HIVE: a novel tool for Horizontal multi-omics Integration analysis using Variational AutoEncoders

Wassila Kathir[1,2*], Giulia Calia[1*], Anjana Bhat[1], Mame Seynabou Fall[1], Jeromine Carret[2], Barbara Bardoni[2], Carole Gwizdek[2], Silvia Bottini[1]

[1] Université Côte d'Azur, INRAE, ISA
[2] Université Côte d'Azur, CNRS, IPMC
* Equal co-authors

Silvia Bottini, PhD
Junior professor chair INRAE

Statomique

Lille, 21st November 2023

The history of FMRP protein, a Multifunctional RNA-binding Protein

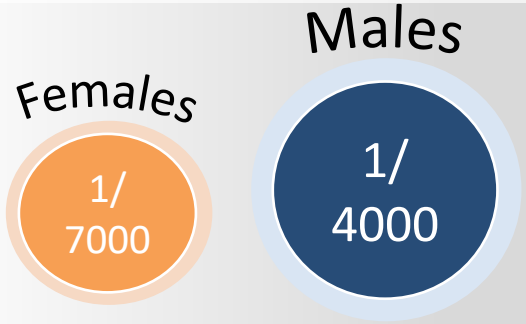
FMRP and the fragile X syndrome

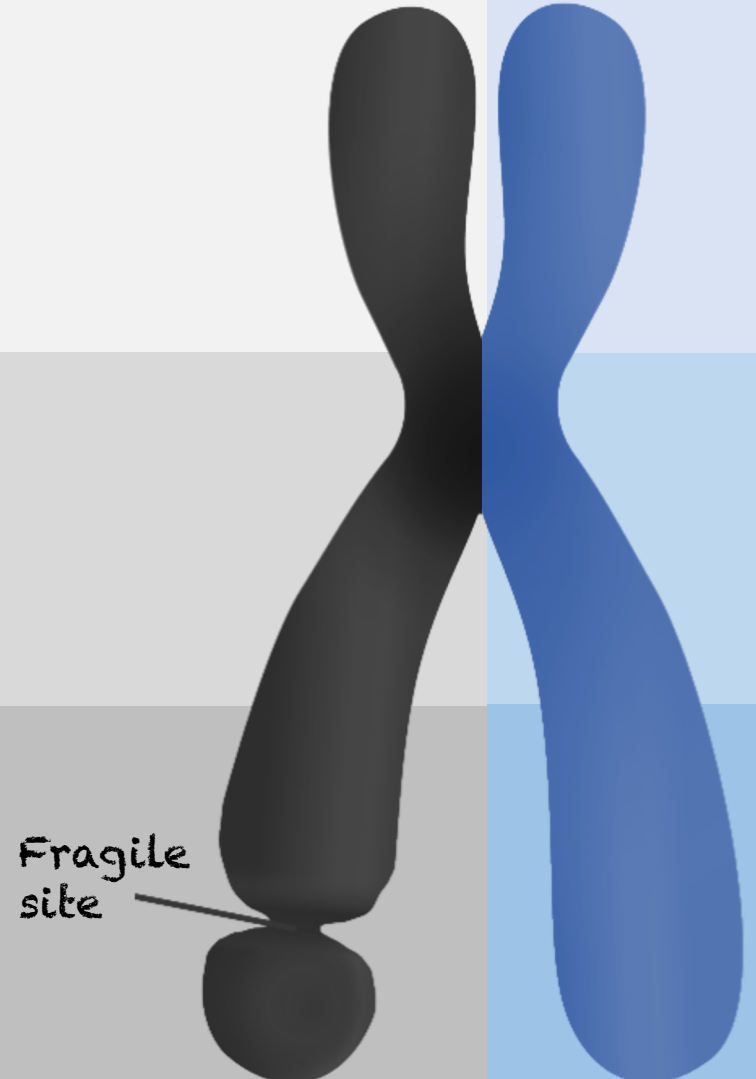
Intellectual disabilities

Physical anomalies

Behavioural characteristics

35% Autism Spectrum Disorder

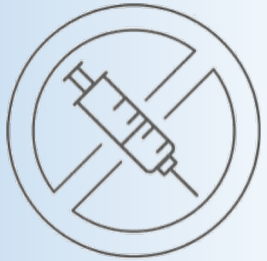




FMRP: Fragile X Messenger Ribonucleoprotein 1

NO treatment

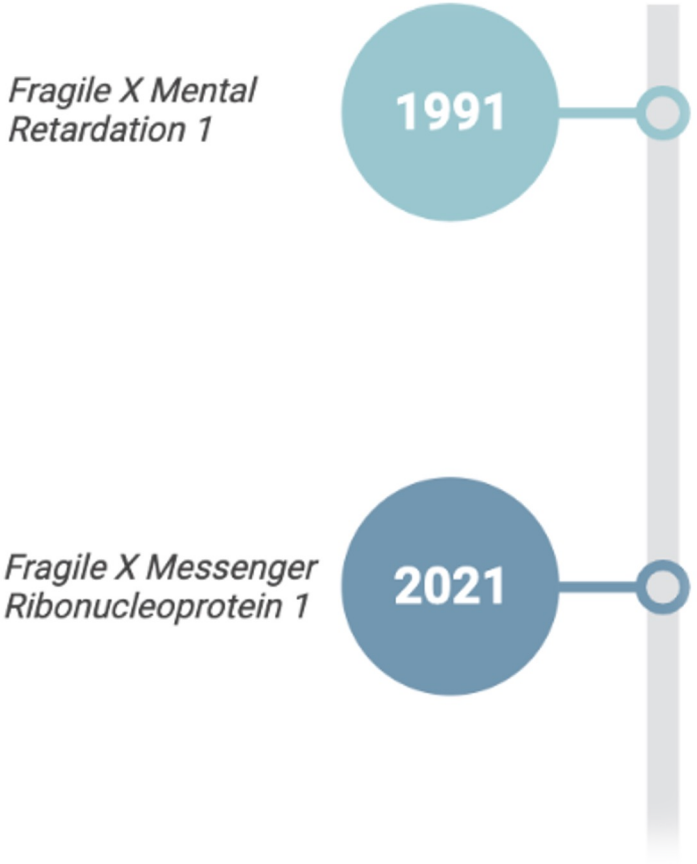
Educational therapies



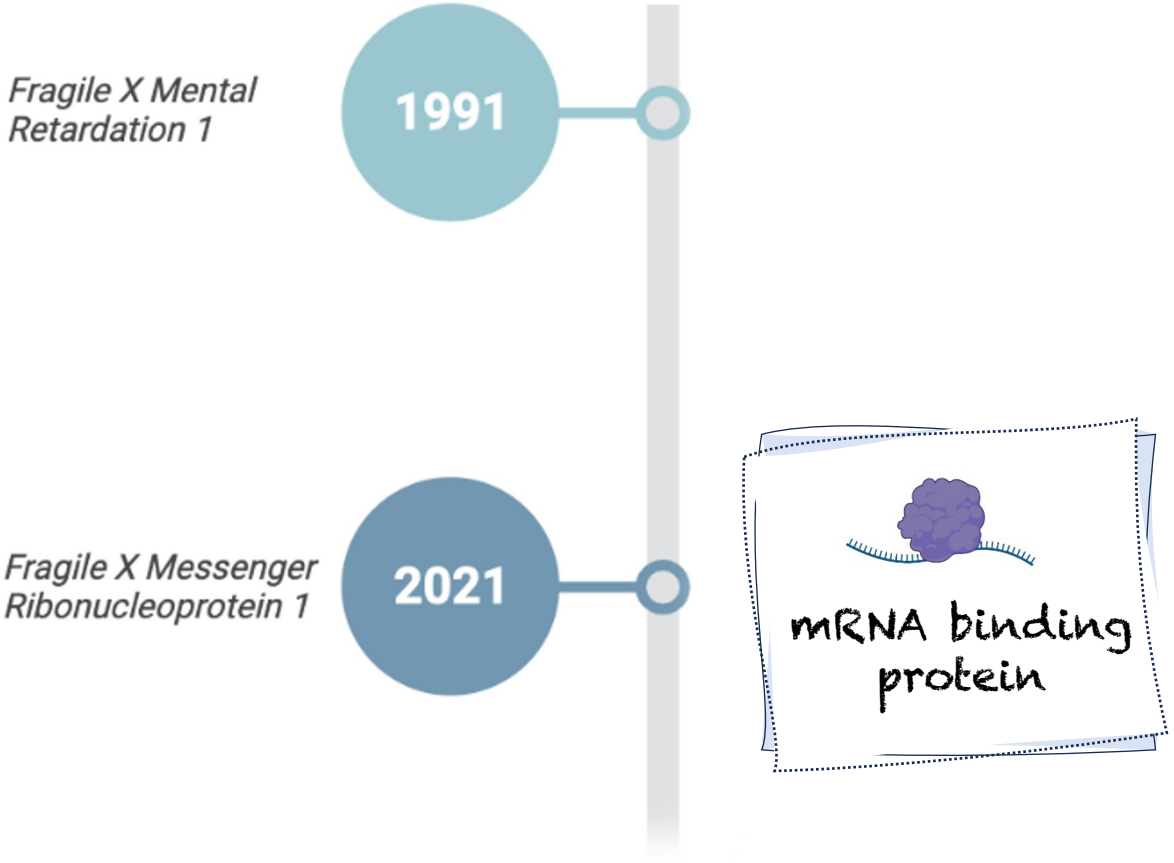
A **mutation** in the ***FMR-1*** **gene** yielding the **absence of FMRP protein**

The optimal intervention would be to **restore the function of the FMRP protein.**

The history of FMRP protein, a Multifunctional RNA-binding Protein



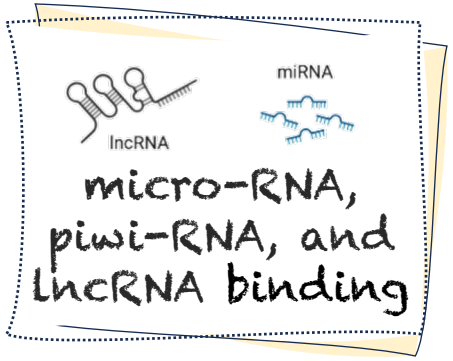
The history of FMRP protein, a Multifunctional RNA-binding Protein



The history of FMRP protein, a Multifunctional RNA-binding Protein

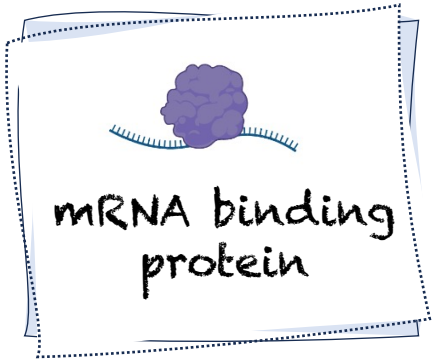
Fragile X Mental Retardation 1

1991



Fragile X Messenger Ribonucleoprotein 1

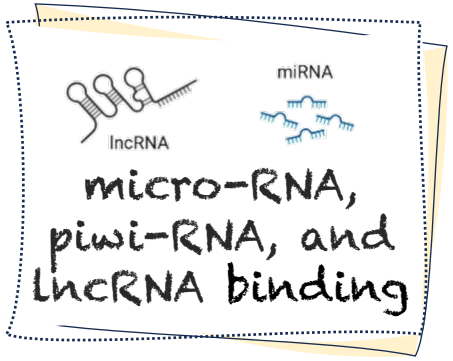
2021

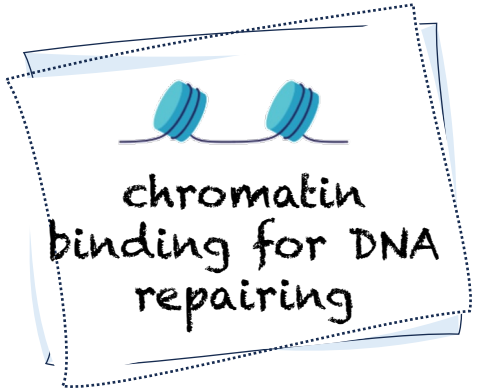


The history of FMRP protein, a Multifunctional RNA-binding Protein

Fragile X Mental Retardation 1

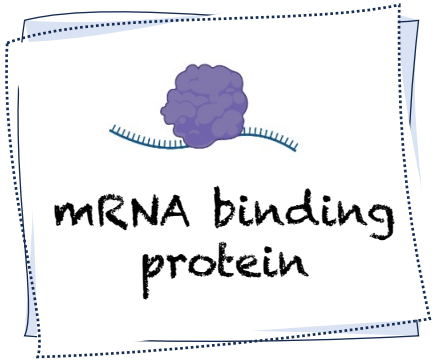
1991





Fragile X Messenger Ribonucleoprotein 1

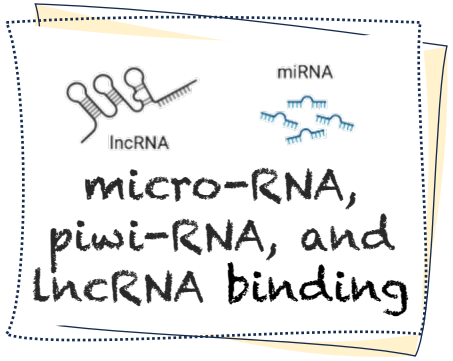
2021

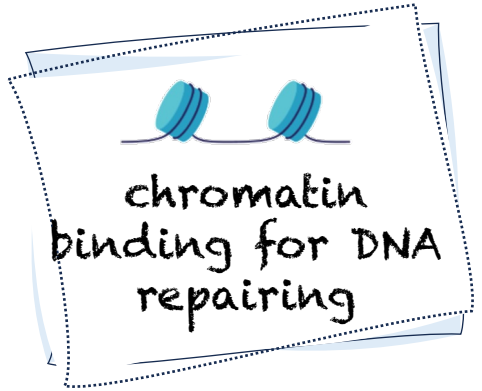


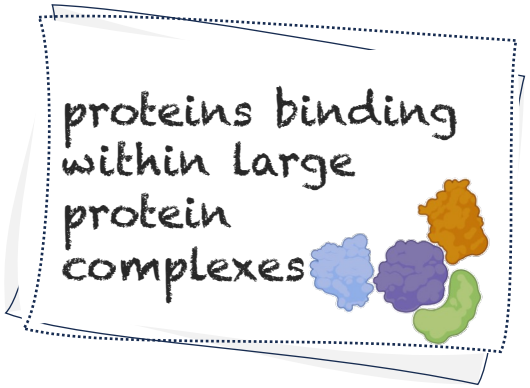
The history of FMRP protein, a Multifunctional RNA-binding Protein

Fragile X Mental Retardation 1

1991

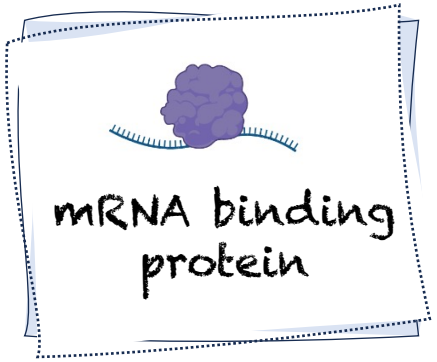






Fragile X Messenger Ribonucleoprotein 1

2021



The history of FMRP protein, a Multifunctional RNA-binding Protein

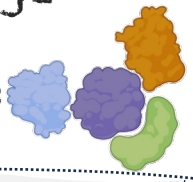
Fragile X Mental Retardation 1

1991

micro-RNA, piwi-RNA, and lncRNA binding

chromatin binding for DNA repairing

proteins binding within large protein complexes



Fragile X Messenger Ribonucleoprotein 1

2021

mRNA binding protein

critical component of neuronal granules

The history of FMRP protein, a Multifunctional RNA-binding Protein

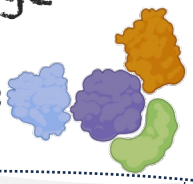
Fragile X Mental Retardation 1

1991

micro-RNA, piwi-RNA, and lncRNA binding

chromatin binding for DNA repairing

proteins binding within large protein complexes



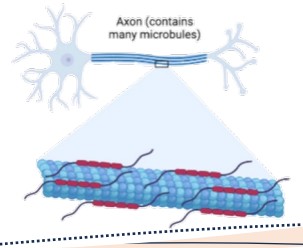
Fragile X Messenger Ribonucleoprotein 1

2021

mRNA binding protein

critical component of neuronal granules

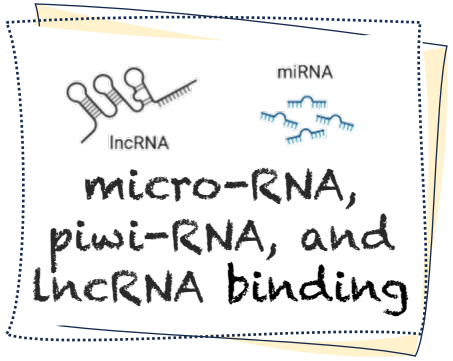
bridge between these granules and microtubule elements

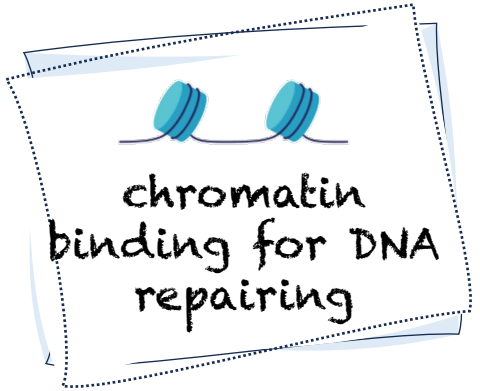


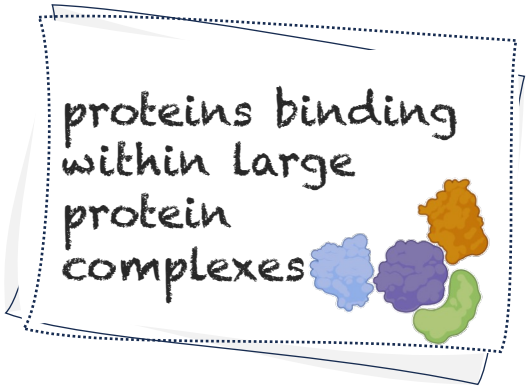
The history of FMRP protein, a Multifunctional RNA-binding Protein

