# Computational Analysis of Sequence Motifs for Discriminating Different ChIP-exo Profiles of Related RBF Proteins

Shamba S. Mondal[1], Yiliang Wei[2], David N. Arnosti[2] and Bartek Wilczynski[3]

[1]Laboratory of Molecular Neurobiology and Laboratory of Bioinformatics, Nencki Institute of Experimental biology, Polish Academy of Sciences, Warsaw, Poland
[2]Department of Biochemistry and Molecular Biology, Michigan State University, Michigan, USA
[3]Institute of Informatics, University of Warsaw, Warsaw, Poland
Email: s.mondal@nencki.gov.pl
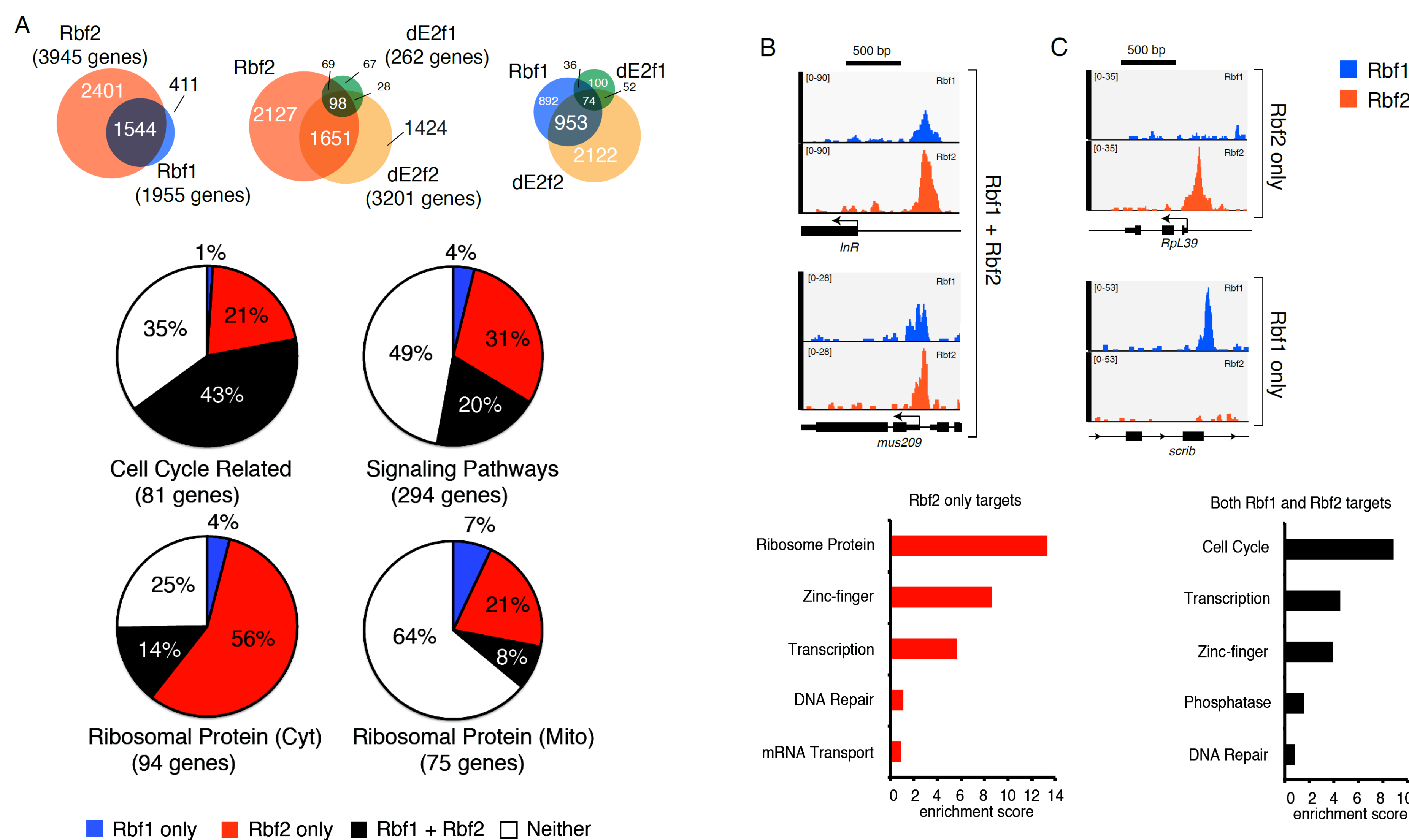
## Background and motivation

### Binding of Retinoblastoma corepressor proteins- Rbf1 and Rbf2 in Drosophila genome and the dREAM/Myb-MuvB complex

Rbf2 is a recently evolved Retinoblastoma family member in Drosophila, differing from Rbf1 especially in the C-terminus. The activity of the activator E2F (dE2F1) is repressed by Rbf1, whereas the repressor E2F (dE2F2) resides in a multi-subunit complex, dREAM/Myb-MuvB consisting of: E2F/DP heterodimer, Rbf1 and/or Rbf2, dMyb, and a set of dMyb-interacting proteins (Mips).

