



Published in final edited form as:

Chromosoma. 2020 June ; 129(2): 121–139. doi:10.1007/s00412-020-00734-9.

Distinct features of nucleolus-associated domains in mouse embryonic stem cells

Aizhan Bizhanova, Aimin Yan, Jun Yu, Lihua Julie Zhu*, Paul D. Kaufman*

Department of Molecular, Cellular and Cancer Biology, University of Massachusetts Medical School, Worcester, MA 01605 USA

Abstract

Background: Heterochromatin in eukaryotic interphase cells frequently localizes to the nucleolar periphery (nucleolus-associated domains, NADs) and the nuclear lamina (lamina-associated domains, LADs). Gene expression in somatic cell NADs is generally low, but NADs have not been characterized in mammalian stem cells.

Results: Here, we generated the first genome-wide map of NADs in mouse embryonic stem cells (mESCs) via deep sequencing of chromatin associated with biochemically-purified nucleoli.

As we had observed in mouse embryonic fibroblasts (MEFs), the large Type I subset of NADs overlaps with constitutive LADs and is enriched for features of constitutive heterochromatin, including late replication timing and low gene density and expression levels. Conversely, the Type II NAD subset overlaps with loci that are not lamina-associated, but in mESCs, Type II NADs are much less abundant than in MEFs. mESC NADs are also much less enriched in H3K27me3 modified regions than are NADs in MEFs. Additionally, comparison of MEF and mESC NADs revealed enrichment of developmentally regulated genes in cell type-specific NADs. Together, these data indicate that NADs are a developmentally dynamic component of heterochromatin.

Conclusions: These studies implicate association with the nucleolar periphery as a mechanism for developmentally-regulated gene expression, and will facilitate future studies of NADs during mESC differentiation.

Keywords

nucleolus associated domains; NADs; embryonic stem cell; mouse; histone modification

*Contact: Julie.zhu@umassmed.edu; paul.kaufman1@umassmed.edu, University of Massachusetts Medical School, Department of Molecular, Cellular and Cancer Biology, 364 Plantation St., Worcester, MA 01605 USA.

Authors contributions: AB performed all the wet lab experiments, and analyzed and interpreted the data. AY and JY performed bioinformatic analyses. LJZ directed and conducted the bioinformatics analyses, and PDK directed and analyzed the wet lab experimentation. AB, LJZ and PDK wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The datasets supporting the conclusions of this article are publicly available at the 4D Nucleome Data Portal (<https://data.4dnucleome.org/>). The RNAseq data is at <https://data.4dnucleome.org/experiment-set-replicates/4DNESDHILYLU/#raw-files>. The DNAseq data for the total genomic samples is at <https://data.4dnucleome.org/experiment-set-replicates/4DNESUJZ5FL2/>, and the DNAseq data for the nucleolar samples is at <https://data.4dnucleome.org/experiment-set-replicates/4DNESXE9K9DB/>.

Competing interests: The authors declare that they have no competing interests.

Introduction

Eukaryotic genomes are broadly subdivided into more accessible, transcriptionally active euchromatin, and less accessible, less active heterochromatin. These functional classifications are accompanied by spatial separation: heterochromatin is mainly found at the nuclear periphery and nucleolar periphery, where they comprise nucleolus-associated domains (NADs) (Németh et al. 2010; van Koningsbruggen et al. 2010) and lamina-associated domains (LADs) (Pickersgill et al. 2006; Guelen et al. 2008; Peric-Hupkes et al. 2010), respectively. Studies in multiple organisms indicate that sequestration of heterochromatin to the nuclear and nucleolar peripheries contributes to gene silencing (Fedoriw et al. 2012b; Zullo et al. 2012; Jakociunas et al. 2013). Therefore, there is great interest in discovering the molecular bases for these localizations. Notably, some trans-acting factors that specifically affect lamina (Zullo et al. 2012; Harr et al. 2015) or nucleolar (Yusufzai et al. 2004; Zhang et al. 2007; Mohammad et al. 2008; Padeken and Heun 2014; Smith et al. 2014; Matheson and Kaufman 2017; Singh et al. 2018) associations have been reported, suggesting that distinct mechanisms contribute at the two locations.

Both NADs and LADs are enriched for silent genes and histone modifications characteristic of constitutive heterochromatin, e.g. H3K9me2 and H3K9me3 (Matheson and Kaufman 2016; van Steensel and Belmont 2017). LADs have been mapped and studied in multiple species and cell types (Pickersgill et al. 2006; Guelen et al. 2008; Peric-Hupkes et al. 2010; Kind et al. 2013; Borsos et al. 2019). In contrast, NADs have been characterized in a few human somatic cell lines (Németh et al. 2010; van Koningsbruggen et al. 2010; Dillinger et al. 2017), in the plant *Arabidopsis thaliana* (Pontvianne et al. 2016), and recently, in mouse embryonic fibroblasts (MEFs) (Vertii et al. 2019). Several experiments indicate that LADs can be redistributed to the nucleolar periphery after passage through mitosis, and vice versa (van Koningsbruggen et al. 2010; Kind et al. 2013). However, the extent of overlap between LADs and NADs is unknown in most organisms and cell types.

Here, we mapped and characterized NADs in mouse embryonic stem cells (mESC), a tractable system for studying how NADs change during differentiation. As in MEFs (Vertii et al. 2019), we identified a large subset of mESC NADs that overlap with LADs (Type I NADs), and a smaller subset of NADs that do not overlap LADs (Type II NADs). However, Type II NADs are less prevalent in mESCs than in MEFs. mESC NADs are also notably less enriched in H3K27me3 modifications. Comparisons of MEF and mESC NADs also revealed enrichment of developmentally regulated genes in cell type-specific NADs. These analyses will facilitate future studies of genome dynamics during stem cell differentiation.

Results

Isolation of nucleoli from crosslinked F121-9 mESCs.

We isolated nucleoli from formaldehyde-crosslinked hybrid F121-9 mES cells using methods previously shown to yield reproducible data using MEF cells (Vertii et al. 2019). In those studies, crosslinked and non-crosslinked MEFs were directly compared, and shown to yield highly overlapping results, with crosslinked samples detecting a greater proportion of the genome associated with nucleoli (Vertii et al. 2019). This suggests crosslinking could

assist detection of weak or transient nucleolar interactions. Therefore, we used crosslinking for all nucleoli isolation experiments here (Fig. 1A). The purity of isolated nucleoli was confirmed using phase-contrast microscopy (Supplemental Fig. S1A). Immunoblot analysis of nucleolar fractions showed that they were enriched for nucleolar protein fibrillarin relative to beta-actin (Supplemental Fig. S1B). Because our previous experiments showed that H3K27 methylation was important for NAD-nucleolar interactions in MEF cells (Vertii et al. 2019), we measured the ratio of heterochromatin mark H3K27me3 to total H3 protein levels, and found it was modestly enriched in nucleolar fractions relative to the total extracts (1.5–1.6 fold enrichment, Supplemental Fig. S1C). Quantitative PCR analysis revealed 9–18-fold enrichment of 45S rDNA sequences in purified nucleolar DNA relative to genomic DNA (Supplemental Fig. S1D). These results indicated the enrichment of nucleoli in our preparations, hence we proceeded with whole-genome sequencing of nucleolar DNA.

Bioinformatic analysis of NADs.

We performed two biological replicate preparations of crosslinked F121-9 mESC nucleoli. In each replicate experiment, we extracted nucleolar-associated DNA from nucleoli, along with genomic DNA from whole cells from the same population of cells. We sequenced approximately 50 million paired-end reads from each nucleolar and genomic DNA sample. We note that subsampling analyses of larger MEF datasets previously showed that the number of peaks detected had reached a plateau at this sequencing depth (Vertii et al. 2019). Genomic reads were mostly uniformly distributed across the genome, whereas nucleolar reads contained well-defined peaks and valleys, with peaks overlapping known heterochromatic regions, such as constitutive LADs (cLADs) (Peric-Hupkes et al. 2010) and late replicating regions (Hiratani et al. 2010) (Fig. 1B, C). cLADs were previously defined as LADs that are lamina-associated in mESCs, and also in neural precursor cells (NPCs) and astrocytes differentiated from these mESCs (Peric-Hupkes et al. 2010). Previous studies of NADs have identified frequent overlap of NADs with LADs (van Koningsbruggen et al. 2010; Németh et al. 2010; Dillinger et al. 2017; Vertii et al. 2019) and with late-replicating regions (Dillinger et al. 2017; Vertii et al. 2019), thus we concluded that the nucleolar reads are enriched with bona fide nucleolar heterochromatic regions in F121-9 mESCs.

Calculating the log ratio of nucleolar reads to genomic reads resulted in a raw metric of nucleolar association across the genome (Nucleolus/gDNA ratio tracks in Fig. 1B, C). As in MEFs, visual inspection of the nucleolus/genomic ratio in mESC revealed a negative slope across most chromosomes, especially noticeable on large chromosomes (Fig. 1C). Mouse chromosomes are acrocentric, i.e. the centromere is found at one end of a chromosome, and by convention these are annotated on the left. Because pericentromeric regions frequently associate with nucleolar periphery (Ragoczy et al. 2014), nucleolar associations on centromeric end of chromosomes are usually more frequent. Peak calling based only on nucleolar/genomic ratios would identify peaks mostly at the centromere-proximal end, with fewer peaks at the chromosome-distal end. For this reason, we used our previously described Bioconductor package named *NADfinder* (Vertii et al. 2019) to call NAD peaks in F121-9 mESCs. This software uses local background correction, which was important for detection of validated NAD peaks distal from centromeres in MEFs (Vertii et al. 2019). *NADfinder* peak calling was performed using the default settings with a 50kb window size, a testing

threshold of $\log_2(1.5)$ for background corrected $\log_2(\text{nucleolar/genomic})$ ratio to define the null hypothesis, and adjusted p -value < 0.05 (Vertii et al. 2019). Potential peaks were further filtered to be > 50 kb long and to have \log_2 ratio > 1.7 .

The length of the identified F121-9 NADs ranges up to 8 Mb (Fig. 1D), with median length 1.1 Mb, which is slightly larger than median length of MEF NADs, 0.7 Mb (Vertii et al. 2019). We noted that NADs in F121-9 cells covered 31% of the non-repetitive genome, a smaller percentage than observed in crosslinked MEF NADs (41%) (Vertii et al. 2019). The 31% fraction of the mESC genome in NADs is also smaller than the fraction of the mouse genome in LADs, either for embryonic stem cells or somatic cells ($\sim 40\%$) (Peric-Hupkes et al. 2010), or during early mouse embryogenesis (~ 40 – 60%) (Borsos et al. 2019) (see Discussion).

3D immuno-FISH confirmation of NAD peaks in F121-9 mESCs.

To validate associations of NADs with nucleoli by an orthogonal method, we performed 3D immuno-FISH experiments, scoring association of BAC DNA probes with nucleolar marker proteins fibrillarin or nucleophosmin (Figs. 1–2). We tested the association of a euchromatic negative control probe, pPK871, which lacks nucleolar association in MEFs (Vertii et al., 2019) and did not contain a peak in our F121-9 NAD-seq data. The mean frequency of nucleolar association for this probe was $\sim 24\%$ (Fig. 3A). Three additional non-NAD BAC probes (pPK825, pPK1000, and pPK1003) displayed similar levels of nucleolar association (Fig. 3A). The average association frequency for these non-NAD probes in F121-9 cells was 22%, similar to the 20% frequency observed in MEF cells (Vertii et al., 2019). These observations result from stochastic positioning of loci within the nuclear volume. We note that pPK825 was also not associated with nucleoli in MEFs, whereas pPK1000 and pPK1003 had not been tested in MEFs (Vertii et al. 2019).

We also analyzed BAC probes pPK915 and pPK999 (Figs. 2A, B and Fig. 3), two centromere-proximal probes which were expected to strongly associate with nucleoli. Indeed, both of these probes displayed more frequent nucleolar association than did the set of non-NAD probes ($p < 0.0001$). We note that probe pPK915 overlaps a ciLAD region (Fig. 2A), which means that it was not observed to associate with lamina in mESCs or MEFs (Peric-Hupkes et al. 2010). Thus, this is an example of preferential nucleolar association, which will be discussed below. We also note that pPK999 overlaps both a NAD and a LAD in mESCs, but not in MEFs (pPK999, Fig. 2B). Furthermore, pPK999 contains the *Egfr* gene, for which transcript levels are higher in MEFs (FPKM value 51.5) (Delbarre et al. 2017) compared to mESCs (FPKM value 0.2 (Supplemental Table 1)). This is an example of a genomic locus that is nucleolar-associated and transcriptionally repressed in mESCs, and which is no longer associated and becomes more active in MEFs. In sum, these FISH data demonstrate that the identified NADs include bona fide nucleolar heterochromatic regions in F121-9 mESCs, and that these localizations can vary during cell differentiation.

