Computational prediction of intrinsic disorder

FROM HUMBLE BEGINNINGS TO MODERN RESOURCES

VCU
College of Engineering
http://biomine.cs.vcu.edu

Lukasz Kurgan
Virginia Commonwealth University
PSB 2020
January 7th, 2019

Intrinsic disorder

Intrinsically disordered proteins[1] **(IDPs)** have at least one **intrinsically disordered region (IDR)**.
IDRs lack a stable structure in isolation and take form of conformational ensembles. "*Intrinsic*" reflects the fact that the disordered state is encoded in the sequence.

15 superimposed NMR structures of the methylosome subunit of spliceosomal assembly protein ICLN
(chaperone for spliceosomal small nuclear ribonucleoproteins)

[1] also called floppy, pliable, rheomorphic, flexible, mobile, partially folded, natively denatured, natively unfolded, natively disordered, intrinsically unstructured, intrinsically denatured, intrinsically unfolded, …

Intrinsic disorder

Abundant in nature

- computational studies estimate that 30 to 50% of eukaryotic proteins have one or more long IDR

Involved in numerous molecular functions

- signaling, regulation, assembly, molecular recognition (proteins and nucleic acids), translation and transcription

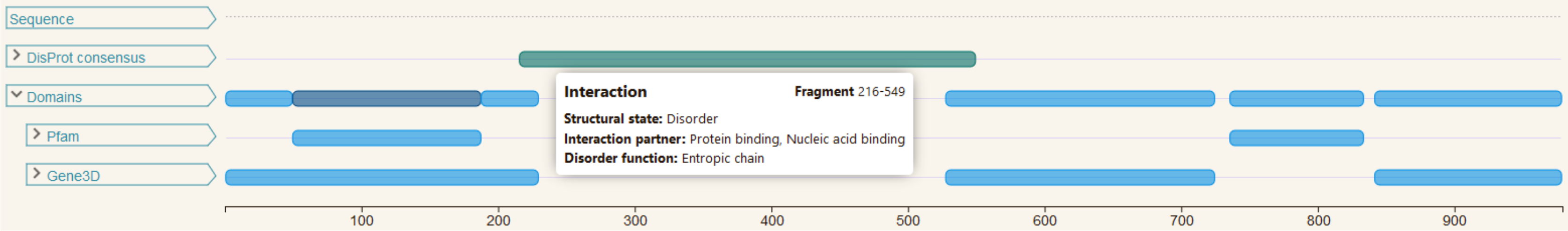
Only about 1,600 IDPs (3,500 IDRs) are annotated experimentally

- experimental data are used to develop and validate computational models
- computational prediction often plays a key role to produce initial hypothesis about presence and function of disorder

Prediction of intrinsic disorder

P06701 (UniPRot) / DP00530 (DisPRot): **silent information regulator (Sir3p) from yeast**

