 738 genes, biological process enrichment performed for: (i) left hippocampus (HCP atlas, data taken from Table S4), (ii) middle insular area (HCP atlas) (Table S5), (iii) brain-wide functional connectivity association study of autism (Table S7), and (iv) chronic schizophrenia (Table S8).
10. Kuncheva et al. [29]: 16 906 pre-selected genes, enrichment on clusters of spatial expression networks (SEns) for GO biological process (uncorrected threshold,  $p < 0.001$ ) with categories reduced using *REVIGO*. Full results not provided; summaries taken from text.
11. Ritchie et al. [30]: 13 384 genes, enrichment on spatial correlation (Spearman’s  $\rho$ ) between expression patterns and the T1w:T2w ratio, using the AUROC method, ranked genes by correlation, including GO categories with between 10 and 200 genes after intersection with AHBA data (6885 GO categories). Data taken as  $p_{\text{FWER}}$  for biological processes: negative correlations (Table 2) and positive correlations (Table 3).
12. Diez and Sepulcre [31]: 3719 neural genes (determined from browsing *AmiGO*), enrichment on correlation between stepwise functional connectivity maps and cortical expression

profiles, focusing *a priori* on brain-related categories. Enrichment performed as over-representation analysis on biological processes using *PANTHER* 13.1 at a threshold FDR  $< 0.005$ . Data taken from Supplementary Tables 3–6 as FDR-corrected  $p$ -values.

13. Wen et al. [32] analyzed for gene modules that are correlated to R2t\* using *ToppGene*, with additional testing using *DAVID*. We noted GO Biological Process categories with  $p < 0.05$  from the top 30 Annotation Clusters (up to an enrichment score of 1.27) from Dataset S3.
14. Liu et al. [33] analyzed how dynamic brain networks (analyzed from the viewpoint of the chronnectome) are spatially configured, and how these spatial maps are associated with gene transcription. We took data from Table S5 (Biological Process: PLS1), and Table S9 (Biological Process, PLS3).
15. Anderson et al. [34] analyzed enrichment for limbic network-biased genes using *ToppGene*. We took Biological Process results from the Supplementary Data 1 (sheet: `ToppGene_limbic_n505`).
16. Anderson et al. [35] analyzed enrichment for genes highly correlated to interneuron markers PVALB and SST using *ToppGene*. We took GO:BP data from Table 1.
17. Betzel et al. [36] used *GOrilla* to understand gene subsets that exhibit correlated expression patterns that are most strongly related to ECoG FC, providing results in frequency bands 1 – 4 Hz (biological process results taken from Table S5) and 4 – 8 Hz (biological process results taken from Table S7).
18. Meijer et al. [37] used *GOrilla* to find genes that are differentially expressed genes in the stress network (brain areas activated by stress in individuals with low or high stress sensitivity). Data was taken from Table S4.
19. Romero-Garcia et al. [38] performed PLS to find genes associated with changes in cortical thickness in autism, using *Enrichr* for enrichment in GO biological processes. Data was taken from the main text.
20. Liu et al. [39] investigated how emotion regulation and memory control related fMRI task activation maps correlate with gene expression. 1061 ‘inhibition-related genes’ are correlated with all four tasks (memory control, emotion regulation, stop-signal, and go/no-go). Enrichment done using *GOATOOLS*. Data taken from Table S3.
21. Grothe et al. [40] analyzed differentially-expressed gene sets in Alzheimer’s disease vulnerable brain regions using GSEA (gene set enrichment analysis) using all 20 737 genes (1036 GO-based gene sets were included). Enrichment results for GO biological processes were taken from Table 3 (differentially expressed gene sets in neurodegeneration-vulnerable brain regions).
22. Vidal-Pineiro et al. [41] investigated transcriptional patterns related to cortical thinning across the lifespan. GO-term enrichment was performed with Visual Annotation Display (VLAD). Data was taken from Table S1.
23. Yao et al. [42] proposed a method to perform enrichment by jointly considering gene sets (GS) and brain circuits (BC) to examine if a GS–BC pair is enriched in a list of gene–neuroimaging quantitative traits (QT, such as the average amyloid deposition). Enrichment results for GO biological processes were taken from Table 3.

