Cell

# Control of Cell Identity Genes Occurs in Insulated Neighborhoods in Mammalian Chromosomes

Jill M. Dowen,[1,6] Zi Peng Fan,[1,2,6] Denes Hnisz,[1,6] Gang Ren,[3,4,6] Brian J. Abraham,[1] Lyndon N. Zhang,[1,5] Abraham S. Weintraub,[1,5] Jurian Schuijers,[1] Tong Ihn Lee,[1] Keji Zhao,[3,*] and Richard A. Young[1,5,*]
[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA
[2]Computational and Systems Biology Program, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Systems Biology Center, NHLBI, National Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20892, USA
[4]College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, P.R. China, 712100
[5]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[6]Co-first author
*Correspondence: zhaok@nhlbi.nih.gov (K.Z.), young@wi.mit.edu (R.A.Y.)
http://dx.doi.org/10.1016/j.cell.2014.09.030

## SUMMARY

The pluripotent state of embryonic stem cells (ESCs) is produced by active transcription of genes that control cell identity and repression of genes encoding lineage-specifying developmental regulators. Here, we use ESC cohesin ChIA-PET data to identify the local chromosomal structures at both active and repressed genes across the genome. The results produce a map of enhancer-promoter interactions and reveal that super-enhancer-driven genes generally occur within chromosome structures that are formed by the looping of two interacting CTCF sites co-occupied by cohesin. These looped structures form insulated neighborhoods whose integrity is important for proper expression of local genes. We also find that repressed genes encoding lineage-specifying developmental regulators occur within insulated neighborhoods. These results provide insights into the relationship between transcriptional control of cell identity genes and control of local chromosome structure.

## INTRODUCTION

Embryonic stem cells depend on active transcription of genes that play prominent roles in pluripotency (ES cell identity genes) and on repression of genes encoding lineage-specifying developmental regulators (Ng and Surani, 2011; Orkin and Hochedlinger, 2011; Young, 2011). The master transcription factors (TFs) OCT4, SOX2, and NANOG (OSN) form super-enhancers at most cell identity genes, including those encoding the master TFs themselves; these super-enhancers contain exceptional levels of transcription apparatus and drive high-level expression of associated genes (Hnisz et al., 2013; Whyte et al., 2013). Maintenance of the pluripotent ESC state also requires that genes encoding lineage-specifying developmental regulators remain repressed, as expression of these genes can stimulate differentiation and thus loss of ESC identity. These repressed lineage-specifying genes are occupied by polycomb group proteins in ESCs (Boyer et al., 2006; Lee et al., 2006; Margueron and Reinberg, 2011; Squazzo et al., 2006). The ability to express or repress these key genes in a precise and sustainable fashion is thus essential to maintaining ESC identity.

Recent pioneering studies of mammalian chromosome structure have suggested that they are organized into a hierarchy of units, which include topologically associating domains (TADs) and gene loops (Figure 1A) (Dixon et al., 2012; Filippova et al., 2014; Gibcus and Dekker, 2013; Naumova et al., 2013; Nora et al., 2012). TADs, also known as topological domains, are defined by DNA-DNA interaction frequencies, and their boundaries are regions across which relatively few DNA-DNA interactions occur (Dixon et al., 2012; Nora et al., 2012). TADs average 0.8 Mb, contain approximately seven protein-coding genes, and have boundaries that are shared by the different cell types of an organism (Dixon et al., 2012; Smallwood and Ren, 2013). The expression of genes within a TAD is somewhat correlated, and thus some TADs tend to have active genes and others tend to have repressed genes (Cavalli and Misteli, 2013; Gibcus and Dekker, 2013; Nora et al., 2012).

Gene loops and other structures within TADs are thought to reflect the activities of transcription factors (TFs), cohesin, and CTCF (Baranello et al., 2014; Gorkin et al., 2014; Phillips-Cremins et al., 2013; Seitan et al., 2013; Zuin et al., 2014). The structures within TADs include cohesin-associated enhancer-promoter loops that are produced when enhancer-bound TFs bind cofactors such as Mediator that, in turn, bind RNA polymerase II at promoter sites (Lee and Young, 2013; Lelli et al., 2012; Roeder, 2005; Spitz and Furlong, 2012). The cohesin-loading factor NIPBL binds Mediator and loads cohesin at these enhancer-promoter loops (Kagey et al., 2010). Cohesin also becomes associated with CTCF-bound regions of the genome, and some of these cohesin-associated CTCF sites facilitate gene activation while others may function as insulators (Dixon et al.,