Native annotation



Interaction **Fragment** 216-549

Structural state: Disorder

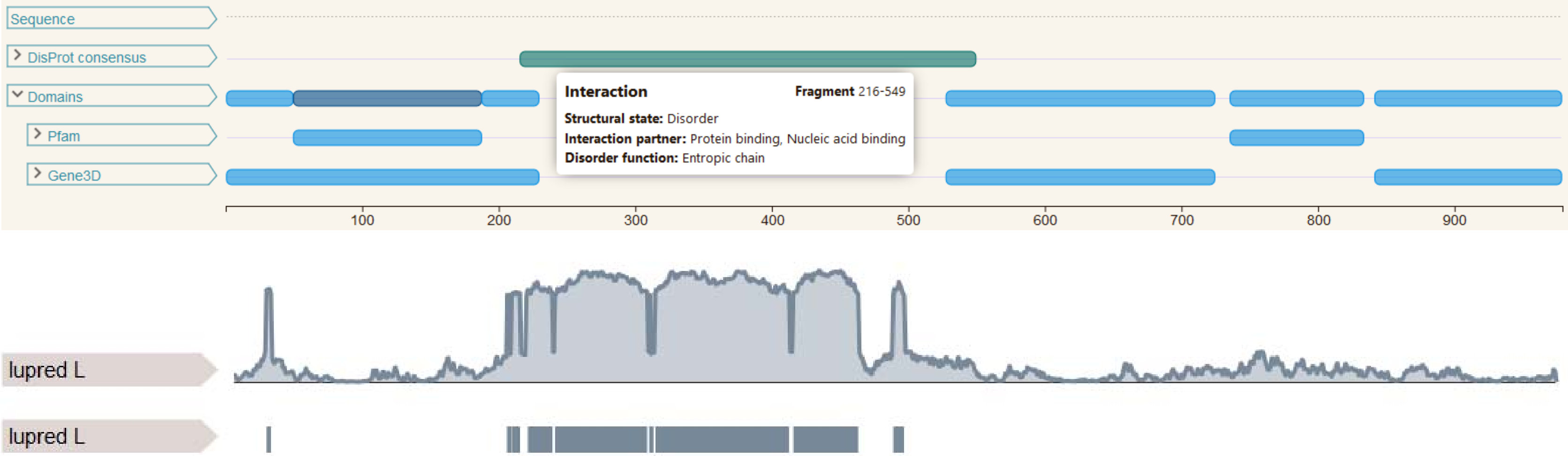
Interaction partner: Protein binding, Nucleic acid binding

Disorder function: Entropic chain

Prediction of intrinsic disorder

P06701 (UniPRot) / DP00530 (DisPRot): **silent information regulator (Sir3p) from yeast**

