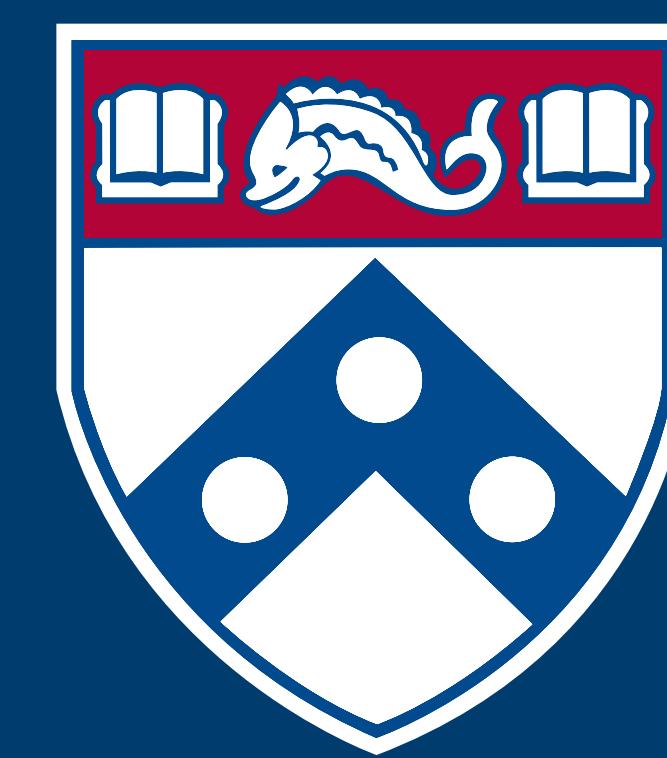


# Characterizing genetic variation in disordered regions associated with Alzheimer's disease

Jason E. Miller, Scott M. Dudek, Alex T. Frase, and Marylyn D. Ritchie

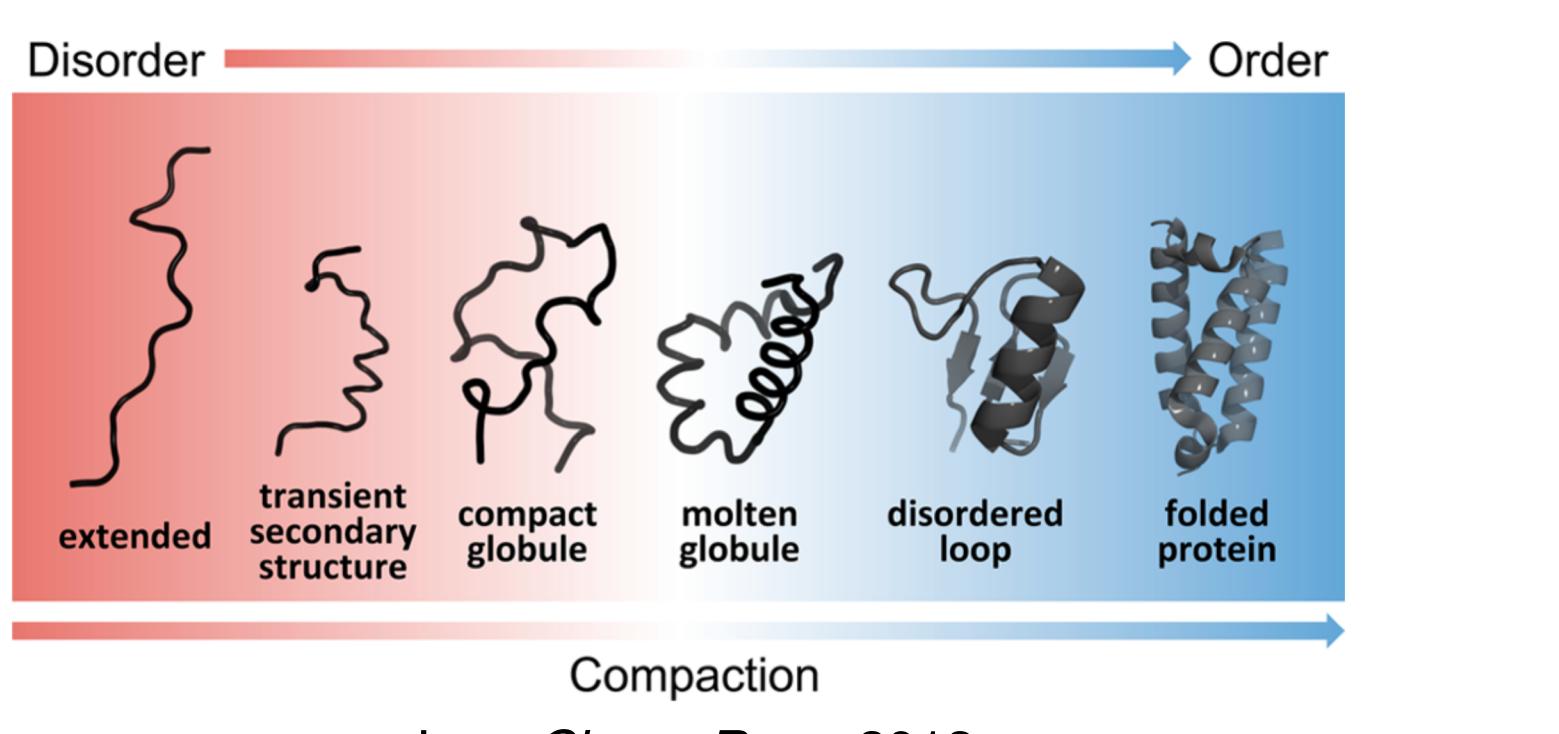
Department of Genetics, Institute for Biomedical Informatics, University of Pennsylvania, Philadelphia, PA, USA



