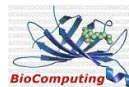# Critical Assessment of protein Intrinsic Disorder prediction (CAID) - Results of round 3

Silvio Tosatto

University of Padova

CASP-16, 1-4 December 2024, Punta Cana

People

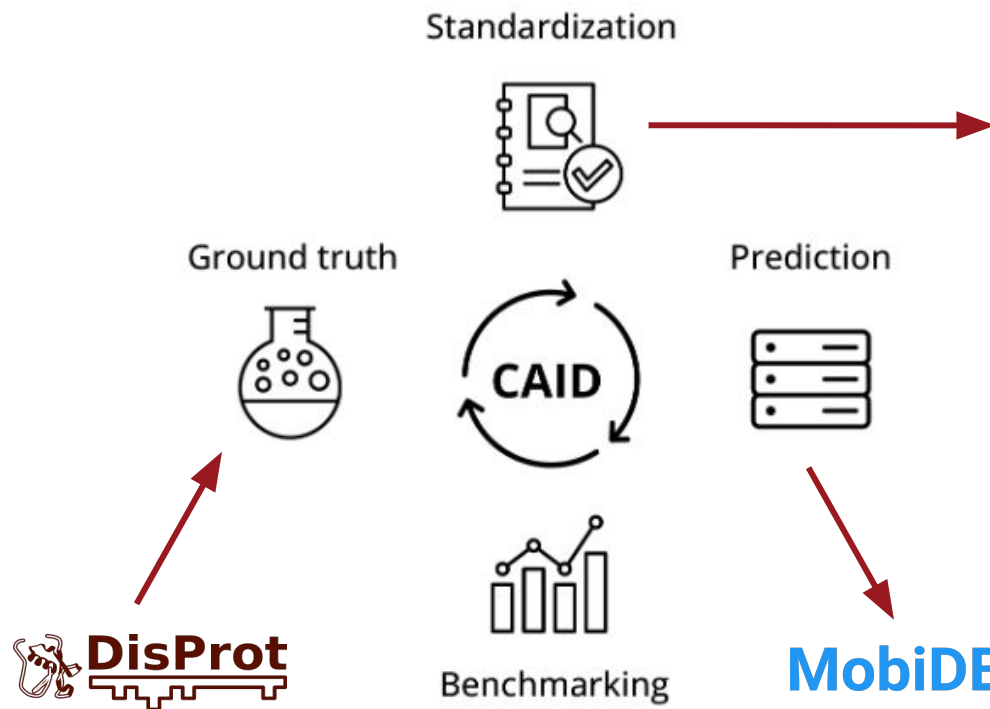**Organizers**
Silvio CE Tosatto
Damiano Piovesan

Assessors
Alessio Del Conte
Mahta Mehdiabadi

DisProt Curators
Victoria Nugnes
Federica Quaglia
Maria Cristina Aspromonte

The CAID cycle

Standardization

Ground truth

CAID

Prediction

Benchmarking

DisProt

MobiDB

CAID Prediction Portal

Sequence based prediction of **intrinsically disordered regions** (IDRs) and **binding regions** inside IDRs.

The server executes disorder prediction methods in parallel and applies optimized thresholds based on the results of the **Critical Assessment of Intrinsic protein Disorder (CAID)** - More info about the **CAID optimization strategy** Currently, the server can run **33 disorder** and **13 binding** methods - **Full list of methods**
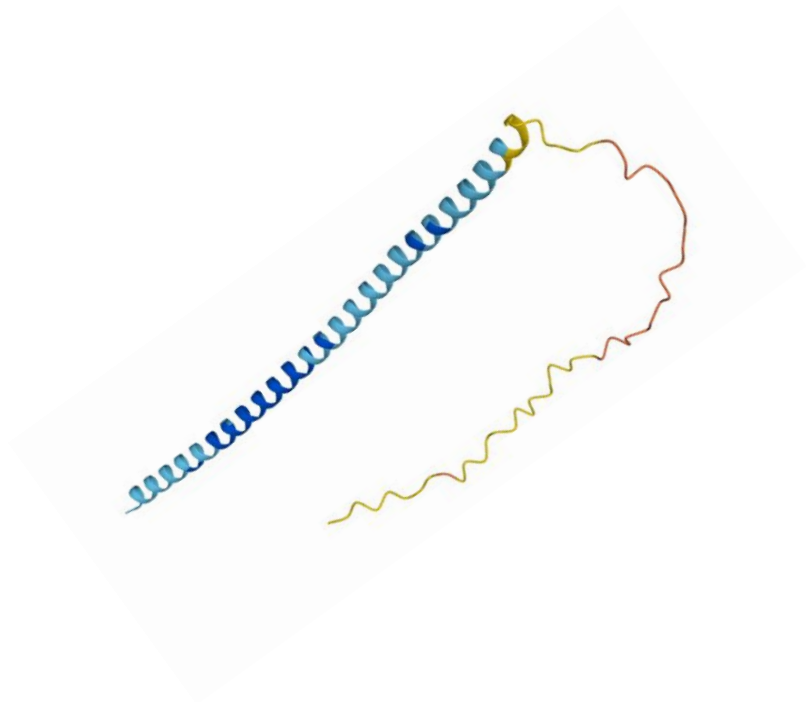
CAID Prediction Portal: A Comprehensive Service for Predicting Intrinsic Disorder and Binding Regions in Proteins
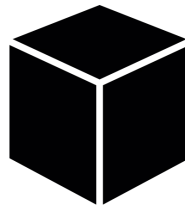*Del Conte et al.* **2023 Nucl Acids Res (Web Server Issue)**

https://caid.idpcentral.org/

3

CAID3 timeline
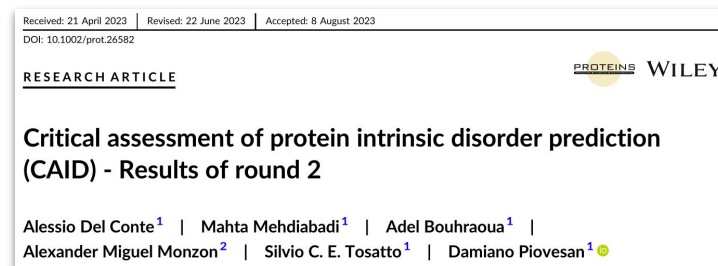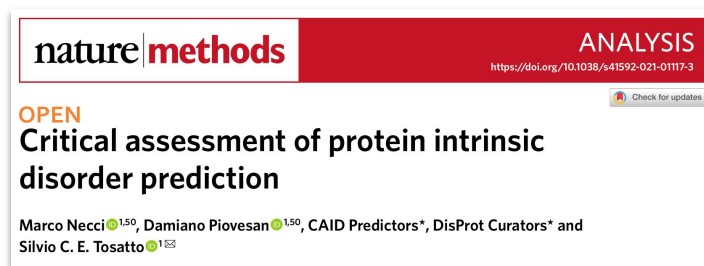
The challenge

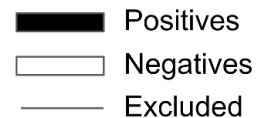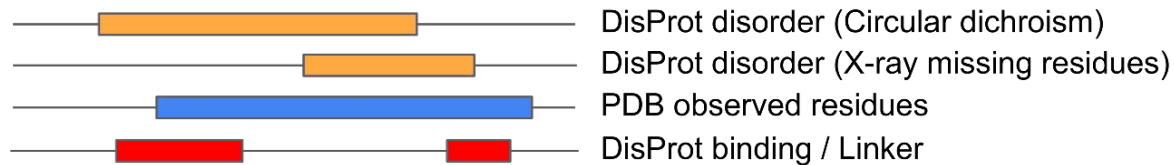**Sequence + Optional input**

Predictors

- **69 predictors**
  - 46 already in CAID2
  - **19 new predictors**
    - 15 submitted as Docker containers (DockerHub)
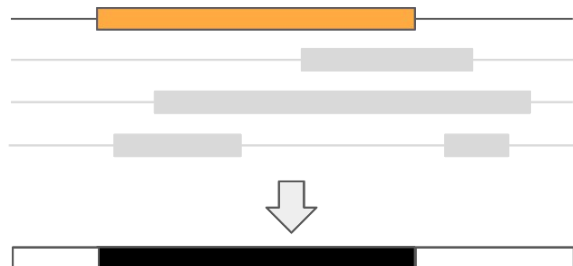   - Only 2 predictors had problems with installation (solved)

Additional inputs

- **PSI-Blast** on Uniref90 → PSSM

- **MMseqs2** on UniRef30 → MSA

- **HHblits** on Uniclust30_2017_10 → HMM

- **SPIDER2** → SS, ASA, torsion angles

- **Embeddings**

- Rost lab embeddings
  - **Prot-t5-xl-uniref50** used by 8 methods
  - **ProstT5** used by 1 method

