Supplementary Material

funMotifs: Tissue-specific transcription factor motifs

Husen M. Umer[1,2], Karolina Smolinska-Garbulowska[1], Nour-al-dain Marzouka[3], Zeeshan Khaliq[1], Claes Wadelius[4], Jan Komorowski*[1,5,6,7]

[1]Science for Life Laboratory, Department of Cell and Molecular Biology, Uppsala University, [2]Science for Life Laboratory, Department of Oncology-pathology, Karolinska Institutet, Sweden; [3]Department of Clinical Sciences, Faculty of Medicine, Lund University; [4]Science for Life Laboratory, Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden; [5]Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland

[6]Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland

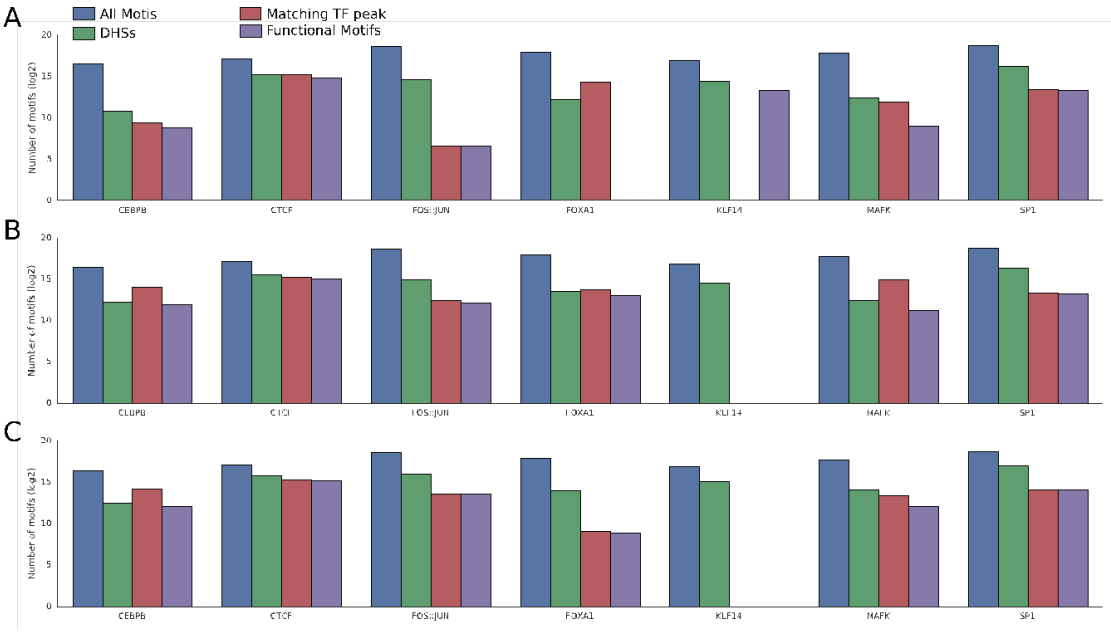
[7]Washington National Primate Research Center, Seattle, WA, USA

*To whom correspondence should be addressed. Tel: +46 (0) 18 471 66 92; Email: jan.komorowski@icm.uu.se

Data File 1 is available on:

