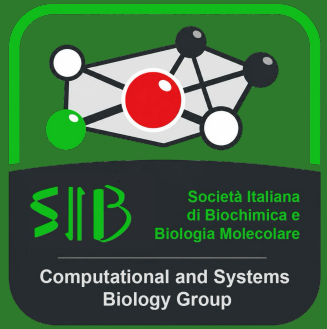
S4 · 16:00–16:15

Mapping the Neurodevelopmental Disorder Gene Landscape: From Data Integration to Multi-Ontology Candidate Prioritization

Chiara Rivi

University of Bologna

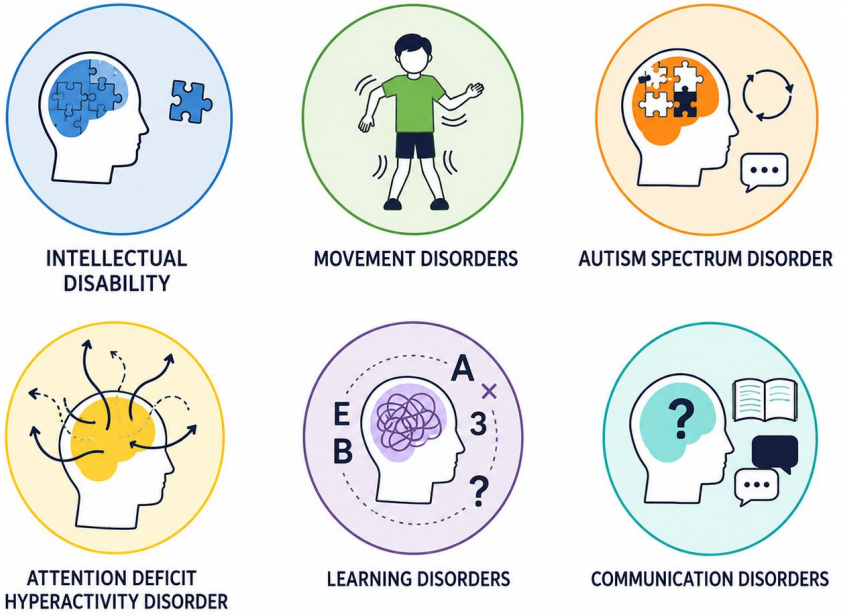


What are Neurodevelopmental Disorders?

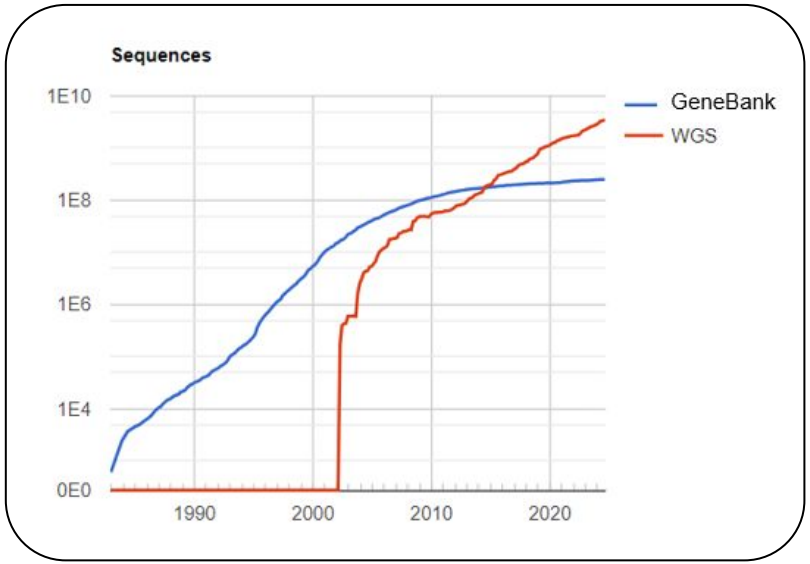
Complex, **early-onset** conditions characterised by impairments in **brain functions**

~4%
Global Prevalence

Males more affected



Diagnostic Difficulties



Multifactorial Contribution

Environmental factors, common and rare genetic variants contribute to the phenotype.

