Novel Autism Risk Genes Inferred by Comparative Sociogenomics and Molecular Network Analysis

Chiodi A[1], Cupaioli FA[1], Mosca E[1], **Mezzelani A**[1*]

(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 behavior 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
- STRING Protein-Protein Interaction network (17'288 genes)

Network Diffusion: dmFind R package

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

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)

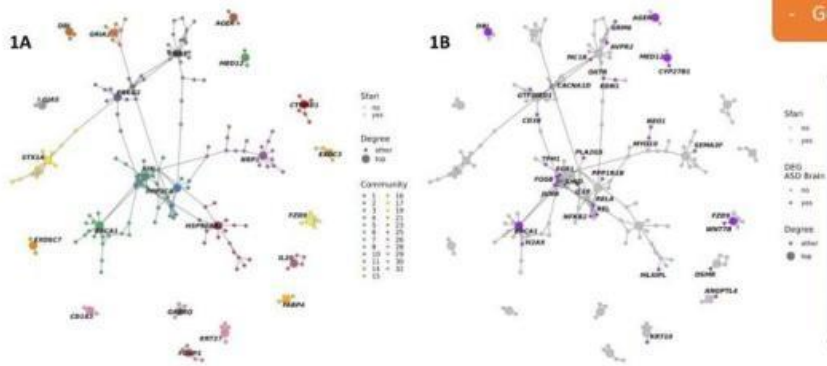
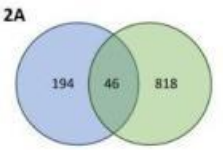


Figure 1 – Top network analysis results. The figure shows the top network resulting from network analysis with shape = Presence in SFARIdb, non-syndromic genes, and size = top genes by degree for each communities. The network are colored by **1A)** communities; **1B)** DEG in ASD Brain. Labels are A)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



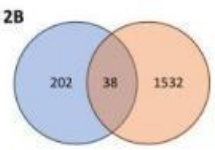


Figure 2 – Venn diagrams. The diagrams shows **2A)** overlap between top network genes (blue) and SFARI non syndromic genes (green), **2B)** DEG in ASD brain (red). Both overlaps are significant (A: p < 10^-15; B: p = 3.16 * 10^-4)

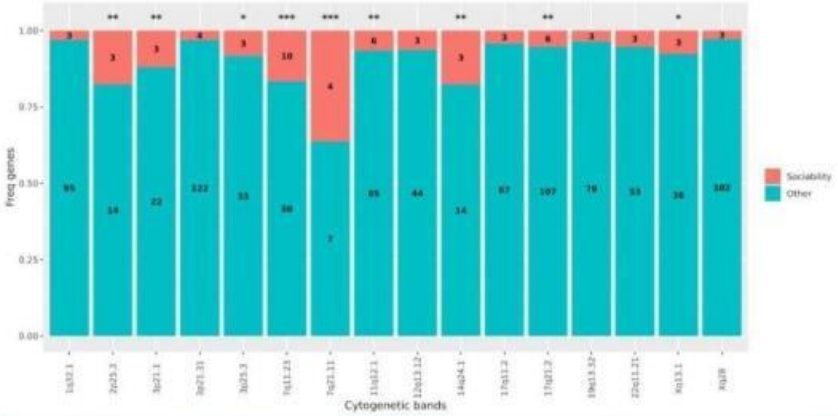


Figure 3 – Frequency of top network genes per chromosome bands.
Values in bars indicate the number of genes detected in the top network (red) mapped in the chromosome band and the remaining genes in the band (light blue). Significance levels: *p ≤ 0.25, ** p ≤ 0.1, *** p ≤ 0.05.

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.

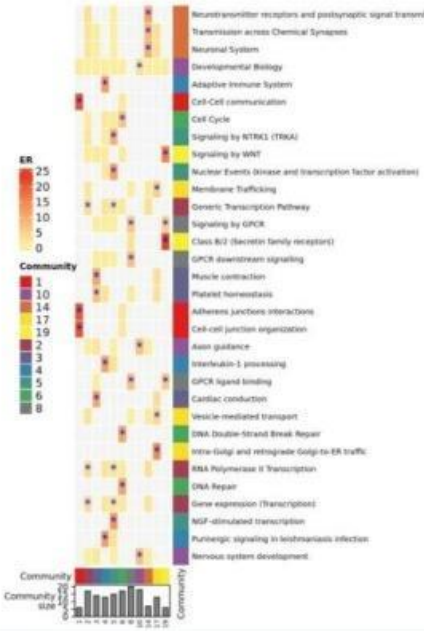


Figure 4 – ORA analysis in the communities. The heatmap shows the top 3 significant (*) pathways in the communities composed by more than 5 genes, coloured by enrichment score (er). On the column (below) colours indicate the communities, while barplot the dimension (number of genes). On the columns colour indicates in which community the pathway is among the top three significant.

Conclusion: Comparative sociogenomics, coupled with advanced bioinformatic methods, allows the identification of conserved gene networks involved in sociability and ASD. Notably, cell cell communication, adherents junctions interaction and cell cell junction organization are among the more enriched pathways and belong to the same community (#1). This approach also lead 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)