A sub-set of guanine- and cytosine-rich genes are actively transcribed at the nuclear Lamin B1 region

Gayan I. Balasooriya[1], Tse-Luen Wee[1] and David L. Spector[1*]

[1]Cold Spring Harbor Laboratory, One Bungtown Road, Cold Spring Harbor, New York 11724, USA.

***Corresponding author: spector@cshl.edu**

Abstract

Chromatin organization in the mammalian cell nucleus plays a vital role in the regulation of gene expression. The lamina-associated domain at the inner nuclear membrane has been proposed to harbor heterochromatin, while the nuclear interior has been shown to contain most of the euchromatin. Here, we show that a sub-set of actively transcribing genes, marked by RNA Pol II pSer2, are associated with Lamin B1 at the inner nuclear envelop in mESCs and the number of genes proportionally increases upon *in vitro* differentiation of mESC to olfactory precursor cells. These nuclear periphery-associated actively transcribing genes primarily represent housekeeping genes, and their gene bodies are significantly enriched with guanine and cytosine compared to genes actively transcribed at the nuclear interior. We found the promoters of these genes to also be significantly enriched with guanine and to be predominantly regulated by zinc finger protein transcription factors. We provide evidence supporting the emerging notion that the Lamin B1 region is not solely transcriptionally silent.

Introduction

Euchromatin and heterochromatin of mammalian nuclei are compartmentalized into two main nuclear territories[1-4]. Early microscopic observations have shown that the heterochromatic histone signatures mark denser chromatin structures at the inner nuclear envelope (NE) and internal heterochromatin domains, and the chromatin at the nuclear interior (NI) is marked by the euchromatic histone signatures that are present on transcriptionally active genes[3-5]. Furthermore, chromatin capture techniques (e.g., Hi-C and 5C) also provide evidence that chromatin in the nucleus is compartmentalized into two Hi-C-defined nuclear compartments, the transcriptionally active gene containing nuclear compartment A and the transcriptionally silent gene containing nuclear compartment B[6]. The inner NE's lamina meshwork primarily contains Lamin A/C and Lamin B, mainly Lamin B1(LB1)[7,8], as well as inner NE proteins such as Lap2α, Emerin and Man1[9]. The chromatin at the NE has been shown to interact with the nuclear lamina[10] and therefore, the nuclear lamina plays a vital role in chromatin organization and the regulation of gene expression in the cell nucleus.

Further, the Dam-ID technique, used to label chromatin interacting with LB1 (lamina-associated domains or LADs), has also shown that the B compartment is predominantly associated with LADs at the NP[11]. However, the nuclear pore regions lacking the lamina meshwork have been shown to "gate" transcriptionally active genes[12,13] and thus, the NE does not represent a completely heterochromatic region[14-16]. Further, by performing structure-resolving image analysis for LB1, recent studies have shown that primed enhancers, predominantly marked by H3K9me2, are present at areas of the inner NE lacking LB1[17].

It has been shown that when cells differentiate, the transcriptional identity and the composition of genes at the LB1 regions are changed[18,19]. Further, the arrangement of euchromatin and heterochromatin in a nucleus can also be changed during cell specification; some genes move away from the LB1 region to be activated, and some move into the LB1 region to be activated[13,20]. For example, most of the euchromatin in olfactory cells, the rod cells, and Lamin B receptor-null thymocytes is present at the inner NE[21-23]. However, the association of those transcriptionally active genes within the different sub-compartments (e.g., LB1, nuclear pores, LB1 adjacent) of the inner NE has generally not been investigated in detail.

To date, the presence of heterochromatic genes at the nuclear periphery (NP) was identified based on data generated from histone signatures, Hi-C, and Dam_ID methods[24-27]. To our knowledge, detailed approaches specifically using transcriptional profiling of the LB1 associated genes based on RNA polymerase II (RNA Pol II) activity has not yet been investigated in detail genome-wide. Although transcriptional profiling by histone signatures provides the overall state of a gene's transcription (active or silent), the different status of a gene's transcription (pre-initiation, promoter-proximal pausing, elongation, termination) is non-informative in those data.

However, occupancy by RNA Pol II and its phosphorylated derivatives at the different regions of a gene can provide more insight into its transcriptional status. Transcription can be explained as a step-by-step process: RNA Pol II recruitment, initiation, elongation, and termination[28]. Although the phosphorylation of Ser5 at the C-terminal domain of the RNA Pol II large subunit defines the transcription initiation of the gene, a gene in such a state can be either in transcription elongation (in 'canonical' transcription) or promoter-proximal pausing [29,30]. However, phosphorylation at Ser2 (pSer2) at the C-terminus of RNA Pol II is consistent with transcriptional elongation and thus can

be used as a molecular signature of actively transcribing genes. To our knowledge, the presence of pSer2 association, and therefore actively transcribing genes, at the LB1 region, has not been investigated in depth. Although the bulk data from RNA Pol II pSer2-associated chromatin profiling using chromatin immunoprecipitation (ChIP) is available[31-33], those data are non-informative for actively transcribed genes residing in different nuclear compartments.

