Genomic Characterization of the Mouse Ribosomal DNA Locus

Gabriel E. Zentner,*,1 Stephanie A. Balow,* and Peter C. Scacheri*,†,2
*Department of Genetics and Genome Sciences and †Comprehensive Cancer Center, Case Western Reserve University, Cleveland, Ohio 44106

ABSTRACT The transcription of rRNA is critical to all living cells and is tightly controlled at the level of chromatin structure. Although the widespread adoption of genomic technologies including chromatin immunoprecipitation with massively parallel short-read sequencing (ChIP-seq) has allowed for the interrogation of chromatin structure on a genome-wide scale, until recently rDNA has not been analyzed by this technique. We extended genomic analysis of rDNA to mouse (*Mus musculus*), in which rDNA is similar in structure but highly divergent in sequence compared with human rDNA. Comparison of rDNA histone marks between mouse embryonic stem cells (mESCs) and more differentiated mouse cell types revealed differences between pluripotent and differentiated states. We also observed substantial divergence in rDNA histone modification patterns between mESCs and human embryonic stem cells (hESCs). Surprisingly, we found that the pluripotency factor OCT4 was bound to rDNA in similar patterns in mESCs and hESCs. Extending this analysis, we found that an additional 17 pluripotency-associated factors were bound to rDNA in mESCs, suggesting novel modes of rDNA regulation in pluripotent cells. Taken together, our results provide a detailed view of rDNA chromatin structure in an important model system and enable high-resolution comparison of rDNA regulation between mouse and human.

KEYWORDS
rDNA
rRNA
ChIP-seq
Oct4
polycomb

rRNA is central to cellular life. Its transcription accounts for more than half of all RNA synthesis in a growing cell (Moss *et al.* 2007) and is rate-limiting for ribosome biogenesis, thereby controlling protein synthesis and downstream processes, including cell proliferation. Disruption of rRNA synthesis and subsequent ribosome biogenesis are linked to pleiotropic growth defects in model organisms from yeast to mouse (Kongsuwan *et al.* 1985; Oliver *et al.* 2004; Deutschbauer *et al.* 2005; Azuma *et al.* 2006; Uechi *et al.* 2006; Danilova *et al.* 2008; Iwanami *et al.* 2008; Chakraborty *et al.* 2009) as well as to congenital anomaly syndromes and cancer in humans (Ridanpää *et al.* 2001; Ruggero and Pandolfi 2003; Ebert *et al.* 2008; Trainor *et al.* 2008; Zentner *et al.* 2010; Dauwerse *et al.* 2011).

In a given mammalian cell, several hundred copies of a single rDNA repeat are present (Prokopowich *et al.* 2003). The mammalian rDNA repeat (~43 kb in human, ~45.3 kb in mouse) is divided into two major portions: the coding region and intergenic spacer (IGS). The coding region, ~13–14 kb in length, contains the sequences of the 18S, 5.8S, and 28S rRNA species as well as several noncoding transcribed spacer sequences. The IGS contains a large number of simple repeats, LINEs, SINEs, and ALU elements, and also harbors an enhancer, spacer promoter, and the core promoter of the adjoining rDNA repeat (Gonzalez and Sylvester 1995; Grozdanov *et al.* 2003). Within a given cell, only a fraction of the rDNA repeats are transcriptionally active and display a euchromatic chromatin structure characterized by histone modifications associated with transcriptional activity (*i.e.*, H3K4me, H3K9ac) and low levels of CpG methylation. The remainder of the repeats adopt a heterochromatic structure containing histone modifications associated with transcriptional repression (*i.e.*, H3K9me, H4K20me) and hypermethylation of CpG dinucleotides (McStay and Grummt 2008). RNA polymerase I (Pol I) transcribes the coding region of each active rDNA repeat into a precursor rRNA (pre-rRNA) transcript containing the 18S, 5.8S, and 28S rRNA sequences as well as the intervening noncoding spacers (Russell and Zomerdijk 2005). The pre-rRNA is subjected to a series of cleavage and chemical modification steps to yield the mature 18S, 5.8S, and 28S rRNA molecules, which are then assembled into ribosomes (Henras *et al.* 2008).

Copyright © 2014 Zentner *et al.*
doi: 10.1534/g3.113.009290
Manuscript received November 4, 2013; accepted for publication December 1, 2013; published Early Online December 17, 2013.
This is an open-access article distributed under the terms of the Creative Commons Attribution Unported License (http://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Supporting information is available online at http://www.g3journal.org/lookup/suppl/doi:10.1534/g3.113.009290/-/DC1.

[1] Present address: Basic Sciences Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, WA 98109.

[2] Corresponding author: BRB 627, Case Western Reserve University, 2109 Adelbert Road, Cleveland, OH 44106. E-mail: pxs183@case.edu

The analysis of rDNA chromatin structure by genomic approaches has been problematic because rDNA is not included in current genome assemblies and sequencing reads corresponding to rDNA are generally discarded during analysis. To facilitate genomic analysis of rDNA, we previously constructed a build of the human genome containing a single copy of rDNA to which we aligned short sequence reads from ChIP-seq and other genomic technologies (Zentner *et al.* 2011a). Using this approach, we described several findings of interest, including previously unknown regions of histone modification within rDNA and the association of the insulator-binding protein CTCF with rDNA. With this approach, active and silent rDNA repeats are sampled together in each ChIP; therefore, signals at rDNA represent an aggregate of signals from all immunoprecipitated repeats. This method could, in principle, be applied to any species for which a genome sequence and sequenced rDNA repeat are available.

Although the structure of the mouse and human rDNA repeats are quite similar, their nucleotide sequences are highly divergent (Gonzalez and Sylvester 1995; Grozdanov *et al.* 2003), potentially suggesting different modes of regulation. To explore this possibility, we extended our previously described method of aligning high-throughput sequencing data to a genome build containing rDNA to the mouse. Using previously generated ChIP-seq data, we analyzed the distribution of histone modifications at rDNA in mouse embryonic stem cells (mESCs), mESC-derived neural precursor cells (mNPCs), and mouse embryonic fibroblasts (MEFs). We found that patterns of rDNA histone modifications in mESCs showed differences from those in mNPCs and MEFs, as well as human embryonic stem cells (hESCs). Strikingly, we found the pluripotency factor OCT4 associated with rDNA in mESCs and hESCs. Extending this analysis further, we observed rDNA association of an additional 14 pluripotency factors as well as three Polycomb proteins in mESCs, suggesting previously unsuspected mechanisms of rDNA regulation. Our results provide insight into chromatin-level regulation of rDNA in an important model organism and allow for comparison of rDNA regulation between human and mouse.

MATERIALS AND METHODS

Datasets

The following datasets were obtained from the SRA: mESC H3K4me1, and H3K4me2 (SRP000230) (Meissner *et al.* 2008); H3K4me3, H3K9me3, H3K27me3, H3K36me3, H4K20me3, and input (SRP000415) (Mikkelsen *et al.* 2007); OCT4, SOX2, NANOG, SMAD1, STAT3, KLF4, c-MYC, n-MYC, and ZFX (SRP000217) (Chen *et al.* 2008); CHD7 and P300 (SRP0002695) (Schnetz *et al.* 2010); BRG1 (SRX003888) (Ho *et al.* 2009); EZH2, SUZ12, and RING1B (SRP000711) (Ku *et al.* 2008); nGFP (SRX207161) (Yamaji *et al.* 2013); CDX2 (SRX012415) (Nishiyama *et al.* 2009); SOX17 (SRX214076) (Aksoy *et al.* 2013); TBX3 (SRP001585) (Han *et al.* 2010); ZC3H11A (SRX188830) and RNA-seq (SRX019275) (Guttman *et al.* 2010); mNPC H3K4me1 and H3K4me2 (SRP000230) (Meissner *et al.* 2008); H3K4me3, H3K9me3, H3K27me3, H3K36me3, and input (SRP000415) (Mikkelsen *et al.* 2007); MEF H3K4me1 (SRX085451), H3K4me3, H3K9me3, H3K27me3, H3K36me3, and input (SRP000415) (Mikkelsen *et al.* 2007); and hESC OCT4 (SRP002512) and input (SRP003670) (Rada-Iglesias *et al.* 2011).

Alignment and analysis of sequencing data

We used our previously described framework for the alignment of sequencing tags to a build of the MM8 genome assembly containing a single mouse rDNA repeat (GenBank accession no. BK000964) added to chromosome 12 (MM8_plus_rDNA) (Zentner *et al.* 2011a). Briefly, unique reads were aligned to the MM8_plus_rDNA genome assembly with Bowtie (Langmead *et al.* 2009), allowing two or fewer mismatches per read and discarding reads with more than one reportable alignment. Peaks were detected with F-seq (Boyle *et al.* 2008) using a fragment size of 200 bp and visualized in R after subtracting input signal at each base. We compared the intensity of ChIP-seq signals at rDNA to those along chromosome 12 via normalized tag density, which was calculated by dividing the mean intensity of all peaks called within rDNA by the mean intensity of all peaks called outside of rDNA on chromosome 12. For correlation analyses, the rDNA locus was divided into 100-bp bins and the median signal in each bin was determined. Least-squares regression was performed using the lm function of R and the associated F-test *P* value was reported. Heatmaps were generated using the gplots R package. Two hESC OCT4 ChIP-seq technical replicates were concatenated, aligned to HG18_plus_rDNA, and analyzed as described (Zentner *et al.* 2011a).