**Existing Model:** Rbf1 binds dE2F1 or dE2F2, while Rbf2 is restricted to binding to dE2F2

**Twist in the tale:** To investigate whether the unique features of Rbf2 contribute to diverse roles in gene regulation, we performed ChIP-Seq for both Rbf2 and Rbf1 in embryos; Rbf2 targets were found to be approximately twice as many genes as Rbf1

## Investigating the basis for differential targeting by Rbf1 and Rbf2

### Dataset Preparation

- **For all (15,829) D. melanogaster genes**, sequence information of 500bp upstream to TSS retreived.
- Same for the **4 functional groups of genes:** Cytoplasmic Ribosomal Protein(CRP), Mitochondrial Ribosomal Protein (MRP), Cell Cycle Genes (CCG) and Signaling Pathway Genes (SPG)
- Genes containing a ChIP enriched region grouped into **4 exclusive classes:** bound by Rbf1 only (208 genes), Rbf2 only (2275), both Rbf1 and Rbf2 (1112) and None(12234).
- The quantitative ChIP enrichments calculated from the .wiggle files.
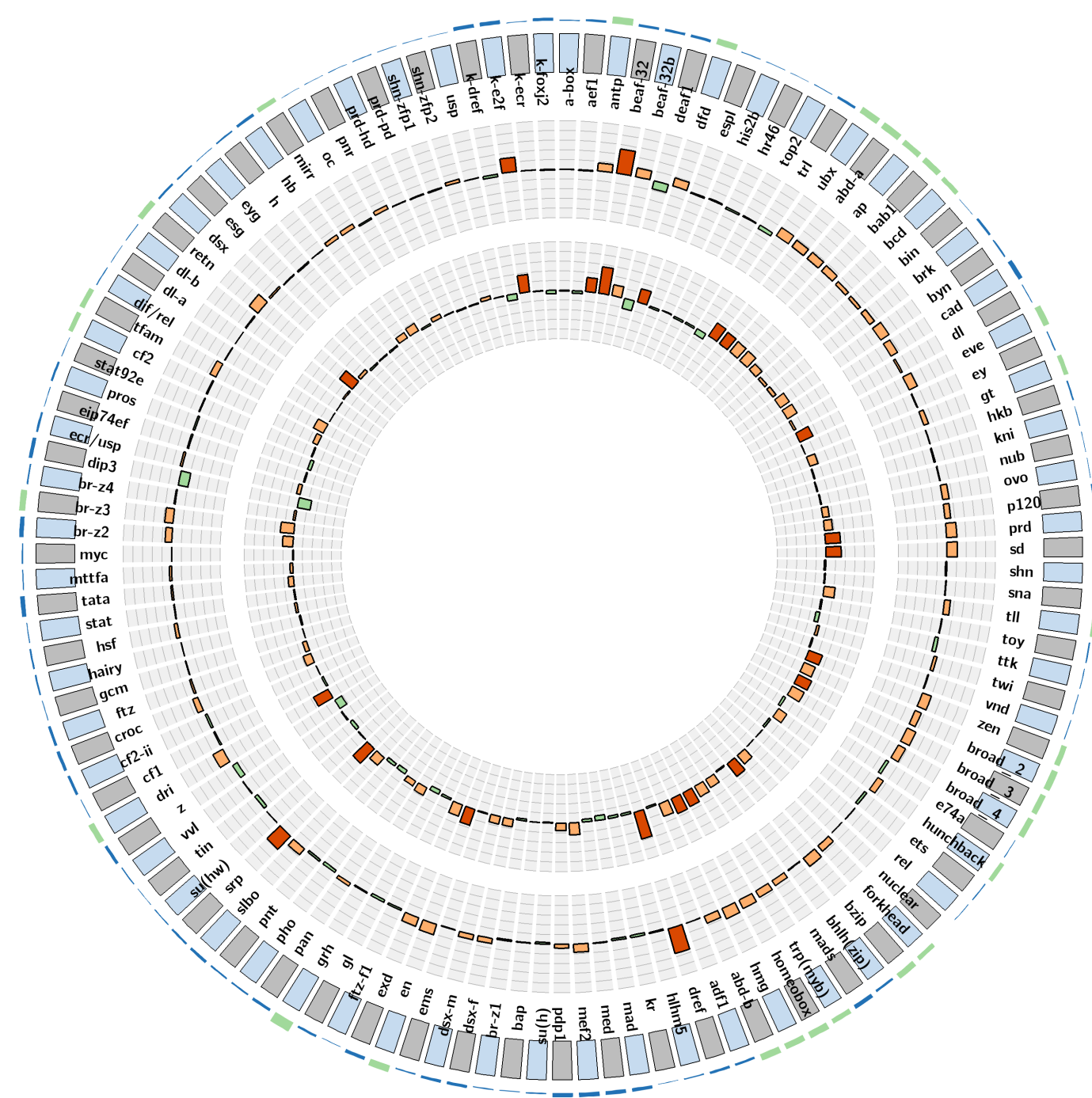- Position Weight Matrices (PWMs) of 127 motifs of TFBS