Native annotation

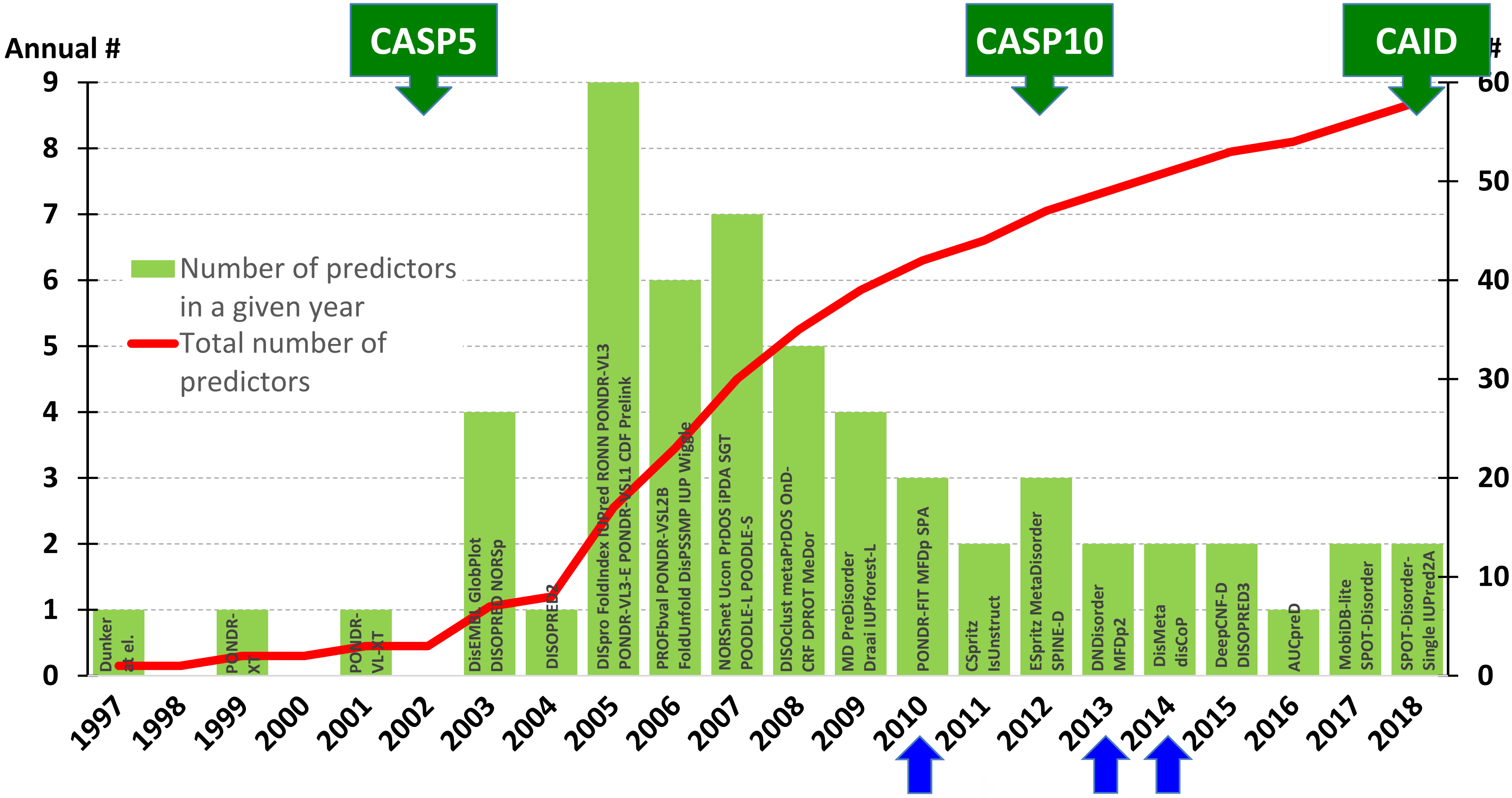


Interaction **Fragment** 216-549
Structural state: Disorder
Interaction partner: Protein binding, Nucleic acid binding
Disorder function: Entropic chain

