The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis

Chia-Feng Liu and Véronique Lefebvre

SUPPLEMENTARY FIGURES

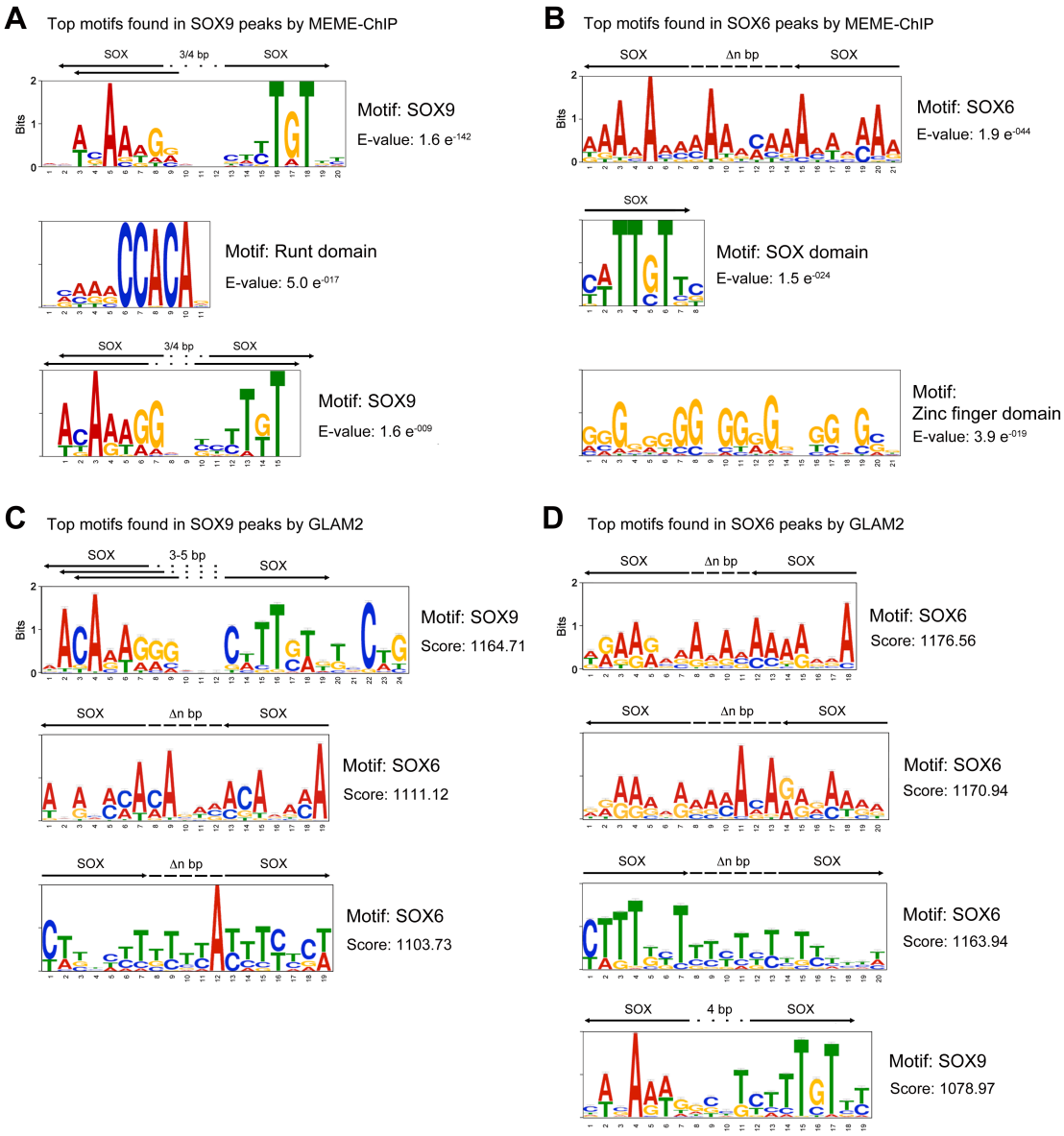


Figure S1. *De novo* discovery of DNA motifs in SOX9 peaks **(A,C)** and SOX6 peaks **(B,D)** using MEME-ChIP **(A,B)** and GLAM2 **(C,D)**. Three motifs identified by MEME-ChIP and the best-scoring motifs identified by GLAM2 (score ≥ 1110) are shown. Black arrows indicate the predicted position and orientation of SOX sites (SOX). Dots schematize intervening nucleotides. Logos labeled with discontinuous lines between black arrows feature multiple SOX-like motifs at overlapping positions. The black arrows show two of these motifs and the discontinuous line indicate nucleotides that could be part of SOX-like binding sites or part of the intervening sequence between two sites.

