



Strengthening

Inference of Possible Novel Autism Risk Genes by Comparative Sociogenomics and Molecular Network Analysis



Consiglio Nazionale delle **Ricerche**

(1) National Research Council, Institute for Biomedical Technologies - Segrate, Italy 🔍 *alessandra.mezzelani@itb.cnr.it

Background: Autism spectrum disorders (ASD) are neurodevelopmental conditions characterized by difficulty with social interaction and communication. Genetics play a key role in ASD aetiology and thousands of autism risk genes have been reported in the SFARI Autism db. Since some of these genes are involved in social behaviour of different animals, including humans, we hypothesized that the genetic underpinnings of sociability should be in common and stable across multiple species and may be implicated in autism.

Methods:

INPUT: • Sociability genes = 659 genes

Network analysis



Network analysis:

- Top genes identification
- Community: gene sharing similar biological function
- Functional characterisation of the communities by using Reactome through Over Representation Analysis (ORA)

STRING Protein-Protein Interaction

network (17'288 genes)



DM find

https://github.com/emosca-cnr/dmfind

Gene analysis:

- Overlap between top genes and SFARIdb
- Overlap between top genes and ASD Brain Differentially Expressed Genes (DEG) in DOI: 10.1038/s41593-023-01361-0
- Gene mapping and enrichment (ORA)

Figure 1 – Top network analysis results. The figure shows the top network resulting from network analysis with shape = Presence in SFARIdb, non-syndromic genes, and size = top genes by degree for each communities. The network are colored by **1A**) communities; **1B**) DEG in ASD Brain. Labels are A) top gene by degree for each communities; B) DEG genes in ASD Brain. Only communities with > 2 genes are shown.

Within DEG in ASD brain, BRCA1, AGER, MED12, FZD9, DBI are also top degree genes.



2A



2B

Figure 2 – Venn diagrams. The diagrams shows **2A**) overlap between top network genes (blue) and SFARI non syndromic genes (green), **2B)** DEG in ASD brain (red). Both overlaps are significant (A: p < 10^-15; B: p = 3.16 * 10^-4)

Sociability

Other

o no



Figure 4 – Functional analysis of the network communities of genes. The heatmap shows the association among genes, topological communities of genes (numbers on the left) and molecular meta-pathways (M1-M5): M1 - cell-cell junction and communication; M2 - inflammatory signaling; M3 - synaptic signaling; M4 - neurotransmitter receptors; M5 - semaphorin signaling.

Only the communities with at least 3 genes and significant association with pathways (hypergeometric test, FDR <



SNAP2

OTOF

Ч Ч Ч Ч Ч Ч Л Л Л П Л Л Л Л Л

Cytogenetic bands

Figure 3 – Frequency of top network genes per chromosome bands.

Values in bars indicate the number of genes detected in the top network (red) mapped in the chromosome band and the remaining genes in the band (light blue). Significance levels: * $p \le 0.25$, ** $p \le 0.1$, *** $p \le 0.05$. Loci 7q11.23, already involved both in Chromosome 7q11.23 deletion syndrome and

Williams Bourden Syndrome, and 7q21.11 have been highlighted with higher statistically significant frequency genes.



Conclusion: Comparative sociogenomics, coupled with advanced bioinformatic methods, allows the identification of conserved gene networks involved in sociability and ASD. Notably, cell-cell junction and communication, inflammatory and synaptic signalling, neurotransmitter receptors, semaphorin signalling are among the more enriched meta-pathways, while the strongest inferred (most interesting) autism risk genes are MED12, FZD9, and DMD. This approach also leads to the prediction of novel autism risk genes that, if validated, could represent novel biomarkers of the condition, shedding light on its complex aetiology.

Authors declare no conflict of interest.

Grants: EU "GEMMA" (825033); EU/MUR "BBMRI.it" (IR0000031)