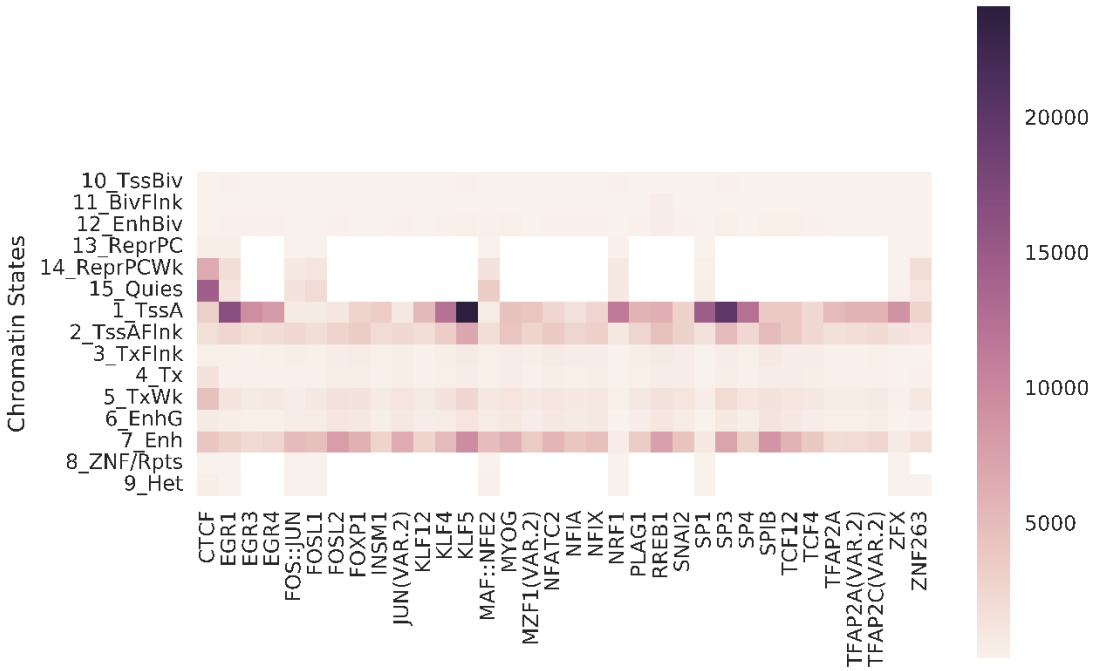
https://figshare.com/s/073c985fd131f78fb9c5

Supplementary Figures

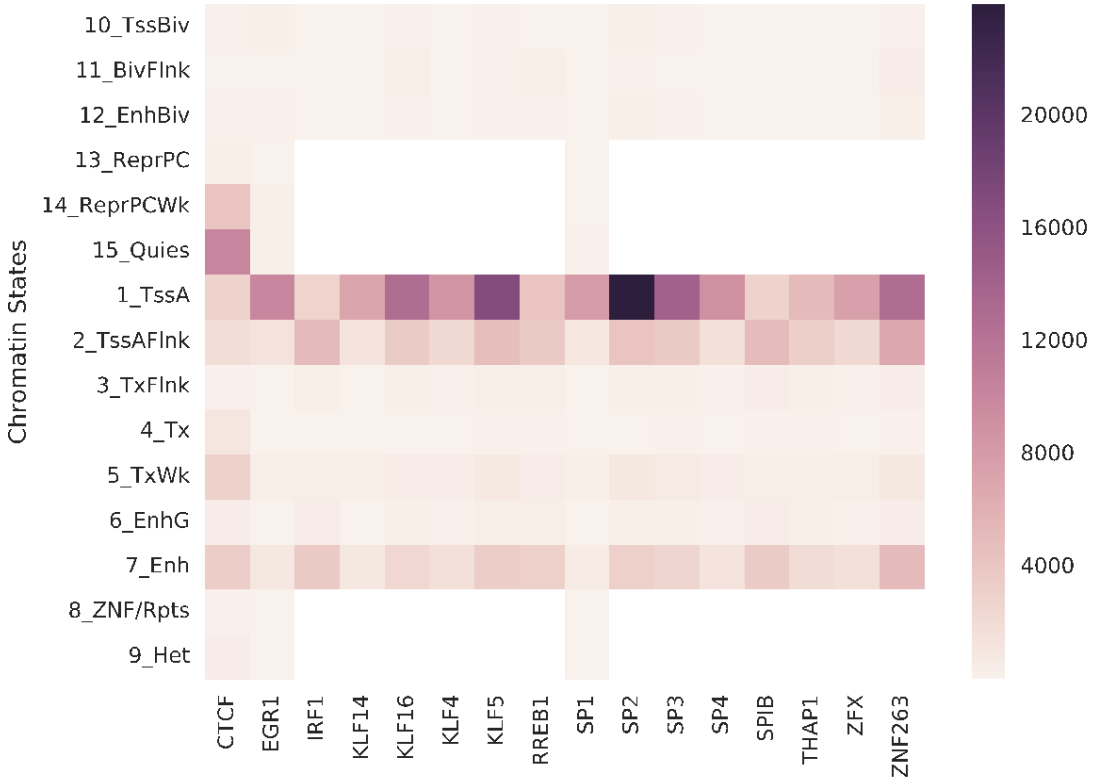


Supplementary Figure 1: Motif-annotation overlaps (log2 scale)

