Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Cell Biology

Unraveling architecture of the pluripotent genome

Jennifer E Phillips-Cremins

Genomes are folded into sophisticated configurations that both shape, and are shaped by, a diverse range of nuclear functions. High-throughput variations of Chromosome-Conformation-Capture-based technologies now enable analysis of architecture at unprecedented resolution and scale. Here I discuss the complex structure–function relationship of the mammalian genome using the model system of embryonic stem cells, and the progression from pluripotency to terminal differentiation and back again

Addresses
Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19103, USA

Corresponding author: Phillips-Cremins, Jennifer E (jcremins@seas.upenn.edu)

Current Opinion in Cell Biology 2014, **28**:96–104

This review comes from a themed issue on **Cell nucleus**

Edited by **Gary H Karpen** and **Michael P Rout**

http://dx.doi.org/10.1016/j.ceb.2014.04.006

Introduction

A fundamental question in genome biology is related to the molecular mechanism(s) by which phenotypic diversity is established and maintained during development. Cellular phenotypes are a function of the lineage-specific genes that are expressed. However, how these gene expression signatures are intricately regulated in space and time as cells progress through development remains a great mystery.

One leading hypothesis is that lineage commitment is regulated by epigenetic mechanisms. Indeed, genome-wide profiling has revealed that hundreds of chromatin modifications and transcription factors across the genome undergo marked changes as cells transition from the totipotent zygote, to pluripotent embryonic stem cells and ultimately terminal differentiation. These potentially heritable modifications function together in a combinatorial manner to create the 'epigenetic code' that ultimately governs cell type specific transcriptional signatures unique to each cell type.

Higher-order folding of chromatin in the 3-D nucleus is one poorly understood feature of the epigenome that could be important for cellular function. In this review, I discuss recent reports that shed light on the role for chromatin architecture in the establishment and maintenance of pluripotency, differentiation and somatic cell reprogramming. Data are consistent with a model in which pluripotent genomes are folded in a nested hierarchy, with different organizing principles and functional outcomes applying at each length scale.

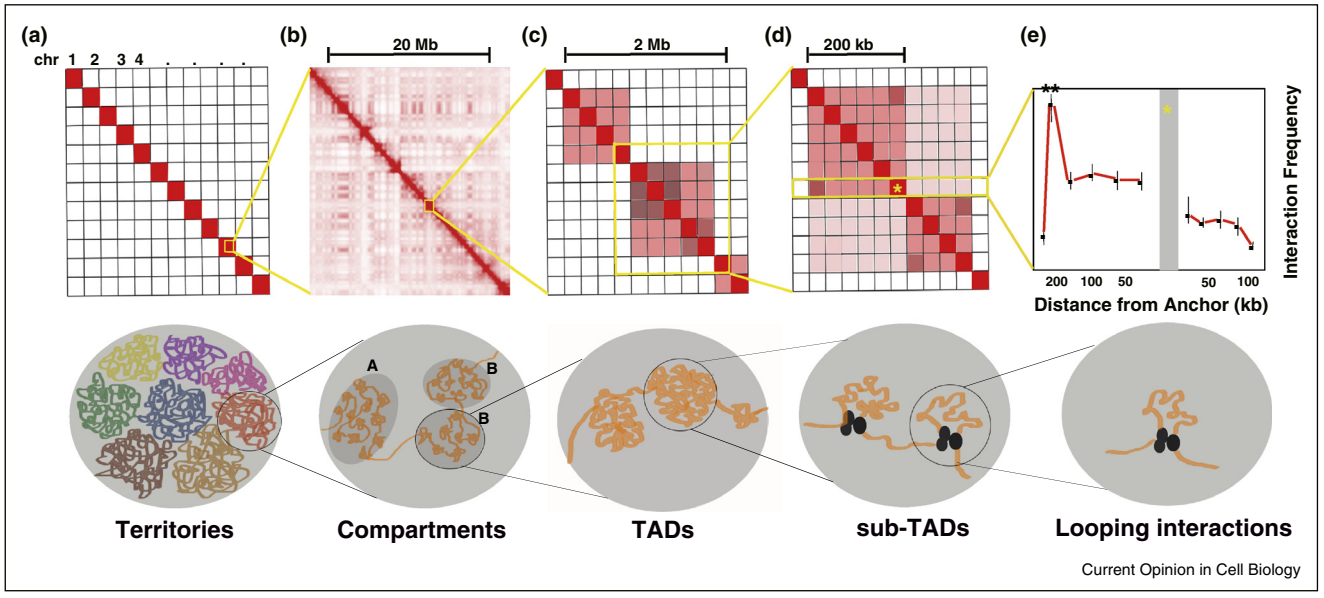
Hierarchical organization of the pluripotent genome

Seminal studies have long established that chromatin is non-randomly organized in eukaryotic nuclei (reviewed in [1]). At the largest length scale, individual chromosomes occupy distinct spatial territories with respect to each other during interphase [2]. Moreover, within territories, at the smallest length scale, long-range intra-chromosomal contacts bring distant regulatory elements in close spatial proximity to their target promoters [3]. In between these two dramatically different scales, however, chromatin's precise architectural features have been poorly understood due to the limitations in throughput and resolution of conventional microscopy techniques.

