# The Binding Profiles of Key Transcriptional Factors on Long Non-coding RNAs during Somatic Cell Reprogramming

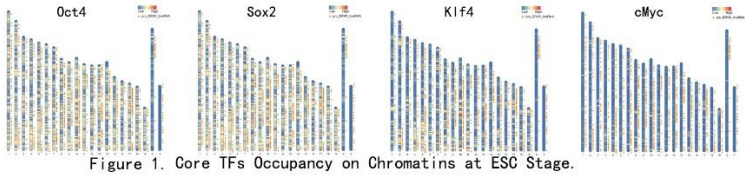


Figure 1. Core TFs Occupancy on Chromatins at ESC Stage.

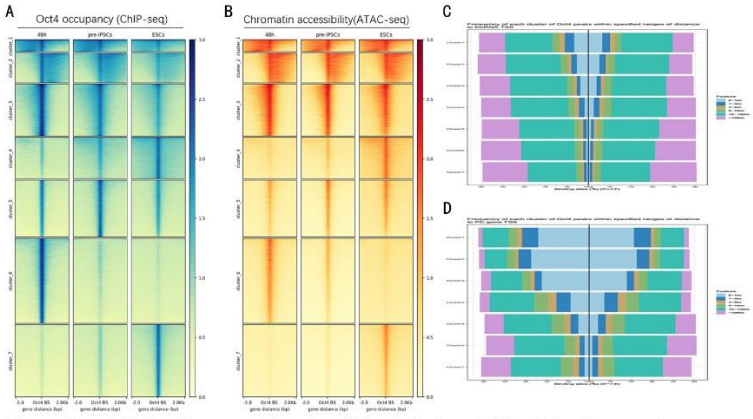


Figure 2. Oct4-Binding Patterns and Dynamics of Chromatin Accessibility during Reprogramming

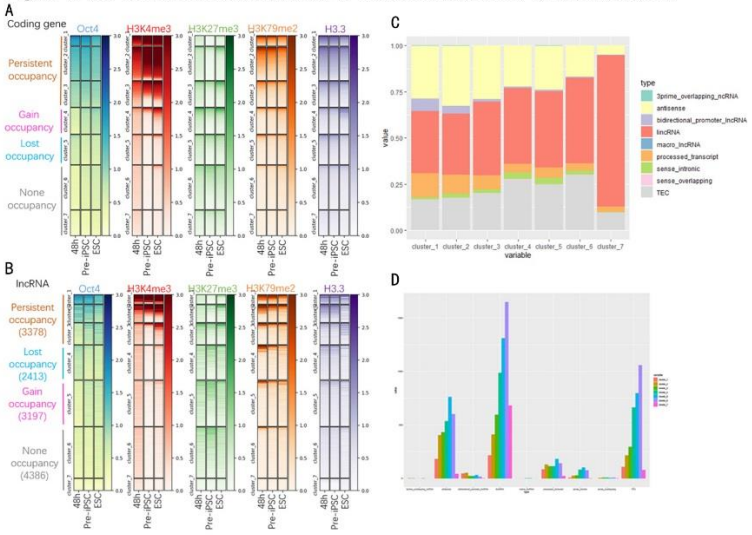


Fig. 3. Dynamics of Oct4 Binding and Chromatin Remodeling in Promoter Regions of Coding Genes and LncRNA Genes。

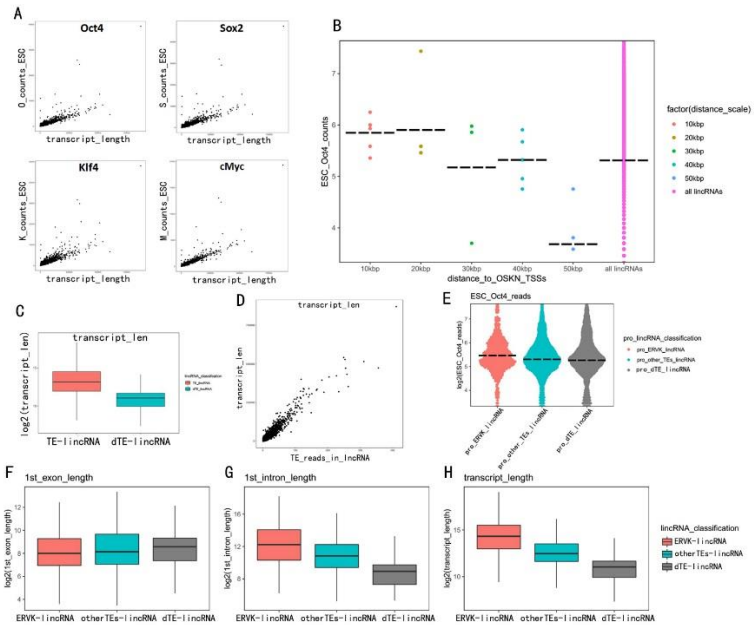


Fig. 4. Relationship between TFs binding signals and types of lncRNA, transcript length, distance to pluripotent genes TSSs, and the content of TEs.

Somatic cells can be reverted to a pluripotent stem cell state by forced expression of Yamanaka factors Oct4, Sox2, Klf4, and c-Myc (OSKM). Understanding the underlying mechanisms driving this process is crucial. Protein coding regions account for only 2% of the human genome, however, 75% are non-coding transcripts. LncRNAs are long non-coding RNAs (>200 nt) that play important roles in various biological processes, but their roles in somatic cell reprogramming remain largely unexplored.

In this study, by using bioinformatics methods, I computed the core TFs OSKM counts on 13,002 lncRNAs at three time points of mouse somatic cell reprogramming based on ChIP-seq data. For the distributions of OSKM-binding on different chromosomes, the results showed that OSKM was more inclined to bind on autosomes (Fig. 1). To investigate patterns of Oct4 binding in the reprogramming process, I clustered Oct4 peaks on genome by K-means and obtained seven Oct4-binding clusters (Fig. 2A). Moreover, the Oct4-binding events correlated with chromatin accessibility (Fig. 2B). The distance of Oct4 peaks to coding-gene TSSs (Fig. 2C) was larger than that to lncRNA TSSs (Fig. 2D), and persistent occupied Oct4 peaks were most in ±2kb from TSSs. Next, I questioned how Oct4 binding in promoter regions regulates gene expression in coordination with histone modification. The results revealed that no matter on lncRNAs or coding-genes, Oct4 dynamic binding and histone modification in promoter regions were well coordinated to regulate gene expression (Fig. 3A, B). And I found that 30% of Oct4-“Persistent Occupancy” lncRNAs are antisense lncRNAs, whereas almost 80% of Oct4-“None Occupancy” lncRNAs were lincRNAs, and bidirectional promoter lncRNAs were enriched in Oct4-“Persistent Occupancy” lncRNAs (Fig. 3C, D). In addition, I explored the relationship between TFs binding signals and types of lncRNA, transcript length, distance to pluripotent genes TSSs, and the content of TEs (Fig. 4). Moreover, I built coding-lncRNAs genes co-expression network using WGCNA, identified some lncRNAs which are important in pluripotency maintainence.