Fragile X Mental Retardation 1

1991

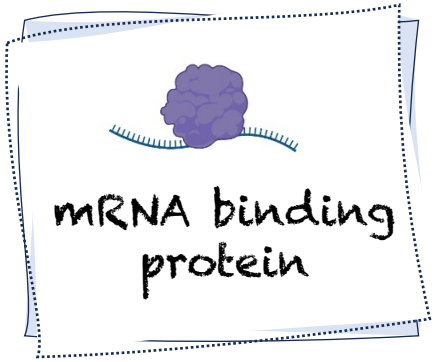




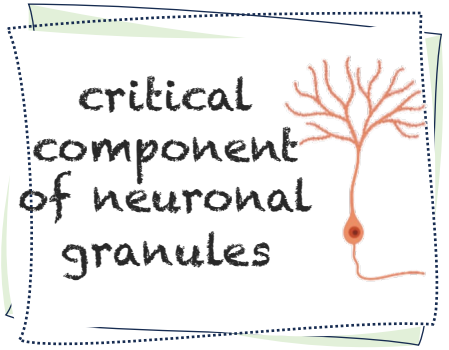


Fragile X Messenger Ribonucleoprotein 1

2021



More than 30 functional categories



The history of FMRP protein, a Multifunctional RNA-binding Protein

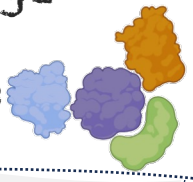
Fragile X Mental Retardation 1

1991

micro-RNA, piwi-RNA, and lncRNA binding

chromatin binding for DNA repairing

proteins binding within large protein complexes



Fragile X Messenger Ribonucleoprotein 1

2021

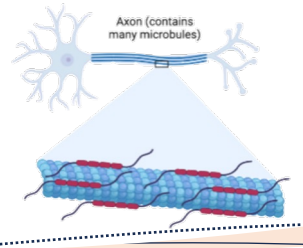
mRNA binding protein

More than 30 functional categories

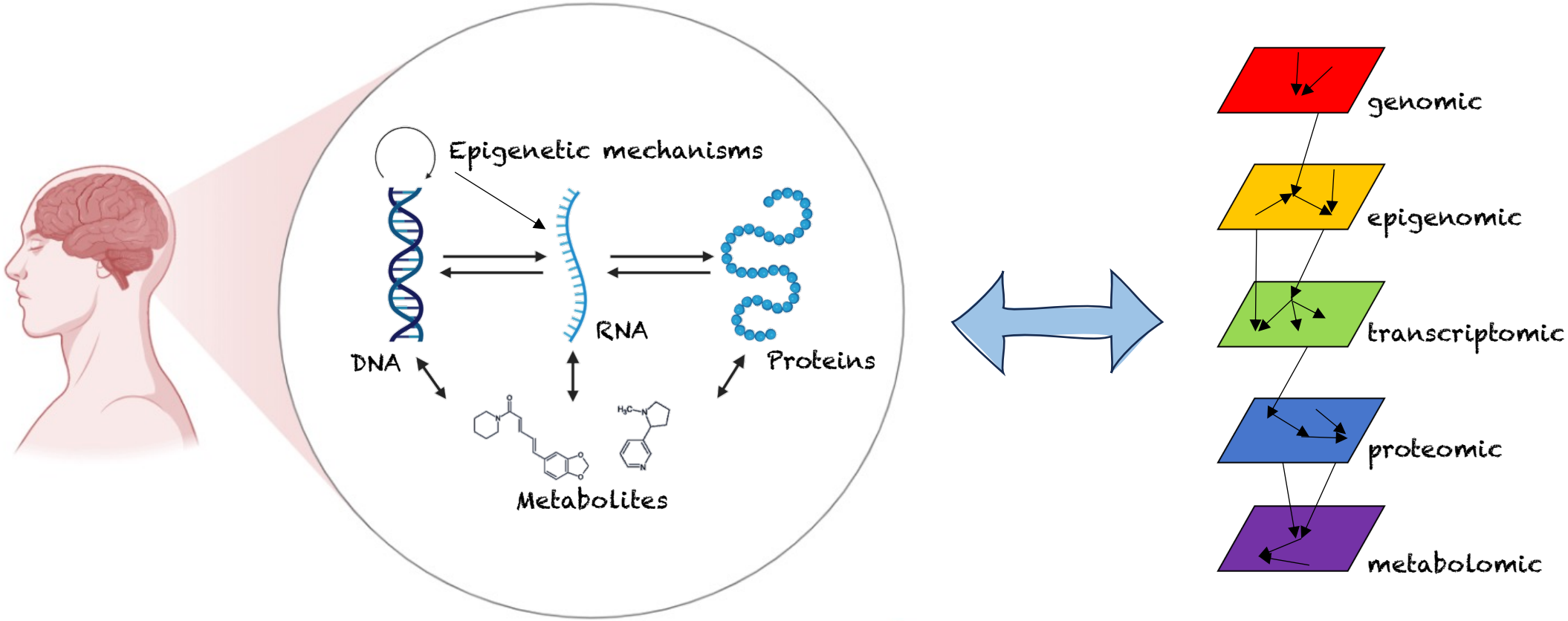
critical component of neuronal granules

Is that all?

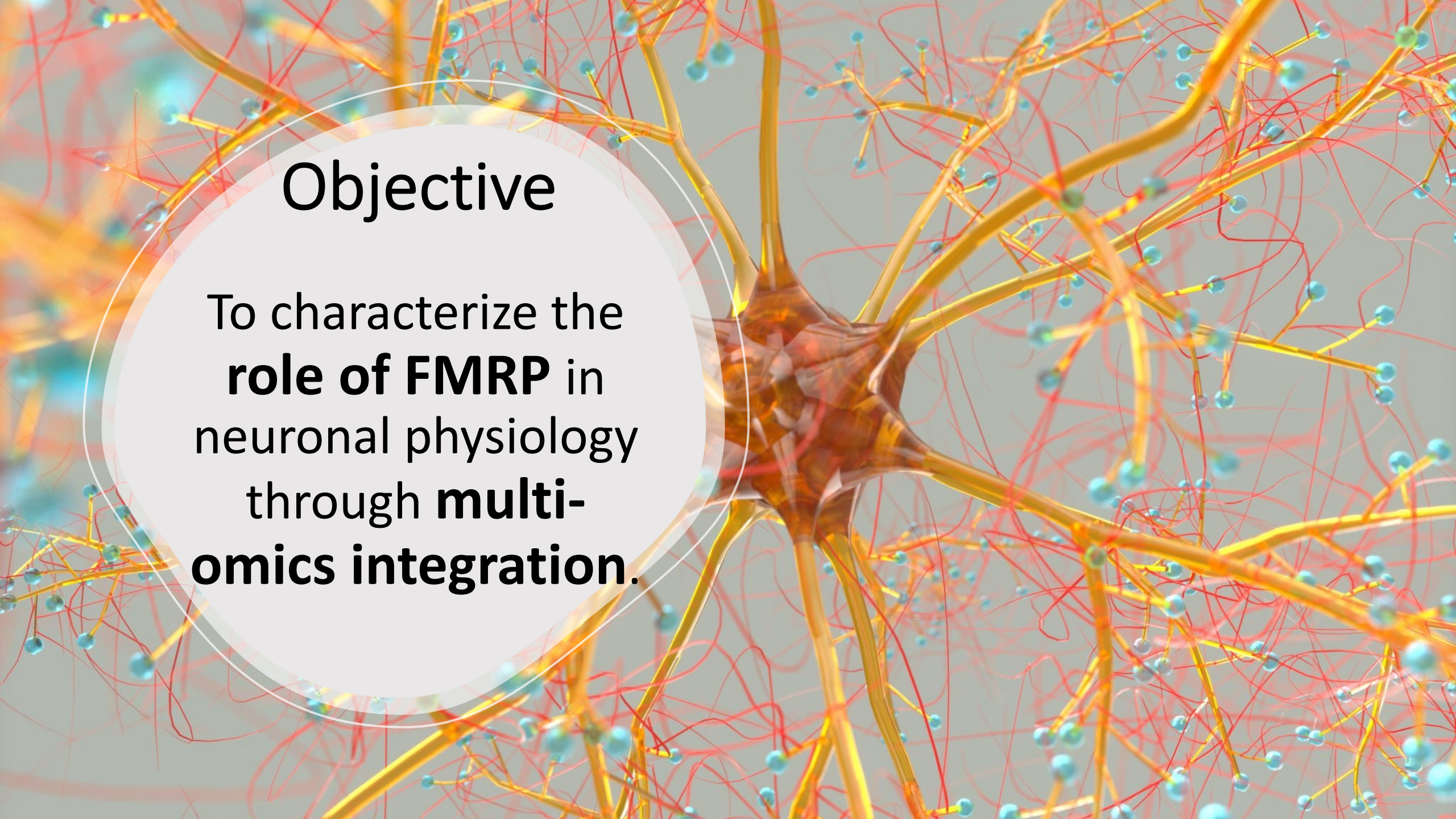
bridge between these granules and microtubule elements



How to characterize the multiple roles of FMRP?



Understanding the multiple roles of FMRP by using multi-omics integration.



Objective

To characterize the **role of FMRP** in neuronal physiology through **multi-omics integration**.

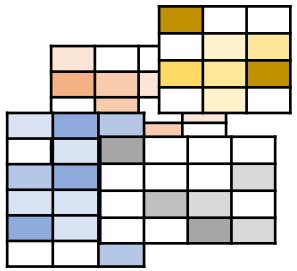
Data collection

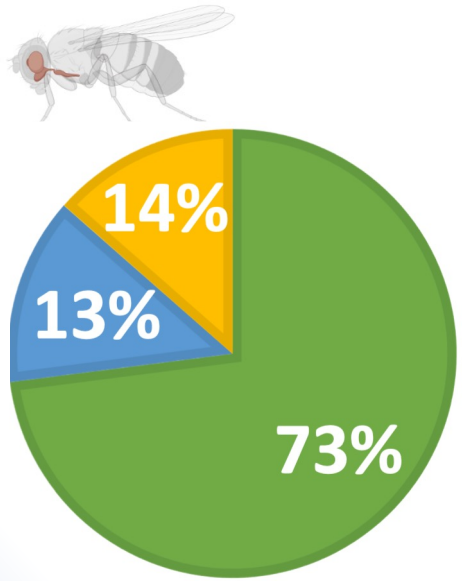
Experimental design

FMRP WT

FMRP KO

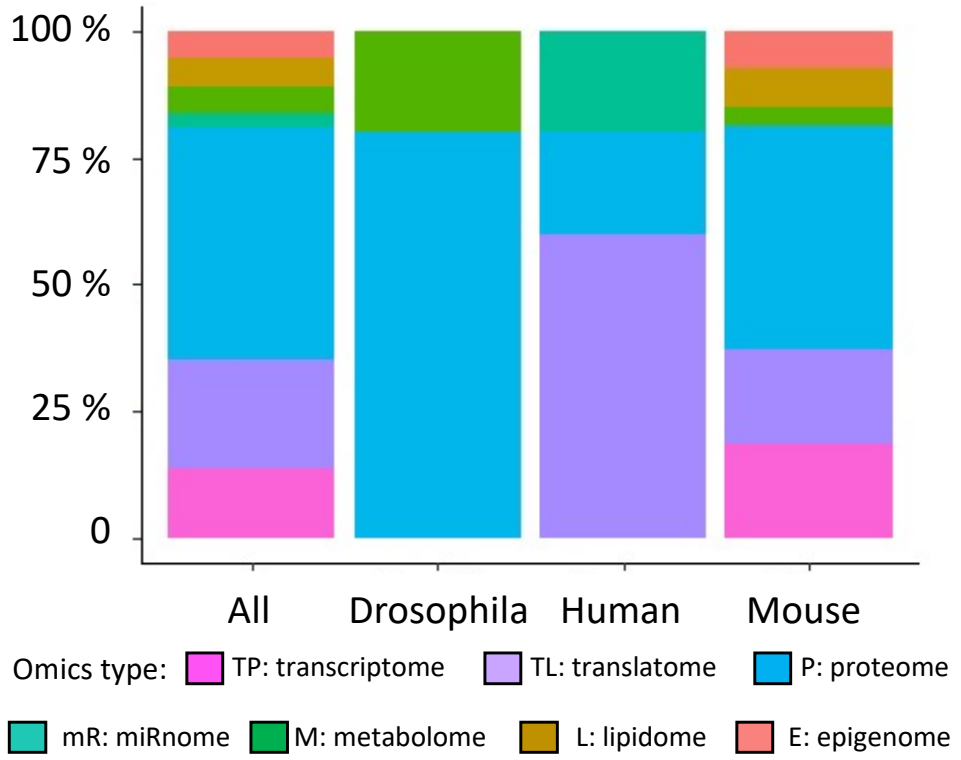
Public databases



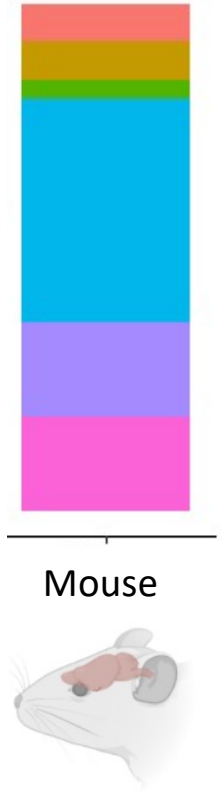


43 Multi-omics datasets

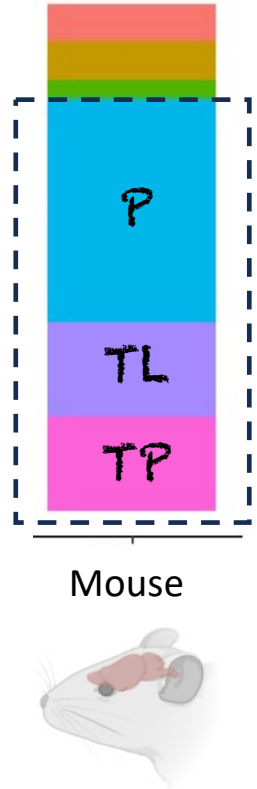
7 different type of omics



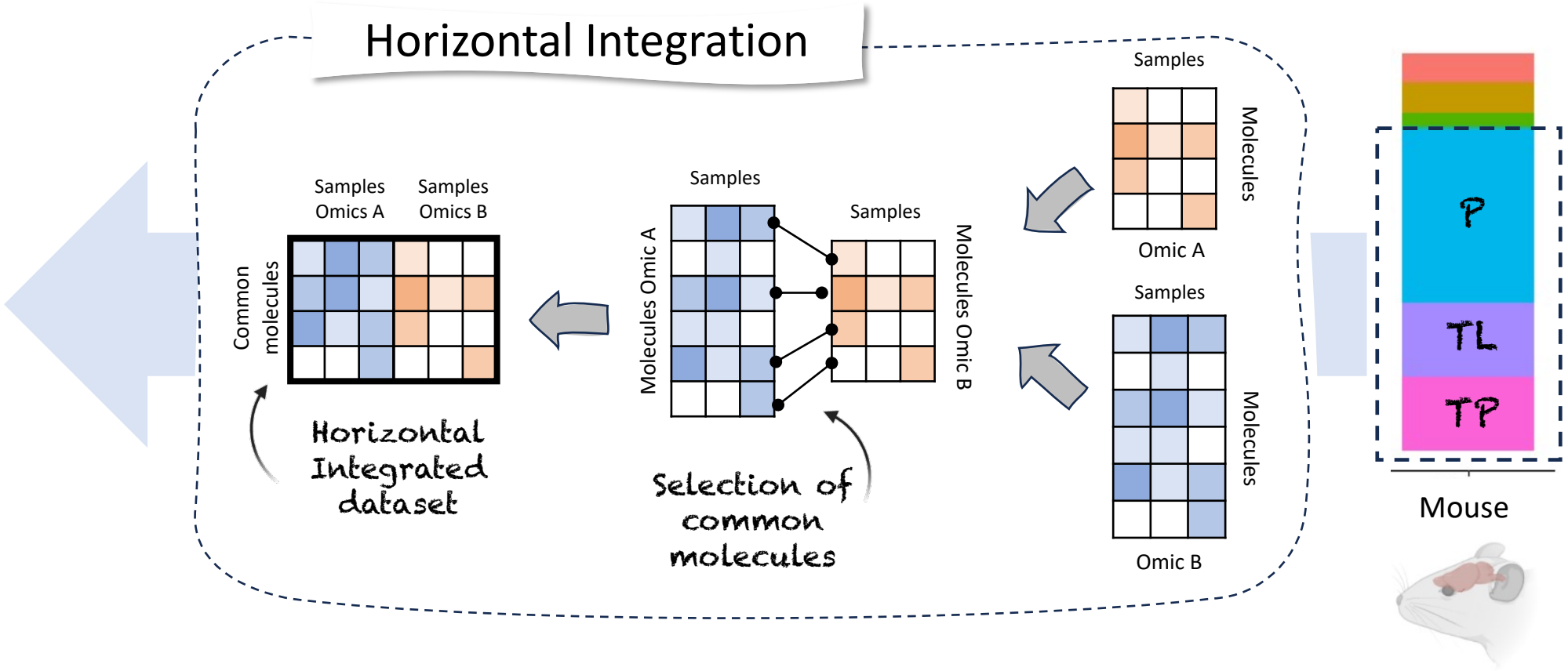
The pilot integrated dataset



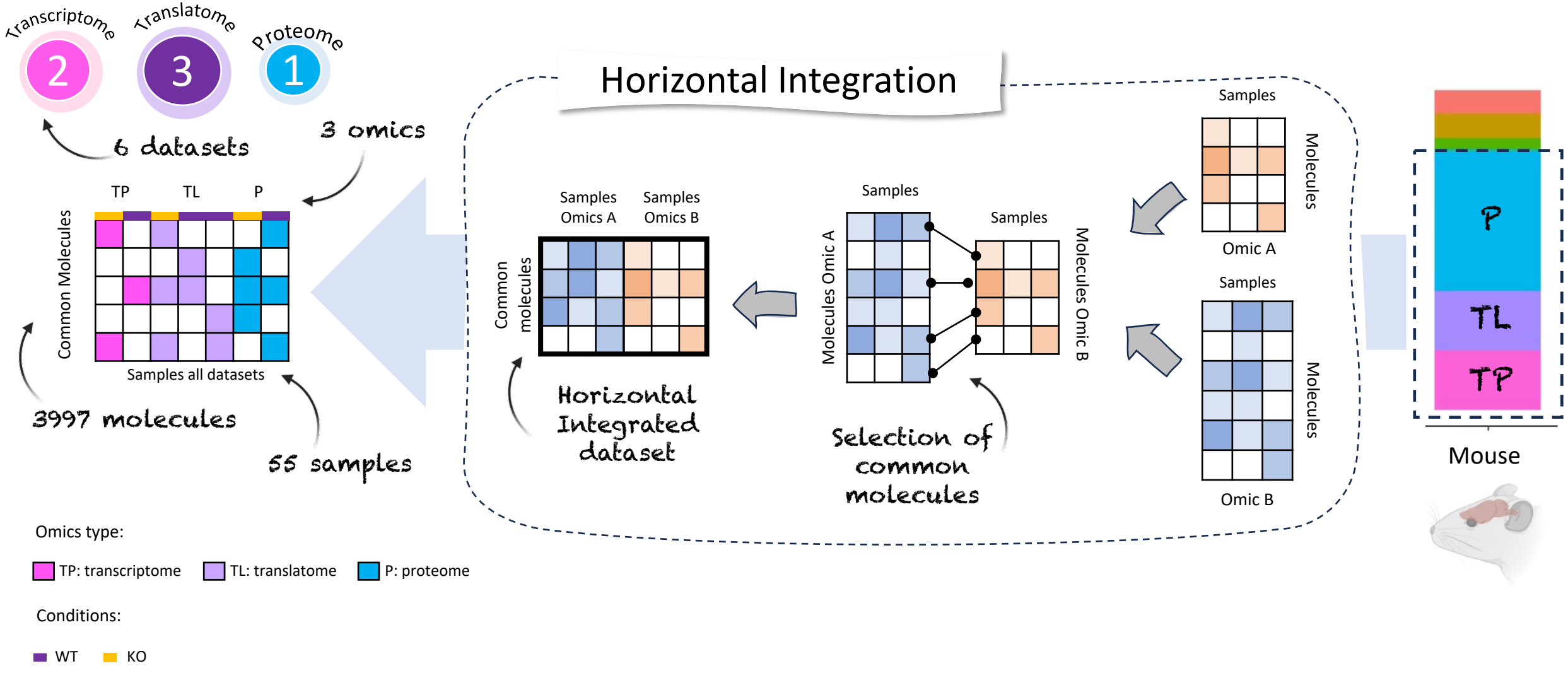
The pilot integrated dataset



The pilot integrated dataset



The pilot integrated dataset

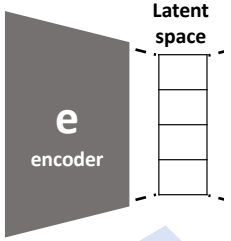


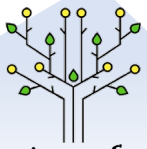
HIVE: a general framework to analyse integrated multi-omics data

