Table S1, related to Figure 1: The total number of sequenced reads obtained from the Oct4, Sox2, Klf4, and c-Myc ChIP reactions in infected and non-infected (mock) cells treated with dox for 48 hr and the total sequenced tags from non-immuno enriched DNA (input) samples are indicated. Multiple reads that aligned to the same genomic regions were not considered in our analysis only unique reads are shown. Genomic coverage is displayed as a fraction of the human genome (hg18 build 36) that aligned with sequenced reads.

	Total read count		Unique reads	
Factor	aligned reads	genomic coverage	aligned reads	genomic coverage
Input 1	35206971	0.39 X	33516500	0.37 X
Input 2	36887920	0.41 X	32849269	0.37 X
Oct4	38052089	0.42 X	35865026	0.4 X
Sox2	30030469	0.33 X	29208966	0.32 X
Klf4	36521658	0.41 X	33573804	0.37 X
cMyc	36969964	0.41 X	34070915	0.38 X
Oct4 mock	34523589	0.38 X	22168416	0.25 X
cMyc mock	32605607	0.36 X	24482375	0.27 X

Notes for Tables S2, S3, and S4, as separate Excel files:

- Table S2, related to Figures 2, 3, 6, 7 and S2B, S3C, S5B, and S7F (excel file): The DNA oligonucleotides used as primers for ChIP-qPCR are listed along with the amplified genomic regions in the hg-18 build 36 human genome and the associated RefSeq genes, and the validated OSKM occupancy at 48 hr. *Asterisks denote sites used for ChIP-qPCR experiments shown in Figures 7B and S7F.

- **Table S3, related to Figures 1 and 2 (excel file):** Genes' sites of occupancy (bound vs. not bound) by O, S, K, M at 48 hr, as summarized in Figure 1C, along with gene expression correlations in uninfected human fibroblasts, human iPS cells, and human ES cells (Lowry et al. 2008) (see Figure 2A).

Note that we observed a few peaks in the coding sequences of the *POU5f1 (Oct4), SOX2, KLF4, and c-MYC* genes, which could be due to contamination with lentivirus proviral DNA that contains the coding sequences in the infected cells. Thus, these genes were provisionally classified as "not bound" in Table S3.

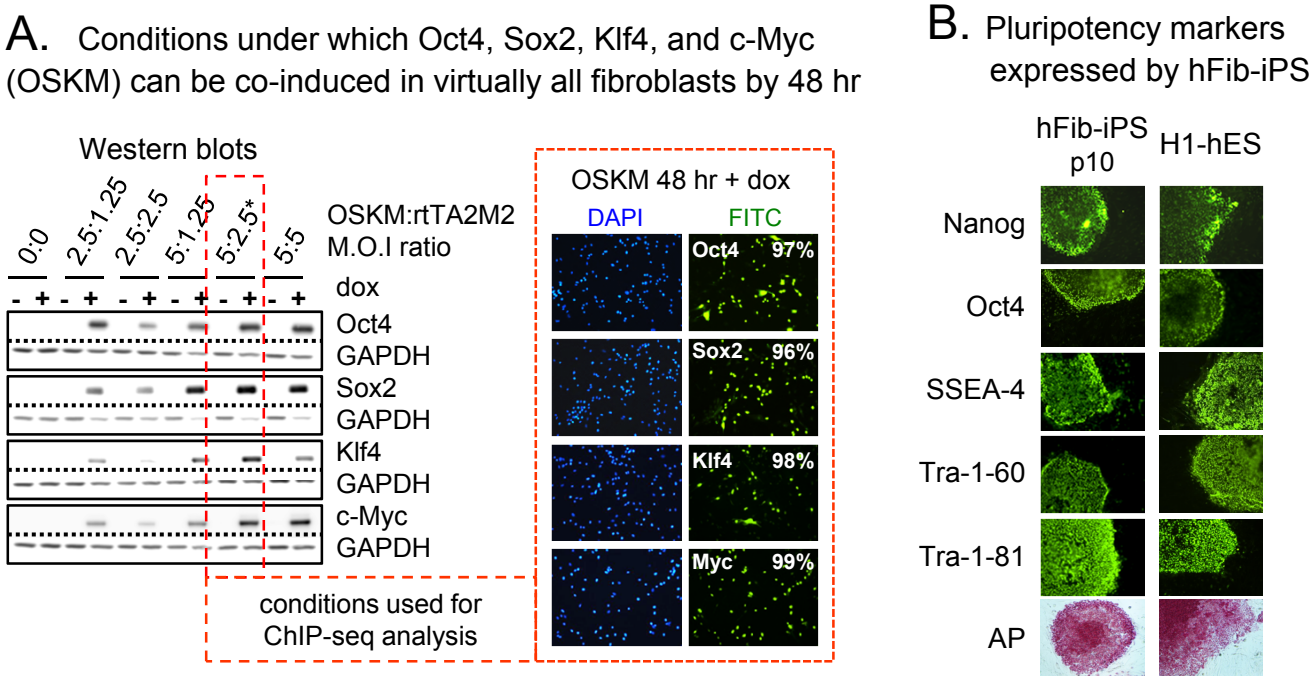
- Table S4, related to Figures 6 and 7 (excel file): OSKM-DBR data, including:

- Tab 1: - genome location of 264 OSKM-DBRs (HG18)
  - genome location of overlapping iPS DMRs (Lister et al. 2011)
  - genome location and extent of overlap with lamin-associated domains (LADs) (Guelen et al. 2008)
- Tab 2: - RefSeq genes within OSKM-DBRs
- Tab 3: - Gene Ontology (GO) analysis of genes within OSKM-DBR
- Tab 4: - Repeat element enrichment within OSKM-DBRs

Table S5, related to Figures 1 and S1C: The DNA oligonucleotides used as primers for RT-qPCR are listed along with the associated genes' names

gene	forward Primer	reverse Primer
AFP	AGC TTG GTG GTG GAT GAA AC	CCC TCT TCA GC AAA GCA GAC
Endogenous OCT4	GAC AGG GGG AGG GGA GGA GCT AGG	CTT CCC TCC AAC CAG TTG CCC CAA AC
Ker	CAG ATG CTG TGT CCC TGT GT	TTC AGA TCC AGA AGG GGA TG
NANOG	CAG CCC CGA TTC TTC CAC CAG TCC C	CGG AAG ATT CCC AGT CGG GTT CAC C
NCAM	ATG GAA ACT CTA TTA AAG TGA ACC TG	TAG ACC TCA TAC TCA GCA TTC CAG T
Pax6	CCA GAA AGG ATG CCT CAT AAA GG	TCT GCG CGC CCC TAG TTA
RUNX1	CCC TAG GGG ATG TTC CAG AT	TGA AGC TTT TCC CTC TTC CA
T	ACCCAGTTCATAGCGGTGAC	CCATTGGGAGTACCCAGGTT
GAPDH	TGT TGC CAT CAA TGA CCC CTT	CTC CAC GAC GTA CTC AGC G

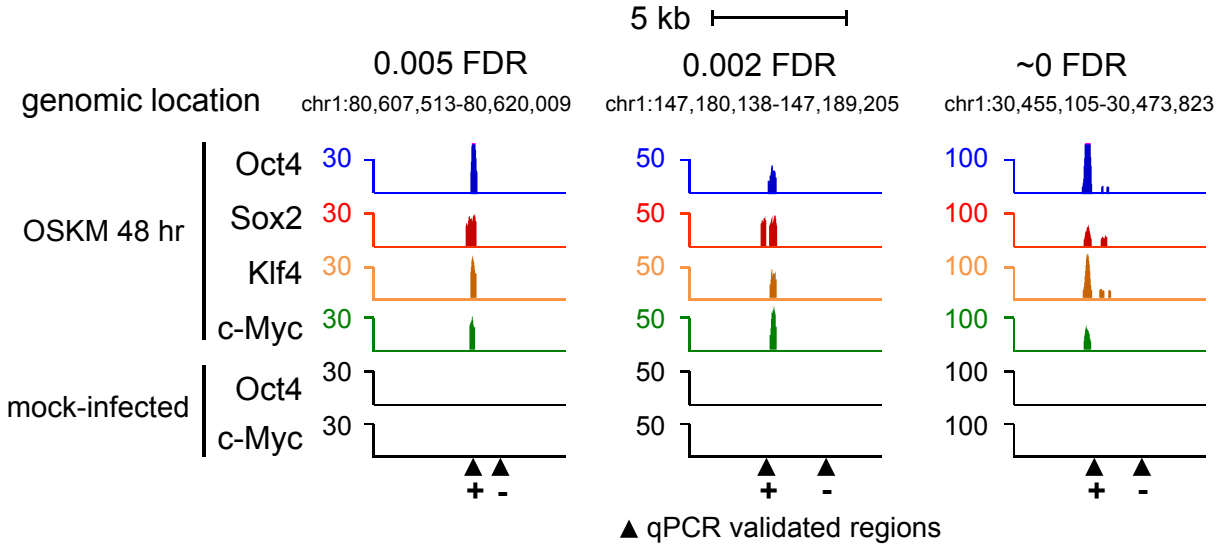
Soufi/Zaret Supplementary Figure 1, related to Figure 1