## S4 Supplementary Figures

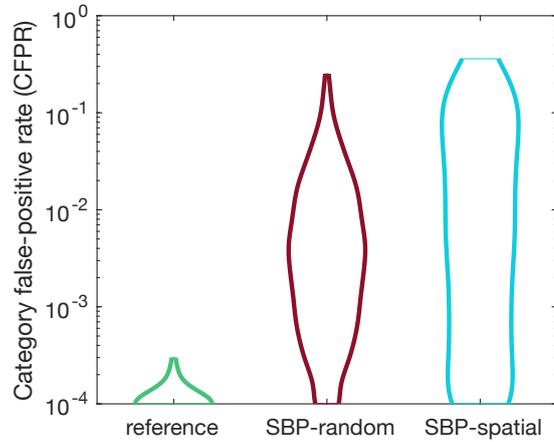


Figure S1: **Category-level false-positive rate (CFPR) across three null ensembles of 10 000 human cortical maps.** See Fig. 2A for information.

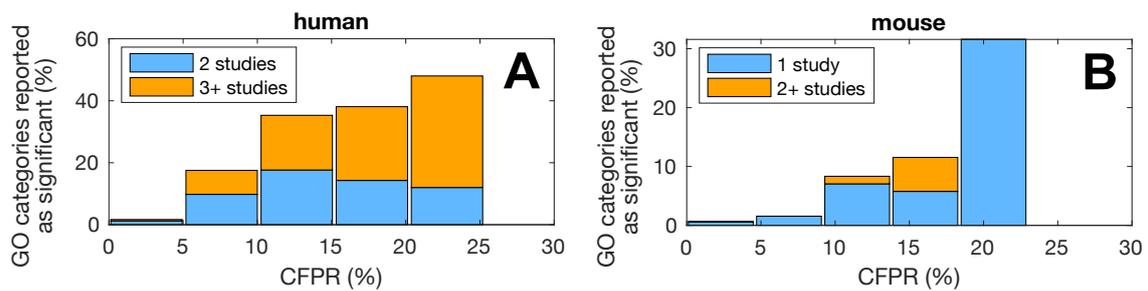
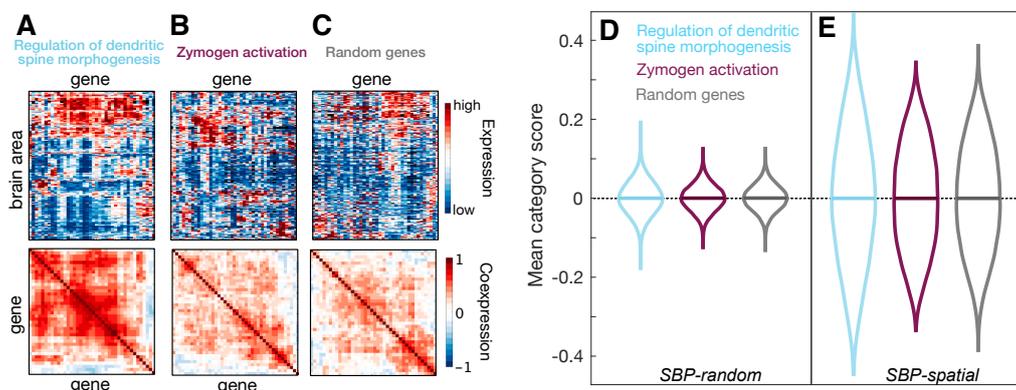


Figure S2: Across five equally-spaced bins of CFPR, bars show the percentage of GO categories that have been reported as significant in the literature survey of **A** human, and **B** mouse. The larger number of human studies allowed us to distinguish between GO categories reported in two, or three or more studies; in mouse we distinguish between one, or two or more studies.



**Figure S3: GO categories with high within-category gene–gene coexpression are most prone to false-positive enrichment.** We plot expression (brain area  $\times$  gene) and coexpression (gene  $\times$  gene) heat maps for two example GO categories in the mouse brain: **A** ‘regulation of dendritic spine morphogenesis’ (40 genes); and **B** ‘zymogen activation’ (42 genes); as well as **C** a random set of 40 genes, for comparison. Each gene’s expression is normalized (low to high) for visualization purposes. Genes annotated to ‘regulation of dendritic spine morphogenesis’ display a more characteristic spatial patterning and hence have higher coexpression. Distributions of each category’s score (mean correlation between the genes in that category and a phenotype) across an ensemble of null phenotypes, are plotted as violin plots for: **D** the SBP-random ensemble of random-number phenotypes [cf. Fig. 2A(ii)], and **E** the SBP-spatial ensemble of random spatially autocorrelated phenotypes [cf. Fig. 2A(iii)]. The mean of each distribution is annotated with a horizontal line.

