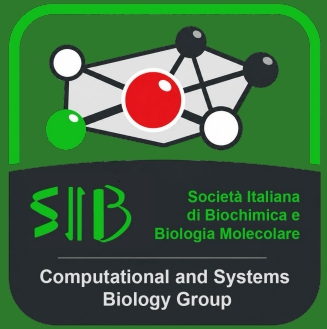
S4 · 16:15–16:30

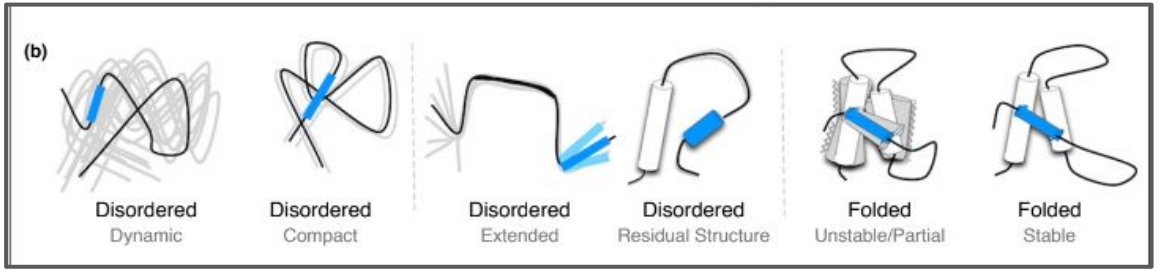
Localized Missense Constraint Reveals Functional Regions in Intrinsically Disordered Proteins Associated with Neurodevelopmental Disorders



Maria Cristina Aspromonte

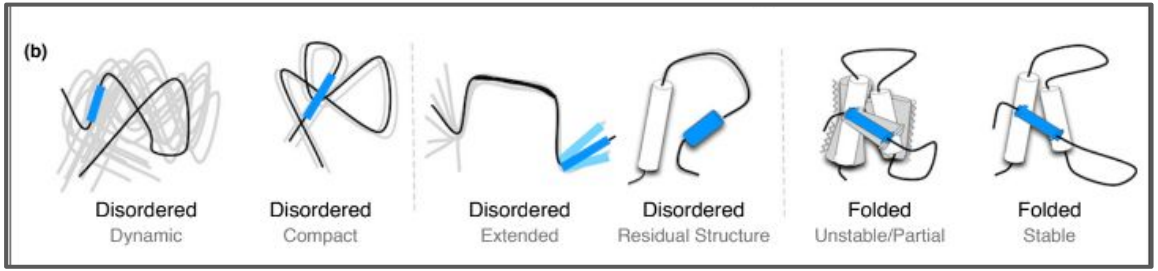
University of Padova

Intrinsically Disordered Proteins and Regions (IDPs/IDRs)

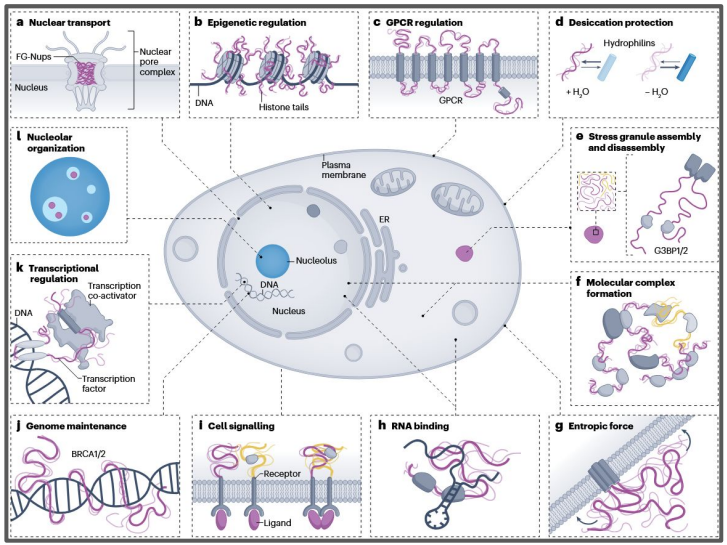


Proteins and protein regions lack a rigid 3D structure under physiological conditions.