We also examined whether the gaps between called peaks in centromere-proximal regions (e.g. Fig. 2C, D and Fig. 3) were associated with nucleoli. In our previous analysis of MEF NADs, we showed that such regions, which we termed “NAD splitting regions (NSRs)” displayed much higher gene expression than do the flanking NADs (Vertii et al. 2019). This

indicated that the decreased nucleolar sequencing read numbers from NSRs reflect a distinct chromatin state. To test directly whether such regions display reduced nucleolar association, we analyzed two BAC probes (pPK1006 and pPK1007) that display canonical NSR characteristics: they overlap ciLADs and early replicating regions, and exhibit high transcriptional activity (Fig. 2C, D). FISH assays established that these two probes are more frequently associated with nucleolar periphery compared to non-NAD probes ($p < 0.0001$), hence these two BAC regions are NADs (Fig. 3A). However, either with or without background correction, *NADfinder* did not predict either pPK1006 or pPK1007 to be NADs (Fig. 2C, D). We hypothesize that there are both bioinformatic as well as biochemical contributions to these discrepancies in the analysis of the centromere-proximal NADs (see Discussion).

We also considered the case of regions at the distal end of chromosomes, far from centromeres. We analyzed two BAC probes, pPK914 and pPK1012, and observed that both these were significantly more frequently associated with nucleoli than were non-NAD probes ($p < 0.0001$, Fig. 2E, F). These regions were classified as NADs by *NADfinder* when background correction method was included, whereas in the absence of local background correction, neither were predicted to be NADs (Fig. 2E, F). Hence, *NADfinder* clearly benefits from the use of background correction method in identifying distal NADs.

Two types of NADs in F121-9 mESCs.

In our previous analysis of MEF data, we had defined a “Type I” class of NADs as those overlapping LADs (Vertii et al. 2019). Additionally, a contrasting “Type II” class of NADs was defined which overlaps “constitutive interLADs” (ciLADs), the regions defined as those which were not lamina-associated during multiple steps of cellular differentiation (Peric-Hupkes et al. 2010). In MEFs, Type I NADs are approximately five-fold more abundant, and tend to replicate late; in contrast, the less abundant Type II NADs more frequently overlap with early replicating regions (Vertii et al. 2019). In F121-9 mESC NADs, we also observed abundant Type I NADs that overlap with cLADs (421 Mb of the total 845Mb NAD population; Fig. 4A). However, Type II NADs that overlap with ciLADs comprise only 77 Mb, much less than the 147 Mb observed in similarly crosslinked MEFs (Fig. 4A; Vertii et al. 2019). Visual inspection of the distribution of the two classes in a genome browser illustrated the greater size of the Type I subset compared to Type II regions (Fig. 4B, E, F). Despite the small size of the F121-9 Type II NAD subset, we note that we have validated nucleolar association of two Type II NAD probes (pPK914, pPK915; Figs. 2–3). These two probes lack significant lamina association in MEFs (Vertii et al. 2019), and both overlap ciLAD regions (Fig. 2A, E), indicating that they lack lamina association during multiple steps in the process of differentiation from mES cells to astrocytes (Peric-Hupkes et al. 2010; Meuleman et al. 2013). We conclude that in mES cells, as in MEFs, a large proportion of NADs overlap LAD regions, but that the amount of ciLAD overlap in mES cells is smaller.

We then analyzed gene density and gene expression characteristics of the different NAD subsets from F121-9 cells. As we had observed in MEFs (Vertii et al., 2019), gene density of Type II NADs was greater than that of NADs as a whole, which in turn have higher gene

density compared to Type I NADs (Fig. 4C). Using RNA-seq data we obtained from the same preparations of F121-9 cells that were used for nucleolar purification, we analyzed genomic trends in steady-state mRNA levels by plotting the distributions of the FPKM values. As in MEFs (Vertii et al. 2019), F121-9 NADs displayed lower FPKM values than the genome-wide average ($p < 0.0001$). In addition, FPKM values for the Type I NAD subset were significantly lower than those for NADs as a whole ($p < 0.0001$) (Fig. 4D). Thus, Type I NADs in both MEFs and F121-9 cells display low gene expression levels characteristic of heterochromatin. In contrast, in F121-9 cells Type II NADs displayed mean gene expression levels that are slightly higher than those observed in the whole genome ($p < 0.0003$) or even in non-NAD regions ($p < 0.0233$) (Fig. 4D). Therefore, in both F121-9 cells and MEFs (Vertii et al., 2019), Type II NADs can become associated with nucleoli without adopting the highly silenced status of Type I NADs.

However, F121-9 NADs displayed a prominent difference from those in MEFs, regarding overlap with H3K27me3 peaks. We note that H3K27me3 is functionally important for heterochromatin localization because Ezh2 inhibitors that block this modification decrease laminar and nucleolar associations by heterochromatin (Harr et al. 2015; Vertii et al. 2019). In MEFs, we observe frequent overlap of H3K27me3 peaks (Delbarre et al. 2017) with both Type I (117 Mb out of 567 Mb) and Type II NADs (101 Mb out of 147 Mb) (Fig. 4G, H and Supplemental Figure 2; Vertii et al. 2019). In contrast, in F121-9 cells we observed that overlap of NADs with H3K27me3-enriched domains (Cruz-Molina et al. 2017) was much smaller than observed in MEFs: only 9 Mb of the 421 Mb of Type I NADs and 22 Mb of 77 Mb of Type II NADs overlap with H3K27me3 domains (Fig. 4E, F). These differences likely reflect the lower abundance of repressive histone marks in mESCs compared to differentiated cells; this includes H3K27me3, which becomes more abundant during differentiation ((Martens et al. 2005; Hawkins et al. 2010; Atlasi and Stunnenberg 2017; Zhang et al. 2020); see Discussion). Indeed, our analysis of an mESC data set (GSM2416833; Cruz-Molina et al. 2017; see Methods) detected 517 Mb of H3K27me3 peak regions in F121-9 cells, and an almost two-fold larger amount (990 Mb) was found in MEFs (GSM1621022; Delbarre et al. 2017). However, we note that the amount of H3K27me3 peaks in NADs is much more than two-fold greater in MEFs (417 Mb, Fig. 4G, H) than in F121-9 cells (66 Mb, Fig. 4E, F). Together, these data suggest that H3K27 methylation is a key aspect of NAD chromatin maturation that has not yet occurred fully in mES cells (see Discussion).

Cell type-specific and conserved NADs.

We compared F121-9 stem cell NADs with crosslinked MEF NADs (Vertii et al. 2019), defining overlapped regions on a nucleotide-by-nucleotide basis (e.g. Fig. 5A). Close to 80% (660 Mb) of nucleotides in stem cell NADs overlap with nucleotides in MEF NADs (Fig. 5A). We designate NADs shared by MEFs and F121-9 stem cells as “conserved NADs”. Analysis of the intersection of conserved NADs with cLAD and ciLAD regions revealed that more than half of conserved NADs overlap cLADs (370 Mb; Fig. 4A), which are the most gene-poor subset of LADs and are generally poorly expressed, constitutive heterochromatin (Peric-Hupkes et al. 2010; Meuleman et al. 2013; van Steensel and Belmont 2017). Consistent with these trends, Jaccard similarity coefficient analysis indicated high

correlation of conserved NADs with cLADs and late replicating regions (Marchal et al. 2018) (Fig. 5B). Furthermore, the conserved NADs display the lowest transcript levels in both cell types (Fig. 5C–F), as expected due to the constitutive heterochromatic features of these regions.

We next turned our attention to NADs found only in one of the two analyzed cell types. The Jaccard analysis indicated that these cell type-specific NAD regions (i.e. “MEF-specific NADs” and “F121-9-specific NADs”) are distinct from the conserved NADs, clustering separately from conserved NADs, cLAD and late replicating regions (Fig. 5B). We analyzed steady-state mRNA levels in conserved and cell type-specific NADs by using FPKM values from F121-9 and MEF (Delbarre et al. 2017) RNA-seq data (Fig. 5C, D). As we expected, MEF RNA-seq data revealed lower levels of transcripts from genes within MEF-specific NADs than from F121-9-specific NADs (p-value < 0.0001) (Fig. 5C), indicating that in MEFs, nucleolar association correlates with transcriptional silencing. In contrast, our RNA-seq data from F121-9 cells showed that transcript levels within both the MEF-specific NADs and the F121-9-specific NADs are statistically indistinguishable (p-value = 0.82) (Fig. 5D). We observed similar trends in independent sets of MEF and mESC RNA-seq data from the literature (Lowe et al. 2015; Chronis et al. 2017) (Fig. 5E, F). These observations were unexpected in that the MEF-specific NADs are not nucleolar-associated in the F121-9 cells, yet are on average less highly expressed than non-NAD genes in these cells. These data suggest that in F121-9 stem cells, gene repression could precede localization to the nucleolar periphery that occurs later during cellular differentiation (see Discussion).

Gene Ontology analysis of conserved and cell-type specific NADs.

To further characterize the conserved NADs, we next analyzed enriched GO-terms within these. The most significantly enriched Molecular Functions term was “Response to smell detection” (Fig. 6A; Supplemental Table 3), including olfactory receptor (OR) and vomeronasal receptor genes. These clustered genes are not expressed in either stem cells or fibroblasts and are frequently within NADs in both F121-9 stem cells and MEFs (e.g. the OR genes on chr11, Fig. 6B). Among other well-represented gene families in conserved NADs were cytochrome P450 family members: *Cyp2a12*, *Cyp2b10*, *Cyp2c50* (“heme-interacting genes” in Fig. 6A), which are responsible for breaking down toxins, as well as synthesizing steroid hormones, fats and acids, and are most highly expressed in liver (Hannemann et al. 2007). Neurotransmitter receptors were also enriched for conserved NADs, for example, genes that encode for glutamate receptors (*Gria2*, *Grid2*, etc.), GABA-A receptors (*Gabra5*, *Gabrb1*, etc.) and glycine receptors (*Glr1*, *Glr2*, etc.). The common thread among these gene classes is in that they are developmentally regulated, and most strongly induced in lineages not represented by embryonic stem cells or fibroblasts.

We next analyzed the F121-9-specific NADs. Among these, chemotactic cytokines were the GO-derived “Molecular Functions” class with the lowest q-value (Fig. 6C; Supplemental Table 4). The majority of these chemokines are represented by the CC chemokine ligand family, a cluster of which is shown in Fig. 6D. This cluster of *Ccl2*, *Ccl12* and *Ccl11* genes has heterochromatic features in the F121-9 cells: late replication timing, no steady-state mRNA transcripts, presence within both LAD and NAD regions. In contrast, in MEFs this

gene cluster is within neither NAD nor LAD sequences and has euchromatic features, including early replication timing and high gene transcript levels. This is an example of a genomic region in which multiple features are altered, becoming more euchromatic upon differentiation.

We then considered the converse case, the MEF-specific NADs. Among these, the “Biological Processes” GO classifications included genes responsible for differentiation along the anterior-posterior axis (Fig. 6E; Supplemental Table 5), an example of which is *Pcska6* gene (Fig. 6F). This genomic region displays euchromatic features (overlapping a ciLAD region, early replicating timing and high transcript levels: FPKM value 22.2) in mESCs, befitting the need for anterior-posterior axis establishment factors at this early developmental stage. In MEFs, this locus displays altered features, becoming nucleolar-associated, and generating reduced transcript levels (FPKM value 6.6) (Delbarre et al. 2017). In general, both conserved and cell type-specific NADs generally include genes that display reduced expression levels, suggesting that nucleolar localization could contribute to (or be a consequence of) the transcriptional silencing of resident genes. A major question remains as to how functionally distinct classes of NADs (e.g. Type I and Type II NADs) are targeted to nucleoli, and how this has distinct transcriptional consequences in each case (e.g. Fig. 4D; see Discussion).

Allele-specific NADs.

The F121-9 mES cell line we analyzed was derived from a blastocyst that resulted from a cross between two different mouse sub-species (*129x Castaneus*) (Rivera-Mulia et al. 2018). The genomes of these subspecies have a high single nucleotide polymorphism (SNP) density, with a genome-wide average of 1 SNP per 150 bp (Rivera-Mulia et al. 2018). To determine if we could detect allele-specific differences in our data, we assigned our NADseq data to one of the two genomes using *SNPsplit_v0.3.2* for allele-specific splitting of alignments (Krueger and Andrews 2016). We obtained approximately 50 million reads in each replicate of the nucleolar DNA samples. We then identified the reads with informative SNPs, yielding approximately 14–16 million allele-specific reads per genome. These allele-specific reads were used for the identification of allele-specific NAD peaks by *NADfinder*. However, visual inspection of the allele-specific peaks did not reveal correlations with allelic differences in previously published F121 subclone data sets, such as replication timing and ATAC-seq (Rivera-Mulia et al. 2018), Hi-C (Giorgetti et al. 2016) and our F121-9 RNA-seq data sets. However, previous subsampling analysis of larger MEF NAD-seq datasets revealed that the number of peaks detected plateaued at approximately 50 million reads (Vertii et al. 2019). Hence, to detect the maximum number of allele-specific features, we predict needing to obtain ~ 160–180 million reads per sample prior to assigning to specific genomes.

