



Research article

GNet: An integrated context-aware neural framework for transcription factor binding signal at single nucleotide resolution prediction

Jujuan Zhuang, Kexin Feng*, Xinyang Teng and Cangzhi Jia*

School of Science, Dalian Maritime University, Dalian, Liaoning 116026, China

* **Correspondence:** Email: kexin1004_0919@163.com, cangzhijia@dlmu.edu.cn.

Supplementary

| | | HOCOMOCO | GNet | FCNsignal | BPNet | STREME | MEME |
|---------|-------|----------|--|--|--|--|--|
| GM12878 | MEF2A | | p-value:1.32E-18 E-value:5.31E-16 q-value:1.05E-15 | p-value:8.39E-15 E-value:3.37E-12 q-value:6.60E-12 | p-value:1.88E-16 E-value:7.53E-14 q-value:1.48E-13 | p-value:8.05E-05 E-value:3.23E-02 q-value:1.56E-02 | p-value:1.35E-05 E-value:5.42E-03 q-value:2.36E-03 |
| | ELK1 | | p-value:5.06E-08 E-value:2.03E-05 q-value:4.58E-06 | p-value:7.55E-07 E-value:3.03E-04 q-value:1.35E-04 | p-value:2.29E-06 E-value:9.17E-04 q-value:7.71E-04 | p-value:1.50E-02 E-value:6.03E+00 q-value:2.59E-01 | p-value:1.63E-04 E-value:6.52E-02 q-value:5.10E-03 |
| | EBF1 | | p-value:5.50E-15 E-value:2.21E-12 q-value:2.04E-12 | p-value:8.58E-15 E-value:3.44E-12 q-value:3.24E-12 | p-value:7.20E-12 E-value:2.89E-09 q-value:2.85E-09 | p-value:1.44E-08 E-value:5.76E-06 q-value:5.71E-06 | p-value:3.53E-05 E-value:6.41E-02 q-value:1.33E-02 |
| HeLa-S3 | JUND | | p-value:4.72E-09 E-value:1.89E-06 q-value:5.86E-07 | p-value:3.15E-08 E-value:1.26E-05 q-value:3.03E-06 | p-value:2.05E-08 E-value:8.23E-06 q-value:3.76E-06 | p-value:2.84E-09 E-value:1.14E-06 q-value:5.39E-07 | p-value:1.60E-04 E-value:1.42E-02 q-value:1.06E-02 |
| | MAX | | p-value:6.53E-07 E-value:2.62E-04 q-value:1.19E-04 | p-value:5.17E-06 E-value:2.07E-03 q-value:1.89E-03 | p-value:2.75E-07 E-value:1.10E-04 q-value:1.02E-04 | p-value:1.66E-04 E-value:6.67E-02 q-value:8.94E-03 | — |
| K562 | NFYA | | p-value:9.21E-12 E-value:3.69E-09 q-value:1.77E-09 | p-value:1.45E-08 E-value:5.81E-06 q-value:2.90E-06 | p-value:2.35E-03 E-value:9.44E-01 q-value:1.32E-01 | p-value:1.47E-12 E-value:5.90E-10 q-value:2.94E-10 | p-value:1.53E-05 E-value:6.14E-03 q-value:6.13E-03 |
| | FOSL1 | | p-value:1.20E-08 E-value:4.83E-06 q-value:2.32E-06 | p-value:2.15E-09 E-value:8.61E-07 q-value:2.02E-07 | p-value:6.26E-08 E-value:2.51E-05 q-value:1.92E-05 | p-value:2.50E-08 E-value:1.00E-05 q-value:3.17E-06 | p-value:1.40E-07 E-value:5.63E-05 q-value:1.98E-05 |
| | CBFB | | p-value:7.27E-07 E-value:2.92E-04 q-value:2.59E-04 | p-value:8.40E-04 E-value:3.37E-01 q-value:2.58E-01 | — | — | — |

Figure S1: Comparison of motifs predicted by different models on ChIP-seq datasets, and compared with the experimentally verified motif in the standard database HOCOMOCO.

Table S1. Details of all data sets.