Intrinsically Disordered Proteins and Regions (IDPs/IDRs)



Proteins and protein regions lack a rigid 3D structure under physiological conditions.



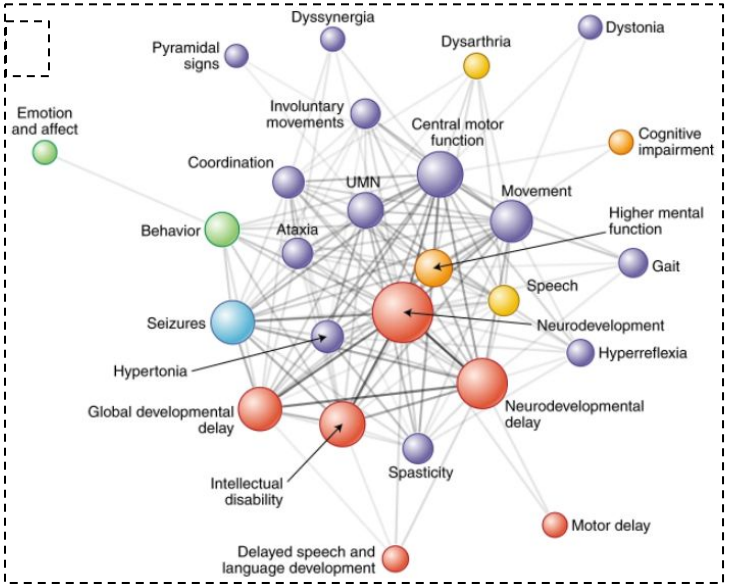
Diverse and Dynamic Functions:

- Signaling and Assembly
- Modification and Molecular Recognition
- Phase Separation

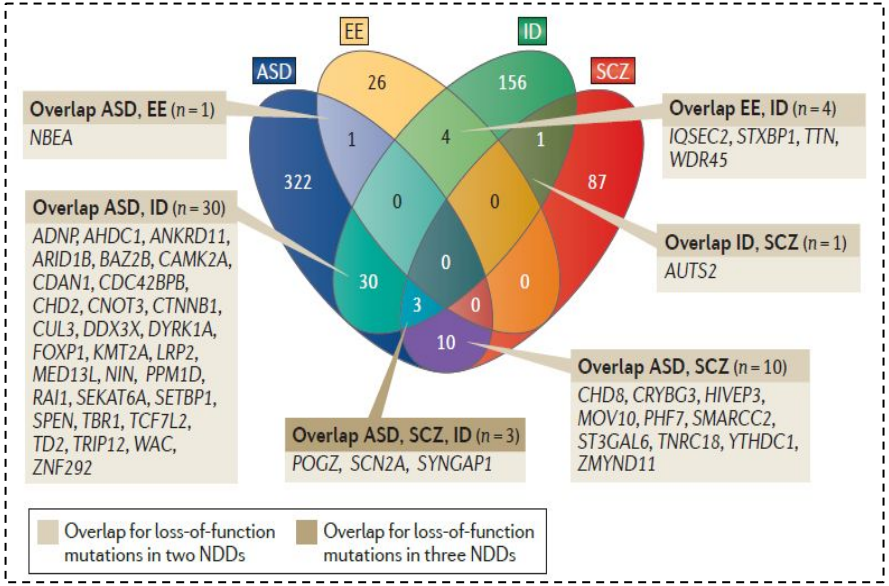
A.S Haleouse et al. Nature Reviews Molecular Cell Biology 2024

Neurodevelopmental disorders (NDDs)

NDDs are characterized by significant comorbidities such as intellectual disability (ID), Autism Spectrum Disorders (ASD), and epilepsy



