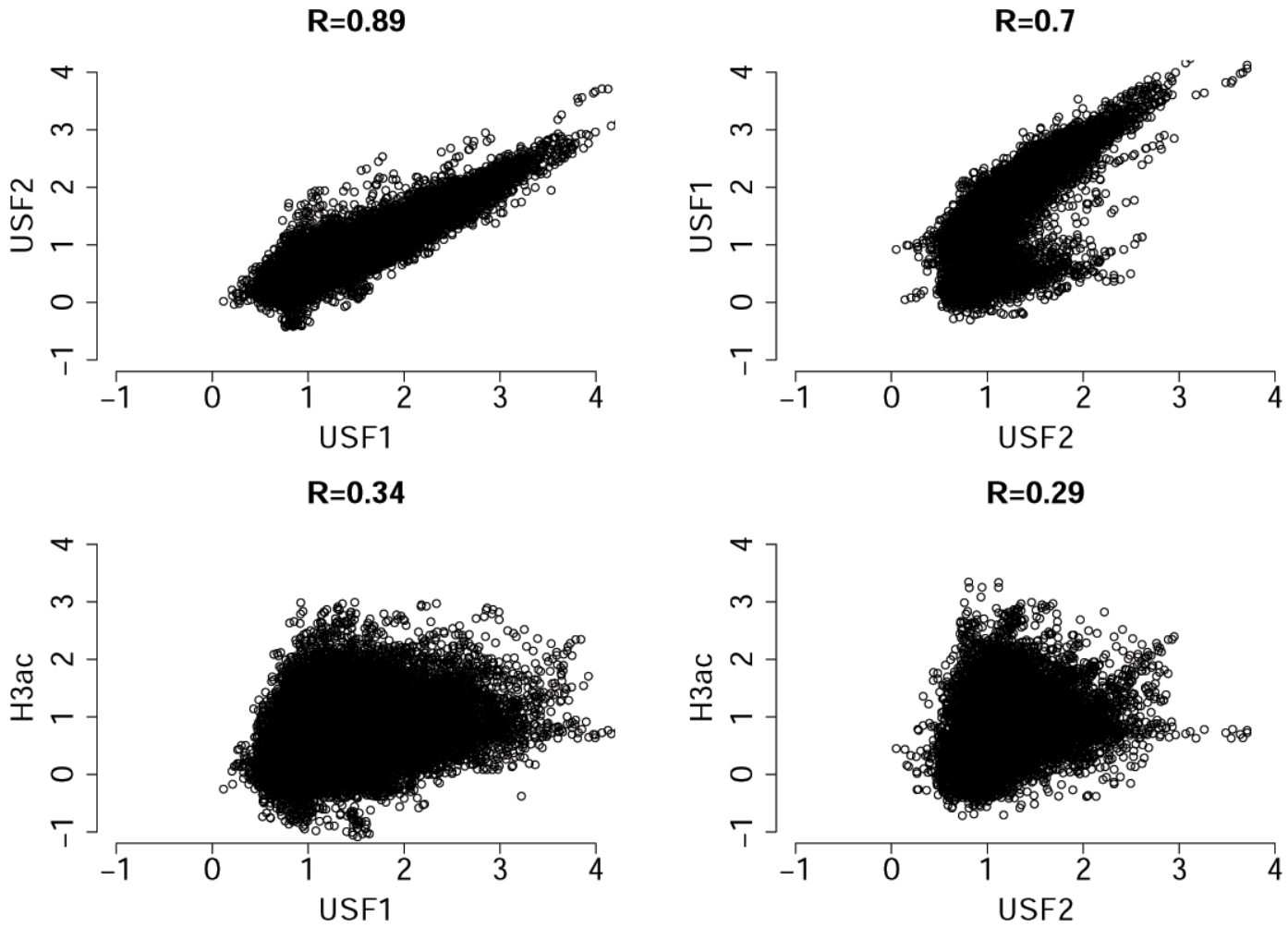
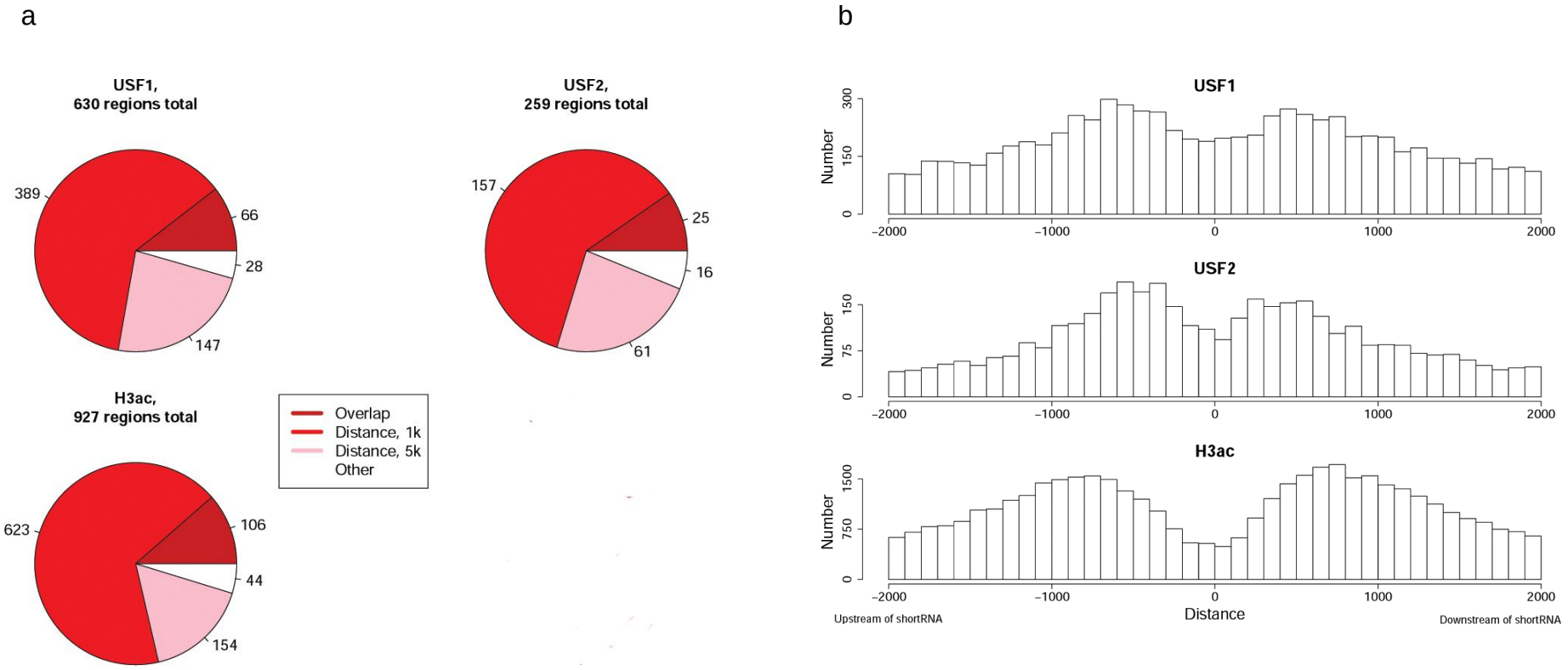
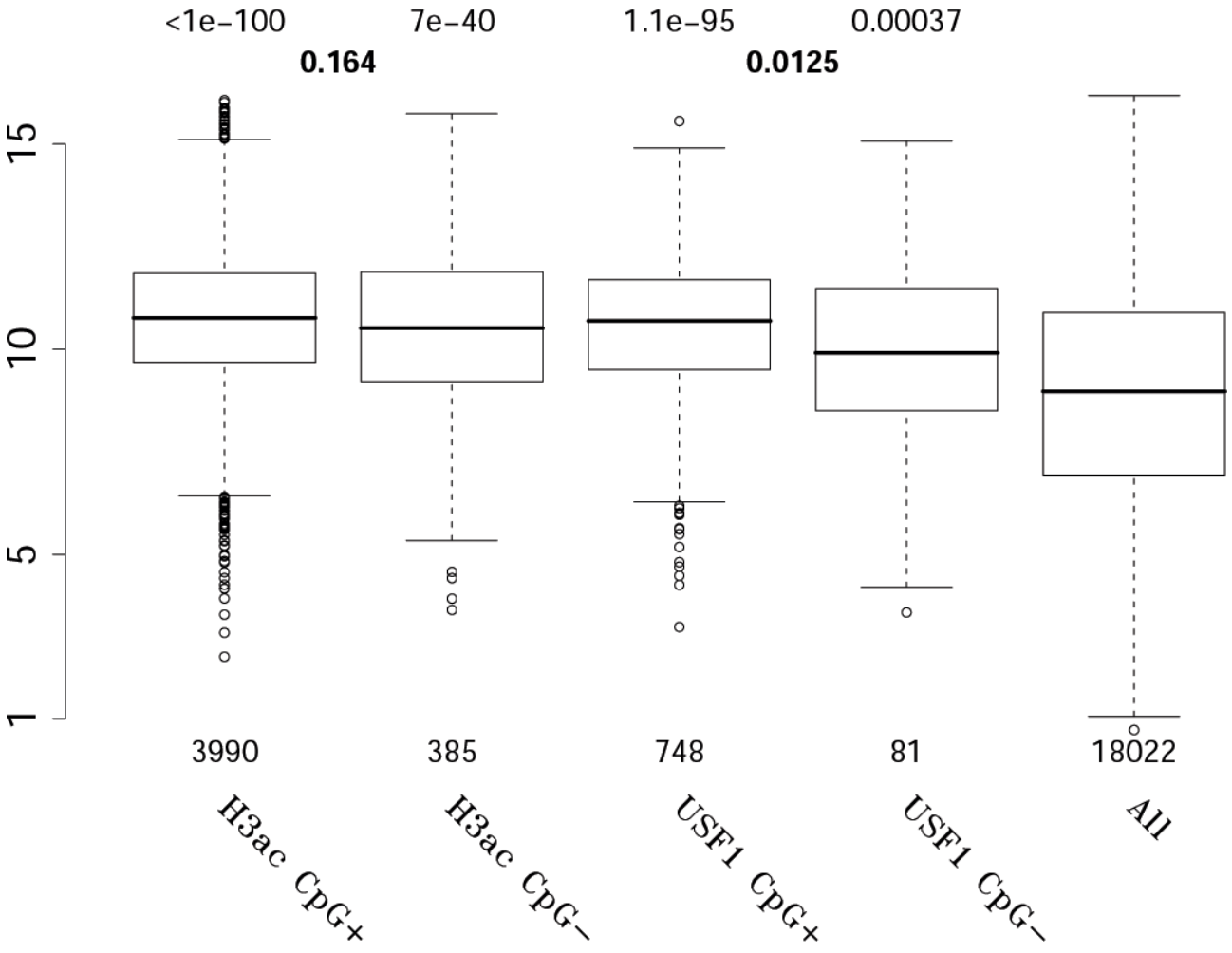