The mappability of mouse rDNA was assessed using bias elimination algorithm for deep sequencing (BEADS) (Cheung *et al.* 2011). Mouse rDNA was divided into 36-bp fragments at 1-bp intervals and fragments were aligned to MM8_plus_rDNA as described. Mapped reads were extended to 200 bp in the 5′ and 3′ directions to simulate the estimated 200-bp fragment size. Extended fragments were concatenated and mappability was defined as the number of fragments overlapping each base. Because reads may be positive-stranded or negative-stranded, the maximum number of 200-bp fragments that can overlap a given genomic position is 400 (100% mappability). The mappability value for each base of mouse rDNA is presented in Supporting Information, Table S1. Base positions with a mappability of <25% were considered to be poorly mappable because they were discarded from analysis in the original BEADS study (Cheung *et al.* 2011).

ChIP

The mESCs and mNPCs were cultured as previously described (Schnetz *et al.* 2009). ChIP was performed as described (Schmidt *et al.* 2009) from 5×10^6 to 1×10^7 crosslinked cells. PCR reactions were performed in triplicate on the ABI7300 real-time PCR system using Sybr Green chemistry (ABI). The following antibodies were used for ChIP: rabbit anti-H3K4me1 (8 μg/ChIP; Abcam #8895) and goat anti-OCT4 (10 μg/ChIP; Santa Cruz #8628). Primers used for ChIP-PCR are detailed in Table S2.

qRT-PCR

RNA was extracted from mESCs and mNPCs using TRIzol (Invitrogen), and cDNA was prepared using the High-Capacity cDNA Archive Kit (ABI). Triplicate PCR reactions using Sybr Green (ABI) on an ABI 7300 real-time thermal cycler. *GAPDH* was used for endogenous control qRT-PCR reactions. Primers used for qRT-PCR are detailed in Table S3.

RESULTS

Analysis of sequencing data

To analyze mouse rDNA by ChIP-seq, we constructed a build of the mouse genome containing a single copy of the mouse rDNA repeat. The mouse rDNA repeat was added to the proximal end of chromosome 12, on which rDNA is located endogenously, to allow comparison of rDNA signals to those on nucleoplasmic chromatin. This genome build was designated "MM8_plus_rDNA." We discarded duplicate reads and reads aligning to multiple locations in the genome to reduce false-positives. After alignment, experimental data were normalized to input genomic DNA to correct for systematic biases.

We also analyzed the mappability of mouse rDNA because there are many rDNA pseudogenes throughout the genome that might complicate analysis (Gonzalez and Sylvester 1997). Mappability is a measure of the uniqueness of a given genomic sequence based on the number of sequenced fragments uniquely alignable to that sequence. We analyzed the mappability of mouse rDNA using BEADS (Cheung *et al.* 2011), a bias-correction software suite for sequencing data. We divided the rDNA into 36-mers at 1-bp intervals and mapped these fragments to MM8_plus_rDNA as described. We then extended the mapped reads to 200 bp to match our estimated fragment size and determined the number of fragments overlapping each base. Regions of robust mappability were detected within the coding region, at the IGS from ~14 kb to 26 kb, and at the spacer promoter. Poorly mappable regions (those defined as having a mappability <25%) (Cheung *et al.* 2011) were found from ~2 kb to 5 kb of the rDNA coding region and at various points throughout the 18S coding region, as well as the IGS. A region of the IGS from ~27 kb to 42 kb showed particularly poor mappability (Figure 1A), likely a result of the high repeat content of this region (Grozdanov *et al.* 2003). Overall, 13,371 out of 45,309 (29.5%) bases in the mouse rDNA repeat displayed a mappability >25%. The coding region displayed a substantially higher overall mappability (45.3%), whereas the IGS showed generally lower mappability (22.8%).

Distribution of histone modifications at rDNA in mESCs

We first examined the distribution of seven histone modifications on rDNA (depicted schematically in Figure 1A; this schematic is included above all rDNA plots for reference) in mESCs: H3K4me1, generally associated with enhancers (Barski *et al.* 2007; Heintzman *et al.* 2007, 2009; Ernst *et al.* 2011; Rada-Iglesias *et al.* 2011; Zentner *et al.* 2011b); H3K4me2, found at promoters and enhancers (Barski *et al.* 2007; Lupien *et al.* 2008; Ernst *et al.* 2011); H3K4me3, enriched at transcription start sites (TSSs) (Barski *et al.* 2007; Heintzman *et al.* 2007; Ernst *et al.* 2011); H3K9me3 and H4K20me3, associated with heterochromatin (Mikkelsen *et al.* 2007); H3K27me3, the signature of Polycomb repression (Cao *et al.* 2002; Müller *et al.* 2002); and H3K36me3, primarily associated with the bodies of transcribed genes (Vakoc *et al.* 2006). We detected three major areas of histone modification enrichment: within the coding region, which spans 0–13.4 kb of the repeat; a broad region from ~14 kb to 26 kb; and proximal to the spacer promoter, located ~2 kb upstream of the rDNA TSS. Notably, these are regions of high mappability (Figure 1A).

Within the coding region, we detected modest enrichment of H3K4me2 and H3K4me3 at ~9 kb into the repeat. H3K27me3 was also broadly distributed over this region, potentially indicating Polycomb repression of a subset of rDNA repeats (Figure 1A). Within the broad region from ~14 kb to 26 kb, several modifications were detected. H3K4me1 was robustly enriched throughout this region, with particularly strong enrichment at ~21 kb. H3K4me2 followed a similar distribution in this region, but with less robust signal intensity, whereas H3K4me3 was virtually absent. Varying degrees of enrichment of H3K9me3, H3K36me3, and H4K20me3 were also detected in this region, with maximal signal generally seen at ~21 kb. H3K27me3 showed some regions of moderate enrichment within this region (Figure 1A). At the spacer promoter, all methylated forms of H3K4 as well as H3K9me3 were detected, reflecting the mixture of active and silent rDNA repeats sampled in these ChIP experiments (Figure 1A).

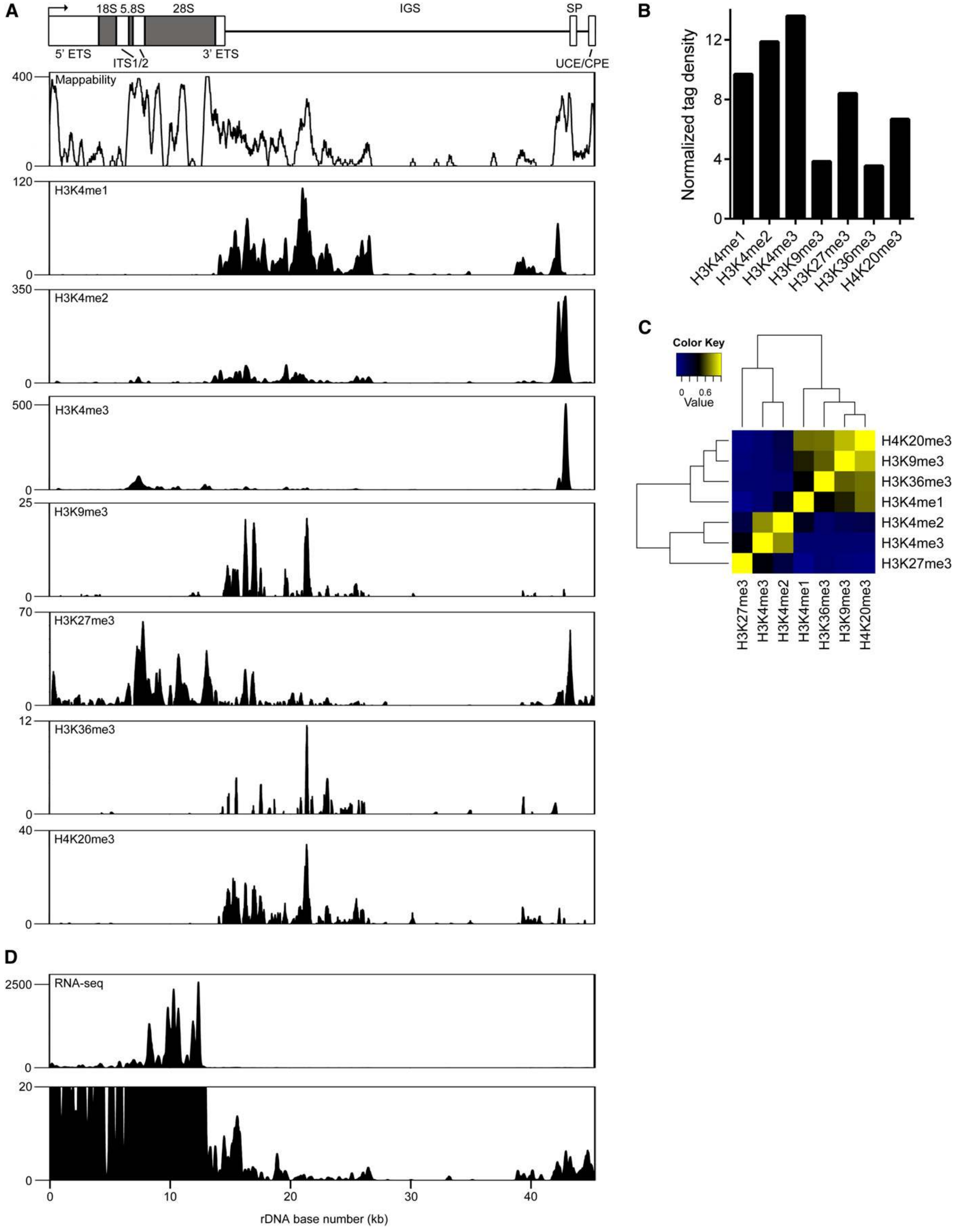
Given the high copy number of rDNA, it might be expected that signals at rDNA would appear inflated compared with those on nucleoplasmic chromatin when reads are aligned to a genome build containing a single rDNA repeat. However, many factors may influence the signal at rDNA, including the numbers of active and silent repeats, ChIP efficiency, and signal scaling at high copy sequences by peak calling algorithms. Nevertheless, we would expect rDNA signals to be at least as intense as those on nucleoplasmic chromatin. We compared the intensity of histone modification signals at rDNA to those along chromosome 12 by generating a normalized tag density score for each mark (see *Materials and Methods*). Histone modification signals were ~3.5–13.6 times as intense as those seen along chromosome 12 (Figure 1B).