Nevertheless, we did observe chromosome-scale differences in the nucleolar associations in the two genomes. Although the numbers of allele-specific sequencing reads from the genomic samples were comparable for the two alleles throughout the genome, for some chromosomes the numbers of nucleolar sequencing reads differed for the two alleles. For example, more nucleolar reads were detected from the *Castaneus* genome for chromosomes 10 and 11 (Fig. 7A, B), whereas more reads from the *129* genome were detected for

chromosome 16 (Fig. 7C). We note that these differences were more pronounced on the centromere-proximal side of the chromosome. Therefore, one possible explanation for these allele-specific differences could be distinctions in chromosomes bearing nucleolar organizer regions (NORs) between these two mouse genomes. Nucleoli are formed by NORs, which contain arrays of ribosomal DNA (rDNA) repeats (McStay 2016). The identity of NOR-bearing chromosomes vary among mouse subspecies and strains (Britton-Davidian et al. 2012). *Mus musculus castaneus* subspecies harbor NORs on chromosome 11, and less frequently, on chromosome 10, but not on chromosome 16 (Suzuki et al. 1990; Britton-Davidian et al. 2012). Conversely, a laboratory inbred strain *129* has been reported to contain a NOR on chromosome 16 (Kurihara et al. 1994), but there are no reports of NORs on chromosomes 10 and 11 in this strain.

We hypothesized that higher nucleolar signals of a chromosome would correlate with the higher frequency of nucleolar contact. To test our hypothesis, we performed immuno-FISH using whole chromosome paint probes for chromosomes 5, 10, 12 and nucleophosmin antibody to label nucleoli (Fig. 7D–F). We chose to analyze chromosome 10 because the difference between allele-specific nucleolar reads was most noticeable (Fig. 7A). We used chromosome 5 as our negative control for NOR-bearing chromosome, due to it being a large chromosome with no literature reports of containing a NOR and exhibiting low nucleolar read intensity in both genomes according to NAD-seq (Supplemental Fig. S3A). Chromosome 12 has been shown to contain a NOR both in the *129* strain (Kurihara et al. 1994) and in *Castaneus* subspecies (Suzuki et al. 1990; Britton-Davidian et al. 2012). Consistent with these reports, both *129* and *Castaneus* genomes showed robust nucleolar signal for chromosome 12 (Supplemental Fig. S3B). Thus, we utilized chromosome 12 as our positive control for NOR-bearing chromosome. We predicted that chromosome 10 would show intermediate results between chromosome 5 and 12, due to only *Castaneus* genome showing robust nucleolar signal. Indeed, the mean frequency of chromosome 10 contacting the nucleoli was 49.7%, whereas for chromosome 5 it was 36.8%, and for chromosome 12 it was 74.3% (Fig. 7G). The statistical difference between the association frequencies of chromosomes 5 and 10 ($p=0.0179$) and chromosomes 10 and 12 ($p=0.0014$) was statistically significant. We conclude that NAD-seq was able to distinguish different allele-specific frequencies of nucleolar association across broad chromosomal domains.

Discussion

Heterochromatin formation during differentiation.

Several types of evidence indicate that compared to differentiated cells, chromatin in mESCs is less condensed, and the ratio of euchromatin to heterochromatin is higher (Gaspar-Maia et al. 2011). For example, fluorescence recovery after photobleaching experiments demonstrated that mESCs display more highly mobile core and linker histones, as well as Heterochromatin Protein 1 (HP1 α) than do differentiated cells. These features are thought to contribute to the transcriptional hyperactivity in pluripotent stem cells (Meshorer et al. 2006; Bhattacharya et al. 2009), including many repetitive elements that are silent in somatic cells but are transcribed in mESCs (Efroni et al. 2008). Additionally, microscopy studies showed that electron-dense heterochromatic structures are less condensed and less frequently

localize near nuclear lamina in mESCs compared to heterochromatin in differentiated cells (Hiratani et al. 2010; Ahmed et al. 2010; Mattout et al. 2015). Particularly relevant to our studies, more prominent electron-dense perinucleolar heterochromatin-like structures have been observed in differentiated cells, such as NPCs, compared to mESCs (Savi et al. 2014). In concert with changes in the appearance and localization of heterochromatin, the abundance of heterochromatic marks such as H3K27me3 and H3K9me3 increases during differentiation (Lee et al. 2004; Martens et al. 2005; Meshorer et al. 2006; Wen et al. 2009; Hawkins et al. 2010). Some of this regulation is locus-specific, as in a recent report indicating that PRC2 is prevented from generating repressive H3K27me3 marks at rDNA as part of a mechanism to promote high levels of ribosome biogenesis in pluripotent cells (Zhang et al. 2020). Together, these data are consistent with our observation that NADs in mESCs comprise a smaller fraction of the genome compared to MEFs (31 vs. 41%). Likewise, genome coverage by LADs increases during differentiation. For example, a recent study shows that LADs are first established immediately after fertilization, preceding TAD formation and instructing A/B compartment establishment (Borsos et al. 2019).

The Type II class of NADs is different in stem cells and fibroblasts.

Two functionally distinct classes of NADs have recently been reported in mouse embryonic fibroblasts (Vertii et al. 2019). Here, we show that in F121-9 mESCs, Type I NADs that overlap LAD regions are frequently the same as those found in MEFs (Fig. 4A), and exhibit similar low gene expression levels as expected for constitutive heterochromatin (Fig. 4D). In contrast, the Type II NADs defined by their overlap with ciLAD regions is much smaller in F121-9 than in MEF cells (Fig. 4A). We also note that NADs in F121-9 cells display much less overlap with H3K27me3 peaks than do MEF NADs (Fig. 4E–H). Together, these data suggest that acquisition of H3K27me3, the hallmark of facultative heterochromatin (Trojer and Reinberg 2007) by NADs is part of the process of cellular differentiation. Indeed, we note that GO analysis of MEF Type II NADs showed enrichment for developmentally regulated GO terms, for example, organ morphogenesis and sensory organ development (Vertii et al. 2019). Thus, stem cells prevent developmentally important genes from acquiring characteristics of facultative heterochromatin including nucleolar association, whereas these genes can become NADs after they are no longer required during development.

How are NADs targeted to nucleoli?

The precise mechanisms for targeting the two distinct classes of NADs to nucleoli remain unclear. Several studies implicate phase separation in the formation of heterochromatin domains (Larson et al. 2017; Strom et al. 2017; Shin et al. 2018) and nuclear bodies, such as nucleoli (Brangwynne et al. 2011; Feric et al. 2016; Mitrea et al. 2016). Our recent data suggest that Type II NADs are more sensitive than Type I NADs to hexanediol treatment (Vertii et al. 2019). Hexanediol perturbs phase separation, likely due to interfering with weak hydrophobic interactions that are important for liquid-like condensate formation (Ribbeck and Görlich 2002). Liquid-liquid demixing reactions frequently involve proteins that have intrinsically disordered regions (IDR) and RNA recognition motifs (Feric et al. 2016), as found for example in nucleolar proteins fibrillarin (Fbl) and nucleophosmin (Npm1). Notably, depletion of Nlp, the *Drosophila* homolog of Npm1, led to declustering of

centromeres and decreased association of centromeres with nucleolar periphery (Padeken et al. 2013). Therefore, it is possible that Type II NADs are specifically targeted to nucleolar periphery through the interactions between nucleolar proteins with IDRs (e.g. Npm1) with RNA species that are yet to be identified. Additionally, Polycomb repressive complex 1 (PRC1) protein chromobox 2 (CBX2) undergoes phase separation and forms liquid-like condensates in mESCs (Tatavosian et al. 2019), and Polycomb proteins are part of the MiCee complex that together with let-7 family miRNAs confers nucleolar association to specific loci (Singh et al. 2018). Therefore, Polycomb group (PcG) proteins are good candidates for nucleolar targeting of Type II NADs via phase separation. This may be especially important during differentiation, when PcG proteins gain special importance (Aloia et al. 2013; Lavarone et al. 2019). However, inhibition of PRC2 enzymatic activity decreases both nucleolar (Singh et al. 2018; Vertii et al. 2019) and laminar heterochromatin localizations (Harr et al. 2015), making it unlikely that PRC2 can target loci to a unique destination. Additionally, nucleolar localization of the *Kcnq1* locus can occur in cells lacking functional Polycomb complexes (Fedoriw et al. 2012a), indicating that multiple mechanisms likely exist. Other candidate trans-acting factors that could specifically target genomic regions to the nucleolar periphery are the proteins Ki-67 and the p150 subunit of Chromosome Assembly Factor-1 (CAF-1) (Smith et al. 2014; Matheson and Kaufman 2017), and the *Kcnq1ot1* (Mohammad et al. 2008) and Firre (Yang et al. 2015) long non-coding RNAs.

Evaluation of the nucleolar association of “NAD splitting regions” in centromere-proximal regions.

We have analyzed the localization of two centromere-proximal probes, pPK1006 and pPK1007, which overlap sequences identified as “NAD-splitting regions (NSRs)” by *NADfinder* (Fig. 2C, D). NSRs are regions between adjacent NADs; previous analysis of these in mouse embryonic fibroblasts indicate that these display much greater levels of gene expression than neighboring NAD sequences (Vertii et al. 2019). However, both of these probes displayed significant nucleolar association in immuno-FISH experiments (Fig. 3). We note that both pPK1006 and pPK1007 cover regions with much fewer nucleolar reads than neighboring NADs, and therefore appear as “valleys”, especially in the case of pPK1007 (Fig. 2C, D). These observations led us to consider whether background correction could have eliminated positive regions, such as pPK1006 and pPK1007, which are much less frequently associated than neighboring sequences, but are still more frequently associated than the genome-wide average. However, neither pPK1006 nor pPK1007 become positive when background correction is omitted from peak calling by *NADfinder* (Fig. 2C, D). Therefore, we hypothesize that discrepancies could arise from the biochemical purification of nucleoli inherent to our approach. That is, it is possible that centromere-proximal NSR loci represent sonication-sensitive chromatin loops that are frequently lost during purification, but which indeed are often nucleolar-associated in intact cells. Comprehensive testing of this idea would best come from future studies using an orthogonal method for identifying NADs that does not rely on biochemical purification of sonicated nucleoli.

Anomalies of MEF-specific NADs in stem cells.

One question of interest is whether nucleolar association leads to, or is a consequence of, transcriptional repression. Notably, previous studies have shown that tethering of loci to the nucleolar periphery via 5S rDNA sequences results in transcriptional silencing (Fedoriw et al. 2012b), so at least in that case a causal relationship has been established.

In MEF cells, genes in the MEF-specific NADs display mean expression levels lower than genes in the F121-9-specific NADs ($p < 0.0001$) (Fig. 5C, E). This is the expected situation, in which genes that had been in NADs earlier in development (e.g. in stem cells) become derepressed if that localization is lost. In contrast, in F121-9 cells, genes within MEF-specific NADs showed similar transcript levels as genes within F121-9-specific NADs ($p = 0.82$, Fig. 5D); the same was true in other mES cells analyzed ($p = 0.13$, Fig. 5F). Why aren't the MEF-specific NADs more transcriptionally active in stem cells, since they haven't yet acquired nucleolar association? This could be due to other repressive mechanisms acting on regions within MEF-specific NADs, for example, lamina association: 40% of MEFs-specific NADs overlap with cLADs (Fig. 4A). Alternatively, additional factors contributing to transcriptional repression may precede (and perhaps contribute to) nucleolar association. Development of reagents allowing control of perinucleolar associations will be key to exploring the relationship between nucleolar localization and transcriptional repression.

Materials and Methods

F121-9 mESC cell culture and isolation of crosslinked nucleoli.—The F121 mouse embryonic stem cell (mESC) line is a female cell line derived from a cross between a male *Castaneus* and a female *129* mouse (Monkhorst et al. 2008), and F121-9 was subcloned subsequently (Rivera-Mulia et al. 2018). F121-9 cells were obtained from the Gilbert lab (Florida State University) at passage 8. The cells were grown on gelatin-coated plates and cultured in 2i medium. Accutase (EMD Millipore SF006) was used to detach cells from plates and passage into new dishes. Prior to seeding cells, dishes were coated with 0.1% gelatin (EMD Millipore, SF008) for at least 25 min at room temperature, after which gelatin was aspirated. Dishes were rinsed with DPBS (Gibco, 14190144), which was aspirated, and cells were seeded in these dishes. 2i medium was obtained as described previously (Vertii et al. 2019). Cells were passaged at $3 \times 10^4/\text{cm}^2$ density. 2X HyCryo-STEM cryopreservation medium (GE Healthcare, SR30002.02) was used to freeze cells. For each preparation of nucleoli, cells were grown in eleven 15-cm plates and harvested one or two days after seeding them, with total cell numbers of $3\text{--}5 \times 10^8$ per preparation. One hour prior to nucleoli isolation, old cell culture medium was replaced with fresh medium. Plates grown in parallel were used for genomic DNA extraction (DNeasy Blood & Tissue kit, Qiagen), and RNA extraction (TRIZol, ThermoFisher Scientific and RNeasy mini kit, Qiagen). Crosslinking followed by isolation of nucleoli was done as described previously (Vertii et al. 2019)).

Quantitative PCR.—DNA was extracted from input whole cells and purified nucleoli using DNeasy Blood & Tissue kit (Qiagen). Quantitative PCR analysis was done as outlined previously (Vertii et al. 2019).

