Expanded View Figures

Figure EV1. *In vitro* EMT system used to study HOTAIR function.

A Epi and Mes cell lines corresponding to HA5-early and HA5-late, respectively (Castro-Vega *et al*, 2013), stained for F-actin fibers by Phalloidin-TRITC and for DNA by DAPI (scale bar, 16 μm).

B HOTAIR levels quantified in Epi and Mes cells by random-primed RT–qPCR, relative to RPL11 mRNA (gray dots); bars represent mean ± standard deviation for three replicates.

C Protein levels of EMT markers in whole protein extracts of parental Epi and Mes cells assessed by Western blot.

D HOTAIR quantification by random-primed RT–qPCR in Epi cell lines normalized to RPL11 mRNA and represented as a fold change to CTR. Each dot corresponds to independent replicate; bars represent mean ± standard deviation values.

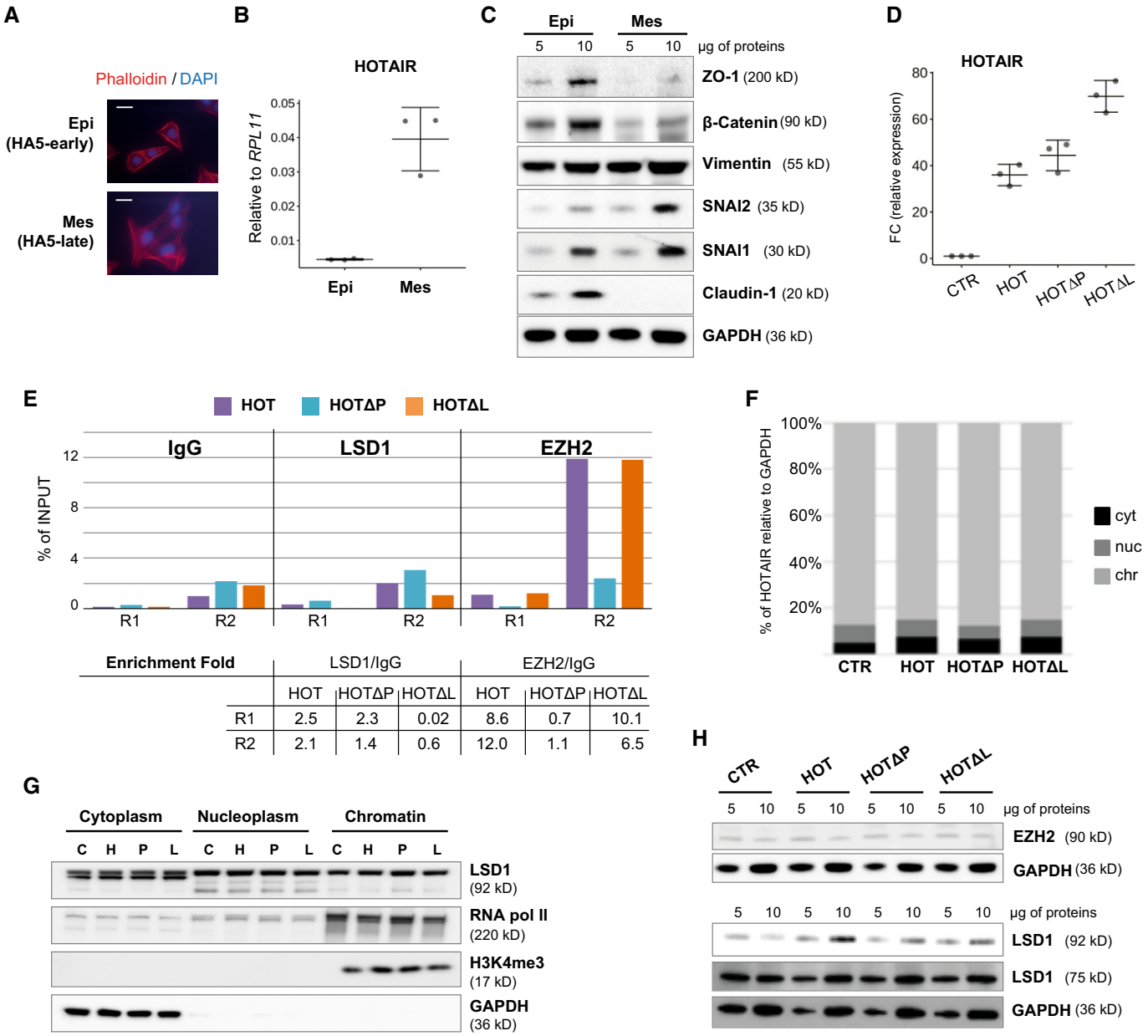
E Quantification of HOTAIR variants associated with LSD1 and EZH2 by RT–qPCR in two independent RIP experiments (R1 and R2), represented as a % of input, with fold-enrichment.

