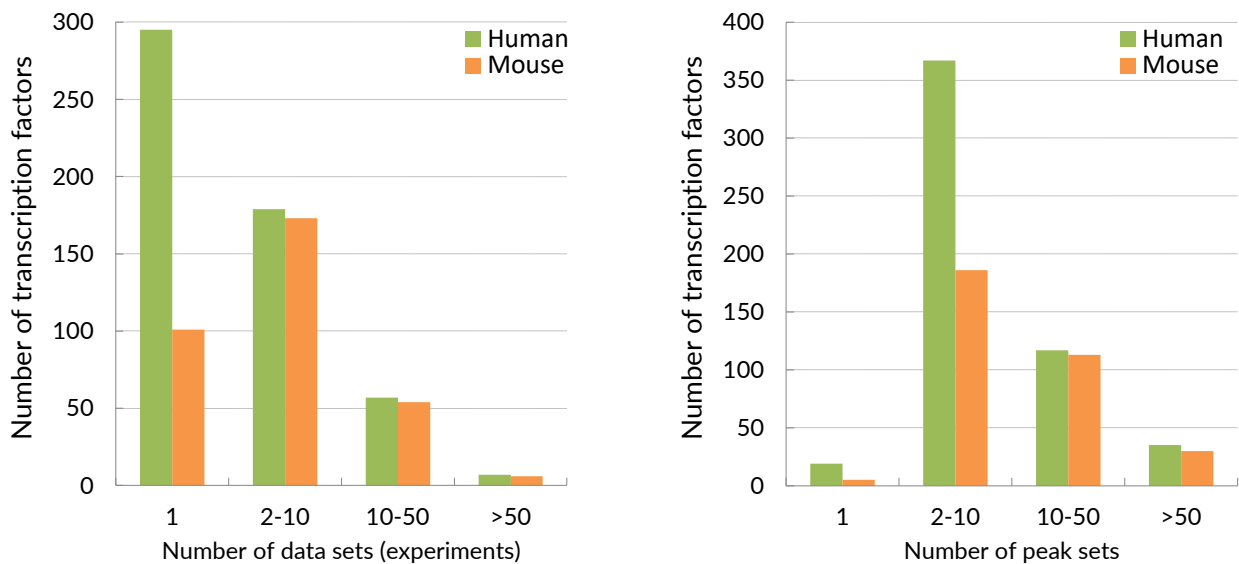
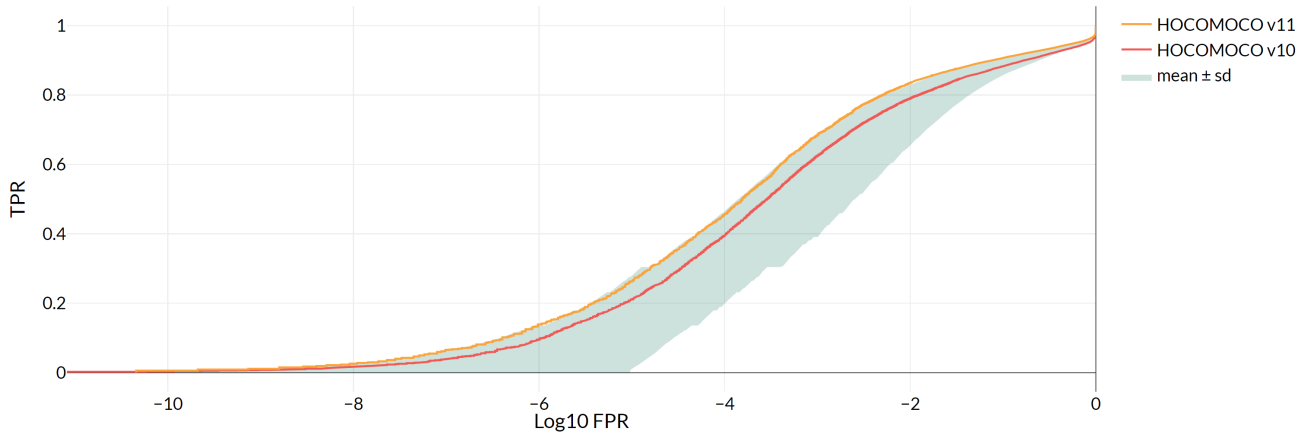


Supplementary figure 1. Overview of the workflow used to compile HOCOMOCO v11.