- Facebook embeddings
  - **esm1b_t33_650M_UR50S** used by 1 method
  - **esm2_t33_650M_UR50D** used by 2 methods
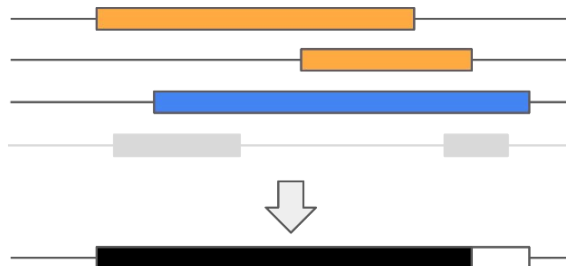 - **esm_msa1b_t12_100M_UR50S** used by 3 methods (embeddings from MSA)

Ground truth

DisProt disorder (Circular dichroism)
DisProt disorder (X-ray missing residues)
PDB observed residues
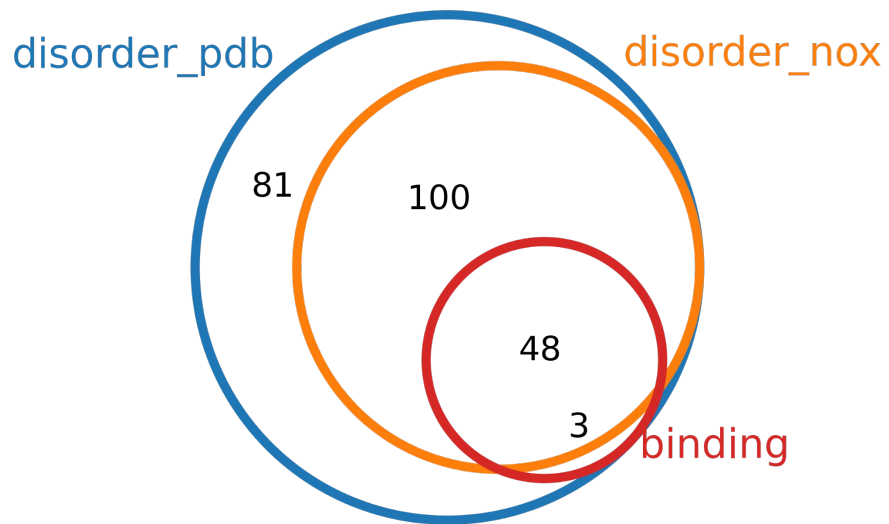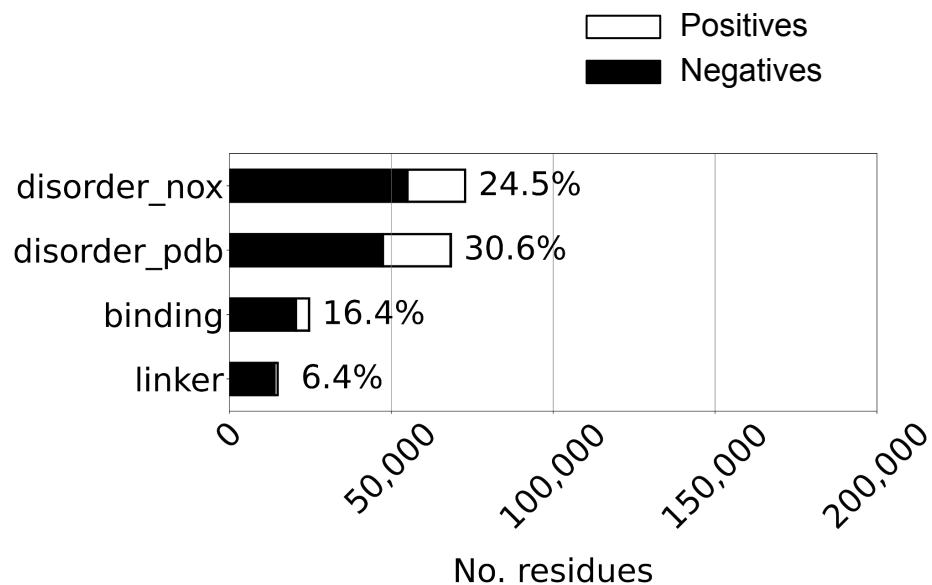DisProt binding / Linker

Positives
Negatives
Excluded

Disorder-NOX

Disorder-PDB

Binding / Linker

disorder_pdb

disorder_nox

81   100

48

3

binding

Disorder-PDB	232
Disorder-NOX	148
Binding	51
Linkers	20

7 isoforms!

Positives
Negatives

disorder_nox — 24.5%
disorder_pdb — 30.6%
binding — 16.4%
linker — 6.4%

0   50,000   100,000   150,000   200,000

No. residues

9

Results

Disorder PDB - top 10 predictors
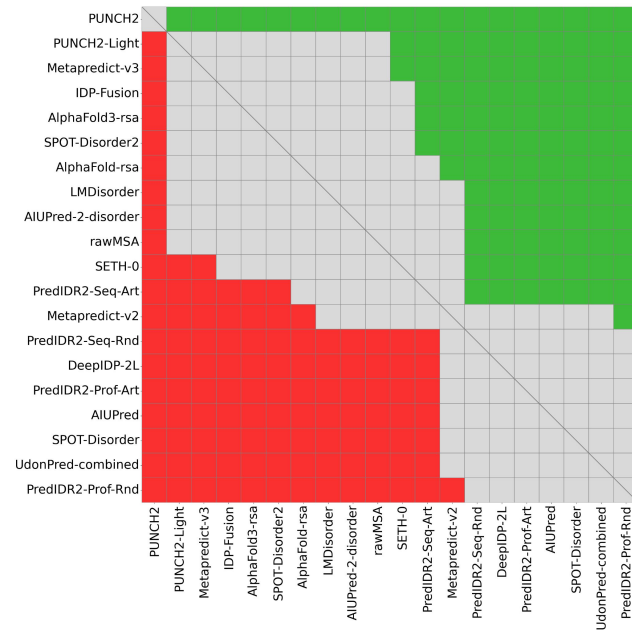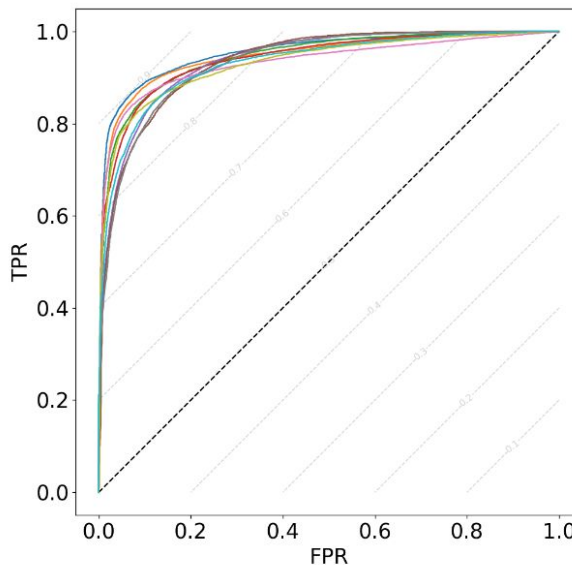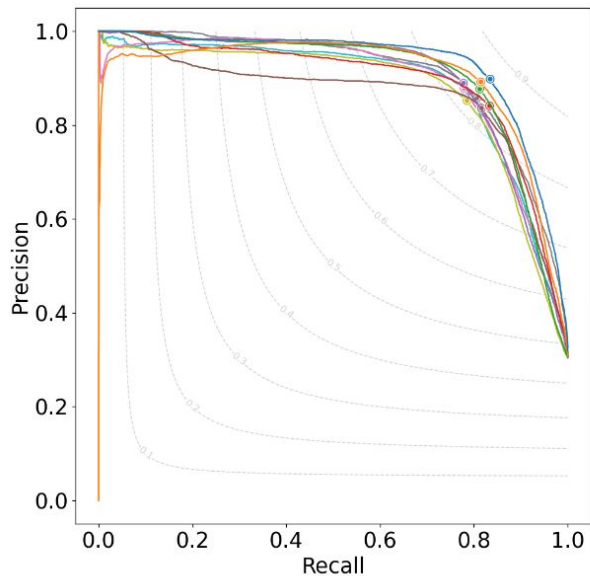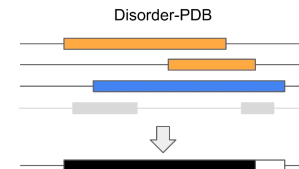
Disorder-PDB

