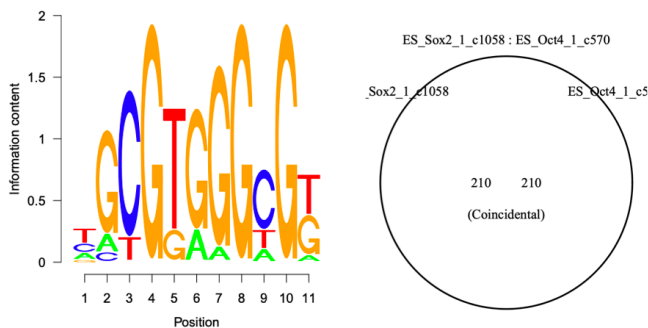


MotifFunc

Minh An Ho

Existing Workflows (R Bioconductor Packages):

- MotIV
- motifStack
- rGADEM
- MotifDb
- Motifmatchr
- motifRG
- DiffLogo
- Cormotif
- universalmotif
- transite



(a) Mmusculus-JASPAR_CORE-Egr1-MA0162.1

Figure 1 (Left): seqLogo used in conjunction with MotifDB**

Figure 2 (Right): venndiagram function output from cobindR**

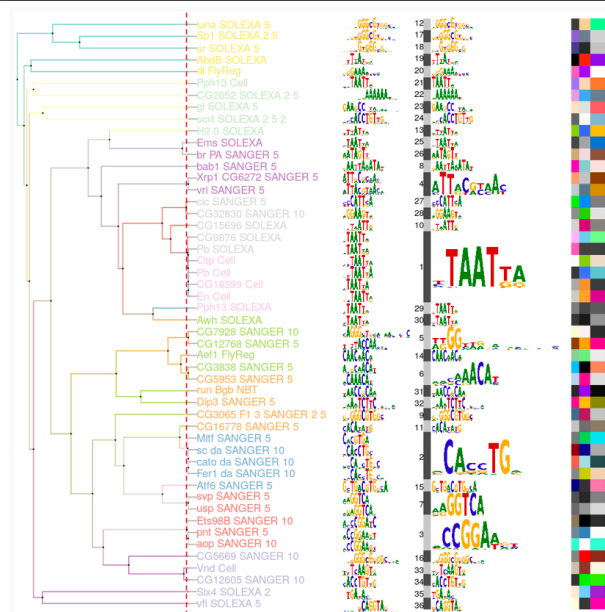


Figure 3: motifPiles function output from motifStack**

External Workflows:

- MOTIF (GenomeNet)
- PDBeMotif

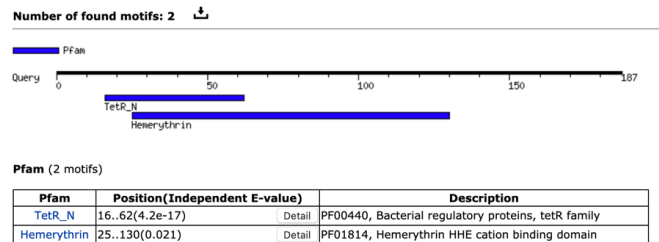


Figure 4: Output of MOTIF on protein FASTA sequence of NP_215990

Unmet Needs:

- Functionality and disease associations of classified binding sites not well catalogued
- A visual/graphical representation of prevalence of certain groupings of functions

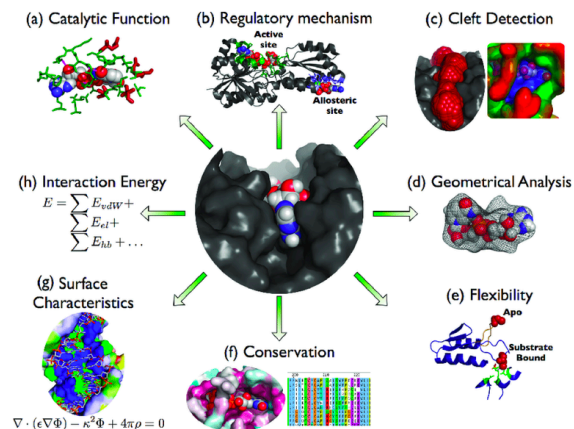


Figure 5: Image taken from *Methods to Characterize the Structure of Enzyme Binding Sites* (2008)

General Idea:

Build off of pre-existing packages to handle genomic/proteomic sequence sets, classify existing binding sites/motifs, categorize binding sites based on function and present in a graphical format inspired by to sequence logos and “word cloud” visualizations