Clinical Challenge

The **difficulty in obtaining a precise genetic diagnosis** constitutes a huge drawback in clinical practice.

Shift in Methodology

Research has shifted to genome-wide approaches, especially after the introduction of **next-generation sequencing (NGS)** techniques.

Aims & Research Objectives

Functional Analysis

Functional characterization of genes associated with NDDs.

Expand the list of causative genes

Expand the list of known causative genes, contributing to the improvement of the diagnostic process.

Prioritize candidate NDD genes

Develop a Multi-Ontology Enrichment (MOE) score to rank candidate genes by functional similarity to the curated set.

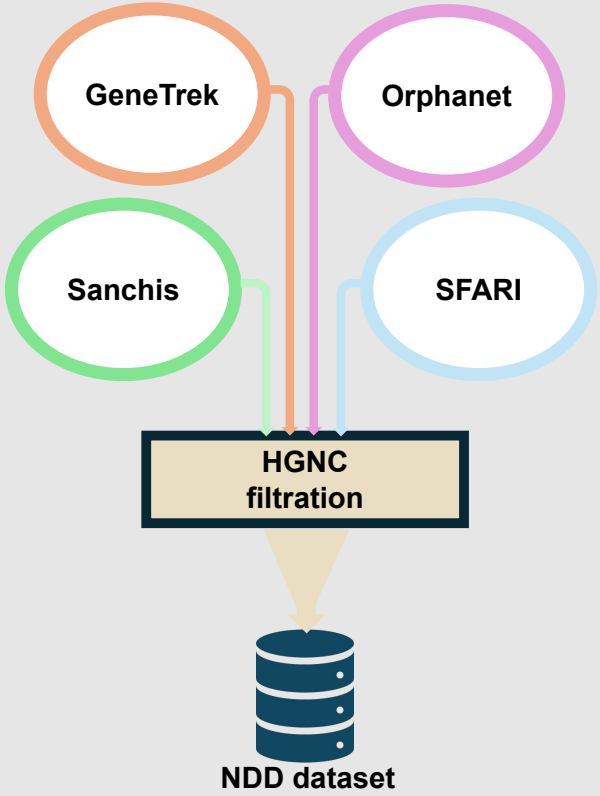
Dataset Construction

Resource Integration

Integrated information from: **Orphanet**, **SFARI**, **GeneTrek**, and **Sanchis et al.** (PMID: 37541188).

Data Refinement

dataset was further refined to homogenize the identifiers exploiting **HGNC** filtration to ensure high-quality, standardized data.



Dataset scoring

Dataset Distribution

SCORE 2
2,639

SCORE 1
6,856

SCORE 0
9,859

TOTAL GENES
19,354

Scoring Methodology

A **score based on the consistency of annotations** across different sources was used to divide genes into sets with varying degrees of association to NDDs.

Negative Set (Set 0)

The **rest of human genes**, not related to NDDs, were categorized as a negative control set under the label **'set 0'**.

Functional enrichment analysis

Program

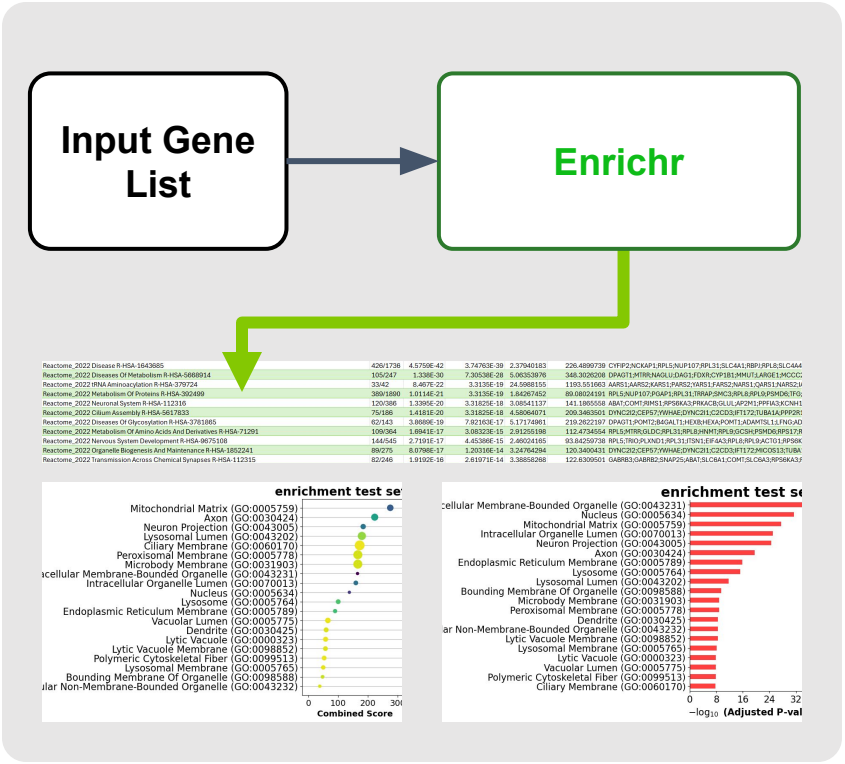
Python package **GSEApy**, a wrapper for **Enrichr**.