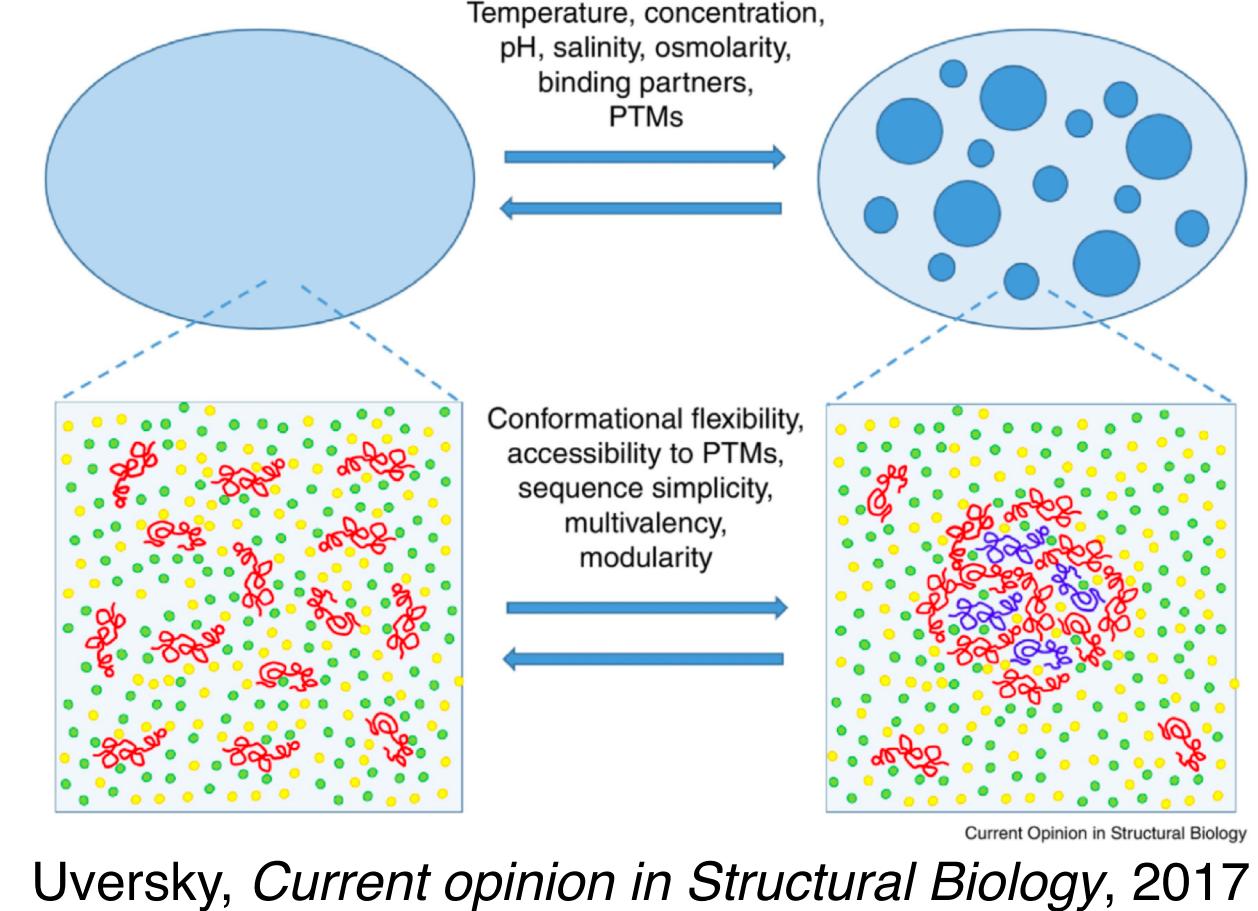
Perelman  
School of Medicine  
UNIVERSITY of PENNSYLVANIA