— PUNCH2 - F-MAX: 0.86, C: 0.97	— UdonPred-combined - F-MAX: 0.83, C: 1.00	— PUNCH2 - AUC: 0.956
— AlphaFold-rsa - F-MAX: 0.85, C: 0.93	— AIUPred-2-disorder - F-MAX: 0.83, C: 1.00	— AlphaFold-rsa - AUC: 0.947
— SETH-0 - F-MAX: 0.84, C: 1.00	— SPOT-Disorder2 - F-MAX: 0.83, C: 0.94	— SPOT-Disorder2 - AUC: 0.945
— ESMDisPred-2PDB - F-MAX: 0.84, C: 0.85	— SETH-1 - F-MAX: 0.82, C: 1.00	— LMDisorder - AUC: 0.941
— Metapredict-v3 - F-MAX: 0.83, C: 1.00	— IDP-Fusion - F-MAX: 0.82, C: 0.97	— PredIDR2-Seq-Art - AUC: 0.939

| |
|---|
| — PredIDR2-Prof-Art - AUC: 0.935 |
| — SETH-0 - AUC: 0.933 |
| — PredIDR2-Seq-Rnd - AUC: 0.933 |
| — AIUPred-2-disorder - AUC: 0.932 |
| — rawMSA - AUC: 0.932 |

11

Time vs AUC of methods - Disorder PDB
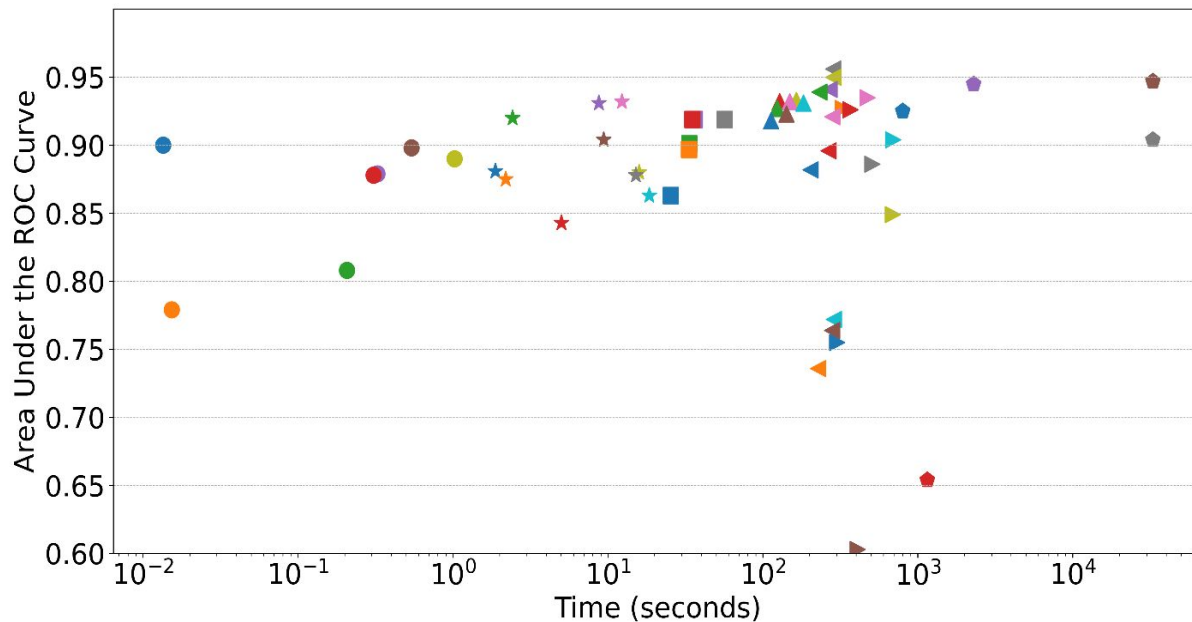
Top 25 by ROC AUC

◄ PUNCH2 - AUC: 0.956, 4.7m
◄ PUNCH2-Light - AUC: 0.950, 4.7m
● AlphaFold-disorder - AUC: 0.947, 9.1h
⬟ SPOT-Disorder2 - AUC: 0.945, 38.1m
◄ LMDisorder - AUC: 0.941, 4.5m
◄ PredIDR2-Seq - AUC: 0.939, 3.9m
► PredIDR2-Prof - AUC: 0.935, 7.9m
▲ SETH-0 - AUC: 0.933, 2.7m
★ AIUPred-2 - AUC: 0.932, 12.3s
▲ rawMSA - AUC: 0.932, 2.1m
▲ ESMDisPred - AUC: 0.932, 2.5m
▲ IDP-Fusion - AUC: 0.931, 3.0m
★ Metapredict - AUC: 0.931, 8.7s
► DisorderUnetLM - AUC: 0.927, 5.5m
▲ UdonPred - AUC: 0.927, 2.1m
► PredIDR - AUC: 0.926, 6.1m
⬟ SPOT-Disorder - AUC: 0.925, 13.3m
▲ DisoPred - AUC: 0.923, 2.4m
◄ DisoFLAG - AUC: 0.921, 4.7m
★ DARUMA-beta - AUC: 0.920, 2.4s
■ SPOT-Disorder-Single - AUC: 0.919, 36.4s
■ DeepIDP-2L - AUC: 0.919, 56.8s
■ SETH-1 - AUC: 0.919, 35.1s
▲ AUCpred - AUC: 0.918, 1.9m
► NeProc2-beta - AUC: 0.904, 11.5m

Disorder NOX - top 10 predictors
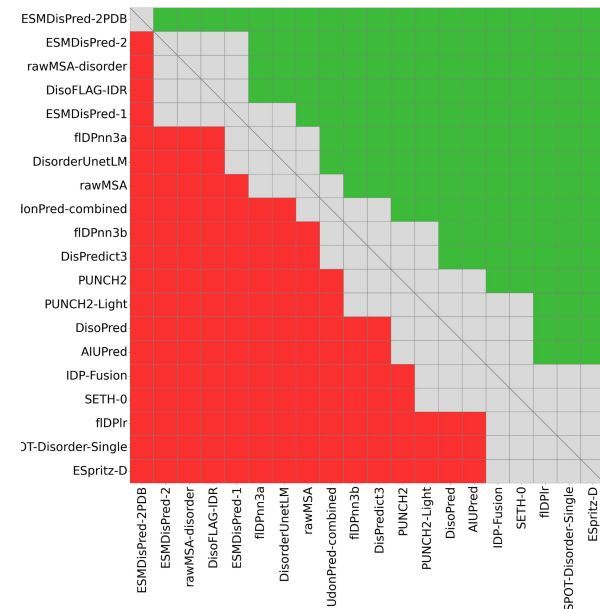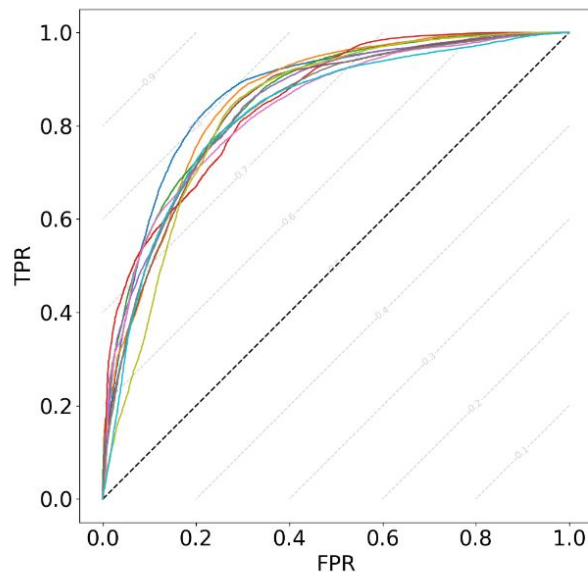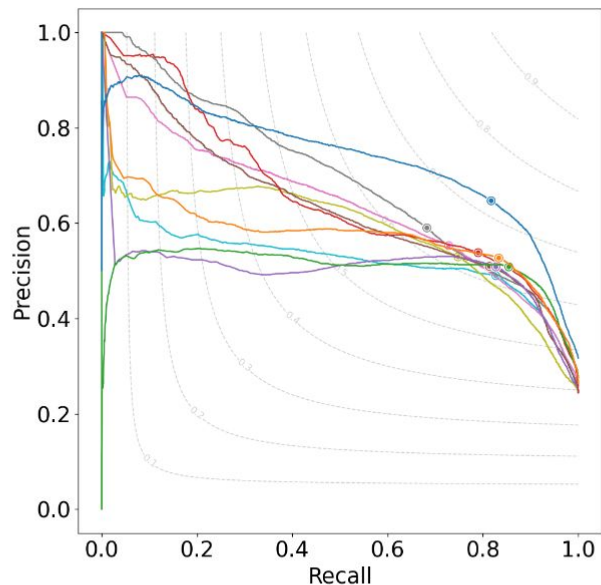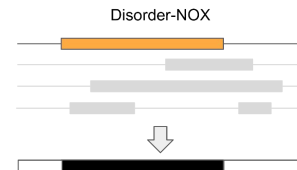
ESMDisPred-2PDB - F-MAX: 0.72, C: 0.86
PUNCH2 - F-MAX: 0.65, C: 0.98
AlphaFold-rsa - F-MAX: 0.64, C: 0.91
rawMSA-disorder - F-MAX: 0.64, C: 0.99
Metapredict-v3 - F-MAX: 0.63, C: 1.00

