

FCNGRU: Locating Transcription Factor Binding Sites by combining Fully Convolutional Neural Network with Gated Recurrent Unit

Siguo Wang, Ying He, Zhanheng Chen, Qinhu Zhang*

Supplemental materials

Catalogue

Supplementary Note1. Basic concepts and formulations.....	2
Supplementary Note2. Locating TFBSs and visualizing motifs	2
Supplementary Note3. Experiments on in vivo datasets.....	2
Supplementary Fig. 1. The distribution of IOU values on each in vitro dataset	3
Supplementary Table S1. Average PCC of various methods on 45 uPBM datasets, and the best scores are highlighted in bold.	3
Supplementary Table S2. Average IOU of various methods on 45 uPBM datasets, and the best scores are highlighted in bold.	4
Supplementary Table S3. Average PCC of FCNGRU-double on 45 uPBM datasets	5
Supplementary Table S4. Comparison of the identified motif by different methods.....	7
Supplementary Table S5. Experimental results of the single-task model and the dual-task model on the in vivo datasets	15
Supplementary Table S6. Comparison of the identified in vivo motifs by using in vitro trained models	15

Supplementary Note1. Basic concepts and formulations

These two loss functions are defined as follows:

$$BCE = -Y_i * \log(\tilde{Y}_i) - (1 - Y_i) * \log(1 - \tilde{Y}_i) \quad (1)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (\tilde{Y}_i - Y_i)^2 \quad (2)$$

where Y_i denotes original values and \tilde{Y}_i denotes predicted value of the i -th sample.

Two indicators are defined as follows:

$$IOU = \frac{C \cap G}{C \cup G} \quad (3)$$

$$PCC(y, Y) = \sum_i \frac{y_i - \bar{y}}{\sqrt{(y_i - \bar{y})^2}} \cdot \frac{Y_i - \bar{Y}}{\sqrt{(Y_i - \bar{Y})^2}} \quad (4)$$

where C indicates the candidate and G indicates ground truth bound; where y_i, Y_i, \bar{y} and \bar{Y} respectively represent the observed, predicted, average observed and average predicted binding intensity score.

Supplementary Note2. Locating TFBSs and visualizing motifs

Methods: First, in order to roughly locate the TFBSs, a 20bp sliding window was used to select region containing the maximum number of label 1 from the model output, which are regarded as located regions. Since the convolution kernel is similar to the motif detector in the neural network, which scans all sequences to locate the activation position. Then the trained weights of the first convolutional layer were used to score each sub-region of the located regions, from which the ones with the highest score were then selected. These selected sub-regions were aligned and calculated to obtain the PFMs (Positional Frequency Matrix). Next, PFMs are converted to PWMs based on Equation (5):

$$W(b, i) = \log_2 \frac{f_{b,i}}{p_b} \quad (5)$$

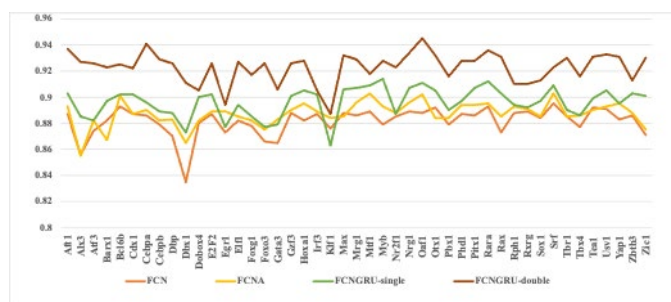
where, $f_{b,i}$ indicates the frequency of base b at position i , p_b is the background frequency of base b in the genome and we assume that all four bases occur equally ($p_b = 0.25$). Finally, compared the learned PWMs representing the predicted motifs with the known motifs in the standard database through TOMTOM.

Supplementary Note3. Experiments on in vivo datasets

Methods: For each in vivo TF dataset, we extracted 101bp peaks as positive samples, and the same length upstream of peaks as negative samples. Then we used a TF's corresponding PWM to annotate each nucleotide of positive samples as 0 or 1 in which 1 means that the nucleotide belongs to TFBSs. Compared with the framework for in vitro datasets, the two-task framework FCNGRU-double for in vivo datasets is composed of two classification branches: a classification task at nucleotide level for

locating TFBSs and a classification task at sequence level for discriminating binding or non-binding sequences. IOU and F1-score are separately used to evaluate the performance of two tasks.

Supplementary Fig. 1. The distribution of IOU values on each *in vitro* data set



Supplementary Fig. 1. The distribution of IOU values on each *in vitro* dataset

Supplementary Table S1. Average PCC of various methods on 45 uPBM datasets, and the best scores are highlighted in bold.

Dataset	DeepBind	DeeperBind	DeepSEA	FCNGRU-double
Aft1	0.609	0.647	0.625	0.696
Alx3	0.355	0.385	0.371	0.44
Atf3	0.037	0.065	0.127	0.067
Barx1	0.386	0.491	0.42	0.673
Bcl6b	0.112	0.187	0.151	0.14
Cdx1	0.158	0.262	0.258	0.364
Cebpa	0.096	0.096	0.101	0.355
Cebpb	0.099	0.161	0.089	0.236
Dbp	0.047	0.043	0.057	0.153
Dbx1	0.476	0.495	0.486	0.58
Dobox4	0.541	0.625	0.543	0.753
E2F2	0.417	0.424	0.378	0.593
Egr1	0.396	0.529	0.479	0.616
Elf1	0.004	0.009	0.003	0.037
Foxg1	0.215	0.348	0.371	0.278
Foxo3	0.407	0.412	0.409	0.403
Gata3	0.149	0.193	0.172	0.602
Gzf3	0.21	0.19	0.16	0.653
Hoxa1	0.137	0.147	0.123	0.315
Irf3	0.069	0.097	0.074	0.468
Klf1	0.354	0.357	0.333	0.43
Max	0.464	0.497	0.501	0.443
Mrg1	0.469	0.533	0.517	0.623

Mtfl	0.15	0.151	0.133	0.199
Myb	0.163	0.237	0.255	0.696
Nr2f1	0.309	0.327	0.313	0.282
Nrg1	0.32	0.431	0.407	0.498
Oaf1	0.631	0.647	0.641	0.679
Otx1	0.319	0.352	0.341	0.442
Pbx1	0.353	0.389	0.387	0.702
Phd1	0.385	0.381	0.353	0.404
Pitx1	0.271	0.353	0.3	0.705
Rara	0.419	0.481	0.462	0.54
Rax	0.247	0.399	0.369	0.485
Rph1	0.178	0.237	0.203	0.287
Rxrg	0.088	0.229	0.202	0.355
Sox1	0.161	0.302	0.345	0.126
Srf	0.204	0.363	0.348	0.431
Tbr1	0.17	0.18	0.165	0.222
Tbx4	0.182	0.273	0.268	0.182
Tea1	0.417	0.425	0.418	0.44
Usv1	0.34	0.447	0.422	0.505
Yap1	0.277	0.336	0.304	0.363
Zbtb3	0.284	0.338	0.304	0.345
Zic1	0.466	0.505	0.482	0.481
Average value	0.279	0.332	0.315	0.429

Supplementary Table S2. Average IOU of various methods on 45 uPBM datasets, and the best scores are highlighted in bold.