| | Dataset | URL | Signal | p-value | IDR thresholded peaks | |
|------------------|------------------|---|---|-------------|-----------------------|-------------|
| ChIP-seq GM12878 | ARID3A | https://www.encodeproject.org/experiments/ENCSR778UBR | ENCFF503CWF | | ENCFF531CLQ | |
| | BACH1 | https://www.encodeproject.org/experiments/ENCSR636MKU | ENCFF859VLX | | ENCFF631RZH | |
| | BATF | https://www.encodeproject.org/experiments/ENCSR000BGT | ENCFF716MIY | | ENCFF728KFD | |
| | BCL11A | https://www.encodeproject.org/experiments/ENCSR000BHA | ENCFF197TWD | | ENCFF824QXX | |
| | BHLHE40 | https://www.encodeproject.org/experiments/ENCSR987MTA | ENCFF673GWX | | ENCFF445XYV | |
| | CEBPB | https://www.encodeproject.org/experiments/ENCSR000BRX | ENCFF836LCO | | ENCFF955YFB | |
| | CTCF | https://www.encodeproject.org/experiments/ENCSR000AKB | ENCFF803OXW | | ENCFF797SDL | |
| | CUX1 | https://www.encodeproject.org/experiments/ENCSR000DYR | ENCFF833GJC | | ENCFF451AII | |
| | E2F4 | https://www.encodeproject.org/experiments/ENCSR000DYY | ENCFF337MYH | | ENCFF744QAC | |
| | EBF1 | https://www.encodeproject.org/experiments/ENCSR000BGU | ENCFF810XRY | | ENCFF895MHN | |
| | ELK1 | https://www.encodeproject.org/experiments/ENCSR000DZB | ENCFF385JTY | | ENCFF164MPE | |
| | ETV6 | https://www.encodeproject.org/experiments/ENCSR626VUC | ENCFF203NIA | | ENCFF151UJT | |
| | JUND | https://www.encodeproject.org/experiments/ENCSR000DYS | ENCFF948QYK | | ENCFF134BQO | |
| | MEF2A | https://www.encodeproject.org/experiments/ENCSR000BKB | ENCFF497DHP | | ENCFF826GQU | |
| | MXI1 | https://www.encodeproject.org/experiments/ENCSR000DZI | ENCFF568SNH | | ENCFF376AEL | |
| | NFYB | https://www.encodeproject.org/experiments/ENCSR000DNM | ENCFF225JEM | | ENCFF156MUM | |
| | NR2C1 | https://www.encodeproject.org/experiments/ENCSR784VIQ | ENCFF986VHR | | ENCFF626EEU | |
| | PBX3 | https://www.encodeproject.org/experiments/ENCSR000BGR | ENCFF480ZQT | | ENCFF402DQD | |
| | SRF | https://www.encodeproject.org/experiments/ENCSR000BMI | ENCFF042HPQ | | ENCFF114CWH | |
| | USF1 | https://www.encodeproject.org/experiments/ENCSR000BGI | ENCFF902UDF | | ENCFF879TPT | |
| | ZEB1 | https://www.encodeproject.org/experiments/ENCSR000BND | ENCFF553ZPL | | ENCFF703XCL | |
| | ChIP-seq HeLa-S3 | CTCF | https://www.encodeproject.org/experiments/ENCSR000DUB | ENCFF959LLF | | ENCFF030QAM |
| | | E2F1 | https://www.encodeproject.org/experiments/ENCSR000EVM | ENCFF955LBT | | ENCFF872RSG |
| | | E2F6 | https://www.encodeproject.org/experiments/ENCSR000EVK | ENCFF573ZIC | | ENCFF160MCO |
| ELK1 | | https://www.encodeproject.org/experiments/ENCSR000ECI | ENCFF647ESC | | ENCFF279OML | |
| ELK4 | | https://www.encodeproject.org/experiments/ENCSR000EVI | ENCFF632TTT | | ENCFF582BHA | |

Continued on next page

| | Dataset | URL | Signal | p-value | IDR thresholded peaks |
|------------------|---------|---|-------------|---------|-----------------------|
| ChIP-seq HeLa-S3 | JUND | https://www.encodeproject.org/experiments/ENCSR000EDH | ENCFF756JBL | | ENCFF995QMN |
| | MAFF | https://www.encodeproject.org/experiments/ENCSR140DSL | ENCFF320OIL | | ENCFF280WZN |
| | MAX | https://www.encodeproject.org/experiments/ENCSR000ECN | ENCFF323ASO | | ENCFF666XCS |
| | MAZ | https://www.encodeproject.org/experiments/ENCSR000ECL | ENCFF036NCL | | ENCFF424JHD |
| | REST | https://www.encodeproject.org/experiments/ENCSR000BMN | ENCFF103IAQ | | ENCFF096JDA |
| | RFX5 | https://www.encodeproject.org/experiments/ENCSR000ECX | ENCFF482YCH | | ENCFF438ZEN |
| | TBP | https://www.encodeproject.org/experiments/ENCSR000EDD | ENCFF936GLP | | ENCFF363NQY |
| ChIP-seq K562 | ARNT | https://www.encodeproject.org/experiments/ENCSR155KHM | ENCFF430XHC | | ENCFF728ITJ |
| | CBFB | https://www.encodeproject.org/experiments/ENCSR116NDV | ENCFF829AHU | | ENCFF802NHC |
| | CREB1 | https://www.encodeproject.org/experiments/ENCSR000BSO | ENCFF014CIM | | ENCFF193LLN |
| | ETS1 | https://www.encodeproject.org/experiments/ENCSR000BKQ | ENCFF490BCW | | ENCFF159OUK |
| | FOSL1 | https://www.encodeproject.org/experiments/ENCSR000BMV | ENCFF970NEI | | ENCFF004HXL |
| | GATA2 | https://www.encodeproject.org/experiments/ENCSR257RKC | ENCFF343XFN | | ENCFF772OKO |
| | JUN | https://www.encodeproject.org/experiments/ENCSR000EFS | ENCFF756UVW | | ENCFF190CGV |
| | JUNB | https://www.encodeproject.org/experiments/ENCSR525VAT | ENCFF114SLT | | ENCFF510IDP |
| | JUND | https://www.encodeproject.org/experiments/ENCSR000EGN | ENCFF253NTK | | ENCFF306SZL |
| | KLF16 | https://www.encodeproject.org/experiments/ENCSR397DQC | ENCFF374RSO | | ENCFF585OJO |
| | MXI1 | https://www.encodeproject.org/experiments/ENCSR000EGZ | ENCFF845HTM | | ENCFF068IGH |
| | MYC | https://www.encodeproject.org/experiments/ENCSR744JJU | ENCFF670SRV | | ENCFF114VAI |
| | NEUROD1 | https://www.encodeproject.org/experiments/ENCSR986CDX | ENCFF112BYB | | ENCFF625QHR |
| | NFYA | https://www.encodeproject.org/experiments/ENCSR000EGR | ENCFF643SJZ | | ENCFF908HSL |
| | NR2C1 | https://www.encodeproject.org/experiments/ENCSR742IDN | ENCFF525KET | | ENCFF469ZBB |
| | NR2F6 | https://www.encodeproject.org/experiments/ENCSR128KLX | ENCFF918KYJ | | ENCFF647RGI |
| | TCF7 | https://www.encodeproject.org/experiments/ENCSR863KUB | ENCFF986NMS | | ENCFF736XUU |
| | YY1 | https://www.encodeproject.org/experiments/ENCSR000BMH | ENCFF927TDC | | ENCFF398UQZ |
| | ZBTB33 | https://www.encodeproject.org/experiments/ENCSR876GXA | ENCFF288PSZ | | ENCFF917RIN |
| | ZNF740 | https://www.encodeproject.org/experiments/ENCSR532EMP | ENCFF600BUX | | ENCFF004YCK |