UdonPred-combined - F-MAX: 0.63, C: 0.99
DisPredict3 - F-MAX: 0.63, C: 0.96
DisoFLAG-IDR - F-MAX: 0.63, C: 0.97
AIUPred - F-MAX: 0.62, C: 0.99
AIUPred-2-disorder - F-MAX: 0.62, C: 1.00

ESMDisPred-2PDB - AUC: 0.868
rawMSA-disorder - AUC: 0.858
DisoFLAG-IDR - AUC: 0.857
DisorderUnetLM - AUC: 0.85
rawMSA - AUC: 0.847

UdonPred-combined - AUC: 0.841
flDPnn3b - AUC: 0.836
DisPredict3 - AUC: 0.835
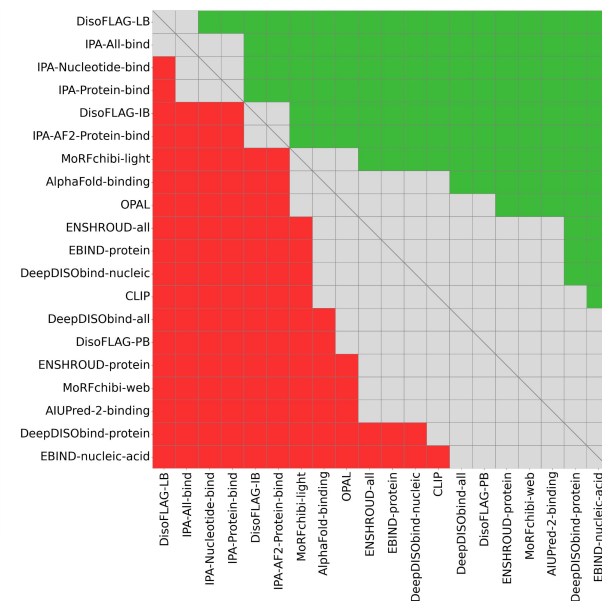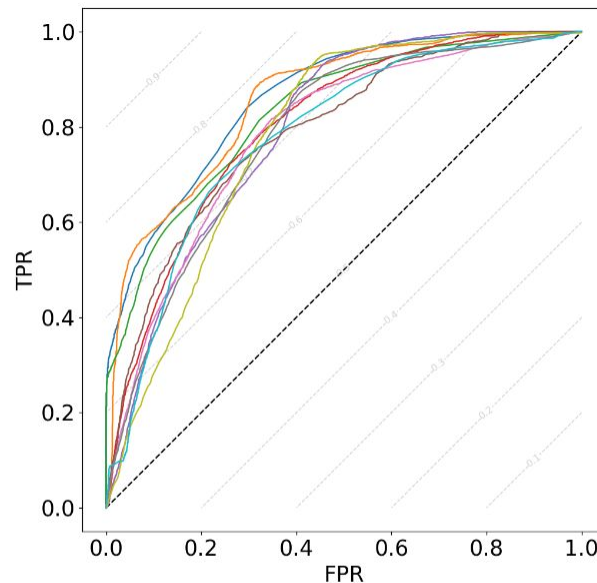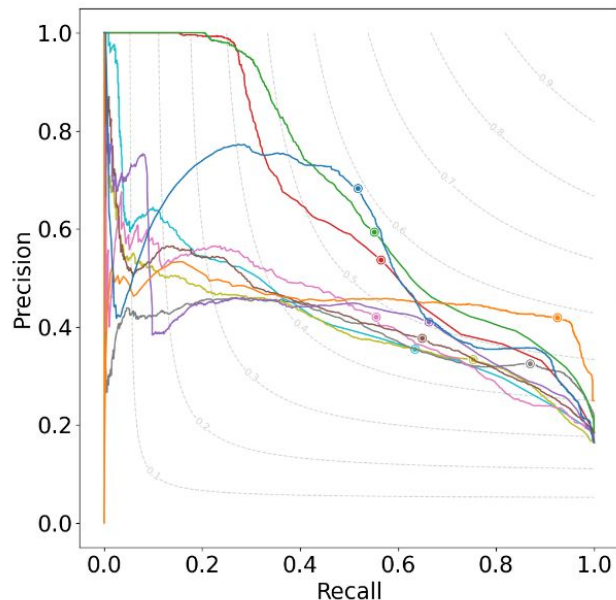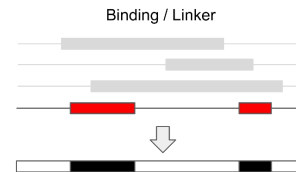PUNCH2 - AUC: 0.832
AIUPred - AUC: 0.828

13

Time vs AUC of methods - Disorder NOX
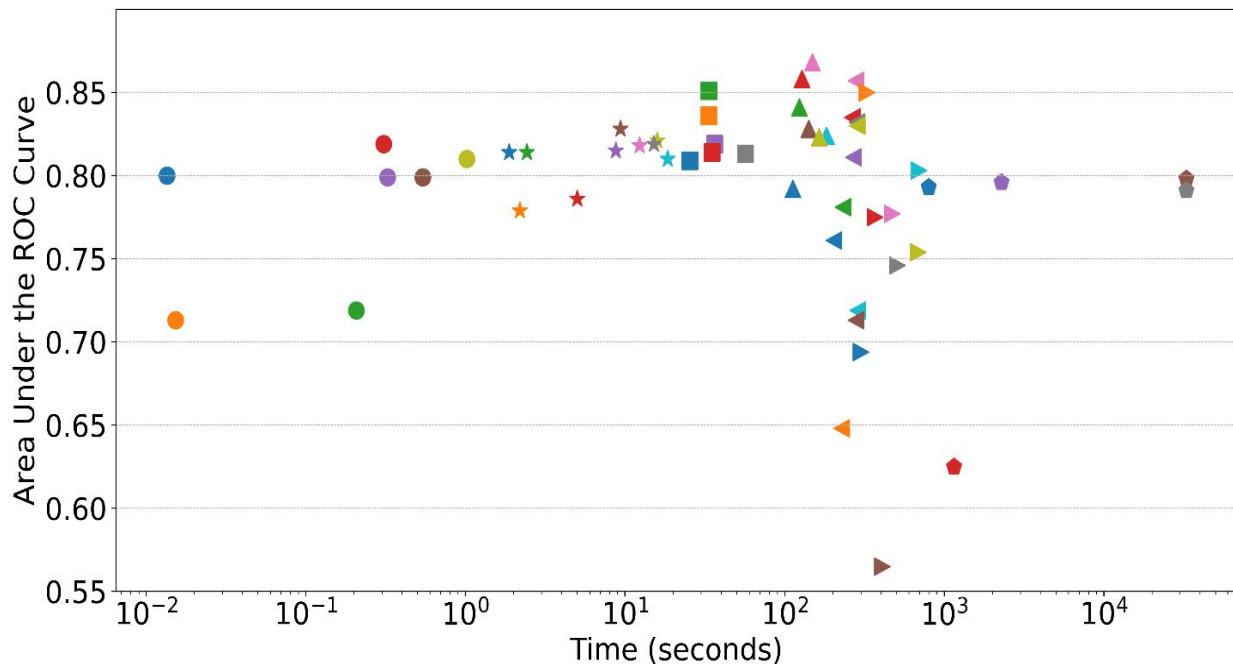
Top 25 by ROC AUC