Dataset	FCN	FCNA	FCNGRU-single	FCNGRU-double
Aft1	0.887	0.893	0.903	0.937
Alx3	0.856	0.855	0.885	0.927
Atf3	0.874	0.882	0.882	0.926
Barx1	0.882	0.867	0.897	0.923
Bcl6b	0.893	0.901	0.902	0.925
Cdx1	0.887	0.887	0.902	0.922
Cebpa	0.886	0.89	0.896	0.941
Cebpb	0.879	0.882	0.889	0.929
Dbp	0.87	0.883	0.888	0.926
Dbx1	0.835	0.865	0.873	0.911
Dobox4	0.88	0.882	0.9	0.905
E2F2	0.887	0.889	0.902	0.926
Egr1	0.873	0.889	0.877	0.894
Elf1	0.882	0.885	0.894	0.927
Foxg1	0.878	0.882	0.885	0.917

Foxo3	0.866	0.875	0.877	0.926
Gata3	0.865	0.883	0.879	0.906
Gzf3	0.888	0.89	0.901	0.926
Hoxa1	0.882	0.895	0.905	0.928
Irf3	0.887	0.889	0.902	0.905
Klf1	0.876	0.884	0.863	0.887
Max	0.888	0.885	0.906	0.932
Mrg1	0.886	0.896	0.907	0.929
Mtf1	0.889	0.903	0.909	0.918
Myb	0.879	0.893	0.914	0.928
Nr2f1	0.885	0.888	0.887	0.923
Nrg1	0.889	0.896	0.907	0.934
Oaf1	0.888	0.902	0.911	0.945
Otx1	0.892	0.884	0.905	0.932
Pbx1	0.879	0.884	0.89	0.916
Phd1	0.887	0.894	0.897	0.928
Pitx1	0.886	0.894	0.907	0.928
Rara	0.893	0.895	0.912	0.936
Rax	0.873	0.885	0.903	0.931
Rph1	0.888	0.893	0.894	0.91
Rxrg	0.889	0.891	0.892	0.91
Sox1	0.884	0.885	0.897	0.913
Srf	0.895	0.903	0.909	0.923
Tbr1	0.885	0.885	0.89	0.93
Tbx4	0.877	0.886	0.886	0.916
Tea1	0.892	0.89	0.899	0.931
Usv1	0.891	0.893	0.905	0.933
Yap1	0.883	0.895	0.895	0.931
Zbtb3	0.886	0.888	0.903	0.913
Zic1	0.871	0.875	0.901	0.93
Average IOU value	0.881	0.887	0.896	0.922

Supplementary Table S3. Average PCC of FCNGRU-double on 45 uPBM datasets

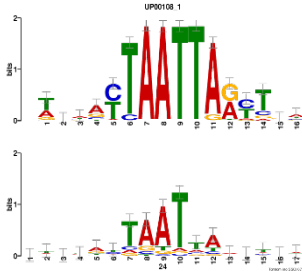
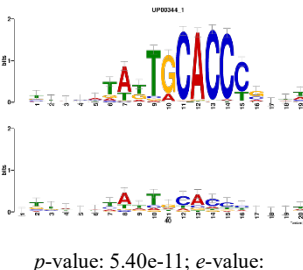
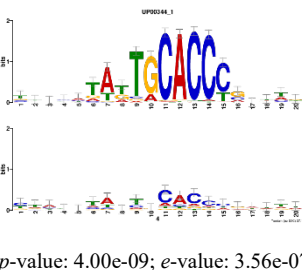
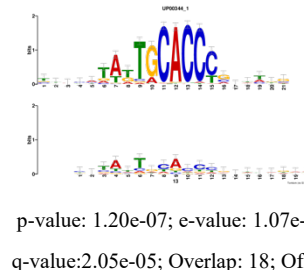
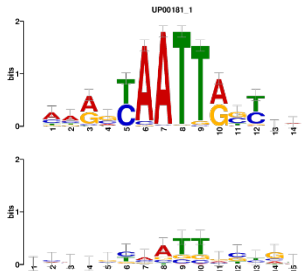
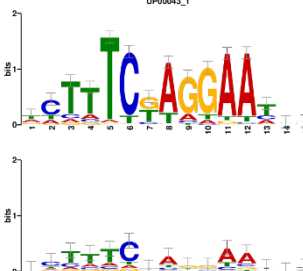
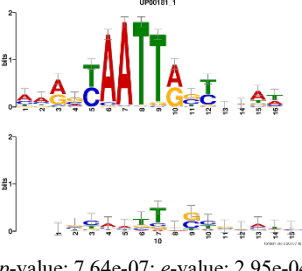
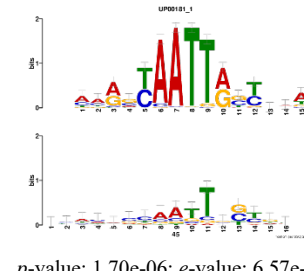
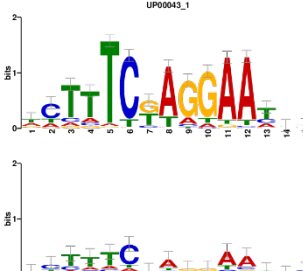
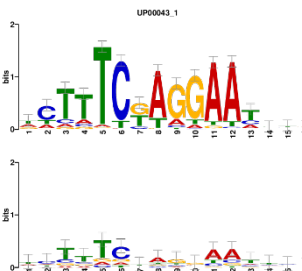
dataset	PCC
Aft1	0.696
Alx3	0.44
Atf3	0.039
Barx1	0.673
Bcl6b	0.045
Cdx1	0.364
Cebpa	0.355
Cebpb	0.236