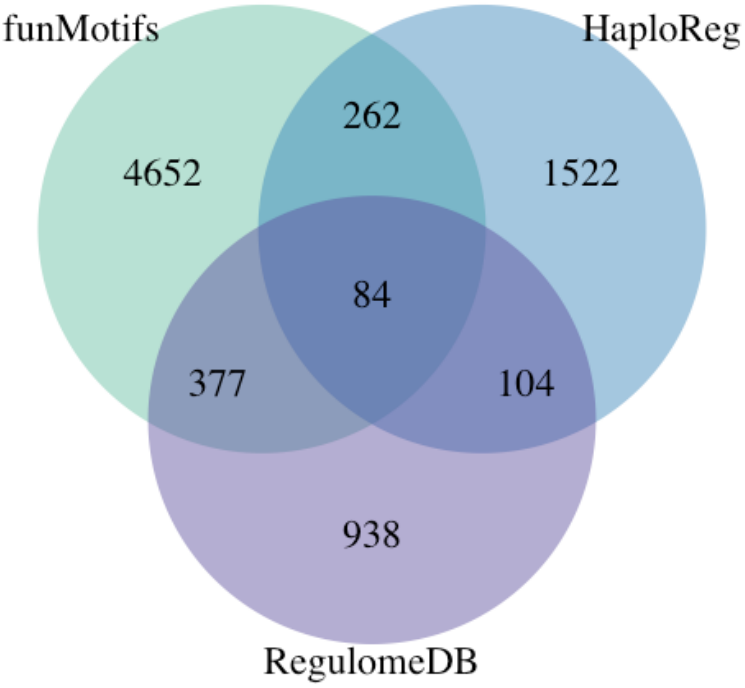
The columns represent number of sequences predicted motifs, motifs overlapping DHSs, motifs overlapping TF peaks and functional motifs respectively in (A) blood, (B) brain and (C) liver tissues.



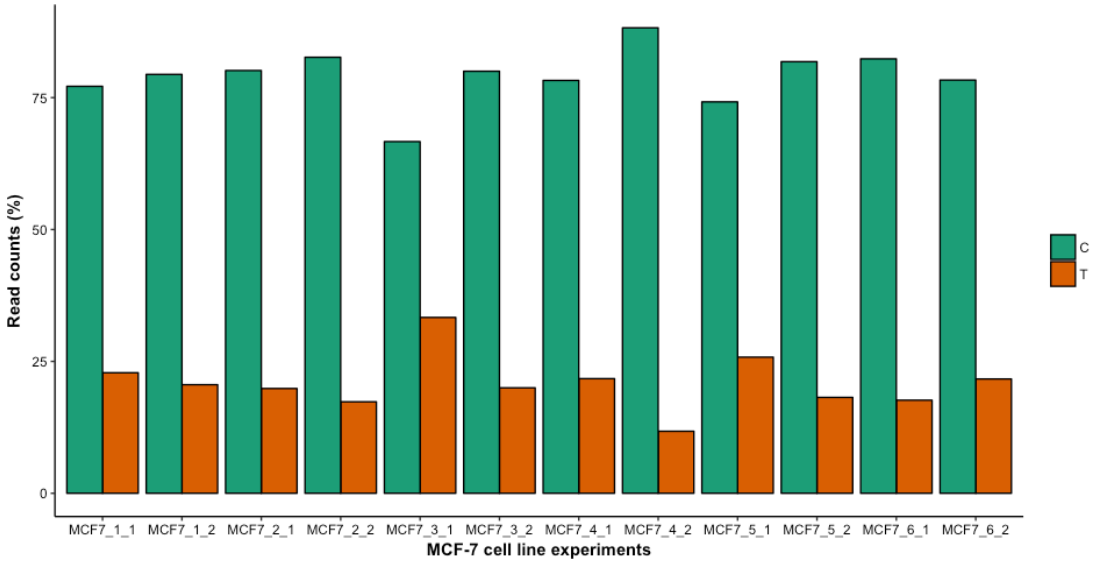
Supplementary Figure 2: Enrichment of functional motifs in chromatin states in myeloid.