Secret code
=
Multi-omics data

HIVE

Variational Autoencoder



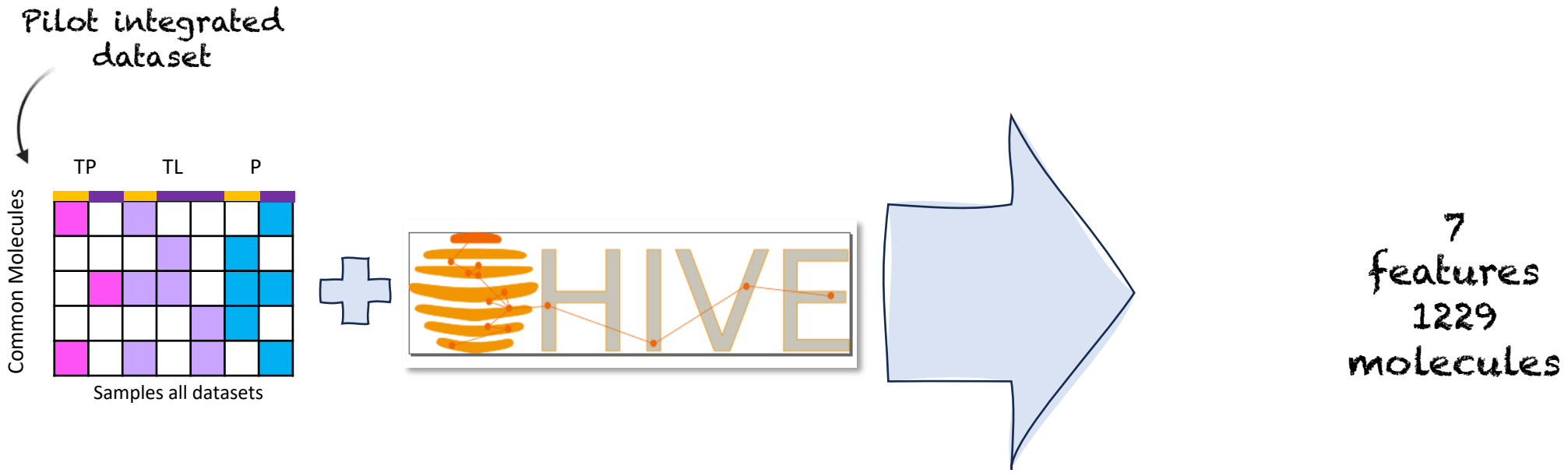


Random forest regression

Molecules contribution to each latent feature
=
The key to convert the secret code

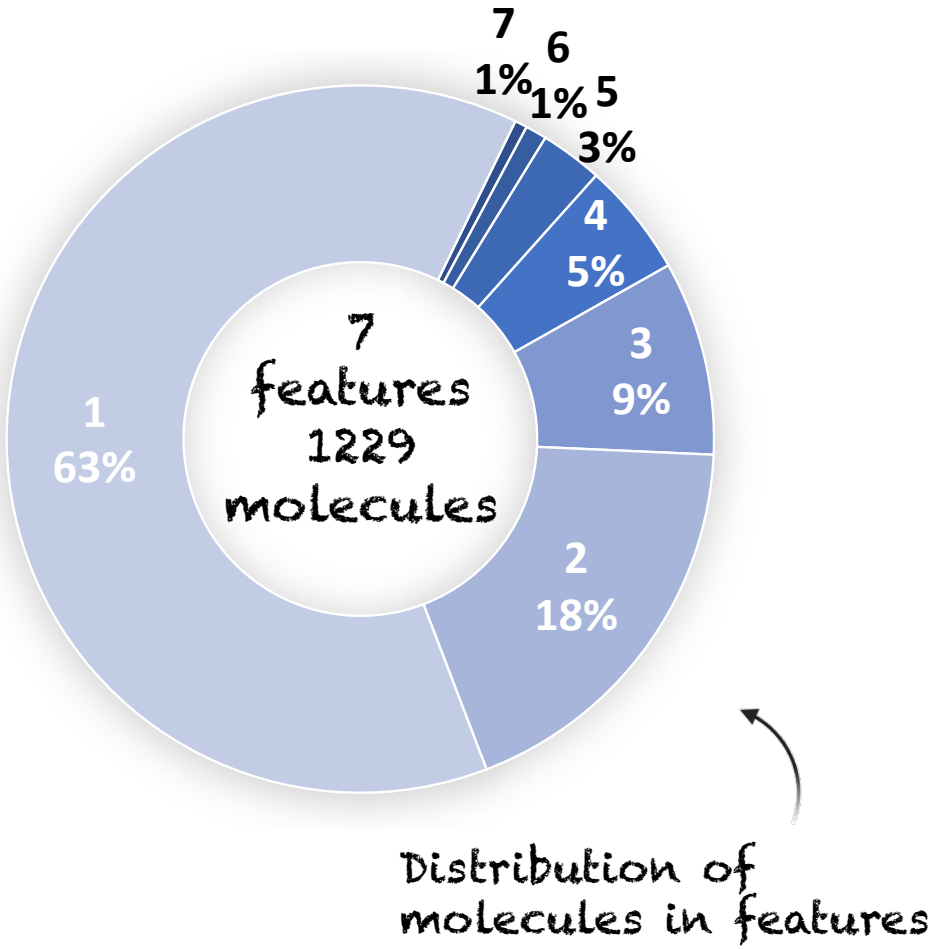
Decrypted code
=
Latent features

Application of HIVE on the pilot integrated dataset



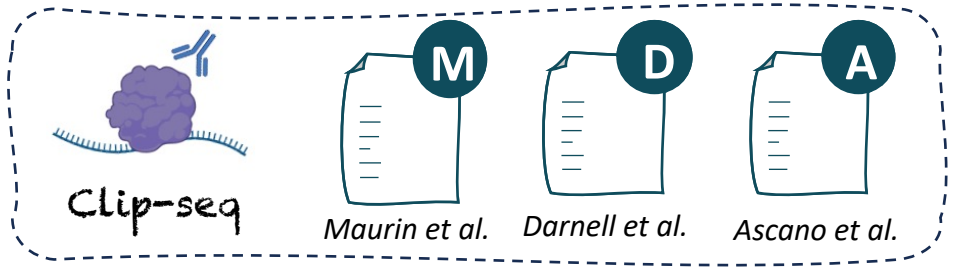
Each feature identifies a specific set of molecules that might correspond to different biological meanings.

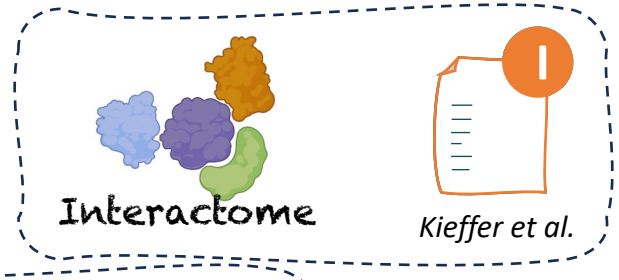
Application of HIVE on the pilot integrated dataset

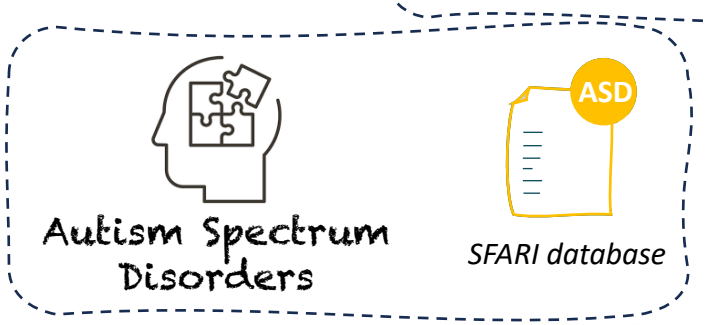


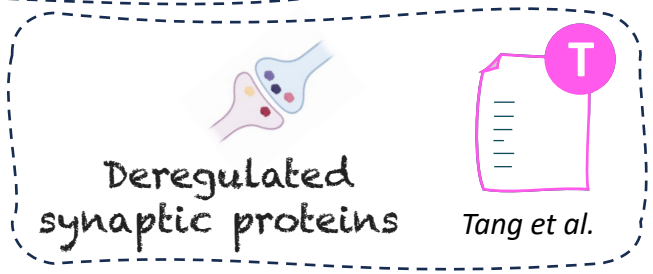
Each feature identifies a specific set of molecules that might correspond to different biological meanings.

Characterization of the molecules found by HIVE

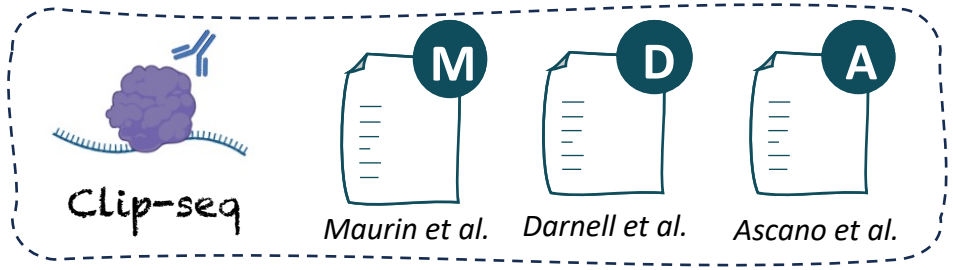


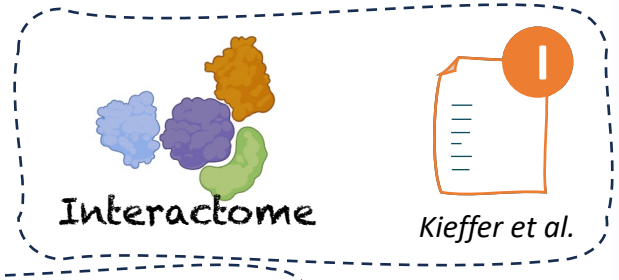


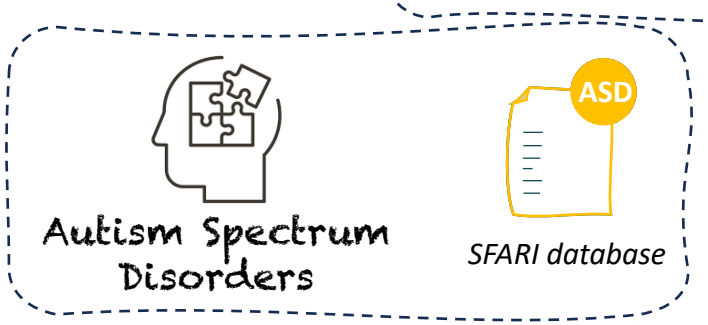


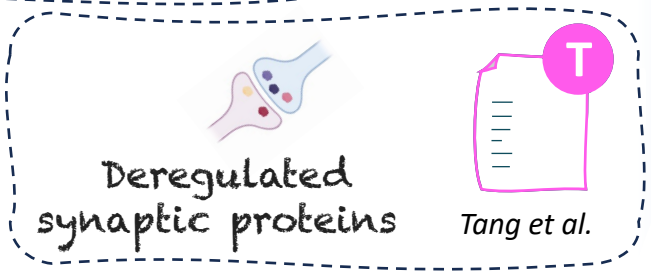


Characterization of the molecules found by HIVE



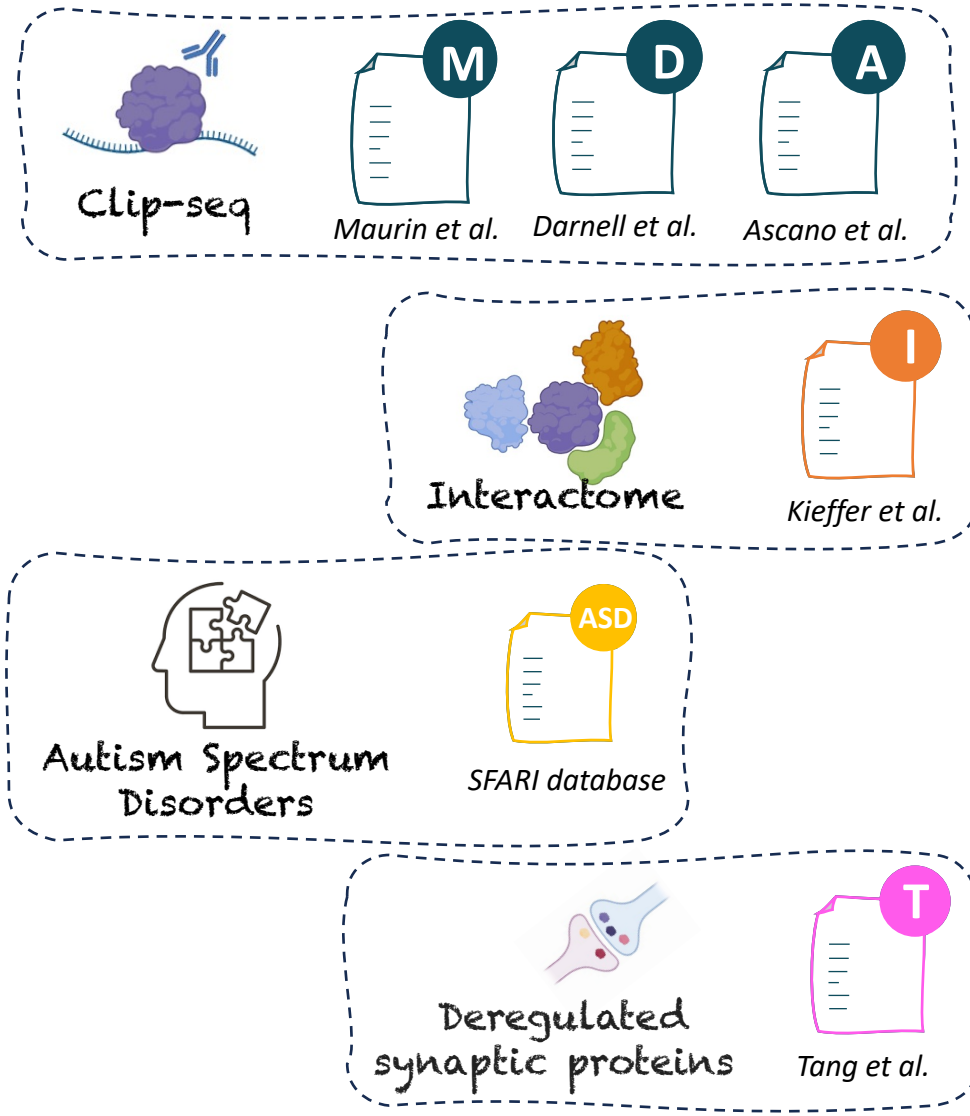


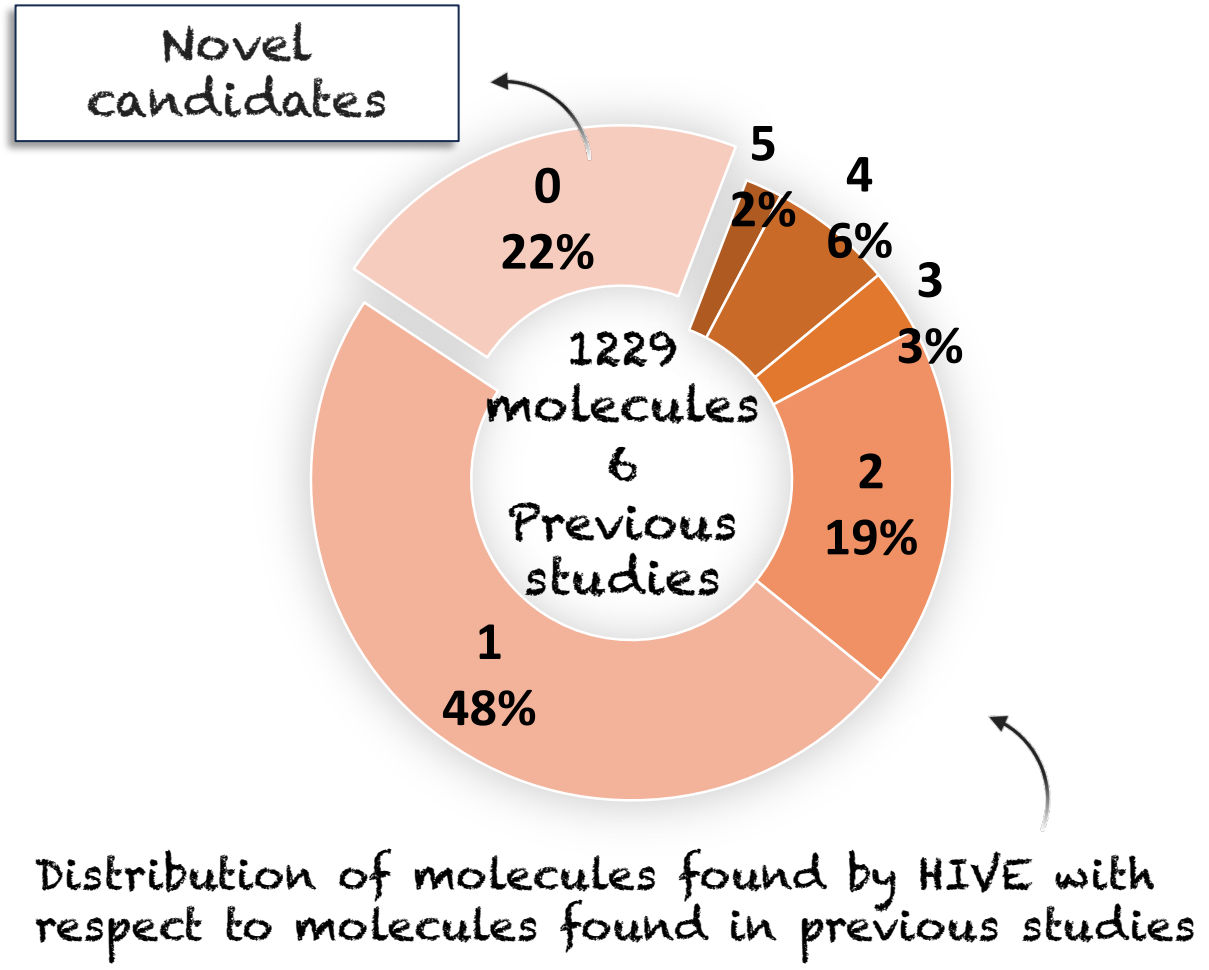




1229 molecules

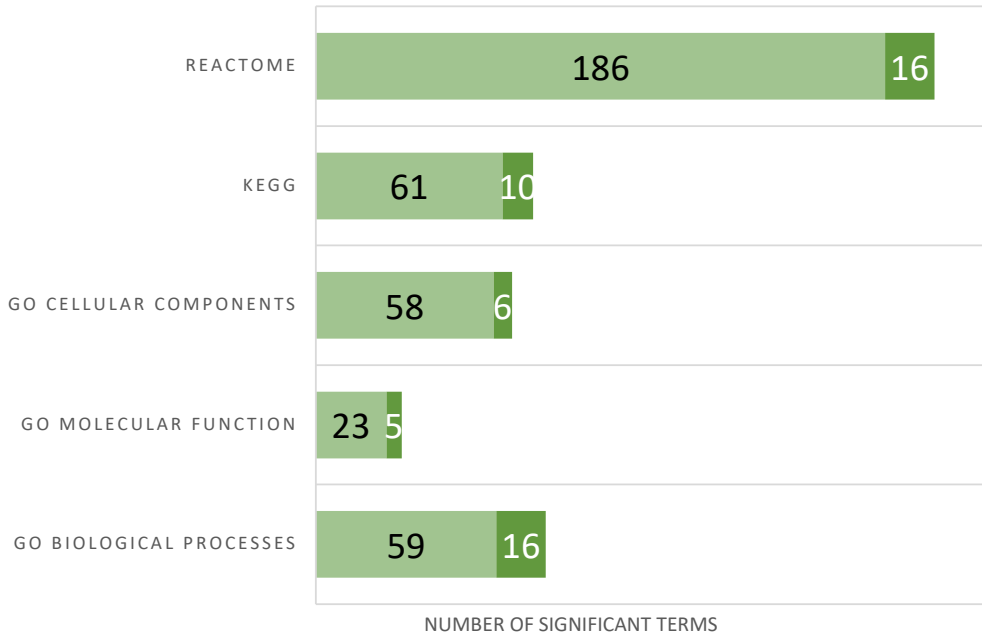
Characterization of the molecules found by HIVE