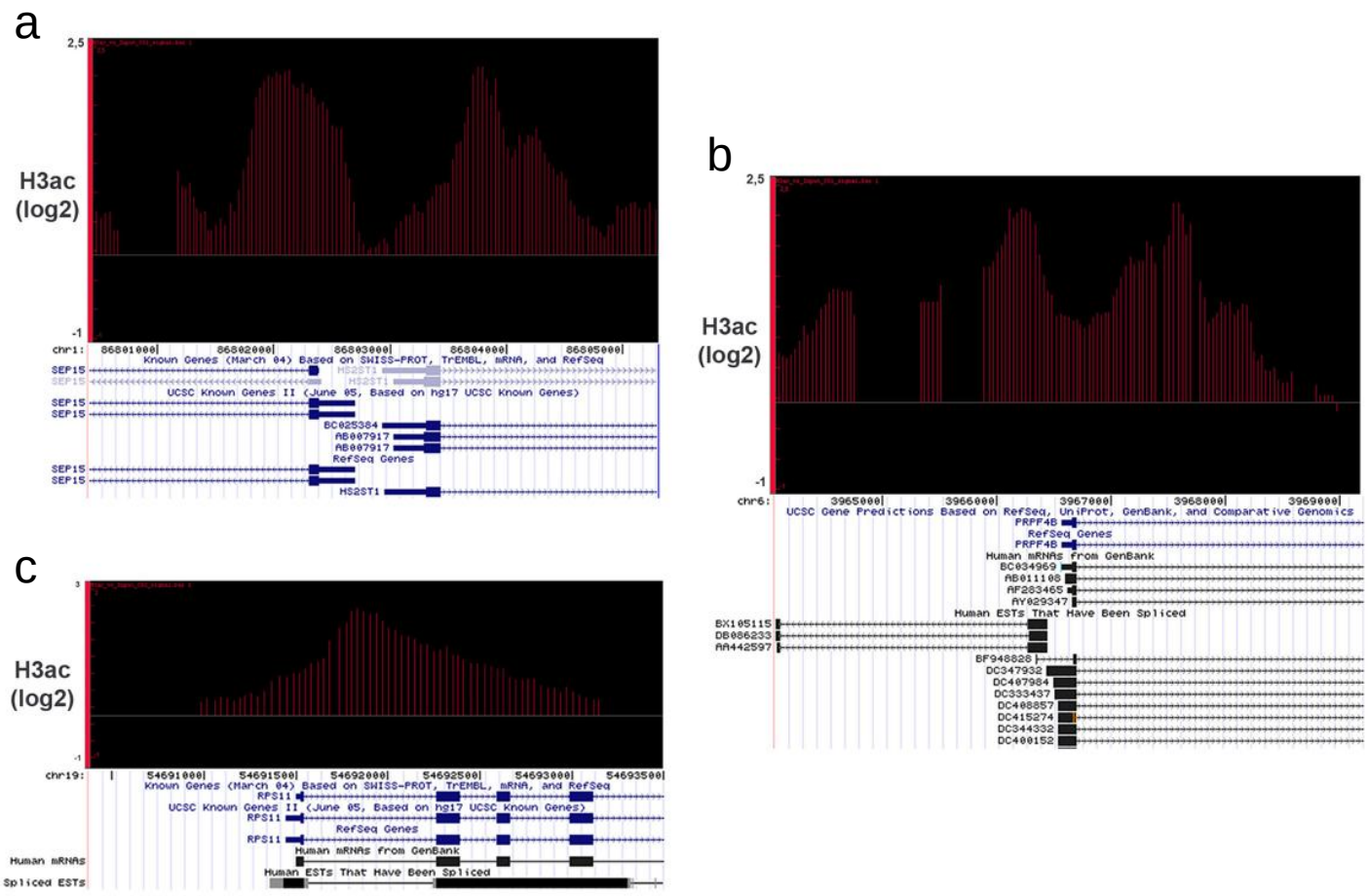
Supplementary Figure 7. Correlation between different ChIP-chip data sets.
All enriched spots under our relaxed criteria were considered. In each case, all enriched spots for the protein displayed in the X-axes were compared to the signals of all those spots for the protein presented in the Y-axes. The numbers on top of each plot are R Person coefficient values.