Databases

Gene Ontology (GO), **KEGG PATHWAY**, and **Reactome** pathway.

Statistical analysis

p-value computed using **Fisher's exact test** and corrected via **Benjamini-Hochberg (BH)**. Cut-off set at **0.01**.



What the enrichment reveals

GO Biological Process

- tRNA aminoacylation: #1 ranked → RNA-metabolism
- Nervous/Brain system development
- Transcription regulation & RNA biosynthesis

GO Molecular Function

- Binding of DNA, RNA and ATP
- Histone modification (chromatinopathies)
- Ion channel activity: synaptic transmission

GO Cellular Component

- Synaptic and neuronal compartments
- Nuclear / chromatin-associated localizations
- mitochondrial matrix

Pathways (KEGG & Reactome)

- Lysosome & GPI-anchor biosynthesis
- Cancer shared genes
- Metabolism diseases & N-glycosylation

Key Summary

Curated set converges on core NDD biology. Recurrent across ontologies: tRNA biology, cell cycle and cilium terms.

Prioritizing Candidate genes: MOE Score

The Idea

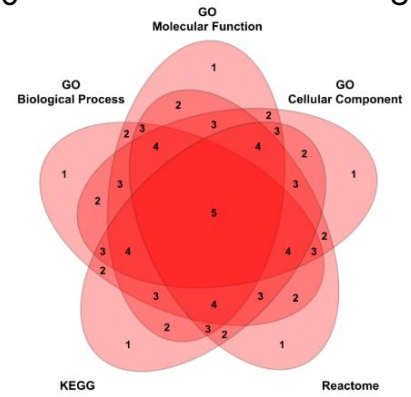
A candidate gene is more likely to be a true NDD gene the more its functional annotation resembles that of the curated set.

MOE = Multi-Ontology Enrichment

How it works (step-by-step)

- Take significantly enriched terms (adj. p ≤ 0.01) in the **curated** set.
- Consider 5 categories: 3 GO domains + KEGG + Reactome.
- Check if candidate gene carries curated-enriched terms.
- **+1 point per category → MOE score 0 to 5.**

MOE	Total Genes
5	534
4	1,097
3	1,827
2	1,568
1	1,032
0	801



MOE score: MONDO disease validation

Does the MOE score actually capture disease relevance? → validation against MONDO

Top MONDO Terms (MOE ≥ 4)

- Development / nervous system disorders
- ***Newly emerging:*** mitochondrial & metabolic disease