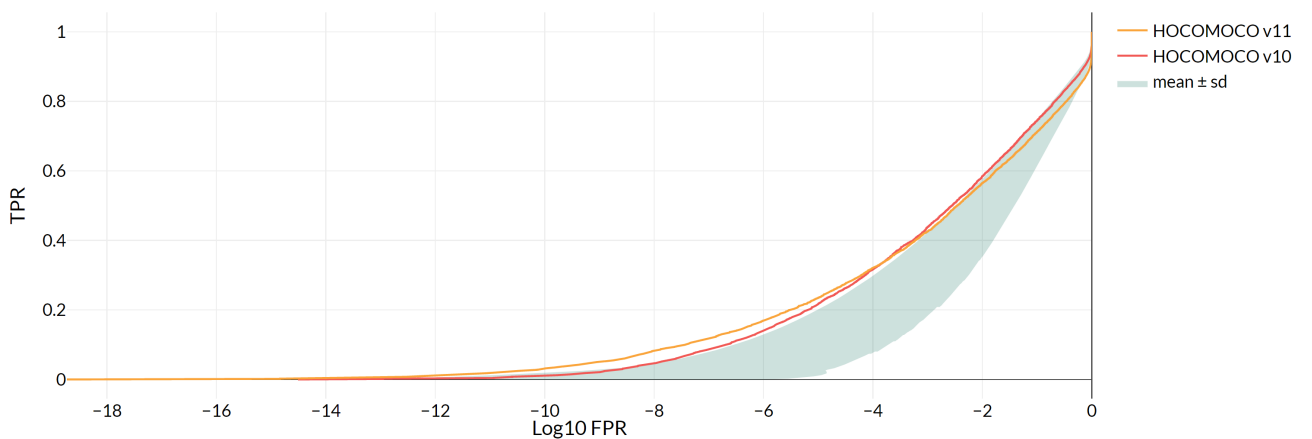


Supplementary figure 2. The number of transcription factors with a given number of available ChIP-Seq experiments (left panel) and peak sets (right panel). Only filtered data used for motif discovery and benchmarking is shown.

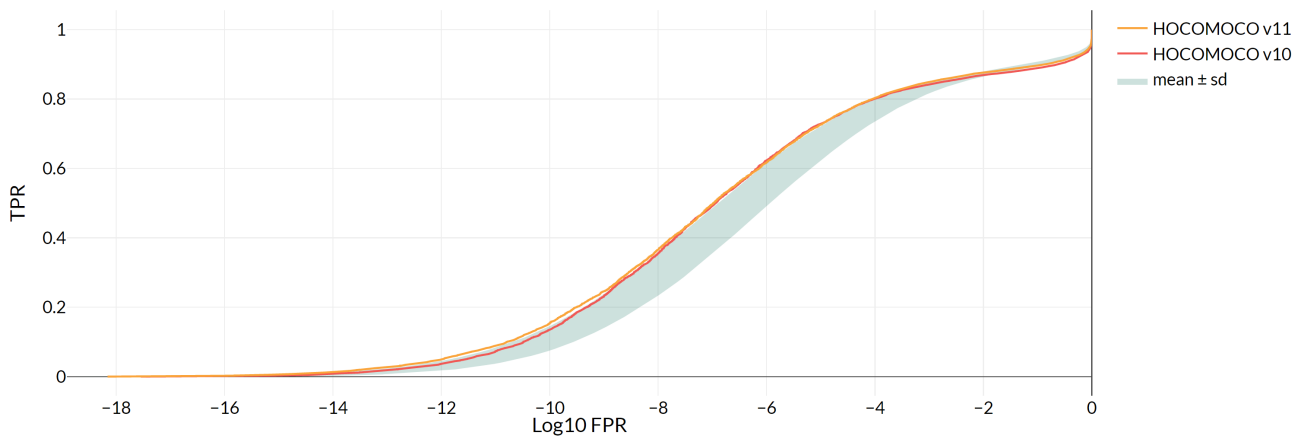
FOXA1 TFBS models performance evaluation