Continued on next page

| | Dataset | URL | Signal | p-value | IDR thresholded peaks |
|-------------|---------|---|---|-------------|-----------------------|
| ATAC-seq | A549 | https://www.encodeproject.org/experiments/ENCSR032RGS | ENCFF399KCR | | ENCFF674RFR |
| | GM12878 | https://www.encodeproject.org/experiments/ENCSR095QNB | ENCFF180ZAY | | ENCFF945SYZ |
| | K562 | https://www.encodeproject.org/experiments/ENCSR483RKN | ENCFF600FDOO | | ENCFF925CYR |
| | IMR90 | https://www.encodeproject.org/experiments/ENCSR200OML | ENCFF770EAV | | ENCFF114GDS |
| | HepG2 | https://www.encodeproject.org/experiments/ENCSR042AWH | ENCFF285FQS | | ENCFF438JMM |
| | MCF7 | https://www.encodeproject.org/experiments/ENCSR422SUG | ENCFF976UNK | | ENCFF882OVP |
| Mus C2C12 | FOSL1 | https://www.encodeproject.org/experiments/ENCSR000AIK | ENCFF983CVX | | ENCFF278EHP |
| | MAX | https://www.encodeproject.org/experiments/ENCSR000AIO | ENCFF877FZW | | ENCFF041FSC |
| | CEBPB | https://www.encodeproject.org/experiments/ENCSR000AIV | ENCFF688ZJR | | ENCFF174GMB |
| | USF1 | https://www.encodeproject.org/experiments/ENCSR000AIQ | ENCFF747WZN | | ENCFF333AWB |
| Mus G1E | GATA2 | https://www.encodeproject.org/experiments/ENCSR000DIE | ENCFF737BQS | | ENCFF832ZPK |
| Mus CH12.LX | JUN | https://www.encodeproject.org/experiments/ENCSR000ERO | ENCFF401NRB | | ENCFF937POX |
| | BHLHE40 | https://www.encodeproject.org/experiments/ENCSR000ERC | ENCFF229EVK | | ENCFF918PVE |
| | TBP | https://www.encodeproject.org/experiments/ENCSR000ERP | ENCFF883ODR | | ENCFF523MOI |
| | MYC | https://www.encodeproject.org/experiments/ENCSR000ERN | ENCFF901VFB | | ENCFF342ASE |
| | MXI1 | https://www.encodeproject.org/experiments/ENCSR000ERE | ENCFF439CGA | | ENCFF172JAS |
| | MEF2A | https://www.encodeproject.org/experiments/ENCSR806JZK | ENCFF458UDN | | ENCFF028EUT |
| | Mus MEL | JUND | https://www.encodeproject.org/experiments/ENCSR000ETZ | ENCFF886PYK | |
| CTCF | | https://www.encodeproject.org/experiments/ENCSR000ETQ | ENCFF993KMG | | ENCFF223ASW |
| MAX | | https://www.encodeproject.org/experiments/ENCSR000ETX | ENCFF933SHE | | ENCFF262ITC |
| MAZ | | https://www.encodeproject.org/experiments/ENCSR000ESL | ENCFF008RGX | | ENCFF980MPW |
| E2F4 | | https://www.encodeproject.org/experiments/ENCSR000ETY | ENCFF927MMJ | | ENCFF734ZCR |
| MXI1 | | https://www.encodeproject.org/experiments/ENCSR000ETN | ENCFF989MCQ | | ENCFF067LMI |
| USF1 | | https://www.encodeproject.org/experiments/ENCSR705HGT | ENCFF750OCK | | ENCFF281FXQ |

Table S2. The comparison results of the first two tasks on 53 ChIP-seq datasets.