Distribution of molecules found by HIVE with respect to molecules found in previous studies

Global characterization of functional processes



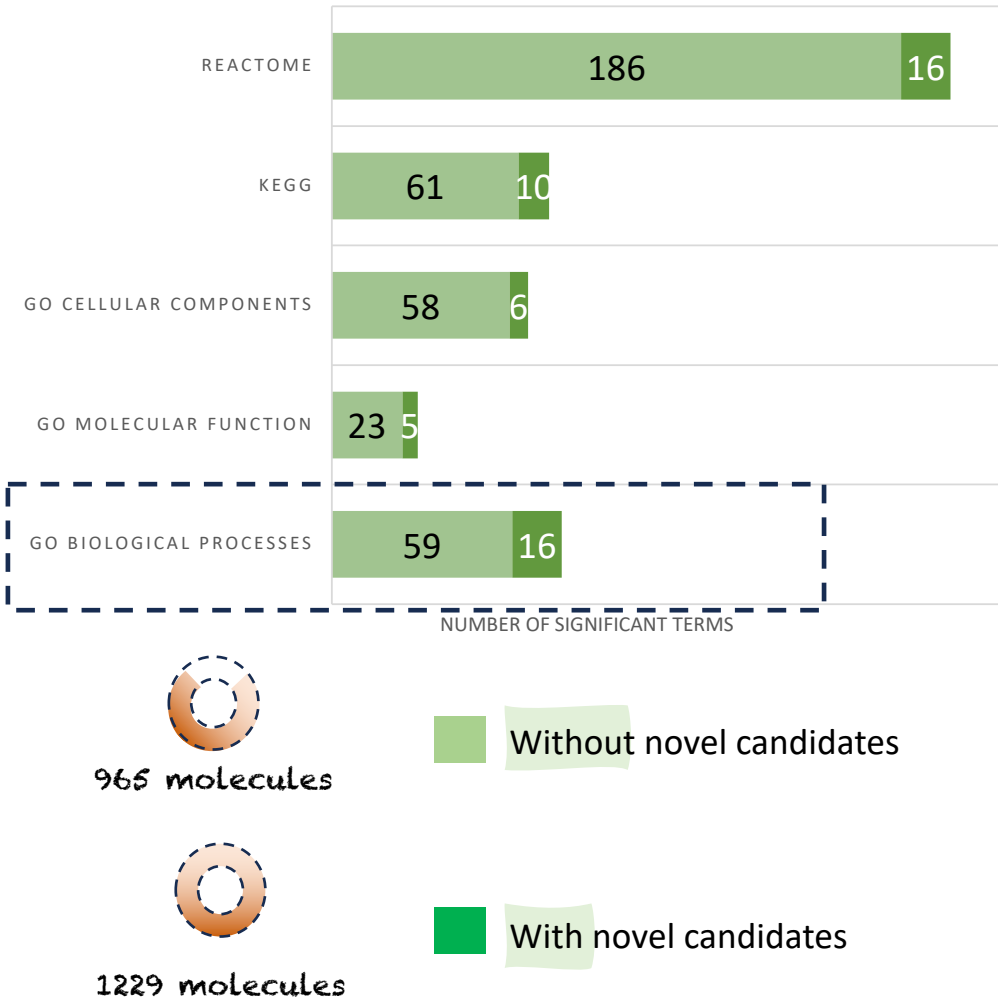
965 molecules

Without novel candidates

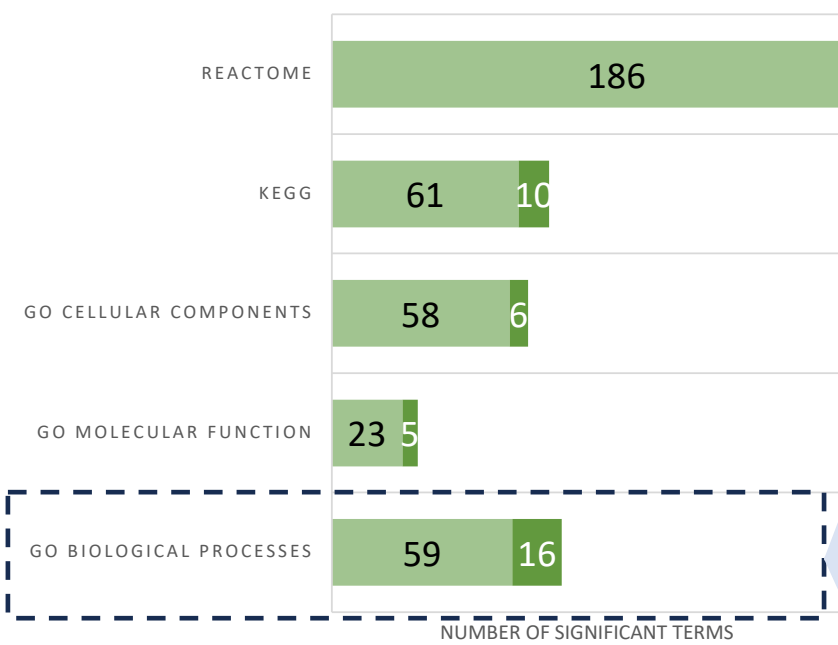
1229 molecules

With novel candidates

Global characterization of functional processes



Global characterization of functional processes



965 molecules

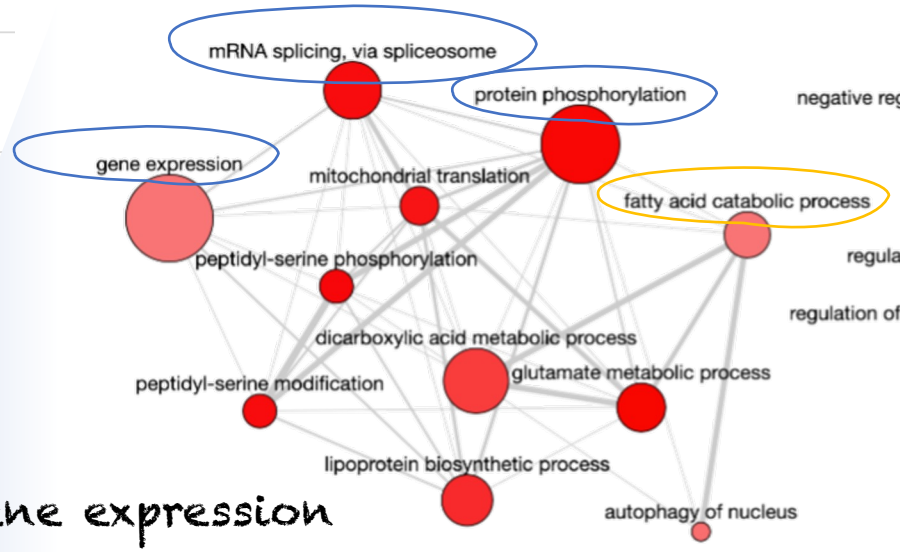
Without novel candidates

1229 molecules

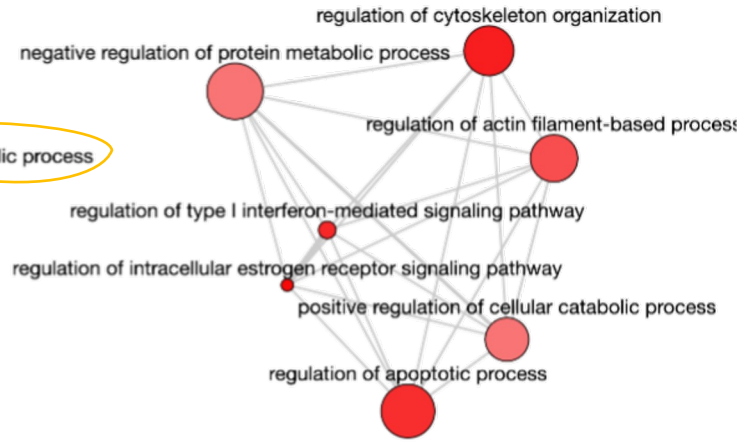
With novel candidates

Known term

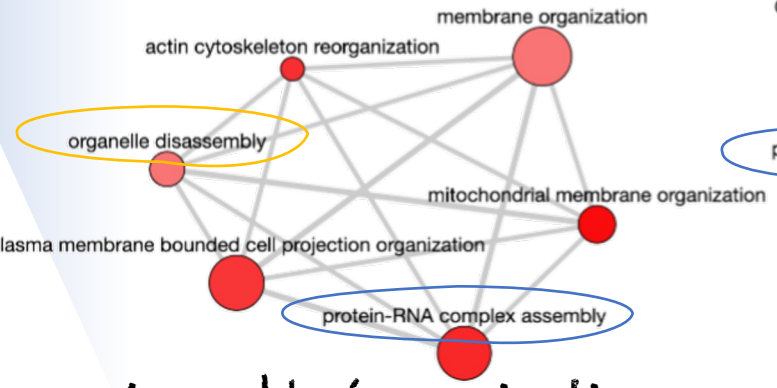
Term found with novel candidtes



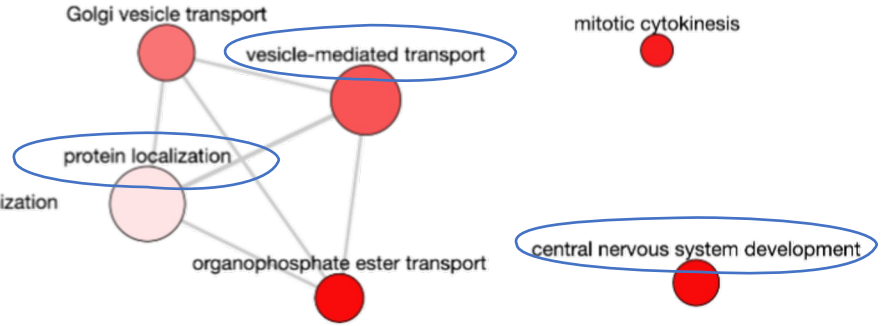
Gene expression



Regulatory mechanisms

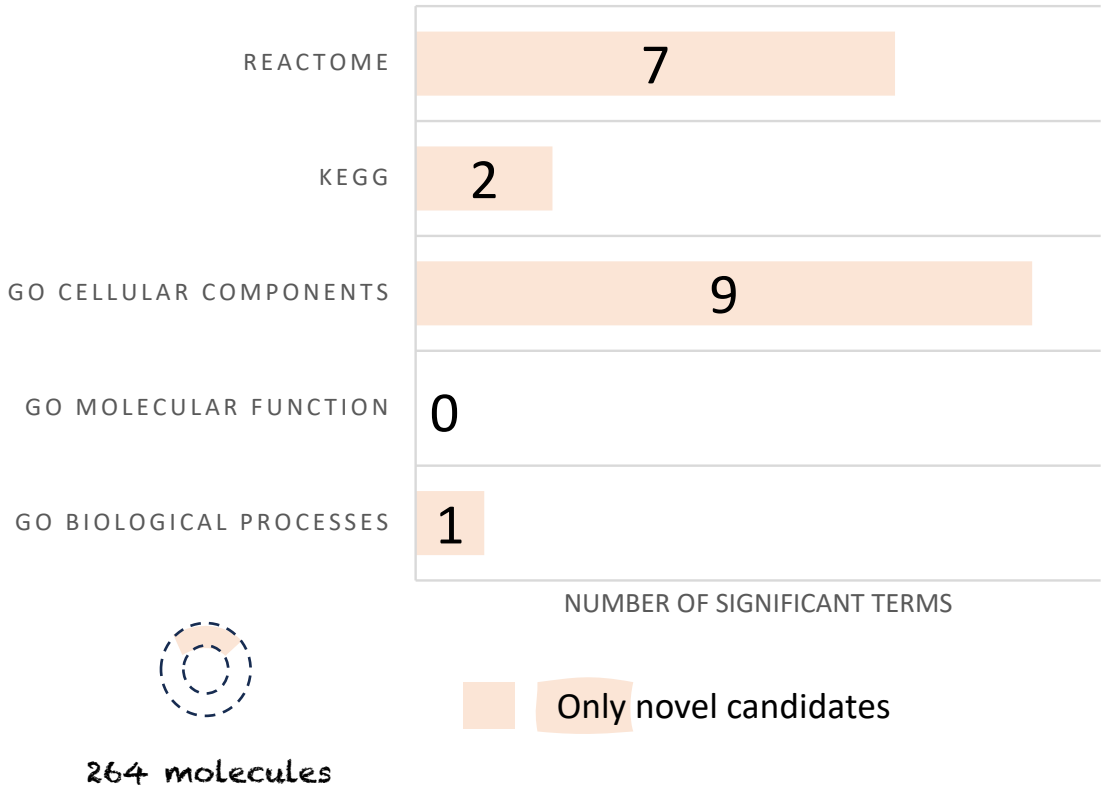


Assembly/organization



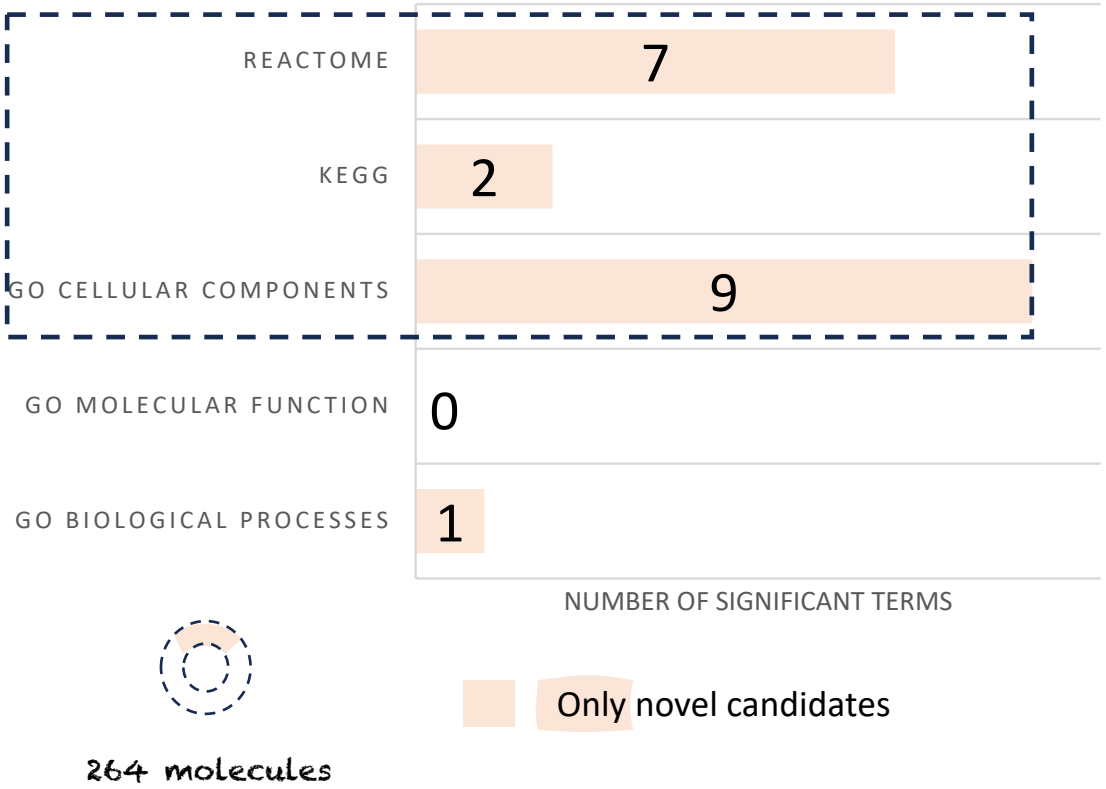
Transport

Global characterization of functional processes

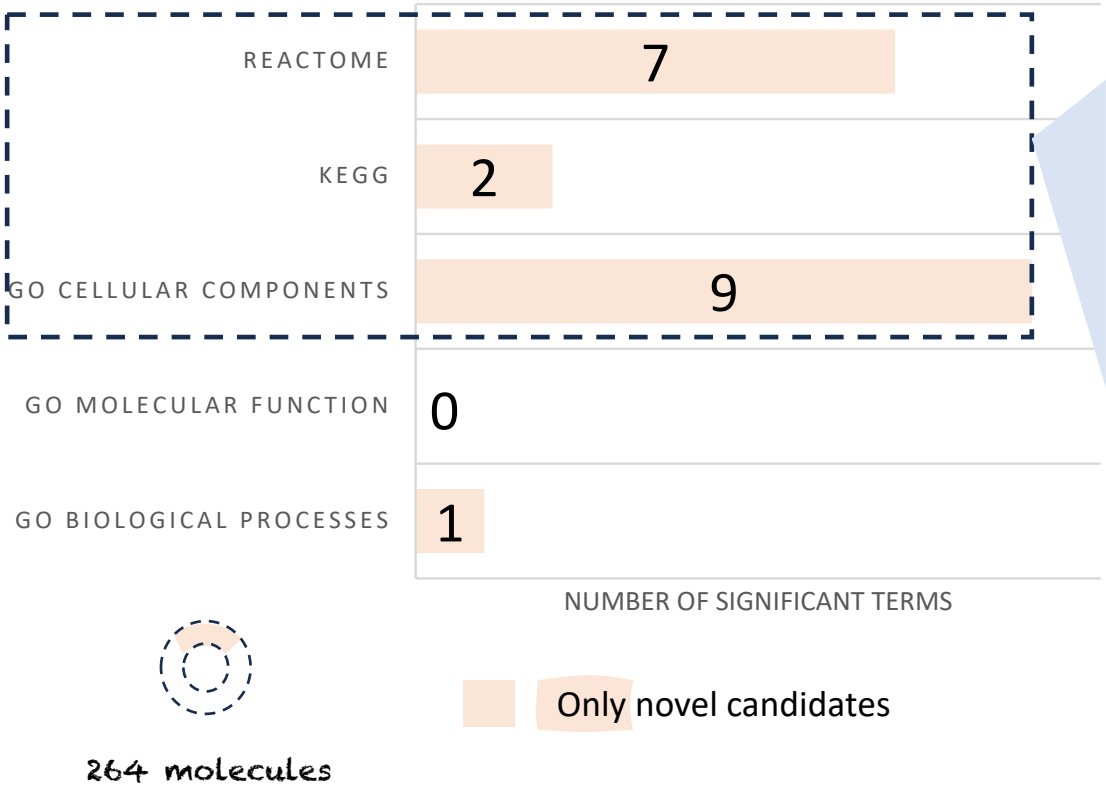


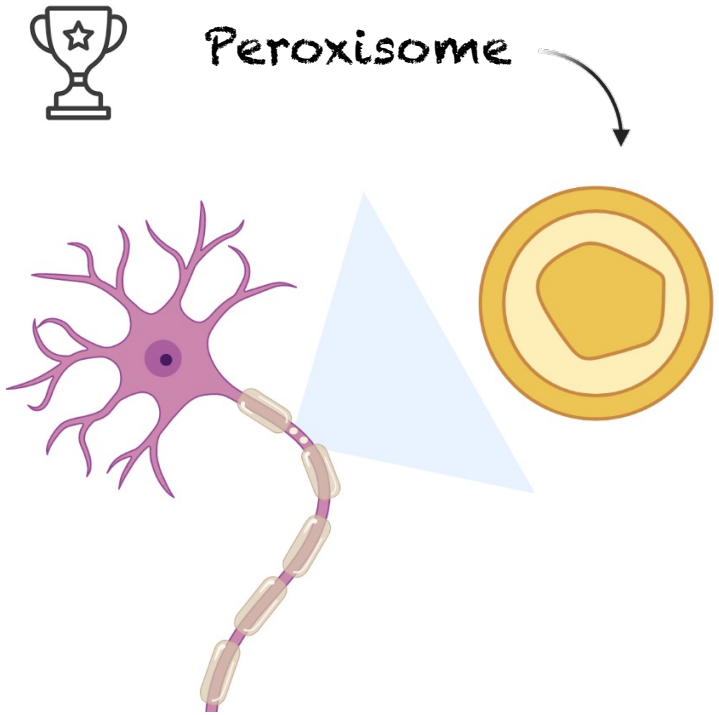
264 molecules

Global characterization of functional processes

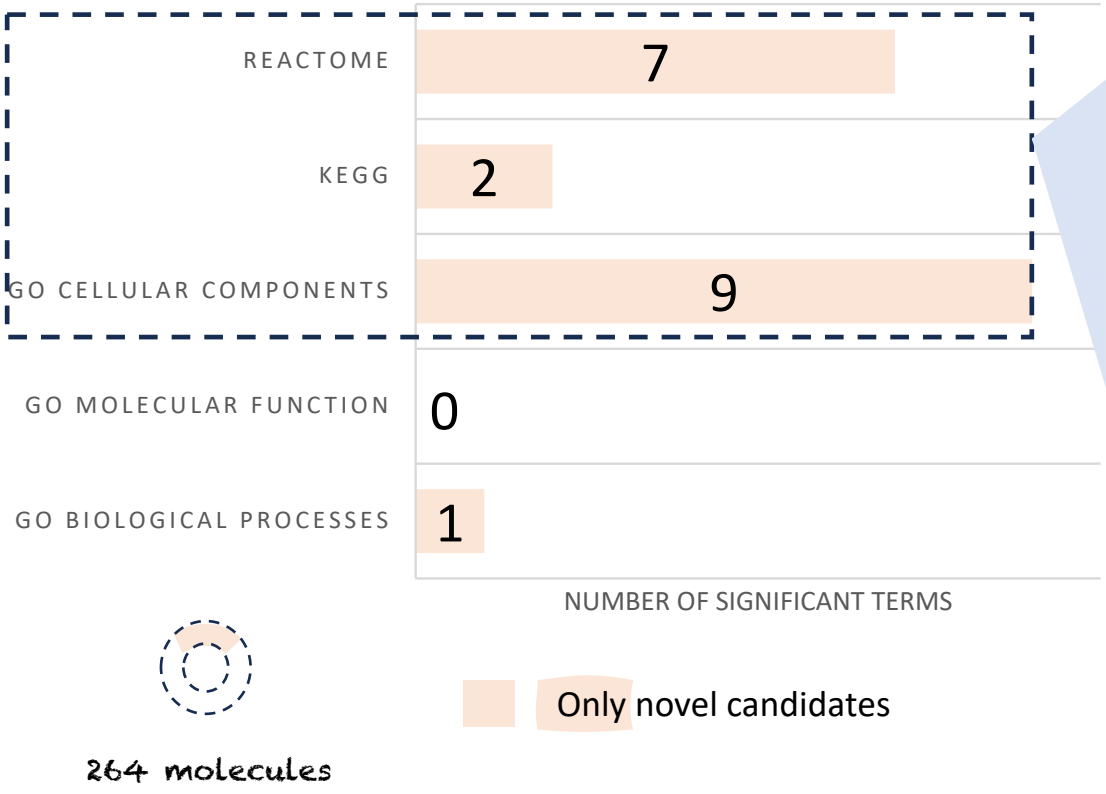


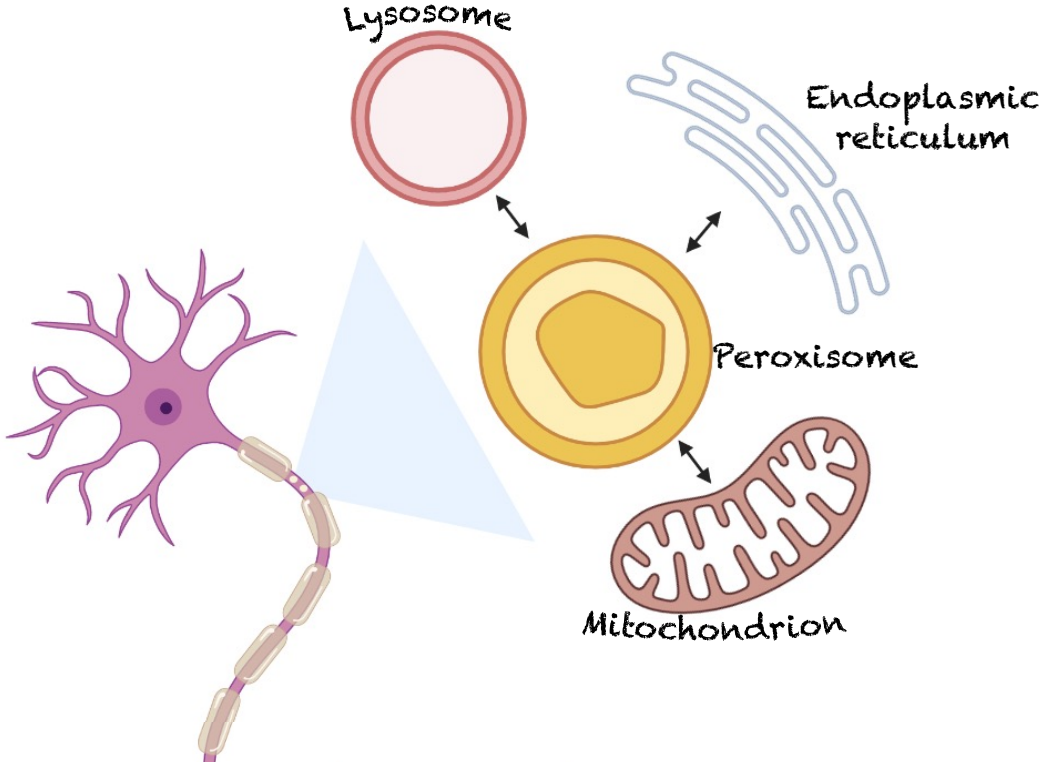
Global characterization of functional processes





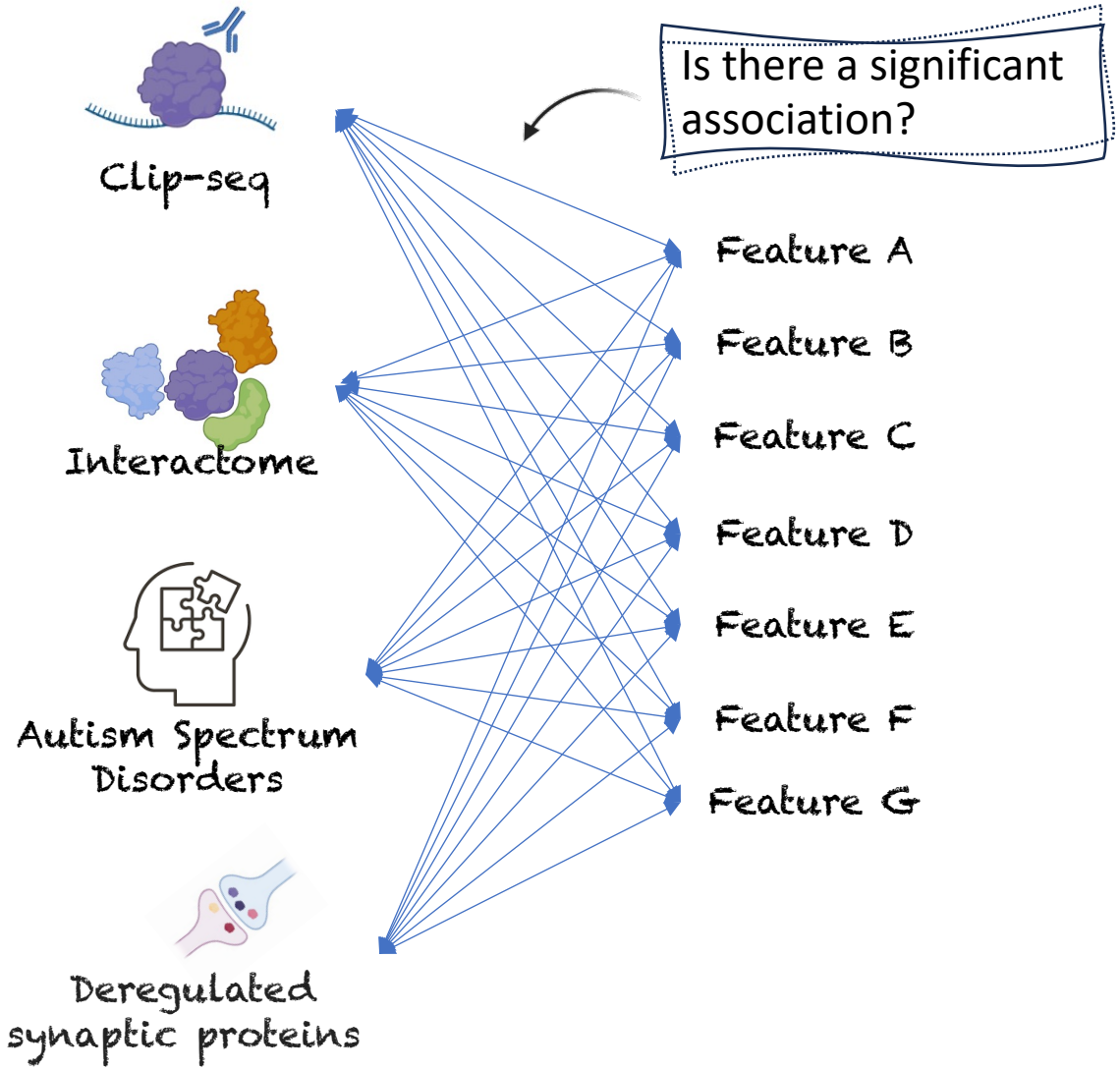