- ESMDisPred - AUC: 0.868, 2.5m
- rawMSA - AUC: 0.858, 2.1m
- DisoFLAG - AUC: 0.857, 4.7m
- fIDPnn3a - AUC: 0.851, 33.5s
- DisorderUnetLM - AUC: 0.850, 5.5m
- UdonPred - AUC: 0.841, 2.1m
- fIDPnn3b - AUC: 0.836, 33.3s
- DisPredict3 - AUC: 0.835, 4.4m
- PUNCH2 - AUC: 0.832, 4.7m
- PUNCH2-Light - AUC: 0.830, 4.7m
- AIUPred - AUC: 0.828, 9.4s
- DisoPred - AUC: 0.828, 2.4m
- IDP-Fusion - AUC: 0.824, 3.0m
- SETH-0 - AUC: 0.823, 2.7m
- fIDPlr - AUC: 0.821, 15.9s
- SPOT-Disorder-Single - AUC: 0.819, 36.4s
- ESpritz - AUC: 0.819, 307.7ms
- fIDPnn2 - AUC: 0.819, 15.2s
- AIUPred-2 - AUC: 0.818, 12.3s
- Metapredict - AUC: 0.815, 8.7s
- SETH-1 - AUC: 0.814, 35.1s
- DARUMA-beta - AUC: 0.814, 2.4s
- pyHCA - AUC: 0.814, 1.9s
- DeepIDP-2L - AUC: 0.813, 56.8s
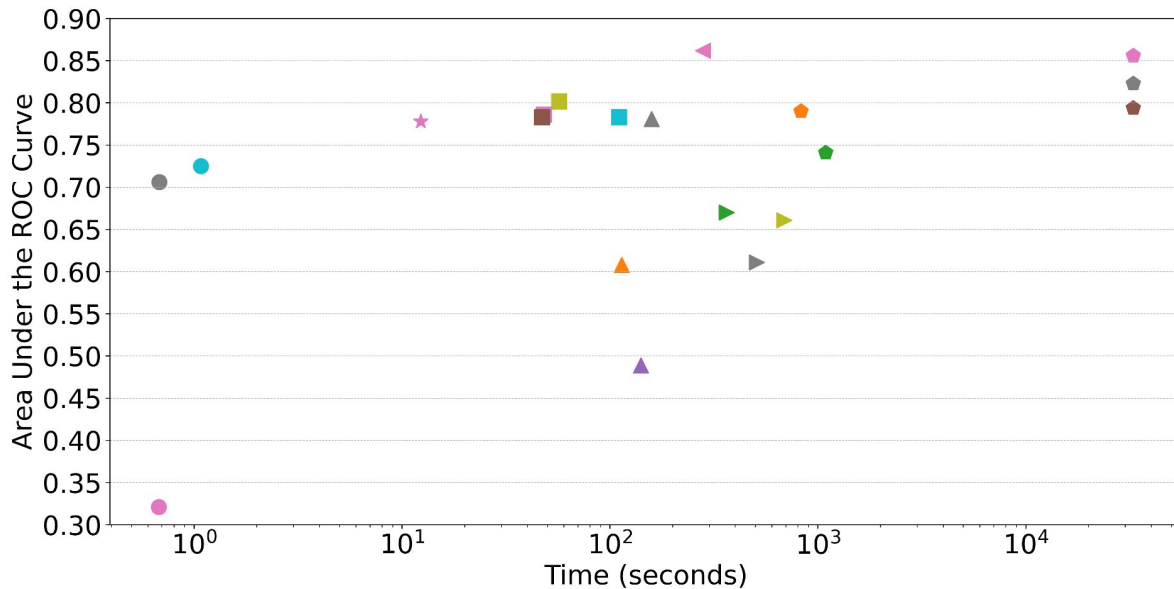- LMDisorder - AUC: 0.811, 4.5m

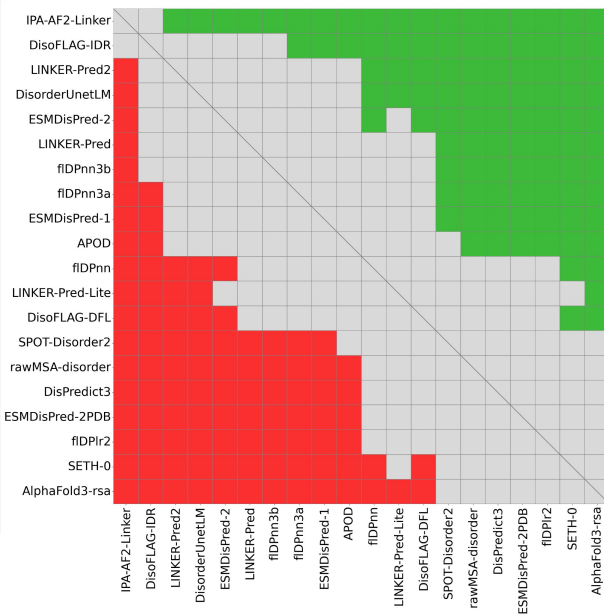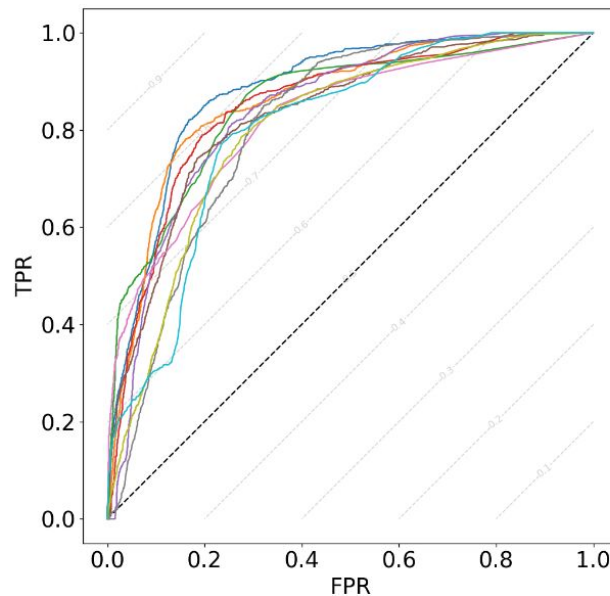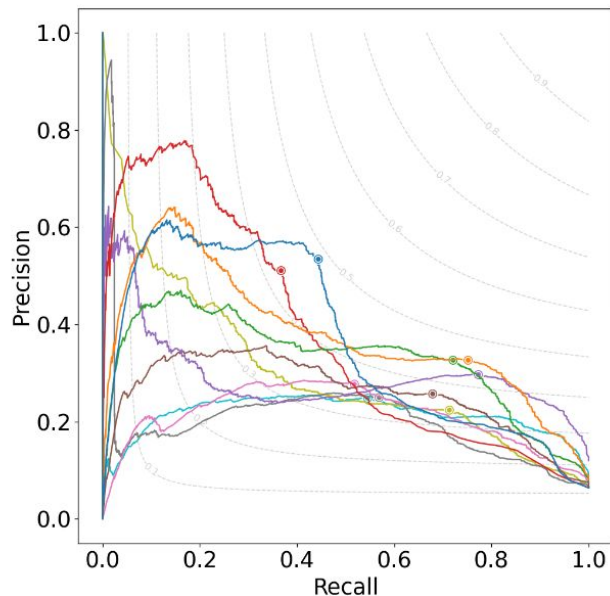# Binding - top 10 predictors
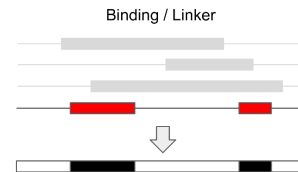
Time vs AUC of methods - Binding