F Distribution of full-length and truncated variants of HOTAIR between cytoplasm (cyt), nucleoplasm (nuc), and chromatin (chr) fractions in Mes and transduced Epi cells (HOT, HOTΔP, HOTΔL) assessed by RT–qPCR relative to GAPDH mature mRNA.

G Subcellular distribution and protein levels of LSD1, RNA Pol II, H3K4me3, and GAPDH in cytoplasm, nucleoplasm, and chromatin fractions in Epi cells expressing CTR (C), HOT (H), HOTΔP (P), and HOTΔL (L).

H Protein levels of LSD1, EZH2, and GAPDH in whole protein extracts assessed by Western blot in Epi-CTR, -HOT, -HOTΔP, and -HOTΔL cells. * The Lsd1 signal at 75 kD was detected only in total and cytoplasm protein fractions and may correspond to either degradation product or Lsd1 variant.

Source data are available online for this figure.



Enrichment Fold	LSD1/IgG			EZH2/IgG		
	HOT	HOTΔP	HOTΔL	HOT	HOTΔP	HOTΔL
R1	2.5	2.3	0.02	8.6	0.7	10.1
R2	2.1	1.4	0.6	12.0	1.1	6.5

Figure EV1.

Figure EV2. H3K27ac discriminates low- from high-migration phenotype.

A H3K27ac reads distribution across distinct genomic features, and H3K27ac peaks identification and assignment to gene promoters in Epi-CTR and cells expressing HOT, HOTΔP, and HOTΔL variants.

B Box-plot distribution of H3K27ac mid-peak to the closest TSS in Epi-CTR (N = 5,770), -HOT (N = 4,103), HOTΔP (N = 4,825), and HOTΔL (N = 5,282). Central bars represent the median, boxes—lower and upper quartiles and whiskers are set to 1.5*interquartile range (IQR) above the third and below the first quartiles.

C Metagene profiling within the 10-kb window around TSS of H3K27ac, LSD1, and H3K4me2 over IgG in Epi cells for three classes of gene promoters: enriched in H3K27ac in low-migrating CTR and HOTΔL cells (Class 1, red) or in high-migrating HOT and HOTΔP cells (Class 2, gray) and all other cases (Class 3, yellow). Red and yellow parallel lines correspond to peak summit values in Epi-CTR cells for genes of Class 1 and Class 2, respectively. Epi cells expressing none (CTR) or different HOTAIR variants (HOT, HOTΔP, and HOTΔL).

D Box-plot of Log_2 fold changes (FC) of H3K27ac mean density of each peak in pairwise comparisons between four Epi cell lines expressing CTR (C), HOT (H), HOTΔP (P), and HOTΔL (L) for Class 1 (N = 7,467), Class 2 (N = 426), and Class 3 (N = 7,722) transcripts. Central bars represent the median, boxes—lower and upper quartiles and whiskers are set to 1.5*IQR above the third and below the first quartiles.