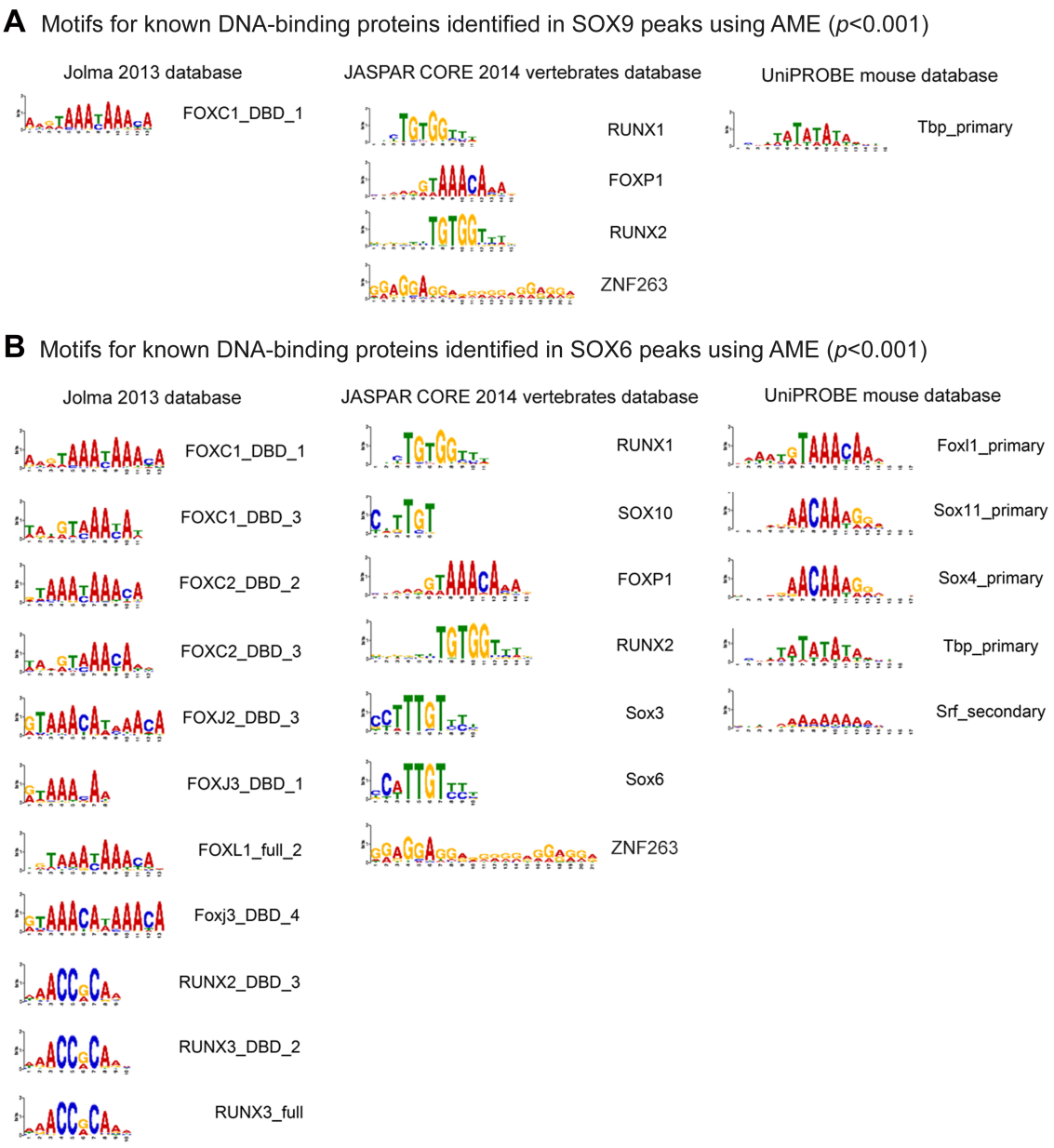