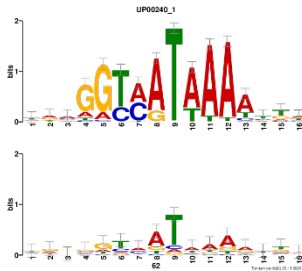
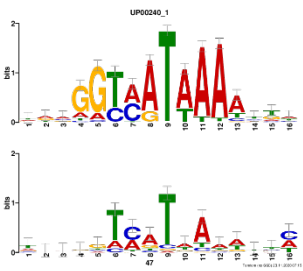
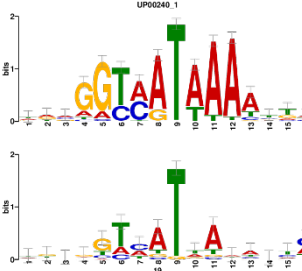
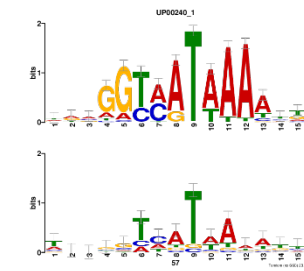
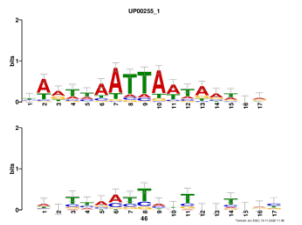
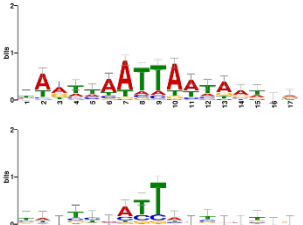
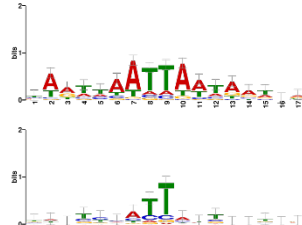
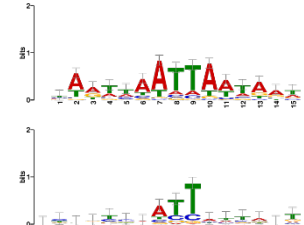
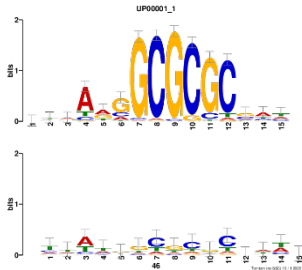
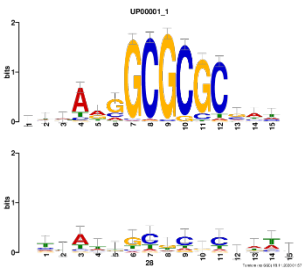
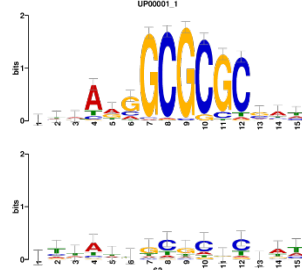
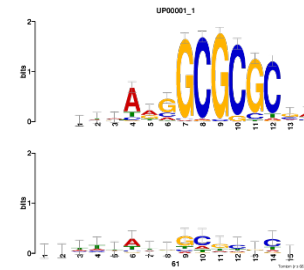
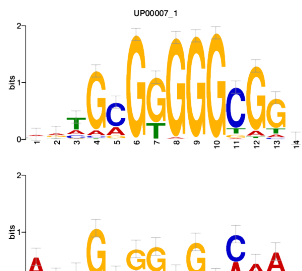
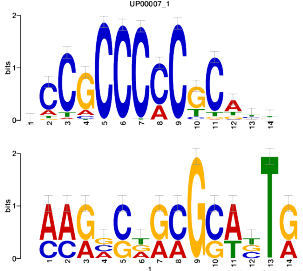
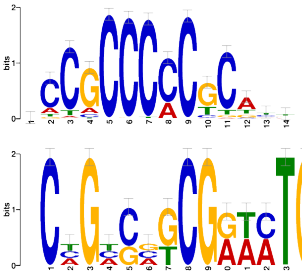
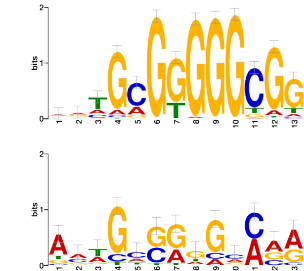
Dbp	0.153
Dbx1	0.58
Dobox4	0.753
E2F2	0.593
Egr1	0.616
Elf1	0.037
Foxg1	0.168
Foxo3	0.379
Gata3	0.602

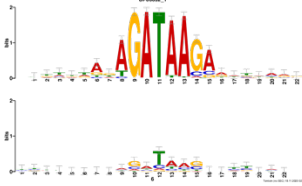
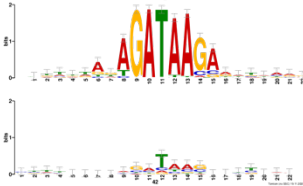
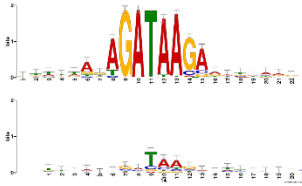
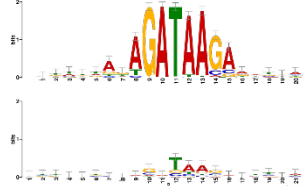
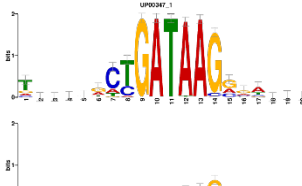
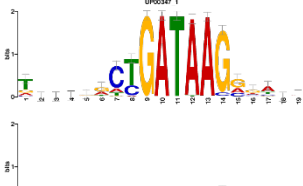
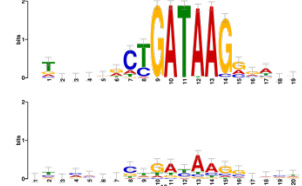
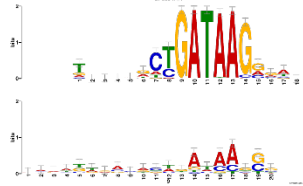
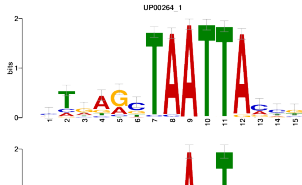
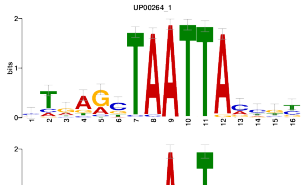
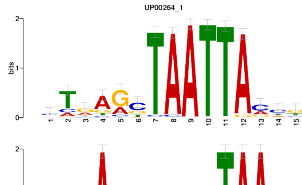

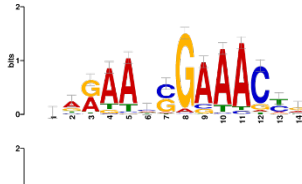

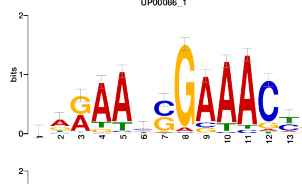
Gzf3	0.653
Hoxa1	0.315
Irf3	0.468
Klf1	0.16
Max	0.443
Mrg1	0.623
Mtf1	0.199
Myb	0.696
Nr2f1	0.282
Nrg1	0.498
Oaf1	0.679
Otx1	0.276
Pbx1	0.702
Phd1	0.122