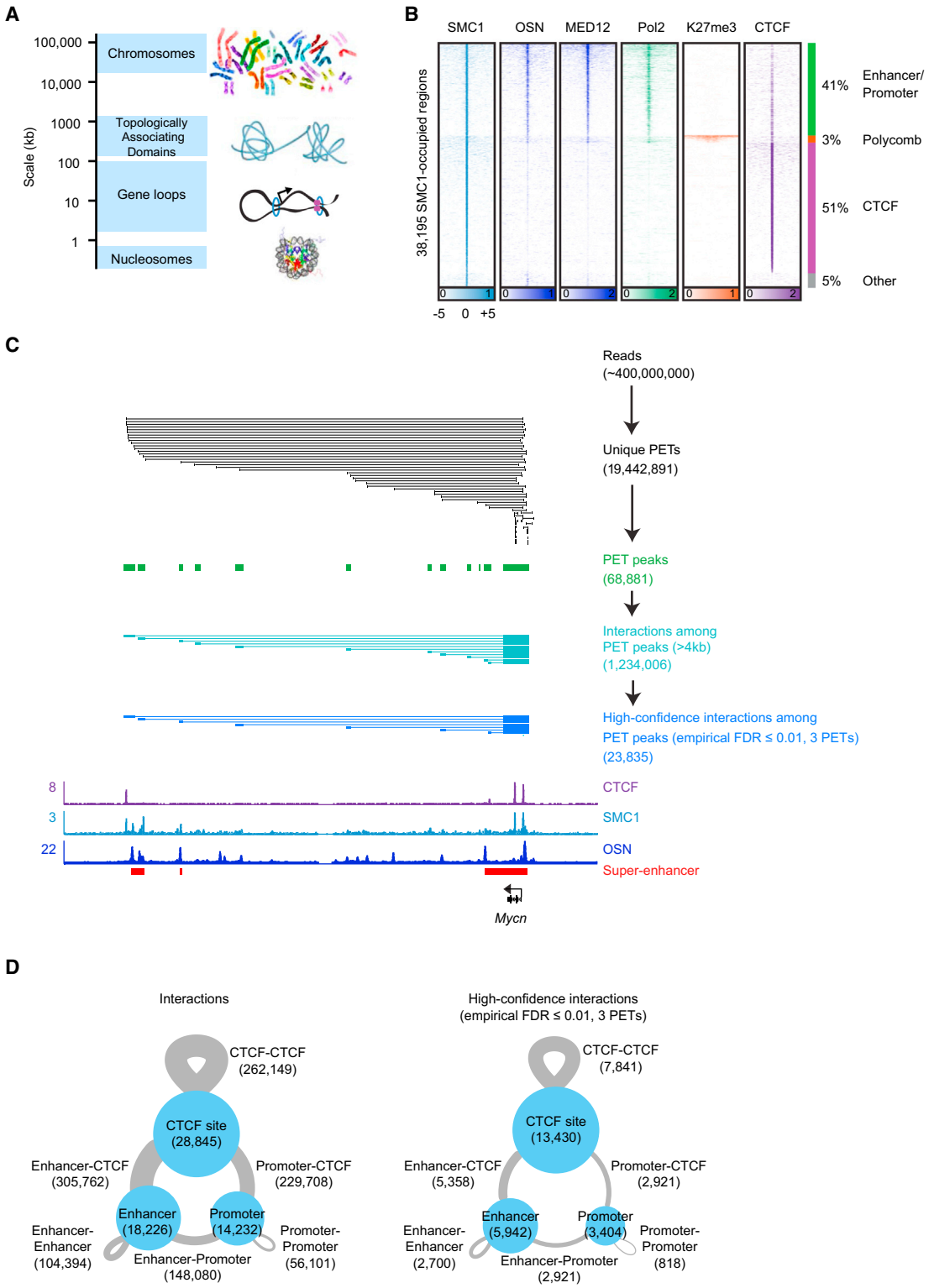


**Figure 1. DNA Interactions Involving Cohesin**

(A) Units of chromosome organization. Chromosomes consist of multiple topologically associating domains (TADs). TADs (image adapted from Dixon et al., 2012) contain multiple genes with DNA loops involving interactions between enhancers, promoters, and other regulatory elements, which are mediated by cohesin (blue ring) and CTCF (purple balls). Nucleosomes represent the smallest unit of chromosome organization.

*(legend continued on next page)*

Download English Version:

https://daneshyari.com/en/article/2035493

Download Persian Version:

https://daneshyari.com/article/2035493

Daneshyari.com