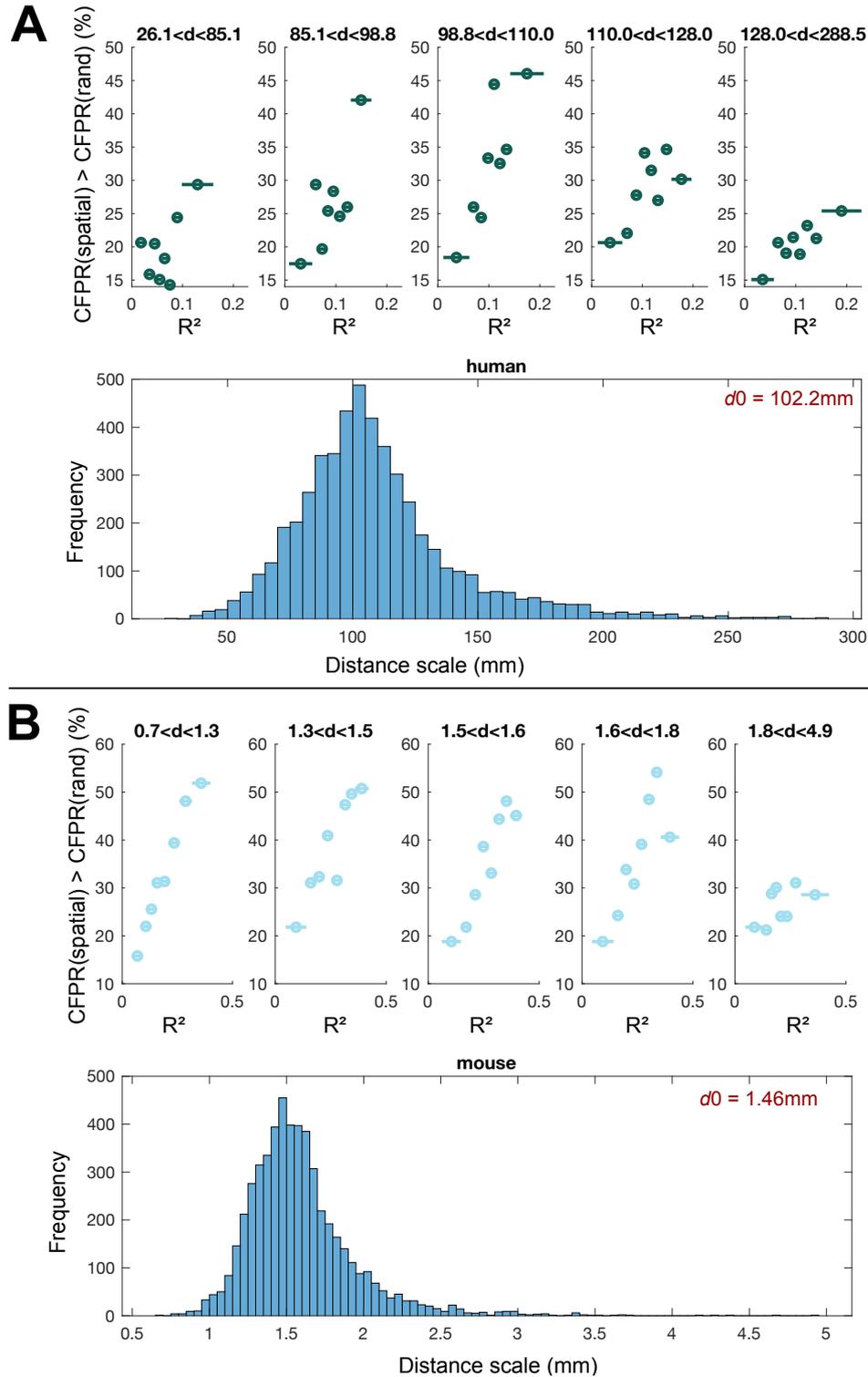


Figure S4: **CFPR (SBP-spatial) > CFPR (SBP-rand) is greatest for GO categories with strong spatial autocorrelation ( $R_{\text{exp}}^2$ ) and a similar distance scale of autocorrelation.** For **A** human, and **B** mouse, we show a distribution of fitted distance scales,  $\lambda$ , across all GO categories. For each category, this was estimated from an exponential fit to correlated gene expression (of a given category of genes) as a function of distance. The global value,  $d_0$  (obtained from including all genes in the correlated gene expression calculation, and used to construct the SBP-spatial ensemble) is annotated in the upper right of the distribution plot. Across equiprobable bins of  $R_{\text{exp}}^2$  and for a given distance range, upper plots show the percentage of GO categories in each bin that display an increase in CFPR under the SBP-spatial null ensemble relative to the SBP-random null ensemble (as Fig. 3B).

