REPTILE

- Regulatory Element Prediction based on TIssue-specific Local Epigenetic marks

https://github.com/yupenghe/REPTILE

REPTILE is a tool to identify the precise location of enhancers by integrating histone modification data and base-resolution DNA methylation profiles.

Please contact Yupeng He for for feedbacks, questions or bugs.

PRETRAINED MODEL (MOUSE)

Pretrained enhancer models are provided and the files are in `model/`. Each file corresponds the enhancer model of one combination of epigenetic marks (order matters):

- `mm_model_sevenMarks.reptile` - Meth, H3K4me1, H3K4me2, H3K4me3, H3K27ac, H3K27me3, H3K9ac
- `mm_model_coreMarks.reptile` - Meth, H3K4me1, H3K4me3, H3K27ac
- `mm_model_sixHisMod.reptile` - H3K4me1, H3K4me2, H3K4me3, H3K27ac, H3K27me3, H3K9ac
- `mm_model_coreHisMod.reptile` - H3K4me1, H3K4me3, H3K27ac

The performance of these models on experimentally validated regions is available in `model_accuracy.tsv`.

These models were trained on EP300 data in mouse embryonic stem cells (mESCs). The training dataset contains 40,000 genomic regions (5,000 EP300 binding sites as positives; 5,000 promoter regions and 30,000 randomly chosen genomic 2kb regions as negatives). The performance of each model was tested on 545 elements from VISTA enhancer browser. These elements were experimentally validated using transgenic reporter assay. We compared the predictions with experimental results to evaluate the accuracy in the below 6 tissues from mouse embryo at E11.5 development stage:

- E11_5_HT: E11.5 heart
- E11_5_LM: E11.5 limb
- E11_5_FB: E11.5 forebrain
- E11_5_MB: E11.5 midbrain
- E11_5_HB: E11.5 hindbrain
- E11_5_NT: E11.5 neural tube

The performance was measured using several metrics. AUROC is short for The area under the receiver operating characteristic curve, which AUPR is short for area under the precision-recall curve. They are two metrics of prediction accuracy. "top5", "top10" and "top20" are the percentage of true positives in the top 5, 10 and 20 predictions respectively.

PRETRAINED MODEL (HUMAN)

On the to-do list!