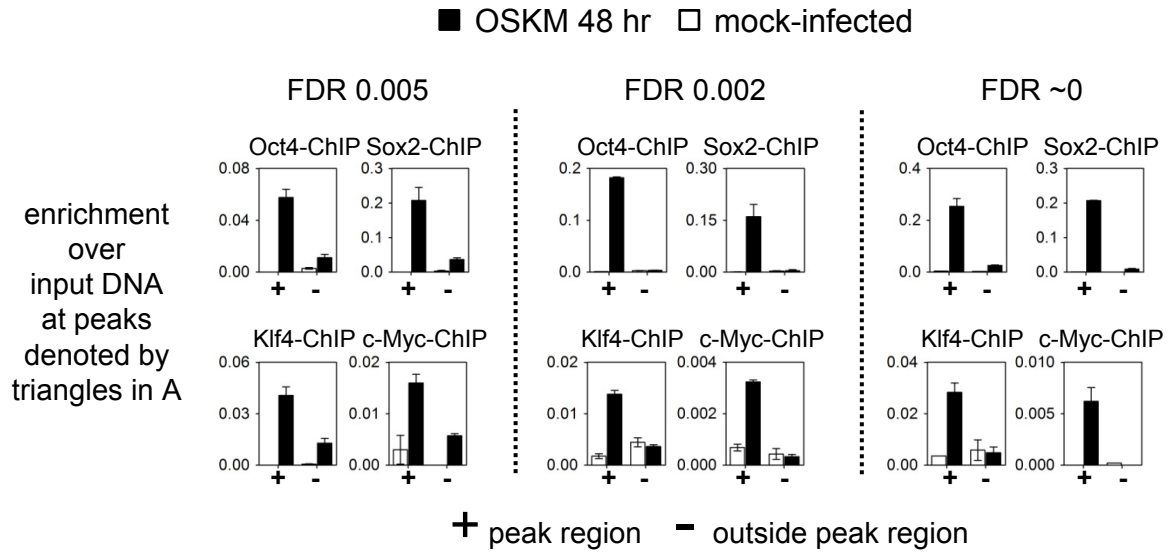
Supplemental Figure 2

Soufi/Zaret Supplementary Figure 2, related to Figures 1 and 2

A. Input normalized ChIP-seq peaks with different FDR values, showing quality of peaks out to 0.005 FDR; the threshold used for analysis in this study

