S6 Table: 4conv2pool4norm recall for regulatory sequence prediction for different cell lines. Ten CNN models of the 4conv2pool4norm architecture were trained each on DHS datasets (positive) and corresponding negative sets of $k$-mer shuffled sequences ( $k=2, k=7$ ) or genomic background sequences ( $t_{G C=0.02 \text { ) for A549 or MCF-7 cells. A549 and MCF-7 cell lines are represented in our data with two training datasets }}$ each, which are labeled as $A$ and $B$, respectively. Model performance was evaluated based on recall for hold-out sets (chromosome 8). The table summarizes mean and standard deviation across ten trained models. There are seven different hold-out sets derived from different cell lines and we assess model generalization across cell-types. Datasets are named according to S1 Table. Respective results for the gkm-SVM models are available Table 1, results for CNN models of 2conv2norm architecture are available in S5 Table.

|  |  | Shuffled ( $k=2$ ) |  |  |  | Model |  |  |  | Genomic background ( $t_{G C}=0.02$ ) |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Shuffled ( $k=7$ ) |  |  |  |  |
|  |  | A549 (A) | A549 (B) | MCF-7 (A) | MCF-7 (B) | A549 (A) | A549 (B) | MCF-7 (A) | MCF-7 (B) | A549 (A) | A549 (B) | MCF-7 (A) | MCF-7 (B) |
|  | A549 (A) |  |  |  |  | $\begin{array}{r} 0.887 \\ \pm 0.034 \end{array}$ | $\begin{array}{r} 0.868 \\ \pm 0.032 \end{array}$ | $\begin{array}{r} 0.869 \\ \pm 0.038 \end{array}$ | $\begin{array}{r} 0.860 \\ \pm 0.030 \end{array}$ | $\begin{array}{r} 0.667 \\ \pm 0.017 \end{array}$ | $\begin{array}{r} 0.614 \\ \pm 0.022 \end{array}$ | $\begin{array}{r} 0.639 \\ \pm 0.027 \end{array}$ | $\begin{array}{r} 0.621 \\ \pm 0.025 \end{array}$ | $\begin{array}{r} 0.840 \\ \pm 0.034 \end{array}$ | $\begin{array}{r} 0.744 \\ \pm 0.035 \end{array}$ | $\begin{array}{r} 0.795 \\ \pm 0.033 \end{array}$ | $\begin{array}{r} 0.742 \\ \pm 0.055 \end{array}$ |
|  | A549 (A) | $\begin{array}{r} 0.867 \\ \pm 0.033 \end{array}$ | $\begin{array}{r} 0.880 \\ \pm 0.031 \end{array}$ | $\begin{array}{r} 0.856 \\ \pm 0.037 \end{array}$ | $\begin{array}{r} 0.847 \\ \pm 0.032 \end{array}$ | $\begin{array}{r} 0.581 \\ \pm 0.017 \end{array}$ | $\begin{array}{r} 0.575 \\ \pm 0.016 \end{array}$ | $\begin{array}{r} 0.557 \\ \pm 0.024 \end{array}$ | $\begin{array}{r} 0.545 \\ \pm 0.022 \end{array}$ | $\begin{array}{r} 0.813 \\ \pm 0.031 \end{array}$ | $\begin{array}{r} 0.795 \\ \pm 0.029 \end{array}$ | $\begin{array}{r} 0.784 \\ \pm 0.028 \end{array}$ | $\begin{array}{r} 0.735 \\ \pm 0.051 \end{array}$ |
| $\stackrel{\widetilde{\rightharpoonup}}{ \pm}$ | HeLa-S3 | $\begin{aligned} & 0.858 \\ & \pm 0.04 \end{aligned}$ | $\begin{array}{r} 0.856 \\ \pm 0.038 \end{array}$ | $\begin{array}{r} 0.849 \\ \pm 0.047 \end{array}$ | $\begin{array}{r} 0.835 \\ \pm 0.025 \end{array}$ | $\begin{array}{r} 0.611 \\ \pm 0.017 \end{array}$ | $\begin{array}{r} 0.575 \\ \pm 0.021 \end{array}$ | $\begin{array}{r} 0.598 \\ \pm 0.031 \end{array}$ | $\begin{array}{r} 0.576 \\ \pm 0.024 \end{array}$ | $\begin{array}{r} 0.773 \\ \pm 0.041 \end{array}$ | $\begin{array}{r} 0.694 \\ \pm 0.033 \end{array}$ | $\begin{array}{r} 0.756 \\ \pm 0.040 \end{array}$ | $\begin{array}{r} 0.691 \\ \pm 0.060 \end{array}$ |