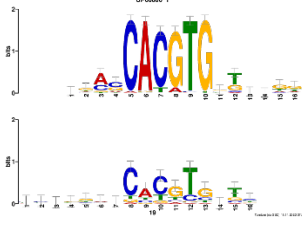
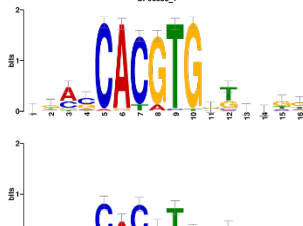
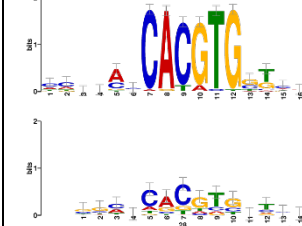
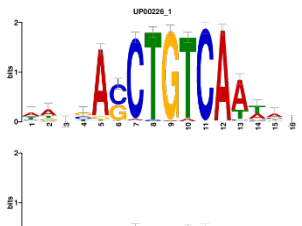
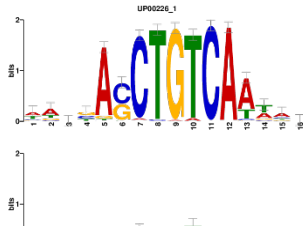
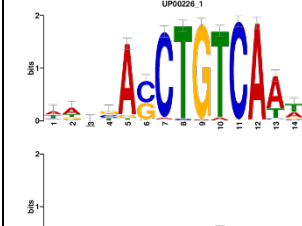
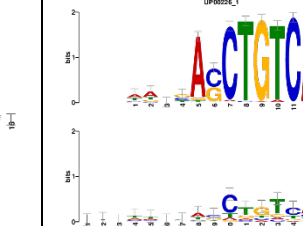
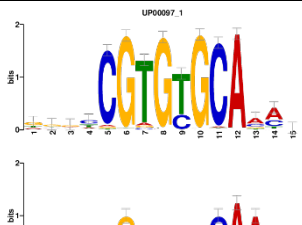
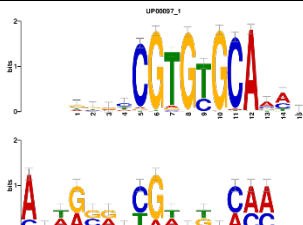
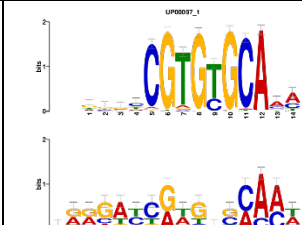
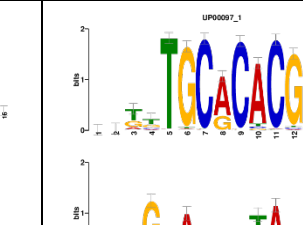
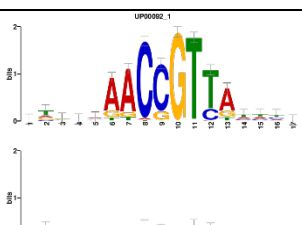
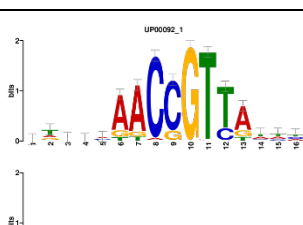
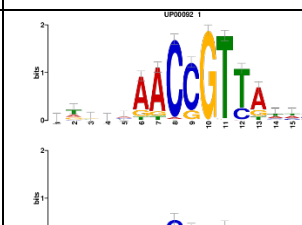
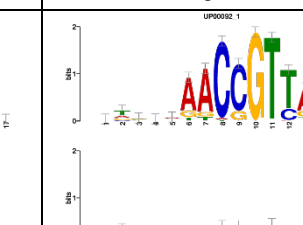
Pitx1	0.705
Rara	0.54
Rax	0.485
Rph1	0.287
Rxrg	0.355
Sox1	0.076
Srf	0.431
Tbr1	0.222
Tbx4	0.162
Tea1	0.44
Usv1	0.505
Yap1	0.363
Zbtb3	0.345
Zic1	0.302

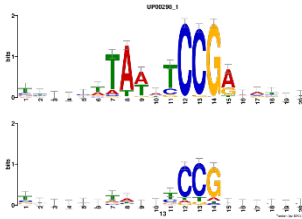
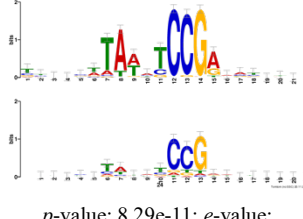
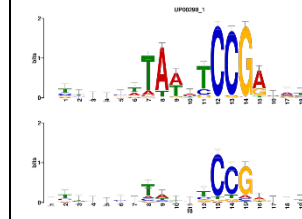
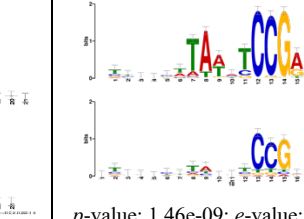
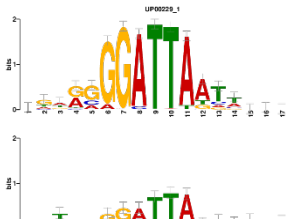
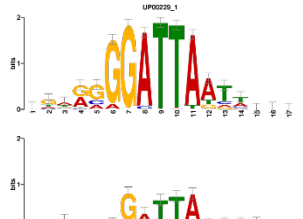
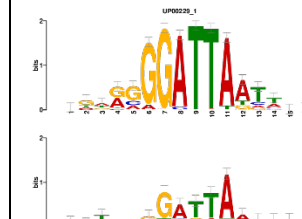
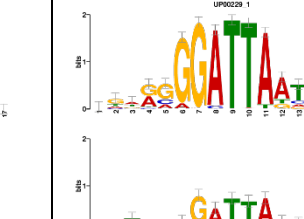
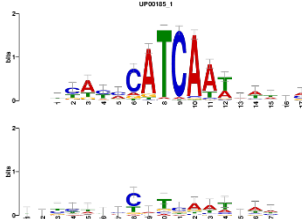
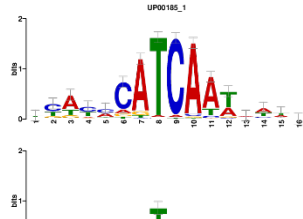
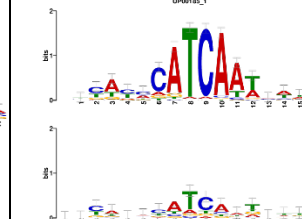
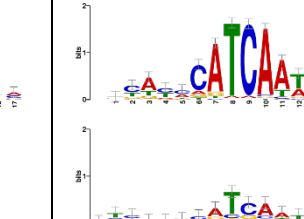
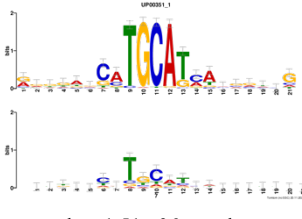
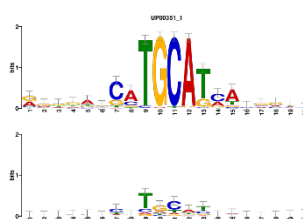
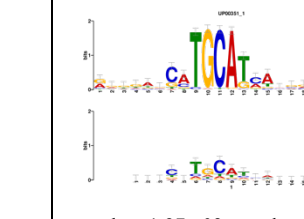
Supplementary Table S4. Comparison of the identified motif by different methods