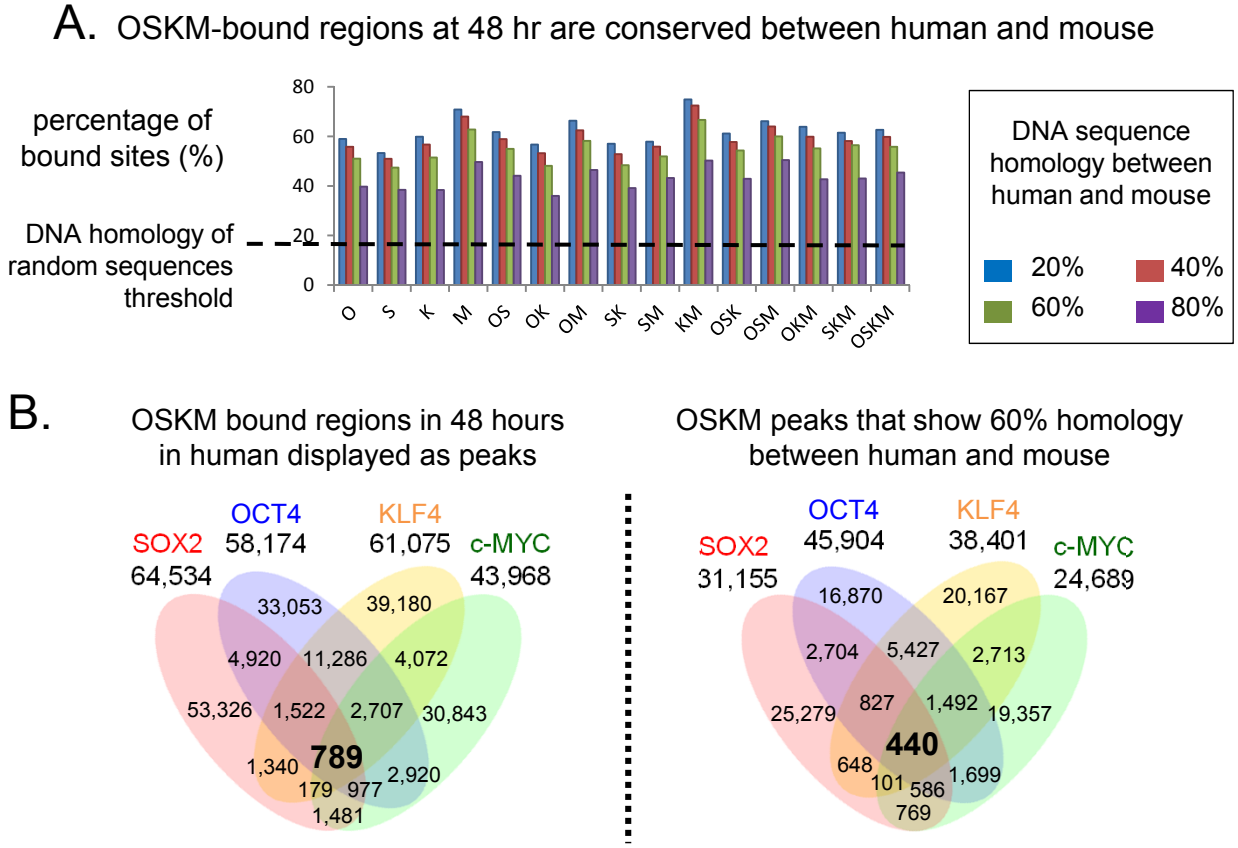


B. ChIP-seq peaks with different FDR values validated by ChIP-qPCR.



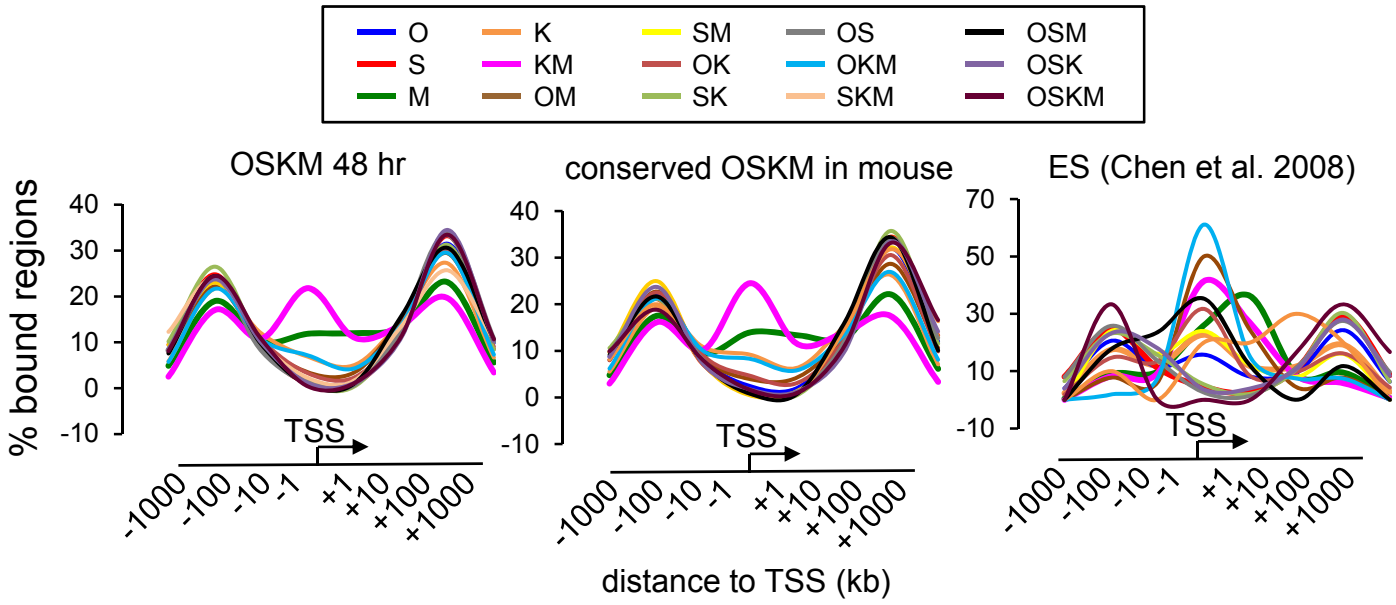
Supplemental Figure 3

Soufi/Zaret Supplementary Figure 3, related to Figures 1 and 2

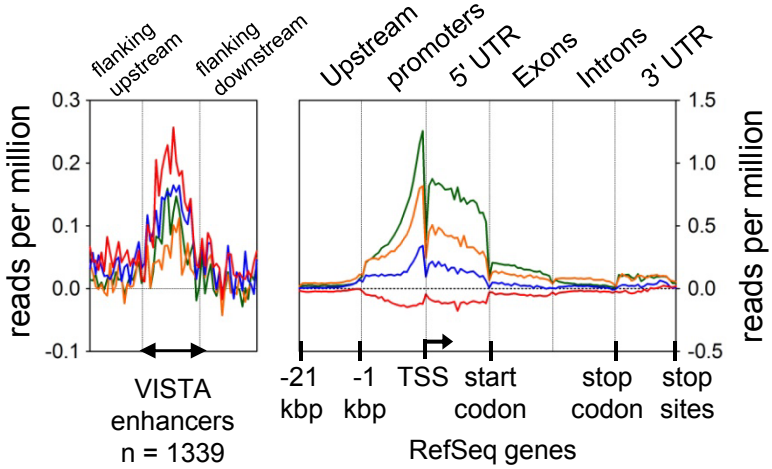


Soufi/Zaret Supplementary Figure 4, related to Figure 2

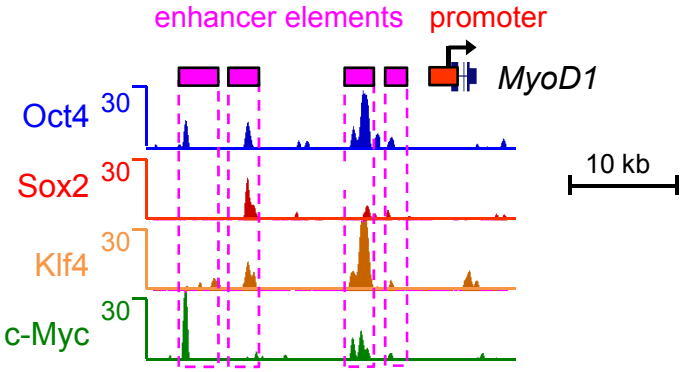
A. Sites bound by OSKM in early reprogramming are predominantly distal to the TSS in both the human and mouse genomes, compared to binding in ES cells



B. By 48 hr, all of O,S,K,M associate with numerous functional enhancers in the VISTA database



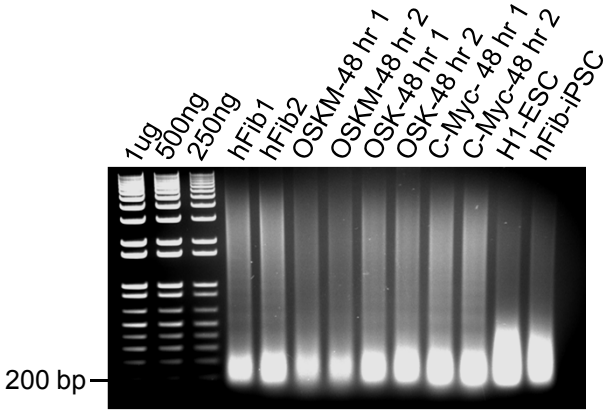
C. OSKM access the *MyoD1* enhancer elements at 48 hr