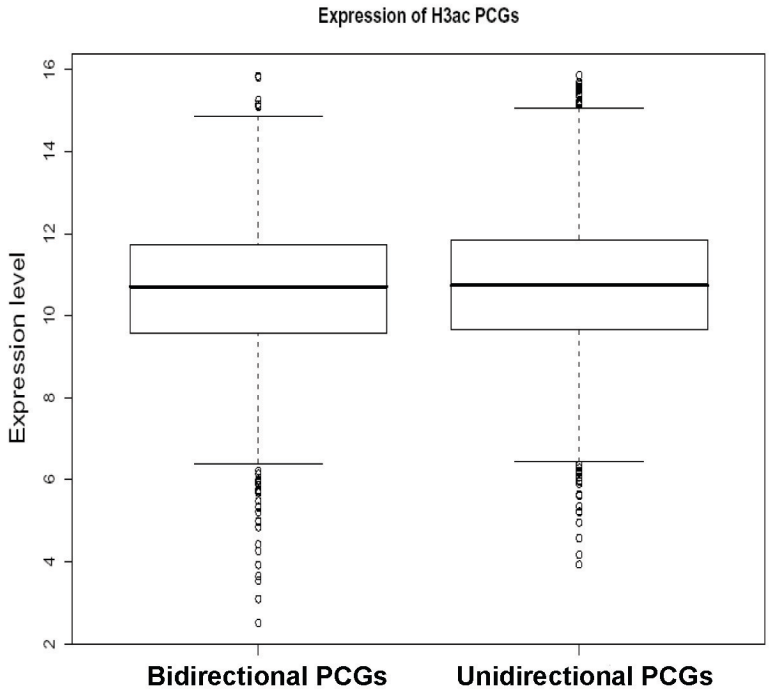
Supplementary Figure 8. USFs and H3ac TSS-distal bound regions are in proximity of transfrags.
(a) Those regions being bound by USF1, USF2 and H3ac and not mapping within 1 Kb of TSSs in any of the transcript annotations employed were selected. Such regions were mapped to the combined genomic positions of expressed sequences (transfrags) generated from transcriptome analysis of long (more than 200 bp) nuclear/cytoplasmic and short (less than 200 bp) RNAs in HepG2 cells. Overlaps were considered to occur when at least there was 1 bp in common between the bound region and one transfrag, otherwise the distance (bp) between the middle positions of the bound region and the transfrag were considered. (b) All USF1, USF2 and H3ac bound regions having at least one short RNA were selected. The distances (bp) between middle position of bound regions and TSS of the nearest short RNA were calculated for each bound region (X-axes, positive and negative numbers being downstream and upstream of short RNAs TSSs respectively). The number of bound regions (Y-axes) with certain distance to short RNAs TSSs is presented every 100 bp, using TSSs of short RNA as reference position (0 in X-axes).



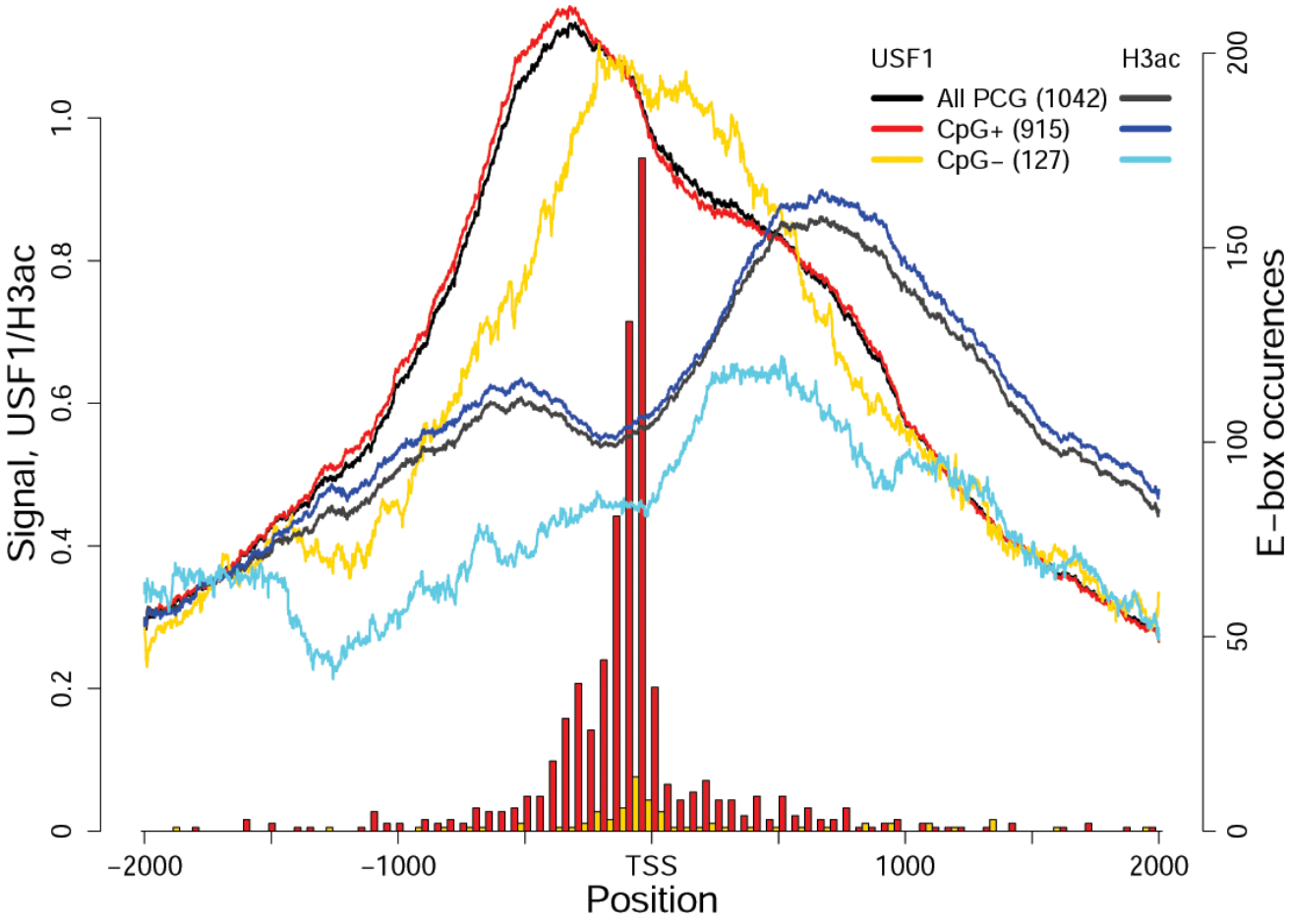
Supplementary Figure 9. Additional features of USF1 and H3ac bound PCG promoters.
The overall expression (in log2 scale and represented as box plots) for 18000 in HepG2 cells , and for those genes (number of genes below box plots) being bound by the indicated proteins in the bottom of the graph within +/- 1 Kb of their TSS was calculated. The numbers on top of each box plot denote the significance values obtained as determined by two-tailed t-test analyses between each group compared to all genes. The bold numbers in between box plots indicate the significance values obtained as determined by two-tailed t-test analyses comparing H3ac CpG+ vs H3ac CpG- and USF1 CpG+ vs USF1 CpG-, respectively



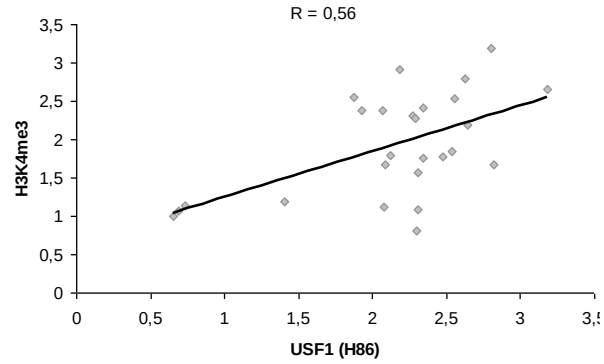
Supplementary Figure 10. H3ac ChIP-chip signal for three different loci that represent a pair of protein coding genes in a bidirectional conformation (a), a protein coding gene and an spliced EST in a bidirectional conformation (b) and a unidirectional protein coding gene (c). In each case the H3ac signal (in red) is presented in log2 scale. Below the H3ac signal a genome browser picture is presented, indicating the genomic location and existing transcripts.