## Question: how are intrinsically disordered proteins linked to Alzheimer's disease?

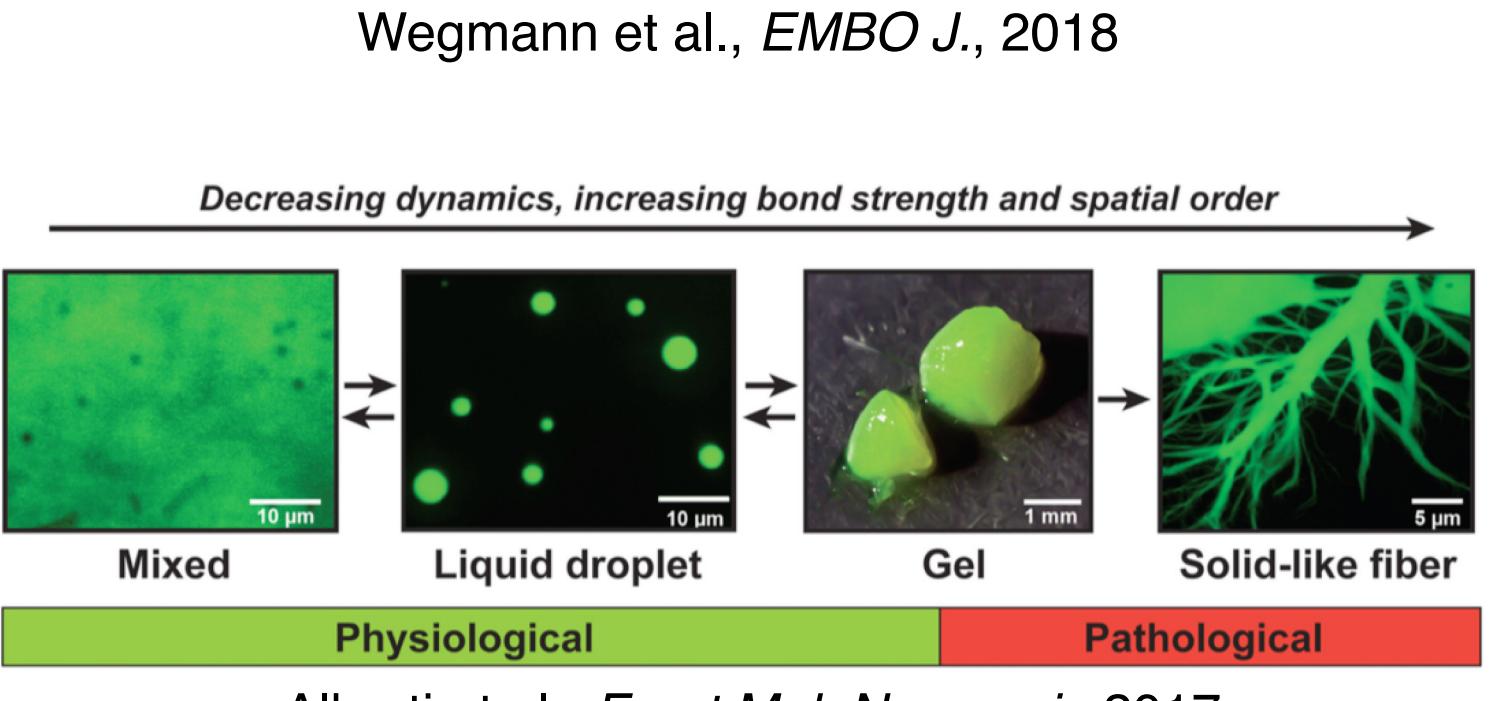
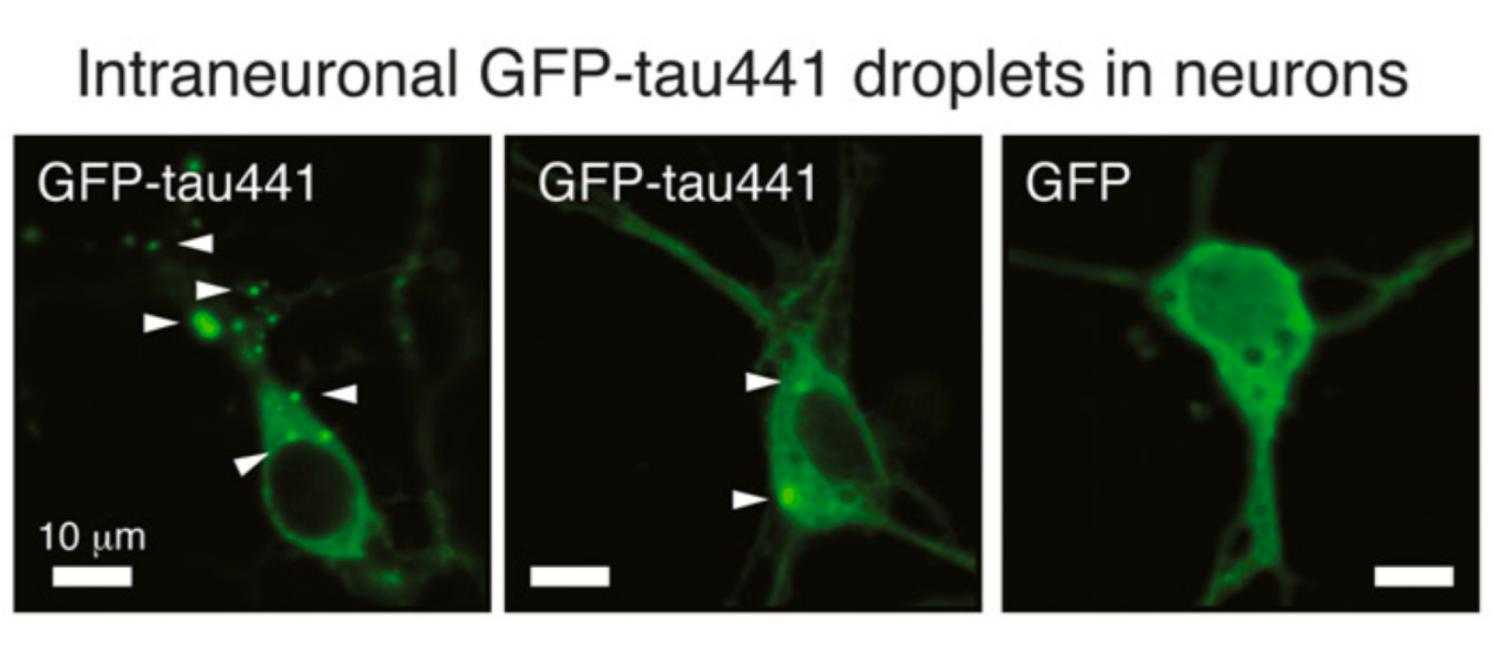
Intrinsically disordered proteins (IDPs) are functional even though they are unable to fold into a stable conformation. They are important for regulating many cellular processes including transcription, translation, and cellular signaling.



