Exploring Molecular Mechanisms of Comorbidity: A Network-Based Analysis of ADHD and Autism

Dewy Nijhof, Oksana Sorokina, J Douglas Armstrong

Manually curated ASD and ADHD gene lists, filtered for synaptic genes and supported by literature evidence, including overlapping and disorder-specific genes.

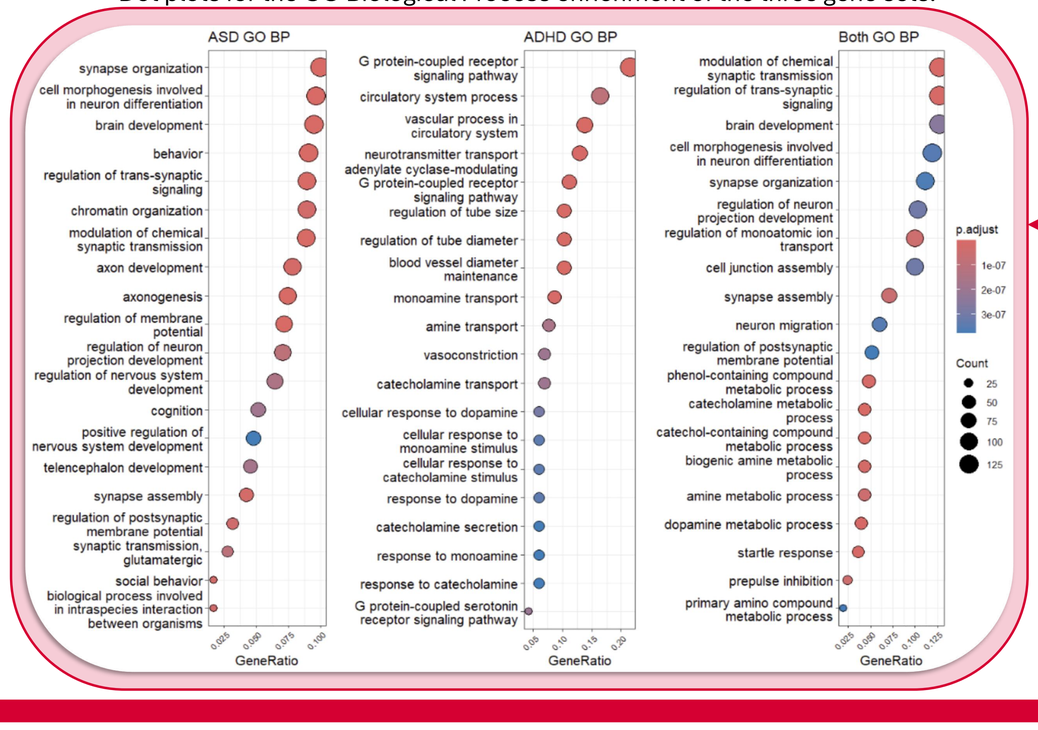
	A	B	C	D	E	F	G	H	I
1	HumanNa	ConditionI	HumanEnt	For.Agains	PMID	DOI	Species	SFARI	ADHDgene
2	AADAT	1	51166	for	23731516	NA	human	no	no
3	AADAT	1	51166	for	39391062	NA	human	no	no
4	ABAT	1	18	for	11918424	NA	human	yes	no
5	ABAT	1	18	for	35982159	NA	human	yes	no
6	ABAT	1	18	for	31398340	NA	human	yes	no
7	ABAT	1	18	for	15830322	NA	human	yes	no
8	ABCA13	2	154664	for	29092799	NA	drosophila	yes	no
9	ABCA13	2	154664	for	38357255	NA	human	yes	no
10	ABCA13	2	154664	for	35813072	NA	human	yes	no

Gene list allows to:
- Select genes by number of evidence
- Check enrichment for the different gene sets (ASD, ADHD, overlapping genes ('Both'))
- Track evidence lines

Majority of evidence for a gene-disease association stems from a single point of research

GO, **KEGG**, and **Reactome** show a pattern of disease-specific and overlapping enrichment terms:
- ASD is driven *by epigenetic and developmental disruption* affecting wiring of the brain
- ADHD is driven by *monoamine and GCPR signalling disruption* affecting modulatory control of communication between cells
- Their comorbidity is driven by *disruption of calcium/ion homeostasis* affecting the core mechanisms of excitability and signalling (linking development and neuromodulation).