We next determined the median signal for each mark across rDNA in 100-bp windows and performed pairwise correlation analysis for the entire locus. Two primary clusters were identified, one containing H3K27me3, H3K4me2, and H3K4me3 and the other containing H4K20me3, H3K9me3, H3K36me3, and H3K4me1 (Figure 1C). This is in contrast to human rDNA, in which modifications associated with transcriptional activation and repression separated into two distinct groups (Zentner *et al.* 2011a). However, within each cluster, highly correlated pairs of modifications associated with transcriptional activity or silencing were detected (H3K4me2 and H3K4me3; H3K9me3 and H4K20me3) (Figure 1C).

The coding region of active rDNA repeats is highly transcribed, containing a very high density of Pol I molecules (~1 molecule/100 bp) (Grummt 1999), and rRNA comprises much more than half of all RNA synthesized by a given cell (Moss *et al.* 2007). We therefore assessed transcription from rDNA using RNA-seq data from mESCs (Guttman *et al.* 2010). Despite using data generated from PolyA-selected RNA, we were able to detect RNA-seq signal originating from rDNA. Very high levels of RNA-seq signal were detected over the coding region of rDNA, consistent with an extremely high volume of Pol I transcription (Figure 1D). In addition to transcription of the rDNA coding region, it has been demonstrated that the rDNA spacer promoter, located ~2 kb upstream of the core promoter region, directs transcription of short noncoding RNA molecules. These ncRNAs are homologous the rDNA core promoter and have been designated promoter-associated RNA (pRNA). pRNA appears to function in the epigenetic regulation of the rDNA locus (Bierhoff *et al.* 2011). We therefore expanded the scale of our RNA-seq data to determine the presence of pRNA and any other low-abundance RNA species originating from rDNA. We observed RNA-seq signal at ~43 kb and ~45 kb into the repeat, indicating the presence of pRNA corresponding to the spacer and core promoter regions (Figure 1D). Several other regions of RNA-seq enrichment outside of the coding region were also detected, although their significance is unclear at this time.

rDNA histone modification patterns in differentiated mouse cell types

We next tried to determine if the distribution of histone modifications at rDNA differed between undifferentiated and differentiated cell types. To this end, we compared rDNA histone modification patterns between mESCs, mESC-derived mNPCs, and MEFs visually and by pairwise linear regression for the entire locus. mNPCs and MEFs lost enrichment of H3K4me1 in the ~14-kb to 26-kb region of enrichment, potentially suggesting that the activity of any functional element within this region is restricted in the mESC state (Figure 2A). This drastic reduction in H3K4me1 in mNPCs and MEFs likely contributed to the lack of correlation between its locus-wide distributions (mESC *vs.* mNPC: R^2 = 0.04, P = 2.08 × 10^{-5}; mESC *vs.* MEF: R^2 = 0.005, P = 0.271). Visually, H3K4me2 also seemed to be lost at


A
18S 5.8S
28S
IGS
SP
5' ETS
ITS1/2
3' ETS
UCE/CPE
Mappability
400
0
H3K4me1
120
0
H3K4me2
350
0
H3K4me3
500
0
H3K9me3
25
0
H3K27me3
70
0
H3K36me3
12
0
H4K20me3
40
0
B
Normalized tag density
0
4
8
12
H3K4me1
H3K4me2
H3K4me3
H3K9me3
H3K27me3
H3K36me3
H4K20me3
C
Color Key
0 0.6
Value
H4K20me3
H3K9me3
H3K36me3
H3K4me1
H3K4me2
H3K4me3
H3K27me3
H3K27me3
H3K4me3
H3K4me2
H3K4me1
H3K36me3
H3K9me3
H4K20me3
D
RNA-seq
2500
0
20
0
0
10
20
30
40
rDNA base number (kb)

this region, but the major site of enrichment, at the spacer promoter, was unaffected, and the locus-wide correlation remained robust and highly significant (mESC *vs.* mNPC: R^2 = 0.87, $P < 2.2 \times 10^{-16}$). H3K4me3 also appeared to decrease at the spacer promoter region, but its locus-wide distribution was not affected (mESC *vs.* mNPC: R^2 = 0.93, $P < 2.2 \times 10^{-16}$; mESC *vs.* MEF: R^2 = 0.94, $P < 2.2 \times 10^{-16}$). The locus-wide distributions of H3K9me3 and H3K27me3 were largely correlated between cell types, with the exception of MEF H3K27me3 (mESC *vs.* mNPC H3K9me3: R^2 = 0.61, $P < 2.2 \times 10^{-16}$; mESC *vs.* mNPC H3K27me3: R^2 = 0.58; mESC *vs.* MEF H3K9me3: R^2 = 0.66, $P < 2.2 \times 10^{-16}$; mESC *vs.* MEF H3K27me3: R^2 = 0.1, $P = 5.98 \times 10^{-12}$). The distributions of H3K36me3 were not correlated between mESCs, mNPCs, and mESCs (mESC *vs.* mNPC: $R^2 = -0.002$, P = 0.58; mESC *vs.* MEF: $R^2 = -0.002$, P = 0.69). Normalized tag densities for mNPC and MEF histone marks ranged from ~2.5 to 9.2 (Figure 2B). Using ChIP-PCR we confirmed that H3K4me1 was reduced within the IGS but maintained at the spacer promoter in mNPCs (Figure 2C).

Because H3K4me1 decreases at rDNA during the transition from the mESC to mNPC state, we wondered if rRNA expression might also decrease. To this end, we measured the expression of the pre-rRNA transcript in mESCs and mNPCs by qRT-PCR. We observed a significant decrease in pre-rRNA expression (P = 0.045 by one-tailed *t* test) in mNPCs *vs.* mESCs (Figure 2D). We speculate that the reduction in H3K4me1 within the spacer promoter in mNPCs may be linked to reduced rRNA expression, although we cannot rule out the influence of other histone modifications or cell type–specific transcription factors not assayed here.

Differences in histone modification patterns at mESC and hESC rDNA

Despite the highly similar structures of mouse and human rDNA, their sequences are highly divergent. We previously analyzed the distributions of several histone modifications on rDNA in hESCs (Zentner *et al.* 2011a). Using these data, we compared patterns of rDNA histone modification in mESCs to those in hESCs. A schematic of the distributions of five histone modifications analyzed in both cells types is presented in Figure 3A. Within the coding region of both mouse and human rDNA, substantial H3K27me3 enrichment was seen; however, the enrichment of H3K4me2 and H3K4me3 seen at ~9 kb in the mouse rDNA repeat was not observed in human rDNA and, overall, H3K27me3 had a broader distribution on human rDNA, encompassing nearly the entire repeat. H3K4me1 enrichment was seen from ~14 kb to 20 kb in the mouse and human repeats, although it extended to ~26 kb in mouse. Whereas discrete peaks of H3K4me2 and H3K4me3 were seen at ~15 kb and ~20 kb in the human repeat, these peaks were not evident in mouse rDNA, although the ~21-kb peak of H3K4me1 in mouse rDNA may correspond to this latter peak. No H3K4me3 enrichment was seen in this region of mouse rDNA. The peak of activation-associated modifications seen at ~27–28 kb in human rDNA was not found in mouse rDNA. Several regions of H3K4me1 and H3K36me3 enrichment in the human IGS were not observed in the mouse IGS. Methylation of H3K4 and H3K9me3 were detected at the spacer promoter region in human and mouse rDNA. Overall, it appears that the distributions of histone marks at rDNA are quite different between mESCs and hESCs.

Given the marked differences in rDNA histone modifications between mESCs and hESCs, we wondered if other factors known to differentially occupy the genome in mESCs and hESCs might also bind rDNA in distinct patterns. We chose to analyze ChIP-seq data for the key pluripotency factor OCT4, which has distinct genomic distributions in mESCs and hESCs (Loh *et al.* 2006) and for which there is some evidence of nucleolar localization (Table 1). Strikingly, comparison of OCT4 ChIP-seq profiles in mESCs and hESCs revealed relatively similar distributions of OCT4, with high enrichment across the coding region and proximal to the spacer promoter (Figure 3B). A notable exception is a peak of OCT4 enrichment at ~27–28 kb into the human rDNA repeat not seen in mouse rDNA. This region is also enriched for active histone modifications in hESCs (Figure 3A) (Zentner *et al.* 2011a). The overall similarity in OCT4 rDNA binding between mESCs and hESCs is surprising given that OCT4 tends to occupy distinct sites in mESCs and hESCs (Loh *et al.* 2006) and suggests that pluripotency factors may regulate rRNA expression.