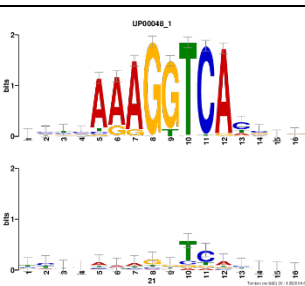
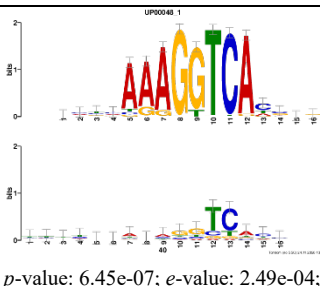
TF name	FCNGRU-double	FCNGRU-single	FCNA	FCN
Alx3 (UP00108)	 <p><i>p</i>-value: 4.23e-08; <i>e</i>-value: 1.63e-05; <i>q</i>-value: 2.63e-05; Overlap: 16; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 7.86e-07; <i>e</i>-value: 1.41e-03; <i>q</i>-value: 2.07e-03; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 8.39e-06; <i>e</i>-value: 3.24e-03; <i>q</i>-value: 3.38e-03; Overlap: 13; Offset: -4; Orientation: Normal</p>	 <p><i>p</i>-value: 3.99e-06; <i>e</i>-value: 1.54e-05; <i>q</i>-value: 2.49e-05; Overlap: 13; Offset: -4; Orientation: Normal</p>
Aft1 (UP00344)	 <p><i>p</i>-value: 7.56e-17; <i>e</i>-value: 6.73e-15; <i>q</i>-value: 1.32e-14; Overlap: 21; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 5.40e-11; <i>e</i>-value: 4.80e-09; <i>q</i>-value: 9.49e-09; Overlap: 20; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 4.00e-09; <i>e</i>-value: 3.56e-07; <i>q</i>-value: 6.94e-05; Overlap: 21; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.20e-07; <i>e</i>-value: 1.07e-05; <i>q</i>-value: 2.05e-05; Overlap: 18; Offset: 3; Orientation: Normal</p>
Barx1 (UP00108)	 <p><i>p</i>-value: 4.48e-09; <i>e</i>-value: 1.73e-06; <i>q</i>-value: 2.90e-06; Overlap: 15; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 5.75e-08; <i>e</i>-value: 2.22e-05; <i>q</i>-value: 3.79e-05; Overlap: 15; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 7.64e-07; <i>e</i>-value: 2.95e-04; <i>q</i>-value: 5.29e-04; Overlap: 14; Offset: 2; Orientation: Normal</p>	 <p><i>p</i>-value: 1.70e-06; <i>e</i>-value: 6.57e-04; <i>q</i>-value: 1.20e-03; Overlap: 14; Offset: -2; Orientation: Normal</p>
Bcl6b (UP00108)	 <p><i>p</i>-value: 5.59e-13; <i>e</i>-value: 2.16e-10; <i>q</i>-value: 4.31e-10;</p>	 <p><i>p</i>-value: 1.13e-12; <i>e</i>-value: 4.37e-10; <i>q</i>-value: 8.74e-10;</p>	 <p><i>p</i>-value: 5.96e-10; <i>e</i>-value: 2.30e-07; <i>q</i>-value: 4.60e-07; Overlap: 16; Offset:</p>	 <p><i>p</i>-value: 1.27e-11; <i>e</i>-value: 4.92e-09; <i>q</i>-value: 9.84e-09; Overlap: 16; Offset:</p>

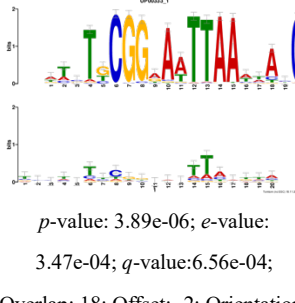
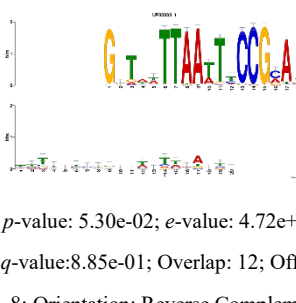
	Overlap: 16; Offset: 0; Orientation: Normal	Overlap: 16; Offset: 0; Orientation: Normal	0; Orientation: Normal	0; Orientation: Normal
Cdx1 (UP00240)	 <p><i>p</i>-value: 2.29e-16; <i>e</i>-value: 8.85e-14; <i>q</i>-value:1.63e-13; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 2.10e-11; <i>e</i>-value: 8.10e-06; <i>q</i>-value:9.75e-06; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 3.90e-09; <i>e</i>-value: 1.51e-06; <i>q</i>-value:2.69e-06; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 3.75e-08; <i>e</i>-value: 1.45e-05; <i>q</i>-value:1.59e-05; Overlap: 16; Offset: 0; Orientation: Normal</p>
Dbx1 (UP00255)	 <p><i>p</i>-value: 1.79e-09; <i>e</i>-value: 6.92e-04; <i>q</i>-value:1.28e-03; Overlap: 16; Offset: 1; Orientation: Normal</p>	 <p><i>p</i>-value: 4.86e-08; <i>e</i>-value: 1.88e-03; <i>q</i>-value:3.58e-03; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 5.76e-06; <i>e</i>-value: 2.22e-03; <i>q</i>-value:4.18e-03; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.08e-04; <i>e</i>-value: 4.17e-02; <i>q</i>-value:2.58e-02; Overlap: 16; Offset: -1; Orientation: Normal</p>
E2f2 (UP00001)	 <p><i>p</i>-value: 2.37e-07; <i>e</i>-value: 9.16e-05; <i>q</i>-value:1.81e-04; Overlap: 14; Offset: 1; Orientation: Normal</p>	 <p><i>p</i>-value: 3.97e-06; <i>e</i>-value: 1.53e-03; <i>q</i>-value:3.02e-03; Overlap: 14; Offset: 1; Orientation: Normal</p>	 <p><i>p</i>-value: 1.50e-05; <i>e</i>-value: 5.77e-03; <i>q</i>-value:1.15e-02; Overlap: 15; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.17e-04; <i>e</i>-value: 4.51e-02; <i>q</i>-value:5.74e-02; Overlap: 13; Offset: -2; Orientation: Normal</p>
Egr1 (UP00007)	 <p><i>p</i>-value: 3.03e-05; <i>e</i>-value: 6.72e-01; <i>q</i>-value:6.72e-01;</p>	 <p><i>p</i>-value: 9.18e-03; <i>e</i>-value: 3.54e+00; <i>q</i>-value:1.00e+00; Overlap: 13; Offset: 1; Orientation:</p>	 <p><i>p</i>-value: 8.64e-03; <i>e</i>-value: 3.34e+00; <i>q</i>-value:1.00e+00; Overlap: 13; Offset: 1; Orientation: Normal</p>	 <p><i>p</i>-value: 2.42e-02; <i>e</i>-value: 9.32e+00; <i>q</i>-value:1.00e+00; Overlap: 13; Offset: 1; Orientation: Normal</p>