All methods by ROC AUC

- DisoFLAG - AUC: 0.862, 4.7m
- IPA - AUC: 0.856, 9.1h
- IPA-AF2 - AUC: 0.823, 9.1h
- MoRFchibi - AUC: 0.802, 57.0s
- AlphaFold-disorder - AUC: 0.794, 9.1h
- OPAL - AUC: 0.790, 13.8m
- ENSHROUD - AUC: 0.786, 48.0s
- EBIND - AUC: 0.783, 1.8m
- DeepDISObind - AUC: 0.783, 47.0s
- CLIP - AUC: 0.781, 2.6m
- AIUPred-2 - AUC: 0.778, 12.3s
- DRPBind - AUC: 0.741, 18.1m
- ANCHOR2 - AUC: 0.725, 1.1s
- DisoRDPbind - AUC: 0.706, 680.3ms
- LIPNet - AUC: 0.670, 6.1m
- NeProc - AUC: 0.661, 11.5m
- DISOPRED3 - AUC: 0.611, 8.5m
- ProBiPred - AUC: 0.608, 1.9m
- bindEmbed21IDR - AUC: 0.489, 2.3m
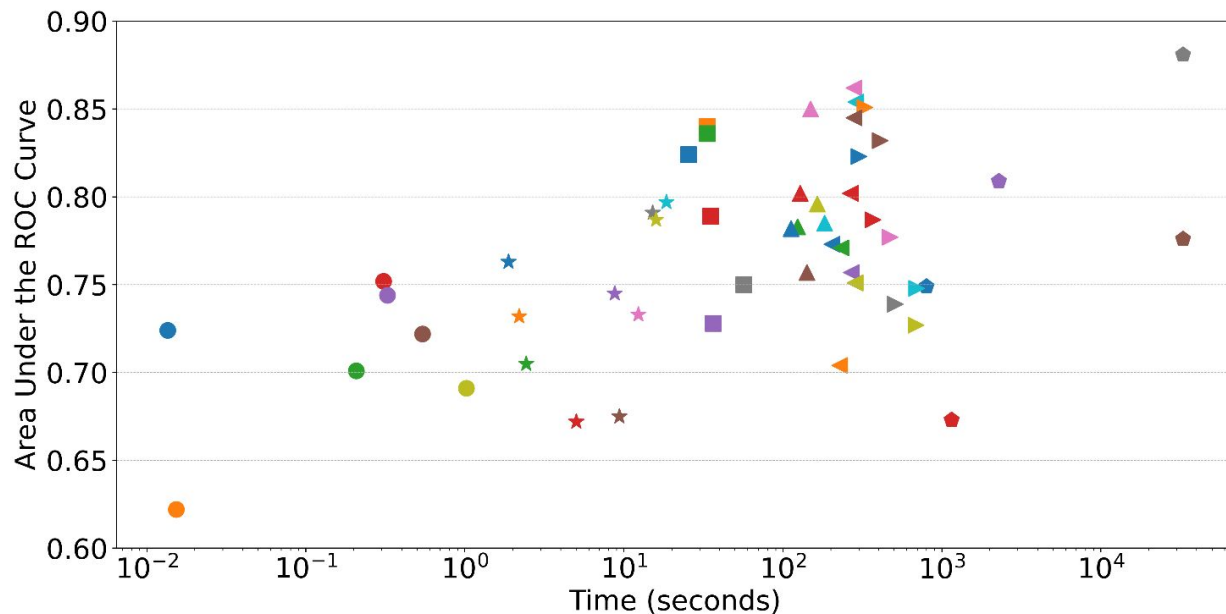- DFLpred - AUC: 0.321, 676.2ms

Linker - top 10 predictors
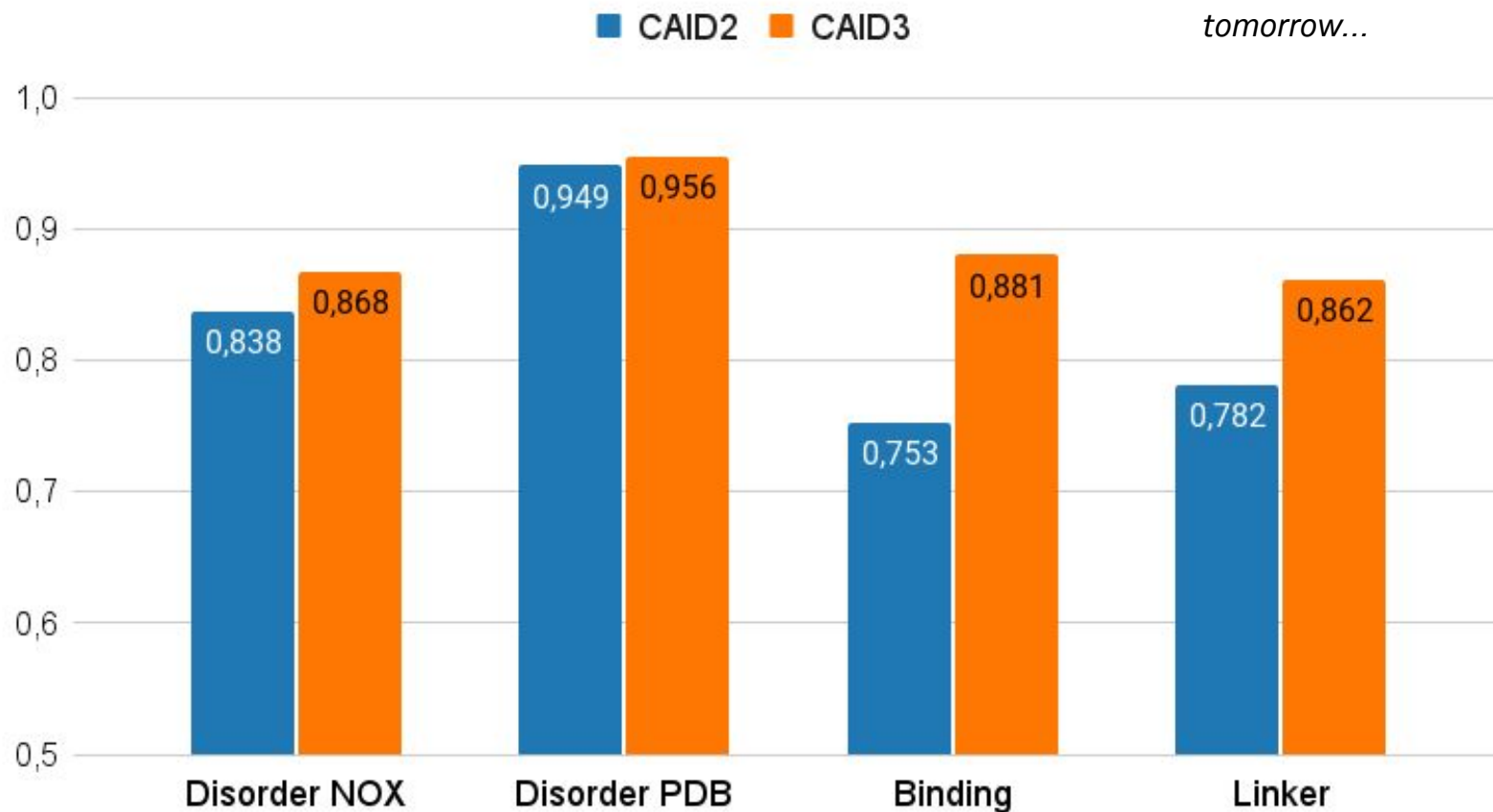
Time vs AUC of methods - Linkers

Top 25 methods by ROC AUC

- ⬠ IPA-AF2 - AUC: 0.881, 9.1h
- ◀ DisoFLAG - AUC: 0.862, 4.7m
- ◀ LINKER-Pred2 - AUC: 0.854, 4.8m
- ▶ DisorderUnetLM - AUC: 0.851, 5.5m
- ▲ ESMDisPred - AUC: 0.850, 2.5m
- ▶ LINKER-Pred - AUC: 0.845, 4.7m
- ■ fIDPnn3b - AUC: 0.840, 33.3s
- ■ fIDPnn3a - AUC: 0.836, 33.5s
- ▶ APOD - AUC: 0.832, 6.8m
- ■ fIDPnn - AUC: 0.824, 25.5s
- ▶ LINKER-Pred-Lite - AUC: 0.823, 5.0m
- ⬠ SPOT-Disorder2 - AUC: 0.809, 38.1m
- ▲ rawMSA - AUC: 0.802, 2.1m
- ◀ DisPredict3 - AUC: 0.802, 4.4m
- ★ fIDPlr2 - AUC: 0.797, 18.5s
- ▲ SETH-0 - AUC: 0.796, 2.7m
- ★ fIDPnn2 - AUC: 0.791, 15.2s
- ■ SETH-1 - AUC: 0.789, 35.1s
- ★ fIDPlr - AUC: 0.787, 15.9s
- ▶ PredIDR - AUC: 0.787, 6.1m
- ▲ IDP-Fusion - AUC: 0.785, 3.0m
- ▲ UdonPred - AUC: 0.783, 2.1m
- ▲ AUCpred - AUC: 0.782, 1.9m
- ▶ PredIDR2-Prof - AUC: 0.777, 7.9m
- ⬠ AlphaFold-disorder - AUC: 0.776, 9.1h

18

CAID2 vs CAID3 max ROC AUC

More CAID-3 results on Damiano's poster tomorrow...