| 苞 | HepG2 | $\begin{array}{r} 0.806 \\ \pm 0.046 \end{array}$ | $\begin{aligned} & 0.818 \\ & \pm 0.04 \end{aligned}$ | $\begin{array}{r} 0.805 \\ \pm 0.044 \end{array}$ | $\begin{array}{r} 0.800 \\ \pm 0.041 \end{array}$ | $\begin{array}{r} 0.476 \\ \pm 0.024 \end{array}$ | $\begin{array}{r} 0.493 \\ \pm 0.019 \end{array}$ | $\begin{array}{r} 0.468 \\ \pm 0.030 \end{array}$ | $\begin{array}{r} 0.465 \\ \pm 0.024 \end{array}$ | $\begin{array}{r} 0.685 \\ \pm 0.040 \end{array}$ | $\begin{array}{r} 0.626 \\ \pm 0.040 \end{array}$ | $\begin{array}{r} 0.668 \\ \pm 0.038 \end{array}$ | $\begin{array}{r} 0.617 \\ \pm 0.063 \end{array}$ |
| $\begin{aligned} & \mathbb{O} \\ & \mathbb{O} \\ & \mathbb{O} \end{aligned}$ | K562 | $\begin{aligned} & 0.848 \\ & \pm 0.04 \end{aligned}$ | $\begin{array}{r} 0.833 \\ \pm 0.038 \end{array}$ | $\begin{array}{r} 0.825 \\ \pm 0.044 \end{array}$ | $\begin{array}{r} 0.819 \\ \pm 0.041 \end{array}$ | $\begin{array}{r} 0.648 \\ \pm 0.021 \end{array}$ | $\begin{array}{r} 0.609 \\ \pm 0.025 \end{array}$ | $\begin{array}{r} 0.609 \\ \pm 0.026 \end{array}$ | $\begin{array}{r} 0.595 \\ \pm 0.026 \end{array}$ | $\begin{array}{r} 0.729 \\ \pm 0.041 \end{array}$ | $\begin{array}{r} 0.635 \\ \pm 0.044 \end{array}$ | $\begin{array}{r} 0.674 \\ \pm 0.037 \end{array}$ | $\begin{array}{r} 0.627 \\ \pm 0.065 \end{array}$ |
|  | MCF-7 (A) | $\begin{aligned} & 0.857 \\ & \pm 0.04 \end{aligned}$ | $\begin{array}{r} 0.845 \\ \pm 0.038 \end{array}$ | $\begin{array}{r} 0.888 \\ \pm 0.038 \end{array}$ | $\begin{array}{r} 0.875 \\ \pm 0.026 \end{array}$ | $\begin{array}{r} 0.621 \\ \pm 0.017 \end{array}$ | $\begin{array}{r} 0.577 \\ \pm 0.022 \end{array}$ | $\begin{array}{r} 0.651 \\ \pm 0.025 \end{array}$ | $\begin{array}{r} 0.631 \\ \pm 0.024 \end{array}$ | $\begin{array}{r} 0.790 \\ \pm 0.038 \end{array}$ | $\begin{array}{r} 0.700 \\ \pm 0.036 \end{array}$ | $\begin{array}{r} 0.858 \\ \pm 0.029 \end{array}$ | $\begin{array}{r} 0.801 \\ \pm 0.049 \end{array}$ |
|  | MCF-7 (B) | $\begin{array}{r} 0.864 \\ \pm 0.037 \end{array}$ | $\begin{array}{r} 0.853 \\ \pm 0.035 \end{array}$ | $\begin{array}{r} 0.894 \\ \pm 0.036 \end{array}$ | $\begin{array}{r} 0.886 \\ \pm 0.026 \end{array}$ | $\begin{array}{r} 0.628 \\ \pm 0.018 \end{array}$ | $\begin{array}{r} 0.590 \\ \pm 0.022 \end{array}$ | $\begin{array}{r} 0.666 \\ \pm 0.027 \end{array}$ | $\begin{array}{r} 0.651 \\ \pm 0.025 \end{array}$ | $\begin{array}{r} 0.816 \\ \pm 0.037 \end{array}$ | $\begin{array}{r} 0.733 \\ \pm 0.035 \end{array}$ | $\begin{array}{r} 0.876 \\ \pm 0.028 \end{array}$ | $\begin{array}{r} 0.839 \\ \pm 0.043 \end{array}$ |