Intrinsically disordered regions (IDRs) often give proteins the ability to generate membrane-less organelles to localize biochemical processes in a process referred to as liquid-liquid phase separation.

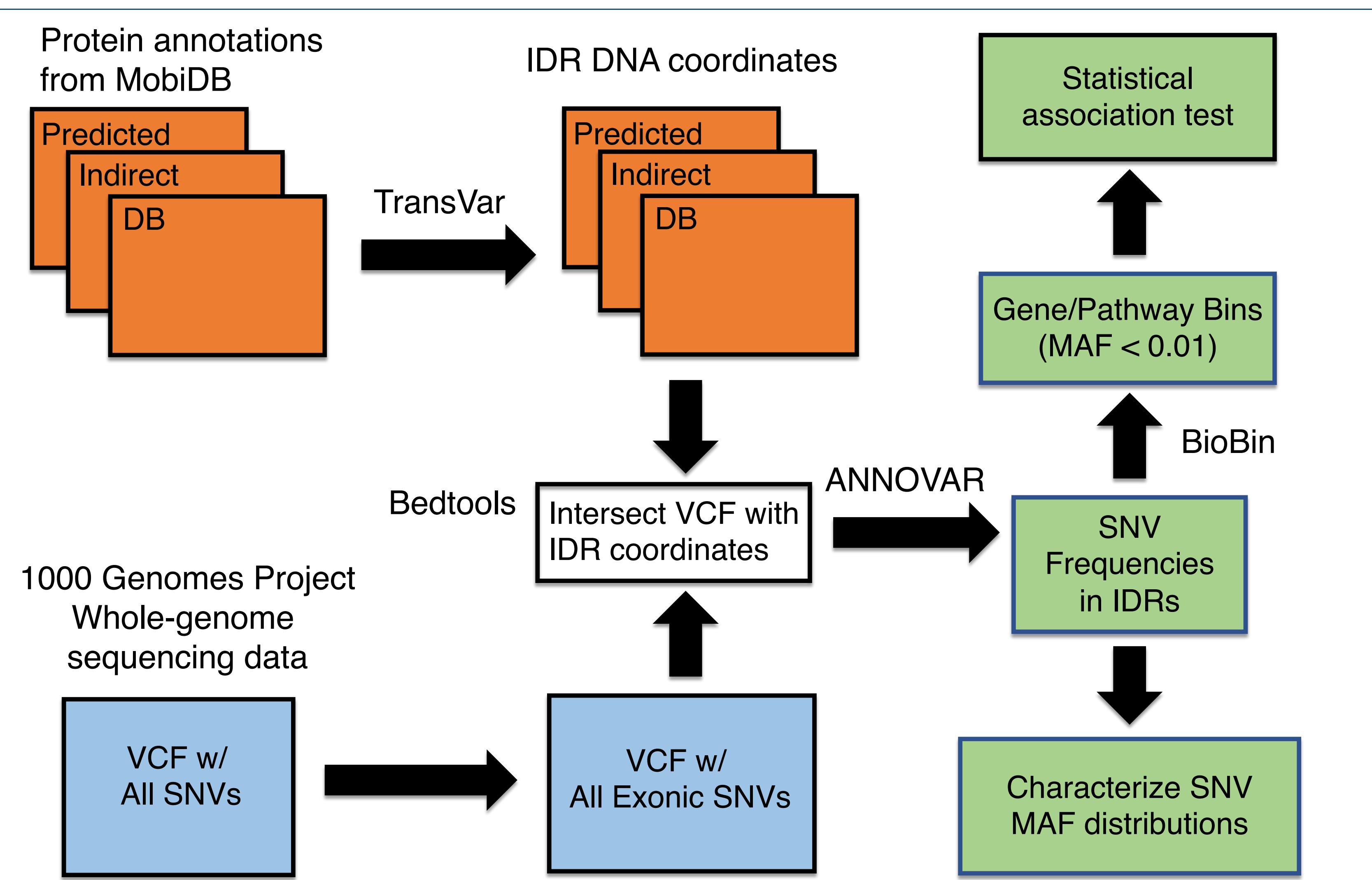


The natural ability of IDRs to mediate interactions with other proteins and generate phase transitions also underlies its connection to protein aggregation in multiple forms of neurodegeneration, including Alzheimer's disease.