### Testing for motif association with ChIP enrichment

- Using the STAP (Sequence to Affinity Prediction) program[1], tested which TFBS affinity scores correlate with ChIP enrichment for the DNA sequences upstream of the TSS.

### Motif enrichment and motif strength assessment

- Individual, and therefrom selected pairwise, motif occurrences obtained by running MAST (Motif Alignment and Search Tool)[2].
- Compared the distribution of the strength of non-overlapping binding sites reported by MAST (as p-values) in **2 groups: co-bound by Rbf1+Rbf2, and bound by Rbf2-only.**
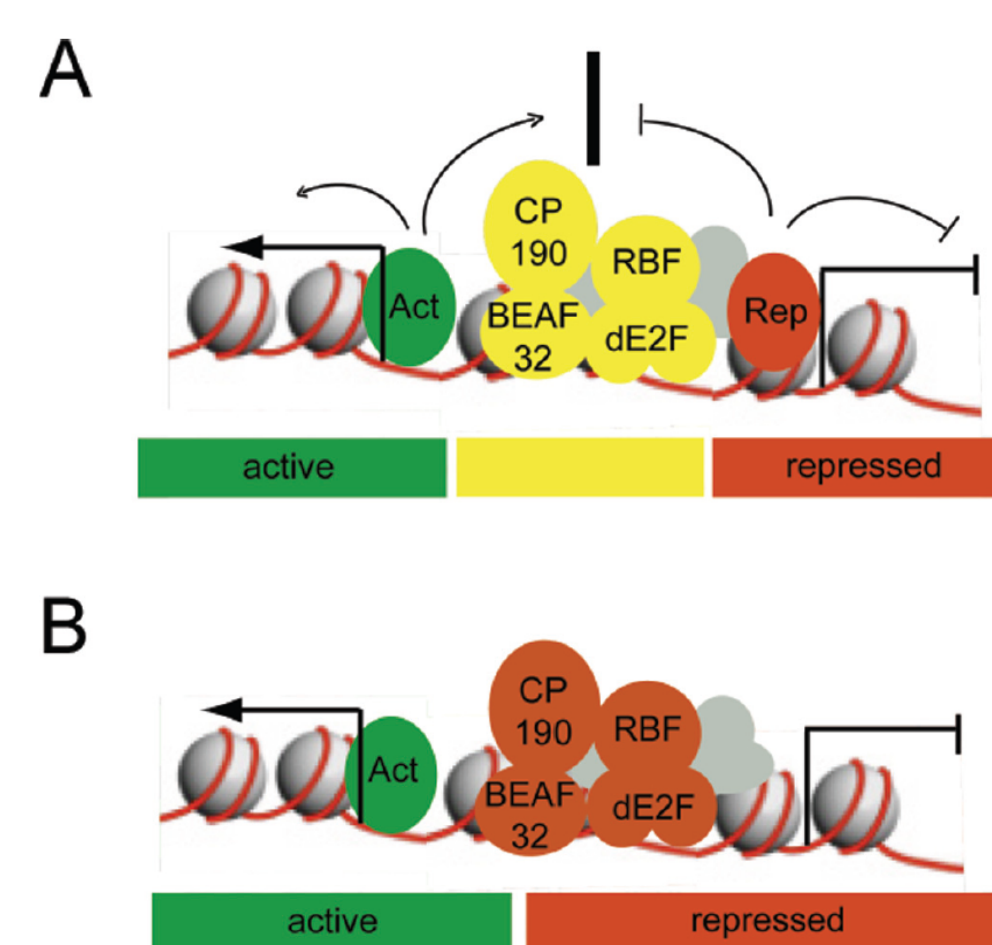
## References