| Cell line | TF | BPNet | | | | FCNsignal | | | | GNet | | | |
|-----------|---------|-------|-------|-------|-------|-----------|-------|-------|-------|-------|-------|-------|-------|
| | | MSE | PCC | AUC | AUPRC | MSE | PCC | AUC | AUPRC | MSE | PCC | AUC | AUPRC |
| K562 | ARNT | 0.187 | 0.714 | 0.778 | 0.792 | 0.158 | 0.741 | 0.87 | 0.876 | 0.15 | 0.756 | 0.877 | 0.88 |
| | CBFB | 0.097 | 0.583 | 0.908 | 0.909 | 0.083 | 0.638 | 0.936 | 0.94 | 0.08 | 0.654 | 0.938 | 0.943 |
| | CREB1 | 0.121 | 0.796 | 0.954 | 0.956 | 0.119 | 0.795 | 0.962 | 0.964 | 0.114 | 0.816 | 0.966 | 0.966 |
| | ETS1 | 0.105 | 0.705 | 0.961 | 0.965 | 0.098 | 0.731 | 0.976 | 0.978 | 0.089 | 0.75 | 0.979 | 0.98 |
| | FOSL1 | 0.108 | 0.833 | 0.928 | 0.925 | 0.073 | 0.888 | 0.974 | 0.976 | 0.068 | 0.888 | 0.973 | 0.973 |
| | GATA2 | 0.106 | 0.817 | 0.927 | 0.932 | 0.105 | 0.823 | 0.941 | 0.947 | 0.11 | 0.815 | 0.935 | 0.94 |
| | JUN | 0.078 | 0.834 | 0.964 | 0.966 | 0.079 | 0.825 | 0.96 | 0.959 | 0.078 | 0.842 | 0.968 | 0.967 |
| | JUNB | 0.155 | 0.761 | 0.835 | 0.826 | 0.117 | 0.799 | 0.945 | 0.954 | 0.108 | 0.849 | 0.947 | 0.956 |
| | JUND | 0.134 | 0.849 | 0.951 | 0.954 | 0.115 | 0.868 | 0.963 | 0.961 | 0.115 | 0.877 | 0.969 | 0.968 |
| | KLF16 | 0.156 | 0.77 | 0.954 | 0.96 | 0.16 | 0.776 | 0.952 | 0.958 | 0.145 | 0.785 | 0.962 | 0.966 |
| | MXI1 | 0.172 | 0.527 | 0.84 | 0.855 | 0.117 | 0.7 | 0.944 | 0.945 | 0.104 | 0.732 | 0.954 | 0.959 |
| | MYC | 0.207 | 0.718 | 0.905 | 0.918 | 0.206 | 0.734 | 0.926 | 0.939 | 0.171 | 0.757 | 0.929 | 0.94 |
| | NEUROD1 | 0.105 | 0.762 | 0.947 | 0.945 | 0.109 | 0.77 | 0.955 | 0.954 | 0.1 | 0.793 | 0.964 | 0.965 |
| | NFYA | 0.185 | 0.633 | 0.848 | 0.852 | 0.098 | 0.822 | 0.978 | 0.983 | 0.092 | 0.841 | 0.983 | 0.985 |
| | NR2C1 | 0.29 | 0.816 | 0.918 | 0.927 | 0.218 | 0.812 | 0.912 | 0.923 | 0.304 | 0.826 | 0.917 | 0.928 |
| | NR2F6 | 0.171 | 0.653 | 0.868 | 0.902 | 0.186 | 0.628 | 0.842 | 0.869 | 0.176 | 0.64 | 0.825 | 0.855 |
| | TCF7 | 0.08 | 0.58 | 0.891 | 0.903 | 0.064 | 0.628 | 0.919 | 0.907 | 0.058 | 0.654 | 0.923 | 0.912 |
| | YY1 | 0.133 | 0.814 | 0.95 | 0.953 | 0.132 | 0.816 | 0.956 | 0.957 | 0.116 | 0.831 | 0.96 | 0.958 |
| | ZBTB33 | 0.192 | 0.897 | 0.961 | 0.959 | 0.212 | 0.888 | 0.961 | 0.961 | 0.201 | 0.895 | 0.961 | 0.96 |
| | ZNF740 | 0.248 | 0.7 | 0.815 | 0.851 | 0.345 | 0.62 | 0.795 | 0.836 | 0.29 | 0.667 | 0.816 | 0.856 |
| HeLa-S3 | CTCF | 0.081 | 0.934 | 0.981 | 0.982 | 0.069 | 0.943 | 0.99 | 0.99 | 0.067 | 0.941 | 0.991 | 0.991 |
| | E2F1 | 0.135 | 0.84 | 0.992 | 0.992 | 0.141 | 0.825 | 0.992 | 0.991 | 0.132 | 0.838 | 0.993 | 0.993 |
| | E2F6 | 0.107 | 0.639 | 0.959 | 0.954 | 0.109 | 0.644 | 0.962 | 0.96 | 0.1 | 0.686 | 0.963 | 0.958 |
| | ELK1 | 0.062 | 0.681 | 0.961 | 0.978 | 0.068 | 0.671 | 0.961 | 0.982 | 0.063 | 0.691 | 0.974 | 0.986 |
| | ELK4 | 0.095 | 0.793 | 0.964 | 0.966 | 0.099 | 0.794 | 0.966 | 0.97 | 0.095 | 0.795 | 0.976 | 0.977 |