**Hypothesis:** characterizing genetic variation in IDRs will lead to the identification of new AD risk loci.

## Linking intrinsically disordered protein regions to DNA coordinates



## Overview of MobiDB: a comprehensive database of IDR annotations

MobiDB is a database that contains three broad types of IDR information.

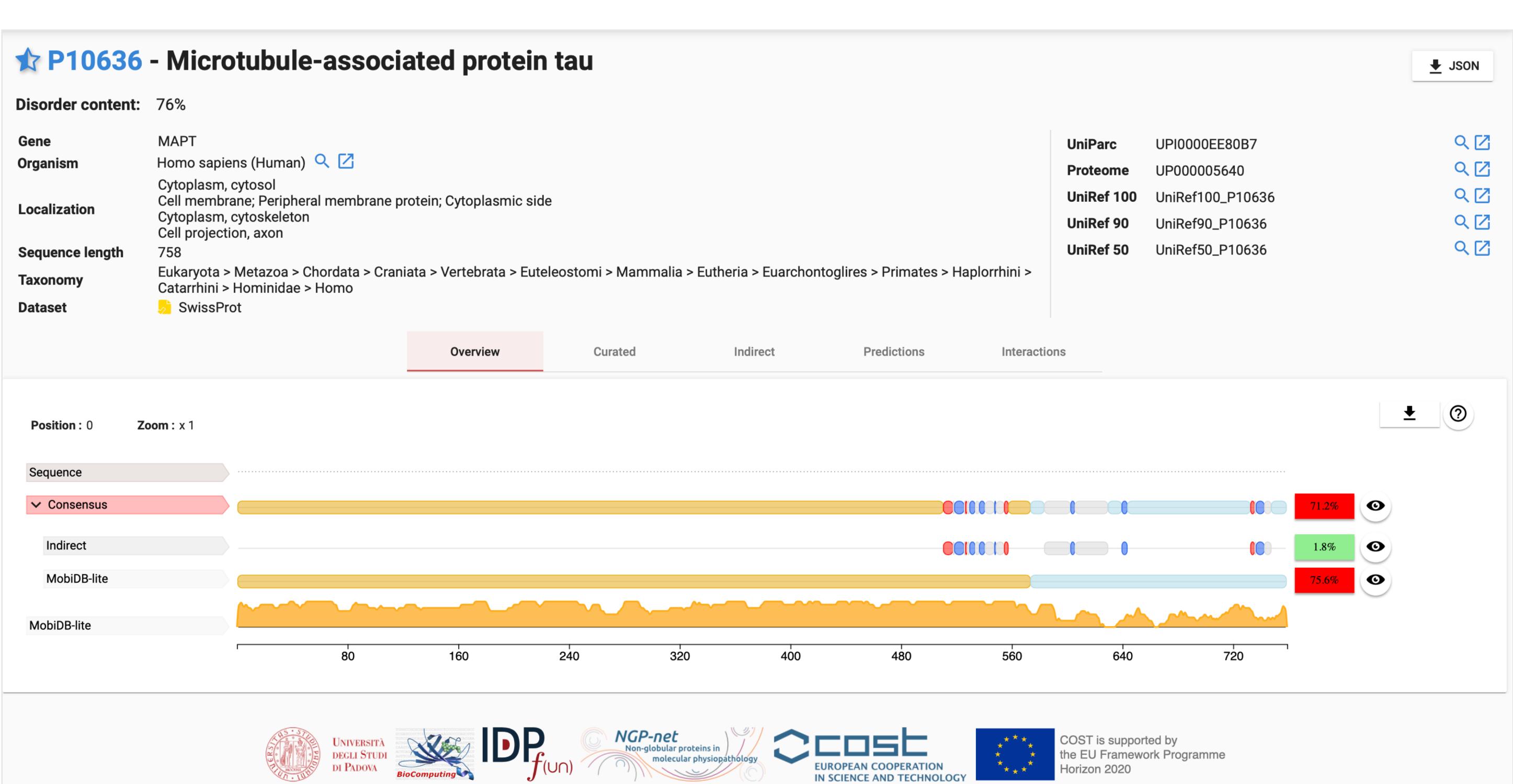
DB: manually curated structural information from databases containing experimental results (e.g. DisProt, FuzDB, and UniProt).