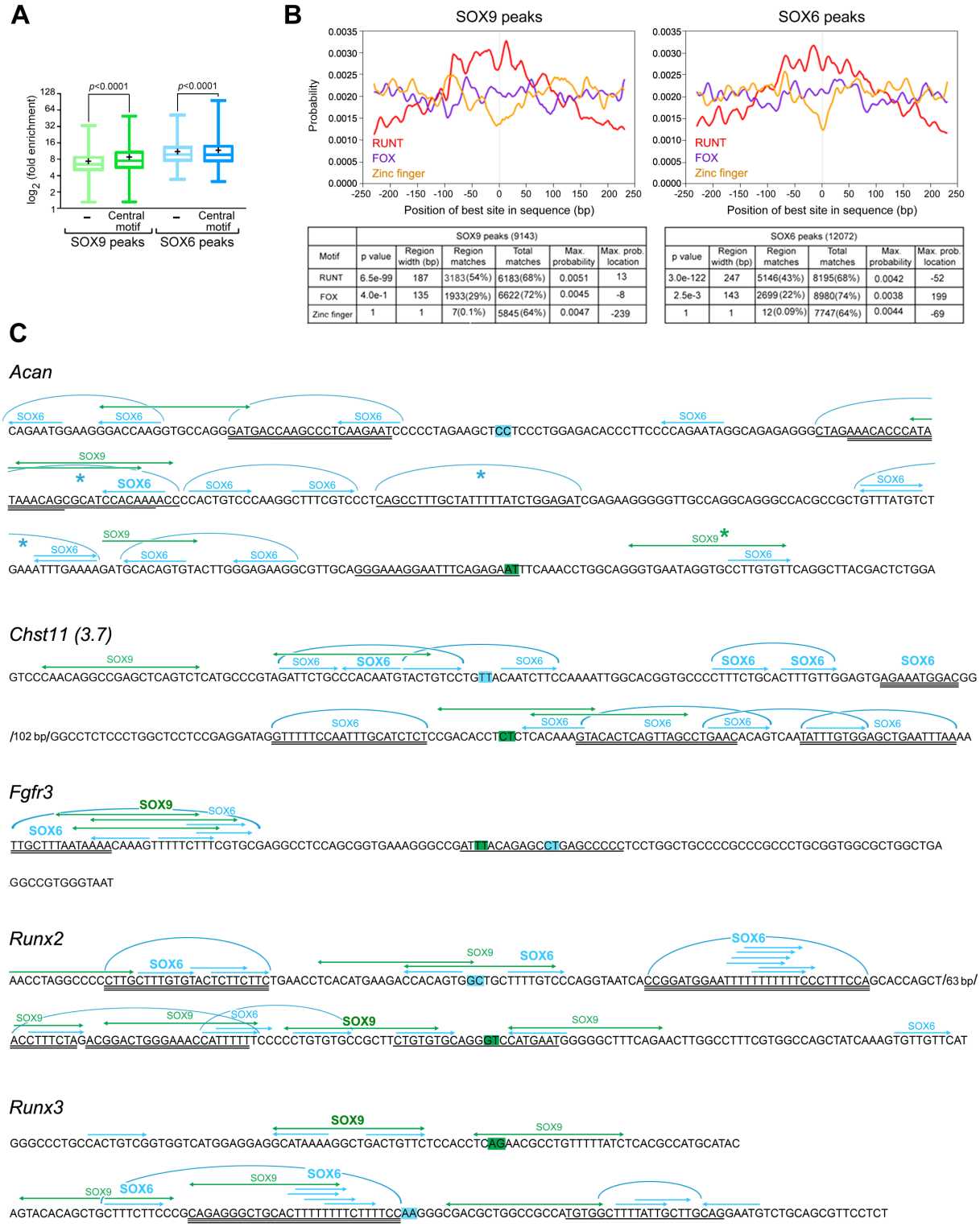