Antibodies.—The following antibodies were used: fibrillarin (Abcam, ab5821), nucleophosmin (Abcam, ab10530), histone H3K27me3 (Active Motif, 39155), histone H3 (Abcam, ab1791) and actin (Sigma-Aldrich, A1978). Secondary antibody for immunofluorescence was Alexa 594-conjugated donkey anti-rabbit (ThermoFisher, A-21207) and Alexa 594-conjugated goat anti-mouse (ThermoFisher, A-11020). For western blots, horseradish peroxidase (HRP) anti-mouse and anti-rabbit secondary antibodies (Jackson ImmunoResearch) were used.

Immunoblotting.—Proteins from total cell lysates and purified nucleoli were analyzed as noted previously (Vertii et al. 2019). Proteins from the same fractions were analyzed on 17% gels and transferred to nitrocellulose (replicate #1) or PVDF (replicate #2) membranes in order to detect H3K27me3 and total H3 protein levels. Ratios of H3K27me3 bands to total H3 bands were calculated by estimating adjusted volumes of bands in Bio-Rad Image Lab software.

DNA isolation, deep sequencing, and read preprocessing and mapping.—Total genomic and nucleolar DNA was purified using DNeasy Blood & Tissue kit (Qiagen). Libraries were generated using NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs). The DNA was fragmented to a size of 350 bp, and these fragments were size selected with sample purification beads. 150 bp paired-end sequencing was performed using Illumina reagents. 52.1 and 51.5 million reads were obtained for two replicates of genomic samples, and 49.4 and 52.8 million reads were obtained for two replicates of nucleolar samples. >95% of nucleolar samples, and >96% of genomic samples were mappable. From nucleolus preparation replicate #1, 13.9 million reads were assigned to the *Castaneus* genome and 14.8 million reads were assigned to the *I29* genome. From replicate #2, 15.1 million reads were assigned to the *Castaneus* genome and 16.1 million reads were assigned to the *I29* genome. The sequencing data files are available at data.4dnucleome.org under accession numbers 4DNESXE9K9DB, 4DNESUJZ5FL2. Trimming and alignment of mapped reads to the mouse genome (mm10) was done as previously described (Vertii et al. 2019).

RNA isolation, deep sequencing, and read preprocessing and mapping.—Total RNA from the two replicate preparations of F121-9 mESC were extracted using TRIzol (ThermoFisher Scientific) and purified using the RNeasy mini kit (Qiagen). Libraries were constructed using the NEBNext Ultra II RNA Library Prep kit for Illumina (New England Biolabs). The mRNA was fragmented, and double-stranded cDNA library synthesized, and completed through size selection and PCR enrichment. 150 bp paired-end sequencing was achieved using Illumina HiSeq 4000 platform. 22.2 and 26.7 million reads were obtained from the two replicates of mESC RNA. >92% of replicate 1, and >86% of replicate 2 were mappable. These sequencing data files are available at data.4dnucleome.org under accession number 4DNESDHILYLU. The quality of the sequencing reads was evaluated with fastqc (0.11.5) (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.) The paired-end reads were aligned to the mouse genome (ensemble GRCm38) using STAR (version 2.5.3a) with ENCODE standard options as `--outFilterMultimapNmax 20, --alignSJoverhangMin 8, --alignSJDBoverhangMin 1, --outFilterMismatchNmax 999, --alignIntronMin 20, --`

alignIntronMax 1000000, and --alignMatesGapMax 1000000. Additional parameter settings are --outFilterMismatchNoverReadLmax 0.04 and --outSAMattributes NH HI NM MD. To visualize the mapped reads, bigwig files were generated using the bamCoverage function in deepTools2 with the parameter setting --normalizeUsingRPKM.

DNA-FISH probes.—The bacterial artificial chromosomes (BACs) were obtained from the BACPAC Resource Center of Children’s Hospital Oakland Research Institute (Oakland, CA). DNA was isolated using BAC DNA miniprep Kit (Zymo Research). BAC probes were labeled using BioPrime Labeling Kit (ThermoFisher). Streptavidin, Alexa Fluor 488 conjugate (ThermoFisher, S-32354) was used to stain biotin-labeled BAC probes. Probes are described in Supplemental Table 2.

3-D DNA FISH/ immunocytochemistry and microscopy.—3-D DNA FISH/ immunocytochemistry-labeling was performed as described previously (Vertii et al. 2019), except that DNA FISH-labeling was done after immunocytochemistry, and coverslips were not treated with RNA removal solution. F121-9 mESC were seeded on 0.1% gelatin-coated 22 × 22 mm coverslips (Corning, 2850–22), with total cell number 150–250 × 10³ cells/ coverslip, and permeabilized and fixed the next day. Nucleoli were stained with anti-fibrillarin antibodies, except anti-nucleophosmin antibodies were used in the third biological replicates of the pPK999 and pPK1000 and all three biological replicates of pPK1006, pPK1007 and pPK1012 immuno-FISH assays. Cells were fixed with 4% Paraformaldehyde in 1X PBS for 10 min at RT for the second time at the end of immunocytochemistry part of the assay, before proceeding to DNA FISH-labeling in order to preserve nucleolar staining. Images were acquired using Zeiss LSM 700 laser scanning confocal microscope and PMT detector (63× 1.40 Oil DIC M27 Plan-Apochromat objective). DNA-FISH probes were counted through z-stacks manually and scored as “associated” if there was no gap between the probe and the nucleolar marker. Each probe was analyzed in at least three biological replicates, with at least 100 alleles scored in each replicate. Z stacks are represented as 2D maximum projections using Fiji software (Schindelin et al. 2012). R, a system for statistical computation and graphics (Ihaka and Gentleman 1996), was used for the analysis of FISH data. Percentage data was first arcsin transformed to homogenize the variance. Levene’s test indicated that the assumption of homogeneity of variances was met. One-way analysis of variance (ANOVA) with Completely Randomized Design was performed using the lsmeans package to test whether there are significant differences among the probes in question. Comparisons are declared statistically significant if Hochberg-adjusted p-value < 0.05 (Hochberg 1988).

Chromosome paints FISH/immunocytochemistry.—Whole chromosome paint probes for Chr 5 (MetaSystems Probes, D-1405–050-FI), Chr 10 (MetaSystems Probes, D-1410–050-FI) and Chr 12 (MetaSystems Probes, D-1412–050-FI) directly conjugated to a green emitting fluorochrome were purchased from MetaSystems Probes and used according to manufacturer’s directions with some modifications. Briefly, after immunocytochemistry, coverslips were processed for DNA FISH-labeling as described previously (Vertii et al. 2019). 5–7 µl of chromosome paint probes were denatured at 75°C for 2 min, after which the probes were hybridized with separately denatured coverslips in a humidified chamber at

37°C overnight. The next day coverslips were washed in 0.4X SSC, pH 7.0–7.5 at 72°C for 2 min, then drained and washed in 2X SSC, 0.05% Tween-20, pH 7.0 at RT for 30 s. The coverslips were rinsed in 1X PBS, and stained with DAPI at RT for 2 min. After rinsing the coverslips in 1X PBS and distilled water, they were mounted on slides with ProLong Gold antifade mountant (ThermoFisher Scientific, P36934). Microscopy and analysis of imaging were done as for the BAC probes, except that chromosomes were counted as contacting nucleoli if chromosome paint probe and nucleolus had at least 3 pixels of overlapping signal. This quantitation criterion was based on chromosome paint immuno-FISH experiments described previously (Strongin et al. 2014)

NAD identification and annotation.—We used the same workflow for NAD-seq data analysis as described previously (Vertii et al. 2019), except that we removed 20 NAD peaks that are less than 50 kb long (totaling 0.74 MB). Because there are 624 peaks totaling 845 Mb in the F121-9 NAD-seq data, this represents 0.087% of the NAD nucleotides. We used version 1.6.1 of *NADfinder* for NAD identification in this manuscript. To identify allele-specific NADs, first we performed a single alignment to the N-masked mouse genome mm10 using *bowtie2/2.3.2* and then assigned the reads to one of the two genomes based on the SNP found in the masked positions using *SNPsplit_v0.3.2* for allele-specific splitting of alignments (Krueger and Andrews 2016).

Overlap analyses.—Nucleotide-level overlap analyses of F121-9 NADs with cLADs, ciLADs (Peric-Hupkes et al. 2010), MEF NADs (Vertii et al. 2019), and H3K27me3-enriched domains (GSM2416833; (Cruz-Molina et al. 2017); GSM1621022; (Delbarre et al. 2017)) were performed using *GenomicRanges* (Lawrence et al. 2013) as described in detail in Vertii et al., 2019. These nucleotide-based overlap analyses in some cases generated small overlapped regions, such that single genes would end up with both Type I and Type II designations, or both MEF-specific and F121-9-specific designations. Because the biology of NADs is centered on large (~1 MB-sized) domains, we removed regions <50 kb in length from overlap analyses of Type I and II NADs and from cell-type-specific NADs to avoid these confounding designations. GO enrichment analyses of conserved and cell type-specific NADs derived from the overlap analysis were performed using *ChIPpeakAnno* (Zhu et al. 2010; Zhu 2013). mESC H3K27me3-enriched domains were identified based on H3K27me3 ChIP-seq data (GSM2416833; Cruz-Molina et al. 2017) using *RSEG* (v0.4.9) with 20 iterations for Baum training. MEF H3K27me3-enriched domains were obtained from GSM1621022 (Delbarre et al. 2017). FPKM values based on MEF RNA-seq data were obtained from GSM1621026 (Delbarre et al. 2017) and GSE90894 (Chronis et al. 2017). FPKM values from mES RNA-seq data were obtained from GSM1418813 (Lowe et al. 2015). Calculations of the statistical significance of pairwise comparisons were performed using Welch's t-test in GraphPad Prism.

The *NADfinder* software is available at: <https://bioconductor.org/packages/release/bioc/vignettes/NADfinder/inst/doc/NADfinder.html>

We calculated Jaccard indexes among NADs, cLAD/ciLAD (Peric-Hupkes et al. 2010), and F121-9 early/late replication timing (GSE95091 (Marchal et al. 2018)). The Jaccard index is the size of the intersect divided by the size of the union of two sets. The higher the Jaccard

index, the higher the extent of the overlap. Boxplots and comparisons of gene densities (genes/Mb) and gene expression distributions were performed using R For statistical comparisons, p-values were calculated using Welch's t-test.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank Takayo Sasaki and David Gilbert (Florida State University) for the kind gift of F121-9 cells, and Anastasiia Vertii for guidance with the nucleolar preparation and DNA-FISH experiments.

Funding: Research reported in this publication was supported by the National Institutes of Health (National Institute on Drug Abuse, U01 DA040588 to P.D.K.) as part of the 4D Nucleome Consortium.

References cited

- Ahmed K, Deghani H, Rugg-Gunn P, et al. (2010) Global Chromatin Architecture Reflects Pluripotency and Lineage Commitment in the Early Mouse Embryo. *PLoS One* 5:e10531. doi: 10.1371/journal.pone.0010531 [PubMed: 20479880]
- Aloia L, Di Stefano B, Di Croce L (2013) Polycomb complexes in stem cells and embryonic development. *Development* 140:2525–2534. doi: 10.1242/dev.091553 [PubMed: 23715546]
- Atlasi Y, Stunnenberg HG (2017) The interplay of epigenetic marks during stem cell differentiation and development. *Nat Rev Genet* 18:643–658. doi: 10.1038/nrg.2017.57 [PubMed: 28804139]
- Bhattacharya D, Talwar S, Mazumder A, Shivashankar GV (2009) Spatio-temporal plasticity in chromatin organization in mouse cell differentiation and during *Drosophila* embryogenesis. *Biophys J* 96:3832–9. doi: 10.1016/j.bpj.2008.11.075 [PubMed: 19413989]
- Borsos M, Perricone SM, Schauer T, et al. (2019) Genome-lamina interactions are established de novo in the early mouse embryo. *Nature* 569:729–733. doi: 10.1038/s41586-019-1233-0 [PubMed: 31118510]
- Brangwynne CP, Mitchison TJ, Hyman AA (2011) Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc Natl Acad Sci* 108:4334–4339. doi: 10.1073/pnas.1017150108 [PubMed: 21368180]
- Britton-Davidian J, Cazaux B, Catalan J (2012) Chromosomal dynamics of nucleolar organizer regions (NORs) in the house mouse: Micro-evolutionary insights. *Heredity (Edinb)* 108:68–74. doi: 10.1038/hdy.2011.105 [PubMed: 22086078]
- Chronis C, Fiziev P, Papp B, et al. (2017) Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* 168:442–459.e20. doi: 10.1016/j.cell.2016.12.016 [PubMed: 28111071]
- Cruz-Molina S, Respuela P, Tebartz C, et al. (2017) PRC2 Facilitates the Regulatory Topology Required for Poised Enhancer Function during Pluripotent Stem Cell Differentiation. *Cell Stem Cell* 20:689–705.e9. doi: 10.1016/j.stem.2017.02.004 [PubMed: 28285903]
- Delbarre E, Ivanauskienė K, Spirkoski J, et al. (2017) PML protein organizes heterochromatin domains where it regulates histone H3.3 deposition by ATRX/DAXX. *Genome Res* 27:913–921. doi: 10.1101/gr.215830.116 [PubMed: 28341773]
- Dillinger S, Straub T, Nemeth A (2017) Nucleolus association of chromosomal domains is largely maintained in cellular senescence despite massive nuclear reorganisation. *PLoS One* 12:1–28. doi: 10.1371/journal.pone.0178821
- Efroni S, Duttagupta R, Cheng J, et al. (2008) Global Transcription in Pluripotent Embryonic Stem Cells. *Cell Stem Cell* 2:437–447. doi: 10.1016/j.stem.2008.03.021 [PubMed: 18462694]
- ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74. doi: 10.1038/nature11247 [PubMed: 22955616]