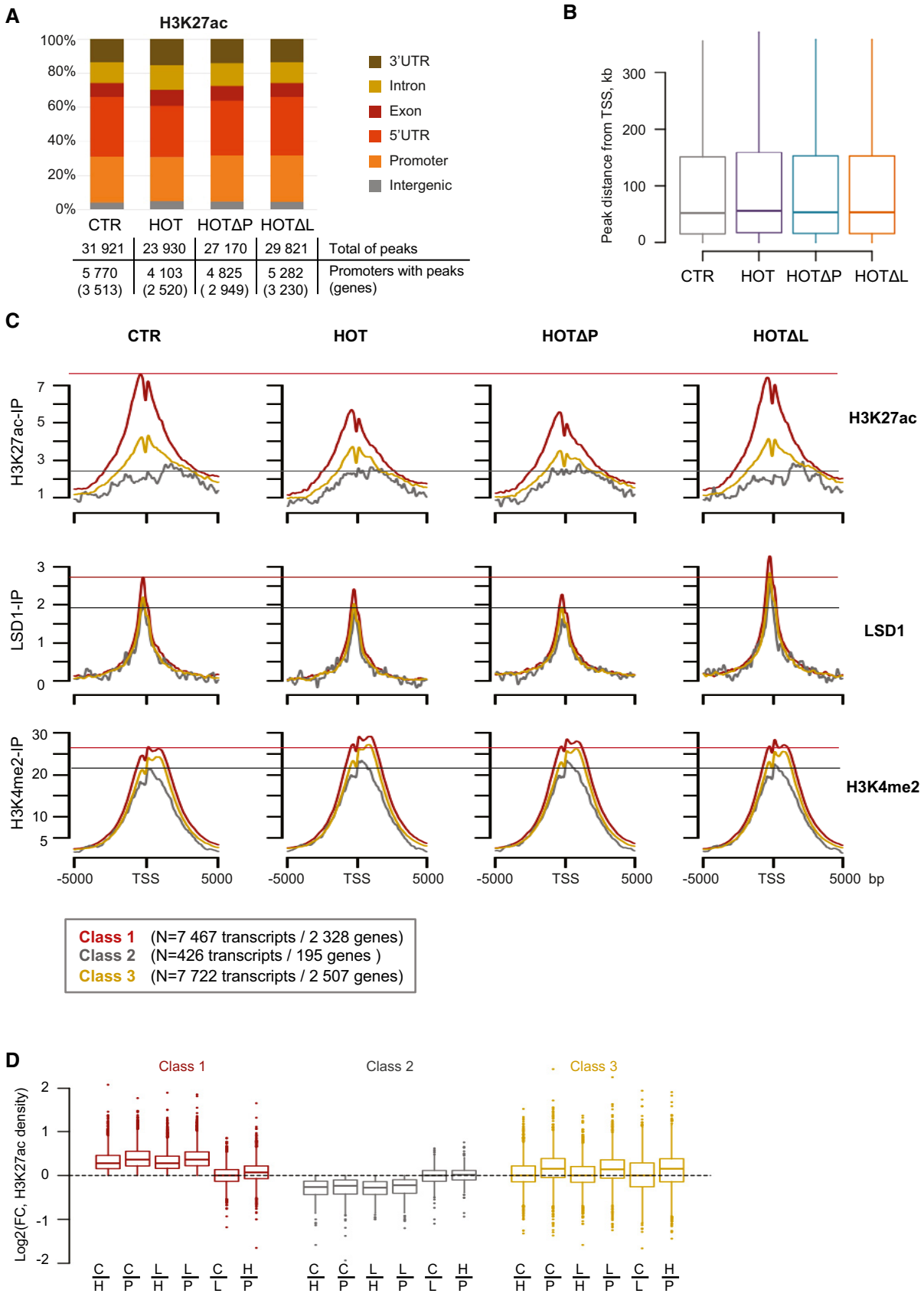

A
H3K27ac
100%
80%
60%
40%
20%
0%
CTR HOT HOTΔP HOTΔL
3'UTR
Intron
Exon
5'UTR
Promoter
Intergenic
31 921 23 930 27 170 29 821 Total of peaks
5 770 (3 513) 4 103 (2 520) 4 825 (2 949) 5 282 (3 230) Promoters with peaks (genes)
B
Peak distance from TSS, kb
300
200
100
0
CTR HOT HOTΔP HOTΔL
C
CTR HOT HOTΔP HOTΔL
H3K27ac-IP
H3K27ac
LSD1-IP
LSD1
H3K4me2-IP
H3K4me2
-5000 TSS 5000 bp
Class 1 (N=7 467 transcripts / 2 328 genes)
Class 2 (N=426 transcripts / 195 genes)
Class 3 (N=7 722 transcripts / 2 507 genes)
D
Class 1
Class 2
Class 3
Log2(FC, H3K27ac density)
C/H C/P L/H L/P C/L H/P

Figure EV2.

Figure EV3. LSD1 and H3K4me2 profiling in Epi cells expressing HOTAIR variants.

A LSD1 tags distribution across distinct genomic features, and a summary of LSD1 peaks, regions, and gene promoters with LSD1 peaks in Epi cells expressing CTR or HOT, HOTΔP, and HOTΔL variants.

B Box-plot distribution of LSD1 mid-peak to the closest TSS in Epi-CTR (*N* = 3,548), -HOT (*N* = 2,797), HOTΔP (*N* = 1,670), and HOTΔL (*N* = 2,754). Central bars represent the median, boxes—lower and upper quartiles and whiskers are set to 1.5*IQR above the third and below the first quartiles.

C Venn diagram representing intersections of LSD1 peak regions in Epi cells.

D H3K4me2 tags distribution across distinct genomic features, and H3K4me2 peaks identification, merging to regions and assignment to gene promoters in Epi-CTR and cells expressing HOT, HOTΔP, and HOTΔL variants.

E Box-plot distribution of H3K4me2 mid-peak to the closest TSS. in Epi-CTR (*N* = 22,172), -HOT (*N* = 21,636), HOTΔP (*N* = 22,199), and HOTΔL (*N* = 21,874). Central bars represent the median, boxes—lower and upper quartiles and whiskers are set to 1.5*IQR above the third and below the first quartiles.

F Venn diagram representing intersections of H3K4me2 peak regions in Epi cells.