Figure S2. Analysis of the enrichment of known DNA-binding motifs in SOX9 peaks **(A)** and SOX6 peaks **(B)**. Motifs were identified by AME. The best-scoring motifs found using three different databases are shown.



	SOX9 peaks (9143)					
Motif	p value	Region width (bp)	Region matches	Total matches	Max. probability	Max. prob. location
RUNT	6.5e-99	187	3183(54%)	6183(68%)	0.0051	13
FOX	4.0e-1	135	1933(29%)	6622(72%)	0.0045	-8
Zinc finger	1	1	7(0.1%)	5845(64%)	0.0047	-239

	SOX6 peaks (12072)					
p value	Region width (bp)	Region matches	Total matches	Max. probability	Max. prob. location	
3.0e-122	247	5146(43%)	8195(68%)	0.0042	-52	
2.5e-3	143	2699(22%)	8980(74%)	0.0038	199	
1	1	12(0.09%)	7747(64%)	0.0044	-69	

Figure S3. CentriMo and motif analysis in SOX9 and SOX6 peaks. **(A)** Global analysis of the fold enrichment of SOX9 and SOX6 peaks presenting or not presenting central enrichment of SOX motifs**. (B)** CentriMo analysis of SOX9 and SOX6 peaks for RUNT, FOX, and zinc-finger

motifs. **(C)** Analysis of SOX9 and SOX6 motifs in enhancers associated with *Acan* (10 kb 5' of the gene), *Chst11* (3.7 kb 5' of the gene), *Fgfr3* (28 kb 5' of the gene), *Runx2* (306 kb 3' of the transcription start site), and *Runx3* (5.5 kb 5' of the gene). The two nucleotides that correspond to SOX9 and SOX6 peak summits are highlighted in green and blue, respectively. The sequences were scanned for SOX9 and SOX6 motifs (shown in Figure 3D and E) using FIMO. SOX9 motifs are shown with double green arrows. SOX6 single-SOX motifs and their orientation are shown with blue arrows. SOX6 A-rich motifs are underlined. Motifs that have a p value <0.001 are labeled using bold characters for SOX9 and SOX6. Motifs that have a p value between 0.005 and 0.001 are labeled with regular characters. Motifs that have a p value between 0.01 and 0.005 are not labeled. Arched lines encompass SOX6 A-rich sites and tandem SOX6 single-SOX motif pairs. Asterisks in the *Acan* sequence indicate the SOX5/6 and SOX9 sites functionally validated in (21). Gaps of 102 and 63 bp are shown in the *Chst11* and *Runx2* sequences. Please note that these sequences are part of the enhancers that were functionally validated in Figure 8.