- Fedoriw AM, Calabrese JM, Mu W, et al. (2012a) Differentiation-driven nucleolar association of the mouse imprinted *Kcnq1* locus. *G3 (Bethesda)* 2:1521–8. doi: 10.1534/g3.112.004226 [PubMed: 23275875]
- Fedoriw AM, Starmer J, Yee D, Magnuson T (2012b) Nucleolar association and transcriptional inhibition through 5S rDNA in mammals. *PLoS Genet* 8:. doi: 10.1371/journal.pgen.1002468
- Feric M, Vaidya N, Harmon TS, et al. (2016) Coexisting Liquid Phases Underlie Nucleolar Subcompartments. *Cell* 165:1686–1697. doi: 10.1016/j.cell.2016.04.047 [PubMed: 27212236]
- Gaspar-Maia A, Alajem A, Meshorer E, Ramalho-Santos M (2011) Open chromatin in pluripotency and reprogramming. *Nat Rev Mol Cell Biol* 12:36–47. doi: 10.1038/nrm3036 [PubMed: 21179060]
- Giorgetti L, Lajoie BR, Carter AC, et al. (2016) Structural organization of the inactive X chromosome in the mouse. *Nature* 535:575–579. doi: 10.1038/nature18589 [PubMed: 27437574]
- Guelen L, Pagie L, Brasset E, et al. (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453:948–951. doi: 10.1038/nature06947 [PubMed: 18463634]
- Hannemann F, Bichet A, Ewen KM, Bernhardt R (2007) Cytochrome P450 systems—biological variations of electron transport chains. *Biochim Biophys Acta - Gen Subj* 1770:330–344. doi: 10.1016/j.bbagen.2006.07.017
- Harr JC, Luperchio TR, Wong X, et al. (2015) Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J Cell Biol* 208:33–52. doi: 10.1083/jcb.201405110 [PubMed: 25559185]
- Hawkins RD, Hon GC, Lee LK, et al. (2010) Distinct Epigenomic Landscapes of Pluripotent and Lineage-Committed Human Cells. *Cell Stem Cell* 6:479–491. doi: 10.1016/j.stem.2010.03.018 [PubMed: 20452322]
- Hiratani I, Ryba T, Itoh M, et al. (2010) Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. *Genome Res* 20:155–69. doi: 10.1101/gr.099796.109 [PubMed: 19952138]
- Hochberg Y (1988) A Sharper Bonferroni Procedure for Multiple Tests of Significance. *Biometrika* 75:800. doi: 10.2307/2336325
- Ihaka R, Gentleman R (1996) R: A Language for Data Analysis and Graphics. *J Comput Graph Stat* 5:299–314. doi: 10.1080/10618600.1996.10474713
- Jakociunas T, Domange Jordö M, Ait Mebarek M, et al. (2013) Subnuclear relocalization and silencing of a chromosomal region by an ectopic ribosomal DNA repeat. *Proc Natl Acad Sci U S A* 110:E4465–73. doi: 10.1073/pnas.1315581110 [PubMed: 24191010]
- Kind J, Pagie L, Ortobozkoyun H, et al. (2013) Single-cell dynamics of genome-nuclear lamina interactions. *Cell* 153:178–192. doi: 10.1016/j.cell.2013.02.028 [PubMed: 23523135]
- Krueger F, Andrews SR (2016) SNPsplit: Allele-specific splitting of alignments between genomes with known SNP genotypes [version 2; referees: 3 approved]. *F1000Research* 5:. doi: 10.12688/F1000RESEARCH.9037.2
- Kurihara Y, Suh D-S, Suzuki H, Moriwaki K (1994) Chromosomal locations of Ag-NORs and clusters of ribosomal DNA in laboratory strains of mice. *Mamm Genome* 5:225–228. doi: 10.1007/BF00360550 [PubMed: 8012113]
- Larson AG, Elnatan D, Keenen MM, et al. (2017) Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* 547:236–240. doi: 10.1038/nature22822 [PubMed: 28636604]
- Lavarone E, Barbieri CM, Pasini D (2019) Dissecting the role of H3K27 acetylation and methylation in PRC2 mediated control of cellular identity. *Nat Commun* 10:1679. doi: 10.1038/s41467-019-09624-w [PubMed: 30976011]
- Lawrence M, Huber W, Pagès H, et al. (2013) Software for Computing and Annotating Genomic Ranges. *PLoS Comput Biol* 9:e1003118. doi: 10.1371/journal.pcbi.1003118 [PubMed: 23950696]
- Lee J-H, Hart SRL, Skalnik DG (2004) Histone deacetylase activity is required for embryonic stem cell differentiation. *genesis* 38:32–38. doi: 10.1002/gene.10250 [PubMed: 14755802]

- Lowe R, Gemma C, Rakyan VK, Holland ML (2015) Sexually dimorphic gene expression emerges with embryonic genome activation and is dynamic throughout development. *BMC Genomics* 16:295. doi: 10.1186/s12864-015-1506-4 [PubMed: 25888192]
- Marchal C, Sasaki T, Vera D, et al. (2018) Genome-wide analysis of replication timing by next-generation sequencing with E/L Repli-seq. *Nat Protoc* 13:819–839. doi: 10.1038/nprot.2017.148 [PubMed: 29599440]
- Martens JHA, O’Sullivan RJ, Braunschweig U, et al. (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J* 24:800–12. doi: 10.1038/sj.emboj.7600545 [PubMed: 15678104]
- Matheson TD, Kaufman PD (2017) The p150N domain of chromatin assembly factor-1 regulates Ki-67 accumulation on the mitotic perichromosomal layer. *Mol Biol Cell* 28:21–29. doi: 10.1091/mbc.e16-09-0659 [PubMed: 27807046]
- Matheson TD, Kaufman PD (2016) Grabbing the genome by the NADs. *Chromosoma* 125:361–371. doi: 10.1007/s00412-015-0527-8 [PubMed: 26174338]
- Mattout A, Cabianca DS, Gasser SM (2015) Chromatin states and nuclear organization in development — a view from the nuclear lamina. *Genome Biol* 16:174. doi: 10.1186/s13059-015-0747-5 [PubMed: 26303512]
- McStay B (2016) Nucleolar organizer regions: Genomic ‘dark matter’ requiring illumination. *Genes Dev.* 30:1598–1610 [PubMed: 27474438]
- Meshorer E, Yellajoshula D, George E, et al. (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Dev Cell*. doi: 10.1016/j.devcel.2005.10.017
- Meuleman W, Peric-Hupkes D, Kind J, et al. (2013) Constitutive nuclear lamina-genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res* 23:270–80. doi: 10.1101/gr.141028.112 [PubMed: 23124521]
- Mitrea DM, Cika JA, Guy CS, et al. (2016) Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *Elife* 5:. doi: 10.7554/eLife.13571
- Mohammad F, Pandey RR, Nagano T, et al. (2008) Kcnq1ot1/Lit1 Noncoding RNA Mediates Transcriptional Silencing by Targeting to the Perinucleolar Region. *Mol Cell Biol* 28:3713–3728. doi: 10.1128/MCB.02263-07 [PubMed: 18299392]
- Monkhorst K, Jonkers I, Rentmeester E, et al. (2008) X Inactivation Counting and Choice Is a Stochastic Process: Evidence for Involvement of an X-Linked Activator. *Cell* 132:410–421. doi: 10.1016/j.cell.2007.12.036 [PubMed: 18267073]
- Németh A, Conesa A, Santoyo-Lopez J, et al. (2010) Initial genomics of the human nucleolus. *PLoS Genet* 6:. doi: 10.1371/journal.pgen.1000889
- Padeken J, Heun P (2014) Nucleolus and nuclear periphery: Velcro for heterochromatin. *Curr Opin Cell Biol* 28:54–60. doi: 10.1016/j.ceb.2014.03.001 [PubMed: 24690547]
- Padeken J, Mendiburo MJ, Chlamydas S, et al. (2013) The Nucleoplasmin Homolog NLP Mediates Centromere Clustering and Anchoring to the Nucleolus. *Mol Cell* 50:236–249. doi: 10.1016/j.molcel.2013.03.002 [PubMed: 23562326]
- Peric-Hupkes D, Meuleman W, Pagie L, et al. (2010) Molecular Maps of the Reorganization of Genome-Nuclear Lamina Interactions during Differentiation. *Mol Cell* 38:603–613. doi: 10.1016/j.molcel.2010.03.016 [PubMed: 20513434]
- Pickersgill H, Kalverda B, de Wit E, et al. (2006) Characterization of the *Drosophila melanogaster* genome at the nuclear lamina. *Nat Genet* 38:1005–1014. doi: 10.1038/ng1852 [PubMed: 16878134]
- Pontvianne F, Carpentier MC, Durut N, et al. (2016) Identification of Nucleolus-Associated Chromatin Domains Reveals a Role for the Nucleolus in 3D Organization of the *A. thaliana* Genome. *Cell Rep* 16:1574–1587. doi: 10.1016/j.celrep.2016.07.016 [PubMed: 27477271]
- Ragoczy T, Telling A, Scalzo D, et al. (2014) Functional redundancy in the nuclear compartmentalization of the Late-Replicating genome. *Nucleus* 5:626–635. doi: 10.4161/19491034.2014.990863 [PubMed: 25493640]

- Ribbeck K, Görlich D (2002) The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J* 21:2664–71. doi: 10.1093/emboj/21.11.2664 [PubMed: 12032079]
- Rivera-Mulia JC, Dimond A, Vera D, et al. (2018) Allele-specific control of replication timing and genome organization during development. *Genome Res* 28:800–811. doi: 10.1101/gr.232561.117 [PubMed: 29735606]
- Savi N, Bär D, Leone S, et al. (2014) lncRNA Maturation to Initiate Heterochromatin Formation in the Nucleolus Is Required for Exit from Pluripotency in ESCs. *Cell Stem Cell* 15:720–734. doi: 10.1016/j.stem.2014.10.005 [PubMed: 25479748]
- Schindelin J, Arganda-Carreras I, Frise E, et al. (2012) Fiji: an open-source platform for biological-image analysis. *Nat Methods* 9:676–682. doi: 10.1038/nmeth.2019 [PubMed: 22743772]
- Shin Y, Chang Y-C, Lee DSW, et al. (2018) Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome. *Cell* 175:1481–1491.e13. doi: 10.1016/J.CELL.2018.10.057 [PubMed: 30500535]
- Singh I, Contreras A, Cordero J, et al. (2018) MiCEE is a ncRNA-protein complex that mediates epigenetic silencing and nucleolar organization. *Nat Genet* 50:990–1001. doi: 10.1038/s41588-018-0139-3 [PubMed: 29867223]
- Smith CL, Matheson TD, Trombly DJ, et al. (2014) A separable domain of the p150 subunit of human chromatin assembly factor-1 promotes protein and chromosome associations with nucleoli. *Mol Biol Cell* 25:2866–2881. doi: 10.1091/mbc.e14-05-1029 [PubMed: 25057015]
- Strom AR, Emelyanov AV, Mir M, et al. (2017) Phase separation drives heterochromatin domain formation. *Nature* 547:241–245. doi: 10.1038/nature22989 [PubMed: 28636597]
- Strongin DE, Groudine M, Politz JCR (2014) Nucleolar tethering mediates pairing between the IgH and Myc loci. *Nucleus* 5:474–481. doi: 10.4161/nucl.36233 [PubMed: 25482199]
- Suzuki H, Kurihara Y, Kanehisa T, Moriwaki K (1990) Variation in the distribution of silver-staining nucleolar organizer regions on the chromosomes of the wild mouse, *Mus musculus*. *Mol Biol Evol* 7:271–82. doi: 10.1093/oxfordjournals.molbev.a040598 [PubMed: 1694258]
- Tatavosian R, Kent S, Brown K, et al. (2019) Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *J Biol Chem* 294:1451–1463. doi: 10.1074/jbc.RA118.006620 [PubMed: 30514760]
- Trojer P, Reinberg D (2007) Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Mol Cell* 28:1–13. doi: 10.1016/J.MOLCEL.2007.09.011 [PubMed: 17936700]
- van Koningsbruggen S, Gierlinski M, Schofield P, et al. (2010) High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Mol Biol Cell* 21:3735–48. doi: 10.1091/mbc.E10-06-0508 [PubMed: 20826608]
- van Steensel B, Belmont AS (2017) Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* 169:780–791. doi: 10.1016/j.cell.2017.04.022 [PubMed: 28525751]
- Vertii A, Ou J, Yu J, et al. (2019) Two contrasting classes of nucleolus-associated domains in mouse fibroblast heterochromatin. *Genome Res* gr.247072118. doi: 10.1101/gr.247072.118
- Wen B, Wu H, Shinkai Y, et al. (2009) Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat Genet* 41:246–250. doi: 10.1038/ng.297 [PubMed: 19151716]
- Yang F, Deng X, Ma W, et al. (2015) The lncRNA Firre anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. *Genome Biol* 16:52. doi: 10.1186/s13059-015-0618-0 [PubMed: 25887447]
- Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G (2004) CTCF Tethers an Insulator to Subnuclear Sites, Suggesting Shared Insulator Mechanisms across Species. *Mol Cell* 13:291–298. doi: 10.1016/S1097-2765(04)00029-2 [PubMed: 14759373]
- Zhang H, Wu Z, Lu JY, et al. (2020) DEAD-Box Helicase 18 Counteracts PRC2 to Safeguard Ribosomal DNA in Pluripotency Regulation. *Cell Rep* 30:81–97.e7. doi: 10.1016/j.celrep.2019.12.021 [PubMed: 31914400]