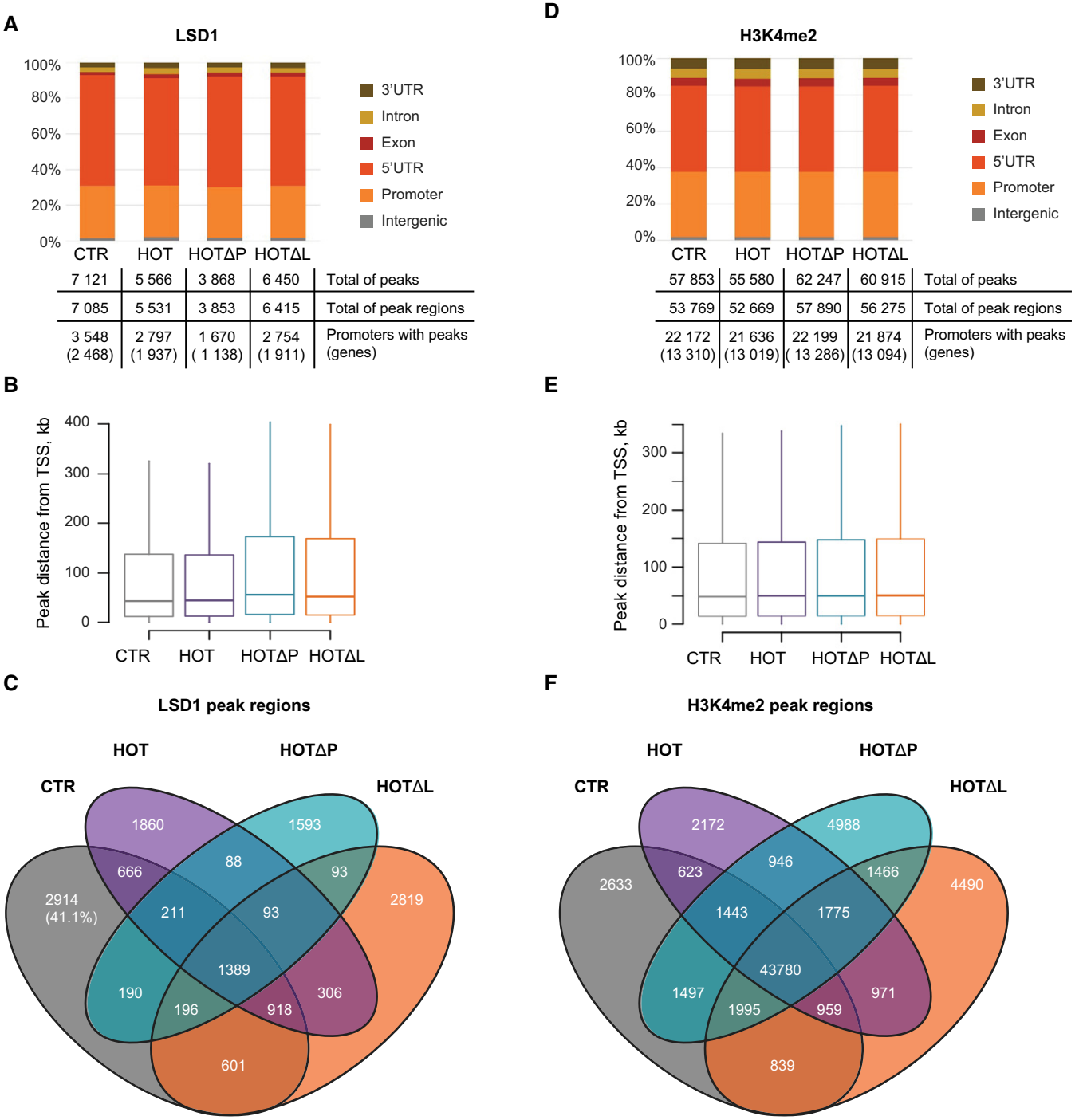

A
LSD1
100%
80%
60%
40%
20%
0%
CTR HOT HOTΔP HOTΔL
3'UTR
Intron
Exon
5'UTR
Promoter
Intergenic
7 121 | 5 566 | 3 868 | 6 450 | Total of peaks
7 085 | 5 531 | 3 853 | 6 415 | Total of peak regions
3 548 (2 468) | 2 797 (1 937) | 1 670 (1 138) | 2 754 (1 911) | Promoters with peaks (genes)
D
H3K4me2
57 853 | 55 580 | 62 247 | 60 915 | Total of peaks
53 769 | 52 669 | 57 890 | 56 275 | Total of peak regions
22 172 (13 310) | 21 636 (13 019) | 22 199 (13 286) | 21 874 (13 094) | Promoters with peaks (genes)
B
Peak distance from TSS, kb
400
300
200
100
0
CTR HOT HOTΔP HOTΔL
E
300
200
100
0
C
LSD1 peak regions
HOT
HOTΔP
CTR
HOTΔL
1860
1593
666
88
93
2914 (41.1%)
211
93
2819
1389
190
196
918
306
601
F
H3K4me2 peak regions
2172
4988
623
946
1466
2633
1443
1775
4490
43780
1497
1995
959
971
839

Figure EV3.