Pluripotency factors associated with rDNA in mESCs

Having determined that OCT4 was associated with rDNA in mESCs and hESCs, we sought to determine if additional pluripotency-associated factors might be bound to rDNA. We obtained ChIP-seq datasets for a number of well-established components of the mESC pluripotency network, SOX2, NANOG, KLF4, STAT3, SMAD1, C-MYC, N-MYC, ZFX, E2F1, ESRRB, TCFCP2L1 (Chen *et al.* 2008), P300 (Schnetz *et al.* 2010), and BRG1 (Ho *et al.* 2009), and aligned them to MM8_plus_rDNA. We also reanalyzed CHD7 mESC ChIP-seq data (Schnetz *et al.* 2010), which we previously aligned to mouse rDNA out of the context of the reference genome (Zentner *et al.* 2010), after alignment to MM8_plus_rDNA. Strikingly, we found that all analyzed factors were associated with rDNA (Figure 4A). We generated normalized tag density scores for each factor and found that the enrichment of each protein at rDNA was more than eight-fold higher than its enrichment along chromosome 12 (Figure 4B). ChIP-PCR confirmed the association of OCT4 with rDNA (Figure S1). Notably, we also found additional experimental evidence of nucleolar localization for nearly all of the tested factors (Table 1).

To ensure that the observed binding of pluripotency factors to rDNA was not simply attributable to the highly accessible chromatin architecture and high number of active rDNA repeats, we first analyzed ChIP-seq data for nuclear GFP (nGFP) (Yamaji *et al.* 2013), because GFP has been shown to give artifactual ChIP signals at highly expressed loci (Teytelman *et al.* 2013). We noted some GFP ChIP signal at the 5′ end of the repeat (Figure S2), but closer

Figure 1 Distribution of histone modifications on rDNA in mESCs. (A) Schematic representation of the mouse rDNA repeat based on GenBank accession no. BK000964. The rDNA TSS is depicted by an arrow. ETS, external transcribed spacer; ITS, internal transcribed spacer; IGS, intergenic spacer; SP, spacer promoter; UCE, upstream control element; CPE, core promoter element. Below the schematic is a mappability track for mouse rDNA as determined by BEADS (Cheung *et al.* 2011), with the mappability score for each base plotted on the Y-axis (maximum mappability = 400). Presented below the mappability track are ChIP-seq plots showing the distributions of H3K4me1, H3K4me2, H3K4me3, H3K9me3, H3K27me3, H3K36me3, and H4K20me3 at rDNA in mESCs. Signal intensity as determined by F-seq is plotted on the Y-axis. (B) Normalized tag density scores for histone modifications at rDNA in mESCs. (C) Heatmap of correlation scores for pairwise comparisons between the median signals for each histone modification at rDNA in mESCs. (D) RNA-seq profiles across rDNA in mESCs. Signal intensity as determined by F-seq is plotted on the Y-axis. Shown below is an RNA-seq plot with a reduced Y-axis scale to display low-intensity RNA-seq signals.

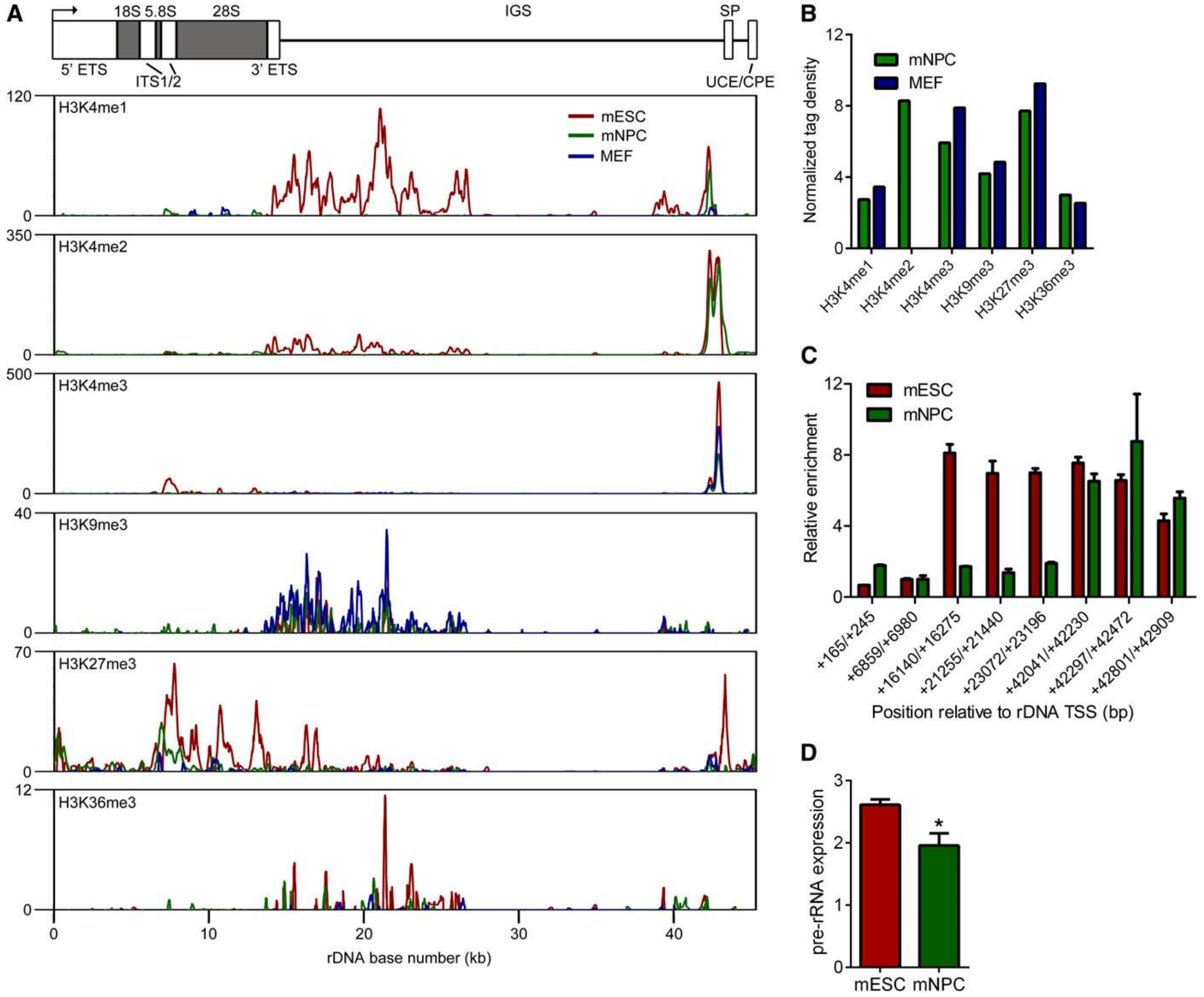


Figure 2 Comparison of rDNA histone modifications between mESCs and mNPCs. (A) ChIP-seq plots comparing the distributions of H3K4me1, H3K4me2, H3K4me3, H3K9me3, H3K27me3, and H3K36me3 at rDNA in mESCs, mNPCs, and MEFs. Signal intensity as determined by F-seq is plotted on the Y-axis. (B) Normalized tag density scores for histone modifications at rDNA in mNPCs and MEFs. (C) ChIP-PCR evaluation of H3K4me1 enrichment at rDNA in mESCs and mNPCs. Enrichment values were normalized to the +6859/+6980 negative control amplicon. Error bars represent mean ± SD for triplicates. (D) Pre-rRNA levels in mESCs and mNPCs as measured by qRT-PCR. Mean ± SEM is shown (n = 2). $^{*}P = 0.045$ by one-tailed *t* test.

inspection revealed that this region showed an irregular, jagged distribution of tags rather than the smooth gradation of tags generally seen with legitimate peak; therefore, we speculate that this is an artifact. We also analyzed ChIP-seq datasets for the CDX2, TBX3, SOX17, and ZC3H11A transcription factors (Nishiyama *et al.* 2009; Han *et al.* 2010; Aksoy *et al.* 2013). Similar to GFP, we noted irregular enrichment at the 5′ end of the repeat as well as a highly discrete region of enrichment at ~15.7 kb in each of the negative control TF datasets (Figure S2). Given that this peak is precisely overlapping between each of the negative control datasets and does not overlap with binding sites for pluripotency factors, we speculate that it is an artifact. From these analyses, we are confident that the association of pluripotency factors with rDNA is legitimate.

In the nucleoplasm, several pluripotency factors often bind to the same genomic regions, designated multiple transcription factor loci (MTLs) (Chen *et al.* 2008). Two major classes of MTLs have been described: an "OCT4-centric" module containing OCT4, SOX2, NANOG, STAT3, SMAD1, and other factors, and a "Myc-centric" module containing c-MYC, n-MYC, E2F1, and ZFX, among other proteins (Chen *et al.* 2008; Schnetz *et al.* 2010; Ng and Surani 2011). We were curious about whether such MTLs might be detected at rDNA. Therefore, we performed correlation analysis for the aforementioned analyzed factors plus Polycomb proteins (Figure 5). This analysis revealed several groups of factors (Figure 4C). One major cluster (top right, outlined in red in Figure 4) comprised two smaller groups, one containing OCT4, STAT3, SOX2, and KLF4 (outlined in blue in Figure 4) and the other containing c-MYC, n-MYC, BRG1, and ZFX (outlined in blue in Figure 4). Notably, NANOG, a component of the core pluripotency machinery, did not cluster with factors such as OCT4 and SOX2; rather, NANOG clustered with P300 and