$p = 1.1 \times 10^{-17}$ · OR = 1.75 → ~2× more likely, not by chance

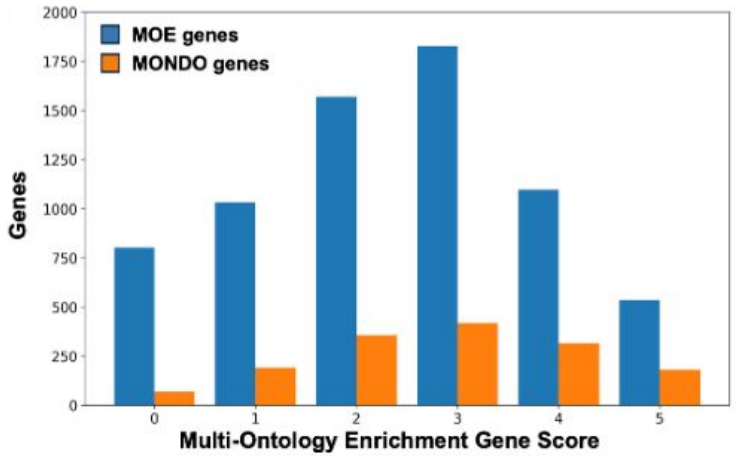
Prevalence & Impact

Higher MOE scores indicate rarer genes (only 534 reach a score of 5).

MONDO-association frequency:

- MOE ≥ 4: >30%
- MOE ≤ 3: <20%

How genes distribute



NDD–Cancer overlap: more than a coincidence

CSB Meeting 2026

Cancer genes overrepresentation

KEGG "Pathways in Cancer" (hsa05200) → 530 genes mapped onto NDD sets.

~2× the background frequency

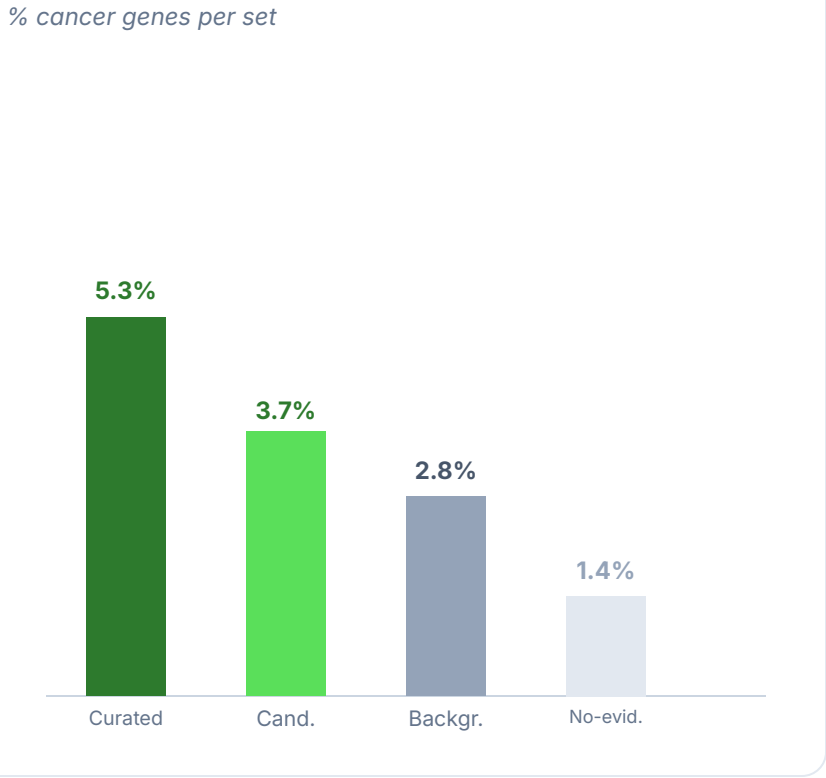
- Curated 5.3% · Candidate 3.7% · Background 2.8%
- No-evidence set → depleted (1.4%)