Since the LB1 region at the NP has been shown to harbor a large amount of DNA[2] we focused on investigating the possibility of actively transcribing genes being localized with LB1 at the NP using immunofluorescence (IF) and sequential ChIP-seq (using LB1 and RNA Pol II pSer2 specific antibodies) assays. In mouse embryonic stem cells (mESCs) and *in vitro* mESCs-derived neural progenitor cells (NPCs), we found that a sub-set of actively transcribing genes were associated with LB1. Further, we demonstrated that upon mESC differentiation to NPCs, the number of actively transcribing genes at the LB1 region was increased. Further, using this approach, we showed that *in vitro* derived NPCs were enriched in active gene signatures similar to olfactory progenitor cells (OPCs). Mature olfactory cells have a reverse nuclear organization whereby euchromatin is present at the NE and heterochromatin at the nuclear interior (NI)[23]. Therefore, the *in vitro* derived NPCs are primed OPCs that are setting up the chromatin organization of mature olfactory cells. Moreover, we found that actively transcribing genes at LB1 regions are significantly enriched with guanine (G) and cytosine (C) and are present in genes with relatively lengthier gene bodies than those that are transcribed at the NI. Although we did not find these genes relevant to the cell identity of OPCs, they primarily represent housekeeping genes. Further, the promoters associated with these genes are also enriched with G and are known to be predominately regulated by zinc-finger transcription factor family proteins. In summary, we

present evidence that actively transcribing genes exist at the LB1 compartment and they exhibit distinct molecular characteristics from genes transcribed at the NI.

Results

The level of RNA Pol II pSer2 increases at the LB1 region in ESC-derived NPCs

To investigate whether actively transcribing genes are associated with LB1, we first performed an immunofluorescence assay in mESCs and *in vitro* derived NPCs using an RNA Pol II pSer2 antibody, a molecular marker of actively transcribed genes, and a LB1 antibody. We observed that the LB1 immunofluorescent signal was punctuated at the NP in mESCs and NPCs (Fig. 1A). Some of the inner NE "pockets" lacking LB1 labeling were positively stained with DAPI (Fig. 1A – white arrowheads), indicating that those LB1 lacking regions contained chromatin and some regions were negative for DAPI (Fig. 1A – yellow arrowheads). In the immunofluorescent images we observed that, although the majority of the RNA Pol II pSer2 immunofluorescent signals were at the NI, fewer RNA Pol II pSer2 signals were localized at the NP (Fig. 1B). Some of the NP localized RNA Pol II pSer2 immunofluorescent signals were also present at regions of the NP lacking LB1 labeling (Fig. 1B – white arrows), while some RNA Pol II pSer2 immunofluorescent signals were seemingly overlapping with LB1 immunofluorescent signals (Fig. 1B – white arrowheads), suggesting the possibility of actively transcribing genes at the LB1 region.

Sequential ChIP-seq analysis revealed that actively transcribed genes are present at the LB1 nuclear region

RNA Pol II pSer2 marks actively transcribing gene bodies[34]. Therefore, based on our immunofluorescence observation, we sought to investigate the RNA Pol II pSer2-associated genes