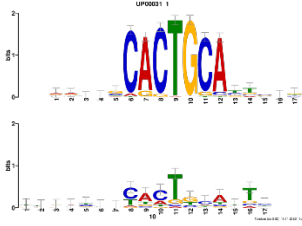
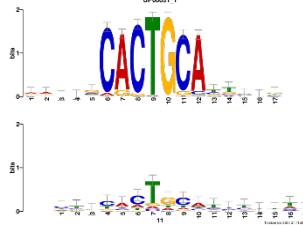
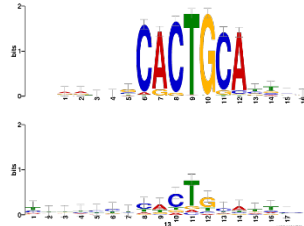
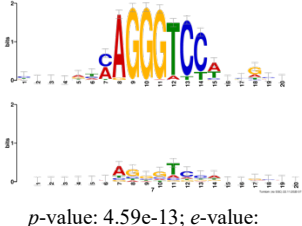
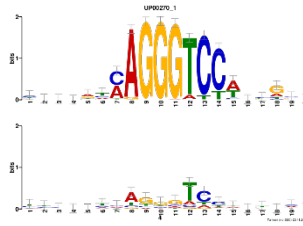
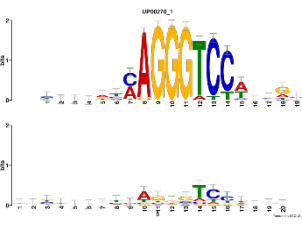
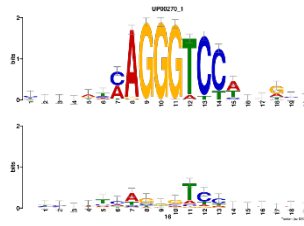
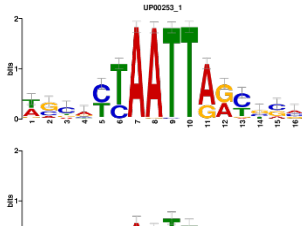
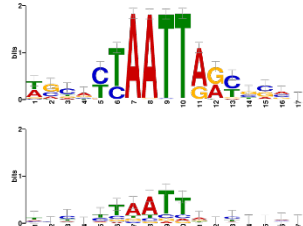
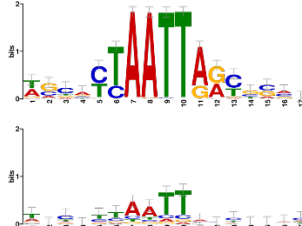
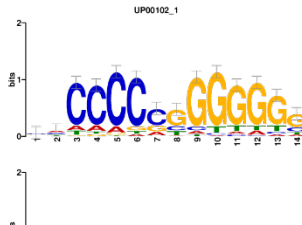
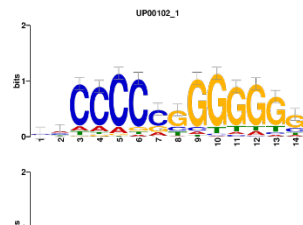
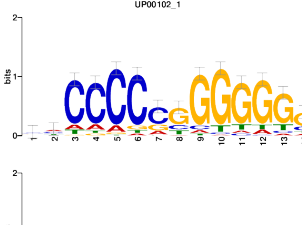
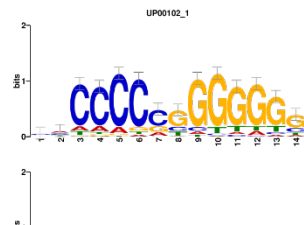
	Overlap: 13; Offset: 1; Orientation: Normal	Normal		
Gata3 (UP00032)	 <p>p-value: 1.93e-10; e-value: 7.46e-06; q-value:1.49e-05; Overlap: 21; Offset: -1; Orientation: Normal</p>	 <p>p-value: 5.57e-08; e-value: 2.15e-05; q-value:4.30e-05; Overlap: 21; Offset: -1; Orientation: Normal</p>	 <p>p-value: 9.21e-07; e-value: 3.55e-04; q-value:7.11e-04; Overlap: 20; Offset: 2; Orientation: Normal</p>	 <p>p-value: 4.60e-07; e-value: 1.78e-04; q-value:3.55e-04; Overlap: 21; Offset: -1; Orientation: Normal</p>
Gzf3 (UP00347)	 <p>p-value: 1.78e-08; e-value: 1.58e-06; q-value:3.13e-06; Overlap: 20; Offset: 0; Orientation: Normal</p>	 <p>p-value: 1.10e-7; e-value: 9.78e-06; q-value:1.94e-05; Overlap: 20; Offset: 0; Orientation: Normal</p>	 <p>p-value: 3.21e-06; e-value: 2.85e-04; q-value:5.67e-04; Overlap: 19; Offset: -1; Orientation: Normal</p>	 <p>p-value: 4.51e-05; e-value: 4.02e-03; q-value:3.94e-03; Overlap: 16; Offset: -4; Orientation: Normal</p>
Hoxa1 (UP00264)	 <p>p-value: 3.07e-06; e-value: 1.19e-03; q-value:1.16e-03; Overlap: 15; Offset: -1; Orientation: Normal</p>	 <p>p-value: 2.64e-05; e-value: 1.02e-02; q-value:9.01e-03; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p>p-value: 3.47e-04; e-value: 1.34e-01; q-value:5.89e-02; Overlap: 15; Offset: -1; Orientation: Normal</p>	 <p>p-value: 6.50e-03; e-value: 2.51e+00; q-value:5.35e-02; Overlap: 14; Offset: -2; Orientation: Normal</p>
Irf3 (UP00086)	 <p>p-value: 2.64e-11; e-value: 1.02e-05; q-value:2.03e-05; Overlap: 13; Offset: -1; Orientation: Normal</p>	 <p>p-value: 2.24e-8; e-value: 8.66e-08; q-value:1.73e-07; Overlap: 14; Offset: 0; Orientation: Normal</p>	 <p>p-value: 9.35e-07; e-value: 3.61e-06; q-value:7.21e-06; Overlap: 14; Offset: 0; Orientation: Normal</p>	None

<p>Max (UP00060)</p>  <p><i>p</i>-value: 5.64e-08; <i>e</i>-value: 2.18e-05; <i>q</i>-value:4.34e-05; Overlap: 13; Offset: -3; Orientation: Normal</p>	 <p><i>p</i>-value: 2.10e-06; <i>e</i>-value: 8.10e-04; <i>q</i>-value:1.620e-03; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.52e-08; <i>e</i>-value: 5.87e-06; <i>q</i>-value:1.17e-05; Overlap: 14; Offset: 2; Orientation: Normal</p>	<p>None</p>
<p>Mrg1 (UP00226)</p>  <p><i>p</i>-value: 7.34e-09; <i>e</i>-value: 2.83e-05; <i>q</i>-value:2.81e-05; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.30e-07; <i>e</i>-value: 5.01e-05; <i>q</i>-value:3.16e-05; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.64e-06; <i>e</i>-value: 6.34e-04; <i>q</i>-value:3.46e-04; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 6.60e-05; <i>e</i>-value: 2.55e-02; <i>q</i>-value:2.54e-02; Overlap: 13; Offset: -3; Orientation: Normal</p>
<p>Mtf1 (UP00097)</p>  <p><i>p</i>-value: 3.14e-07; <i>e</i>-value: 1.21e-04; <i>q</i>-value:2.42e-04; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 6.06e-06; <i>e</i>-value: 2.34e-03; <i>q</i>-value:4.68e-03; Overlap: 13; Offset: -3; Orientation: Normal</p>	 <p><i>p</i>-value: 2.48e-05; <i>e</i>-value: 9.58e-03; <i>q</i>-value:1.92e-02; Overlap: 14; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 2.45e-02; <i>e</i>-value: 9.45e+00; <i>q</i>-value:1.00e+00; Overlap: 16; Offset: 0; Orientation: Reverse complement</p>
<p>Myb (UP00092)</p>  <p><i>p</i>-value: 7.52e-12; <i>e</i>-value: 2.90e-09; <i>q</i>-value:5.80e-09; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.55e-09; <i>e</i>-value: 5.99e-07; <i>q</i>-value:1.20e-07; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.53e-08; <i>e</i>-value: 5.91e-06; <i>q</i>-value:6.73e-06; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 6.04e-08; <i>e</i>-value: 2.33e-01; <i>q</i>-value:4.66e-05; Overlap: 16; Offset: -1; Orientation: Normal</p>