Indirect: information derived from PDB data using missing residues, high temperature residues (X-ray), and mobility (NMR).

Predicted: disorder prediction algorithms that use amino acid sequence information.

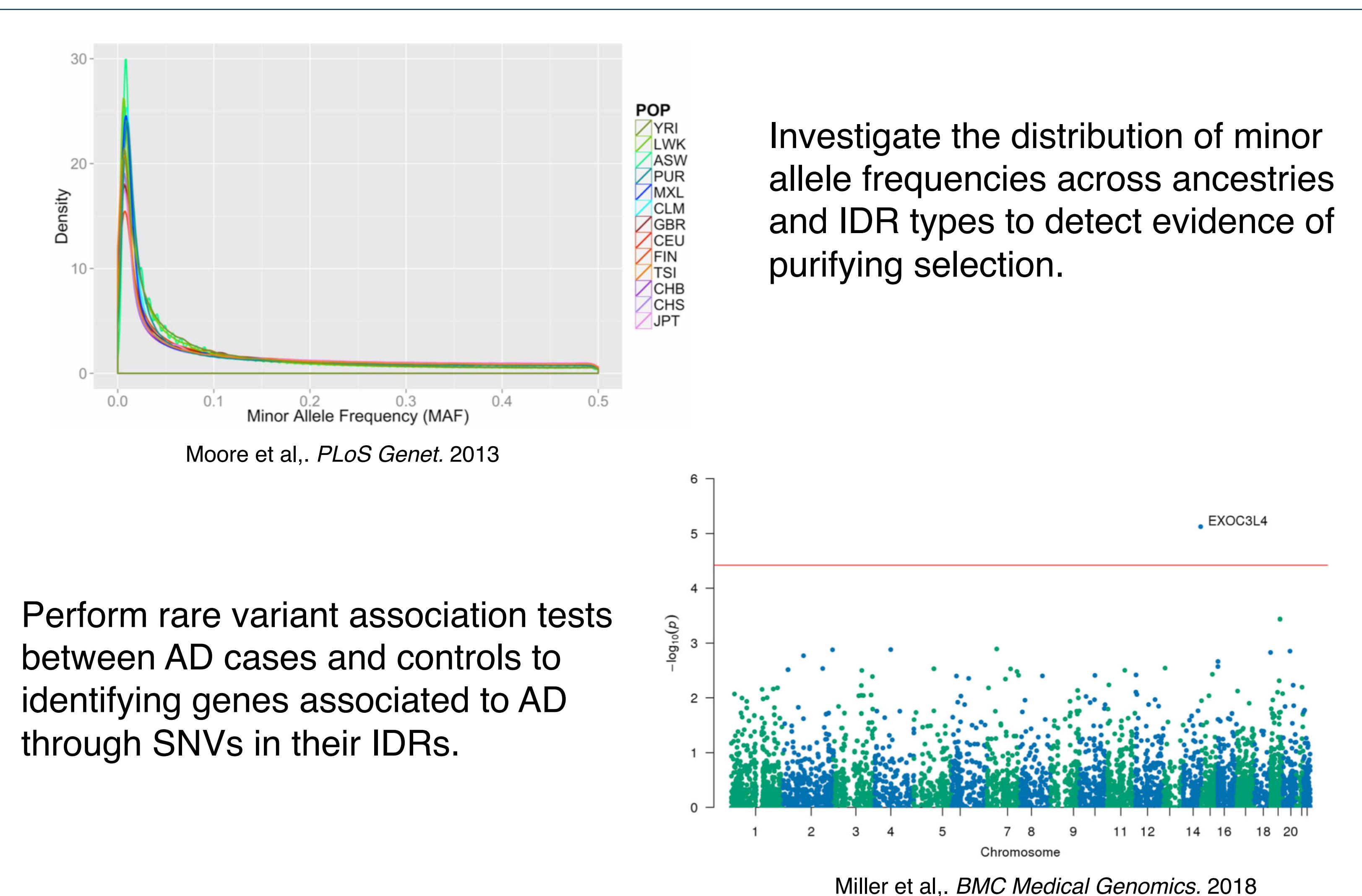
Quality ↑  
↓ Coverage

Piovesan et al., Nucleic Acids Research, 2018



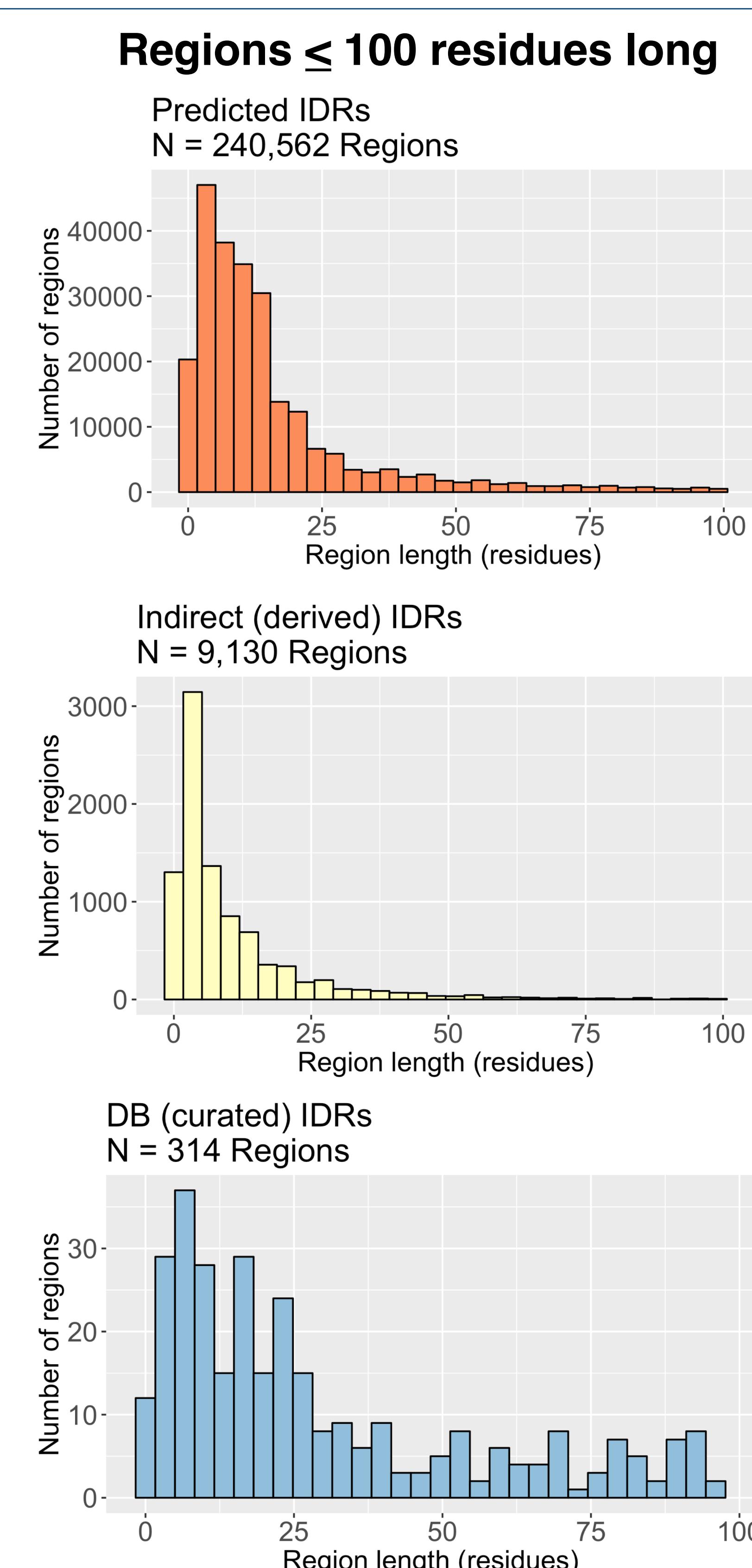
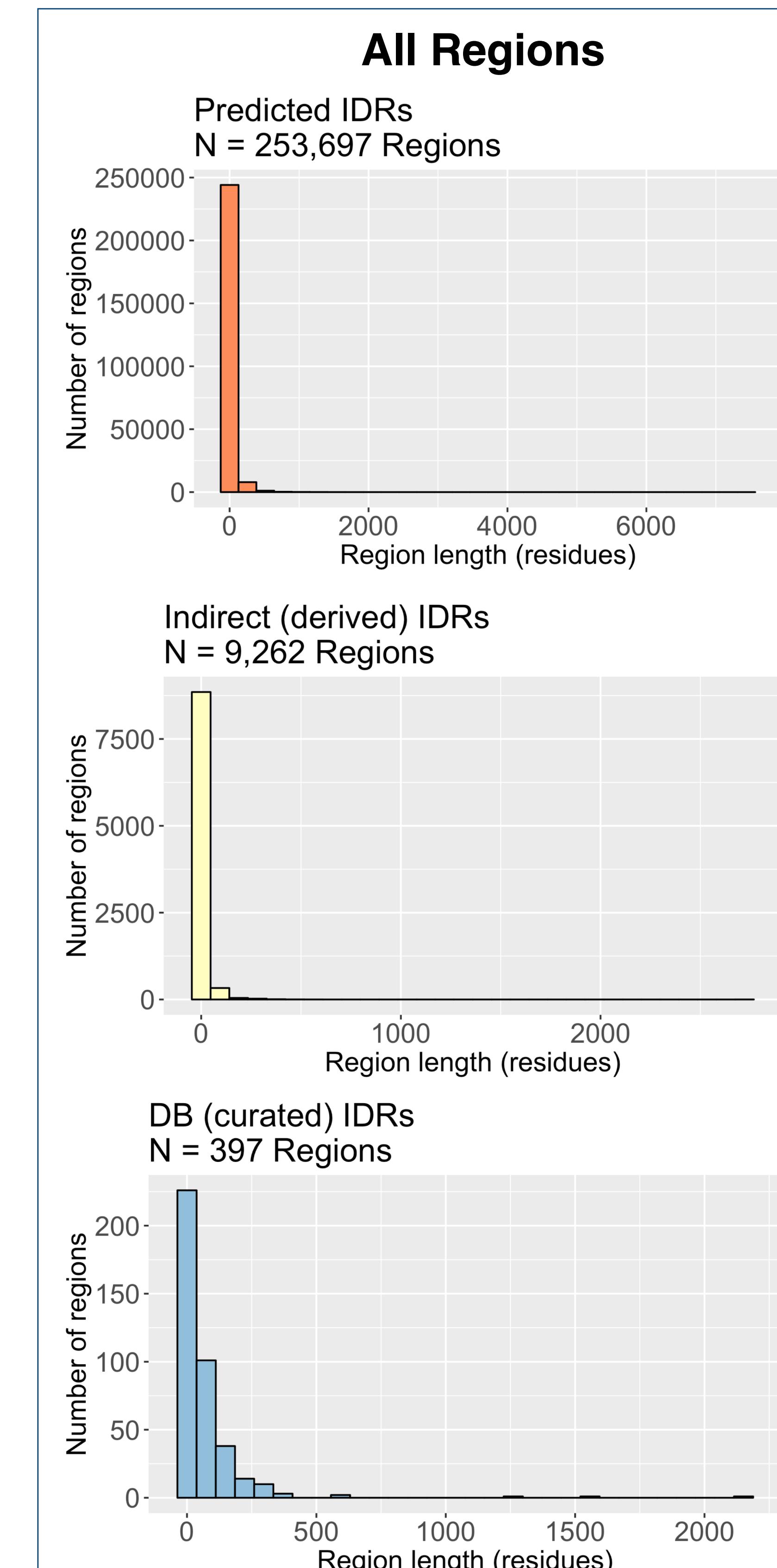
Screenshot of the human Tau protein as visualized in MobiDB. Disordered regions are in orange and red.

## Illustration of potential results



Perform rare variant association tests between AD cases and controls to identifying genes associated to AD through SNVs in their IDRs.

## Characterizing disordered regions using the MobiDB database



Protein IDs were filtered for Swiss-Prot reviewed proteins. All three types of disordered regions have long tails, but are primarily made up of sequences smaller than 100 residues. There are many more predicted regions than derived or curated regions.

## Future Questions

- How does genetic variation differ in known functional regions of IDRs?
- How does genetic variation in IDRs differ from more structured domains?
- Does variation differ between those with/without AD/MCI?
- How does genetic variation in IDRs differ in other diseases?

## Acknowledgments and links