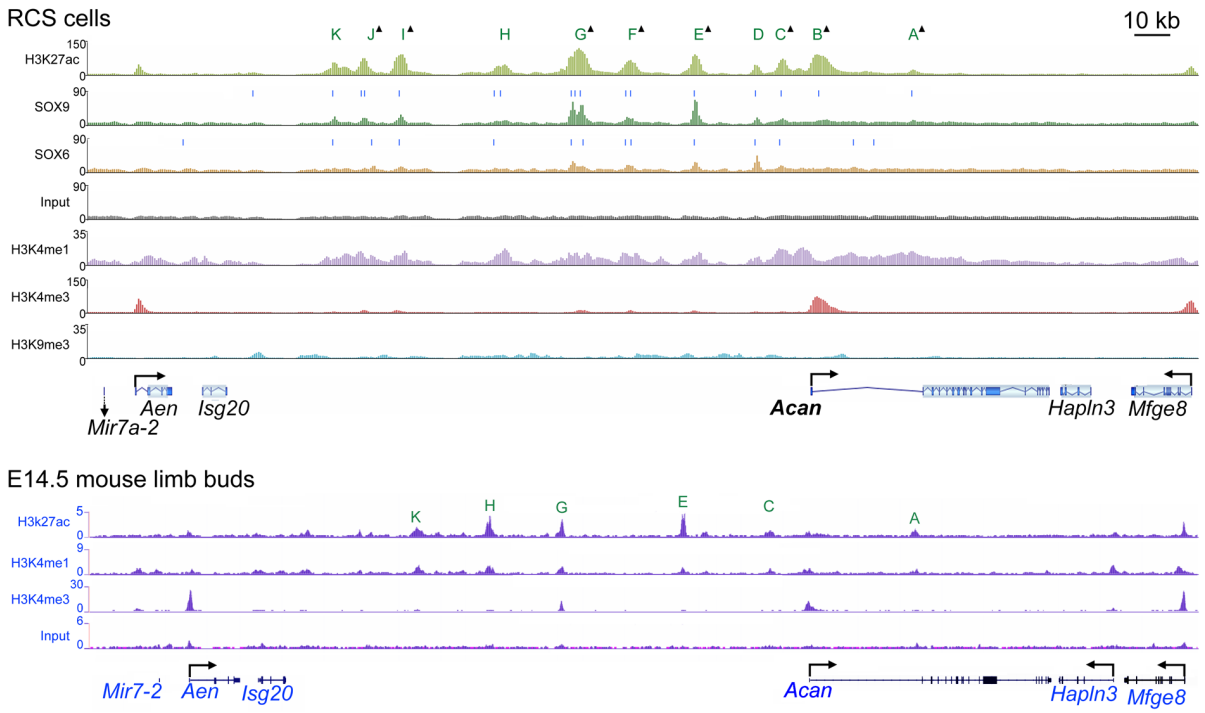


Figure S4. Comparison of newly and previously identified enhancers associated with the *Acan* gene. Profiles of SOX9, SOX6, and histone modification peaks at the *Acan* locus are shown for RCS cells (top) and for E14.5 mouse embryo limbs (bottom). The latter data were downloaded from the UCSC genome browser (mm9). Thirteen regions were identified in the RCS genome that showed highly significant H3K27ac peaks (fold enrichment > 50; labeled A to K). Black triangles in superscript indicate enhancer regions that were previously shown to be capable on their own of activating a reporter transgene in zebrafish or mouse embryos. The mouse orthologues of six of these regions (A, C, E, G, H, and K) carried the most significant H3K27ac peaks in mouse embryo limbs.