<p>OafI (UP00298)</p>	 <p><i>p</i>-value: 3.22e-12; <i>e</i>-value: 2.86e-10; <i>q</i>-value:5.21e-10; Overlap: 21; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 8.29e-11; <i>e</i>-value: 7.38e-09; <i>q</i>-value:1.40e-08; Overlap: 20; Offset: 1; Orientation: Normal</p>	 <p><i>p</i>-value: 5.52e-10; <i>e</i>-value: 4.91e-08; <i>q</i>-value:9.16e-08; Overlap: 20; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 1.46e-09; <i>e</i>-value: 1.30e-07; <i>q</i>-value:2.35e-07; Overlap: 20; Offset: -1; Orientation: Normal</p>
<p>Otx1 (UP00229)</p>	 <p><i>p</i>-value: 2.96e-07; <i>e</i>-value: 1.14e-04; <i>q</i>-value:2.10e-04; Overlap: 15; Offset: 2; Orientation: Normal</p>	 <p><i>p</i>-value: 9.20e-09; <i>e</i>-value: 3.55e-06; <i>q</i>-value:6.50e-06; Overlap: 16; Offset: 1; Orientation: Normal</p>	 <p><i>p</i>-value: 1.23e-07; <i>e</i>-value: 4.74e-05; <i>q</i>-value:8.90e-05; Overlap: 16; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 5.32e-09; <i>e</i>-value: 2.05e-06; <i>q</i>-value:3.81e-06; Overlap: 16; Offset: 1; Orientation: Norma</p>
<p>Pbx1 (UP00185)</p>	 <p><i>p</i>-value: 9.60e-10; <i>e</i>-value: 3.70e-06; <i>q</i>-value:7.35e-06; Overlap: 15; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 3.14e-09; <i>e</i>-value: 1.21e-06; <i>q</i>-value:2.42e-06; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 2.80e-08; <i>e</i>-value: 1.08e-05; <i>q</i>-value:2.16e-05; Overlap: 16; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 7.14e-07; <i>e</i>-value: 2.76e-04; <i>q</i>-value:5.48e-04; Overlap: 16; Offset: -1; Orientation: Normal</p>
<p>Phd1 (UP00298)</p>	 <p><i>p</i>-value: 1.51e-06; <i>e</i>-value: 1.35e-04; <i>q</i>-value:2.66e-04; Overlap: 20; Offset: 1; Orientation: Normal</p>	 <p><i>p</i>-value: 5.22e-05; <i>e</i>-value: 4.65e-03; <i>q</i>-value:9.22e-03; Overlap: 21; Offset: 0; Orientation: Normal</p>	<p>None</p>	 <p><i>p</i>-value: 1.97e-03; <i>e</i>-value: 1.75e-01; <i>q</i>-value:1.74e-01; Overlap: 18; Offset: 3; Orientation: Normal</p>

<p>Pitx1 (UP00153)</p>	 <p><i>p</i>-value: 6.49e-09; <i>e</i>-value: 2.50e-06; <i>q</i>-value:4.94e-06; Overlap: 16; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 7.01e-08; <i>e</i>-value: 3.05e-05; <i>q</i>-value:6.04e-05; Overlap: 15; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 1.02e-06; <i>e</i>-value: 3.94e-10; <i>q</i>-value:7.82e-04; Overlap: 15; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 6.96e-05; <i>e</i>-value: 2.69e-02; <i>q</i>-value:5.33e-02; Overlap: 16; Offset: -1; Orientation: Normal</p>
<p>Rara (UP00048)</p>	 <p><i>p</i>-value: 8.54e-11; <i>e</i>-value: 3.30e-07; <i>q</i>-value:6.59e-07; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 2.14e-09; <i>e</i>-value: 8.25e-07; <i>q</i>-value:1.65e-06; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 6.66e-09; <i>e</i>-value: 2.57e-07; <i>q</i>-value:5.14e-07; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 6.45e-07; <i>e</i>-value: 2.49e-04; <i>q</i>-value:4.98e-04; Overlap: 14; Offset: -2; Orientation: Normal</p>
<p>Rph1 (UP00348)</p>	 <p><i>p</i>-value: 2.26e-06; <i>e</i>-value: 2.02e-03; <i>q</i>-value:1.94e-01; Overlap: 18; Offset: 3; Orientation: Normal</p>	 <p><i>p</i>-value: 1.56e-04; <i>e</i>-value: 1.39e+00; <i>q</i>-value:7.44e-01; Overlap: 19; Offset: 2; Orientation: Normal</p>	 <p><i>p</i>-value: 7.11e-02; <i>e</i>-value: 6.33e+00; <i>q</i>-value:9.99e-01; Overlap: 14; Offset: 7; Orientation: Reverse Complement</p>	 <p><i>p</i>-value: 2.77e-02; <i>e</i>-value: 2.47e+00; <i>q</i>-value:9.30e-01; Overlap: 18; Offset: 3; Orientation: Normal</p>
<p>Sox1 (UP00069)</p>	 <p><i>p</i>-value: 2.73e-10; <i>e</i>-value: 1.05e-07; <i>q</i>-value:2.03e-07; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.75e-08; <i>e</i>-value: 6.74e-06; <i>q</i>-value:1.280e-05; Overlap: 16; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 8.78e-08; <i>e</i>-value: 3.39e-05; <i>q</i>-value:6.57e-05; Overlap: 15; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 4.14e-06; <i>e</i>-value: 1.60e-03; <i>q</i>-value:3.09e-03; Overlap: 15; Offset: -1; Orientation: Normal</p>