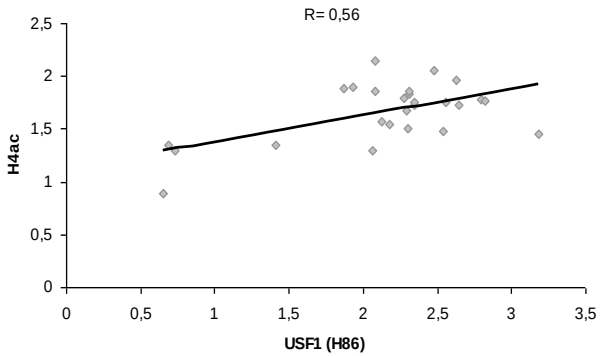


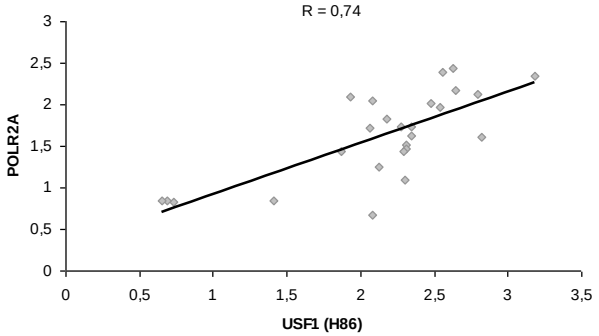
Supplementary Figure 11. Upstream histone acetylation is not due to differences in expression levels. Expression levels (log2 transformed) were calculated for all unidirectional and bidirectional PCGs where expression was available.

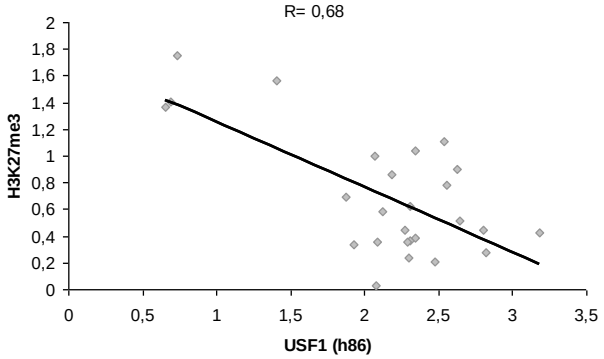


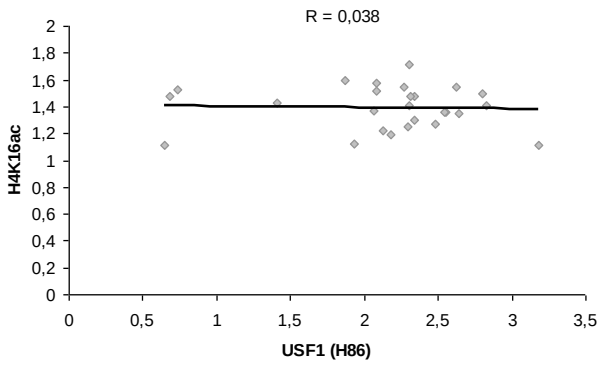
Supplementary Figure 12. Influence of CpG islands in USF and H3ac binding profiles.
All PCGs bound by USF1 within 1Kb of their TSSs were considered as CpG+ (CpG island within 1Kb of TSS) or CpG- (no CpG island within 1Kb of TSS), and USF1 and H3ac binding profiles like in were created for both groups of genes. Occurrence of perfect matches to E-box consensus sequence is also presented for all groups of PCGs.