Continued on next page

| Cell line | TF | BPNet | | | | FCNsignal | | | | GNet | | | |
|-----------|---------|-------|-------|-------|-------|-----------|-------|-------|-------|-------|-------|-------|-------|
| | | MSE | PCC | AUC | AUPRC | MSE | PCC | AUC | AUPRC | MSE | PCC | AUC | AUPRC |
| HeLa-S3 | JUND | 0.082 | 0.9 | 0.961 | 0.966 | 0.097 | 0.894 | 0.969 | 0.972 | 0.097 | 0.892 | 0.968 | 0.97 |
| | MAFF | 0.064 | 0.894 | 0.974 | 0.976 | 0.07 | 0.892 | 0.972 | 0.973 | 0.069 | 0.901 | 0.978 | 0.979 |
| | MAX | 0.102 | 0.794 | 0.956 | 0.953 | 0.112 | 0.771 | 0.945 | 0.944 | 0.107 | 0.795 | 0.955 | 0.953 |
| | MAZ | 0.101 | 0.687 | 0.952 | 0.953 | 0.093 | 0.711 | 0.967 | 0.967 | 0.086 | 0.736 | 0.974 | 0.976 |
| | REST | 0.12 | 0.87 | 0.955 | 0.96 | 0.133 | 0.832 | 0.956 | 0.962 | 0.128 | 0.878 | 0.961 | 0.965 |
| | RFX5 | 0.124 | 0.744 | 0.863 | 0.869 | 0.093 | 0.805 | 0.942 | 0.941 | 0.086 | 0.812 | 0.948 | 0.944 |
| | TBP | 0.174 | 0.712 | 0.906 | 0.917 | 0.15 | 0.742 | 0.935 | 0.94 | 0.156 | 0.763 | 0.948 | 0.951 |
| GM12878 | ARID3A | 0.158 | 0.735 | 0.868 | 0.867 | 0.16 | 0.75 | 0.929 | 0.929 | 0.119 | 0.785 | 0.934 | 0.926 |
| | BACH1 | 0.178 | 0.775 | 0.94 | 0.946 | 0.173 | 0.764 | 0.939 | 0.943 | 0.167 | 0.786 | 0.942 | 0.948 |
| | BATF | 0.113 | 0.89 | 0.969 | 0.969 | 0.123 | 0.888 | 0.967 | 0.967 | 0.116 | 0.895 | 0.97 | 0.972 |
| | BCL11A | 0.115 | 0.829 | 0.945 | 0.946 | 0.113 | 0.823 | 0.955 | 0.954 | 0.105 | 0.844 | 0.959 | 0.957 |
| | BHLHE40 | 0.228 | 0.868 | 0.954 | 0.953 | 0.311 | 0.828 | 0.908 | 0.911 | 0.229 | 0.867 | 0.953 | 0.951 |
| | CEBPB | 0.064 | 0.659 | 0.882 | 0.894 | 0.058 | 0.678 | 0.873 | 0.883 | 0.058 | 0.696 | 0.917 | 0.928 |
| | CTCF | 0.127 | 0.915 | 0.974 | 0.976 | 0.09 | 0.936 | 0.991 | 0.992 | 0.092 | 0.935 | 0.992 | 0.992 |
| | CUX1 | 0.12 | 0.597 | 0.845 | 0.868 | 0.121 | 0.595 | 0.854 | 0.865 | 0.108 | 0.633 | 0.891 | 0.904 |
| | E2F4 | 0.186 | 0.624 | 0.96 | 0.965 | 0.122 | 0.725 | 0.983 | 0.985 | 0.121 | 0.741 | 0.984 | 0.987 |
| | EBF1 | 0.163 | 0.851 | 0.948 | 0.945 | 0.176 | 0.835 | 0.954 | 0.953 | 0.174 | 0.858 | 0.953 | 0.95 |
| | ELK1 | 0.077 | 0.741 | 0.976 | 0.978 | 0.088 | 0.71 | 0.972 | 0.975 | 0.081 | 0.747 | 0.98 | 0.981 |
| | ETV6 | 0.118 | 0.804 | 0.948 | 0.955 | 0.141 | 0.781 | 0.947 | 0.953 | 0.127 | 0.817 | 0.957 | 0.962 |
| | JUND | 0.081 | 0.787 | 0.936 | 0.935 | 0.082 | 0.782 | 0.951 | 0.952 | 0.074 | 0.808 | 0.955 | 0.953 |
| | MEF2A | 0.11 | 0.822 | 0.938 | 0.94 | 0.12 | 0.811 | 0.94 | 0.944 | 0.111 | 0.828 | 0.945 | 0.948 |
| | MXI1 | 0.129 | 0.659 | 0.936 | 0.943 | 0.132 | 0.653 | 0.944 | 0.951 | 0.131 | 0.675 | 0.949 | 0.956 |
| | NFYB | 0.15 | 0.862 | 0.963 | 0.966 | 0.146 | 0.861 | 0.965 | 0.967 | 0.142 | 0.875 | 0.971 | 0.973 |
| | NR2C1 | 0.244 | 0.809 | 0.941 | 0.944 | 0.275 | 0.793 | 0.951 | 0.955 | 0.237 | 0.815 | 0.947 | 0.951 |
| PBX3 | 0.115 | 0.814 | 0.935 | 0.935 | 0.108 | 0.822 | 0.938 | 0.936 | 0.107 | 0.829 | 0.948 | 0.946 | |
| SRF | 0.11 | 0.758 | 0.921 | 0.921 | 0.094 | 0.787 | 0.949 | 0.944 | 0.102 | 0.793 | 0.953 | 0.951 | |