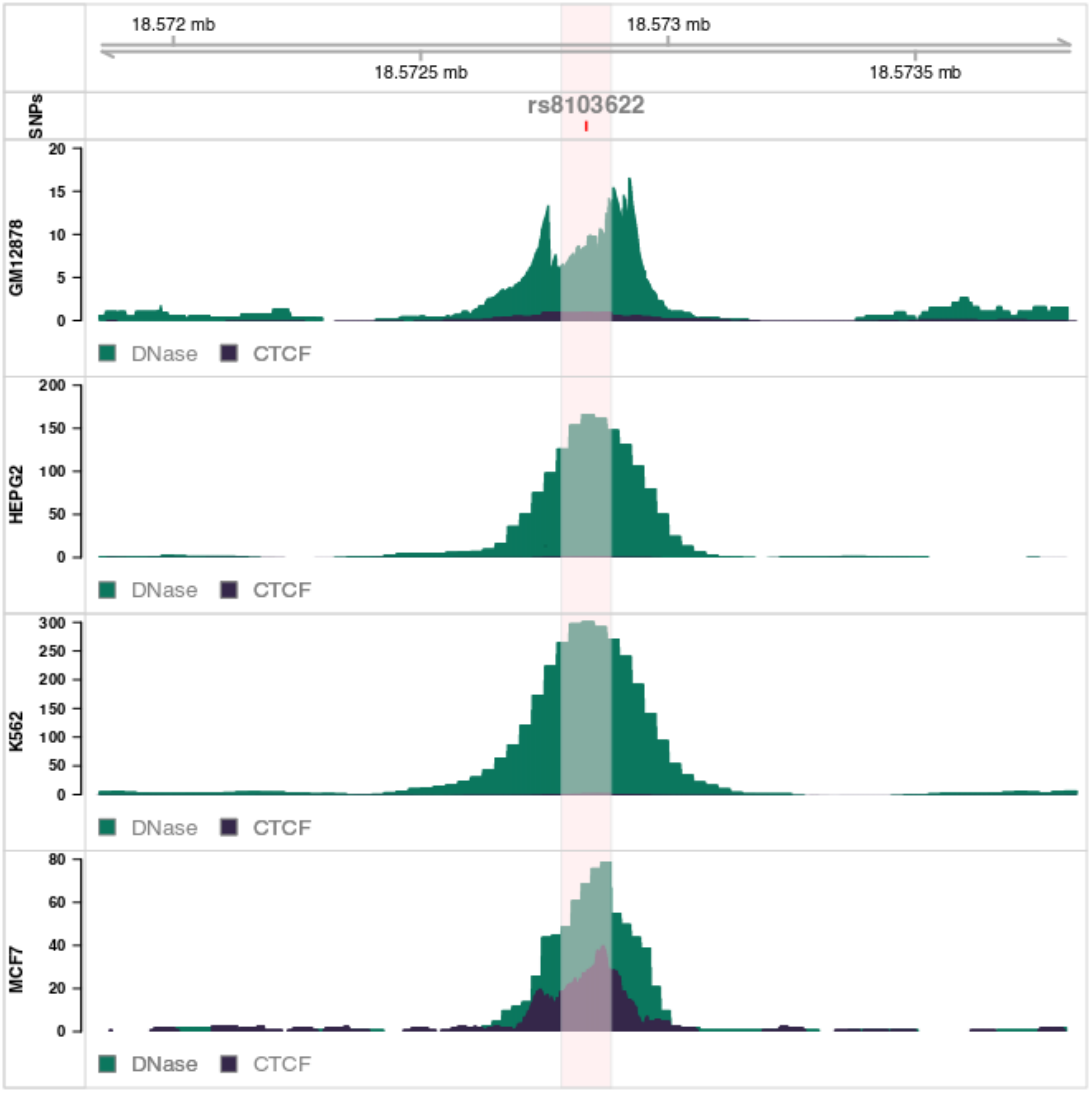
Supplementary Figure 3: Enrichment of functional motifs in chromatin states in blood.