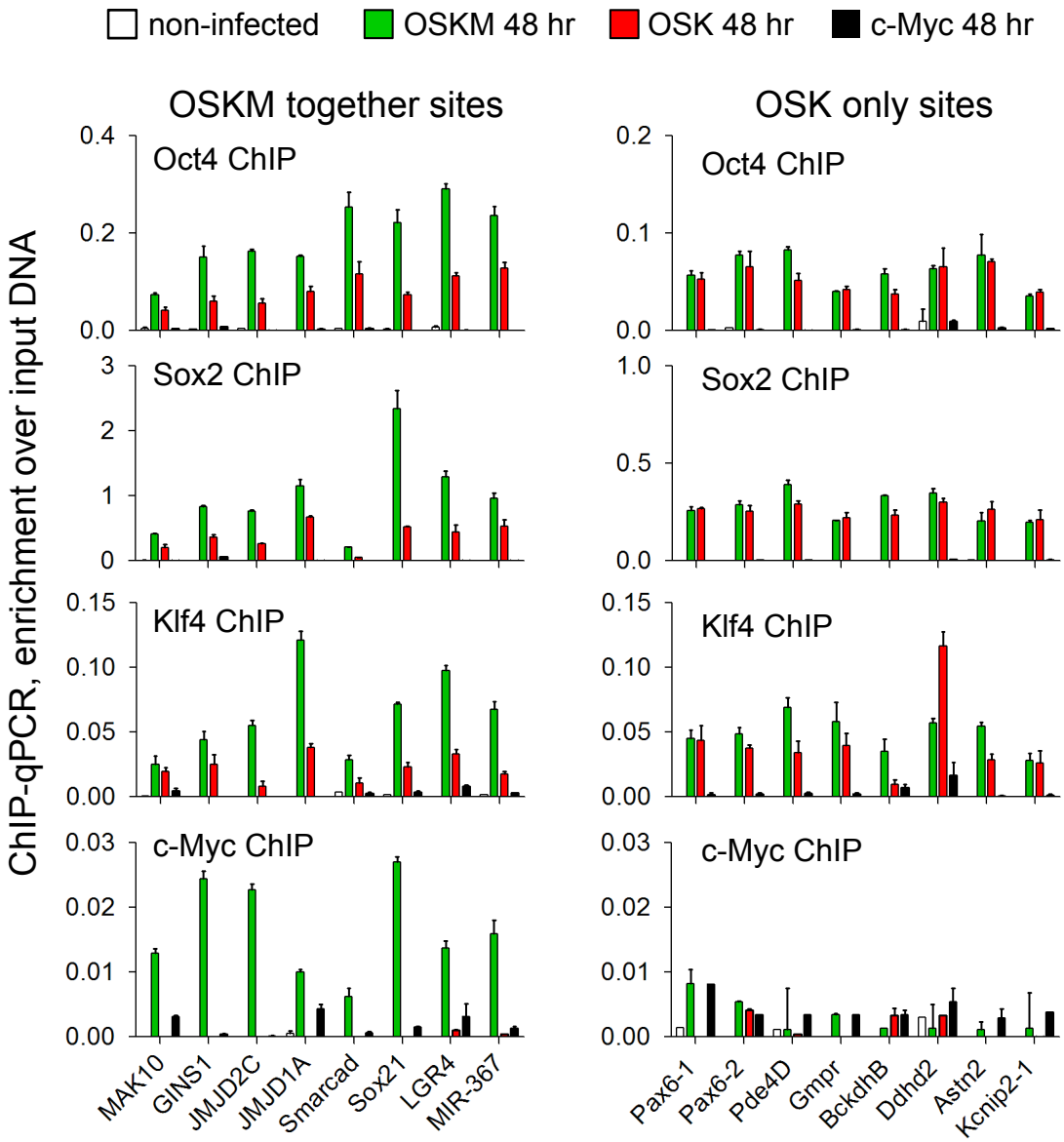
Supplemental Figure 5

Soufi/Zaret Supplementary Figure 5, related to Figure 3

A. Sonicated chromatin used for ChIP-qPCR to investigate the role of c-Myc in early reprogramming

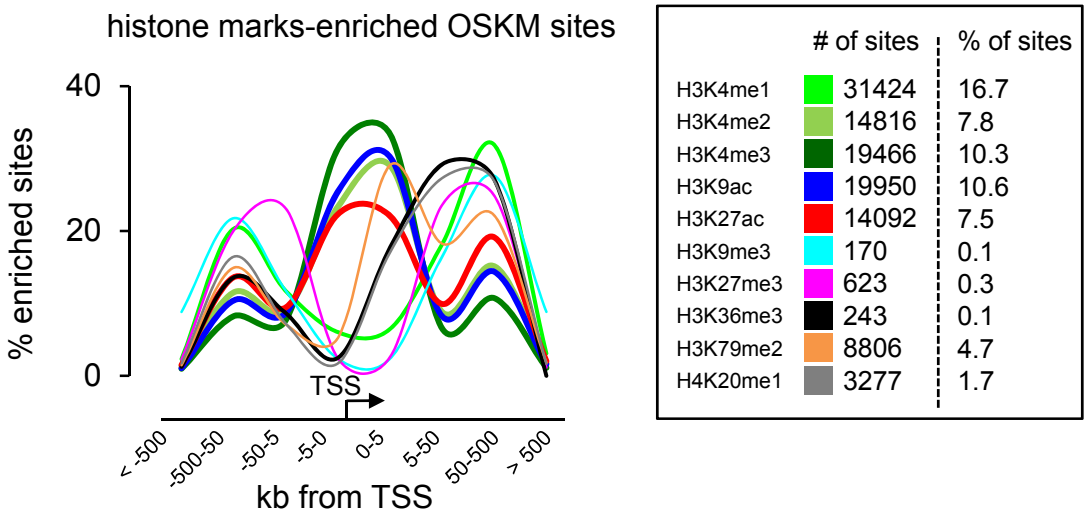


B.