Continued on next page

| Cell line | TF | BPNet | | | | FCNsignal | | | | GNet | | | |
|-----------|------|---------|---------|--------|-------------|-----------|--------|--------|-------------|---------------|---------------|---------------|----------------|
| | | MSE | PCC | AUC | AUPRC | MSE | PCC | AUC | AUPRC | MSE | PCC | AUC | AUPRC |
| GM12878 | USF1 | 0.072 | 0.886 | 0.981 | 0.98 | 0.07 | 0.893 | 0.984 | 0.982 | 0.068 | 0.897 | 0.985 | 0.983 |
| | ZEB1 | 0.109 | 0.784 | 0.957 | 0.957 | 0.108 | 0.789 | 0.96 | 0.959 | 0.112 | 0.791 | 0.961 | 0.96 |
| The mean | | 0.13290 | 0.76828 | 0.9291 | 0.933943396 | 0.1290 | 0.7798 | 0.9448 | 0.948660377 | 0.1214 | 0.7983 | 0.9513 | 0.95432 |
| | | 566 | 3019 | 32075 | | 37736 | 11321 | 67925 | | 33962 | 20755 | 39623 | 0755 |

Table S3. The comparison results on motif recognition tasks for ChIP-seq data.

| Cell line | TF | MEME | | | STREME | | | BPNet | | | FCNsignal | | | GNet | | |
|-----------|---------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | | -log2(p-value) | -log2(E-value) | -log2(q-value) | -log2(p-value) | -log2(E-value) | -log2(q-value) | -log2(p-value) | -log2(E-value) | -log2(q-value) | -log2(p-value) | -log2(E-value) | -log2(q-value) | -log2(p-value) | -log2(E-value) | -log2(q-value) |
| K562 | ARNT | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 7.973 | 0.000 | 1.776 |
| | CBFB | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 10.217 | 1.569 | 1.955 | 20.392 | 11.742 | 11.915 |
| | CREB1 | 0.000 | 0.000 | 0.000 | 17.194 | 8.544 | 10.038 | 16.846 | 8.200 | 9.909 | 18.347 | 9.703 | 11.456 | 18.029 | 9.381 | 11.456 |
| | ETS1 | 7.146 | 0.000 | 3.533 | 8.902 | 0.255 | 4.179 | 0.000 | 0.000 | 0.000 | 16.842 | 8.196 | 10.906 | 9.703 | 1.053 | 4.005 |
| | FOSL1 | 22.768 | 14.117 | 15.624 | 25.253 | 16.610 | 18.267 | 23.929 | 15.282 | 15.669 | 28.793 | 20.147 | 22.239 | 26.312 | 17.660 | 18.717 |
| | GATA2 | 6.288 | 0.000 | 1.431 | 13.712 | 5.064 | 6.825 | 22.305 | 13.657 | 15.271 | 20.454 | 11.807 | 11.813 | 17.639 | 8.995 | 9.587 |
| | JUN | 9.966 | 1.315 | 3.591 | 22.932 | 14.279 | 15.941 | 26.628 | 17.983 | 19.986 | 29.553 | 20.903 | 22.967 | 27.890 | 19.245 | 20.314 |
| | JUNB | 22.116 | 13.467 | 16.054 | 18.621 | 9.974 | 12.619 | 12.813 | 4.164 | 7.114 | 16.642 | 7.995 | 8.395 | 19.934 | 11.288 | 13.929 |
| | JUND | 6.506 | 0.000 | 0.974 | 30.990 | 22.343 | 22.454 | 26.349 | 17.705 | 18.542 | 27.865 | 19.218 | 19.917 | 30.420 | 21.773 | 21.826 |
| | KLF16 | 14.918 | 6.276 | 8.988 | 20.139 | 11.493 | 13.295 | 10.113 | 1.466 | 5.601 | 12.007 | 3.358 | 6.030 | 10.912 | 2.265 | 6.573 |
| | MXI1 | 5.363 | 0.000 | 2.351 | 13.551 | 4.904 | 6.793 | 11.813 | 3.158 | 5.195 | 23.103 | 14.453 | 14.502 | 27.827 | 19.183 | 19.290 |
| | MYC | 6.806 | 0.000 | 2.272 | 11.937 | 3.293 | 5.695 | 7.306 | 0.000 | 0.336 | 12.131 | 3.485 | 4.503 | 17.584 | 8.937 | 9.032 |
| | NEUROD1 | 0.000 | 0.000 | 0.000 | 5.630 | 0.000 | 0.706 | 6.190 | 0.000 | 1.035 | 11.207 | 2.556 | 2.599 | 8.506 | 0.000 | 1.613 |
| | NFYA | 15.996 | 7.348 | 7.350 | 39.307 | 30.659 | 31.663 | 8.733 | 0.083 | 2.921 | 26.039 | 17.393 | 18.396 | 36.660 | 28.014 | 29.074 |
| | NR2C1 | 5.351 | 0.000 | 0.766 | 9.828 | 1.181 | 5.118 | 16.145 | 7.501 | 9.610 | 16.829 | 8.179 | 9.909 | 17.534 | 8.889 | 11.284 |
| NR2F6 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | |
| TCF7 | 12.210 | 3.563 | 4.673 | 19.436 | 10.789 | 10.815 | 0.000 | 0.000 | 0.000 | 12.197 | 3.553 | 3.682 | 32.578 | 23.929 | 24.003 | |