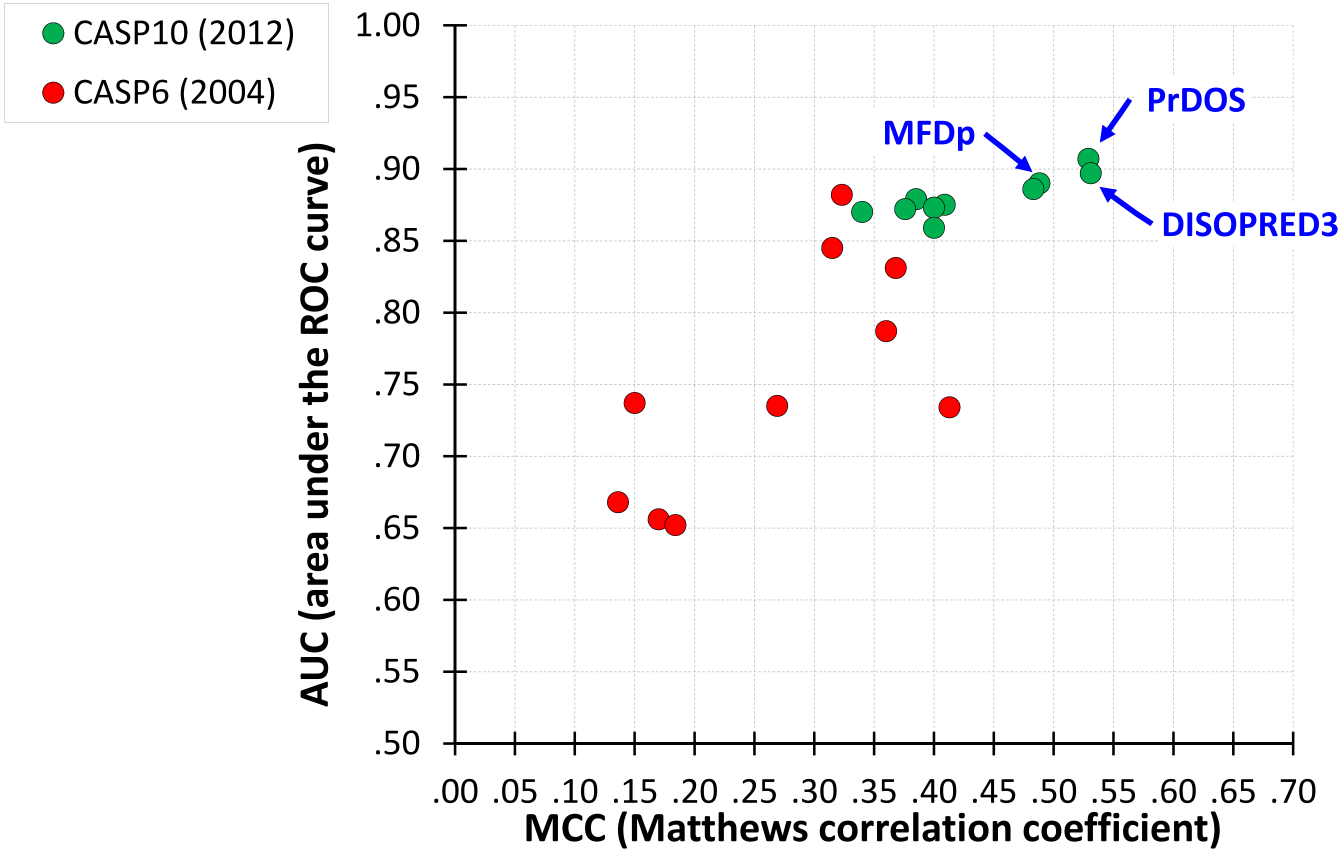
Prediction from sequence

Dosztányi Z, Csizmók V, Tompa P, Simon I. The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Molecular Biology* 347:827-39, 2005

Prediction of intrinsic disorder



Prediction of intrinsic disorder

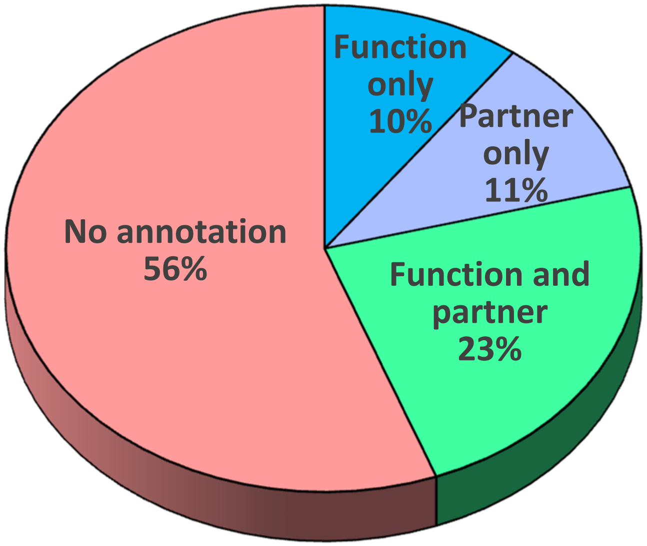


Jin Y, Dunbrack RL. Assessment of disorder predictions in CASP6. *Proteins* 61(Suppl 7):167-75, 2005
Monastyrskyy B, Kryshtafovych A, Moult J, Tramontano A, Fidelis K. Assessment of protein disorder region predictions in CASP10. *Proteins* 82(Suppl 2):127-37, 2014

Functions of intrinsic disorder

Molecular partners

Protein (417 IDRs)
DNA (76 IDRs)
RNA (33 IDRs)
Lipid (30 IDRs)
Metal (39 IDRs)
Small molecule (32 IDRs)
Inorganic Salt (1 IDR)