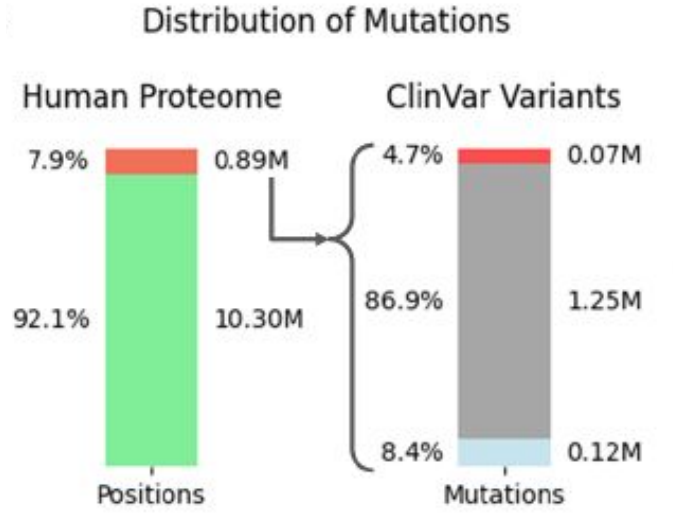
Sanders et al., Nat Med 2019



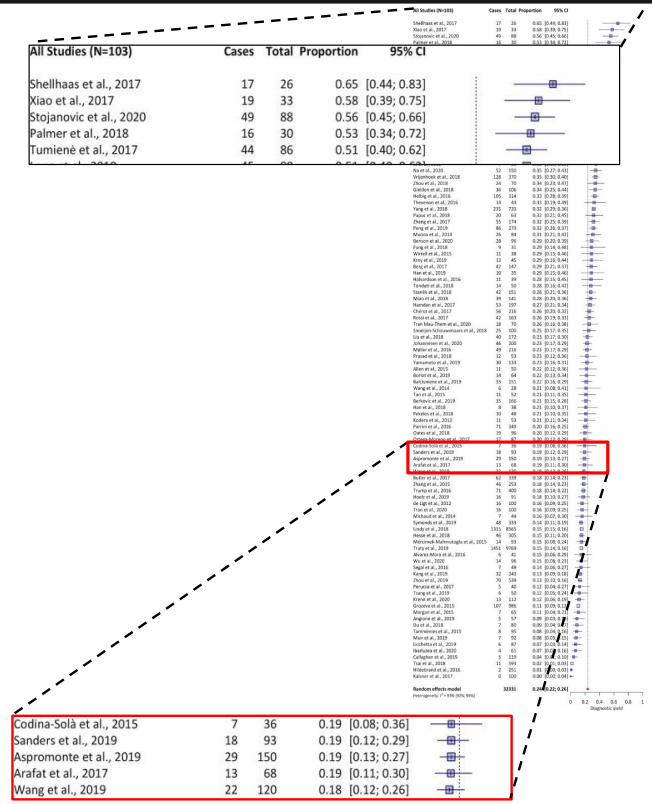
Genes associated with different NDDs are shared among these disorders and converge on common functional pathways

Possible effects of IDR variants

- **20-60 %** diagnostic yield for Rare diseases (by disorder, by age onset, by sequencing technology)
- **2015** Most of the studies started to apply the ACMG guidelines for variant interpretation
- **> 70 %** are classified as Variant of uncertain significance (VUS)