[1] X. He, C. Chen, F. Hong, F. Fang, S. Sinha, H. Ng and S. Zhong, *PLos One*, 2009, **4(12):** e8155

[2] T. Bailey, M. Gribskov, *Bioinformatics*, 1998, **14(1):**, 48–54

[3] M. Korenjak, E. Kwon, R.T. Morris, A. Anderssen, A. Amzallag et al., *Nucleic Acid Res.*, 2014, **42:**, 8939–8953

## Ribosomal protein genes highly enriched among the Rbf2 targets



## Insulator protein BEAF-32 motifs enriched in Rbf bound promoters

### (1) STAP: Pearson correlation between predicted and observed binding
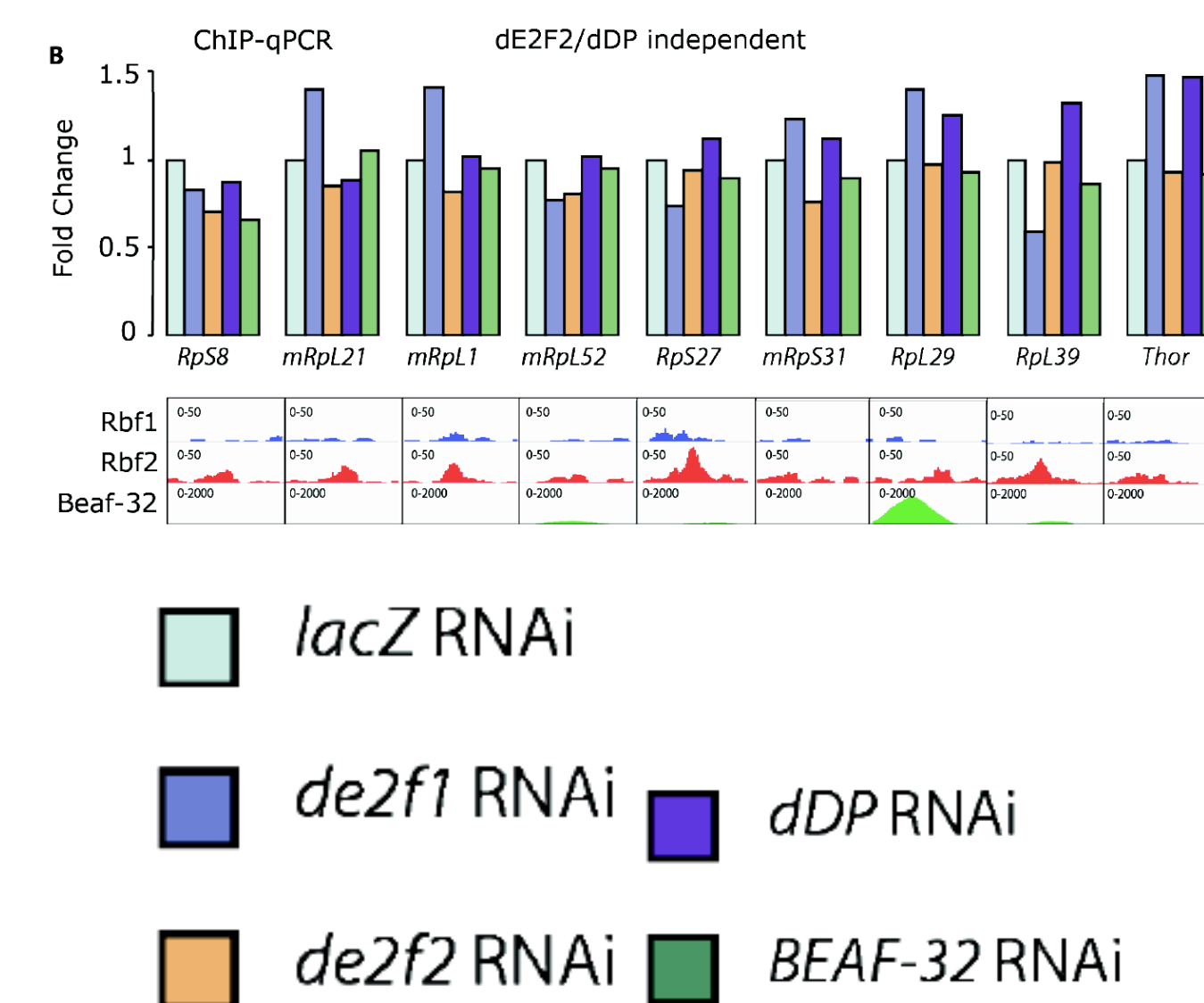


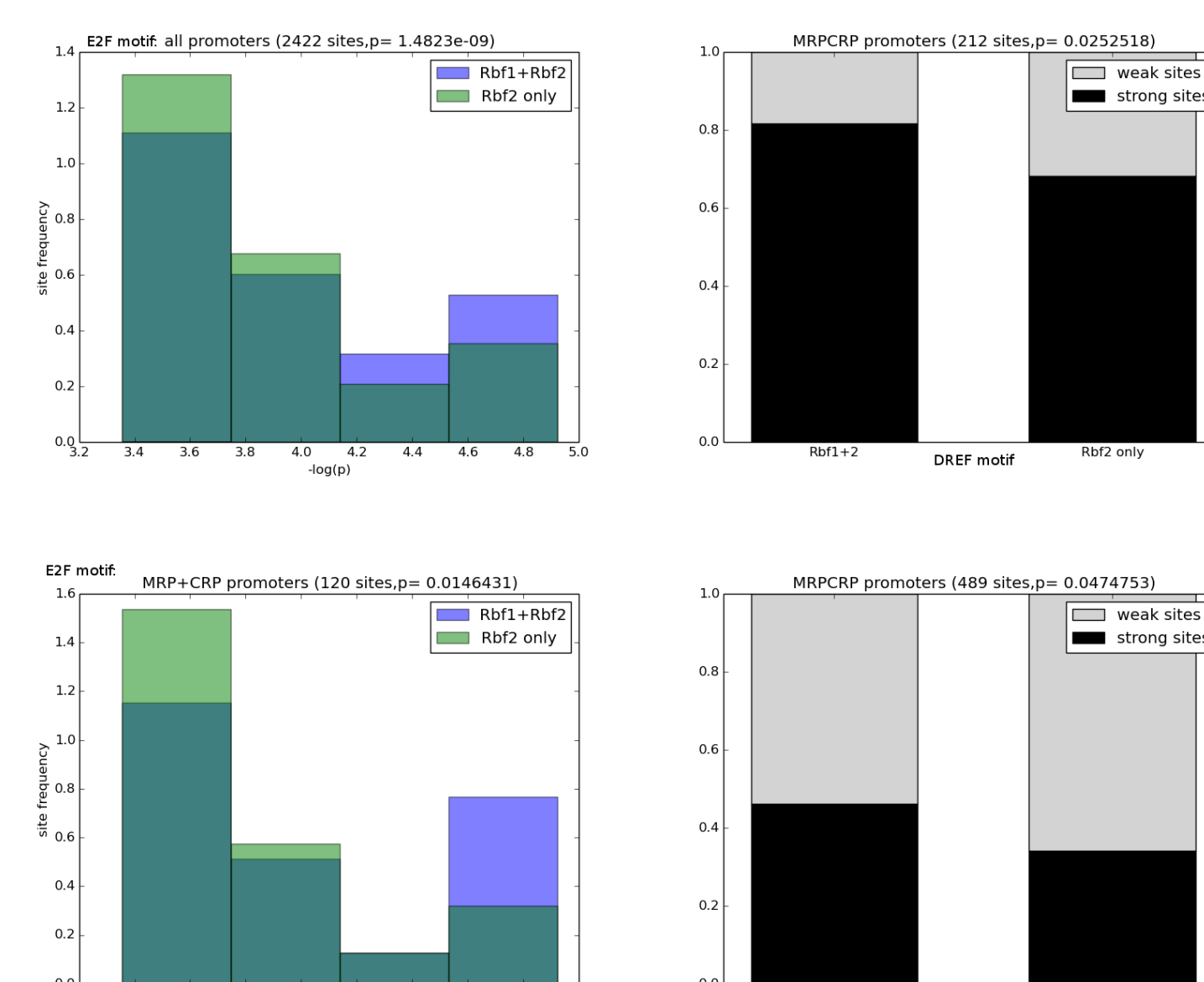### (2) Motif enrichment analysis with MAST
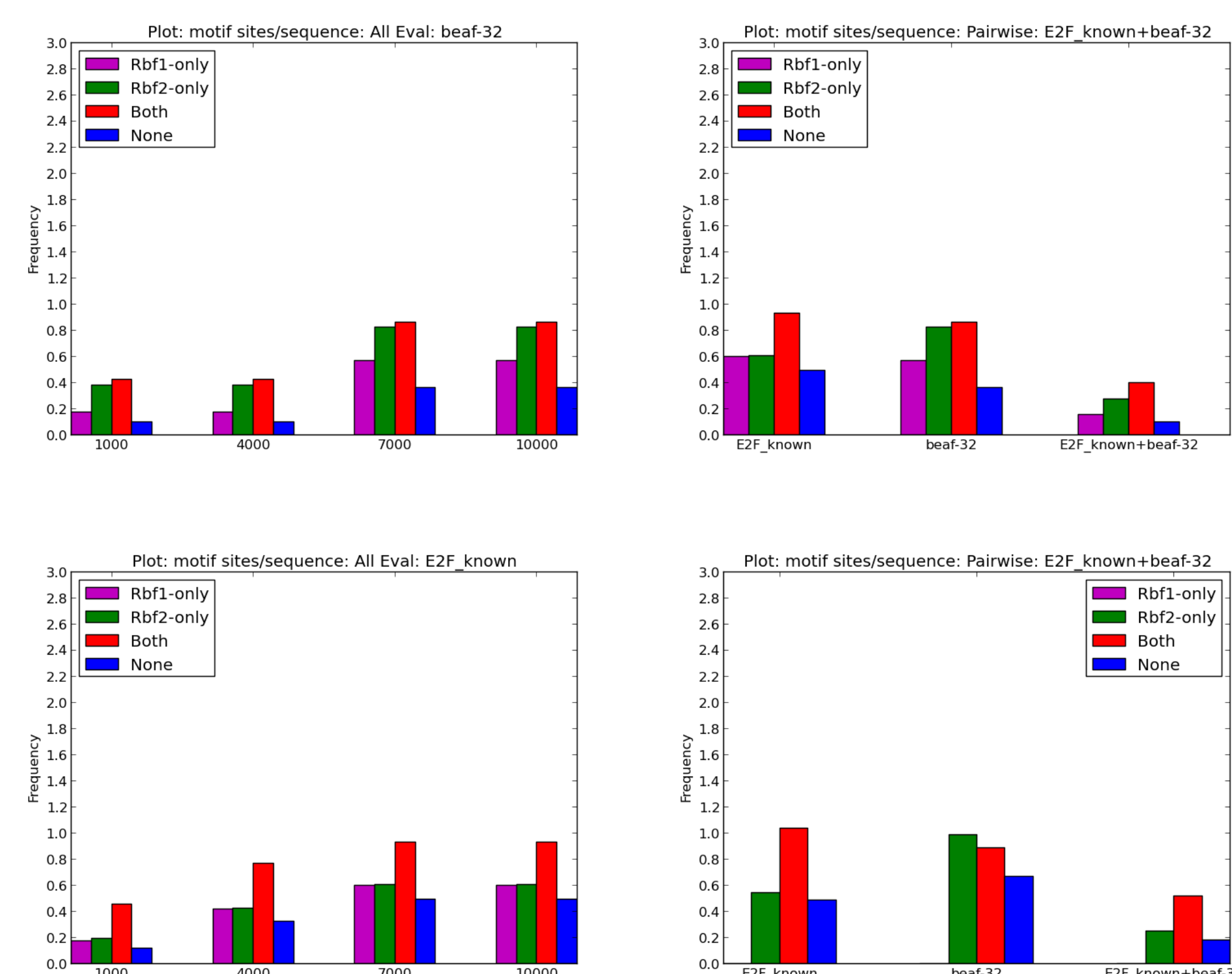


## BEAF-32 is recruited by dREAM complex[3], does not bind Rbf proteins



Adapted from Reference [3]

## Rbf2-only promoters harbors weaker sequence motifs



## Conclusion

- BEAF-32 is associated with dREAM-complex containing Rbf, but **it does not recruit Rbf.**

- From motif strength analysis: **Rbf2-specific promoters** have different preferred motif affinities for multiple factotrs, **suggesting unique targeting mechanisms** based on cooperativity of mutiple weakly bound TFs as opposed to strong binding of a few TFs in RBF1 bound promoters.

- Association of Rbf2 with Rbf2-specific promoters is **dE2F2/dDP-independent**, courtesy distinct sequence motifs.