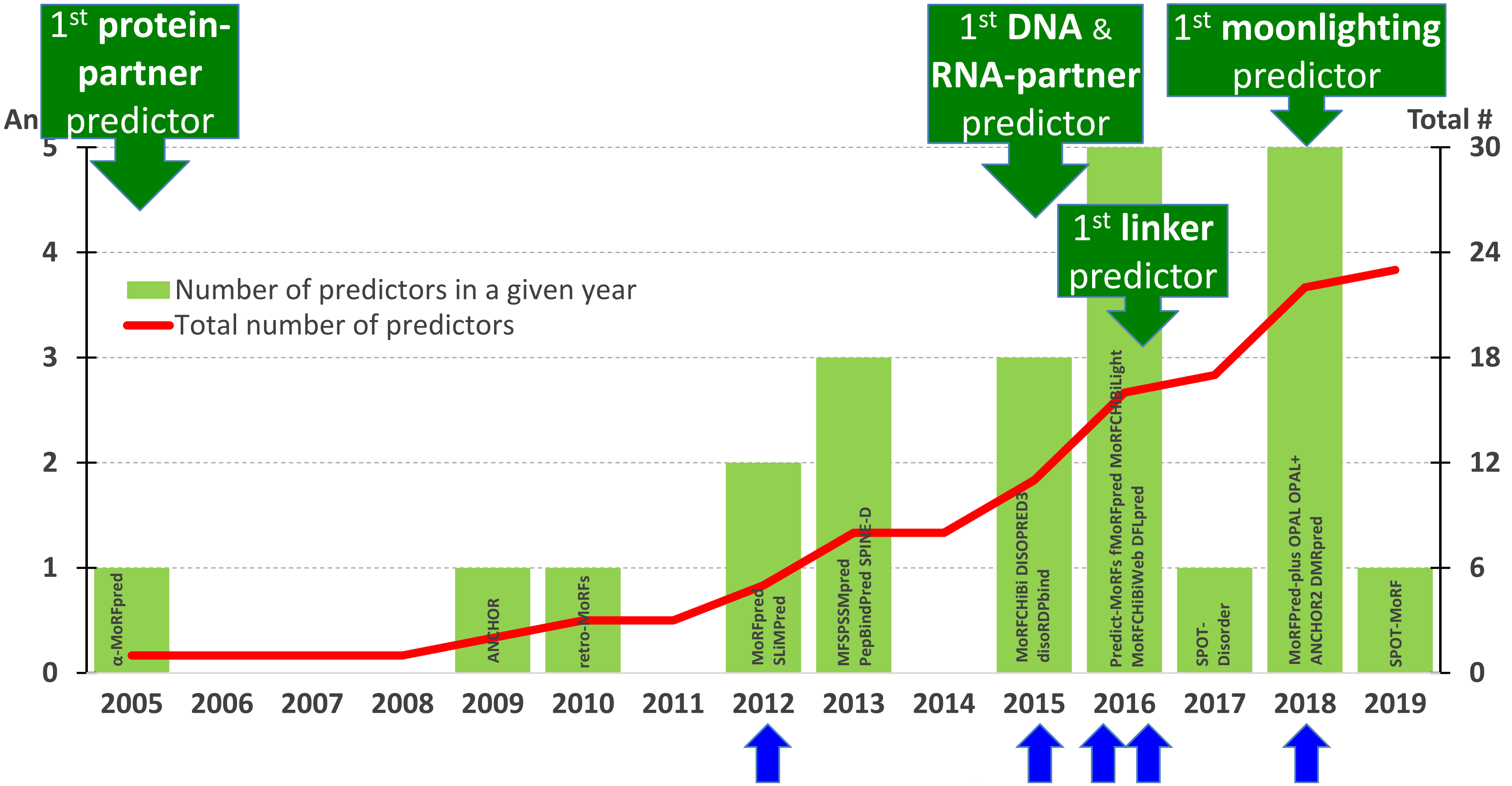


Molecular functions

Entropic Chain (186 IDRs)
Molecular Recognition: Assembler (232 IDRs)
Molecular Recognition: Effector (181 IDRs)
Molecular Recognition: Chaperone (30 IDRs)
Molecular Recognition: Display Site (23 IDRs)
Molecular Recognition: Scavenger (17 IDRs)