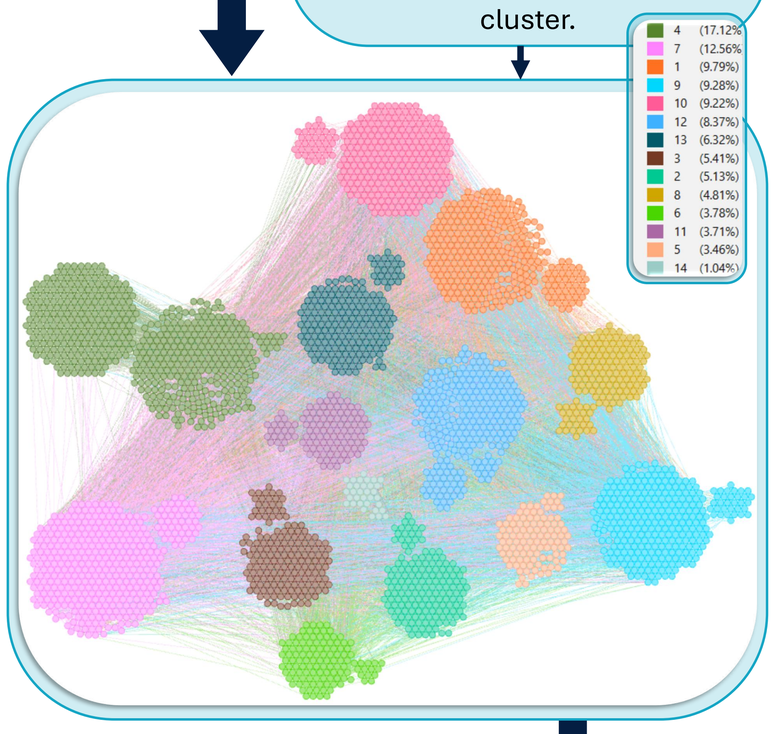
Dot plots for the GO Biological Process enrichment of the three gene sets.



SynaptomeDB is a curated resource of mammalian synaptic proteins and their high-confidence interactions. Covering postsynaptic, presynaptic synaptic vesicle, and synaptosome compartments, with annotations for *function*, *brain region*, and *disease*. It allows users to build customisable synaptic networks and integrate gene or experimental data without needing database expertise.

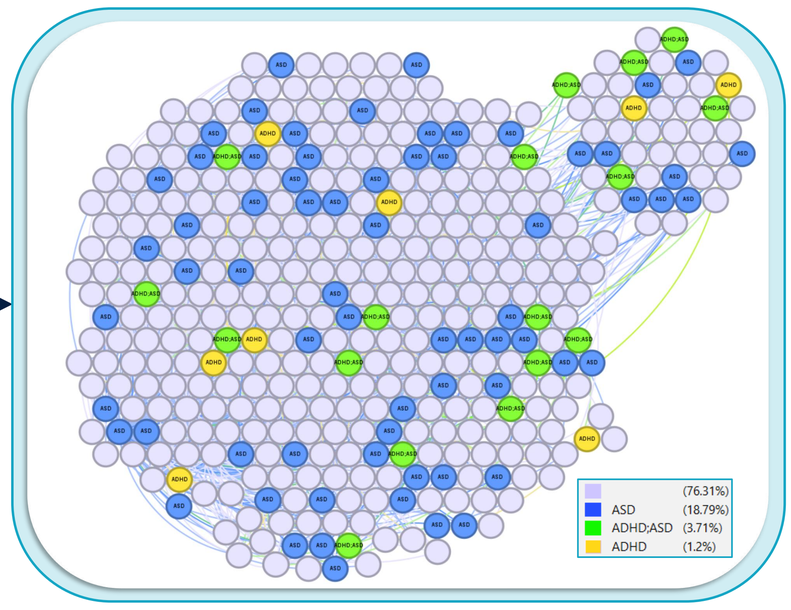
QR code links to SynaptomeDB package.