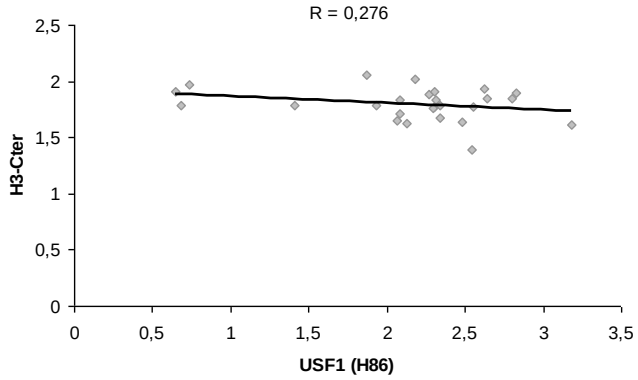


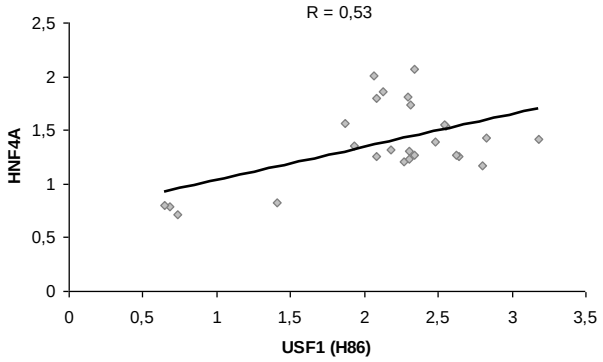




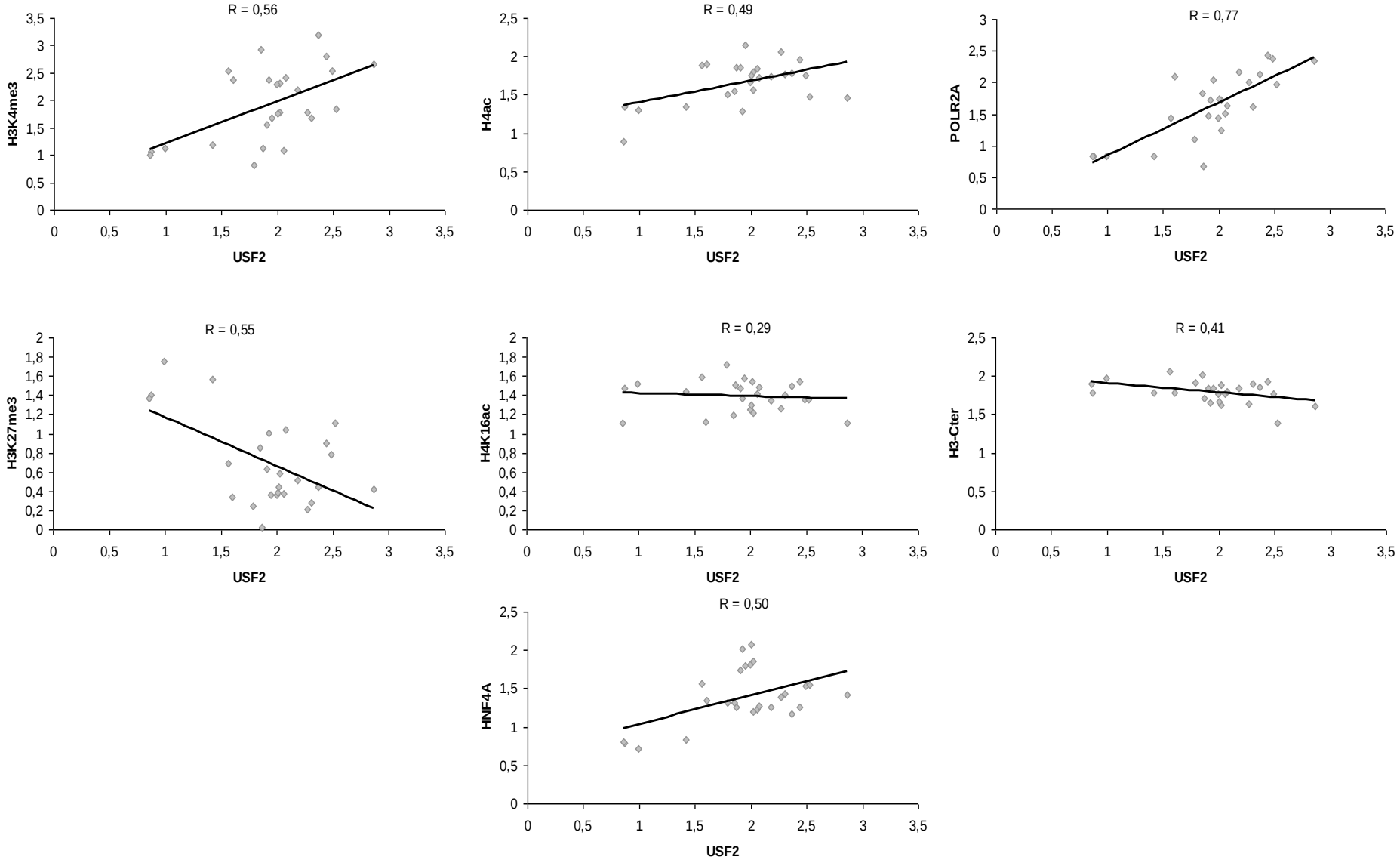




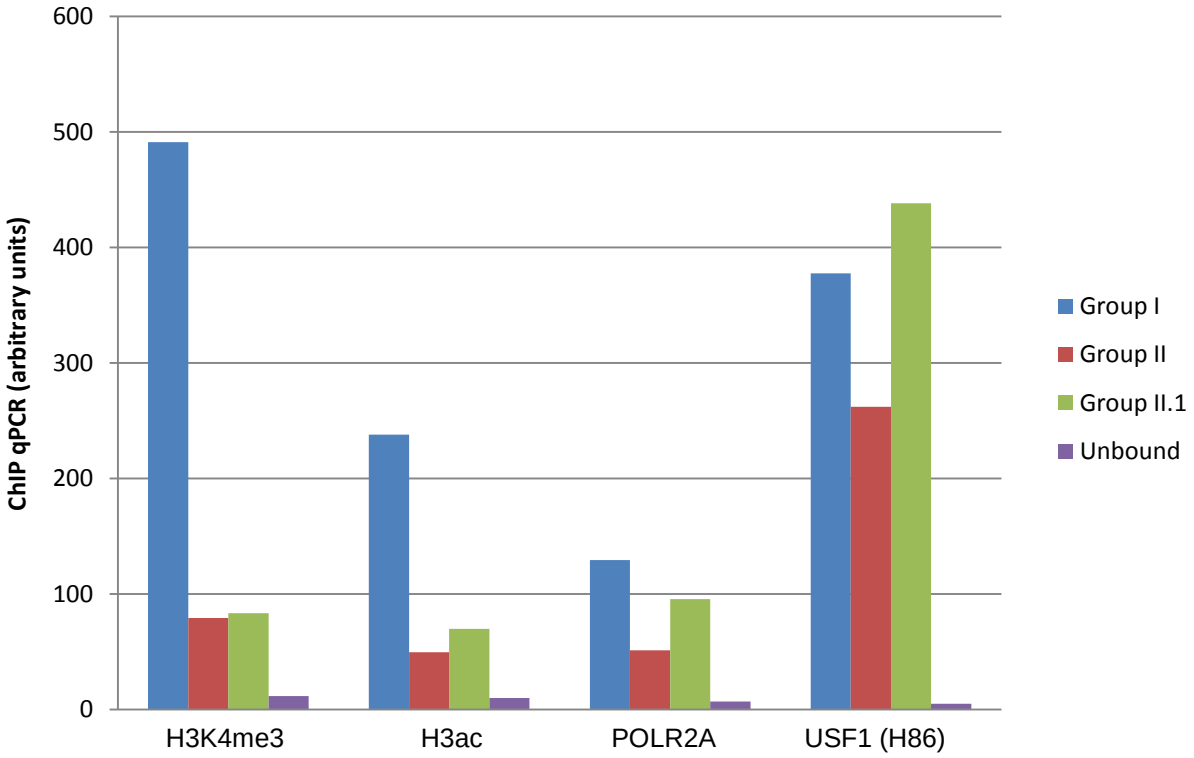




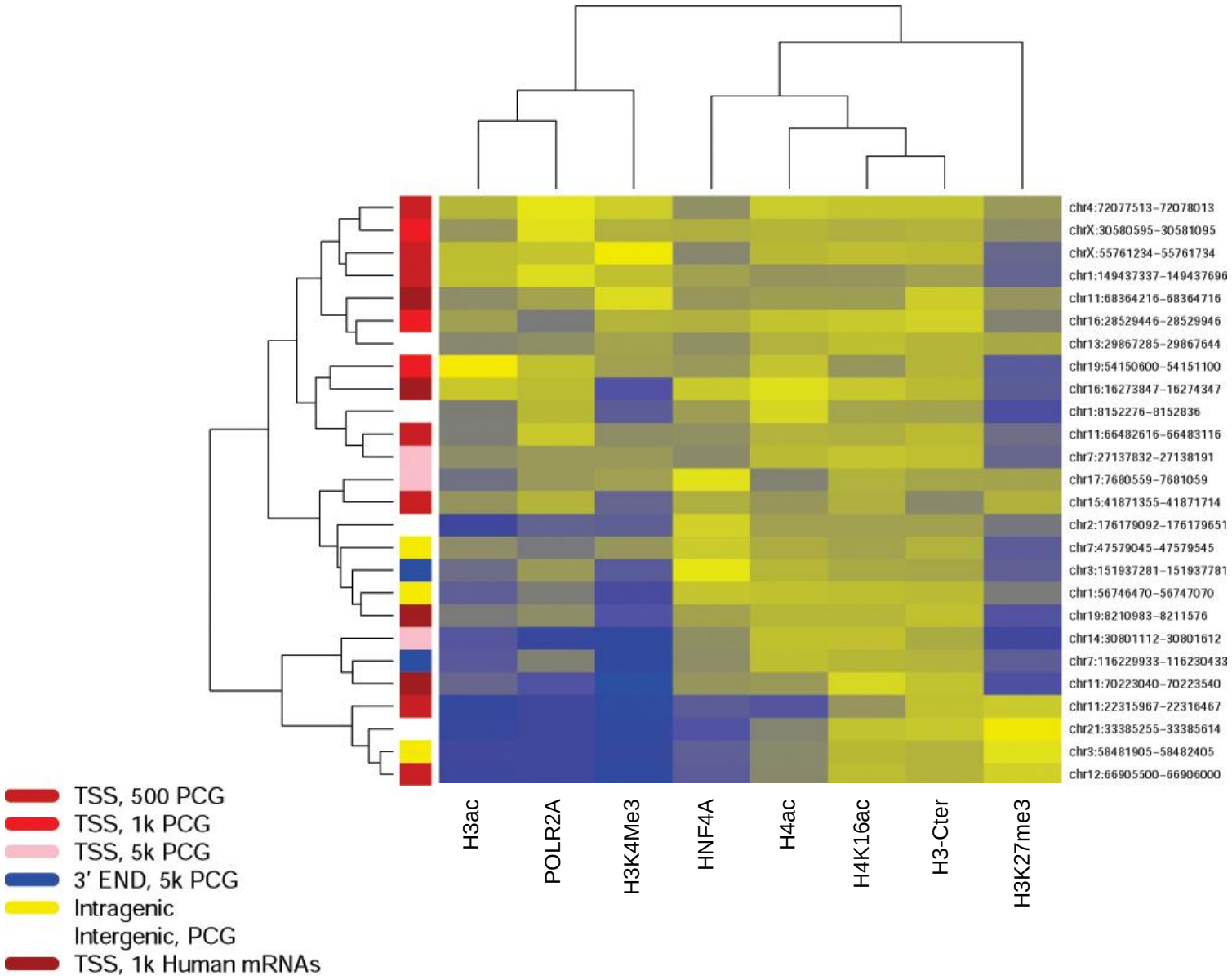
Supplementary Figure 13. Correlation between binding of USF1 and histone modifications using qPCR data.
Correlation between USF1 binding affinity and histone modifications. Person Correlation (R values) between binding profiles of USFs and the indicated proteins were obtained using the corresponding qPCR signals (log10 transformed) in each case.