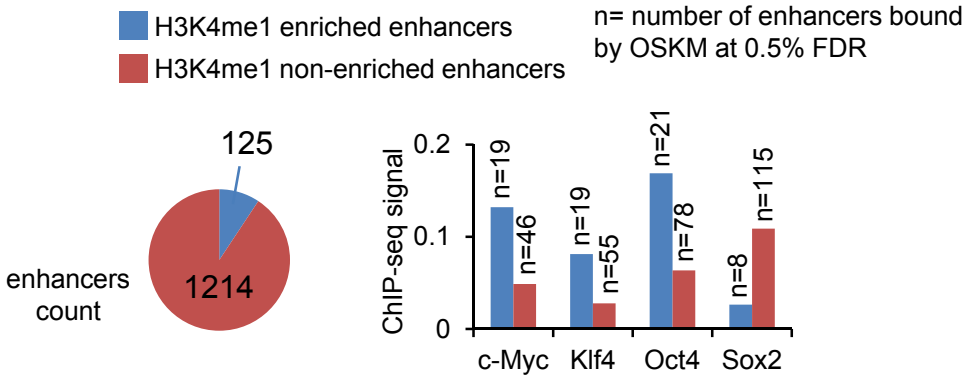


Soufi/Zaret Supplementary Figure 6, related to Figure 5

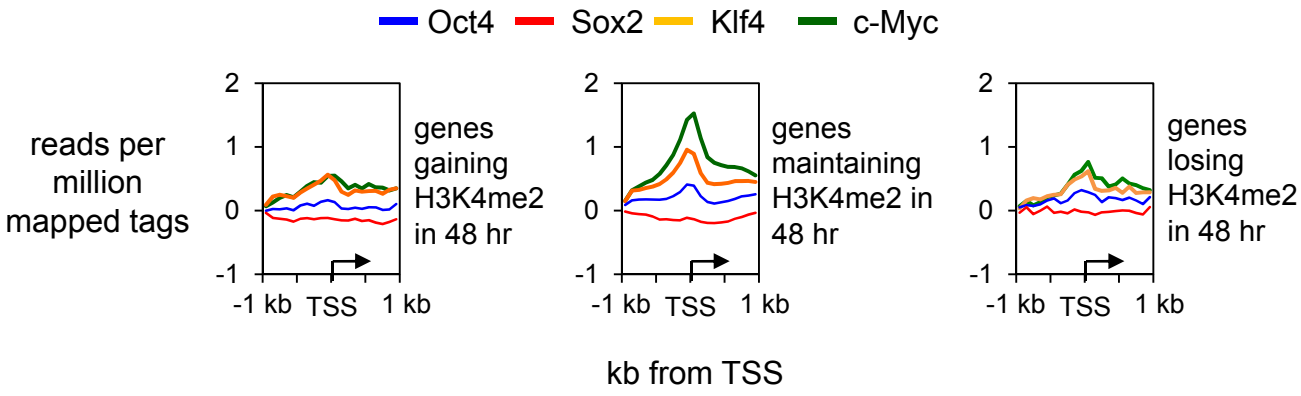
A. Sites enriched for open chromatin marks that OSKM initially bind in fibroblasts are predominantly distributed around promoters (TSS)