Legend: CAID2, CAID3
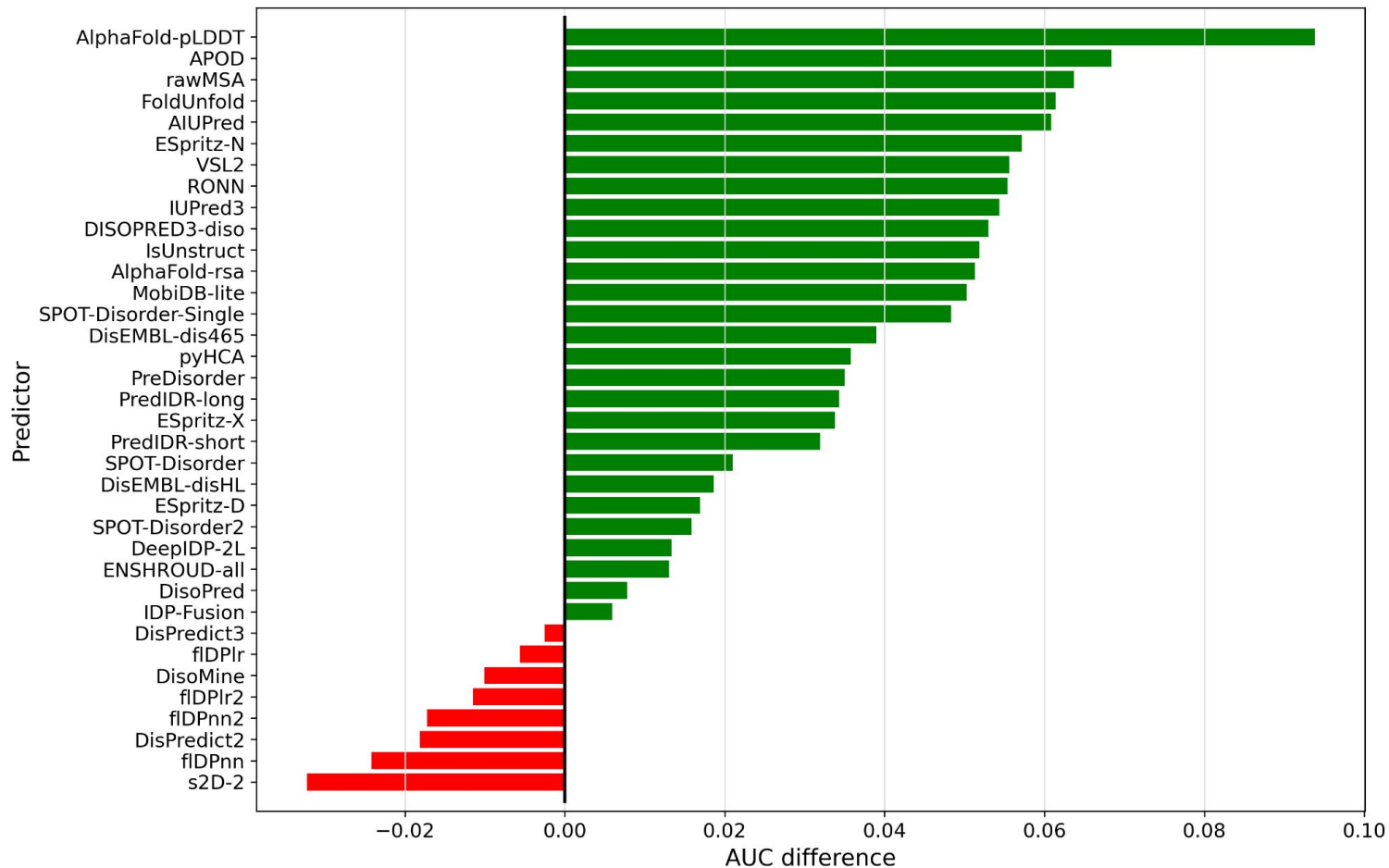
Category	CAID2	CAID3
Disorder NOX	0,838	0,868
Disorder PDB	0,949	0,956
Binding	0,753	0,881
Linker	0,782	0,862

Acknowledgements

BioComputing UP
University of Padova
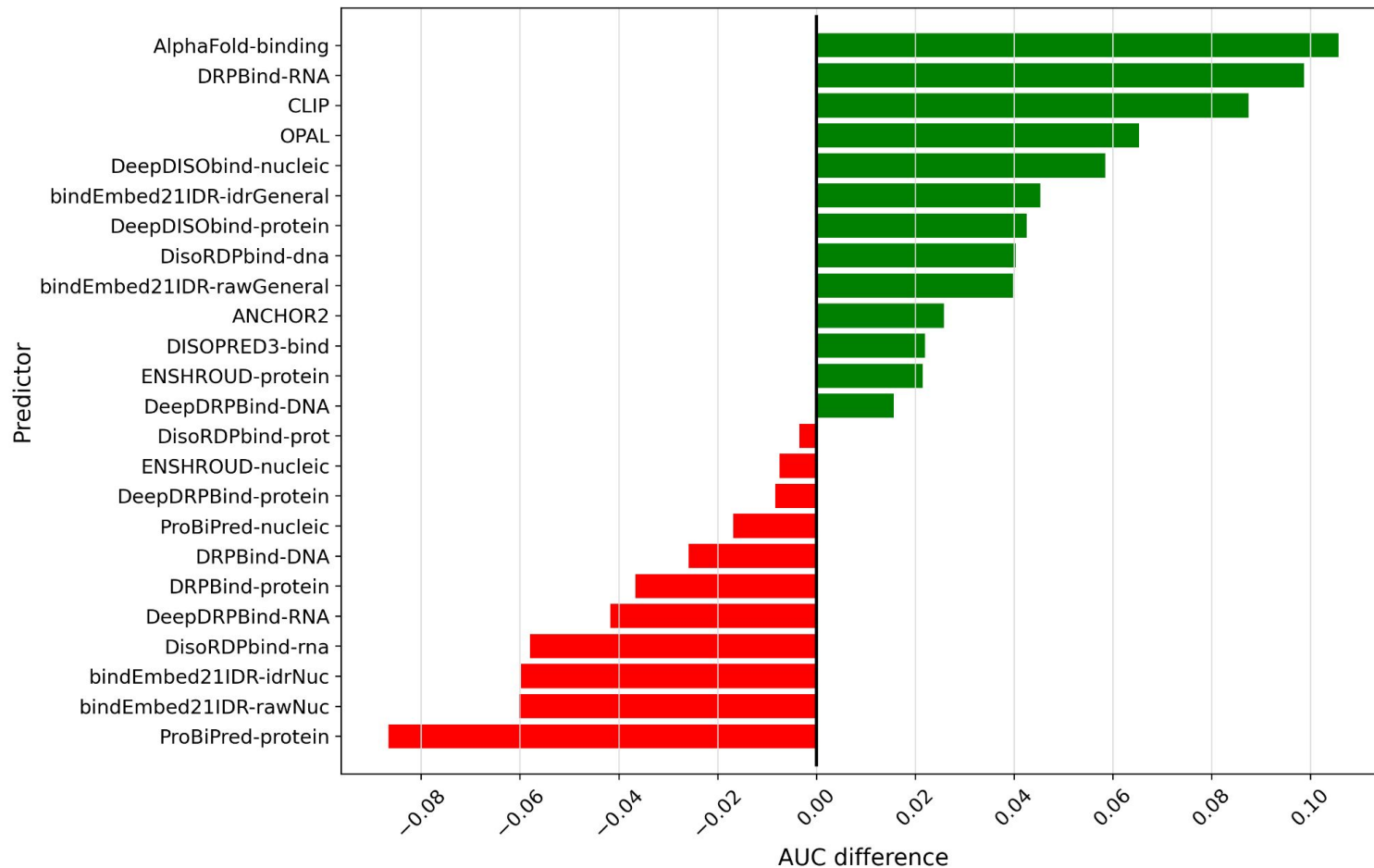Dept. of Biomedical Sciences