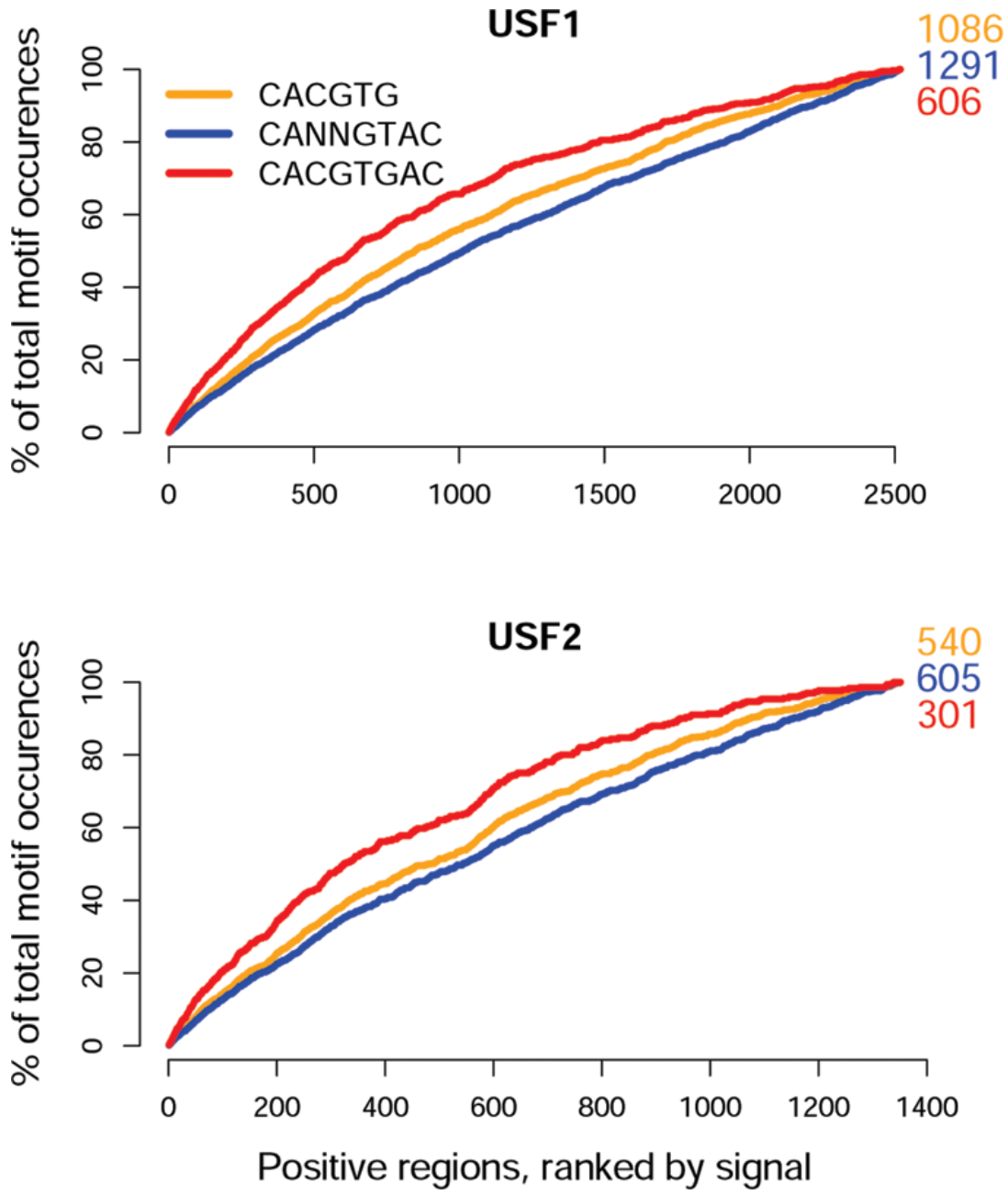
Supplementary Figure 14. Correlation between binding of USF2 and histone modifications using qPCR data.
Correlation between USF2 binding affinity and histone modifications. Person Correlation (R values) between binding profiles of USFs and the indicated proteins were obtained using the corresponding qPCR signals (log10 transformed) in each case.



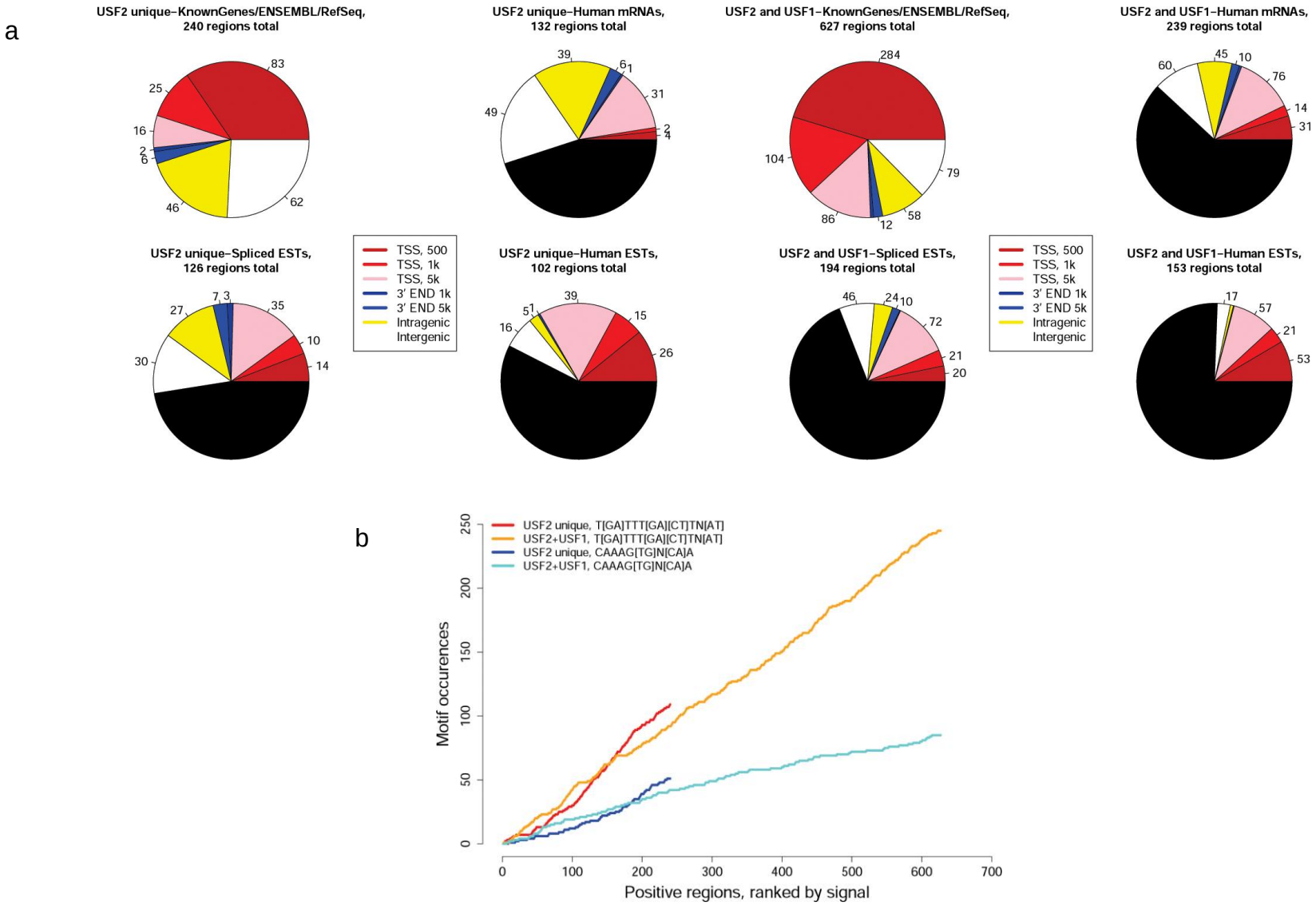
Supplementary Figure 15. H3ac, H3K4me3 and Pol2 ChIP qPCR signal at different USF1 bound loci. Different regions as presented in Figure 6 were clustered based on ChIP qPCR signals for various proteins, and several groups could be distinguished (I, II, II.1, ubound). Here are presented the average ChIP qPCR signals (in arbitrary units) for H3K4me3, H3ac, Pol2 and USF1 (H86) in these four different groups of USF1 bound loci.