- Zhang LF, Huynh KD, Lee JT (2007) Perinucleolar Targeting of the Inactive X during S Phase: Evidence for a Role in the Maintenance of Silencing. *Cell* 129:693–706. doi: 10.1016/j.cell.2007.03.036 [PubMed: 17512404]
- Zhu LJ (2013) Integrative analysis of ChIP-chip and ChIP-seq dataset. *Methods Mol Biol* 1067:105–24. doi: 10.1007/978-1-62703-607-8_8 [PubMed: 23975789]
- Zhu LJ, Gazin C, Lawson ND, et al. (2010) ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* 11:237. doi: 10.1186/1471-2105-11-237 [PubMed: 20459804]
- Zullo JM, Demarco IA, Piqué-Regi R, et al. (2012) DNA sequence-dependent compartmentalization and silencing of chromatin at the nuclear lamina. *Cell* 149:1474–1487. doi: 10.1016/j.cell.2012.04.035 [PubMed: 22726435]

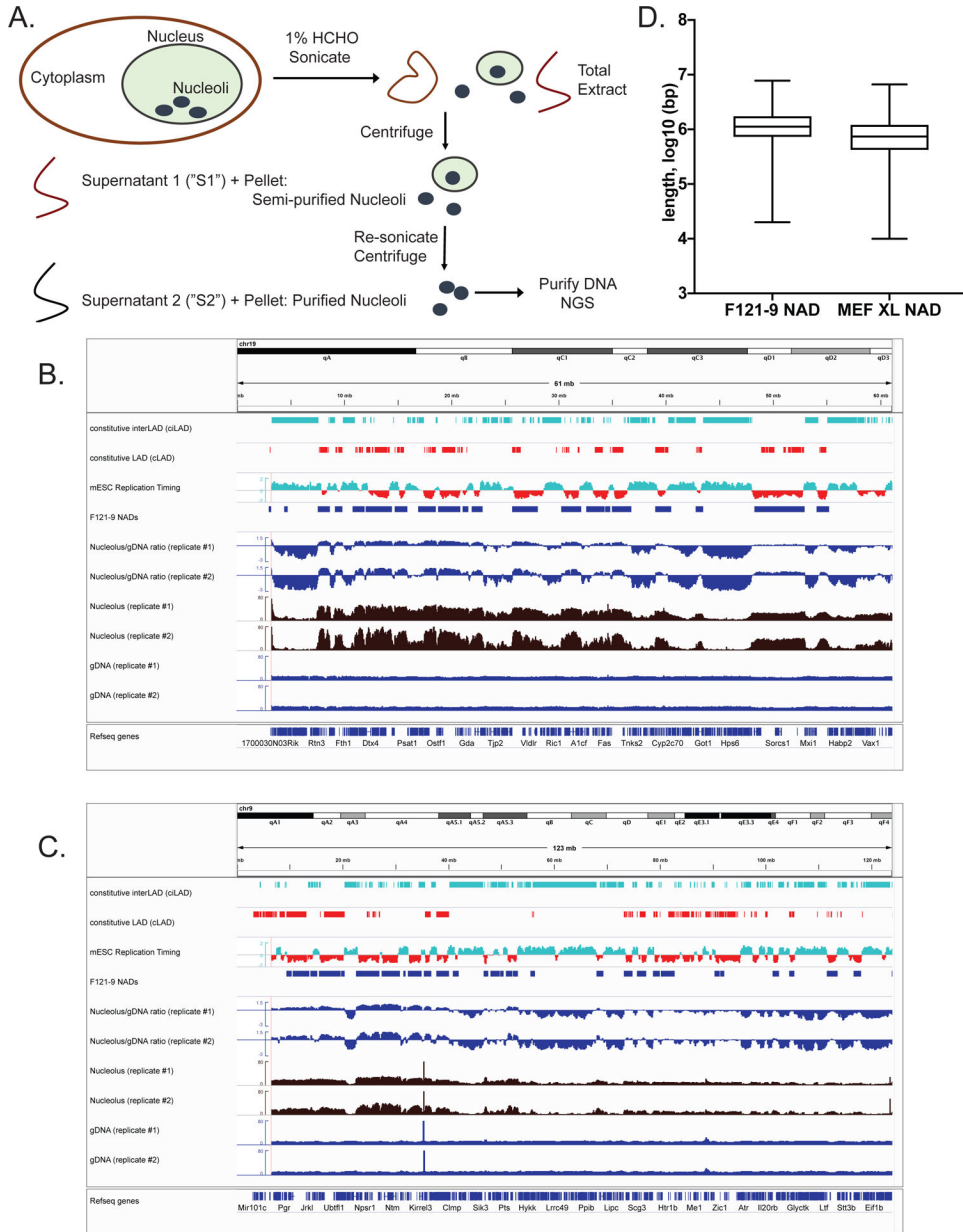


Figure 1. Analysis of F121-9 NAD sequencing data and comparison with heterochromatin.

A. Schematic diagram of nucleoli isolation from crosslinked cells.

B. All of chromosome 19 is shown, which contains strongly nucleoli-associated regions. From the top, tracks shown are: Constitutive interLADs (ciLADs, cyan) and Constitutive LADs (cLADs, red) (Peric-Hupkes et al. 2010); mESC replication timing (Hiratani et al. 2010, early replicating regions in cyan and late replicating regions in red); F121-9 cell NAD peaks (“F121-9 NADs”, called using *NADfinder* software based on two replicate experiments); Nucleolar/gDNA ratios, shown for both replicate experiments; raw read counts from both replicates for nucleoli-associated DNA (“Nucleolus”, brown) and total genomic DNA (“gDNA”, dark blue).

C. As in panel A, with all of chromosome 9 shown.

D. Length distribution of F121-9 NADs, compared to those from crosslinked MEF cells (Vertii et al. 2019).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

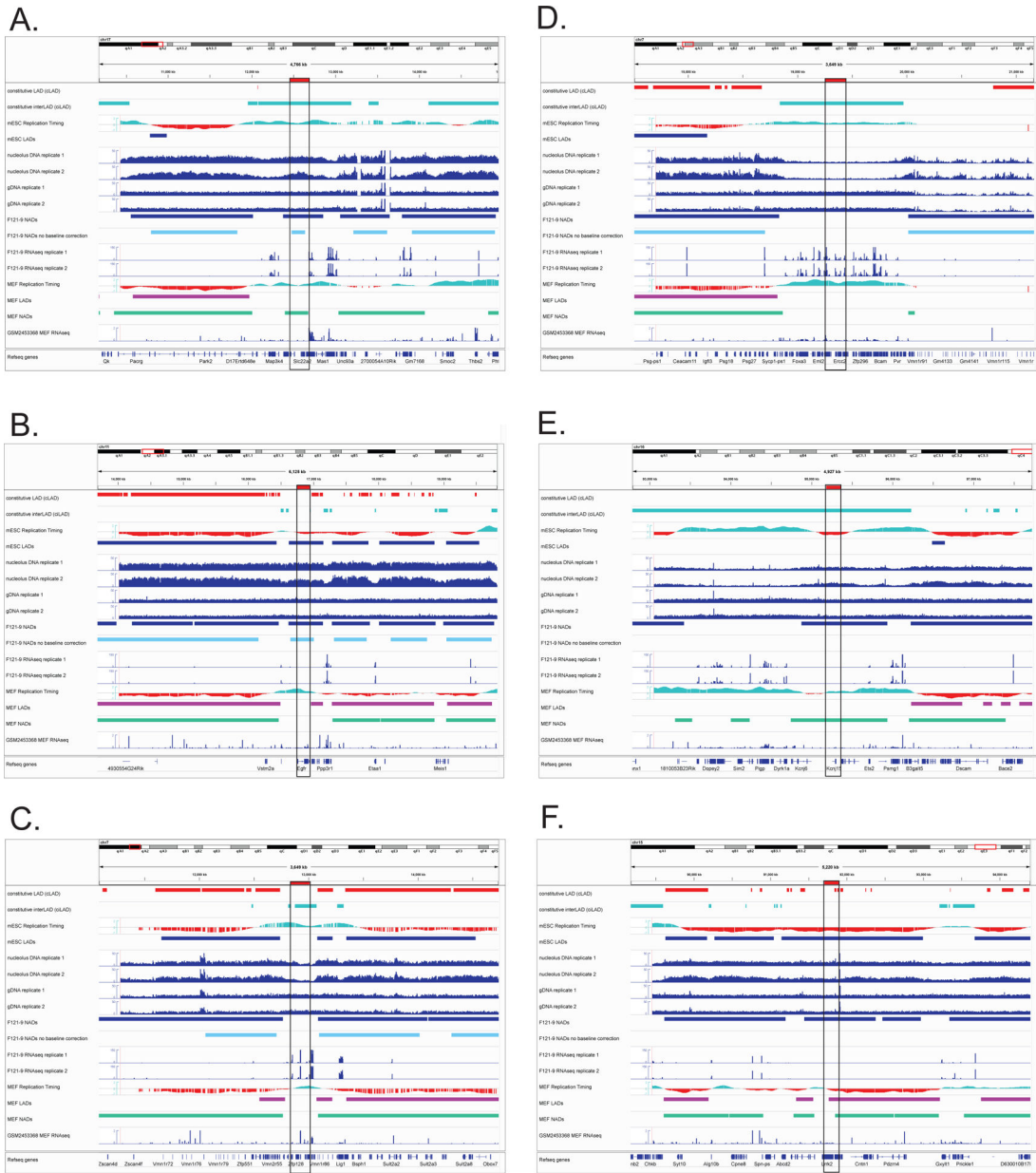


Figure 2. Genomic locations of BACs used for FISH experiments. For each panel, BAC locations are outlined by a black box and indicated with a red horizontal bar above the top track. From the top, tracks include cLADs (red) and ciLADs (cyan) (Peric-Hupkes et al. 2010), followed by mESC replication timing (Hiratani et al. 2010) and mESC LADs (Peric-Hupkes et al. 2010). The next tracks are raw read distributions from both replicate preparations of nucleolus-associated DNA (“nucleolus”) and total genomic DNA (“gDNA”); F121-9 NAD peaks called by *NADfinder* using background correction (dark blue) followed F121-9 NAD peaks predicted by *NADfinder* in the absence of background correction (light blue). These are followed by RNA-seq data from the same preparations of F121-9 cells used to generate the NAD data. At the bottom are data from MEF cells for comparison: replication timing (Hiratani et al. 2010), LADs

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

(Peric-Hupkes et al. 2010), NAD peaks from crosslinked cells (Vertii et al. 2019) and RNA-seq (GSM2453368 (ENCODE Project Consortium 2012)).

A. pPK915. This ciLAD-overlapped BAC is a NAD in both F121-9 and MEF cells, encoding solute carrier membrane transport proteins (Slc22a1, 2, 3) and plasminogen (Plg).

B. pPK999. This BAC overlaps a late-replicating LAD that contains the gene encoding epidermal growth factor receptor (Egfr). Note that in MEF cells this region is not identified as a NAD, is early replicating, and displays greater expression of Egfr.

C. pPK1006. This ciLAD-overlapping BAC is in a centromere-proximal region that exhibits reduced nucleolar reads compared to neighboring regions (see “Nucleolus” tracks), and was not classified as a peak by *NADfinder*.

D. pPK1007. This is another centromere-proximal, ciLAD-overlapping BAC that has low raw nucleolar read numbers.

E. pPK914. This BAC is within a region distal from the centromere. Its identification by *NADfinder* required the background correction feature. This region is within a NAD in both F121-9 and MEF cells, and its overlap with a ciLAD region (cyan) indicates a lack of lamina association in these cell types. This NAD contains ion channel genes (Kcnj6, Kcnj15) and Ets-family transcription factors (Erg, Ets2).

F. pPK1012. This is another centromere-distal BAC region that is called as a NAD only when local background correction method is used in *NADfinder* software.

