An interactome for the analysis of host-microbiota multi-omics data: a proof-of-concept in autism spectrum disorder.

CNR-ITB
Consiglio Nazionale delle Ricerche
Istituto di Tecnologie Biomediche

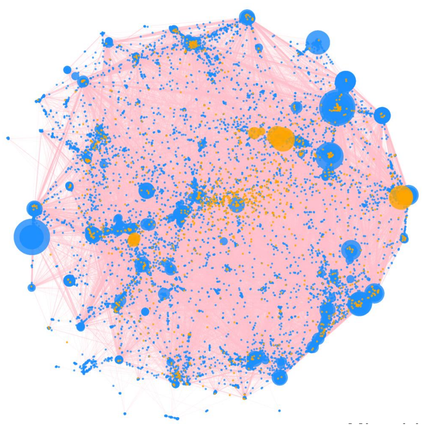
Alice Chiodi[1], Hamed Emamian[2], Ada Sula[1], Andrea Manconi[1], Alessandra Mezzelani[1], **Ettore Mosca**[1*]
(1) Istituto di Tecnologie Biomediche, CNR, Segrate (Milano), Italy.
(2) Dipartimento di Bioscienze, Università degli Studi di Milano, Milano, Italy.
*corresponding author: ettore.mosca@itb.cnr.it

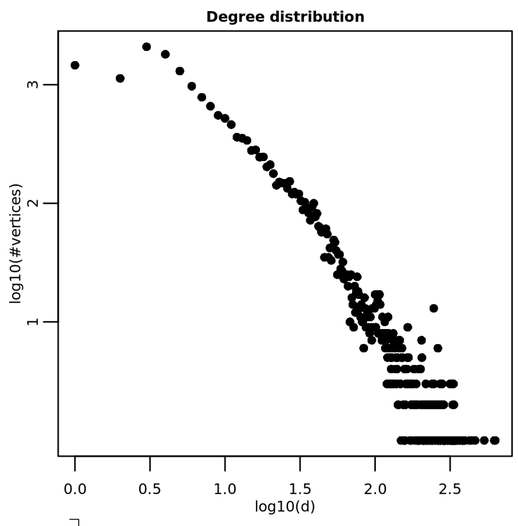
Background & Motivation

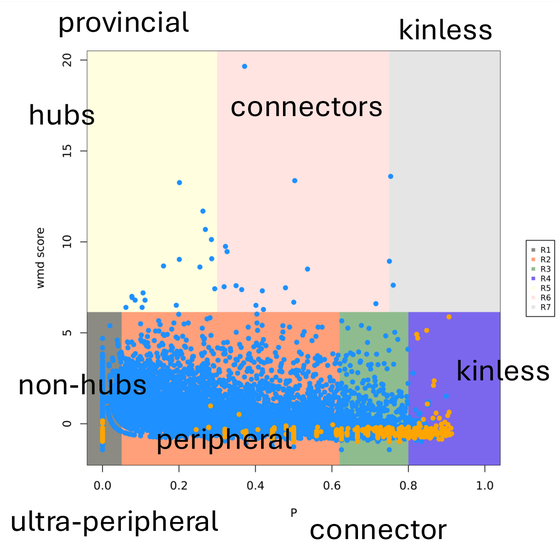
Molecular networks ("**interactomes**") are an **essential tool** in **biomedical research** for **translating multi-omics data** in **valuable insights**. However, the existence of many possible versions of gene-centric interactomes and **the lack of models that go beyond protein-coding genes**, including other biological entities (e.g. non-coding genes, metabolites, microbiota gene products), constitute **major bottlenecks** in the **network-based analysis of multi-omics**. Here, we present an interactome for the analysis of host-microbiota multi-omics data and a proof-of-concept of its use in a multi-omics study on autism spectrum disorder.

Human-microbiome interactome

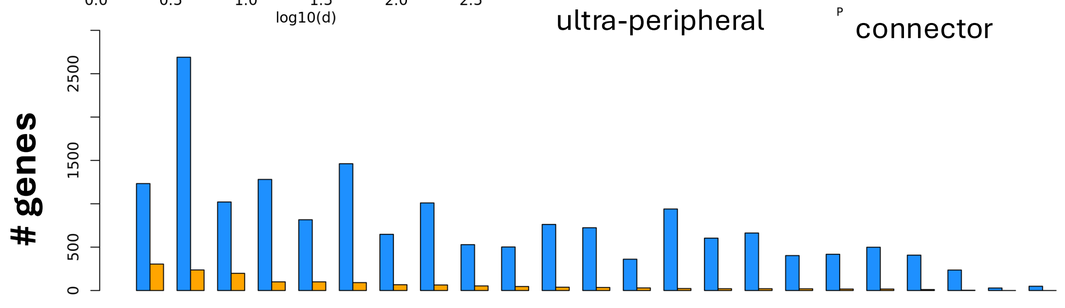
STRING: Protein-Protein Interaction Networks
https://string-db.org, v12
Ensembl protein identifiers mapped to Entrez Gene identifiers
no co-citations, confidence ≥ 700 and top 3 links confidence ≥ 400
HbNet: manually curated, evidence-based host-microbiome interactions
(Zhou et al., Genome Biol, 2022)
Uniprot identifiers mapped to KEGG Orthologues (microbiome) or Entrez Gene identifiers (human)