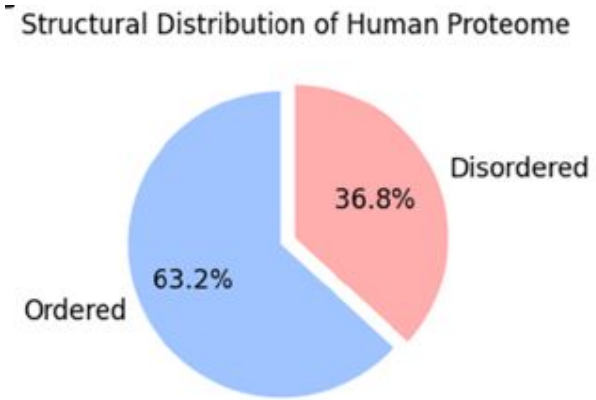


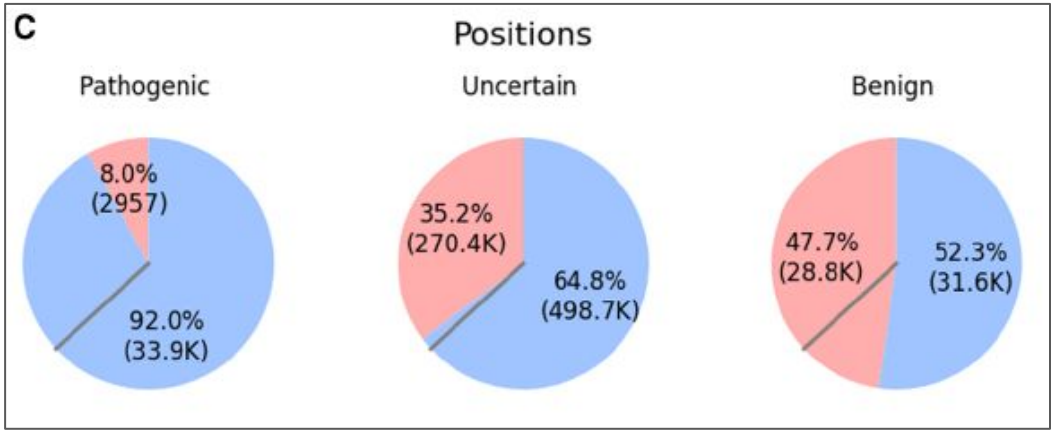
(Deutsch et al., iScience 2026)



Stefanski et al., Epilepsia, 2020

~35 % VUS missense variants map to IDR

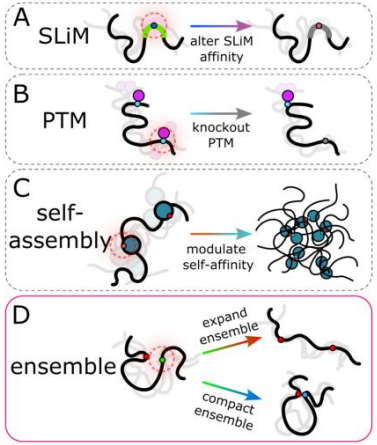




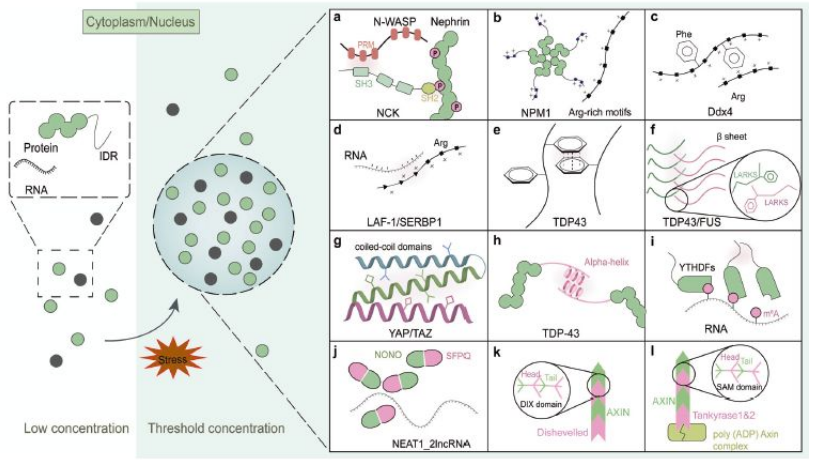
(Deutsch et al., iScience 2026)

- IDR prediction method: MobiDB experimental, AlphaFold2-RSA (>0.582), IUPred3 (>0.4).
- Mutation source: ClinVar
- Variant type: missense

Possible effects of IDR variants



Flores et al, Cur Opin in Struct Biol 2025



Tong et al., Signal Transduction and Targeted Therapy 2022

IDR-containing disease genes often fail through:

- altered motif recognition
- rewired signaling
- abnormal PTMs
- gain/loss of transient interactions
- ensemble conformation changes

Impact on:

- Multivalent electrostatic interactions
- Aromatic–charged residue patterning
- Hydrophoobic interactions
- remove/add regulatory switches
- RNA binding

Our Goals

Structural and Functional Characterization of NDD Proteins

Identify functionally constrained regions within intrinsically disordered proteins associated with neurodevelopmental disorders.

Genetic Constraint Investigation

Determine whether missense constraint can reveal pathogenic hotspots in disordered and structurally flexible regions.

