## Short Linear Motifs and the “Eukaryotic Linear Motif” resource

Gibson team

### Exercices

Objective: Get familiar with the [ELM](http://elm.eu.org) (Eukaryotic Linear Motif) prediction tool.

**1. Search in** [**ELM**](http://elm.eu.org) **by copy/pasting the following sequence and using the following parameters:**

**> P12931**

MGSNKSKPKDASQRRRSLEPAENVHGAGGGAFPASQTPSKPASADGHRGPSAAFAPAAAE

PKLFGGFNSSDTVTSPQRAGPLAGGVTTFVALYDYESRTETDLSFKKGERLQIVNNTEGD

WWLAHSLSTGQTGYIPSNYVAPSDSIQAEEWYFGKITRRESERLLLNAENPRGTFLVRES

ETTKGAYCLSVSDFDNAKGLNVKHYKIRKLDSGGFYITSRTQFNSLQQLVAYYSKHADGL

CHRLTTVCPTSKPQTQGLAKDAWEIPRESLRLEVKLGQGCFGEVWMGTWNGTTRVAIKTL

KPGTMSPEAFLQEAQVMKKLRHEKLVQLYAVVSEEPIYIVTEYMSKGSLLDFLKGETGKY

LRLPQLVDMAAQIASGMAYVERMNYVHRDLRAANILVGENLVCKVADFGLARLIEDNEYT

ARQGAKFPIKWTAPEAALYGRFTIKSDVWSFGILLTELTTKGRVPYPGMVNREVLDQVER

GYRMPCPPECPESLHDLMCQCWRKEPEERPTFEYLQAFLEDYFTSTEPQYQPGENL

Using the following parameters:

* Cell Compartment: **Not specified**
* Motif Probability Cutoff: **100**
* Context information: **(leave blank)**

1. How many instances you find?
2. What can you say about the structure of the protein?
   1. Do you find any domains?
   2. Do you find any disordered regions?

**2. Repeat the previous search (again accession P12931) using these parameters:**

* Cell Compartment: **cytosol**
* Motif Probability Cutoff: **0.01**
* Context information: **Homo sapiens**

1. How many instances (roughly) do you find now?
2. How many of the instances are 'annotated'?
3. Do the structural predictors/filters (SMART, GlobPlot, IUPRED, Secondary Structure) agree in terms of which regions are structured/disordered?
4. Compare the location of the annotated instances with structural information at hand (IUPRED, Secondary Structure).

**3. [extra] Use the ELM predictor for NCOA2\_HUMAN, using default parameters.**

1. There are three annotated instances of the ELM class LIG\_NRBOX in the protein NCOA2\_HUMAN. Do they reside in ordered or disordered regions (according to IUPred and SMART)?

**4. [extra] Get all annotated instances for ”Homo sapiens” that contain the search term ”cilium”**

Hint: Use the instance browse page: <http://elm.eu.org/elms/browse_instances.html>

1. How many are there?
2. Which experimental evidence is annotated and how reliable is this evidence?
3. Try to get these instances TSV-file (tab-separated values)

**5. [extra] Search protein UNG\_HUMAN (P13051) for ELMs.**

1. For the annotated instances, which of these ELM classes require a phosphorylation at a certain residue of the motif? (Hint: This information can be found in the description of the ELM class)
2. Which amino acid residues in UNG\_HUMAN correspond to these and can you find evidence for phosphorylations of these residues (using the [Phospho.ELM](http://phospho.elm.eu.org) database)?

### References

* Davey NE, Roey KV, Weatheritt RJ, Toedt G, Uyar B, Altenberg B, Budd A, Diella F, Dinkel H, Gibson TJ: Attributes of short linear motifs. *Mol. BioSyst.* 2012; 8:268–281
* Davey NE, Travé G, Gibson TJ: How viruses hijack cell regulation. *Trends in Biochemical Sciences* 2011; 36:159–169
* Diella F, Via A, Chica C, Luck K, Gould C, Gibson T: The Eukaryotic Linear Motif Resource (ELM): Regulatory Sites in Proteins. *Nature Precedings* 2009;
* Dinkel H, Van Roey K, Michael S, Kumar M, Uyar B, Altenberg B, Milchevskaya V, Schneider M, Kühn H, Behrendt A, Dahl SL, Damerell V, Diebel S, Kalman S, Klein S, Knudsen AC, Mäder C, Merrill S, Staudt A, Thiel V, Welti L, Davey NE, Diella F, Gibson TJ: ELM 2016—data update and new functionality of the eukaryotic linear motif resource. *Nucleic Acids Research* 2015; 44
* Dinkel H, Roey KV, Michael S, Davey NE, Weatheritt RJ, Born D, Speck T, Krüger D, Grebnev G, Kubań M, Strumillo M, Uyar B, Budd A, Altenberg B, Seiler M, Chemes LB, Glavina J, Sánchez IE, Diella F, Gibson TJ: The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Research* 2013; 42
* Gibson TJ, Dinkel H, Roey KV, Diella F: Experimental detection of short regulatory motifs in eukaryotic proteins: tips for good practice as well as for bad. *Cell Communication and Signaling* 2015; 13
* Gouw M, Sámano-Sánchez H, Roey KV, Diella F, Gibson TJ, Dinkel H: Exploring Short Linear Motifs Using the ELM Database and Tools. *Current Protocols in Bioinformatics* 2017;
* Gouw M, Sámano-Sánchez H, Roey KV, Diella F, Gibson TJ, Dinkel H: Exploring Short Linear Motifs Using the ELM Database and Tools. *Current Protocols in Bioinformatics* 2017;
* Gouw M, Michael S, Sámano-Sánchez H, Kumar M, Zeke A, Lang B, Bely B, Chemes LB, Davey NE, Deng Z, Diella F, Gürth C-M, Huber A-K, Kleinsorg S, Schlegel LS, Palopoli N, Roey KV, Altenberg B, Reményi A, Dinkel H, Gibson TJ: The eukaryotic linear motif resource – 2018 update. *Nucleic Acids Research* 2017;