Continued on next page

| Cell line | TF | MEME | | | STREME | | | BPNet | | | FCNsignal | | | GNet | | |
|--------------|---------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | | -log2(p- value) | -log2(E- value) | -log2(q- value) | -log2(p- value) | -log2(E- value) | -log2(q- value) | -log2(p- value) | -log2(E- value) | -log2(q- value) | -log2(p- value) | -log2(E- value) | -log2(q- value) | -log2(p- value) | -log2(E- value) | -log2(q- value) |
| K562 | YY1 | 6.935 | 0.000 | 1.727 | 31.312 | 22.669 | 22.678 | 28.949 | 20.301 | 20.324 | 31.286 | 22.640 | 22.669 | 26.596 | 17.950 | 17.990 |
| | ZBTB33 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 12.224 | 3.575 | 3.647 | 16.863 | 8.213 | 8.222 | 14.977 | 6.334 | 6.430 |
| | ZNF740 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| HeLa- S3 | CTCF | 7.621 | 0.000 | 1.949 | 54.329 | 45.683 | 45.741 | 60.522 | 51.876 | 51.962 | 63.406 | 54.759 | 54.845 | 28.715 | 20.064 | 20.161 |
| | E2F1 | 11.335 | 2.690 | 6.744 | 8.222 | 0.000 | 4.424 | 10.937 | 2.286 | 3.671 | 11.108 | 2.458 | 6.158 | 10.639 | 1.994 | 3.804 |
| | E2F6 | 11.160 | 2.515 | 6.820 | 6.674 | 0.000 | 2.889 | 15.187 | 6.546 | 6.744 | 7.944 | 0.000 | 2.989 | 18.411 | 9.764 | 9.802 |
| | ELK1 | 5.764 | 0.000 | 1.921 | 8.861 | 0.213 | 3.364 | 16.408 | 7.764 | 10.023 | 22.842 | 14.198 | 15.299 | 26.771 | 18.124 | 18.946 |
| | ELK4 | 7.959 | 0.000 | 4.049 | 10.742 | 2.095 | 4.471 | 32.886 | 24.236 | 24.349 | 32.921 | 24.271 | 24.380 | 34.523 | 25.871 | 26.000 |
| | JUND | 12.610 | 3.964 | 6.560 | 28.391 | 19.743 | 20.823 | 25.540 | 16.891 | 18.021 | 24.920 | 16.276 | 18.332 | 27.659 | 19.013 | 20.703 |
| | MAFF | 11.476 | 2.826 | 4.316 | 41.077 | 32.429 | 34.248 | 31.002 | 22.358 | 22.540 | 36.315 | 27.662 | 27.865 | 37.530 | 28.883 | 28.987 |
| | MAX | 8.163 | 0.000 | 3.095 | 12.557 | 3.906 | 6.806 | 21.794 | 13.150 | 13.259 | 17.561 | 8.916 | 9.047 | 20.546 | 11.898 | 13.037 |
| | MAZ | 8.875 | 0.231 | 2.900 | 10.632 | 1.983 | 5.282 | 18.236 | 9.587 | 10.229 | 12.920 | 4.276 | 5.904 | 17.094 | 11.698 | 11.046 |
| | REST | 34.845 | 26.197 | 26.337 | 26.412 | 17.762 | 17.814 | 41.009 | 32.363 | 32.363 | 27.222 | 18.575 | 18.627 | 42.591 | 33.947 | 33.954 |
| | RFX5 | 8.013 | 0.000 | 2.727 | 7.703 | 0.000 | 1.071 | 9.815 | 1.165 | 3.146 | 49.615 | 40.968 | 39.974 | 30.889 | 22.239 | 21.279 |
| | TBP | 6.918 | 0.000 | 2.591 | 7.822 | 0.000 | 0.801 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 6.984 | 0.000 | 0.930 |
| GM12 878 | ARID3A | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 6.868 | 0.000 | 0.582 |
| | BACH1 | 5.673 | 0.000 | 1.494 | 5.449 | 0.000 | 1.667 | 14.411 | 5.764 | 6.925 | 17.320 | 8.673 | 9.512 | 21.607 | 12.954 | 14.285 |
| | BATF | 7.310 | 0.000 | 0.542 | 16.177 | 7.530 | 7.730 | 47.237 | 38.588 | 37.627 | 73.868 | 65.224 | 64.221 | 72.527 | 63.880 | 62.890 |
| | BCL11A | 8.501 | 0.000 | 3.488 | 22.920 | 14.271 | 16.034 | 28.402 | 19.755 | 21.454 | 36.056 | 24.087 | 26.547 | 33.219 | 24.568 | 25.665 |
| | BHLHE40 | 16.994 | 8.348 | 8.538 | 20.257 | 11.610 | 12.055 | 20.412 | 11.762 | 12.505 | 20.068 | 11.420 | 11.476 | 20.781 | 12.131 | 12.259 |
| | CEBPB | 0.000 | 0.000 | 0.000 | 7.374 | 0.000 | 2.411 | 6.357 | 0.000 | 2.293 | 7.292 | 0.000 | 2.279 | 9.324 | 0.676 | 3.035 |
| | CTCF | 54.720 | 46.074 | 46.197 | 58.869 | 50.219 | 50.252 | 53.985 | 45.337 | 44.340 | 57.923 | 49.273 | 48.278 | 56.186 | 47.545 | 46.549 |
| | CUX1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 7.336 | 0.000 | 2.366 |
| | E2F4 | 9.450 | 0.806 | 4.184 | 11.876 | 3.224 | 5.078 | 7.799 | 0.000 | 1.117 | 17.172 | 8.527 | 8.673 | 19.807 | 11.160 | 11.284 |
| | EBF1 | 14.790 | 6.138 | 6.232 | 26.049 | 17.405 | 17.418 | 37.015 | 28.366 | 28.386 | 46.728 | 38.081 | 38.167 | 47.369 | 38.719 | 38.835 |