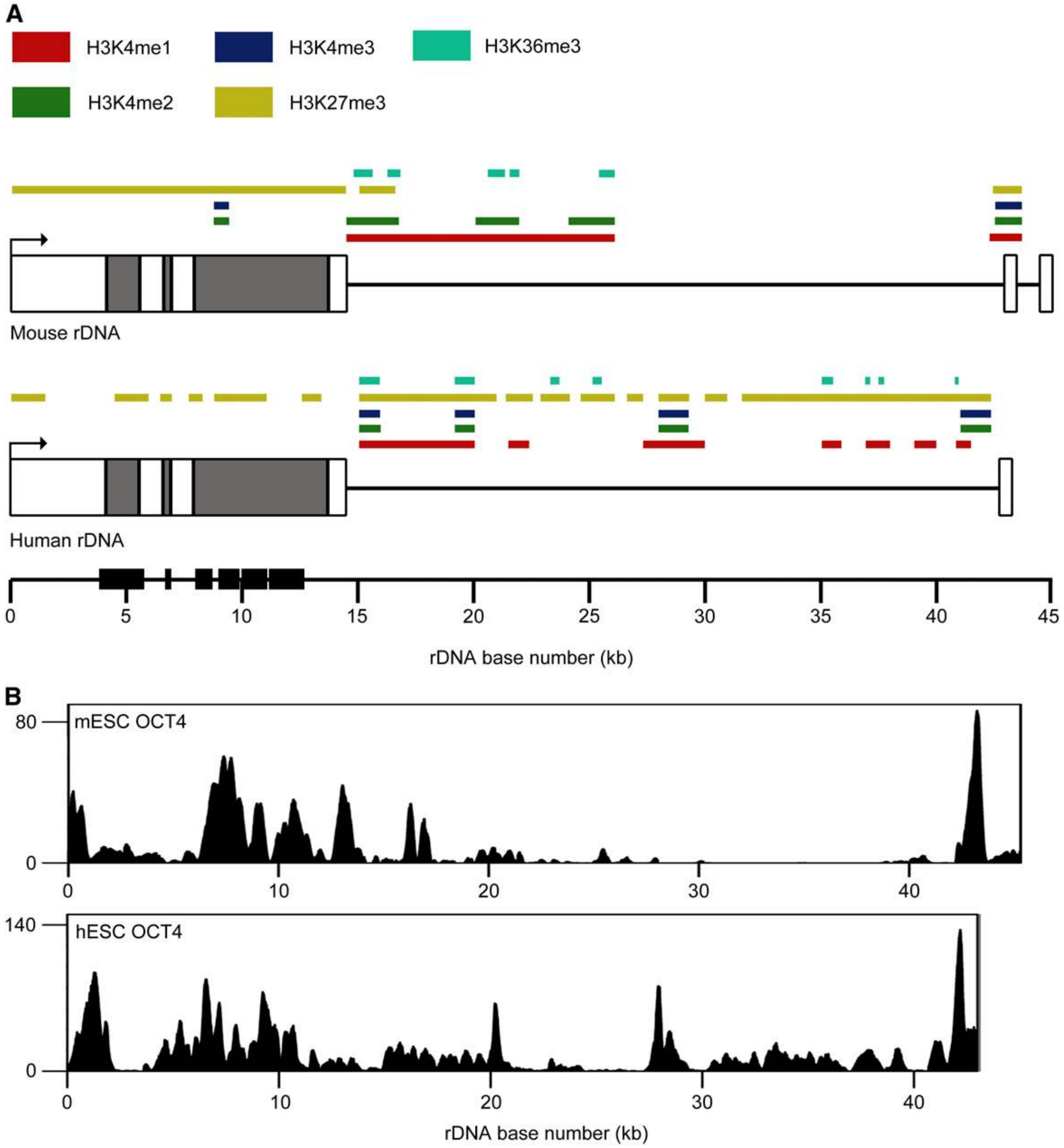


Figure 3 Comparison of rDNA histone modifications between mESCs and hESCs. (A) Schematic representations of the mouse and human rDNA repeats showing sites of histone modification enrichment in mESCs and hESCs as colored boxes. Repeats are depicted approximately to scale (mouse, ~45 kb; human, ~43 kb). Regions of high sequence identity as determined by two-sequence BLAST are depicted as black boxes on the scale bar. (B) ChIP-seq profiles of OCT4 at rDNA in mESCs and hESCs. Signal intensity as determined by F-seq is plotted on the Y-axis.

TCFCP2L1 (center, outlined in red in Figure 4). As expected, the Polycomb group proteins EZH2, SUZ12, and RING1B strongly colocalized throughout the rDNA repeat (bottom left, outlined in red in Figure 4). Although SMAD1 did not belong to any distinct cluster, it showed some degree of correlation with NANOG, TCFCP2L1, and P300, with which it shared a somewhat similar signal distributions (Figure 4A). In general, the groups of factors we detected at rDNA were quite similar to those seen on nucleoplasmic chromatin (Chen *et al.* 2008; Schnetz *et al.* 2010; Ng and Surani 2011).

Polycomb proteins bind rDNA

In this study, we have shown that mouse rDNA is marked by H3K27me3, the chromatin signature of Polycomb-mediated transcriptional repression. We also previously showed that this mark is present throughout the human rDNA repeat in hESCs (Figure 3A) (Zentner *et al.* 2011a). We therefore speculated that Polycomb proteins might be associated with rDNA. To assess this possibility, we made use of previously published mESC ChIP-seq data for the Polycomb group proteins EZH2, SUZ12, and RING1B (Ku *et al.* 2008). We detected binding of these three factors to rDNA in a pattern highly similar to that of H3K27me3 (Figure 5A) (R^2 = 0.39−0.58, $P < 2.2 \times 10^{-16}$ for all), further suggesting that Polycomb silencing is involved in the maintenance of transcriptional silencing at rDNA and indicating that the results of the H3K27me3 ChIP-seq analysis are legitimate. The normalized tag densities for Polycomb proteins were ~6.0–14.6 (Figure 5B).

Table 1 Evidence for nucleolar localization of pluripotency factors in mESCs

Protein	Evidence	Cell Types
OCT4	Nucleolar localization (Parfenov *et al.* 2003; Zuccotti *et al.* 2008); interacts with CHD4, a known regulator of rRNA synthesis (Shimono *et al.* 2005; van den Berg *et al.* 2010); interacts with nucleolar protein nucleophosmin (Johansson and Simonsson 2010)	Mouse oocytes, mESCs
SOX2	Predicted to be nucleolar based on sequence motifs (Bauer *et al.* 2011); interacts with nucleolar protein nucleophosmin (Johansson and Simonsson 2010)	mESCs
NANOG	Nucleolar localization (He *et al.* 2006); interacts with nucleolar protein nucleophosmin (Johansson and Simonsson 2010)	ICM of goat embryos, mESCs
KLF4	Not available	Not available
STAT3	Not available	Not available
SMAD1	Interacts with RUNX2, a known regulator of rRNA synthesis (Zhang *et al.* 2000; Ali *et al.* 2008)	NIH3T3, P19 cell lines
c-MYC	Nucleolar localization and regulation of rRNA transcription (Poortinga *et al.* 2004; Arabi *et al.* 2005; Grandori *et al.* 2005)	WI38, COS-7, HeLa, U2OS, Rat1, CHO, NIH-3T3, HL-60, MPRO cell lines, human primary foreskin fibroblasts
n-MYC	Predicted to be nucleolar based on sequence motifs (Bauer *et al.* 2011)	Not available
ZFX	Not available	Not available
E2F1	Nucleolar localization on ARF expression (Datta *et al.* 2002)	U2OS cell line
ESRRB	Interacts with CHD4, a known regulator of rRNA synthesis (Shimono *et al.* 2005; van den Berg *et al.* 2010)	mESCs
TCFCP2L1	Interacts with CHD4, a known regulator of rRNA synthesis (Shimono *et al.* 2005; van den Berg *et al.* 2010)	mESCs
P300	Closely related protein CBP acetylates Pol I transcription factor UBF to stimulate rRNA transcription (Pelletier *et al.* 2000); interacts with CHD7, a known regulator of rRNA synthesis (Schnetz *et al.* 2010; Zentner *et al.* 2010)	NIH3T3 cell line
BRG1	Interacts with CHD4 and CHD7, known regulators of rRNA synthesis (Shimono *et al.* 2003; Bajpai *et al.* 2010; Zentner *et al.* 2010)	HEK293 cell line, human neural crest–like cells