A series of molecular techniques based on the principles of Chromosome-Conformation-Capture 3C have recently enabled the query of higher-order genome architecture at unprecedented resolution and scale [4–6]. 3C-based technologies broadly fall into two categories: those that utilize an immunoprecipitation step to enrich only for chromatin contacts mediated by a specific protein (i.e. ChIA-PET, ChIP-loop) and those that more generally assay chromatin architecture in a protein-independent manner (i.e. 3C, 4C, 5C, Hi-C). Although each technique offers a unique combination of coverage and resolution, all are based on an experimental paradigm in which physical interactions between chromatin fragments are detected through proximity ligation (reviewed in detail [7]).

Hi-C enables the investigation of chromatin structure in a comprehensive, genome-wide manner without bias toward any particular type of genomic element [8,9]. Initial (Mb-resolution) Hi-C maps confirmed the presence of chromosome territories and also demonstrated that shorter chromosomes are more likely to interact with each other and not with longer chromosomes (Figure 1a). Within chromosomes, Hi-C maps also revealed the compartmentalization of the somatic cell genome into regions of open and closed chromatin, termed 'A' and 'B' compartments, respectively [8,9]. 'A' compartments contain genomic loci that are generally gene rich, transcriptionally

Figure 1



Hierarchical organization of the pluripotent genome. **(a)** Chromosome territories. **(b)** 'A' and 'B' compartments within a chromosome territory. **(c)** Groups of TADs within A/B compartments. (d) sub-TADs within TADs. **(e)** Intrachromosomal 'looping' interactions within and between sub-TADs and TADs. Heatmaps illustrate chromatin interaction data, with signal ranging from lowest (white) to highest (dark red).

active and DNase I hypersensitive; whereas loci found in 'B' compartments are relatively gene poor, transcriptionally silent and insensitive to DNase I (Figure 1b). Compartments were later confirmed in ES cells with Hi-C maps at higher (40 kb) resolution [10••].

Two independent groups employed high-resolution 5C and Hi-C techniques to discover a previously unidentified unit of genome organization termed Topologically Associating Domains or TADs [10••,11••]. TADs are Mb-sized chromatin modules in which genomic fragments have a higher tendency to interact with each other than with loci outside of the organizational unit. These domains span >90% of mammalian genomes and are markedly smaller than A/B compartments (median size 800 kb (TADs) vs. 3 Mb (compartments)). Thus, several TADs are typically found within a compartment, but TADs can also span compartment boundaries (Figure 1c).

Within TADs, sub-TADs and intra-chromosomal 'looping' interactions represent additional pervasive features of genome organization at the sub-Mb scale [12,13••,14••,15]. A recent report used 5C to explore chromatin architecture at single-fragment resolution in pluripotent ES cells and ES-derived neural progenitor cells (NPCs) [13••]. The ultra-high 'single fragment' resolution and medium scale coverage achieved by 5C [16] enabled the observation of an additional nested hierarchy of smaller sub-domains (termed sub-TADs) nested within a single Mb-sized TAD (Figure 1d).

Long-range looping interactions have also been comprehensively identified across the genome within TADs [14••]. Although the large majority of intra-chromosomal interactions occur within TADs, inter-TAD loops have also been identified (Figure 1**e**).

Structure linked to function during lineage commitment and reprogramming

Recent studies have shed new light into the relationship between chromatin structure and cellular function. An emerging theme is that architecture has vastly different dynamics and functional outcomes at each level in the hierarchy in response to different developmental signals.

'A' compartments might reflect transcription factories

A notable proportion of A/B compartments exhibit differences between ES cells and differentiated cells [10••,17•], suggesting that they might reflect an unknown functional significance in the establishment and/or maintenance of cellular phenotype(s). Insight into this functional significance was obtained in studies employing 4C, a genome-scale variant of 3C that focuses on all interactions genome-wide from a single viewpoint. In erythroid cells, 4C analysis identified 3-D contacts between active β-globin genes and numerous other active genes. These 3-D gene clusters co-localize at subnuclear foci consisting of concentrated nodules of hyperphosphorylated RNA Pol II (termed 'transcription factories') [18]. In an independent study in ES cells, the Nanog gene exhibited numerous long-range (>5 Mb) intra-chromosomal and inter-chromosomal