Katuwawala A, Ghadermarzi S, Kurgan L. Computational Prediction of Functions of Intrinsically Disordered Regions, In: Uversky VN, (Ed), *Dancing protein clouds: Intrinsically disordered proteins in health and disease, Part A*, 2019

Prediction of intrinsic disorder function



Modern resources for disorder prediction

Databases of predicted disorder

- MobiDB at http://mobidb.bio.unipd.it/
- D^{2}P^2 at http://d2p2.pro/

Disorder prediction center

- DEPICTER at http://biomine.cs.vcu.edu/servers/DEPICTER/

Quality assessment for disorder predictions

- QUARTER at http://biomine.cs.vcu.edu/servers/QUARTER/

MobiDB

http://mobidb.bio.unipd.it/

Predictions of intrinsic disorder and protein-partner regions for 75 million proteins

P06701 (UniPRot) / DP00530 (DisPRot): **silent information regulator (Sir3p) from yeast**

Native annotation

MobiDB profile

Interaction **Fragment** 216-549

Structural state: Disorder
Interaction partner: Protein binding, Nucleic acid binding
Disorder function: Entropic chain

Piovesan D, Tabaro F, …, Tosatto S. MobiDB 3.0: more annotations for intrinsic disorder, conformational diversity and interactions in proteins. *Nucleic Acids Res.* 46(D1):D471-D476, 2018

DEPICTER

Disorder**E**d **P**redict**I**on cen**TER**

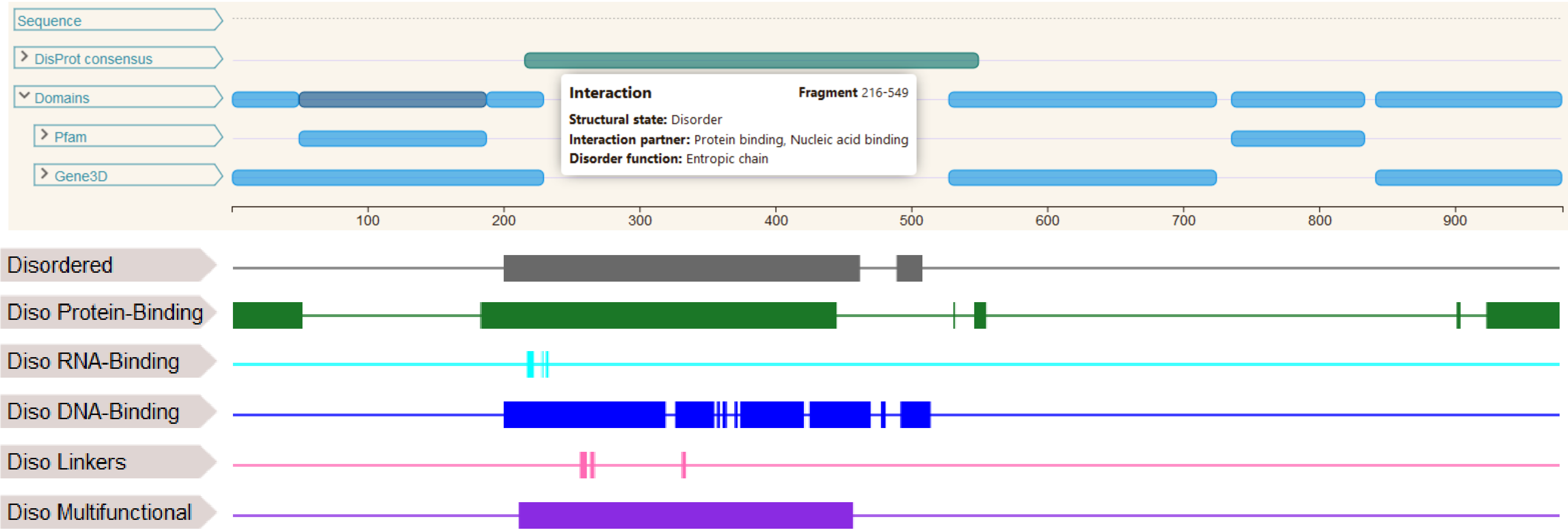
http://biomine.cs.vcu.edu/servers/DEPICTER/