Example of the postsynaptic network clustered with the Louvain algorithm (visualised with Gephi). The network includes 3,178 nodes (2,425 unique genes) and 17,219 edges. Of our curated gene lists, 715 ASD genes, 156 ADHD genes, and 118 overlapping genes are present. Each colour indicates a cluster (see legend), with percentages showing the proportion of proteins per cluster.



Network modules were identified using **BioNAR** with community detection algorithms (e.g., SpinGlass, Louvain, InfoMAP).

Custom disease annotations (ASD, ADHD, overlap) were overlaid to reveal disorder-specific and shared modules. *The figure on the right* shows an example of this overlay on cluster 7 from the network shown above.



QR code links to BioNAR package.

Cluster enrichment was compared to gene list enrichment to explore potential functions of disorder-specific and shared modules.

Module overlap was assessed with **BioNAR** by comparing observed ASD/ADHD co-localization to random networks, and by contrasting our custom annotations with HDO disease terms.

ABSTRACT Autism (ASD) and ADHD often occur together, but their shared molecular mechanisms are not well understood. We curated gene lists for both disorders and integrated them with a synaptic protein-protein interaction network using BioConductor packages **BioNAR** and **SynaptomeDB**, where our curated gene lists served as annotations in the larger network. Subsequent clustering and enrichment analysis are expected to reveal molecular modules specific to each disorder as well as shared modules underlying comorbidity.

METHODS

Gene curation
- Gene lists for ASD and ADHD were compiled from databases (SFARI, ADHDgene, ClinVar) and filtered for synaptic genes. Each gene was manually checked against the literature for evidence of association.

Network construction
- Synaptic protein-protein interaction networks were built using **SynaptomeDB** and **BioNAR**. Only interactions supported by ≥2 publications and located in the postsynaptic compartment were included. The largest connected component (LCC) was used for analysis.

Clustering
- Community detection was performed with **BioNAR** using multiple algorithms (fast-greedy, Walktrap, Louvain, InfoMAP) to identify network modules. Default parameters were used.

Enrichment analysis
- Clusters will be tested for overrepresentation of ASD- or ADHD-associated genes. Gene Ontology enrichment will be applied to interpret cluster function.

Disease overlap analysis
- **BioNAR** will then be used to quantify overlap between disease-associated modules, comparing observed network overlap to randomly permuted networks to assess significance.

CONCLUSION This integrated, data-driven pipeline combines **BioNAR**, **SynaptomeDB**, and curated gene lists to enable systematic exploration of comorbidity at the synaptic and brain region level. By linking evidence from literature, curated proteomic resources, and network analysis, it provides a holistic framework for identifying disorder-specific and shared modules. Easy integration with existing enrichment analyses allow for mechanistic insight. The approach highlights the functionality of these packages for comorbidity research and can be further extended with additional data types (e.g., transcriptomics) to deepen our understanding of shared disease mechanisms.

References:
McLean C, Sorokin A, Simpson TI, Armstrong JD, Sorokina O. BioNAR: an integrated biological network analysis package in bioconductor. Forslund S, editor. Bioinforma Adv. 2023 Jan 5;3(1):vbad137.
Sorokina O, Mclean C, Croning MDR, Heil KF, Wysocka E, He X, et al. A unified resource and configurable model of the synapse proteome and its role in disease. Sci Rep. 2021 May 11;11(1):9967.
Sorokina O, Sorokin A, Armstrong JD. Synaptome.db: a bioconductor package for synaptic proteomics data. Bahar I, editor. Bioinforma Adv. 2022 Jan 10;2(1):vbac086.
Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media.

This work was supported by the United Kingdom Research and Innovation (grant EP/S02431X/1), UKRI Centre for Doctoral Training in Biomedical AI at the University of Edinburgh, School of Informatics.

UK Research and Innovation

Biomedical AI
UKRI CENTRE FOR DOCTORAL TRAINING

QR code links to project GitHub and contact details.