Global characterization of functional processes





Peroxisomes play critical roles in maintaining cellular health, especially in the central nervous system.

Characterization of the features found by HIVE



Characterization of the features found by HIVE

Clip-seq

Interactome

Autism Spectrum Disorders

Deregulated synaptic proteins

Is there a significant association?

Feature A

Feature B

Feature C

Feature D

Feature E

Feature F

Feature G

Market basket analysis

Significant associations between items that are often bought together.

Characterization of the features found by HIVE

Clip-seq

Interactome

Autism Spectrum Disorders

Deregulated synaptic proteins

Is there a significant association?

Feature A

Feature B

Feature C

Feature D

Feature E

Feature F

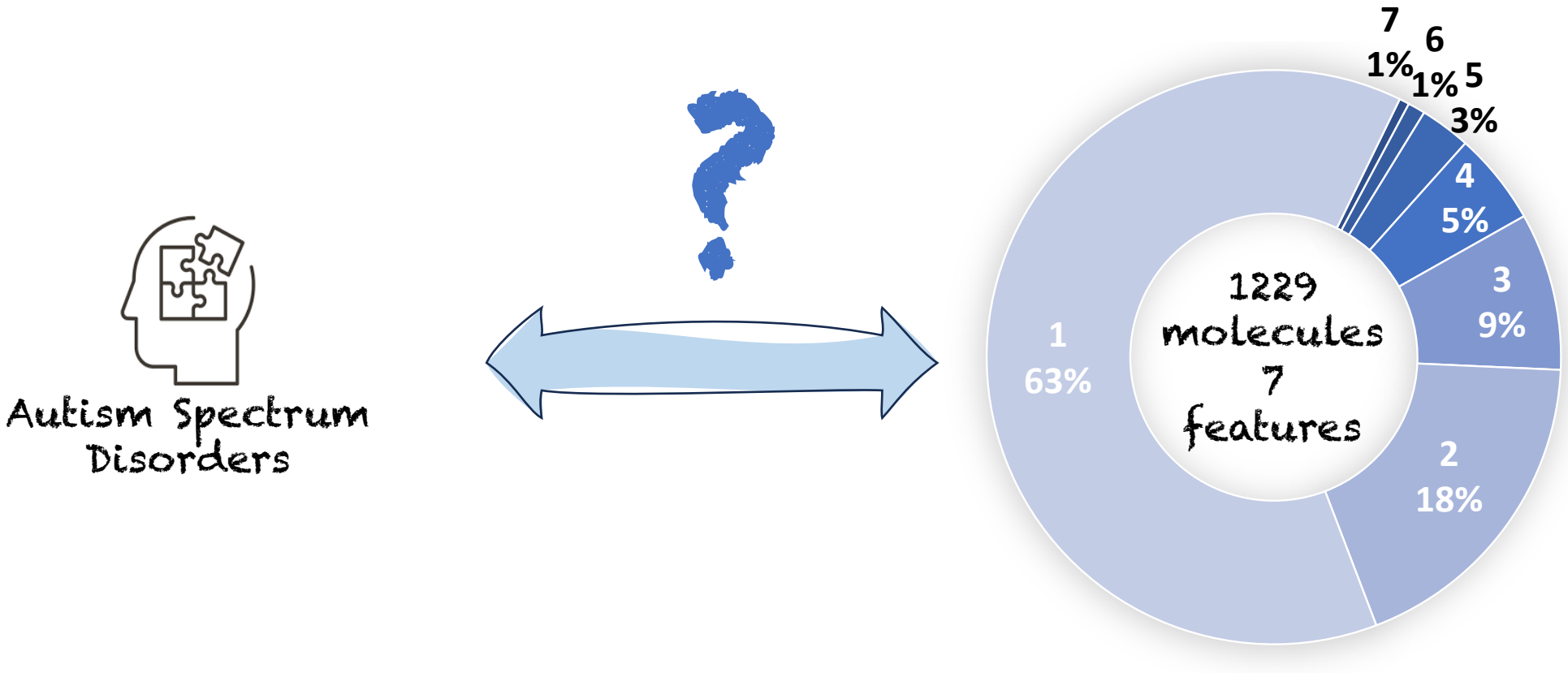
Feature G

Market basket analysis

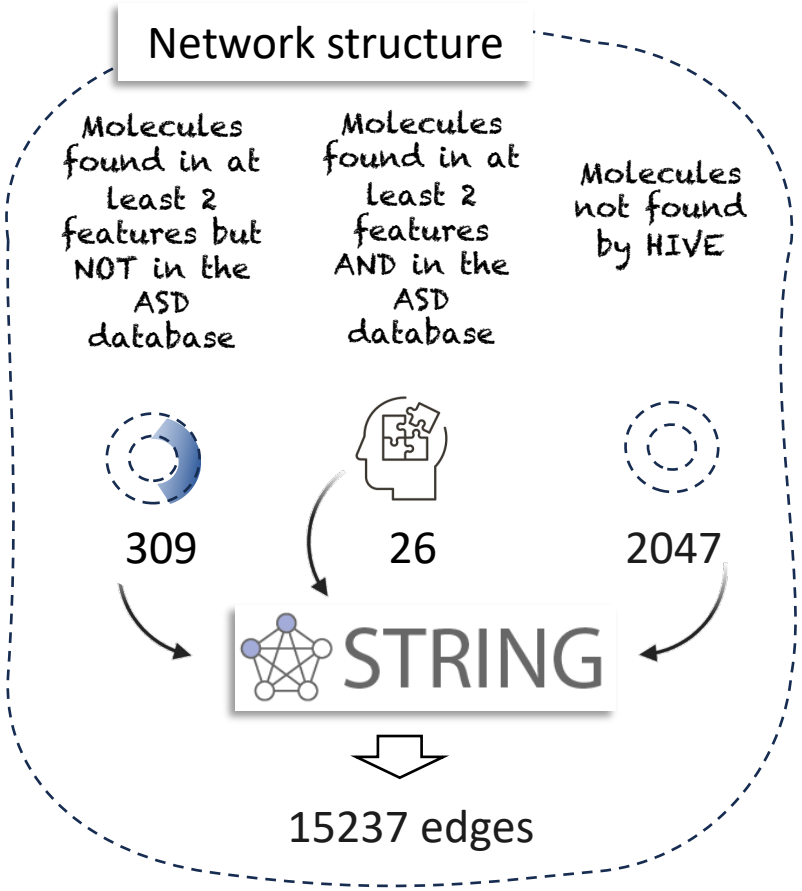
Significant associations between items that are often bought together.

Features encoded in the latent space capture specific biological processes linked to the phenotype under study.

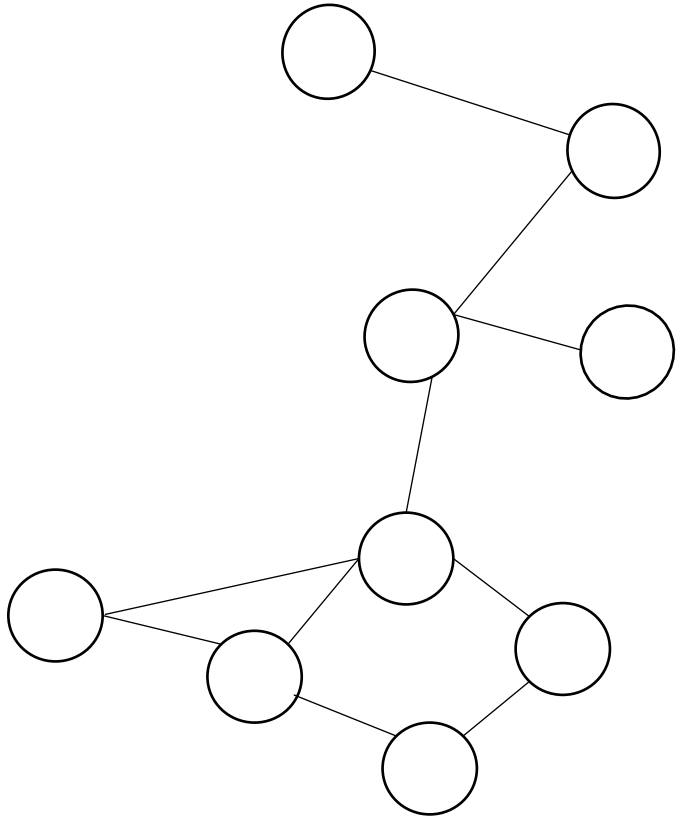
Can we identify novel putative candidate molecules that link ASD and FMRP?