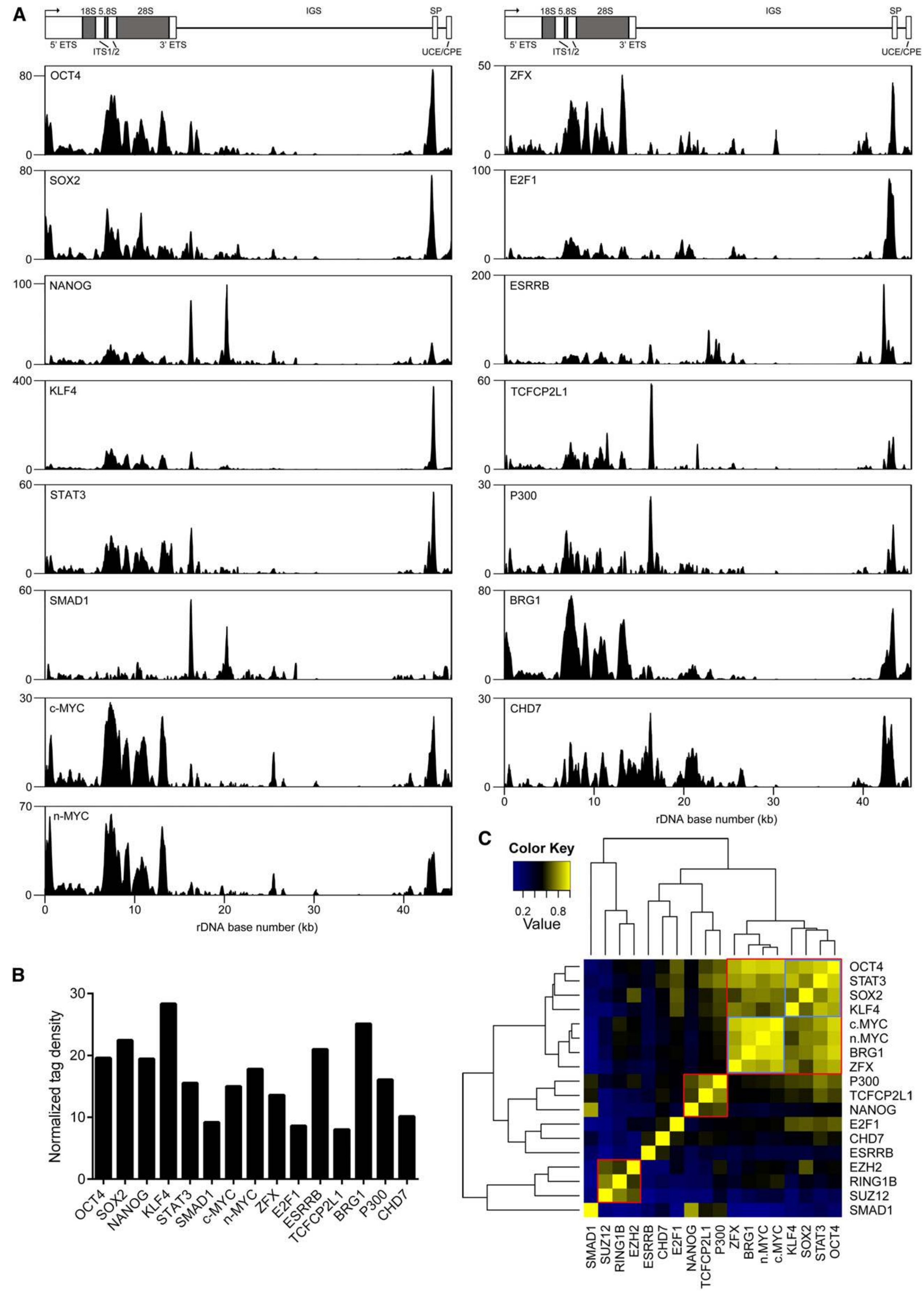
DISCUSSION

The rDNA is a critical genomic region tightly controlled at the level of chromatin structure. Here, we aligned sequencing data from ChIP-seq and other genomic technologies to a build of the mouse genome containing a single rDNA repeat. We provide several lines of evidence that the signals we observed at rDNA are legitimate. First, the ChIP-seq signals we observed fall into regions of high mappability. Second, consistent with the high copy number of rDNA, the ChIP-seq signals at rDNA were all higher than those observed on nucleoplasmic chromatin. Third, the ChIP-seq signals were validated by standard ChIP. Fourth, we failed to detect rDNA association of nGFP or several transcription factors, arguing for specific association of pluripotency factors with rDNA in mESCs. Finally, H3K27me3 enrichment at rDNA coincides with enrichment of the Polycomb proteins EZH2, SUZ12, and RING1B, which are known to contribute to deposition of the H3K27me3 mark and show colocalization with H3K27me3 on nucleoplasmic chromatin.

Our results highlight similarities and substantial differences in the chromatin-level regulation of rDNA between mouse and human. Enrichment of activation and repression-associated marks ~2 kb upstream of the core promoter region is common to both mESCs and hESCs, suggesting that chromatin-level regulation of pRNA production is conserved. The differences in histone modification distributions between mESC and hESC rDNA may reflect species-specific differences or may indicate differences in developmental stage, as previously proposed (Brons *et al.* 2007; Tesar *et al.* 2007). It is also possible that the differences in histone modification patterns seen between mouse and human rDNA could be attributable to the high degree of divergence in the genetic architecture of the repeat between the two species (Gonzalez and Sylvester 1995; Grozdanov *et al.* 2003). Regardless, we speculate that the regions of histone modification in both the mouse and human IGSs are reflective of novel functional elements.

The association of pluripotency factors, including OCT4 and SOX2, with rDNA is particularly interesting. Factors such as MYOD and RUNX2, important for myogenic and osteogenic differentiation, respectively, also associate with rDNA and downregulate rRNA expression during differentiation. Therefore, it has been suggested that a general property of lineage-specific transcription factors is regulation of rRNA transcription as a means to coordinate cell growth and phenotype (Ali *et al.* 2008). Extending this hypothesis to mESCs, it might be speculated that factors that maintain mESC identity serve in the place of lineage-specific transcription factors in regulating rRNA expression. We found that more than a dozen pluripotency-associated factors were bound to rDNA. With some exceptions (*i.e.*, the lack of correlation between NANOG and OCT4/SOX2), the clusters of factors we observed at rDNA ("rDNA MTLs") were remarkably similar to those in the nucleoplasm; for instance, OCT4 and SOX2 were strongly correlated, as were C-MYC, N-MYC, and ZFX. This suggests that pluripotency factors function to regulate transcription in similar combinations in both the nucleoplasm and nucleolus.

The association of Polycomb proteins with rDNA in mESCs is also intriguing, suggesting that Polycomb proteins silence rDNA, similar to their function in the nucleoplasm. Polycomb proteins are repressors of nucleoplasmic gene transcription (Bannister and Kouzarides 2011),


A
18S 5.8S 28S
IGS
SP
5' ETS
ITS1/2
3' ETS
UCE/CPE
OCT4
SOX2
NANOG
KLF4
STAT3
SMAD1
c-MYC
n-MYC
ZFX
E2F1
ESRRB
TCFCP2L1
P300
BRG1
CHD7
0
10
20
30
40
rDNA base number (kb)
B
Normalized tag density
0
10
20
30
OCT4
SOX2
NANOG
KLF4
STAT3
SMAD1
c-MYC
n-MYC
ZFX
E2F1
ESRRB
TCFCP2L1
BRG1
P300
CHD7
C
Color Key
0.2 0.8
Value
OCT4
STAT3
SOX2
KLF4
c.MYC
n.MYC
BRG1
ZFX
P300
TCFCP2L1
NANOG
E2F1
CHD7
ESRRB
EZH2
RING1B
SUZ12
SMAD1

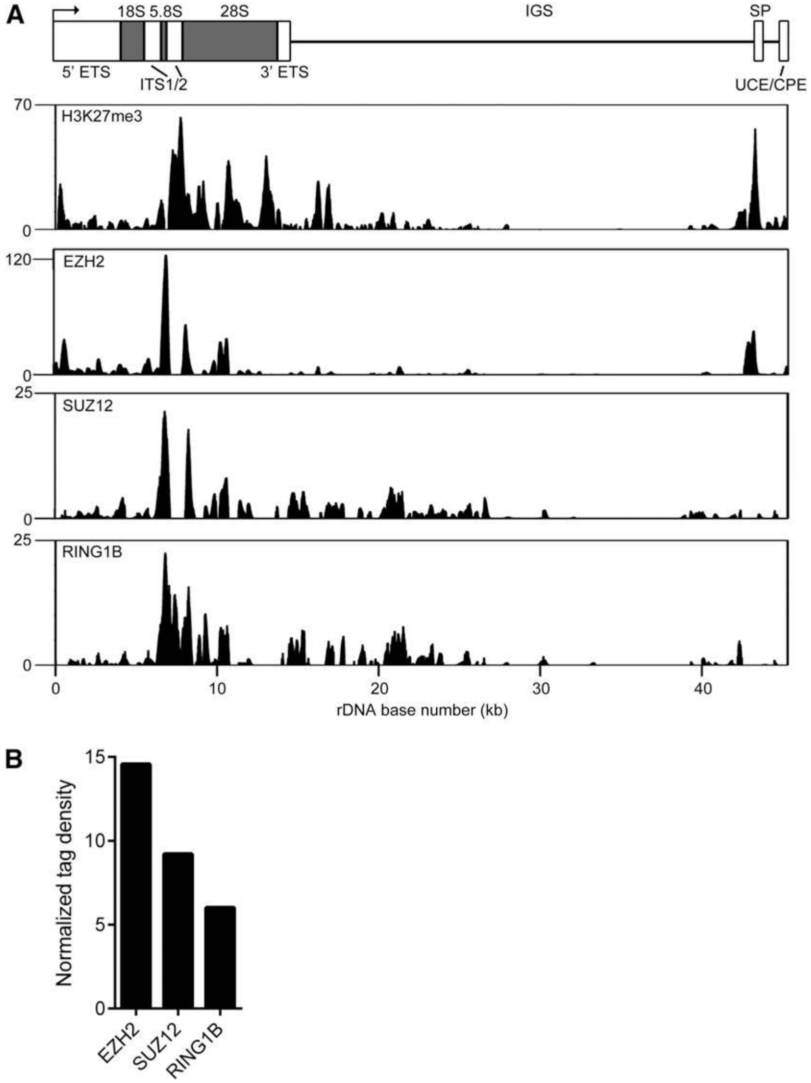


Figure 5 Polycomb proteins bind rDNA. (A) ChIP-seq profiles of H3K27me3, EZH2, SUZ12, and RING1B enrichment at rDNA in mESCs. Signal intensity as determined by F-seq is plotted on the Y-axis. (B) Normalized tag density scores for Polycomb proteins at rDNA in mESCs.

and perhaps they also participate in the repression of Pol I transcription. An additional nonmutually exclusive possibility is that Polycomb proteins may also function in the suppression of cryptic transcription of rDNA and rDNA recombination, both of which have deleterious cellular consequences in yeast and mammalian cells (Kobayashi and Ganley 2005; Gagnon-Kugler *et al.* 2009). Again, functional studies are needed to assess these possibilities.