<p>Srf (UP00077)</p>	 <p><i>p</i>-value: 1.31e-13; <i>e</i>-value: 5.05e-08; <i>q</i>-value: 1.01e-07; Overlap: 14; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.54e-10; <i>e</i>-value: 5.94e-08; <i>q</i>-value: 1.19e-07; Overlap: 14; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 2.94e-8; <i>e</i>-value: 1.13e-07; <i>q</i>-value: 2.27e-07; Overlap: 14; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.87e-08; <i>e</i>-value: 7.22e-06; <i>q</i>-value: 1.44e-05; Overlap: 13; Offset: 1; Orientation: Normal</p>
<p>Tea1 (UP00333)</p>	 <p><i>p</i>-value: 7.85e-08; <i>e</i>-value: 6.99e-06; <i>q</i>-value: 1.26e-05; Overlap: 18; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 3.89e-06; <i>e</i>-value: 3.47e-04; <i>q</i>-value: 6.56e-04; Overlap: 18; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 7.30e-06; <i>e</i>-value: 6.50e-04; <i>q</i>-value: 1.14e-03; Overlap: 18; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 5.30e-02; <i>e</i>-value: 4.72e+00; <i>q</i>-value: 8.85e-01; Overlap: 12; Offset: -8; Orientation: Reverse Complement</p>
<p>Usv1 (UP00343)</p>	 <p><i>p</i>-value: 1.62e-12; <i>e</i>-value: 1.44e-10; <i>q</i>-value: 2.58e-10; Overlap: 20; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 5.31e-11; <i>e</i>-value: 4.72e-09; <i>q</i>-value: 8.77e-09; Overlap: 19; Offset: -1; Orientation: Normal</p>	 <p><i>p</i>-value: 7.27e-10; <i>e</i>-value: 6.47e-08; <i>q</i>-value: 1.19e-07; Overlap: 18; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 3.03e-09; <i>e</i>-value: 2.70e-07; <i>q</i>-value: 5.09e-07; Overlap: 18; Offset: -2; Orientation: Normal</p>
<p>dYap1 (UP00327)</p>	 <p><i>p</i>-value: 5.45e-07; <i>e</i>-value: 4.85e-05; <i>q</i>-value: 9.53e-05; Overlap: 17; Offset: 3; Orientation: Normal</p>	 <p><i>p</i>-value: 7.03e-04; <i>e</i>-value: 6.25e-02; <i>q</i>-value: 1.25e-01; Overlap: 17; Offset: -3; Orientation: Normal</p>	 <p><i>p</i>-value: 1.03e-02; <i>e</i>-value: 9.13e-01; <i>q</i>-value: 5.38e-01; Overlap: 16; Offset: -4; Orientation: Normal</p>	 <p><i>p</i>-value: 2.57e-06; <i>e</i>-value: 2.29e-04; <i>q</i>-value: 4.49e-04; Overlap: 19; Offset: -1; Orientation: Normal</p>

Zbtb3 (UP00327)	 <p><i>p</i>-value: 2.33e-08; <i>e</i>-value: 8.98e-04; <i>q</i>-value:1.80e-03; Overlap: 15; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 5.53e-06; <i>e</i>-value: 2.13e-03; <i>q</i>-value:4.27e-03; Overlap: 15; Offset: 2; Orientation: Normal</p>	None	 <p><i>p</i>-value: 4.07e-04; <i>e</i>-value: 1.57e-01; <i>q</i>-value:3.14e-01; Overlap: 15; Offset: -2; Orientation: Normal</p>
Nrg1 (UP00270)	 <p><i>p</i>-value: 4.59e-13; <i>e</i>-value: 4.09e-11; <i>q</i>-value:8.17e-11; Overlap: 19; Offset: 1; Orientation: Normal</p>	 <p><i>p</i>-value: 1.65e-08; <i>e</i>-value: 1.47e-06; <i>q</i>-value:2.93e-06; Overlap: 20; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 3.46e-06; <i>e</i>-value: 3.08e-04; <i>q</i>-value:6.16e-04; Overlap: 18; Offset: -2; Orientation: Normal</p>	 <p><i>p</i>-value: 7.73e-03; <i>e</i>-value: 6.88e-05; <i>q</i>-value:1.37e-04; Overlap: 19; Offset: 1; Orientation: Normal</p>
Rax (UP00270)	 <p><i>p</i>-value: 4.91e-14; <i>e</i>-value: 1.89e-11; <i>q</i>-value:3.11e-11; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 5.01e-13; <i>e</i>-value: 1.93e-10; <i>q</i>-value:3.21e-10; Overlap: 17; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 7.19e-12; <i>e</i>-value: 2.78e-09; <i>q</i>-value:4.69e-09; Overlap: 17; Offset: 0; Orientation: Normal</p>	None
Zic1 (UP00270)	 <p><i>p</i>-value: 1.24e-14; <i>e</i>-value: 4.78e-09; <i>q</i>-value:9.36e-09; Overlap: 14; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 6.45e-11; <i>e</i>-value: 2.49e-08; <i>q</i>-value:4.87e-08; Overlap: 14; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 4.71e-11; <i>e</i>-value: 1.82e-08; <i>q</i>-value:3.54e-08; Overlap: 14; Offset: 0; Orientation: Normal</p>	 <p><i>p</i>-value: 1.37e-09; <i>e</i>-value: 5.30e-07; <i>q</i>-value:1.04e-06; Overlap: 13; Offset: 1; Orientation: Normal</p>

Supplementary Table S5. Experimental results of the single-task model and the dual-task model on the *in vivo* datasets

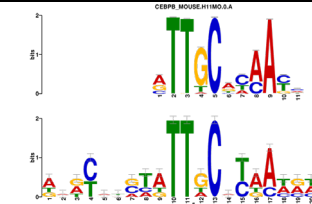
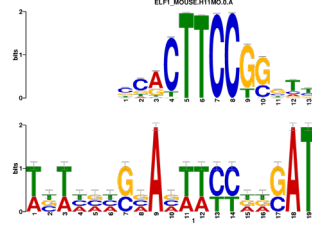
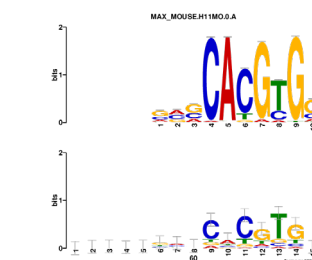
Dataset	IOU	F1-score
Cebpb	0.616	0.451
Elf1	0.635	0.937
Max	0.645	0.852
Myb	0.641	0.658
Srf	0.614	0.833
Average value	0.630	0.746

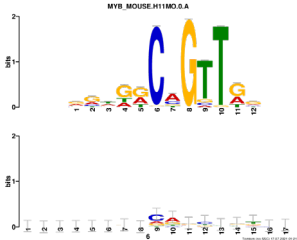
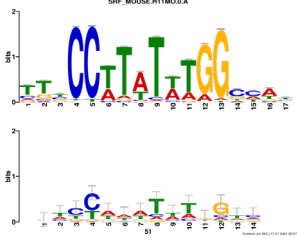
a. IOU and F1-score value of two-tasks model on 5 *in vivo* datasets

Dataset	IOU
Cebpb	0.793
Elf1	0.807
Max	0.81
Myb	0.788
Srf	0.74
Average value	0.788

b. IOU value of single model on 5 *in vivo* datasets

Supplementary Table S6. Comparison of the identified *in vivo* motifs by using *in vitro* trained models

TF name	Identified motifs	Information
Cebpb		p-value: 7.05e-04; e-value: 2.52e-01; q-value: 1.96 e-01; Overlap: 11; Offset: -8; Orientation: Reverse complement
Elf1		p-value: 4.04e-03; e-value: 1.45e+00; q-value: 6.39 e-01; Overlap: 14; Offset: -6; Orientation: Reverse complement
Max		p-value: 2.88e-03; e-value: 1.03e+00; q-value: 3.82 e-01; Overlap: 11; Offset: -5; Orientation: Normal

Myb		<p>p-value: 6.29e-03; e-value: 2.25e+00; q-value:6.43e-01; Overlap: 12; Offset: -3; Orientation: Normal</p>
Srf		<p>p-value: 8.33e-06; e-value: 2.98e-03; q-value:5.95e-03; Overlap: 14; Offset: 1; Orientation: Normal</p>