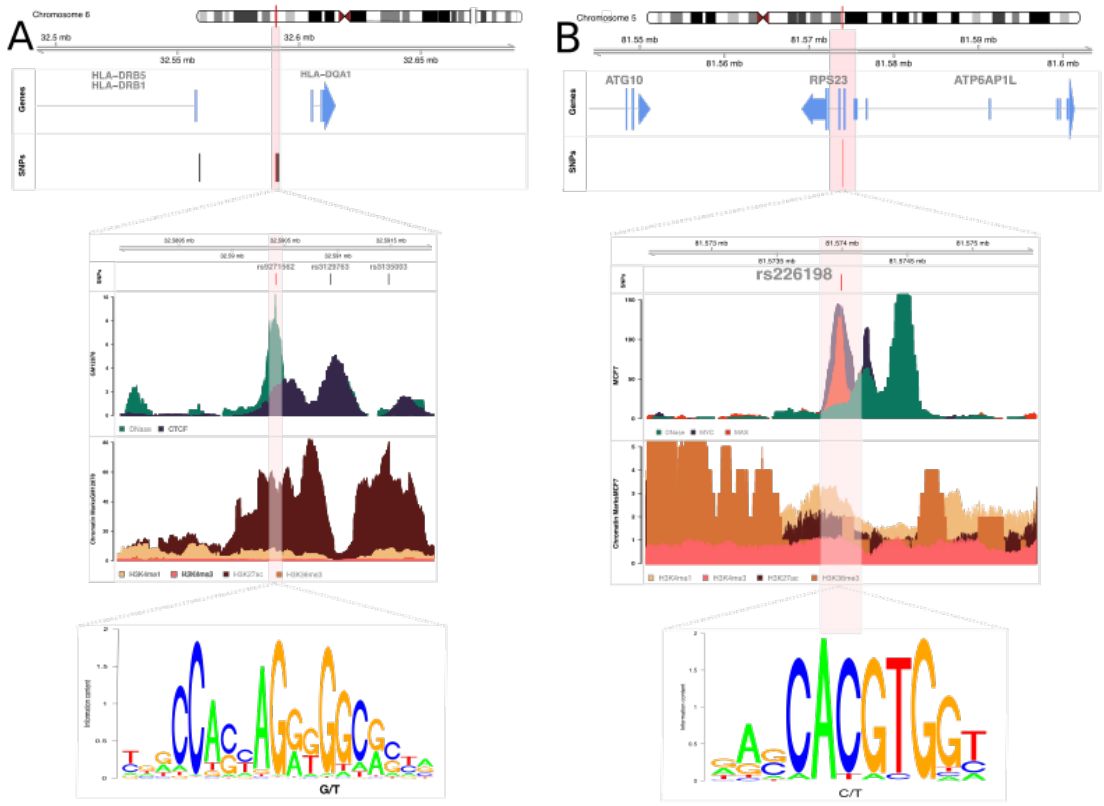
Supplementary Figure 4: The comparison of candidate functional eQTL variants annotated by funMotifs, HaploReg and RegulomeDB.



Supplementary Figure 5: Read counts distribution for the candidate functional GWAS SNP rs8103622 for breast cancer cell line MCF-7 using ChIP-seq data (**Supplementary Table 13**).



Supplementary Figure 6: Presence of DHS and CTCF TF binding signals for the candidate functional GWAS SNP rs8103622 in four tissues: GM12878 - blood, HEPG2 - liver, K562 - myeloid and MCF7 - breast.



Supplementary Figure 7: Overview of the candidate functional GWAS variants: a) rs9271562 in blood and b) rs226198 in breast. Genes box: blue genes are genes associated with the candidate SNP using eQTL variants. SNPs box: red line indicates position of the candidate SNP, black lines are GWAS SNPs in LD with the candidate SNP.