It has been appreciated for some time that the cell uses common epigenetic mechanisms to regulate both Pol I and Pol II transcription (*i.e.*, DNA methylation, histone modifications, chromatin remodeling) (McStay and Grummt 2008). However, there is also accumulating evidence for a large class of transcription factors and chromatin remodelers that dually function in nucleoplasmic and nucleolar transcription regulation (*i.e.*, MYOD, C-MYC, CHD4/7) (Arabi *et al.* 2005; Grandori *et al.* 2005; Shimono *et al.* 2005; Ali *et al.* 2008; Zentner *et al.* 2010). This emerging theme is underscored by our finding that more than a dozen pluripotency-associated transcription factors and chromatin remodelers associate with rDNA in mESCs. Therefore, it may be speculated, despite the distinctive machineries that transcribe rRNA and nucleoplasmic genes, that the cell uses generally similar mechanisms to regulate Pol I and Pol II transcription. From a broader perspective, the similarities in nucleoplasmic and nucleolar transcriptional regulatory may be indicative of the proposed co-evolution of the nucleus and nucleolus (Lo *et al.* 2006). It has been demonstrated that single rDNA repeats are capable of forming mini-nucleoli (Karpen *et al.* 1988; Scheer and Hock 1999); however, only tandemly repeated rDNA arrays form proper nucleoli (Scheer and Hock 1999). Thus, it may be postulated that early in the evolution of the nucleus, the few copies of rDNA in the genome were not contained within any subnuclear structure; however, as genomes became more complex and cellular demand for protein synthesis increased, it became favorable to densely cluster

Figure 4 mESC pluripotency factors are associated with rDNA. (A) ChIP-seq profiles of OCT4, SOX2, NANOG, KLF4, STAT3, SMAD1, c-MYC, n-MYC, ZFX, E2F1, ESRRB, TCFCP2L1, P300, BRG1, and CHD7 at rDNA in mESCs. Signal intensity as determined by F-seq is plotted on the Y-axis. (B) Normalized tag density scores for pluripotency factors at rDNA in mESCs. (C) Heatmap of pairwise correlation scores between pluripotency factors and Polycomb proteins at rDNA.

a large number of rDNA repeats. This may be reflected in the current understanding of the nucleolus as "an organelle formed by the act of building a ribosome" (Mélèse and Xue 1995). Thus, the nuclear regulatory machinery that had participated in modulating rRNA transcription before the emergence of the nucleolus continued to function in this capacity in this novel subnuclear compartment. Although this hypothesis is speculative, further studies into the evolution of transcriptional regulation and the nucleolus will undoubtedly reveal striking insights into these dual-function factors.

ACKNOWLEDGMENTS

We thank Cindy Bartels and Michael Schnetz for assistance with cell culture. This work was supported by grants from the National Institute of General Medical Sciences (5T32GM008613 to G.E.Z.), National Institute of Child Health and Human Development (R01HD056369 to P.C.S.), and the National Human Genome Research Institute (5R01HG004722 to P.C.S.).

LITERATURE CITED

Aksoy, I., R. Jauch, J. Chen, M. Dyla, U. Divakar *et al.*, 2013 Switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. EMBO J. 32: 938–953.

Ali, S. A., S. K. Zaidi, C. S. Dacwag, N. Salma, D. W. Young *et al.*, 2008 Phenotypic transcription factors epigenetically mediate cell growth control. Proc. Natl. Acad. Sci. USA 105: 6632–6637.

Arabi, A., S. Wu, K. Ridderstrale, H. Bierhoff, C. Shiue *et al.*, 2005 c-Myc associates with ribosomal DNA and activates RNA polymerase I transcription. Nat. Cell Biol. 7: 303–310.

Azuma, M., R. Toyama, E. Laver, and I. B. Dawid, 2006 Perturbation of rRNA Synthesis in the *bap28* mutation leads to apoptosis mediated by p53 in the zebrafish central nervous System. J. Biol. Chem. 281: 13309–13316.

Bajpai, R., D. A. Chen, A. Rada-Iglesias, J. Zhang, Y. Xiong *et al.*, 2010 CHD7 cooperates with PBAF to control multipotent neural crest formation. Nature 463: 958–962.

Bannister, A. J., and T. Kouzarides, 2011 Regulation of chromatin by histone modifications. Cell Res. 21: 381–395.

Barski, A., S. Cuddapah, K. Cui, T.-Y. Roh, D. E. Schones *et al.*, 2007 High-resolution profiling of histone methylations in the human genome. Cell 129: 823–837.

Bauer, D. C., K. Willadsen, F. A. Buske, K.-A. Lê Cao, T. L. Bailey *et al.*, 2011 Sorting the nuclear proteome. Bioinformatics 27: i7–i14.

Bierhoff, H., K. Schmitz, F. Maass, J. Ye, and I. Grummt, 2010 Noncoding transcripts in sense and antisense orientation regulate the epigenetic state of ribosomal RNA genes. Cold Spring Harb. Symp. Quant. Biol. 75: 357–364.

Boyle, A. P., J. Guinney, G. E. Crawford, and T. S. Furey, 2008 F-Seq: a feature density estimator for high-throughput sequence tags. Bioinformatics 24: 2537–2538.

Brons, I. G. M., L. E. Smithers, M. W. B. Trotter, P. Rugg-Gunn, B. Sun *et al.*, 2007 Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature 448: 191–195.

Cao, R., L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage *et al.*, 2002 Role of histone H3 lysine 27 methylation in polycomb-group silencing. Science 298: 1039–1043.

Chakraborty, A., T. Uechi, S. Higa, H. Torihara, and N. Kenmochi, 2009 Loss of ribosomal protein L11 affects zebrafish embryonic development through a p53-dependent apoptotic response. PLoS ONE 4: e4152.

Chen, X., H. Xu, P. Yuan, F. Fang, M. Huss *et al.*, 2008 Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell 133: 1106–1117.

Cheung, M.-S., T. A. Down, I. Latorre, and J. Ahringer, 2011 Systematic bias in high-throughput sequencing data and its correction by BEADS. Nucleic Acids Res. 39: e103.

Danilova, N., K. M. Sakamoto, and S. Lin, 2008 Ribosomal protein S19 deficiency in zebrafish leads to developmental abnormalities and defective erythropoiesis through activation of p53 protein family. Blood 112: 5228–5237.

Datta, A., A. Nag, and P. Raychaudhuri, 2002 Differential regulation of E2F1, DP1, and the E2F1/DP1 complex by ARF. Mol. Cell. Biol. 22: 8398–8408.

Dauwerse, J. G., J. Dixon, S. Seland, C. A. L. Ruivenkamp, A. van Haeringen *et al.*, 2011 Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome. Nat. Genet. 43: 20–22.

Deutschbauer, A. M., D. F. Jaramillo, M. Proctor, J. Kumm, M. E. Hillenmeyer *et al.*, 2005 Mechanisms of haploinsufficiency revealed by genome-wide profiling in yeast. Genetics 169: 1915–1925.

Ebert, B. L., J. Pretz, J. Bosco, C. Y. Chang, P. Tamayo *et al.*, 2008 Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. Nature 451: 335–339.

Ernst, J., P. Kheradpour, T. S. Mikkelsen, N. Shoresh, L. D. Ward *et al.*, 2011 Mapping and analysis of chromatin state dynamics in nine human cell types. Nature 473: 43–49.

Gagnon-Kugler, T., F. Langlois, V. Stefanovsky, F. Lessard, and T. Moss, 2009 Loss of human ribosomal gene CpG methylation enhances cryptic RNA polymerase II transcription and disrupts ribosomal RNA processing. Mol. Cell 35: 414–425.

Gonzalez, I. L., and J. E. Sylvester, 1995 Complete sequence of the 43-kb human ribosomal DNA repeat: Analysis of the intergenic spacer. Genomics 27: 320–328.

Gonzalez, I. L., and J. E. Sylvester, 1997 Incognito rRNA and rDNA in databases and libraries. Genome Res. 7: 65–70.

Grandori, C., N. Gomez-Roman, Z. A. Felton-Edkins, C. Ngouenet, D. A. Galloway *et al.*, 2005 c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. Nat. Cell Biol. 7: 311–318.

Grozdanov, P., O. Georgiev, and L. Karagyozov, 2003 Complete sequence of the 45-kb mouse ribosomal DNA repeat: analysis of the intergenic spacer. Genomics 82: 637–643.

Grummt, I., 1999 Regulation of mammalian ribosomal gene transcription by RNA polymerase I. Prog. Nucleic Acid Res. Mol. Biol. 62: 109–154.

Guttman, M., M. Garber, J. Z. Levin, J. Donaghey, J. Robinson *et al.*, 2010 Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat. Biotechnol. 28: 503–510.

Han, J., P. Yuan, H. Yang, J. Zhang, B. S. Soh *et al.*, 2010 Tbx3 improves the germ-line competency of induced pluripotent stem cells. Nature 463: 1096–1100.

He, S., D. Pant, A. Schiffmacher, S. Bischoff, D. Melican *et al.*, 2006 Developmental expression of pluripotency determining factors in caprine embryos: Novel pattern of NANOG protein localization in the nucleolus. Mol. Reprod. Dev. 73: 1512–1522.

Heintzman, N. D., R. K. Stuart, G. Hon, Y. Fu, C. W. Ching *et al.*, 2007 Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat. Genet. 39: 311–318.