Webserver for integrated **prediction of disorder and disorder functions**

P06701 (UniPRot) / DP00530 (DisPRot): **silent information regulator (Sir3p) from yeast**

Native annotation

DEPICTER profile



Interaction **Fragment** 216-549

Structural state: Disorder
Interaction partner: Protein binding, Nucleic acid binding
Disorder function: Entropic chain

Barik A, Katuwawala A, …., Kurgan L. DEPICTER: intrinsic disorder and disorder function prediction server, *J Molecular Biology* https://doi.org/10.1016/j.jmb.2019.12.030, 2020

QUARTER

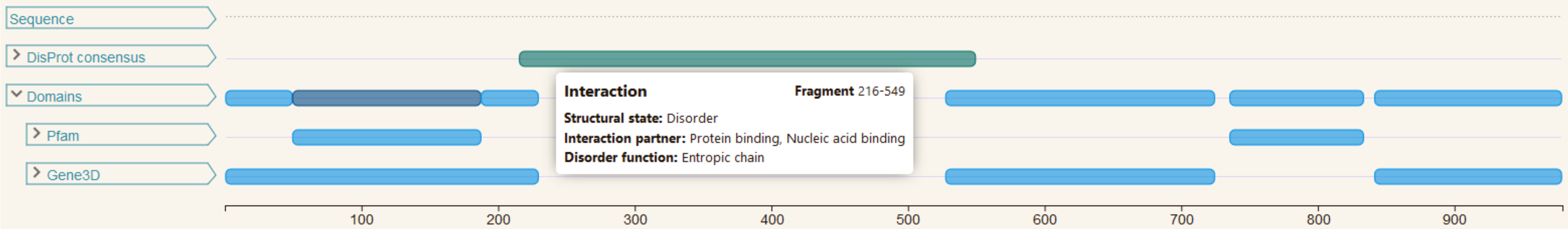
QUality **A**ssessment for p**R**otein in**T**rinsic disord**E**r p**R**edictions

http://biomine.cs.vcu.edu/servers/QUARTER/

Webserver for **quality assessment of disorder predictions**

P06701 (UniPRot) / DP00530 (DisPRot): **silent information regulator (Sir3p) from yeast**

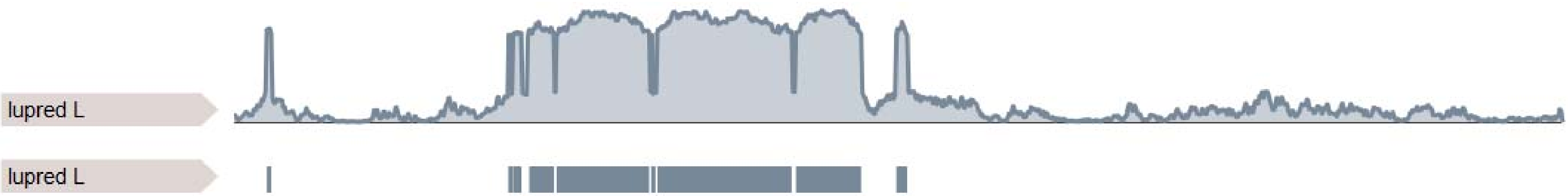
Native annotation



Interaction **Fragment** 216-549

Structural state: Disorder
Interaction partner: Protein binding, Nucleic acid binding
Disorder function: Entropic chain

QUARTER profile



Hu G, Wu Z, Oldfield CJ, Wang C, Kurgan L. Quality assessment for the putative intrinsic disorder in proteins. *Bioinformatics* 35(10):1692-1700, 2019