Alternative Approach

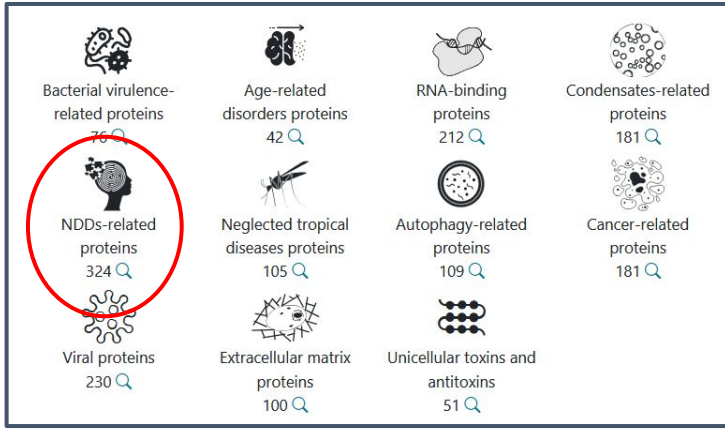
Develop an annotation framework to support the prioritization of VUS in NDD-related proteins.

Methods: Dataset construction

DisProt V9.8

Disorder structural states and functions

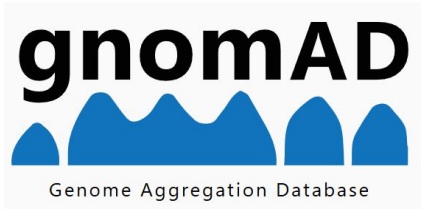
MobiDB
a database of protein disorder and mobility annotations



Ensembl V115

Coding sequences collection

Methods: Dataset construction



V4.1.0

EXTRACTION

730,947 exome and **76,215 whole-genome** sequences
unrelated healthy individuals → GraphQL API / BigQuery pipelines

Variant information:

- ▶ Chromosome position
- ▶ Reference allele
- ▶ Alternative allele
- ▶ Variant type

VCF files extracted from gnomAD →

ANNOTATION ANNOVAR

- ▶ hg38 reference genome
- ▶ *refGeneWithVer* □ transcript consequence
- ▶ *ClinVar* □ clinical significance

Methods: Missense Tolerance Ratio (MTR) calculation

The ratio of *observed missense variants* to the *expected number of missense variants* in a given region of a gene, where the expectation assumes no natural selection

✔ *Variants (**synonymous**; **missense**) were mapped to canonical protein sequences using a sliding window of 31 amino acids, moving one amino acid at a time.*

✔ *MTR score per residue: with lower scores indicating stronger selective constraint.*

$$MTR = \frac{Observed\ missense\ proportion}{Expected\ missense\ proportion}$$

✔ *To identify significantly intolerant regions, an exact binomial test with False Discovery Rate (FDR) correction was applied comparing observed vs expected missense proportions.*

Silk M. et al, NAR 2019

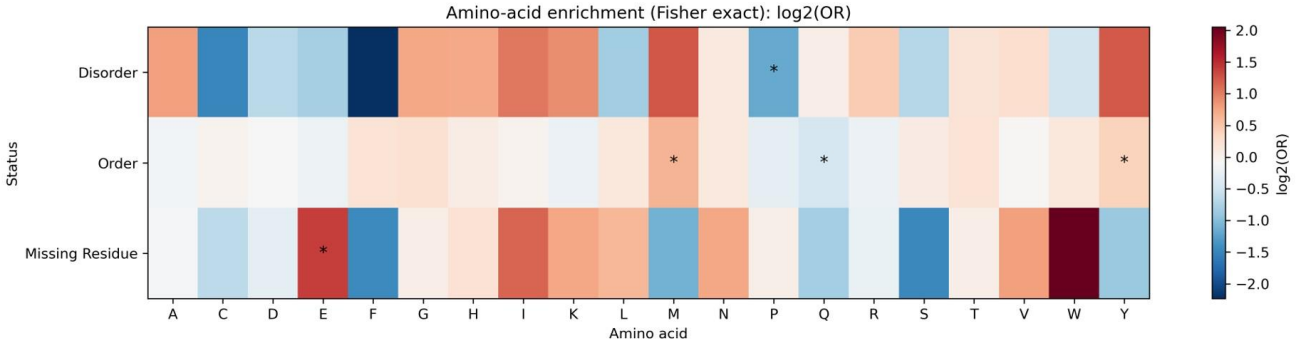
Missense Constraint across all structural states in NDD-associated proteins

Residues classified as Intolerant (MTR < 0.4) and Tolerant (MTR ≥ 0.4), were compared for composition and biophysical properties.

