

Evaluation

- Four real datasets

Dataset	# of sequences	Length	Evaluation
[1] DNA ^[1]	506,018	101	Searched against JASPAR database
[2] DNA ^[1]	10,000	101	Compared with MEME-ChIP
[3] Protein ^[2]	11,642	15	Compared with MEME
[4] Protein ^[3]	119,198	25	Runtime analysis

[1] Alipanahi, Babak, et al. "Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning.", 2015

[2] Liu, Honglei, et al. "Fast motif discovery in short sequences." 2016

[3] <http://pfam.xfam.org/>

Evaluation

Results from DeepMotif

Dataset	# of motifs discovered		Runtime		Evaluation
	DeepMotif	MEME / MEME-ChIP	DeepMotif	MEME / MEME-ChIP	
[1] DNA	158	-	11 mins (GPU) 66 mins (CPU)	>five days (never finished)	All the motifs from DeepMotif can find significant matches in JASPAR CORE 2016
[2] DNA	49	46	1 min (GPU) 2 mins (CPU)	50 mins	All the motifs from MEME-ChIP are recalled by DeepMotif
[3] Protein	40	20	1 min (GPU) 2 min (CPU)	>two weeks	All the motifs from MEME are recalled by DeepMotif
[4] Protein	1773	-	7 mins (GPU) 49 mins (CPU)	-	-

Summary

- DeepMotif is **5 orders of magnitude** faster than MEME
- DeepMotif is **50x** faster than MEME-ChIP(only works on DNA)
- DeepMotif is even **10x** faster than ASC+MEME

Example

- https://www.cs.ucsb.edu/~honglei/abp/test/meme_deepmotif/compare.html