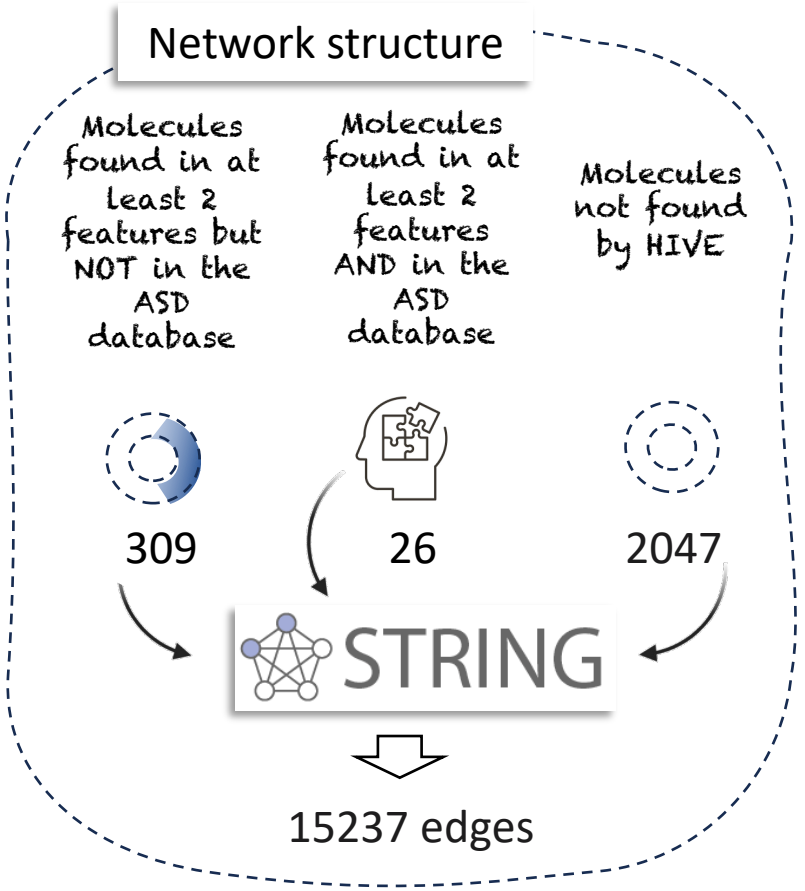


Strategy

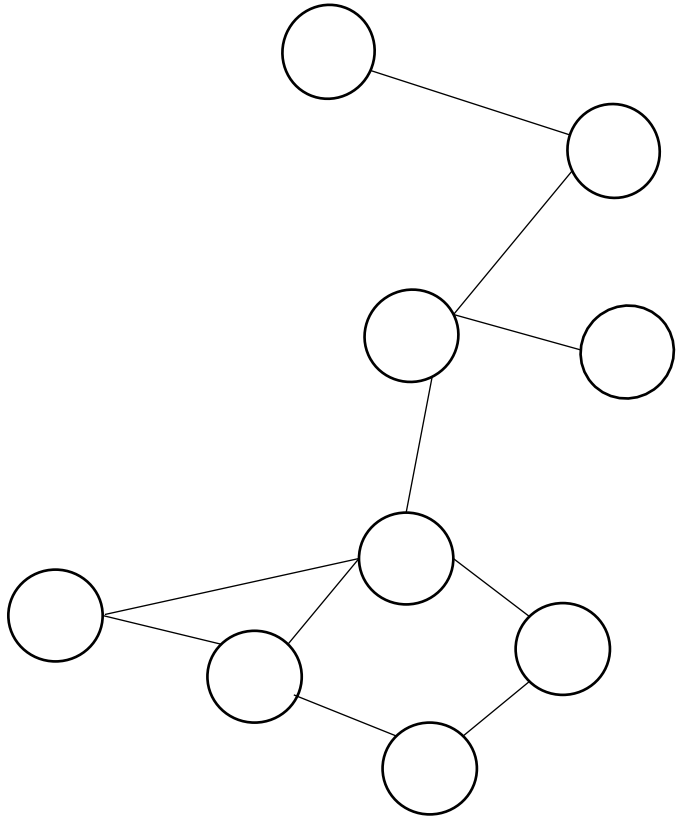

Network structure
Molecules found in at least 2 features but NOT in the ASD database
Molecules found in at least 2 features AND in the ASD database
Molecules not found by HIVE
309
26
2047
STRING
15237 edges

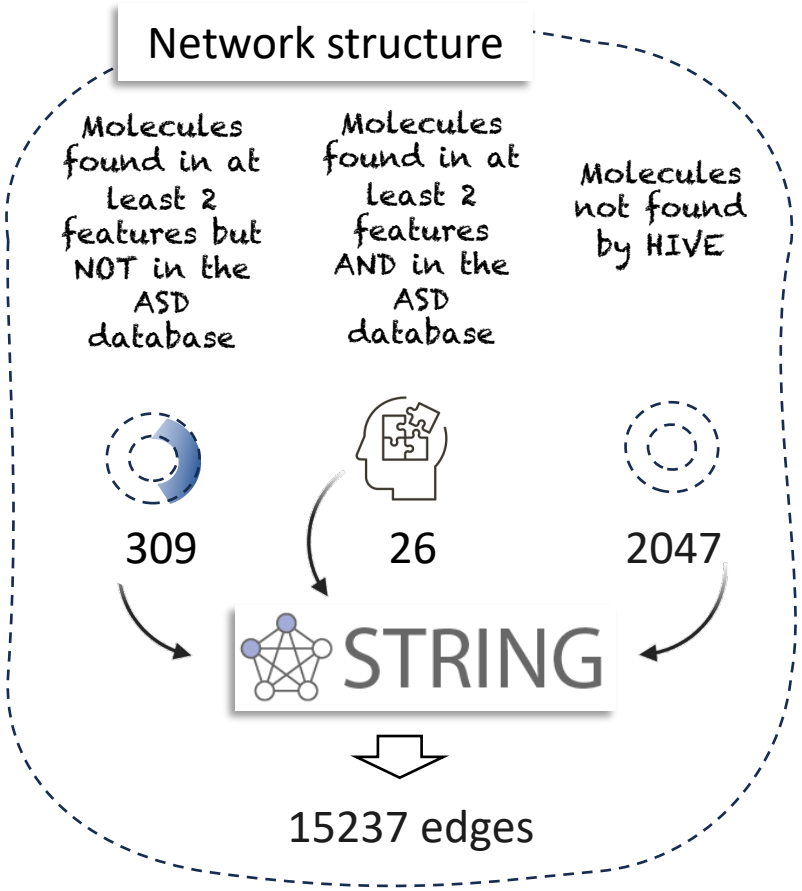
Strategy

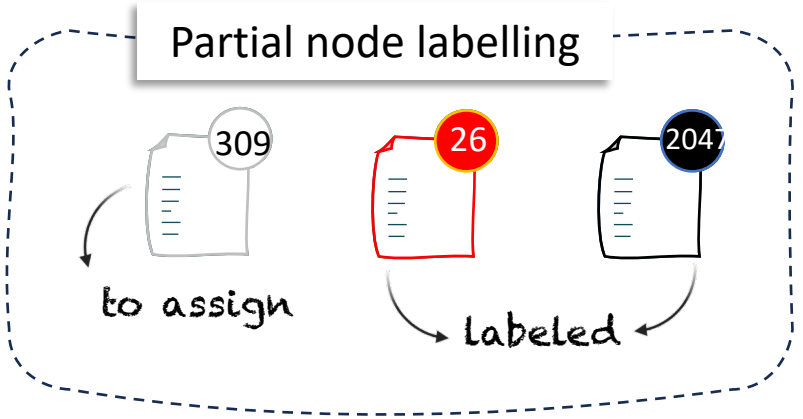



Network structure
Molecules found in at least 2 features but NOT in the ASD database
Molecules found in at least 2 features AND in the ASD database
Molecules not found by HIVE
309
26
2047
STRING
15237 edges

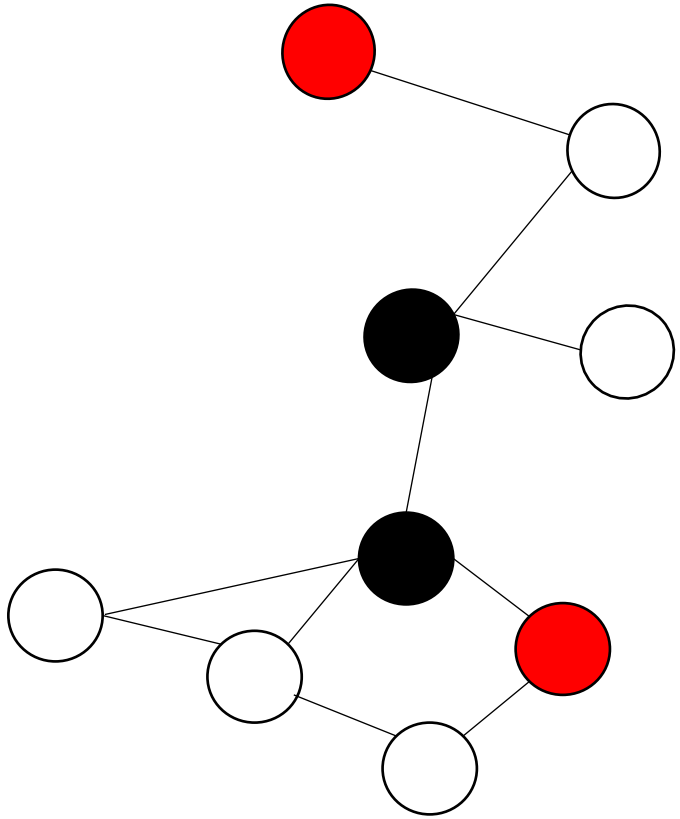
Strategy

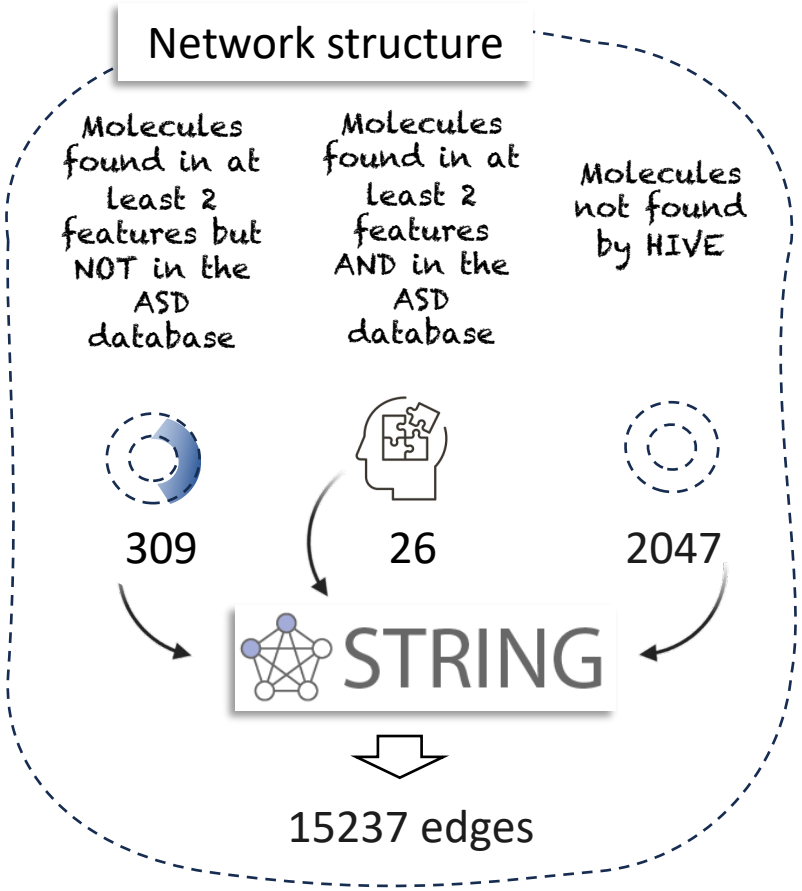


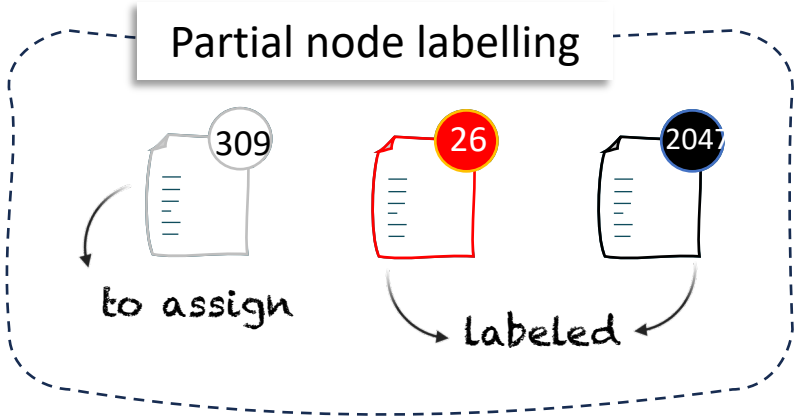

Network structure
Molecules found in at least 2 features but NOT in the ASD database
Molecules found in at least 2 features AND in the ASD database
Molecules not found by HIVE
309
26
2047
STRING
15237 edges


Partial node labelling
309
26
204
to assign
labeled

Strategy

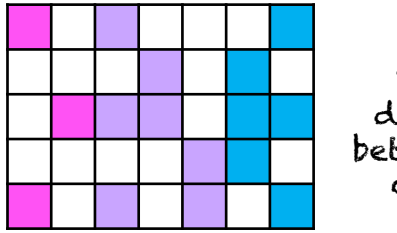



Network structure
Molecules found in at least 2 features but NOT in the ASD database
Molecules found in at least 2 features AND in the ASD database
Molecules not found by HIVE
309
26
2047
STRING
15237 edges


Partial node labelling
309
26
2047
to assign
labeled

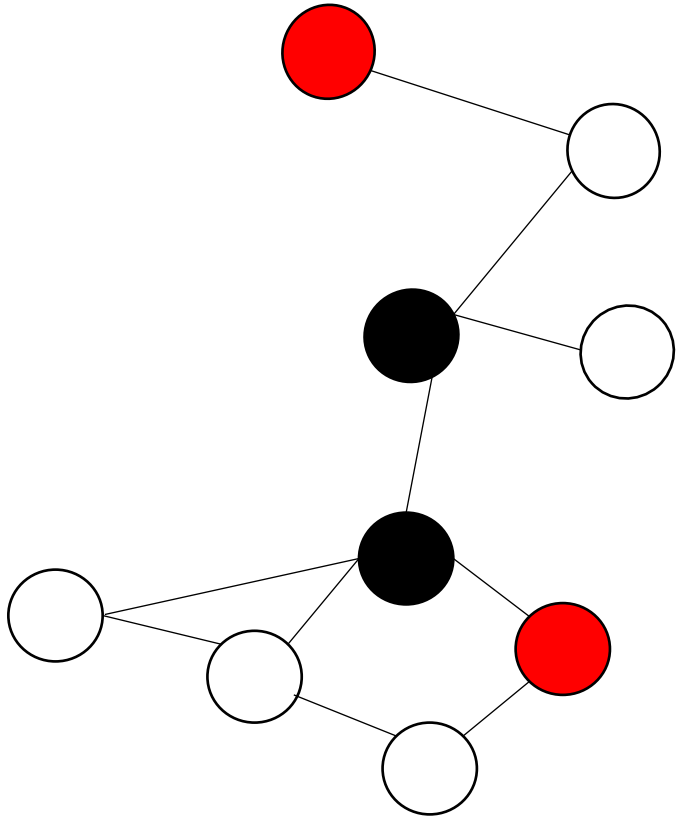
Strategy

Edge weights



Jensen-Shannon divergence between pairs of nodes

Multi-omics integrated dataset



Network structure

Molecules found in at least 2 features but NOT in the ASD database

Molecules found in at least 2 features AND in the ASD database

Molecules not found by HIVE

309

26

2047

STRING

15237 edges

Partial node labelling

309

26

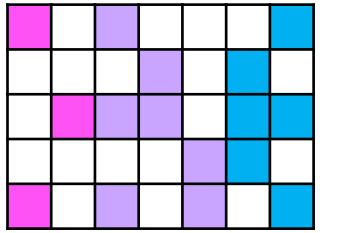
204

to assign

labeled

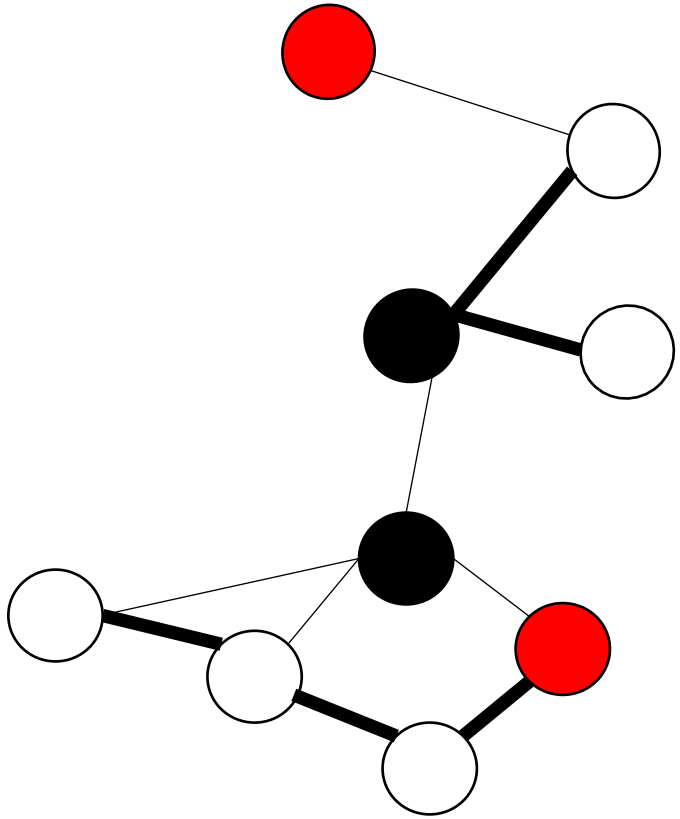
Strategy

Edge weights



Jensen-Shannon divergence between pairs of nodes

Multi-omics integrated dataset



Network structure

Molecules found in at least 2 features but NOT in the ASD database

Molecules found in at least 2 features AND in the ASD database

Molecules not found by HIVE

309

26

2047

STRING

15237 edges

Partial node labelling

309

26

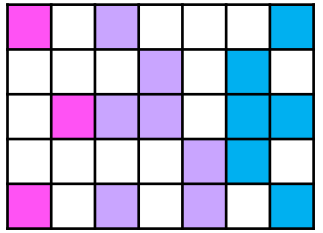
2047

to assign

labeled

Strategy

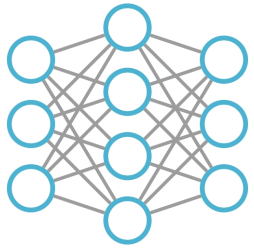
Edge weights



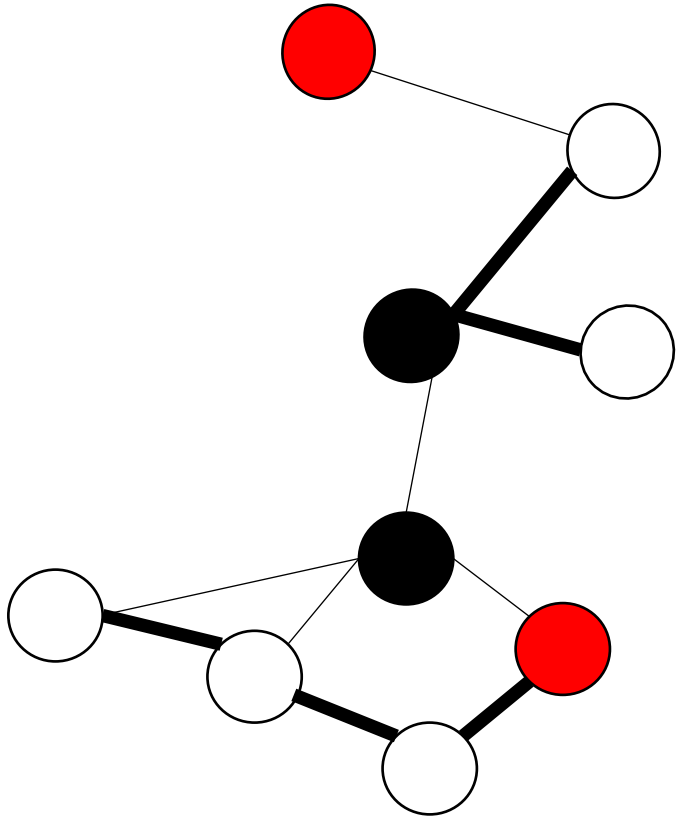
Jensen-Shannon divergence between pairs of nodes