## S5 Supplementary Tables

GO Category	Mouse CFPR [SBP-spatial] (%)	Human CFPR [SBP-spatial] (%)	References
Regulation of synaptic plasticity	23 [37]	10 [7]	[10, 19, 28, 33, 37]
Regulation of postsynaptic membrane potential	19 [35]	10 [17]	[10, 19, 31]
Glutamate receptor signaling pathway	16 [34]	6 [14]	[11, 19, 28, 42]
Respiratory electron transport chain	12 [4]	20 [30]	[10, 19–21, 25]
Learning	20 [34]	5 [2]	[15, 19, 28, 37]

Table S1: **GO categories with high category false-positive rates (CFPRs) are predominantly related to neuronal and metabolic biological function in mouse and human.** We list selected GO categories with amongst the highest CFPRs across mouse and human, across ensembles of spatially independent random phenotypes (SBP-random) and spatially autocorrelated random phenotypes (SBP-spatial). CFPR is listed for SBP-random [and for SBP-spatial in brackets]. A full ordered list is in Supplementary File `CFPRTable.csv`.

## References

- [1] B. Fulcher. "data files to support reproducing analyses in 'overcoming bias in gene category enrichment analyses of brain-wide transcriptomic data' [dataset]". ver. 2 (2021. [Online].).
- [2] M. Ashburner, C. A. Ball, J. A. Blake, et al. Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25 (2000).
- [3] E. Lein, M. J. Hawrylycz, N. Ao, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168 (2007).
- [4] L. French, P. P. C. Tan, and P. Pavlidis. Large-scale analysis of gene expression and connectivity in the rodent brain: insights through data integration. *Front. Neuroinf.* **5**, 12 (2011).
- [5] J. Gillis, M. Mistry, and P. Pavlidis. Gene function analysis in complex data sets using ErmineJ. *Nat. Protoc.* **5**, 1148 (2010).
- [6] L. French and P. Pavlidis. Relationships between gene expression and brain wiring in the adult rodent brain. *PLoS Comp. Biol.* **7**, e1001049 (2011).
- [7] S. Ji, A. Fakhry, and H. Deng. Integrative analysis of the connectivity and gene expression atlases in the mouse brain. *NeuroImage* **84**, 245 (2014).
- [8] A. Fakhry and S. Ji. High-resolution prediction of mouse brain connectivity using gene expression patterns. *Methods* **73**, 71 (2015).
- [9] M. Rubinov, R. J. F. Ypma, C. Watson, and E. T. Bullmore. Wiring cost and topological participation of the mouse brain connectome. *Proc. Natl. Acad. Sci. USA* **112**, 10032 (2015).
- [10] B. D. Fulcher and A. Fornito. A transcriptional signature of hub connectivity in the mouse connectome. *Proc. Natl. Acad. Sci. USA* **113**, 1435 (2016).
- [11] B. D. Mills, D. S. Grayson, A. Shunmugavel, et al. Correlated gene expression and anatomical communication support synchronized brain activity in the mouse functional connectome. *J. Neurosci.* **38**, 2910 (2018).
- [12] Y. Ko, S. A. Ament, J. A. Eddy, et al. Cell type-specific genes show striking and distinct patterns of spatial expression in the mouse brain. *Proc. Natl. Acad. Sci. USA* **110**, 3095 (2013).
- [13] J. D. Cahoy, B. Emery, A. Kaushal, et al. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J. Neurosci.* **28**, 264 (2008).
- [14] M. J. Hawrylycz, E. Lein, A. L. Guillozet-Bongaarts, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391 (2012).
- [15] P. P. C. Tan, L. French, and P. Pavlidis. Neuron-Enriched Gene Expression Patterns are Regionally Anti-Correlated with Oligodendrocyte-Enriched Patterns in the Adult Mouse and Human Brain. *Front. Psychiat.* **7** (2013).
- [16] G. Dennis Jr, B. T. Sherman, D. A. Hosack, and J. Yang. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol.* (2003).
- [17] J. Richiardi, A. Altmann, A.-C. Milazzo, et al. Correlated gene expression supports synchronous activity in brain networks. *Science* **348**, 1241 (2015).