Shared Pathways

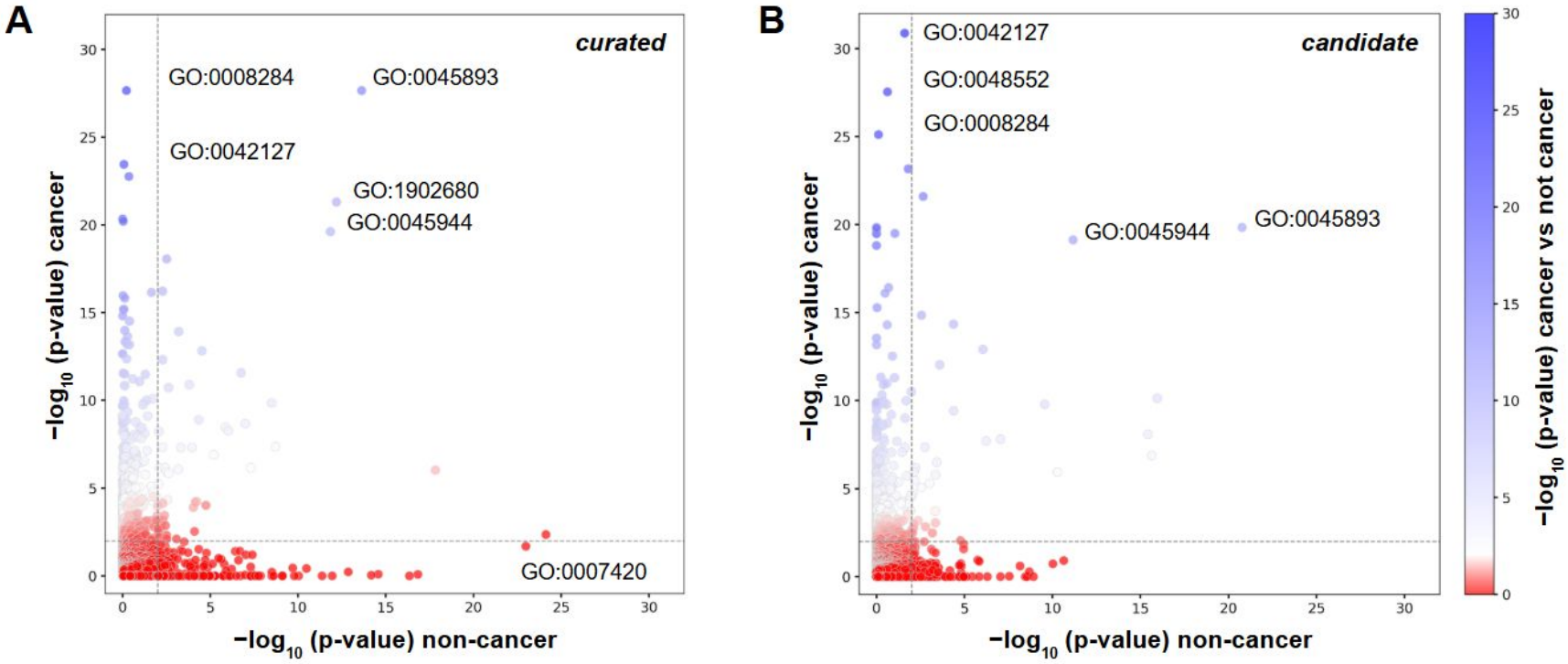
Ras · MAPK · AKT → cell-cycle & proliferation control

Clinical Echo

NDDs (e.g. schizophrenia) → ~50% higher cancer mortality



NDD–Cancer overlap: more than a coincidence



Conclusion and future perspective

Expanded NDD Gene List

Integrating multiple datasets to improve diagnostic accuracy across neurodevelopmental disorders.

Biological Pathways

Identified key associations with cancer, t-RNA, metabolism, and mitochondrial activity.

Candidate prioritization

Prioritized candidate genes as new targets for genetic screening and potential therapies.

Future Perspective

Focus on PPI studies and mapping of pathogenic variants.

Thank you for your attention

Acknowledgement

BioFolD Laboratory Bologna

Emidio Capriotti

Maria Paola Turina

Funding

PNRR-MR1-2022-12376067

Financed by European Union - NextGenerationEU

Annex 1 - NDD association score

Name	Score	Condition
curated	2	$(HC \lor sfari+) \lor (neuro \land develop) \lor [Sanchis \land (neuro \lor develop \lor LC \lor sfari)]$
candidate	1	$(\neg HC \land \neg sfari+) \land (\neg neuro \lor \neg develop) \land [\neg Sanchis \lor (\neg neuro \land \neg develop \land \neg LC \land \neg sfari)] \land [neuro \lor develop \lor Sanchis \lor LC \lor sfari]$
no evidence	0	$\neg HC \land \neg sfari+ \land \neg neuro \land \neg develop \land \neg Sanchis \land \neg LC \land \neg sfari$