A

		Epi-CTR		Epi-HOT		Epi-HOTΔP		Epi-HOTΔL	
		R1	R2	R1	R2	R1	R2	R1	R2
Enhancers	**Total**	12 588	14 848	15 422	8 837	12 841	10 734	13 090	12 320
	Common R1 & R2	**10 640**		**7 382**		**9 047**		**10 008**	
Super-enhancers	**Total**	738	626	879	606	801	691	683	923
	Common R1 & R2	**573**		**577**		**653**		**662**	

B

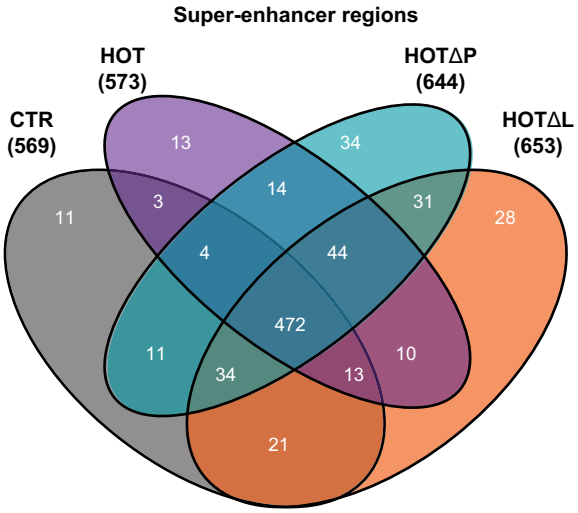


Figure EV4. Enhancers and super-enhancers in Epi cells expressing HOTAIR variants.

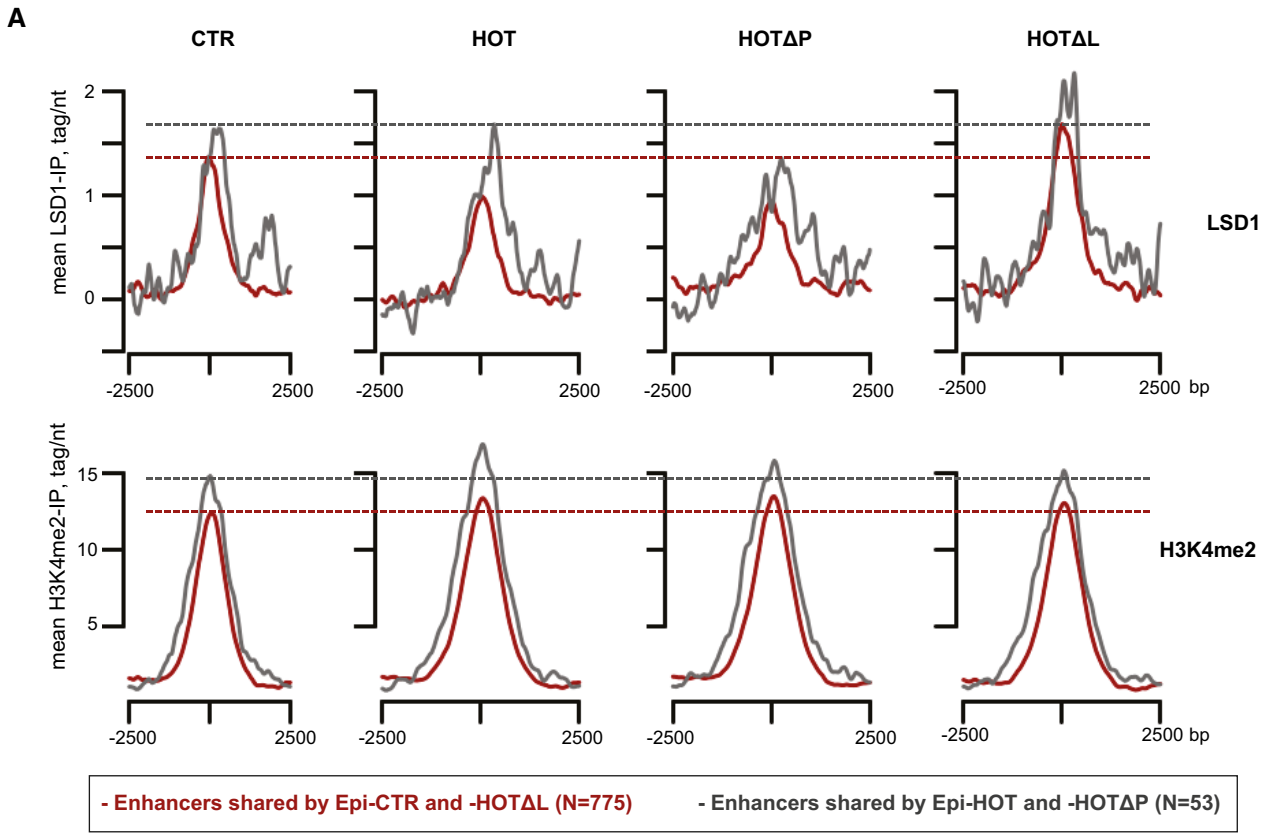
A Annotation of active enhancers and super-enhancers by LILY using H3K27ac as a hallmark, in Epi cells expressing none (CTR), full-length (HOT) and truncated variants (HOTΔP and HOTΔL) of HOTAIR.