Heintzman, N. D., G. C. Hon, R. D. Hawkins, P. Kheradpour, A. Stark *et al.*, 2009 Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459: 108–112.

Henras, A., J. Soudet, M. Gérus, S. Lebaron, M. Caizergues-Ferrer *et al.*, 2008 The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell. Mol. Life Sci. 65: 2334–2359.

Ho, L., R. Jothi, J. L. Ronan, K. Cui, K. Zhao *et al.*, 2009 An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. Proc. Natl. Acad. Sci. USA 106: 5187–5191.

Iwanami, N., T. Higuchi, Y. Sasano, T. Fujiwara, V. Q. Hoa *et al.*, 2008 WDR55 Is a nucleolar modulator of ribosomal RNA synthesis, cell cycle progression, and teleost organ development. PLoS Genet. 4: e1000171.

Johansson, H., and S. Simonsson, 2010 Core transcription factors, Oct4, Sox2 and Nanog, individually form complexes with nucleophosmin

(Npm1) to control embryonic stem (ES) cell fate determination. Aging 2: 815–822.

Karpen, G. H., J. E. Schaefer, and C. D. Laird, 1988 A Drosophila rRNA gene located in euchromatin is active in transcription and nucleolus formation. Genes Dev. 2: 1745–1763.

Kobayashi, T., and A. R. D. Ganley, 2005 Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. Science 309: 1581–1584.

Kongsuwan, K., Q. Yu, A. Vincent, M. C. Frisardi, M. Rosbash *et al.*, 1985 A *Drosophila Minute* gene encodes a ribosomal protein. Nature 317: 555–558.

Ku, M., R. P. Koche, E. Rheinbay, E. M. Mendenhall, M. Endoh *et al.*, 2008 Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet. 4: e1000242.

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg, 2009 Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10: R25.

Lo, S., C. Lee, and H. Lai, 2006 The nucleolus: reviewing oldies to have new understandings. Cell Res. 16: 530–538.

Loh, Y.-H., Q. Wu, J.-L. Chew, V. B. Vega, W. Zhang *et al.*, 2006 The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. 38: 431–440.

Lupien, M., J. Eeckhoute, C. A. Meyer, Q. Wang, Y. Zhang *et al.*, 2008 FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell 132: 958–970.

McStay, B., and I. Grummt, 2008 The epigenetics of rRNA genes: From molecular to chromosome biology. Annu. Rev. Cell Dev. Biol. 24: 131–157.

Meissner, A., T. S. Mikkelsen, H. Gu, M. Wernig, J. Hanna *et al.*, 2008 Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature 454: 766–770.

Mélèse, T., and Z. Xue, 1995 The nucleolus: an organelle formed by the act of building a ribosome. Curr. Opin. Cell Biol. 7: 319–324.

Mikkelsen, T. S., M. Ku, D. B. Jaffe, B. Issac, E. Lieberman *et al.*, 2007 Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448: 553–560.

Moss, T., F. Langlois, T. Gagnon-Kugler, and V. Stefanovsky, 2007 A housekeeper with power of attorney: the rRNA genes in ribosome biogenesis. Cell. Mol. Life Sci. 64: 29–49.

Müller, J., C. M. Hart, N. J. Francis, M. L. Vargas, A. Sengupta *et al.*, 2002 Histone methyltransferase activity of a Drosophila polycomb group repressor complex. Cell 111: 197–208.

Ng, H.-H., and M. A. Surani, 2011 The transcriptional and signalling networks of pluripotency. Nat. Cell Biol. 13: 490–496.

Nishiyama, A., L. Xin, A. A. Sharov, M. Thomas, G. Mowrer *et al.*, 2009 Uncovering early response of gene regulatory networks in ESCs by systematic induction of transcription factors. Cell Stem Cell 5: 420–433.

Oliver, E. R., T. L. Saunders, S. A. Tarlé, and T. Glaser, 2004 Ribosomal protein L24 defect in Belly spot and tail (*Bst*), a mouse *Minute*. Development 131: 3907–3920.

Parfenov, V. N., G. N. Pochukalina, D. S. Davis, R. Reinbold, H. R. Schöler *et al.*, 2003 Nuclear distribution of Oct-4 transcription factor in transcriptionally active and inactive mouse oocytes and its relation to RNA polymerase II and splicing factors. J. Cell. Biochem. 89: 720–732.

Pelletier, G., V. Y. Stefanovsky, M. Faubladier, I. Hirschler-Laszkiewicz, J. Savard *et al.*, 2000 Competitive recruitment of CBP and Rb-HDAC regulates UBF acetylation and ribosomal transcription. Mol. Cell 6: 1059–1066.

Poortinga, G., K. M. Hannan, H. Snelling, C. R. Walkley, A. Jenkins *et al.*, 2004 MAD1 and c-MYC regulate UBF and rDNA transcription during granulocyte differentiation. EMBO J. 23: 3325–3335.

Prokopowich, C., T. Gregory, and T. Crease, 2003 The correlation between rDNA copy number and genome size in eukaryotes. Genome 46: 48–50.

Rada-Iglesias, A., R. Bajpai, T. Swigut, S. A. Brugmann, R. A. Flynn *et al.*, 2011 A unique chromatin signature uncovers early developmental enhancers in humans. Nature 470: 279–283.

Ridanpää, M., H. van Eenennaam, K. Pelin, R. Chadwick, C. Johnson *et al.*, 2001 Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. Cell 104: 195–203.

Ruggero, D., and P. P. Pandolfi, 2003 Does the ribosome translate cancer? Nat. Rev. Cancer 3: 179–192.

Russell, J., and J. C. B. M. Zomerdijk, 2005 RNA-polymerase-I-directed rDNA transcription, life and works. Trends Biochem. Sci. 30: 87–96.

Scheer, U., and R. Hock, 1999 Structure and function of the nucleolus. Curr. Opin. Cell Biol. 11: 385–390.

Schmidt, D., M. D. Wilson, C. Spyrou, G. D. Brown, J. Hadfield *et al.*, 2009 ChIP-seq: Using high-throughput sequencing to discover protein-DNA interactions. Methods 48: 240–248.

Schnetz, M. P., C. F. Bartels, K. Shastri, D. Balasubramanian, G. E. Zentner *et al.*, 2009 Genomic distribution of CHD7 on chromatin tracks H3K4 methylation patterns. Genome Res. 19: 590–601.

Schnetz, M. P., L. Handoko, B. Akhtar-Zaidi, C. F. Bartels, C. F. Pereira *et al.*, 2010 CHD7 targets active gene enhancer elements to modulate ES cell-specific gene expression. PLoS Genet. 6: e1001023.

Shimono, Y., H. Murakami, K. Kawai, P. A. Wade, K. Shimokata *et al.*, 2003 Mi-2β associates with BRG1 and RET finger protein at the distinct regions with transcriptional activating and repressing abilities. J. Biol. Chem. 278: 51638–51645.

Shimono, K., Y. Shimono, K. Shimokata, N. Ishiguro, and M. Takahashi, 2005 Microspherule protein 1, Mi-2β, and RET finger protein associate in the nucleolus and up-regulate ribosomal gene transcription. J. Biol. Chem. 280: 39436–39447.

Tesar, P. J., J. G. Chenoweth, F. A. Brook, T. J. Davies, E. P. Evans *et al.*, 2007 New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature 448: 196–199.

Teytelman, L., D. M. Thurtle, J. Rine, and A. van Oudenaarden, 2013 Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. Proc. Natl. Acad. Sci. U. S. A. 110: 18602–18607.

Trainor, P. A., J. Dixon, and M. J. Dixon, 2008 Treacher Collins syndrome: etiology, pathogenesis and prevention. Eur. J. Hum. Genet. 17: 275–283.

Uechi, T., Y. Nakajima, A. Nakao, H. Torihara, A. Chakraborty *et al.*, 2006 Ribosomal protein gene knockdown causes developmental defects in zebrafish. PLoS ONE 1: e37.

Vakoc, C. R., M. M. Sachdeva, H. Wang, and G. A. Blobel, 2006 Profile of histone lysine methylation across transcribed mammalian chromatin. Mol. Cell. Biol. 26: 9185–9195.

van den Berg, D. L. C., T. Snoek, N. P. Mullin, A. Yates, K. Bezstarosti *et al.*, 2010 An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell 6: 369–381.

Yamaji, M., J. Ueda, K. Hayashi, H. Ohta, Y. Yabuta *et al.*, 2013 PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. Cell Stem Cell 12: 368–382.

Zentner, G. E., E. A. Hurd, M. P. Schnetz, L. Handoko, C. Wang *et al.*, 2010 CHD7 functions in the nucleolus as a positive regulator of ribosomal RNA biogenesis. Hum. Mol. Genet. 19: 3491–3501.

Zentner, G. E., A. Saiakhova, P. Manaenkov, M. D. Adams, and P. C. Scacheri, 2011a Integrative genomic analysis of human ribosomal DNA. Nucleic Acids Res. 39: 4949–4960.

Zentner, G. E., P. J. Tesar, and P. C. Scacheri, 2011b Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. 21: 1273–1283.

Zhang, Y.-W., N. Yasui, K. Ito, G. Huang, M. Fujii *et al.*, 2000 A *RUNX2/PEBP2αA/CBFA1* mutation displaying impaired transactivation and Smad interaction in cleidocranial dysplasia. Proc. Natl. Acad. Sci. USA 97: 10549–10554.

Zuccotti, M., V. Merico, L. Sacchi, M. Bellone, T. Brink *et al.*, 2008 Maternal Oct-4 is a potential key regulator of the developmental competence of mouse oocytes. BMC Dev. Biol. 8: 97.

Communicating editor: A. Gasch