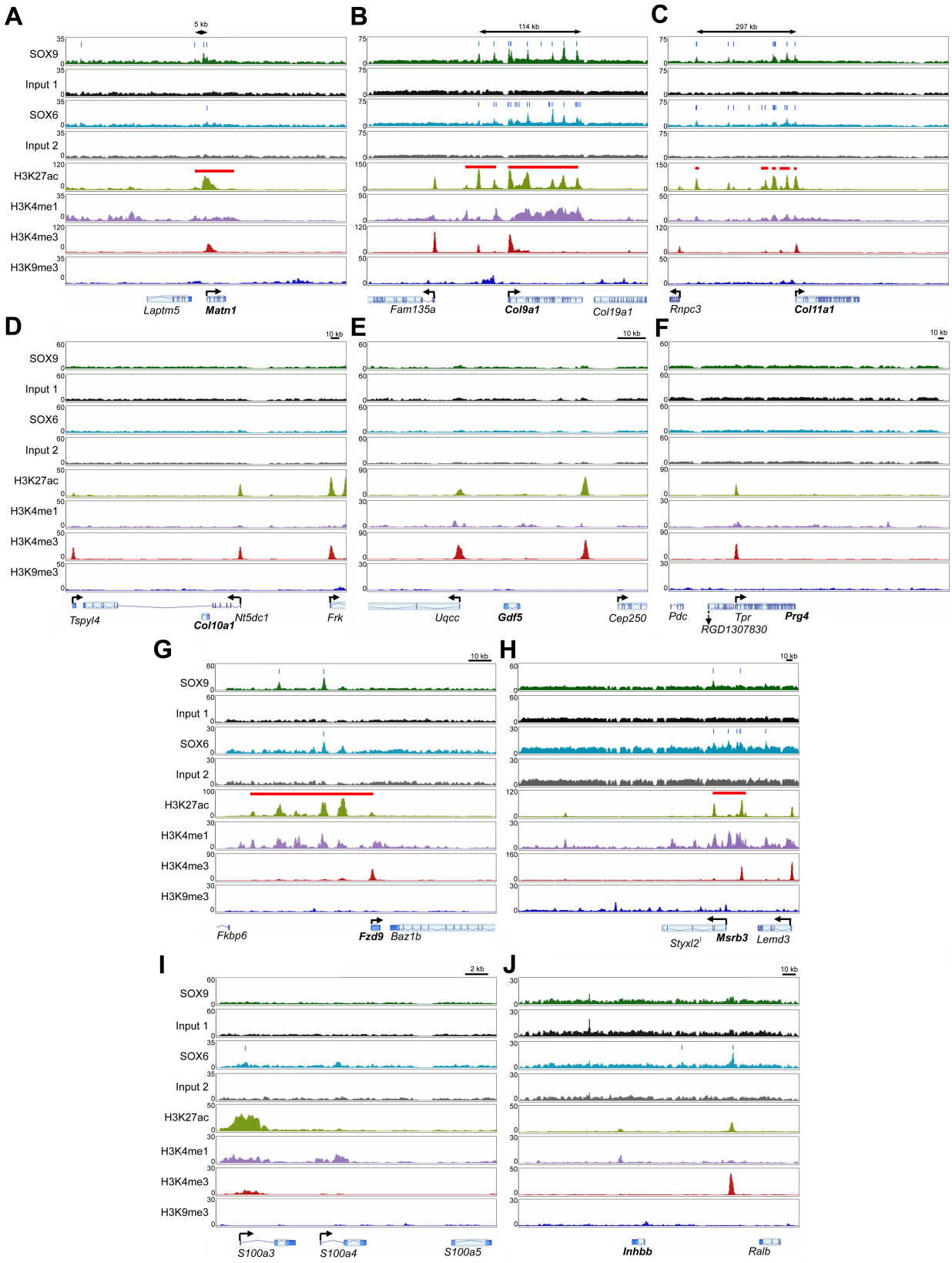

A
B
C
D
E
F
G
H
I
J
SOX9
Input 1
SOX6
Input 2
H3K27ac
H3K4me1
H3K4me3
H3K9me3
Laptm5
Matn1
Fam135a
Col9a1
Col19a1
Rnpc3
Col11a1
Tspyl4
Col10a1
Nt5dc1
Frk
Uqcc
Gdf5
Cep250
Pdc
Tpr
Prg4
RGD1307830
Fkbp6
Fzd9
Baz1b
Styxl2
Msrb3
Lemd3
S100a3
S100a4
S100a5
Inhbb
Ralb

Figure S5. Distribution of SOX peaks and histone modifications within and around genes. Genes of interest are indicated in bold letters. Scale bars are shown as window size reference or to indicate the size of enhancer clusters predicted to influence genes of interest. Blue bars indicate the summits of SOX9 and SOX6 peaks. Red bars indicate super-enhancers.