https://biocomputingup.it

𝕏 @BioComputingUP

Difference in AUC between CAID3 and CAID2 predictors (CAID3 - CAID2) on Disorder NOX reference
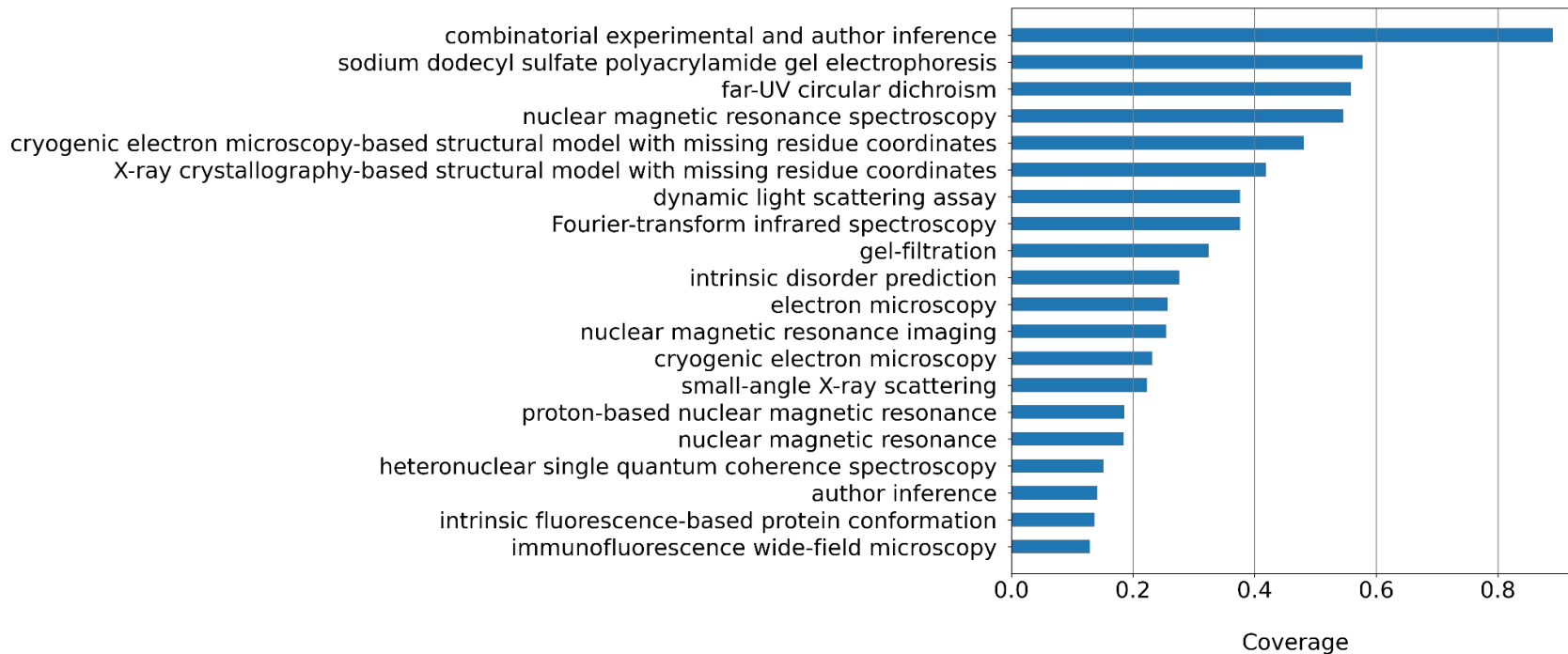
Difference in AUC between CAID3 and CAID2 predictors (CAID3 - CAID2) on Binding reference
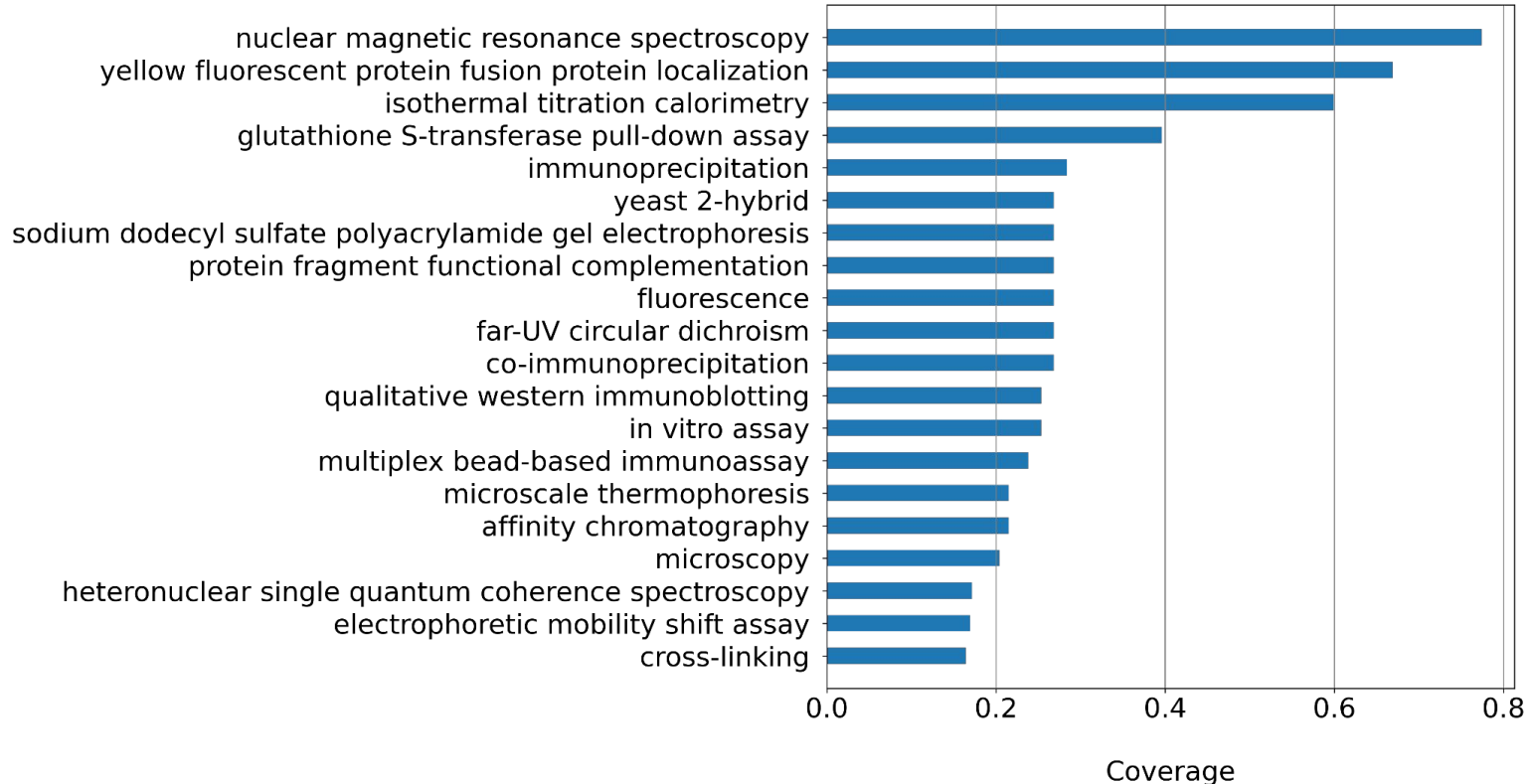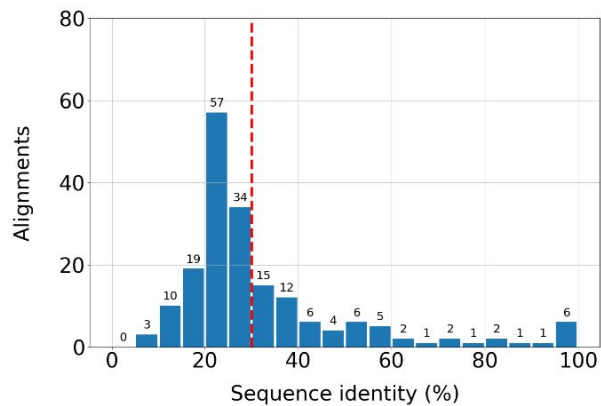
Disorder evidence

Binding evidence

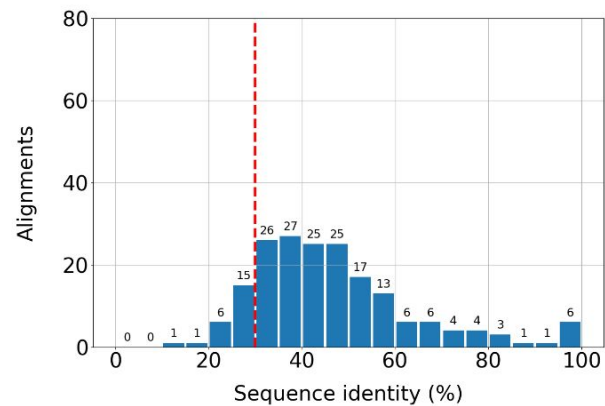Smith–Waterman · Needleman–Wunsch
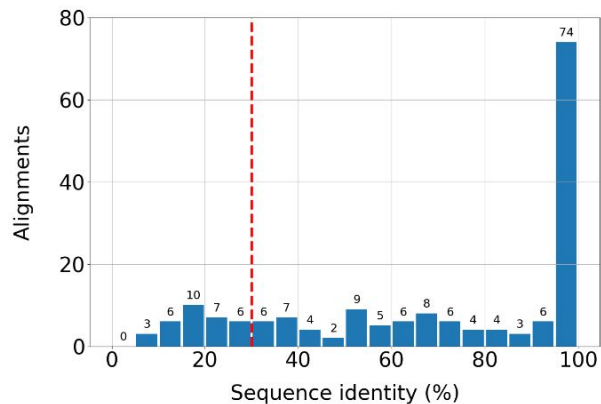
CAID3 Vs. DisProt-Old - Local · CAID3 Vs. DisProt-Old - Global · CAID3 Vs. PDB-Seqres - Local · CAID3 Vs. PDB-Seqres - Global