QUARTER

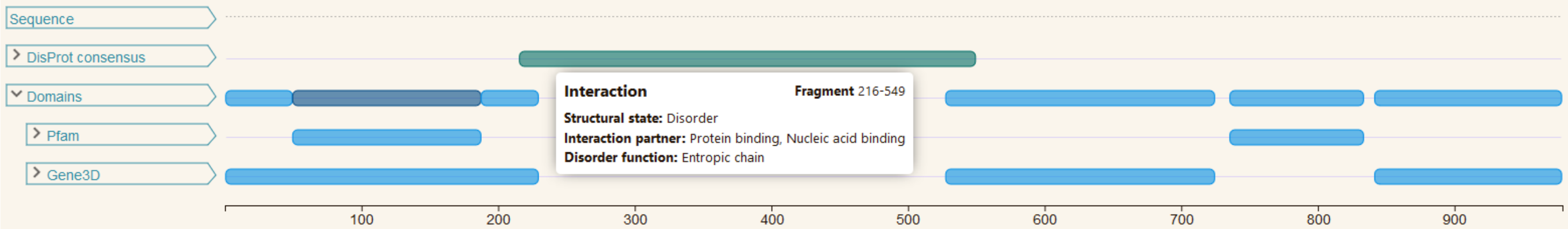
QUality **A**ssessment for p**R**otein in**T**rinsic disord**E**r p**R**edictions

http://biomine.cs.vcu.edu/servers/QUARTER/

Webserver for **quality assessment of disorder predictions**

P06701 (UniPRot) / DP00530 (DisPRot): **silent information regulator (Sir3p) from yeast**

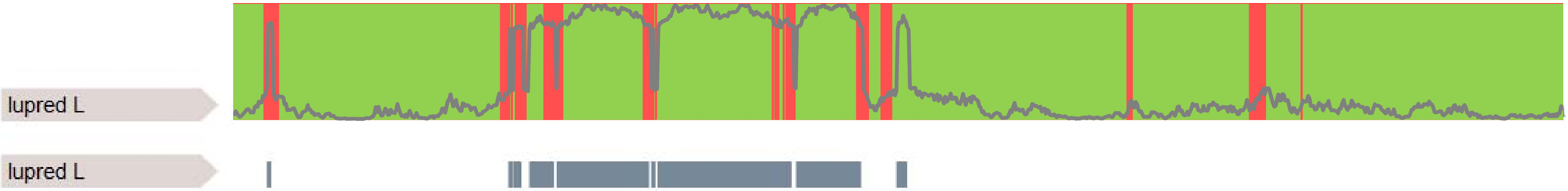
Native annotation



Interaction **Fragment** 216-549

Structural state: Disorder
Interaction partner: Protein binding, Nucleic acid binding
Disorder function: Entropic chain

QUARTER profile



Hu G, Wu Z, Oldfield CJ, Wang C, Kurgan L. Quality assessment for the putative intrinsic disorder in proteins. *Bioinformatics* 35(10):1692-1700, 2019

Summary

80+ predictors, **23 years** of history

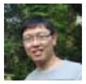
Intrinsic disorder **can be accurately predicted directly from the sequence**

Methods that **predict selected IDR functions** are available

Modern resources emphasize convenience and predictive quality
☐ pre-computed disorder predictions
☐ integrated prediction of disorder and disorder functions
☐ indication of predictive quality/uncertainty at the amino acid level

Acknowledgments


Dr. Amita Barik (NIT Durgapur, India) DEPICTER


Dr. Gang Hu (Nankai University, China) QUARTER


Akila Katuwawala (Virginia Commonwealth University, USA) DEPICTER


Dr. Fanchi Meng (AltaML, Canada) DFLpred DMRpred


Dr. Marcin Mizianty (AltaML, Canada) MFDp MFDp2


Dr. Christopher Oldfield (Virginia Commonwealth University, USA) QUARTER


Dr. Zhenling Peng (Tianjin University, China) DisoRDPbind MFDp2


Dr. Chen Wang (Columbia University, USA) QUARTER DisoRDPbind


Dr. Zhonghua Wu (Nankai University, China) QUARTER

NSERC
CRSNG

NSF