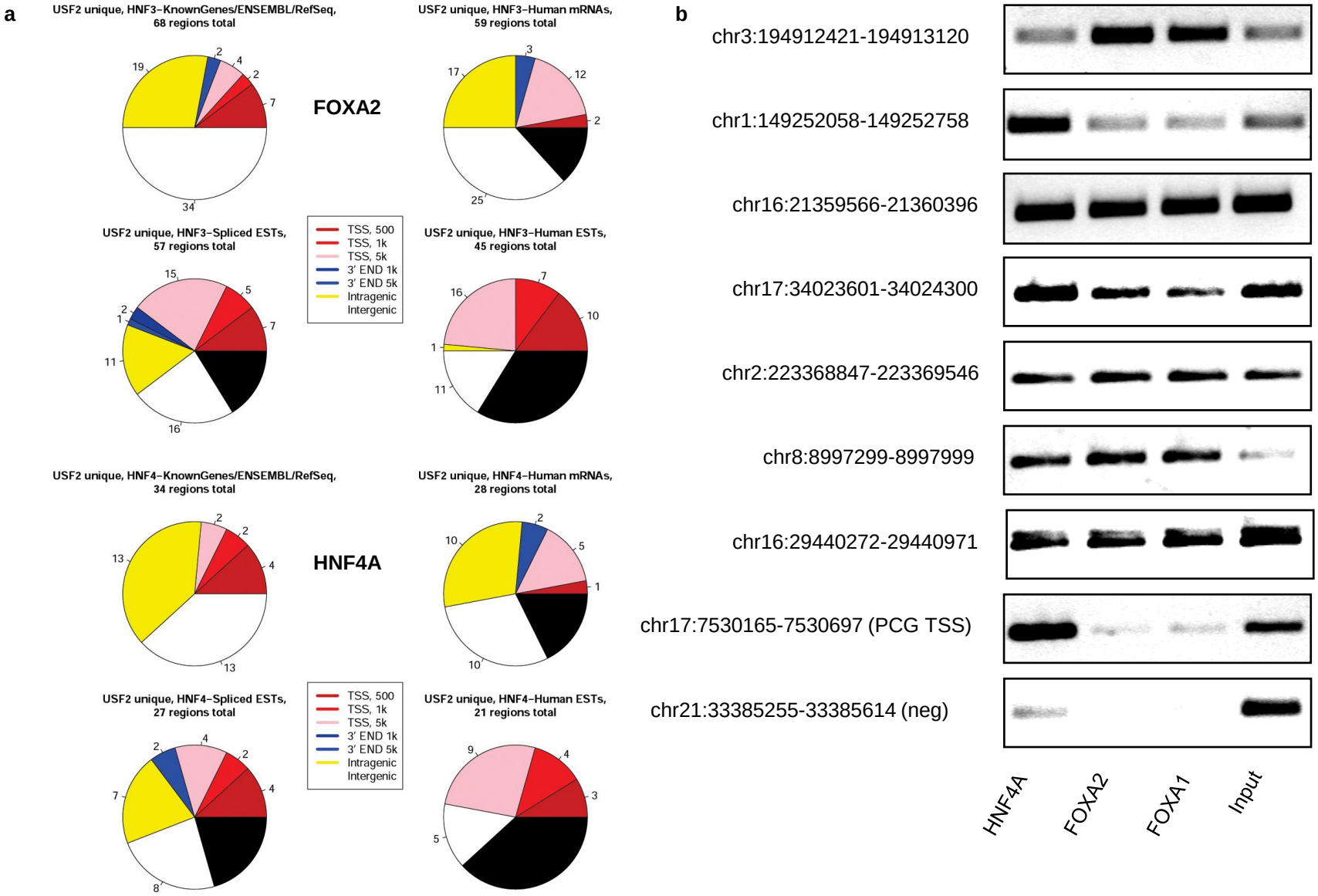
Supplementary Figure 16. USF bound regions display euchromatic characteristics.
26 selected regions were clustered based on qPCR signals for the indicated transcription factors and histone modifications in the lower part of the picture. In all cases the column colours indicates the qPCR signal from high (yellow) to low (blue) values. The analyzed regions are presented to the right, while the genomic location of each region is indicated by the colour to the left.



Supplementary Figure 17. Candidate scanning search of E-box and E-box-like sequences among USF1 and USF2 bound regions.
The total number of E-Box occurrences with a maximum of four per region (right) were considered, using perfect matches to the indicated sequences on top of each graph. USF bound regions were ranked in descending order according their log2 enrichment values.



Supplementary Figure 18. Characteristics of unique USF2 targets.
(a) Genomic distribution of USF2 bound regions divided in USF2 unique and USF2-USF1 is presented. (b) Candidate scanning approaches were used to search for the occurrence of FOXA2 (red, orange) or HNF4A (dark and light blue) consensus sequences among USF2 unique and USF2-USF1 regions, which were ranked in descending order according their log2 enrichment values. The total number of consensus sequence occurrences with a maximum of four per region (right) were considered.



Supplementary Figure 19. HNFs proteins bind unique USF2 targets.
(a) Genomic distribution of USF2 unique regions where at least one perfect match to FOXA2 (HNF3B) or HNF4A consensus sequences was identified. (b) Eight USF2 unique bound regions displaying perfect matches to both HNF4A and FOXA2 consensus sequences were analyzed by semi-quantitative PCR together with a USF2 unbound region. In each case, ChIP DNAs were obtained using antibodies against HNF4A, FOXA2 and FOXA1.