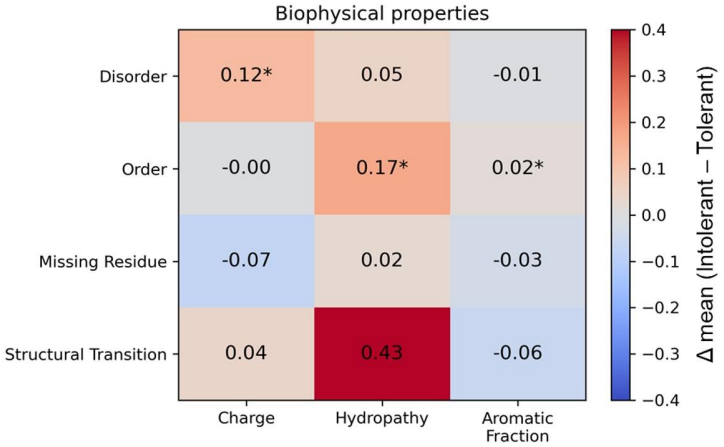
Modest but consistent differences in IDRs, indicating that constrained disordered segments are associated with specific compositional features

Robles et al, Genes 2026

Missense Constraint across all structural states in NDD-associated proteins



Intolerant residues in IDRs are enriched in positive charge, whereas segment-level analyses reveal that constrained regions form short, localized stretches embedded within otherwise tolerant disordered sequences.



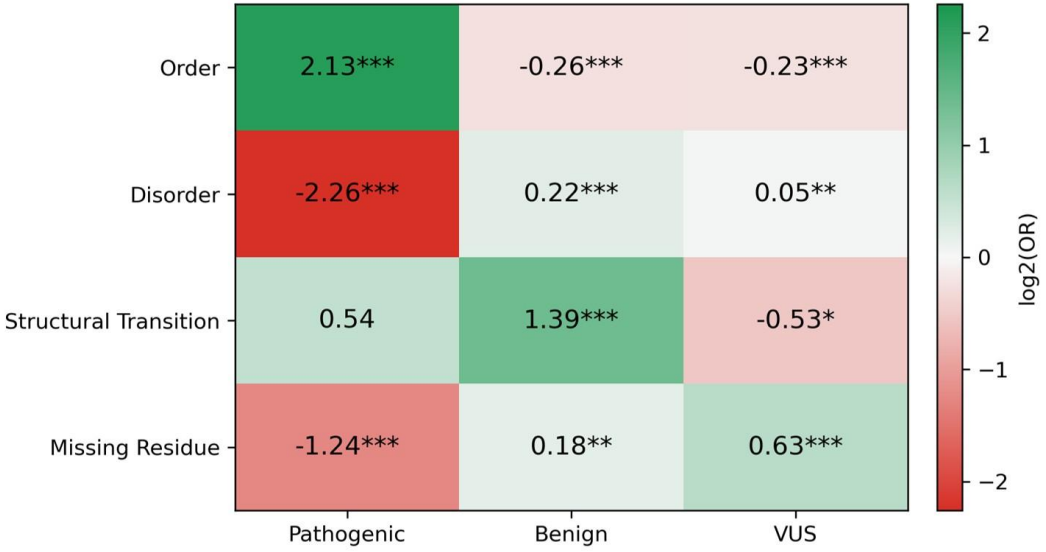
Residues classified as Intolerant (MTR < 0.4) and Tolerant (MTR ≥ 0.4), were compared for composition and biophysical properties.

Modest but consistent differences in IDRs, indicating that constrained disordered segments are associated with specific compositional features.

Red (log2OR > 0.0) = Intolerant, blue (log2OR < 0.0) = Tolerant, white ≈ no difference.