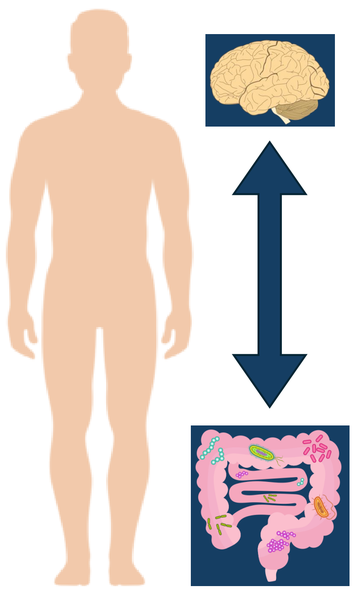
Genes: 18'785 = **17'288 human** + **1'497 microbial**
Gene-gene interactions: 183'654
174'962 human-human + 8'692 human-microbiome

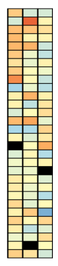


Communities (algorithm: Blondel 2008, J. Stat. Mech.)

Proof-of-concept in ASD datasets with a focus on the gut microbiota-brain axis

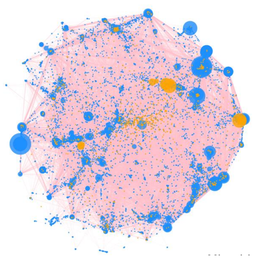
Gut microbiota-brain axis (MGBA)





Data: ASD vs neurotypical subjects
- Predisposition; Genetics (SFARI database)
- Brain gene activity; RNA-seq (post-mortem brain tissues, 4 cohorts)
- Gut microbiota (Shotgun metagenomics from faeces, 3 cohorts)

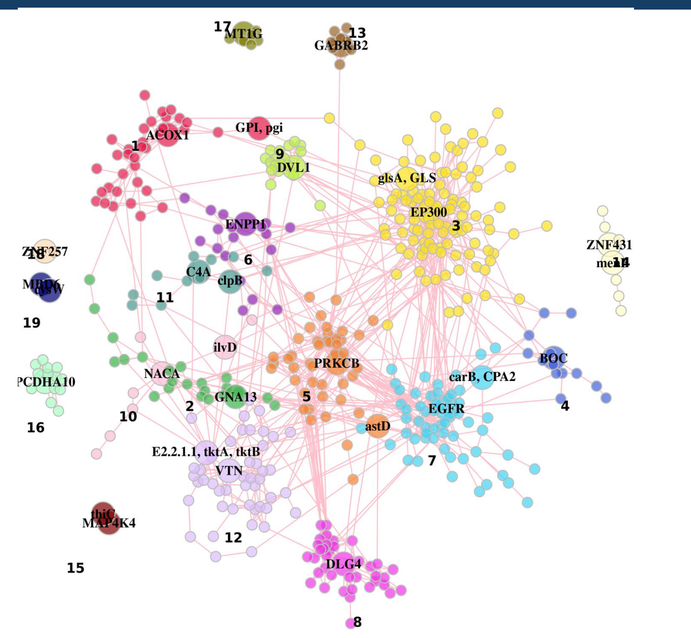
Human-microbiome interactome



Network analysis

$$\frac{\mathrm{d}\mathbf{y}_i}{\mathrm{d}t} = -\mathbf{L}_\alpha \mathbf{y}_i + \beta \mathbf{x}_i$$

mND: Multi-omics network diffusion
https://github.com/emosca-cnr/mND



472 genes, 1470 links, Modularity Q = 0.74

Dataset	# Subjects	Type	AUC
SRP132816	62	RNA-seq (post mortem brain tissue)	0.68
Four RNA-seq datasets	98	RNA-seq (post mortem brain tissue)	0.73
Wang2020	74	Fecal SMS	0.82
Dan2020	60	Fecal SMS	0.82
Averina2020	78	Fecal SMS	0.51
All SMS datasets	212	Fecal SMS	0.55

We reproduced two results of a recent multi-omics study on ASD (Morton 2023, Nat. Neurosci):

1) Over representation analysis of KEGG Pathways
Pathways enriched in microbial genes (FDR < 0.05) control the metabolism of amino acids involved in biosynthesis of neurotransmitters (relevant to the MGBA)

2) Ability to distinguish ASD vs neurotypical subjects
We obtained AUC ranging from 0.5 to 0.8. This variation can be brought back to various sources of heterogeneity (e.g., samples size, batch effects, technological platforms, missing data).

In **conclusion**, despite the limited availability of human-microbiome molecular interaction data**, the proof-of-concept yielded promising results** (pathways and prediction performance), coherent with what a recent study (Morton et al., 2023) found. The results achieved so far encourage the use of our human-microbiome interactome for the analysis of multi-omics data involving human and microbiota.

GEMMA
Multi-Omics Research in Autism

CNR.BiOmics
BIG DATA FOR BETTER LIFE

Strengthening BBMRI.it