Continued on next page

| Cell line | TF | MEME | | | STREME | | | BPNet | | | FCNsignal | | | GNet | | |
|--------------|-------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | | -log2(p- value) | -log2(E- value) | -log2(q- value) | -log2(p- value) | -log2(E- value) | -log2(q- value) | -log2(p- value) | -log2(E- value) | -log2(q- value) | -log2(p- value) | -log2(E- value) | -log2(q- value) | -log2(p- value) | -log2(E- value) | -log2(q- value) |
| GM12 | ELK1 | 12.583 | 3.939 | 7.615 | 6.059 | 0.000 | 1.949 | 18.736 | 10.091 | 10.341 | 20.337 | 11.688 | 12.855 | 24.236 | 15.588 | 17.736 |
| 878 | ETV6 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | JUND | 11.200 | 2.548 | 5.393 | 18.542 | 9.895 | 12.273 | 20.619 | 11.972 | 14.129 | 16.851 | 8.204 | 10.614 | 17.222 | 8.576 | 10.424 |
| | MEF2A | 16.177 | 7.527 | 8.727 | 13.601 | 4.952 | 6.002 | 52.240 | 43.594 | 42.619 | 46.760 | 38.110 | 37.141 | 59.394 | 50.742 | 49.759 |
| | MXI1 | 0.000 | 0.000 | 0.000 | 15.353 | 6.703 | 8.348 | 21.176 | 12.531 | 13.001 | 29.964 | 21.316 | 21.414 | 26.242 | 17.592 | 17.702 |
| | NFYB | 11.064 | 2.419 | 4.016 | 12.259 | 3.612 | 3.638 | 35.214 | 26.561 | 27.570 | 30.959 | 22.312 | 22.603 | 32.170 | 23.526 | 23.531 |
| | NR2C1 | 12.655 | 4.007 | 7.140 | 14.599 | 5.948 | 9.134 | 21.084 | 12.456 | 13.590 | 17.477 | 8.828 | 10.563 | 19.405 | 10.762 | 11.926 |
| | PBX3 | 9.352 | 0.701 | 3.146 | 30.716 | 22.071 | 23.350 | 30.391 | 21.743 | 21.272 | 38.327 | 29.683 | 28.696 | 46.788 | 38.141 | 37.158 |
| | SRF | 10.810 | 2.165 | 5.540 | 10.653 | 2.006 | 2.653 | 23.239 | 14.599 | 13.597 | 29.391 | 20.745 | 19.768 | 27.284 | 18.639 | 17.663 |
| | USF1 | 20.815 | 12.170 | 13.302 | 29.646 | 21.006 | 21.294 | 33.870 | 25.225 | 26.349 | 38.154 | 29.508 | 29.610 | 33.854 | 25.208 | 26.324 |
| | ZEB1 | 6.310 | 0.000 | 1.900 | 10.583 | 1.932 | 3.252 | 12.392 | 3.743 | 6.265 | 11.660 | 3.012 | 3.913 | 16.841 | 8.192 | 8.376 |
| The mean | | 9.538 | 3.422 | 5.034 | 15.908 | 8.835 | 10.177 | 18.835 | 11.951 | 12.733 | 22.518 | 14.989 | 15.659 | 24.005 | 16.014 | 16.713 |

Table S4. Comparison results of attention mechanisms on PBM data.

| uPBM | Dual external attention | | External attention | | Self-Attention | |
|--------------------------|-------------------------|--------|--------------------|-----------|----------------|------------------|
| | R ² | PCC | R ² | PCC | R ² | PCC |
| TF_3_Foxo6_pTH3477_HK | 0.635 | 0.8 | 0.592 | 0.773 | 0.616 | 0.789 |
| TF_10_Nfil3_pTH3041_HK | 0.5024 | 0.7114 | 0.4767 | 0.6949 | 0.5028 | 0.7112 |
| TF_15_Pit1_pTH4326_HK | 0.6745 | 0.8229 | 0.7253 | 0.8539 | 0.7347 | 0.8602 |
| TF_26_Tfec_pTH2885_HK | 0.5849 | 0.768 | 0.5974 | 0.7809 | 0.6284 | 0.7964 |
| TF_36_At4_pTH1014_ME | 0.5283 | 0.7324 | 0.4721 | 0.7093 | 0.4923 | 0.7117 |
| TF_46_Nhlh2_pTH3037_ME | 0.4332 | 0.6663 | 0.3408 | 0.6196 | 0.4218 | 0.6529 |
| TF_51_Pou1f1_pTH3818_ME | 0.6211 | 0.7896 | 0.6168 | 0.7889 | 0.6536 | 0.8101 |
| TF_63_Zkscan5_pTH2283_ME | 0.572 | 0.759 | 0.517 | 0.724 | 0.5 | 0.739 |
| The mean | 0.568925 | 0.7562 | 0.5422625 | 0.7430625 | 0.5687 | 0.7588125 |



AIMS Press

© 2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)