Robles et al, Genes 2026

Clinically Significant Variant Enrichment

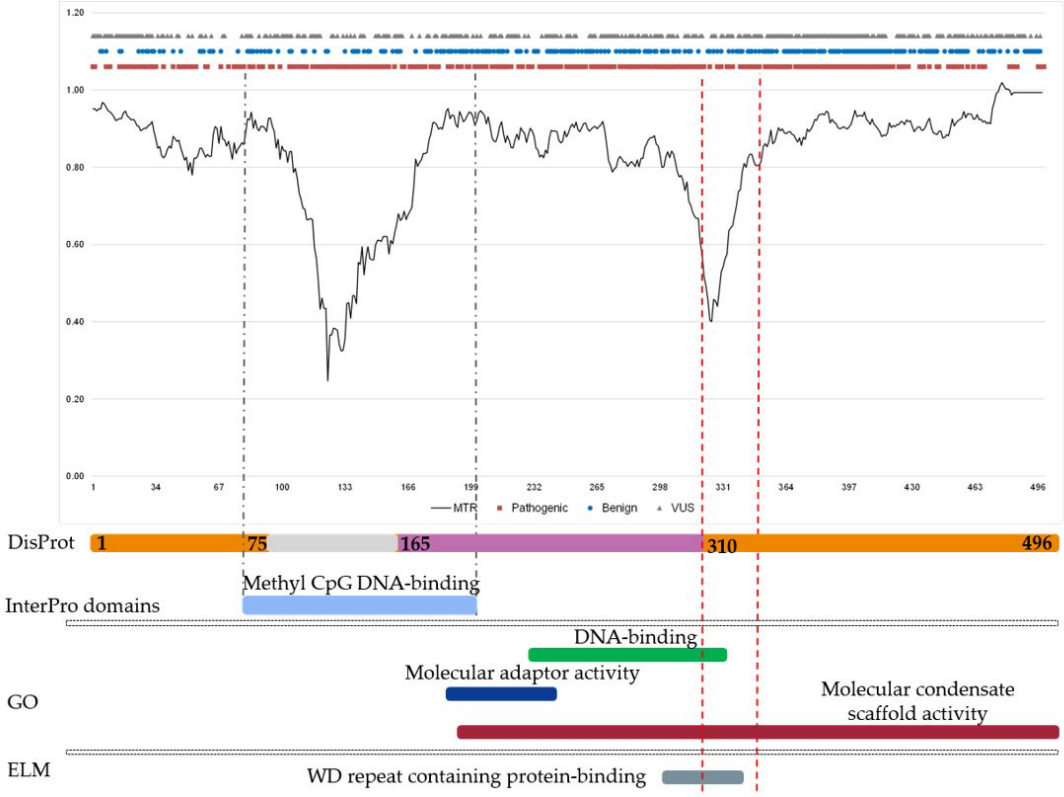


33,124 ClinVar variants mapped in **339 NDD-associated proteins** and tested for enrichment or depletion of pathogenic, benign, and VUS annotations across structural states using Fisher's exact tests.

Pathogenic variants are significantly enriched in ordered regions (log2OR = 2.13, $q < 0.001$) and strongly depleted in IDRs (log2OR = −2.26, $q < 0.001$)

Robles et al, Genes 2026

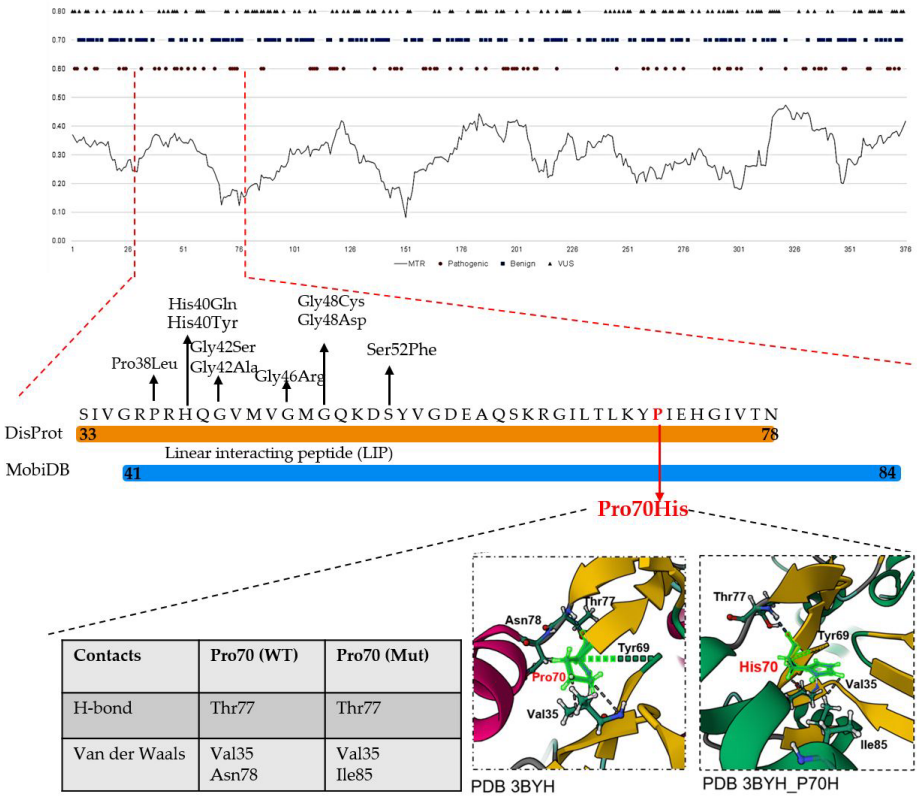
Evaluating missense variations by local constraints in MECP2