- [18] H. Mi, A. Muruganujan, and P. D. Thomas. PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucl. Acid. Res.* **41**, D377 (2013).
- [19] L. French and T. Paus. A FreeSurfer view of the cortical transcriptome generated from the Allen Human Brain Atlas. *Front. Psychiat.* **9**, 25 (2015).
- [20] P. E. Vértes, T. Rittman, K. J. Whitaker, et al. Gene transcription profiles associated with inter-modular hubs and connection distance in human functional magnetic resonance imaging networks. *Phil. Trans. Roy. Soc. B* **371**, 20150362 (2016).
- [21] L. Parkes, B. D. Fulcher, M. Yücel, and A. Fornito. Transcriptional signatures of connectomic subregions of the human striatum. *Genes, Brain and Behavior* **25**, 1176 (2017).
- [22] M. Forest, Y. Iturria-Medina, J. S. Goldman, et al. Gene networks show associations with seed region connectivity. *Hum. Brain Map.* **38**, 3126 (2017).
- [23] P. Langfelder and S. Horvath. Wgcna: an r package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
- [24] P. Shannon, A. Markiel, O. Ozier, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **13**, 2498 (2003).
- [25] K. J. Whitaker, R. Romero-Garcia, F. Váša, et al. Adolescence is associated with genomically patterned consolidation of the hubs of the human brain connectome. *Proc. Natl. Acad. Sci. USA* **113**, 201601745 (2016).
- [26] E. Eden, R. Navon, I. Steinfeld, D. Lipson, and Z. Yakhini. GOrilla: a tool for discovery and visualization of enriched go terms in ranked gene lists. *BMC Bioinformatics* **10**, 48 (2009).
- [27] I. A. C. Romme, M. A. de Reus, R. A. Ophoff, R. S. Kahn, and M. P. van den Heuvel. Connectome Disconnectivity and Cortical Gene Expression in Patients With Schizophrenia. *Biol. Psychiatry* **81**, 495 (2016).
- [28] Z. Liu, E. T. Rolls, J. Zhang, et al. The functional and genetic associations of neuroimaging data: a toolbox. *bioRxiv* p. 178640 (2017).
- [29] Z. Kuncheva, M. L. Krishnan, and G. Montana. Exploring brain transcriptomic patterns: a topological analysis using spatial expression networks. In *PACIFIC SYMPOSIUM ON BIOCOMPUTING 2017*, pp. 70–81. World Scientific (2017).
- [30] J. Ritchie, S. P. Pantazatos, and L. French. Transcriptomic characterization of MRI contrast, focused on the T1-w/T2-w ratio in the cerebral cortex. *NeuroImage* **174**, 504 (2018).
- [31] I. Diez and J. Sepulcre. Neurogenetic profiles delineate large-scale connectivity dynamics of the human brain. *Nat. Comm.* **9**, 1 (2018).
- [32] J. Wen, M. S. Goyal, S. V. Astafiev, M. E. Raichle, and D. A. Yablonskiy. Genetically defined cellular correlates of the baseline brain MRI signal. *Proc. Natl. Acad. Sci. USA* **42**, 201808121 (2018).
- [33] J. Liu, M. Xia, X. Wang, X. Liao, and Y. He. Transcriptional Signatures of the Hierarchical Chronnectome Organization in the Human Brain. *bioRxiv* **39**, 637942 (2019).
- [34] K. M. Anderson, F. M. Krienen, E. Y. Choi, et al. Gene expression links functional networks across cortex and striatum. *Nat. Comm.* **9**, 1428 (2018).

- [35] K. M. Anderson, M. A. Collins, R. Chin, et al. The transcriptional landscape of cortical interneurons underlies in-vivo brain function and schizophrenia risk. *bioRxiv* p. 481036 (2018).
- [36] R. F. Betzel, J. D. Medaglia, A. E. Kahn, et al. Structural, geometric and genetic factors predict interregional brain connectivity patterns probed by electrocorticography. *Nat. Biomed. Eng.* **106**, 1 (2019).
- [37] M. Meijer, A. Keo, J. M. C. van Leeuwen, et al. Molecular characterization of the stress network in the human brain. *bioRxiv* **526**, 661587 (2019).
- [38] R. Romero-Garcia, V. Warrier, E. T. Bullmore, S. Baron-Cohen, and R. A. I. Bethlehem. Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism. *Molecular Psychiatry* **24**, 1053 (2019).
- [39] W. Liu, N. Peeters, G. Fernández, and N. Kohn. Common neural and transcriptional correlates of inhibitory control underlie emotion regulation and memory control. *Social Cognitive and Affective Neuroscience* **15**, 523 (2020).
- [40] M. J. Grothe, J. Sepulcre, G. Gonzalez-Escamilla, et al. Molecular properties underlying regional vulnerability to alzheimer’s disease pathology. *Brain* **141**, 2755 (2018).
- [41] D. Vidal-Pineiro, N. Parker, J. Shin, et al. Cellular correlates of cortical thinning throughout the lifespan. *bioRxiv* p. 585786 (2019).
- [42] X. Yao, J. Yan, S. Kim, et al. Two-dimensional enrichment analysis for mining high-level imaging genetic associations. *Brain Informatics* **4**, 27 (2017).