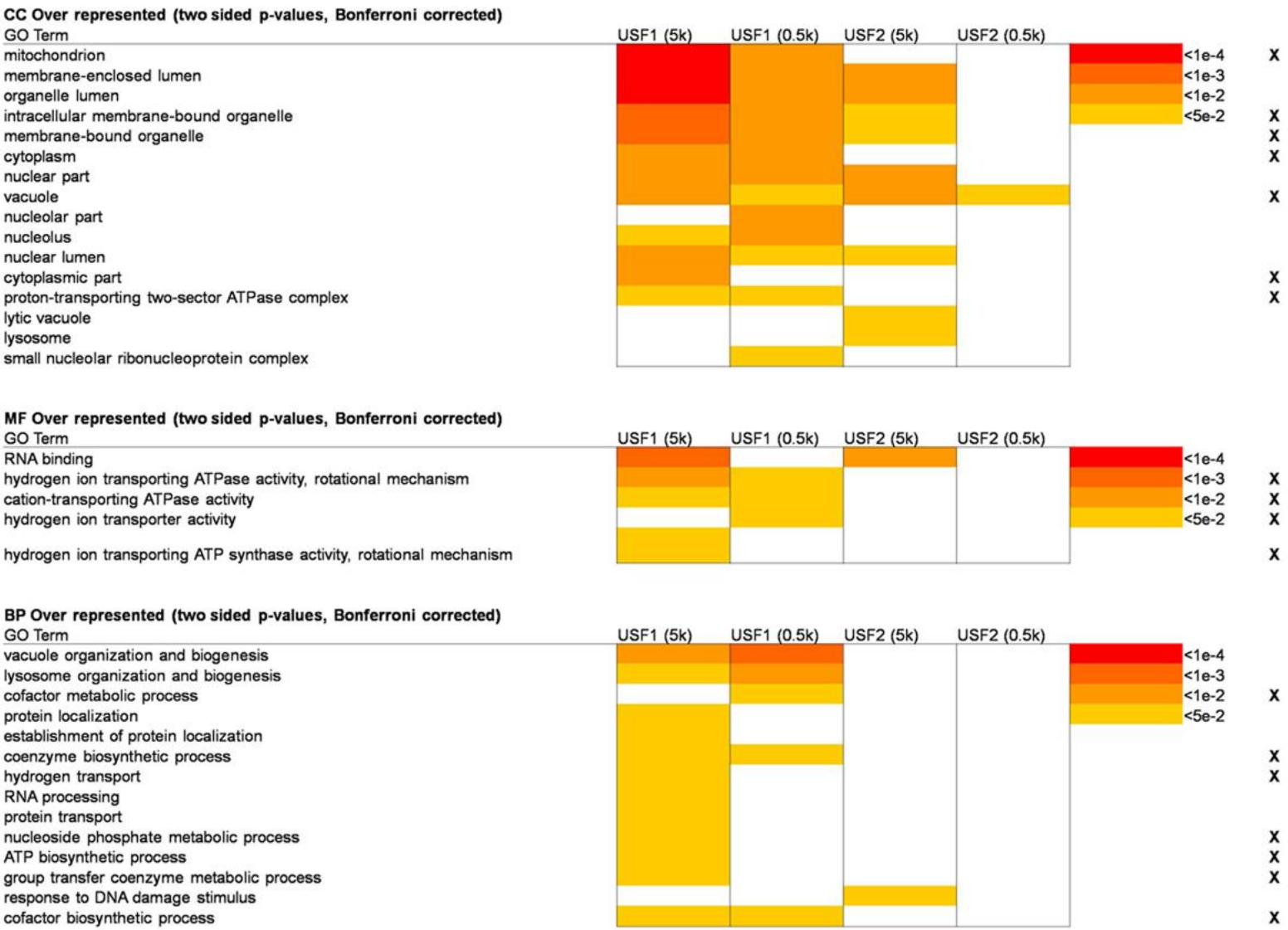
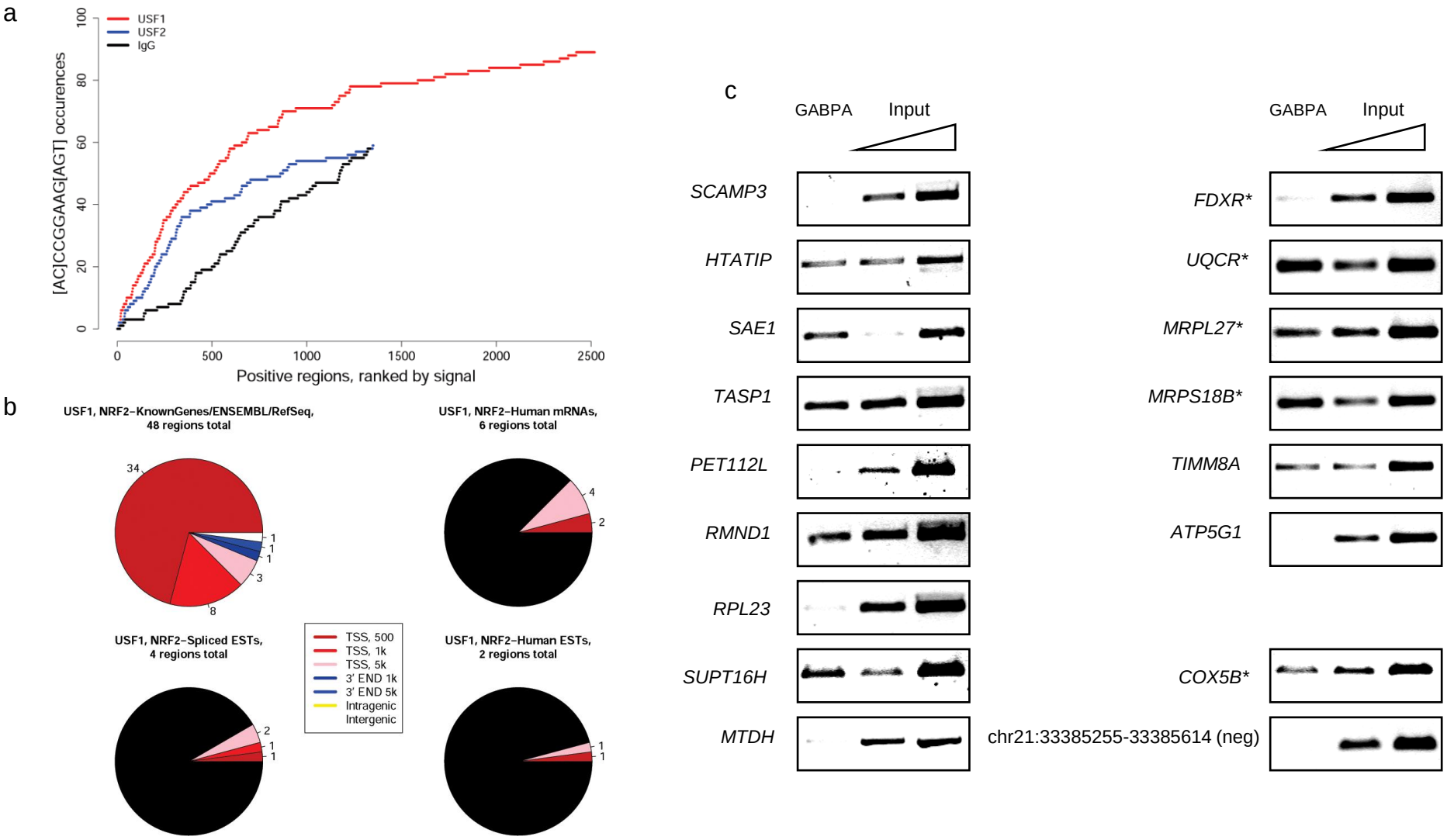


Fig S20. Gene Ontology (GO) analysis of genes bound by USF1 and/or USF2 in proximity of their TSS. Genes being bound by USF1 or USF2 within 5 Kb (5K) or 500 bp (0.5K) were used for GO analysis using the whole human genome as background. Only those GO terms significantly overrepresented after Bonferroni correction for multiple testing for at least one of the USF-binding groups are presented. The colour code indicates different levels of statistical significance. The crosses at the right of the figure indicate those GO categories related with mitochondrial function and/or ATP production. CC: cellular compartment; MF: molecular function; BP: biological process.



Supplementary Figure 21. Cooperativity between USFs and GABPA (NRF-2) at promoter regions.
(a) Candidate scanning approaches were used to search for the occurrence of GABPA (NRF-2) consensus sequences among USF1, USF2 and IgG regions, which were ranked in descending order according their log2 enrichment values. The total number of consensus sequence occurrences with a maximum of four per region (right) were considered. (b) Genomic distribution of USF1 bound regions where at least one perfect match to GABPA (NRF-2) consensus sequences was identified. The distribution for USF2 bound regions was almost identical. (c) Nine USF1 bound regions displaying perfect matches to GABPA consensus sequences were analyzed by semi-quantitative PCR (left). Additionally, six nuclear mitochondrial gene promoters bound by USF1 were also analyzed, together with GABPA positive (*COX5B*) and negative (chr21:33385255-33385614) binding control regions. Those mitochondrial gene promoters with an asterisk contain perfect matches to GABPAconsensus sequence.