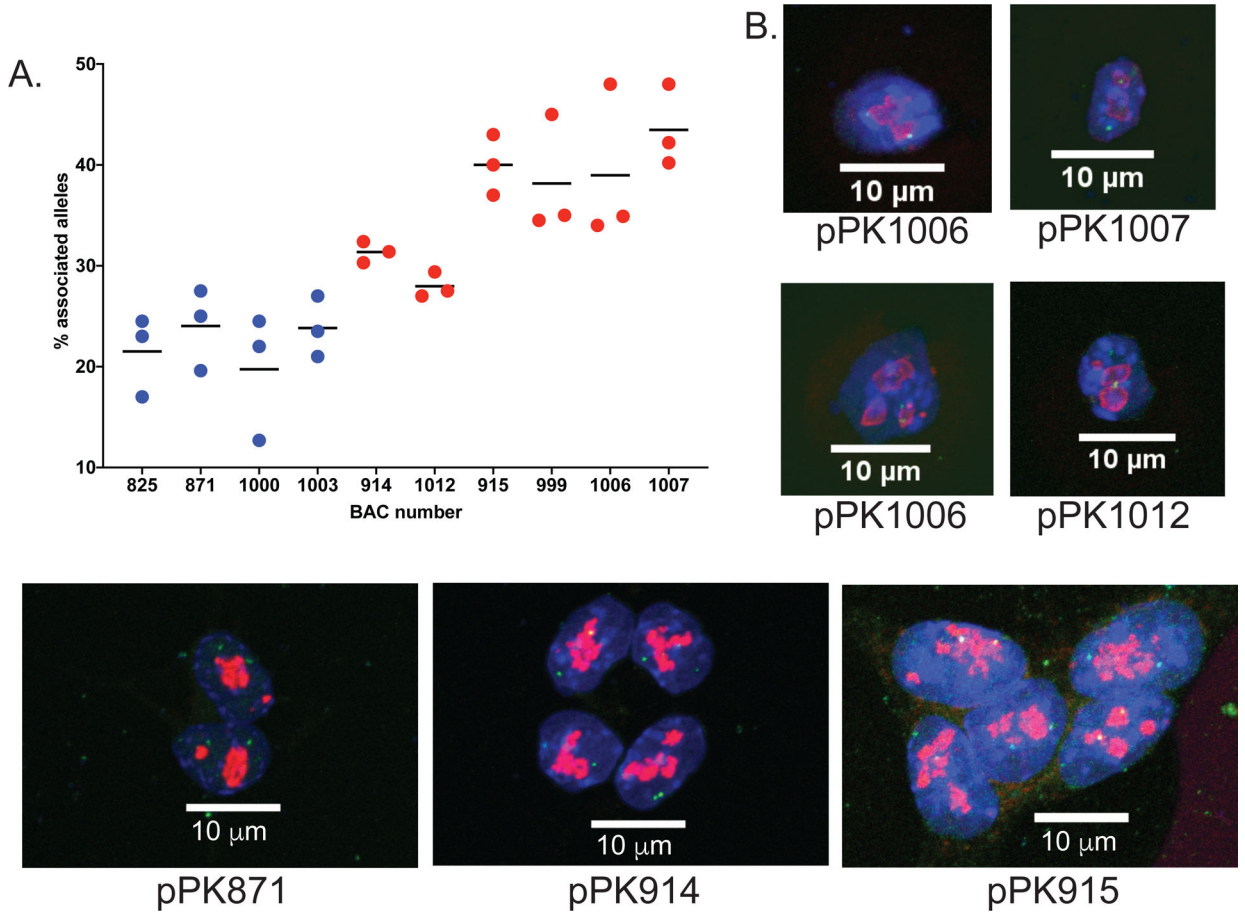


Figure 3. 3D DNA-FISH experiments validate nucleolar association of NADs in F121-9 mESC.

A. Graph of percentage of alleles that are nucleolar-associated (individual values and means are graphed for three biological replicates) for the indicated BAC probes (see Supplemental Table 2). The six probes on the left (red points) all displayed significantly more frequent associations than the non-NAD probes (blue points; $p < 0.0001$ for each red probe).

B. Maximum projection images from 3D immuno-FISH experiments with nuclear DAPI staining in blue, anti-fibrillarin (pPK871, pPK914 and pPK915, lower panels) or anti-nucleophosmin (pPK1006, pPK1007 and pPK1012, upper panels) antibody staining in red, and DNA probes in green. 63x magnification, scale bar 10 μm.

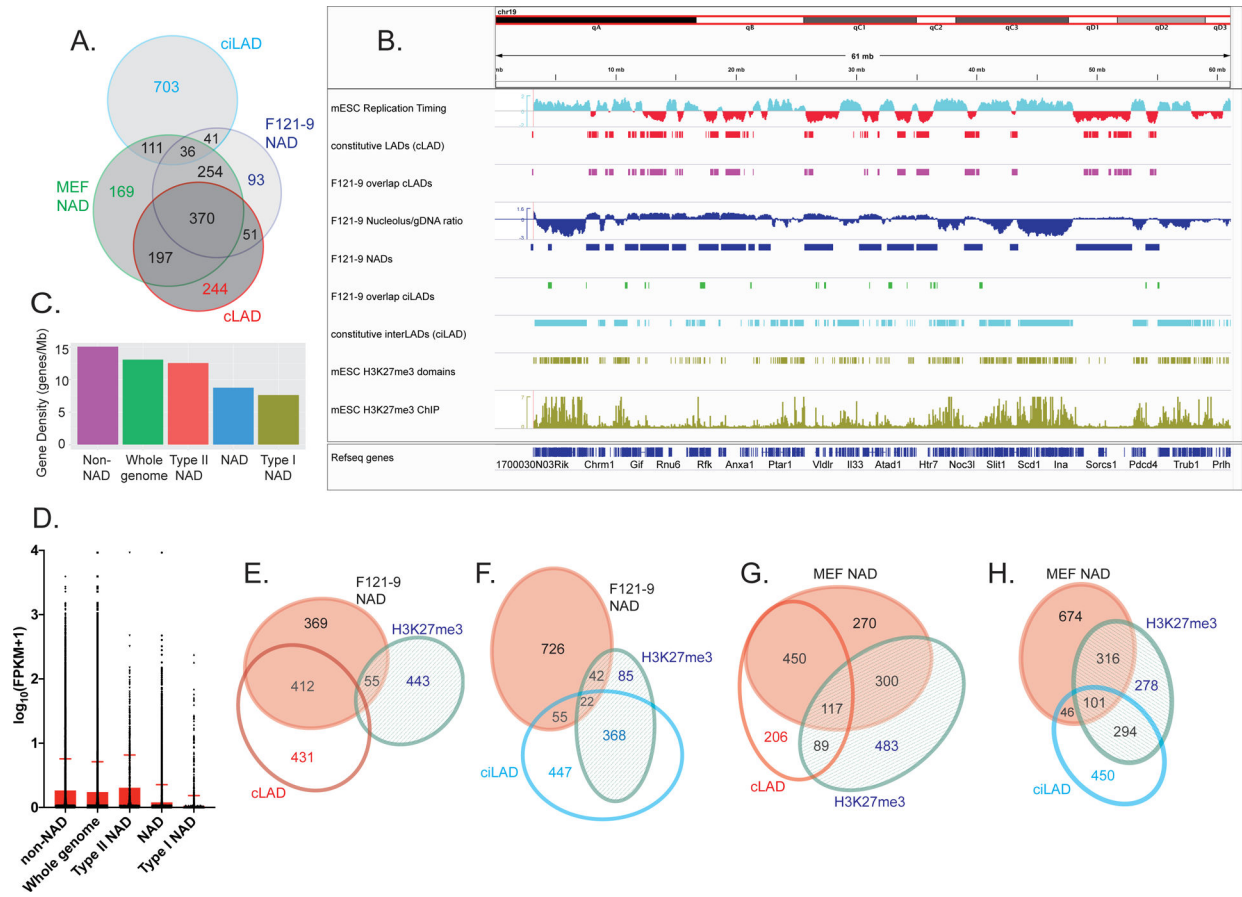


Figure 4. Two types of NADs in F121-9 mESC.

A. Venn diagram illustrating the overlaps among F121-9 NADs, MEF NADs (Vertii et al. 2019), cLAD, and ciLAD regions (Peric-Hupkes et al. 2010). Numbers show the size of the indicated regions in Mb.

B. Chromosomal view of F121-9 NADs overlapping cLADs and ciLADs. The entire chromosome 19 is shown. Euchromatic features (early replication timing, ciLAD) are displayed in cyan, and heterochromatic features (late replication timing, cLAD) are shown in red. From the top, displayed tracks are mESC replication timing (Hiratani et al. 2010), cLAD (Peric-Hupkes et al. 2010), NAD overlap with cLAD (i.e. Type I NADs, magenta), nucleolar/genomic ratio and NAD peaks (blue), NAD overlap with ciLAD (i.e. Type II NADs, green), ciLAD (Peric-Hupkes et al. 2010), H3K27me3 domains, and mESC H3K27me3 ChIP-seq data (Cruz-Molina et al. 2017) used for H3K27me3 domain identification (olive green).

C. Gene densities (genes/Mb) of the indicated regions, ranked left to right. “NAD” indicates all F121-9 NADs.

D. A box plot of gene expression levels from F121-9 RNA-seq data, expressed as $\log_{10}(\text{FPKM}+1)$ for the same indicated genomic regions as in panel C. The top of the red box indicates the mean value for each population, and the standard deviation is marked by the red error bar.

E. Venn diagram illustrating the overlaps among F121-9 NADs, cLADs (Peric-Hupkes et al. 2010) and mESC H3K27me3 domains (Cruz-Molina et al. 2017). Numbers indicate the size

of regions in Mb. The overlaps among all three sets (9 Mb) and between the cLAD and H3K27me3 sets (10 Mb) are left off the diagram because of their small sizes. Diagram was generated using eulerAPE 3.0 .

F. As in panel E, except here the overlap analysis includes ciLADs (Peric-Hupkes et al. 2010) instead of cLADs.

G. As in panel E, except here Venn diagram illustrates the overlaps among crosslinked MEF NADs (Vertii et al. 2019), cLADs (Peric-Hupkes et al. 2010) and MEF H3K27me3 domains (Delbarre et al. 2017).

H. As in panel G, except here the overlap analysis includes ciLADs (Peric-Hupkes et al. 2010) instead of cLADs.

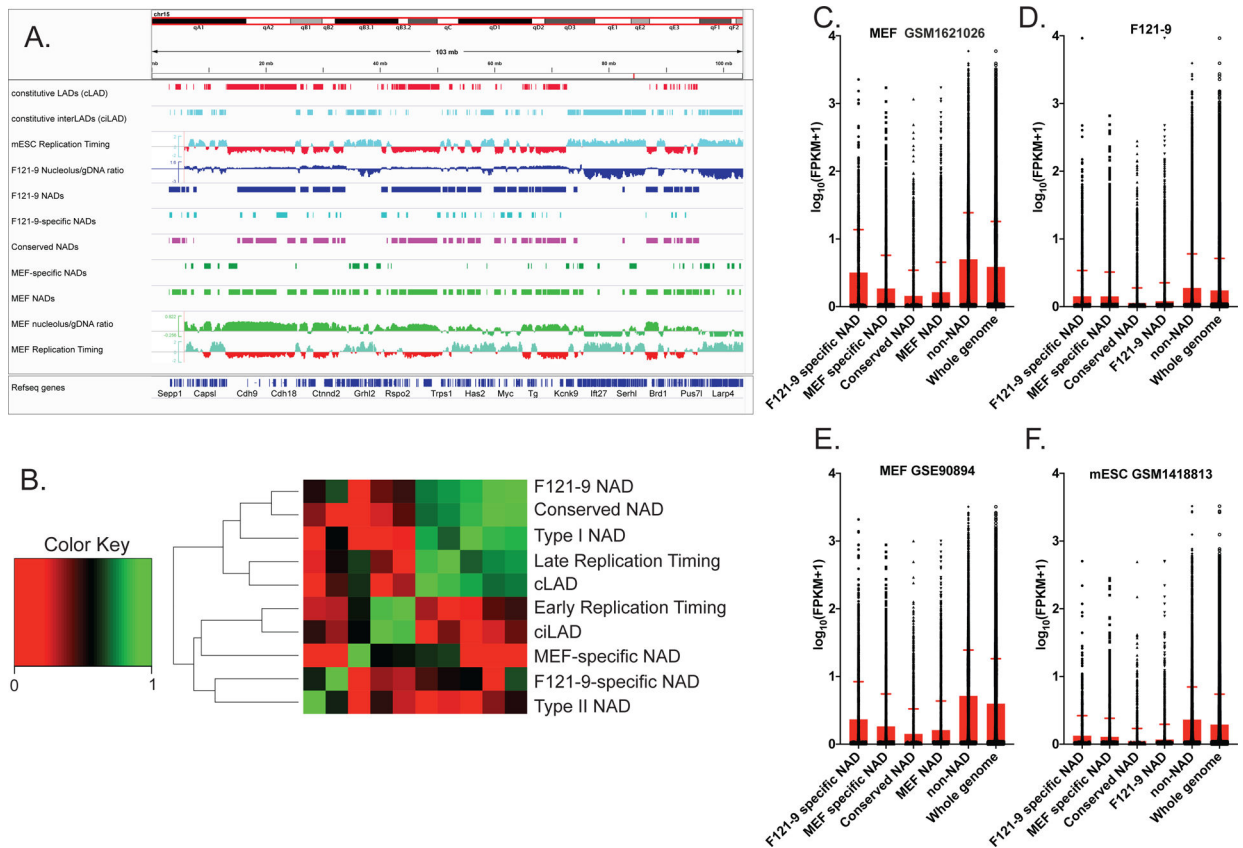


Figure 5. Conserved and cell type-specific NADs.