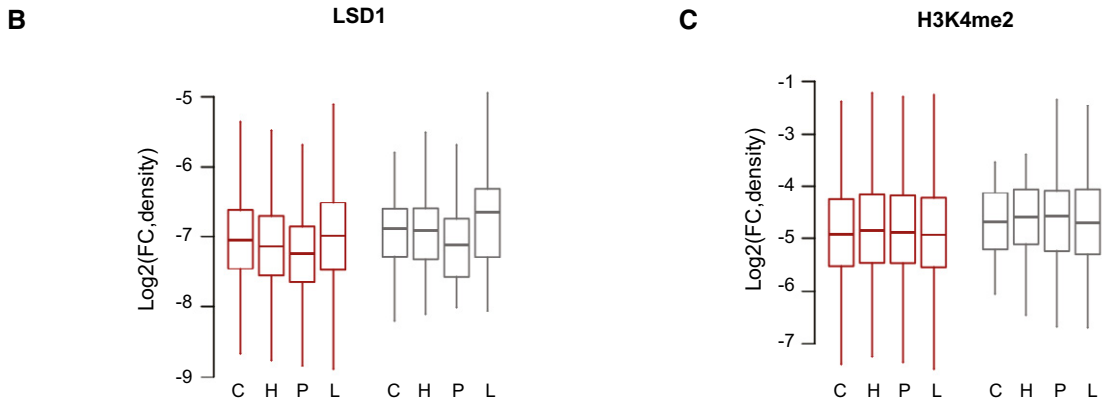
B Venn diagram representing intersections of active super-enhancer regions in Epi cells.

Figure EV5. LSD1 and H3K4me2 levels at enhancers shared by low- or high-migrating cells.

A Metagene profiling of LSD1 and H3K4me2 over IgG tags ratio within the 5-kb window of enhancers shared by low-migrating Epi-CTR and -HOTΔL cells (red) or by high-migrating Epi-HOT and -HOTΔP cells (gray). Parallel lines indicate the peak summit in Epi-CTR, bp stands for base pairs.

B, C Box-plot of Log_2 LSD1 and H3K4me2 densities at enhancers shared by low- (red, $N = 775$) and high-migrating cells (gray, $N = 53$). Central bars represent the median, boxes—lower and upper quartiles and whiskers are set to 1.5*IQR above the third and below the first quartiles. C = CTR, H = HOT, P = HOTΔP, L = HOTΔL. Adj. *P*-value was calculated by Bonferroni method, Wilcoxon test and represented only for significant changes observed for LSD1 at enhancers shared by low-migrating cells.





Adj. *P*-value, for enhancers shared by **Epi-CTR and -HOTΔL**

	CTR	HOT	HOTΔL
HOT	0.049	-	-
HOTΔL	2.8×10^{-9}	2.9×10^{-4}	-
HOTΔP	1	0.002	3.4×10^{-12}

Figure EV5.