B. The OSKM factors are associated with H3K4me1-enriched and non-enriched enhancer elements found in the VISTA database

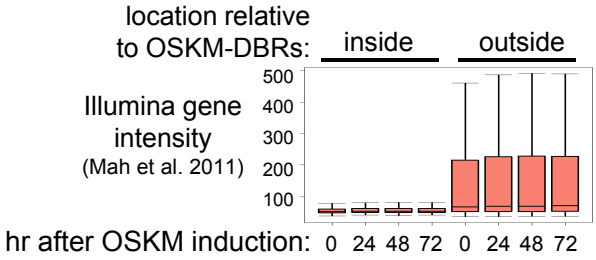


C. A subset of c-Myc and Klf4 bind promoters of genes with pre-existing H3K4me2

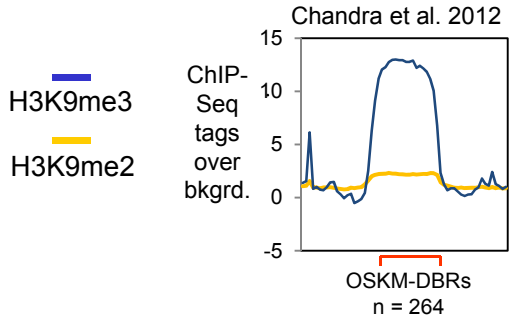


Soufi/Zaret Supplementary Figure 7, related to Figures 6 and 7

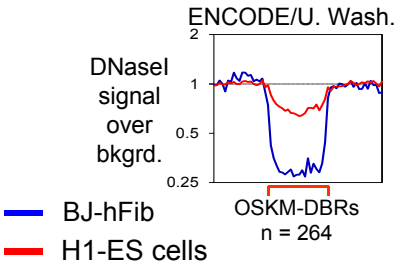
A. Genes within OSKM-DBRs remain silent early in reprogramming

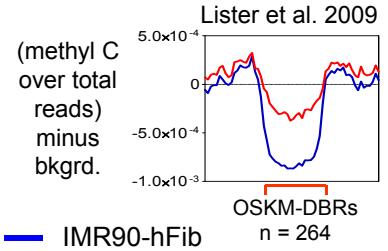


B. H3K9me3 is far more enriched than H3K9me2 at OSKM-DBRs in fibroblasts

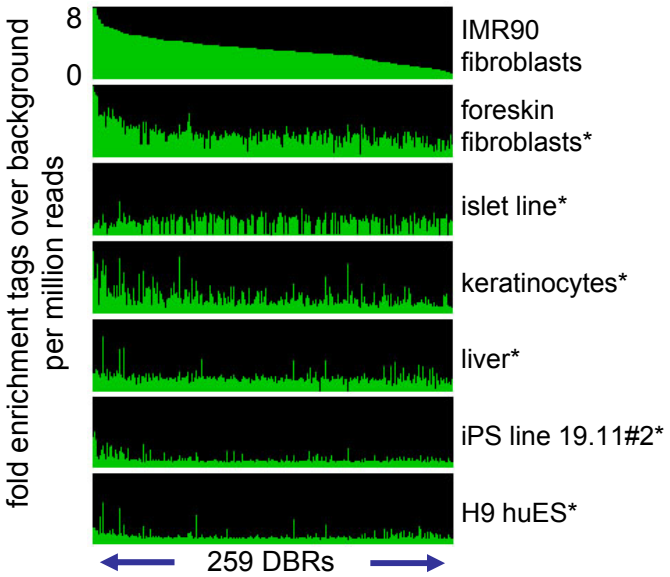


C. OSKM-DBRs are more DNase-resistant and have more demethylated DNA in fibroblasts compared to ES cells.

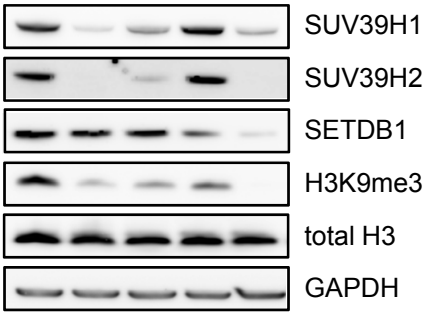




D. OSKM-DBR locations mapped in fibroblasts have varying H3K9me3 levels in different cell types

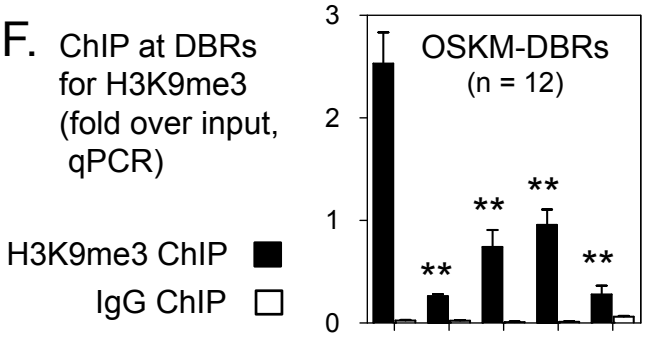


E. Western blots of knock-down for H3K9 methylases



nontargeting (NT) siRNA	+	-	-	-	-
SUV39H1/H2 siRNA #1	-	+	-	-	+
SUV39H1/H2 siRNA #2	-	-	+	-	+
SETDB1 siRNA	-	-	-	+	+

F. ChIP at DBRs for H3K9me3 (fold over input, qPCR)



** = p<0.01 vs NT siRNA