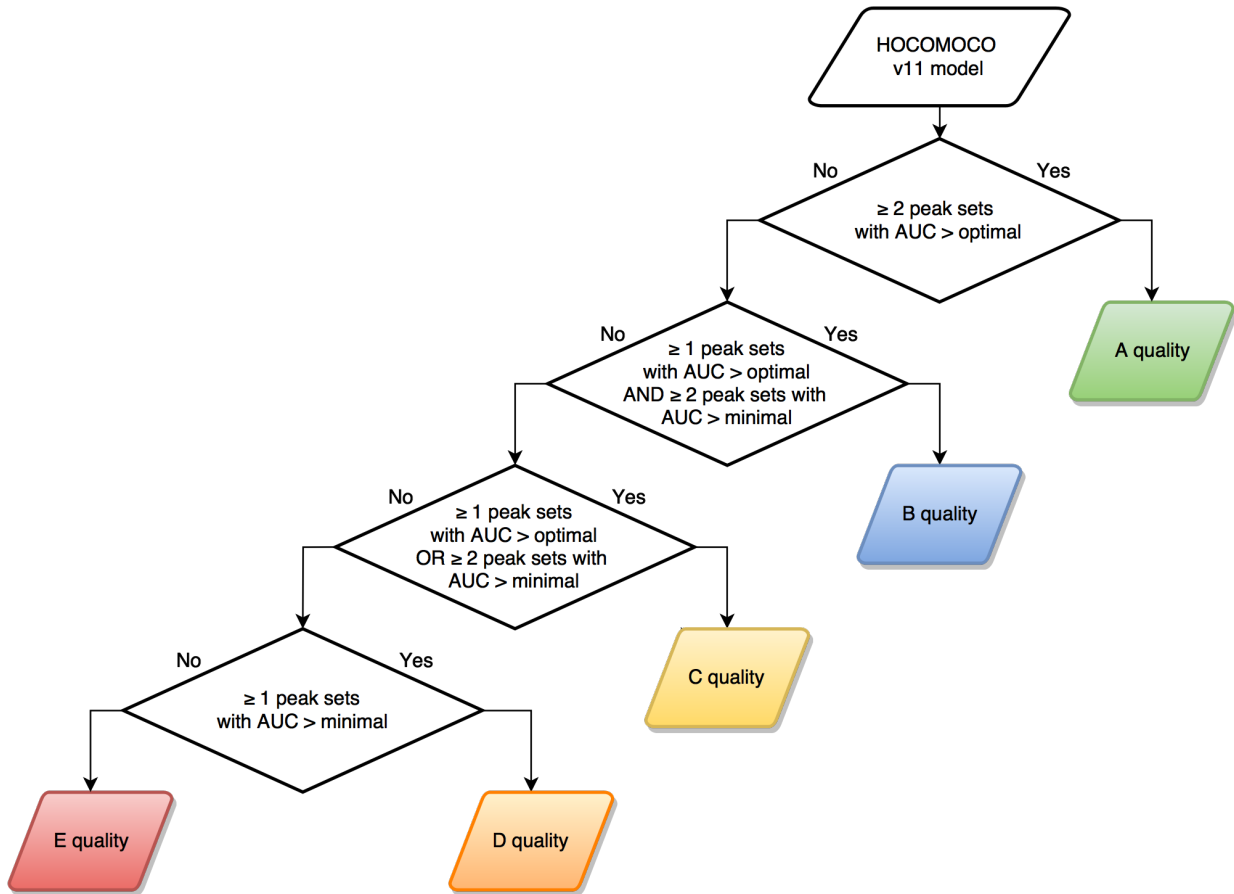
ANDR TFBS models performance evaluation



CTCF TFBS models performance evaluation



Supplementary figure 3. Selected log-ROC curves for TFBS models of human FOXA1 (400 TFBS models), CTCF (1005 models) and ANDR (924 models). Gray areas denotes the mean \pm s.d. estimated over all tested models for the TF under consideration. Orange and red curves correspond to models of HOCOMOCO v11 and v10. HOCOMOCO v11 model has notably better performance than the one from v10 for FOXA1 for a wide range of FPR values; v10 and v11 CTCF models have similar performance; the v11 ANDR model is better than the one from v10 at low FPR. Genome-wide dinucleotide frequencies were used to estimate FPR.



Supplementary figure 4. The workflow used to assign quality ratings to the HOCOMOCO v11 models.

MoLoTool: Transcription Factor Motif Location Toolbox

Mode Hide table Reset Demo Help

COE1_HUMAN.H11MO.0.A EBF1 - Early B-Cell Factor-related factors(6.1.5)

COE1_MOUSE.H11MO.0.A Ebf1 - Early B-Cell Factor-related factors(6.1.5)

-log10(P-value): 4 P-value: 1.00e-4

Submit sequences + Upload file Rewrite: Yes Show sequences Font size: 18px

human SLAMF1 ..
 Show name CAAAAAAGTGATTTAAAGCCTCATGGGAGATGAGCAATCCTCAA

mouse SLAMF1 ..
 Show name TGATAAAGTGATTTAAAGCCTGATCATAAATGAGCAATCCTGGA

Supplementary figure 5. Using MoLoTool to mark EBF1 (COE1) binding site in the human SLAMF1 promoter. The respective binding site is missing in the promoter of the orthologous mouse gene.

The figure contains output as shown on the webpage. This image can be reproduced using MoLoTool 'Demo' function. The gray area in the markup highlights the overlapping predictions from human and mouse TFBS models.