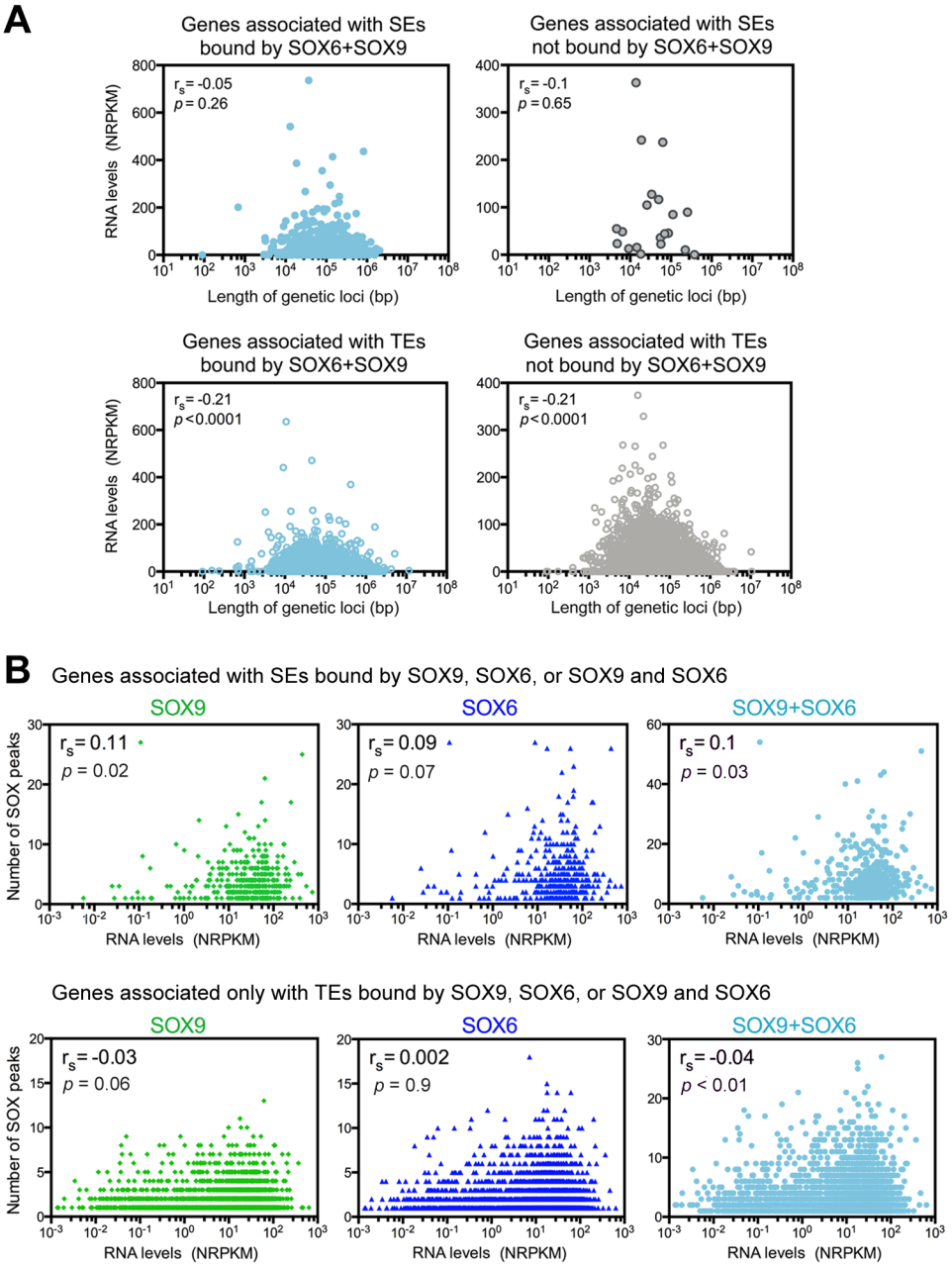


Figure S6. Relationships between RNA levels and genetic locus lengths **(A)**, and between SOX peak numbers and RNA levels **(B)** for genes associated with super-enhancers (SEs) or only with typical enhancers (TEs) in RCS cells. Correlations were determined using the Spearman's correlation test. The Spearman's correlation coefficients (r_s) and *p* values are indicated.