Multi-omics integrated dataset

Node classification



Graph Convolutional network

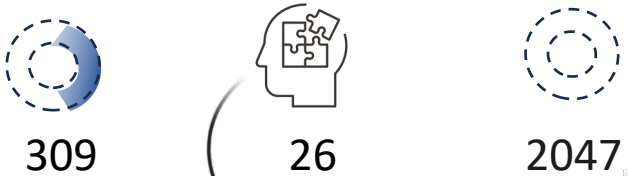


Network structure

Molecules found in at least 2 features but NOT in the ASD database

Molecules found in at least 2 features AND in the ASD database

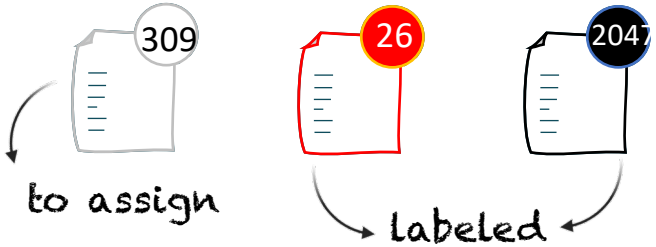
Molecules not found by HIVE



STRING

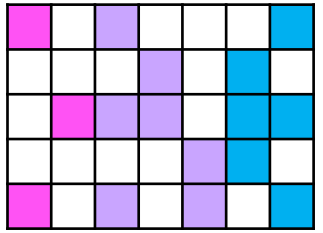
15237 edges

Partial node labelling



Strategy

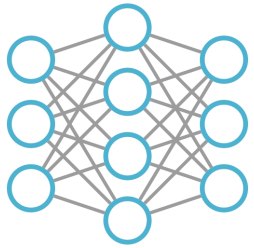
Edge weights



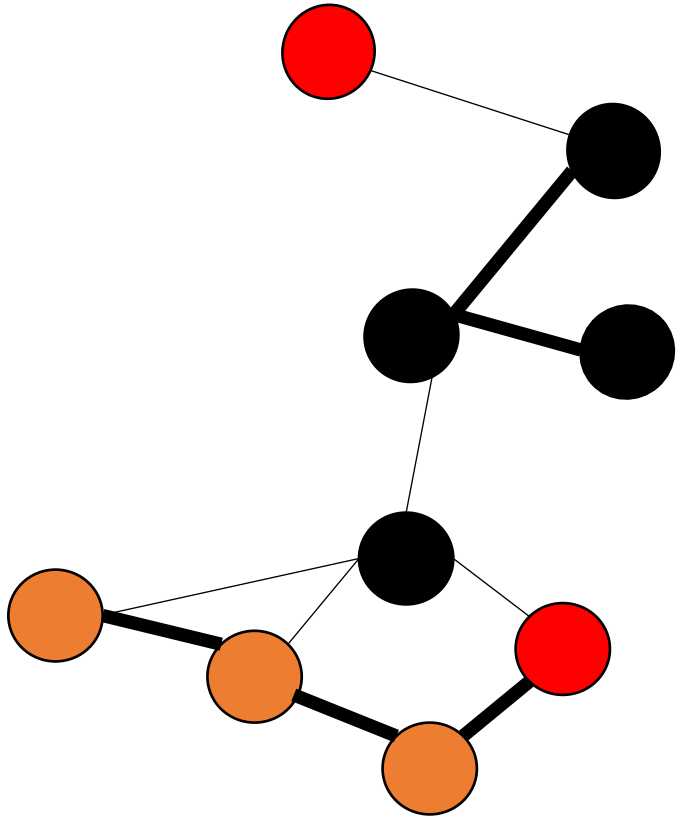
Jensen-Shannon divergence between pairs of nodes

Multi-omics integrated dataset

Node classification



Graph Convolutional network

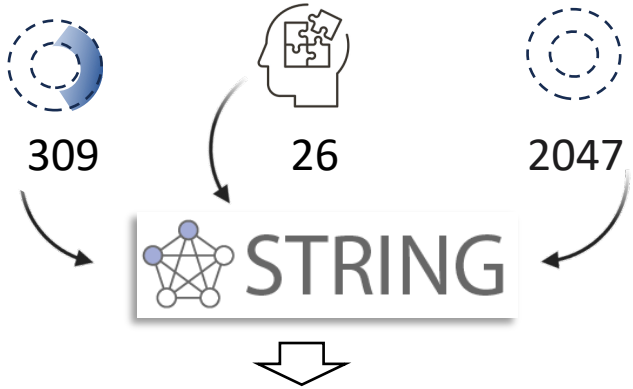


Network structure

Molecules found in at least 2 features but NOT in the ASD database

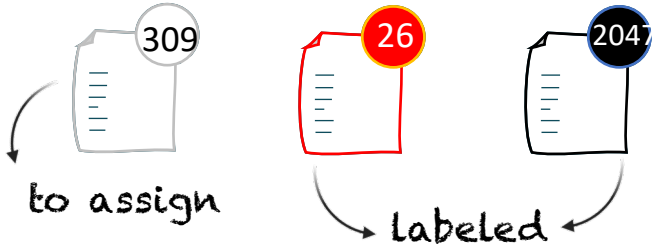
Molecules found in at least 2 features AND in the ASD database

Molecules not found by HIVE



15237 edges

Partial node labelling



Results

Entire network

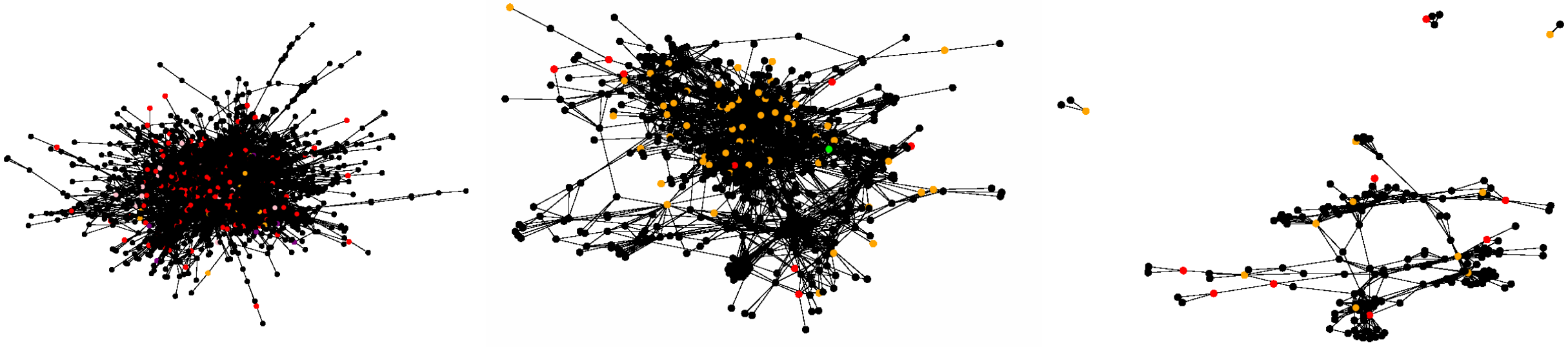
2382 nodes
15237 edges

Second neighbour

nodes
edges

First neighbour

nodes
edges



FMRP

ASD genes from SFARI database

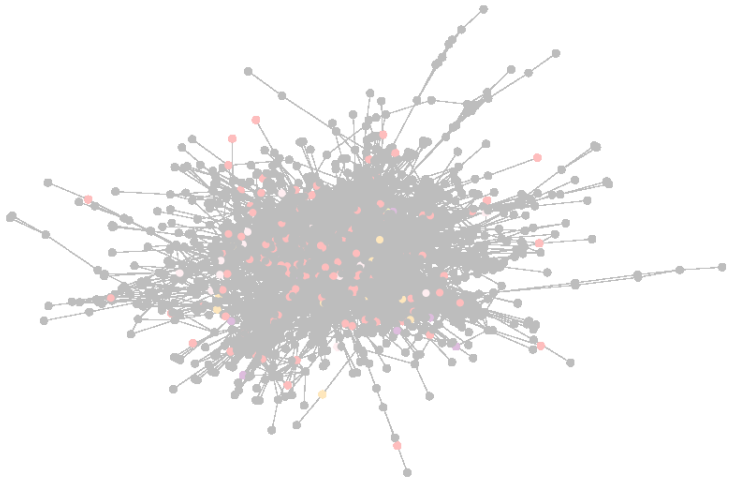
Deregulated in absence of FMRP and predicted as associated to ASD

Predicted as NOT associated to ASD

Results

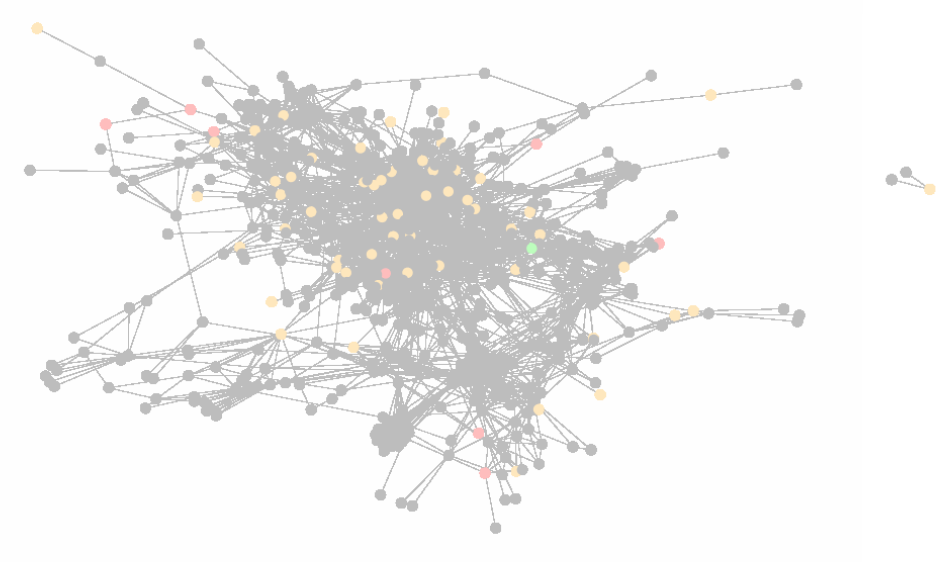
Entire network

2382 nodes
15237 edges



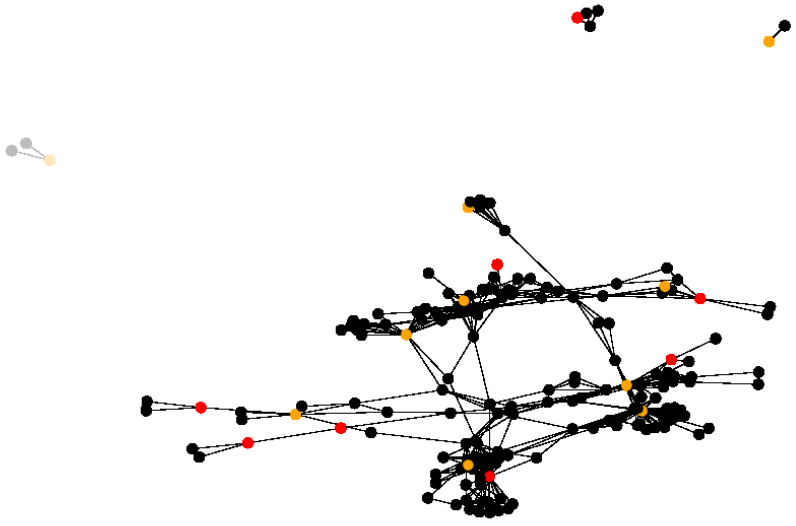
Second neighbour

nodes
edges



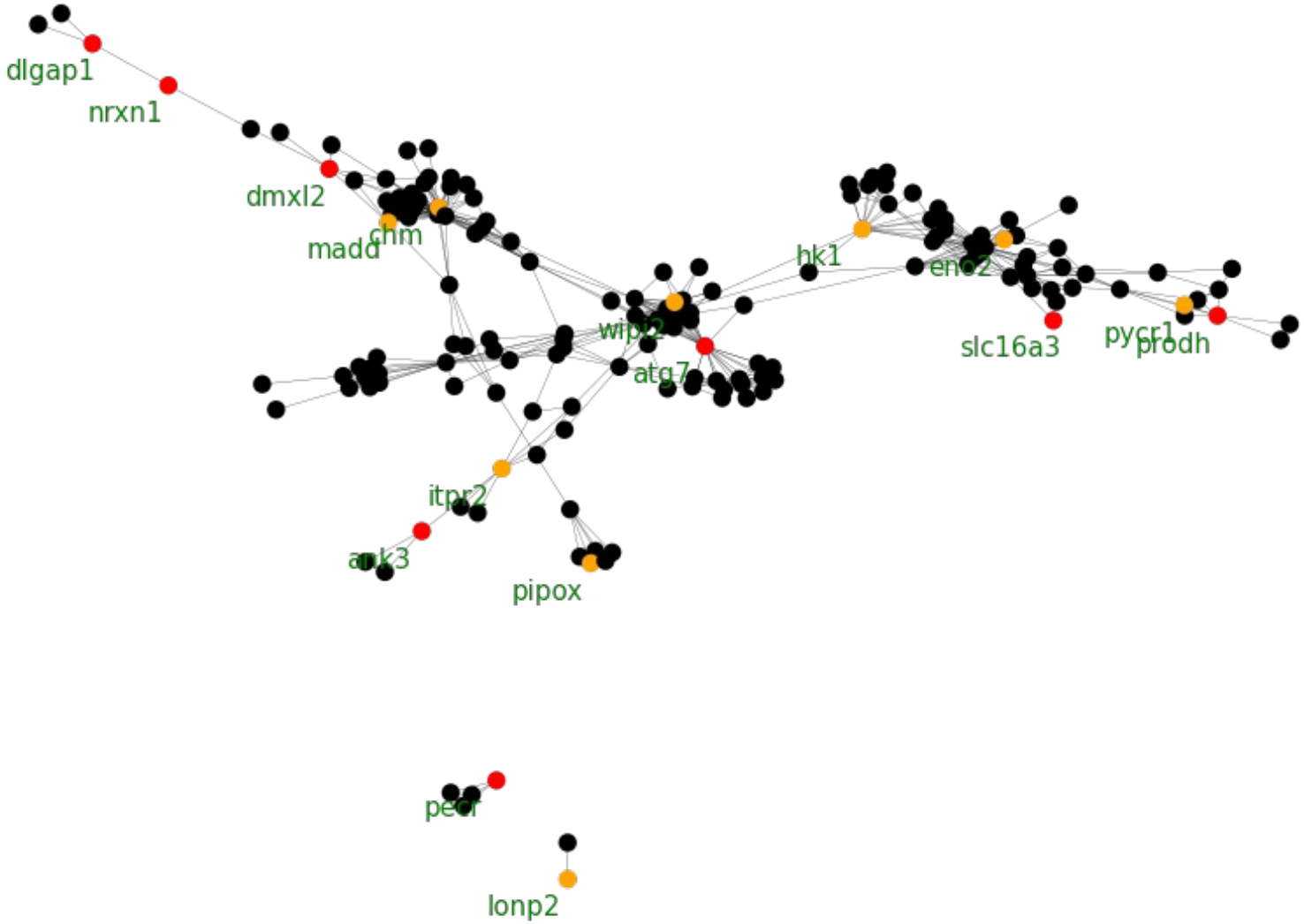
First neighbour

nodes
edges



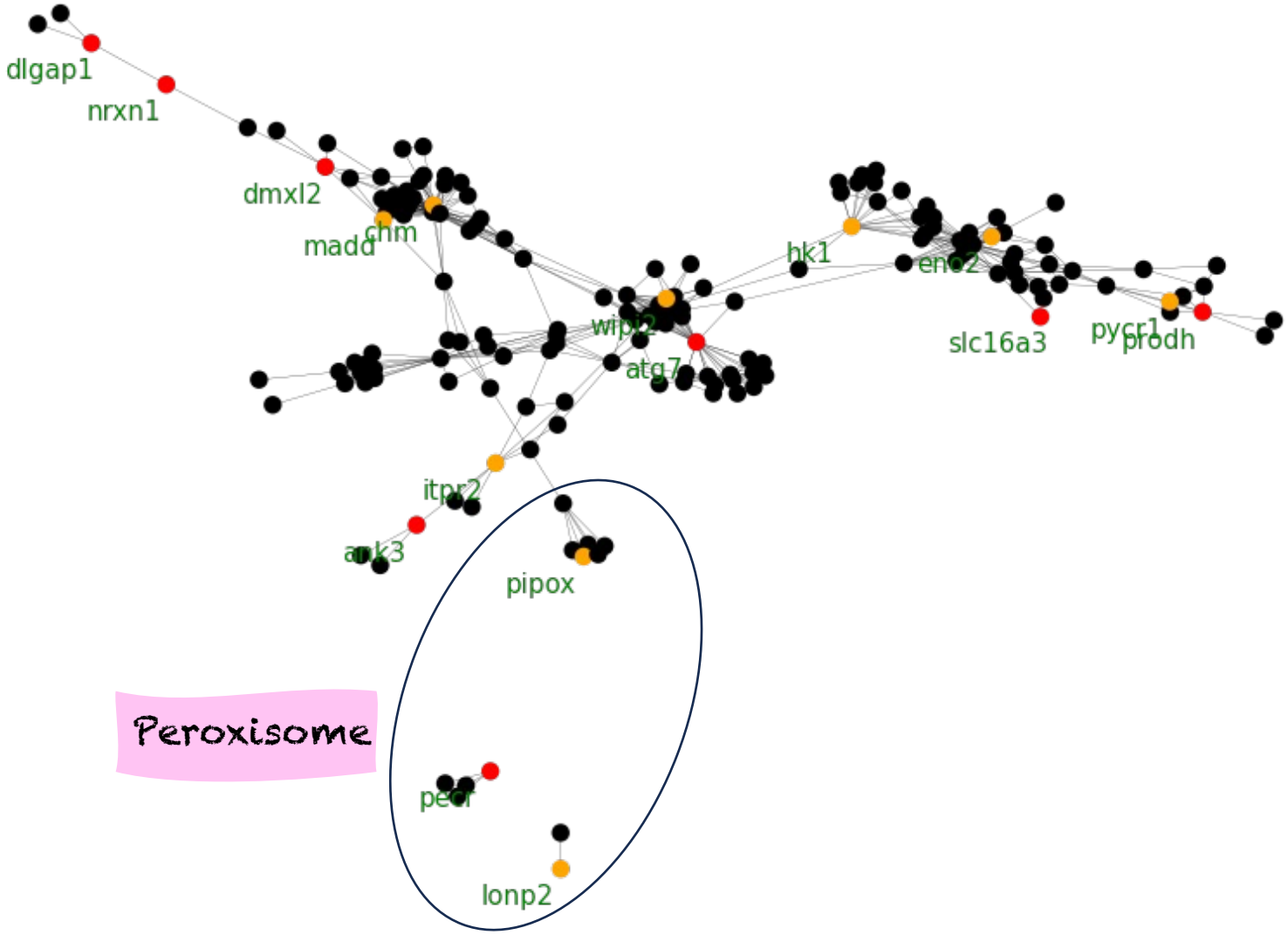
- FMRP
- ASD genes from SFARI database
- Deregulated in absence of FMRP and predicted as associated to ASD
- Predicted as NOT associated to ASD