- Methyl CpG Binding Protein 2 (MECP2), a DNA-binding protein that mediates transcriptironal repression through interactions with histone deacetylases.
- Pathogenic variants in *MECP2* cause Rett syndrome (OMIM #312750).
- Low MTR regions are detected within the Methyl-CpG Binding Domain (residues 78–162) and within a **C-terminal disordered region (residues 310–369)**.

Robles et al, Genes 2026

Evaluating missense variations by local constraints in Actin, cytoplasmic 1



Contacts	Pro70 (WT)	Pro70 (Mut)
H-bond	Thr77	Thr77
Van der Waals	Val35 Asn78	Val35 Ile85

(UniProt: P60709; DisProt: DP03957)

- Actin, cytoplasmic 1 cause Baraitser–Winter syndrome 1 (BRWS1)
- Within the IDR of Actin - cytoplasmic 1 (UniProt: P60709; DisProt: DP03957) corresponding to the lowest MTR values (<0.20)
- Pro70His variant, currently classified as VUS, is absent in gnomAD and is reported in ClinVar in association with Baraitser–Winter syndrome 1 (BRWS1)

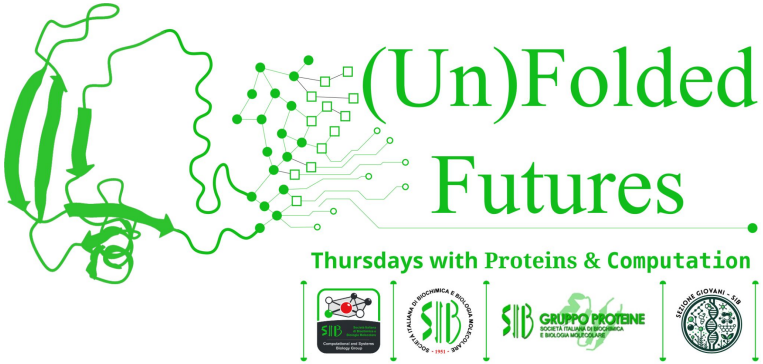
Robles et al, Genes 2026

Constraint is only modestly shifted by global structural class, but highly localized within each class

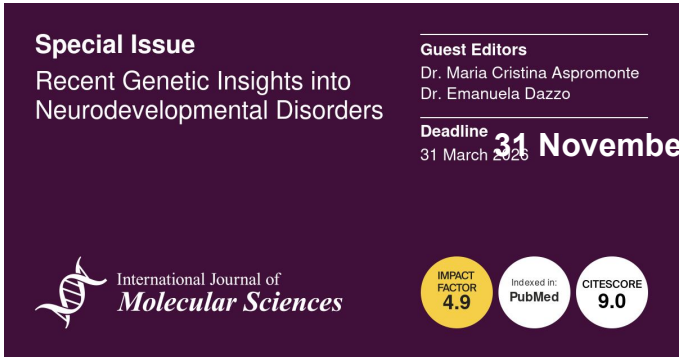
Structural bias in clinical annotation of missense variants

Limitations *in dataset size, clinical annotation and variant interpretation, which disproportionately affect disordered regions and may lead to under-recognition of constrained sites outside ordered domains.*

Acknowledgements



Seminar series for SIB young researchers - from September 2026



Silvio Tosatto

Nazareth Azugaray, Master student

Thank you for your attention