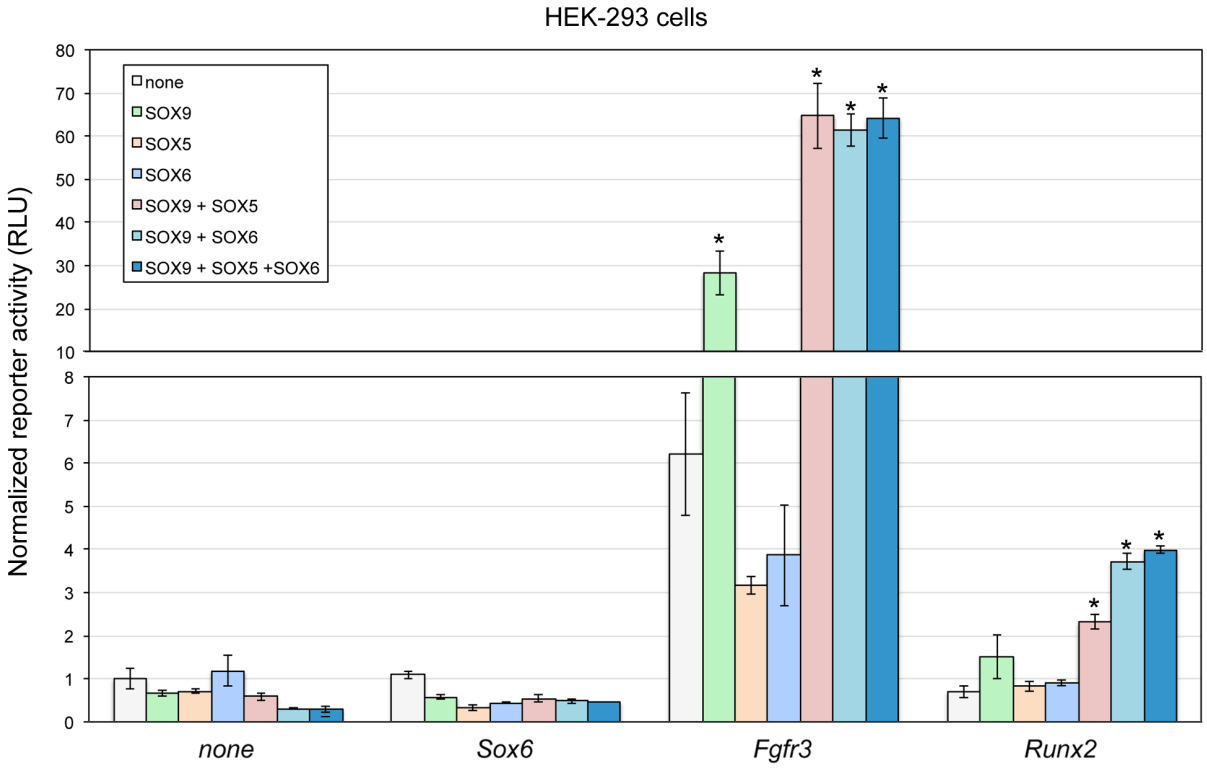


Figure S7. Test of the ability of the SOX trio to activate selected enhancers in non-chondrocytic cells. HEK-293 cells were transiently transfected with reporters harboring a *Col2a1* minimal promoter driven by no enhancer or by an enhancer region associated with the *Sox6*, *Fgfr3*, or *Runx2* genes, as indicated. Various combinations of expression plasmids for the SOX trio members were included in transfection mixtures, also as indicated. Normalized reporter activities are presented as means with standard deviation obtained for triplicate samples in an experiment representative of at least three independent ones. Stars point to reporter activities induced by SOX proteins that are significantly higher than those obtained without SOX proteins (Student's T-test, $p < 0.05$). Data confirm findings previously published for other enhancers that SOX5 and SOX6 are transcriptionally inactive on their own and are similarly capable of cooperating with SOX9 to activate transcription.