Results



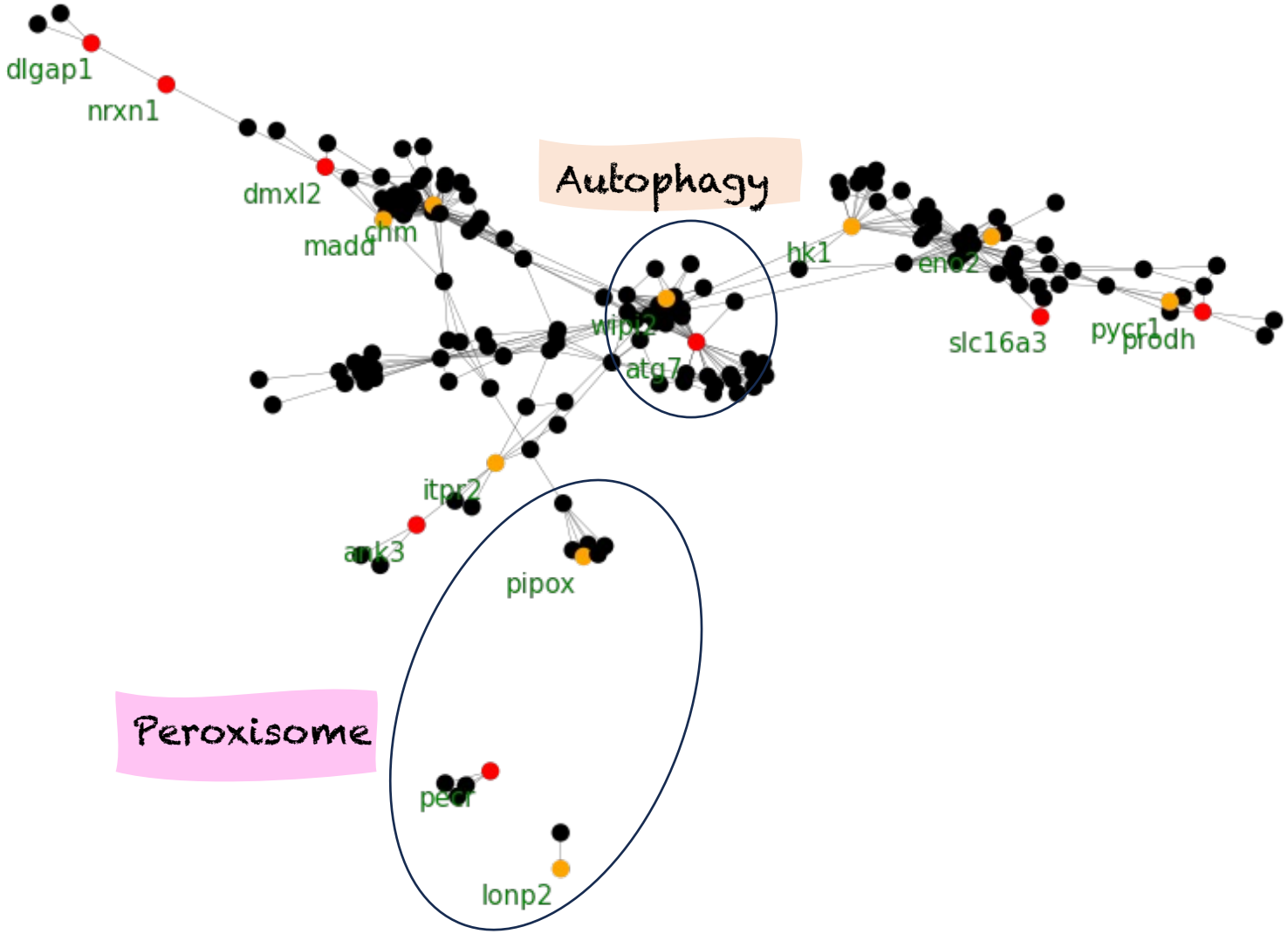
- ASD genes from SFARI database
- Deregulated in absence of FMRP and predicted as associated to ASD
- Predicted as NOT associated to ASD

Results



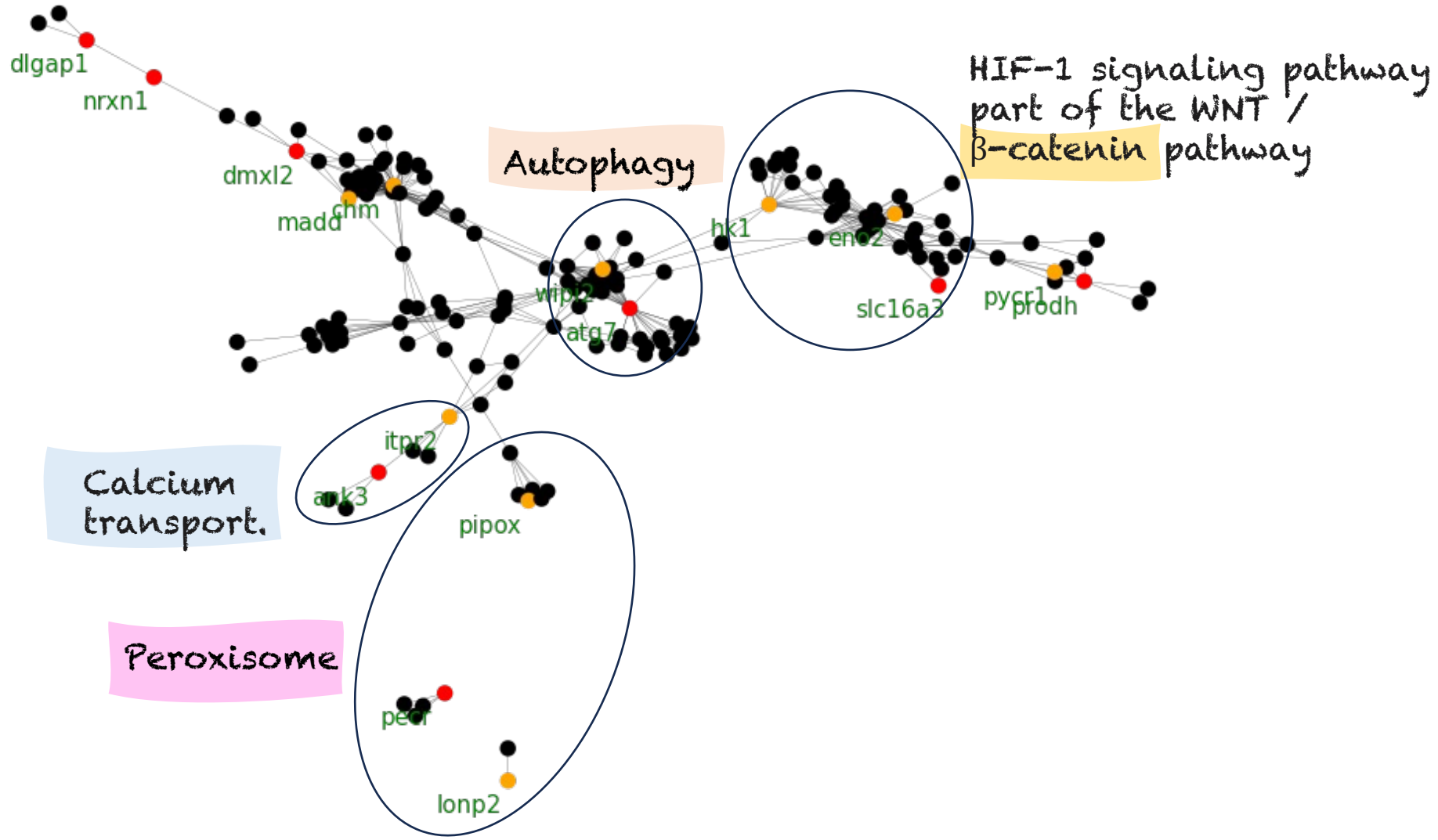
- ASD genes from SFARI database
- Deregulated in absence of FMRP and predicted as associated to ASD
- Predicted as NOT associated to ASD

Results



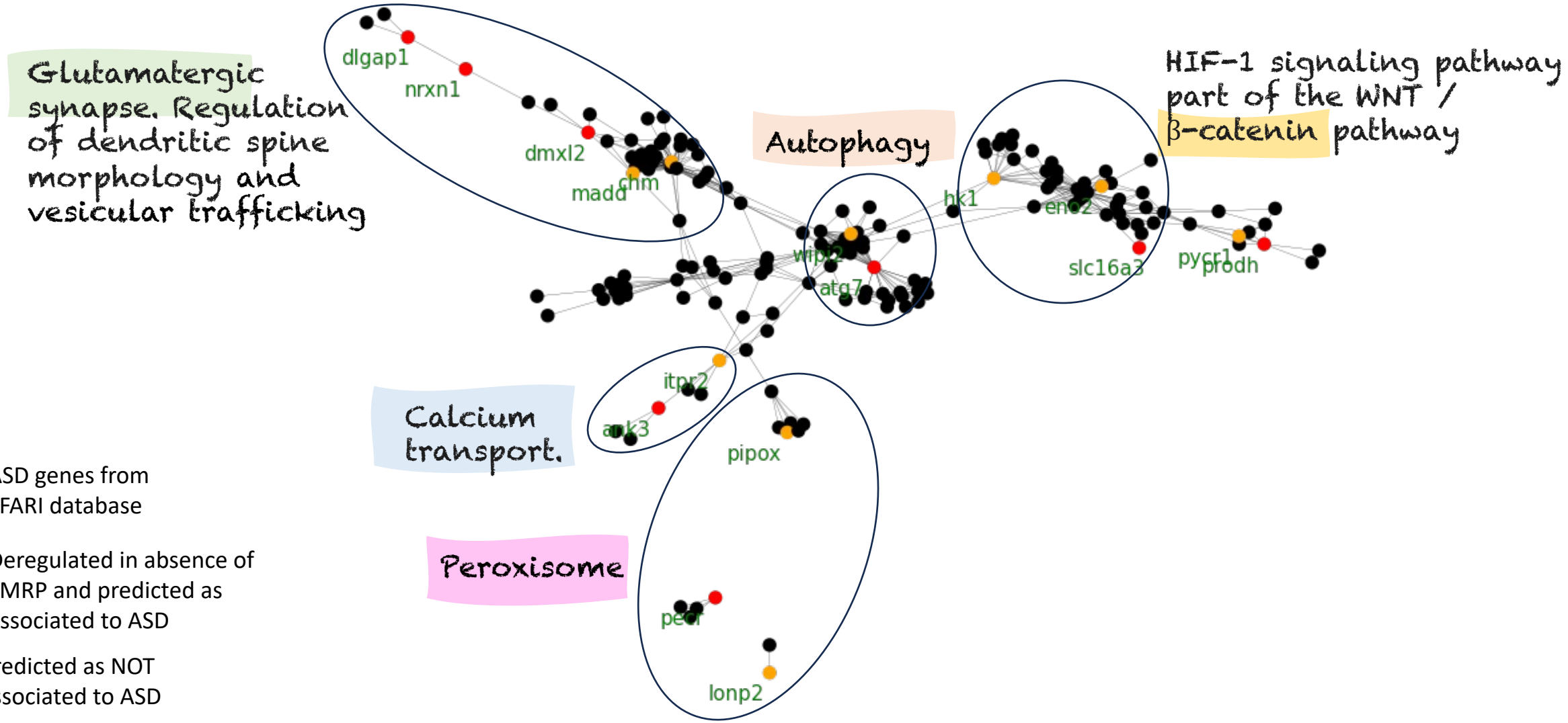
- ASD genes from SFARI database
- Deregulated in absence of FMRP and predicted as associated to ASD
- Predicted as NOT associated to ASD

Results

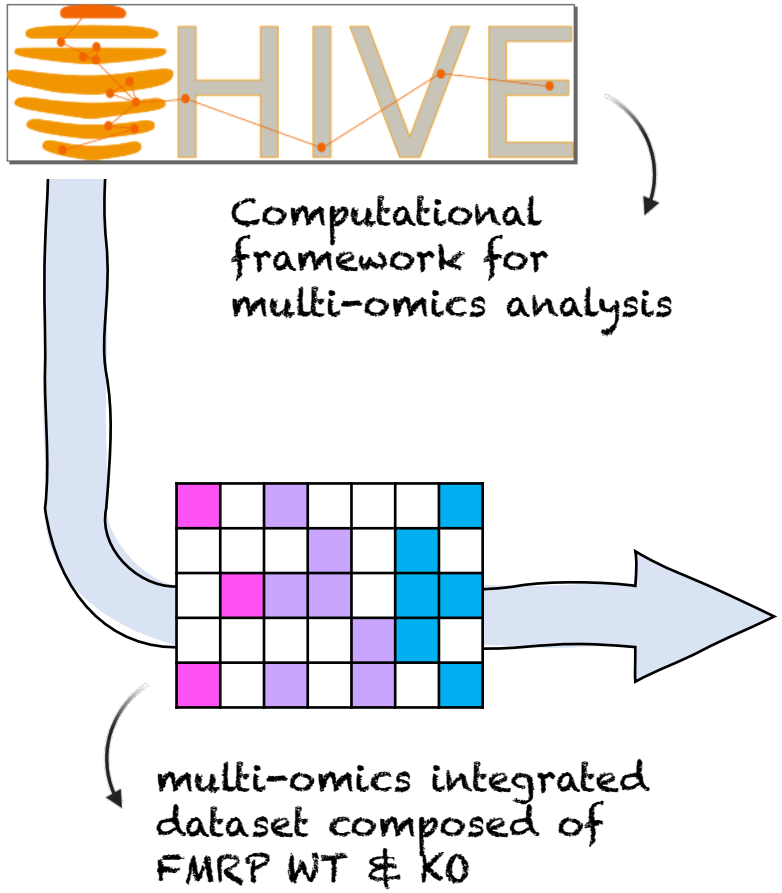


- ASD genes from SFARI database
- Deregulated in absence of FMRP and predicted as associated to ASD
- Predicted as NOT associated to ASD

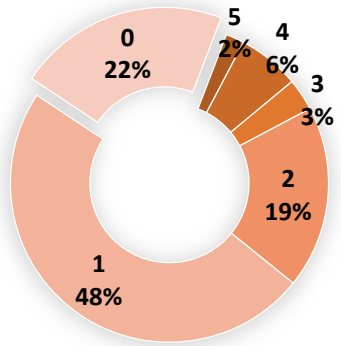
Results


Glutamatergic synapse. Regulation of dendritic spine morphology and vesicular trafficking
dlgap1
nrxn1
dmxl2
madd
chm
Autophagy
wipi2
atg7
HIF-1 signaling pathway part of the WNT / β-catenin pathway
hk1
eno2
slc16a3
pycr1
prodh
Calcium transport.
itpr2
ank3
pipox
Peroxisome
pex1
lonp2
ASD genes from SFARI database
Deregulated in absence of FMRP and predicted as associated to ASD
Predicted as NOT associated to ASD

Conclusions

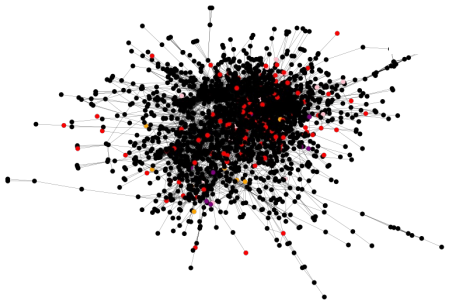


Identify **1229 molecules** deregulated in the absence of FMRP, **22%** of them **not previously** connected to FMRP.

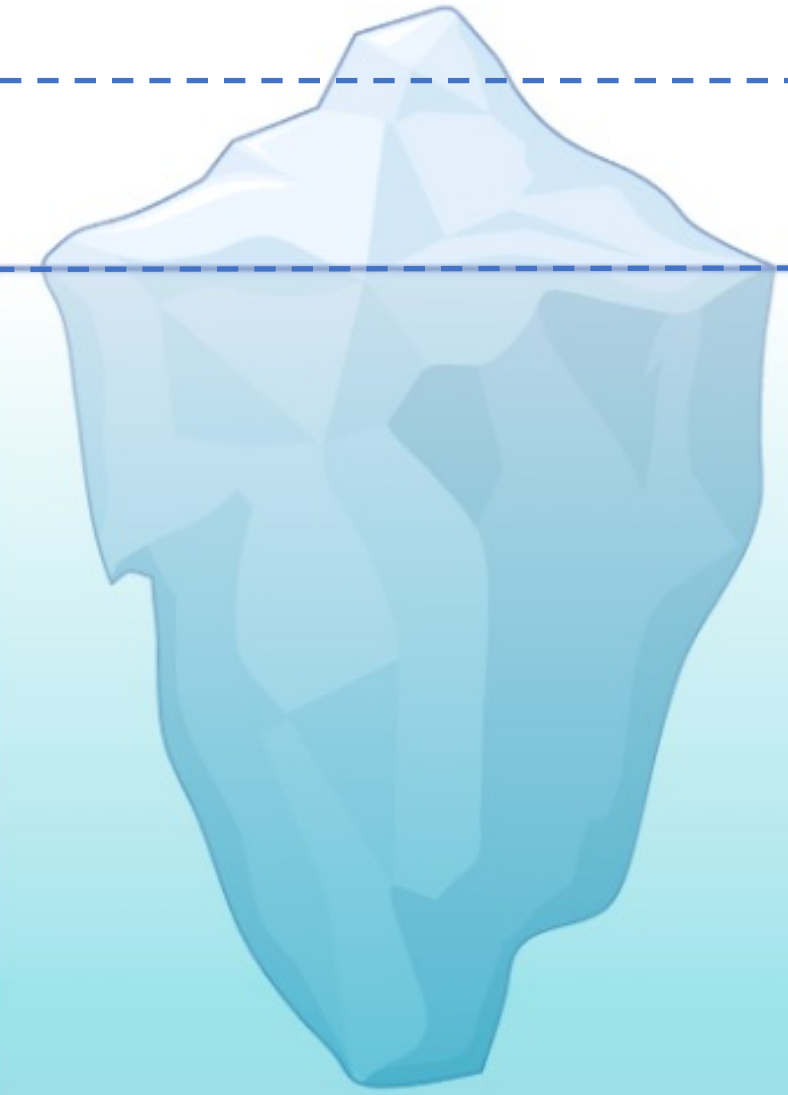


Identify several **novel putative roles** for FMRP including the peroxisome.

Identify of **potential candidates** that link FMRP to ASD.



Take home message



Results showed today

Results obtained with this study

Results that can be potentially obtained by performing

multi-omics,

multi-study,

multi-model

integration analysis.

1 model, 3 omics, 5 datasets

3 models, 7 omics, 43 datasets

Complete understanding of the role of FMRP in neuronal physiology.

Acknowledgements

Giulia Calia
Wassila Kathir
Mame Seynabou Fall
Anjana Bhat
Justine Labory
Djampa Kozlowski
Maxime Multari
Madina Bekbergenova
Youssef Boulaimen
Gauthier Marcovich

Carole Gwizdek
Barbara Bardoni
Jeromine Carret

All the members of the team "RNA Metabolism and Neurodevelopmental Disorders"