A. IGV browser view of entire chromosome 15. Euchromatic features (early replication timing, ciLAD) are displayed in cyan, and heterochromatic features (late replication timing, cLAD) are shown in red. From the top, tracks shown are cLAD, ciLAD (Peric-Hupkes et al. 2010), mESC replication timing (Hiratani et al. 2010), F121-9 nucleolar/genomic ratio and F121-9 NAD peaks (blue), “F121-9 specific NADs”, i.e. NADs found only in F121-9 cells (light blue), “conserved NADs”, or NADs shared between F121-9 and MEFs (magenta), “MEF-specific NADs” (dark green), MEF NAD peaks and MEF nucleolar/genomic ratio (Vertii et al. 2019) in green, and MEF replication timing (Hiratani et al. 2010).

B. Jaccard similarity coefficients were grouped based on similarities among the indicated regions. “F121-9 NAD” indicates all NADs identified in F121-9 cells in this study. “Conserved NAD” indicates NADs shared between F121-9 and MEF NADs (Vertii et al. 2019), whereas “F121-9-specific NAD” indicates NADs detected in F121-9, but not MEF cells. Conversely, “MEF-specific NAD” indicates NADs found in MEFs, but not in F121-9 cells. “Type I NAD” indicates F121-9 NADs that overlap with cLADs, and “Type II NAD” indicates F121-9 NADs that overlap with ciLADs (Peric-Hupkes et al. 2010). “cLAD” and “ciLAD” regions are from Peric-Hupkes et al. 2010, and F121-9 early replication timing and late replication timing regions are from Marchal et al. 2018. Note that F121-9 NADs, conserved F121-9 NADs, cLADs and Type I NADs are highly similar. In contrast, Type II NADs are most similar to F121-9-specific NADs.

C. A box plot of gene expression levels from MEF RNA-seq data (GSM1621026; Delbarre et al. 2017), expressed as $\log_{10}(\text{FPKM}+1)$ for the indicated subsets of NAD, non-NAD and

whole genome regions. The statistical significance of pairwise comparisons were all $p < 0.0001$ (Welch's t-test).

D. As in panel C, except our F121-9 RNA-seq data is used for FPKM analysis. The indicated pairwise comparisons were all statistically significant ($p < 0.0001$), except for that between F121-9 and MEF-specific NADs do not achieve statistical significance ($p = 0.82$).

E. As in panel C, except different MEF RNA-seq data (GSE90894; (Chronis et al. 2017)) was used for FPKM analysis. The changes between cell type-specific NADs achieve statistical significance ($p < 0.0001$, Welch's t-test).

F. As in panel C, except mESC RNA-seq data (GSM1418813; (Lowe et al. 2015)) is used for FPKM analysis. The changes between F121-9 and MEF-specific NADs do not achieve statistical significance ($p = 0.13$).

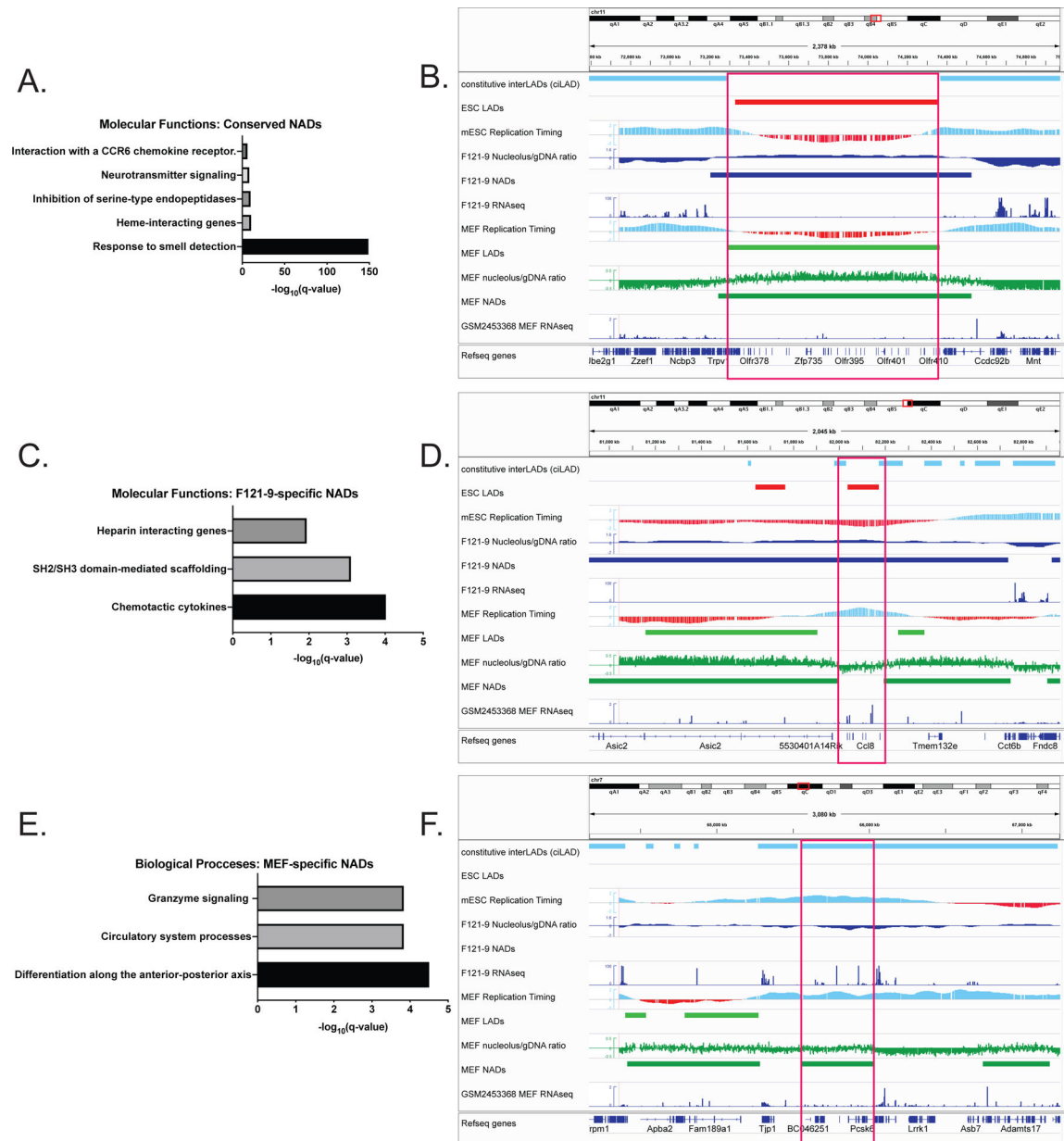


Fig. 6. GO analysis of conserved and cell type-specific NADs.

A. Molecular Functions subset of GO enrichment analysis of conserved NADs, with $-\log_{10}(q\text{-values})$ shown.

B. Genomic region containing NAD peak (red box) conserved in both MEF and F121-9 cells. This peak contains a cluster of olfactory genes on chromosome 11. ciLAD, mESC and MEF replication timing tracks are displayed as in Fig. 5B. The other tracks shown from the top are mESC LADs (Peric-Hupkes et al.2010; red), F121-9 nucleolar/genomic ratio, NADs and RNA-seq data (blue), MEF LADs (Peric-Hupkes et al. 2010), MEF nucleolar/genomic ratio, NADs (Vertii et al. 2019) (green) and RNA-seq (GSM2453368 (ENCODE Project Consortium 2012)) (blue).

C. Molecular Functions subset of GO enrichment analysis of F121-9-specific NADs.

D. As in panel B, showing genomic region corresponding to F121-9-specific NAD (red box), overlapping *Ccl* family of chemokine ligands.

E. Biological Functions subset of GO enrichment analysis of MEF-specific NADs.

F. As in panel B, showing genomic region containing MEF-specific NAD (red box), overlapping the *Pcsk6* gene important for differentiation along anterior-posterior axis.

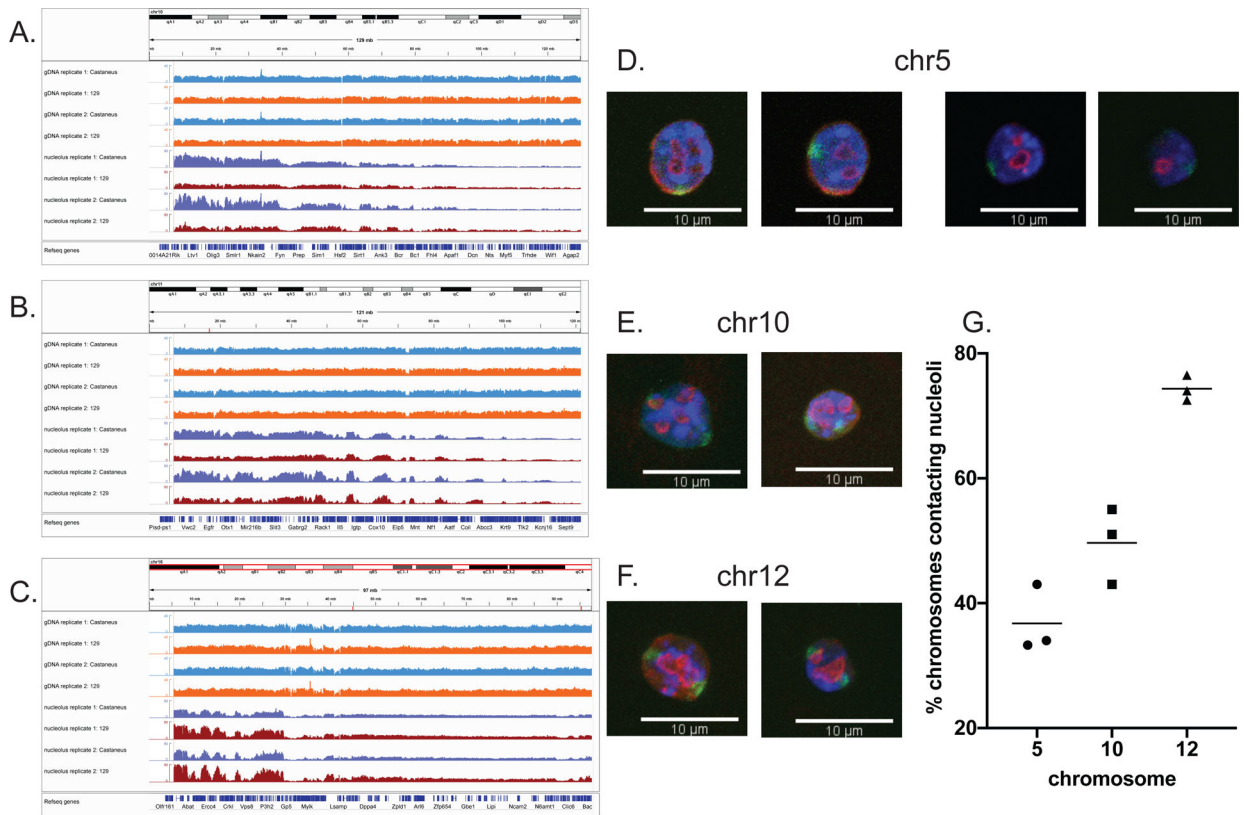


Figure 7. Allele-specific chromosome-nucleolus contacts.

IGV browser view of entire chromosome 10. Allele-specific genomic DNA (“gDNA”) reads assigned to the *Castaneus* genome are shown in light blue and genomic DNA reads assigned to the *129* genome are shown in orange. Nucleolus-associated *Castaneus* reads are in dark blue and those from the *129* genome are in red. Note the higher numbers of nucleolus-associated reads, especially in the centromere-proximal region, assigned to the *Castaneus* genome compared to the *129* genome.

As in panel A, except here chromosome 11 is shown. The *Castaneus* genome has more nucleolar read numbers in the centromere-proximal region than the *129* genome.

As in panel A, except here chromosome 16 is shown. In this case, the *129* genome has more nucleolar read numbers in the centromere-proximal region compared to the *Castaneus* genome.

D. Two pairs of individual z-stack slices from two different cells (left and right pairs of images) from 3D immuno-FISH experiments. Nuclear DAPI staining in blue, anti-nucleophosmin antibody staining in red, and whole chromosome paint probe for Chr 5 in green. 63x magnification, scale bar 10 μ m.

E. As in panel D, except here maximum projection images of two representative nuclei from 3D immuno-FISH experiments with the probe for Chr 10 are shown.

F. As in panel E, except here the probe for Chr 12 was used.

G. Graph of percentage of Chr 5, Chr 10 and Chr 12 territories contacting nucleoli (individual values and mean of three biological replicates). Chr 10 displays significantly different frequencies of nucleolar contacts than Chr 5 ($p = 0.079$) or Chr 12 ($p = 0.0014$).