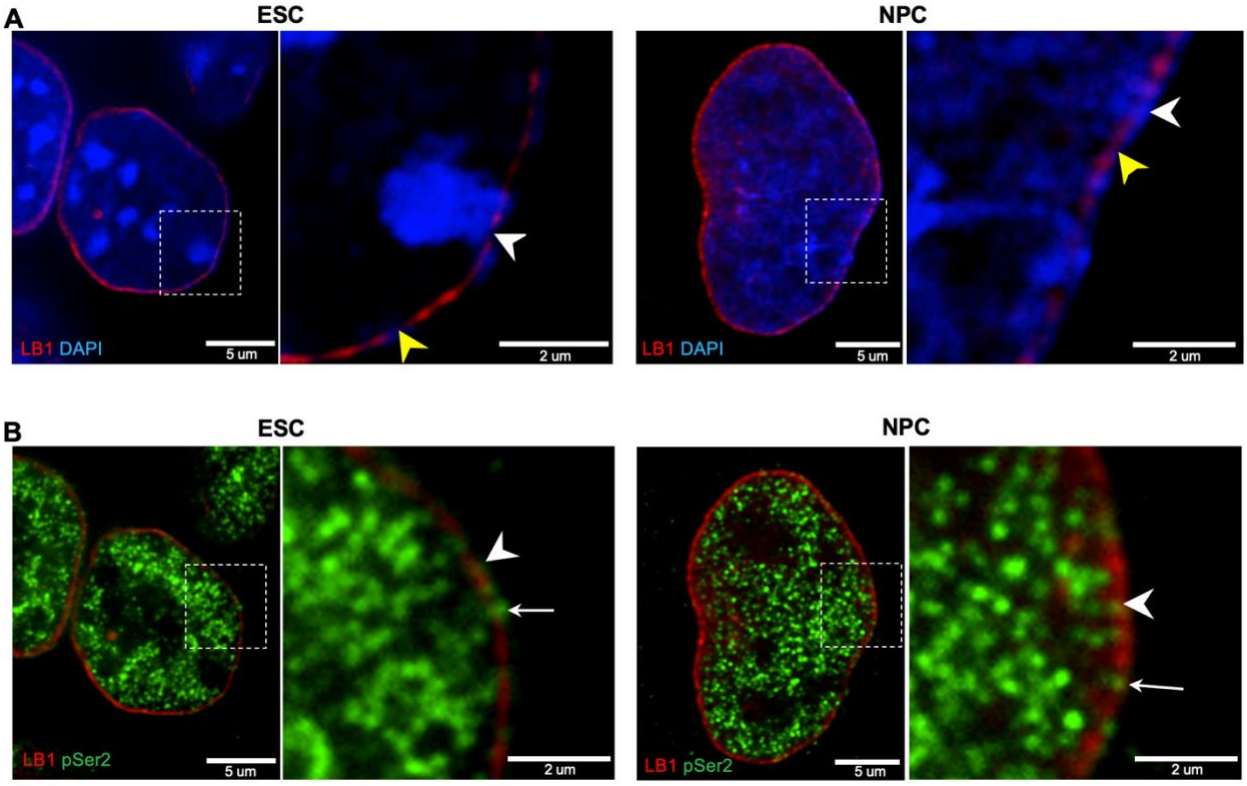


Figure 1. **Presence of LB1 associated RNA Pol II pSer2 increases at the NP in mESCs-derived NPCs** **A**, Confocal microscopy images showing LB1 and DAPI staining in mESCs and NPCs. LB1 is located along the inner nuclear membrane and exhibits a punctuated labeling pattern. (Left panel: Scale bar – 5um, Right panel: Scale bar – 2um). **B**, The same cells in image **A** show RNA Pol II pSer2 and LB1 fluorescent signals. Some RNA Pol II pSer2 fluorescent signals are present at regions lacking LB1at the NP (white arrowhead), and some RNA Pol II pSer2 fluorescent signals are overlapping with LB1 fluorescent signals (white arrows). (Left panel: Scale bar – 5um, Right panel: Scale bar – 2um).

at the LB1 compartment in a high-throughput manner. To do so, we designed a sequential ChIP assay, first to isolate all the genes that were associated with LB1 and then to specifically pull down the genes that interacted with RNA Pol II pSer2, therefore actively transcribing genes in the LB1 compartment (Fig. 2A). In the first step of the sequential ChIP assay, we were able to pull-down ~83% of the total LB1 protein in the input lysate using LB1 antibody (Fig. 2B, C) from mESCs.

We then used the pulled-down LB1-chromatin fraction as input to pull-down the genes that were bound by RNA Pol II pSer2 (Extended Data Fig. 2A). Next, to identify the genes that were enriched in the RNA Pol II pSer2 pull-down chromatin, we constructed DNA libraries (Extended Data Fig. 2B) using the Illumina ChIP-seq kit and sequenced the DNA using the Illumina HiSeq sequencing platform. We used 150 paired-end sequencing to increase the alignment accuracy and minimize the probability of aligning into multiple genome loci when aligning to the reference genome.

In our ChIP-seq analysis, we found that a total of 21,301 genes were expressed (RNA Pol II pSer2 bound) in mESCs and from that, 2.7% (462) genes were actively transcribing at the LB1 region (Fig. 2D). We found that 17,780 genes are common to both LB1 compartment and NI in mESCs. We speculate that this is perhaps due to incomplete LB1 pull-down in the first step of the sequential ChIP assay (Fig. 2C) or an allele of the same gene being expressed from each of the two nuclear regions. Alternatively, some genes may not have preferential compartmentalization (LB1 or NI) in the nucleus when they are transcriptionally active. However, we cannot further clarify the above observation since the assay was performed using bulk RNA-seq from a pool of cells. Therefore, for further analysis we did not include genes common to both the LB1 and NI compartments in downstream analysis to avoid misinterpretation of our ChIP-seq data. To assess the genes that are exclusively expressed in the LB1 region, we intersected the active (LB1 and RNA Pol II pSer2 associated) and inactive (LB1 associated but pSer2 negative) gene sets, and the active (NI and RNA Pol II pSer2 associated) and inactive (NI associated but RNA Pol II pSer2 negative) gene sets at the NI from mESCs (Fig. 2A). We found that both alleles of 274 genes were expressed at the LB1 region, while for an additional 188 genes only a single allele was expressed at the LB1 region in mESCs (Fig. 2E).

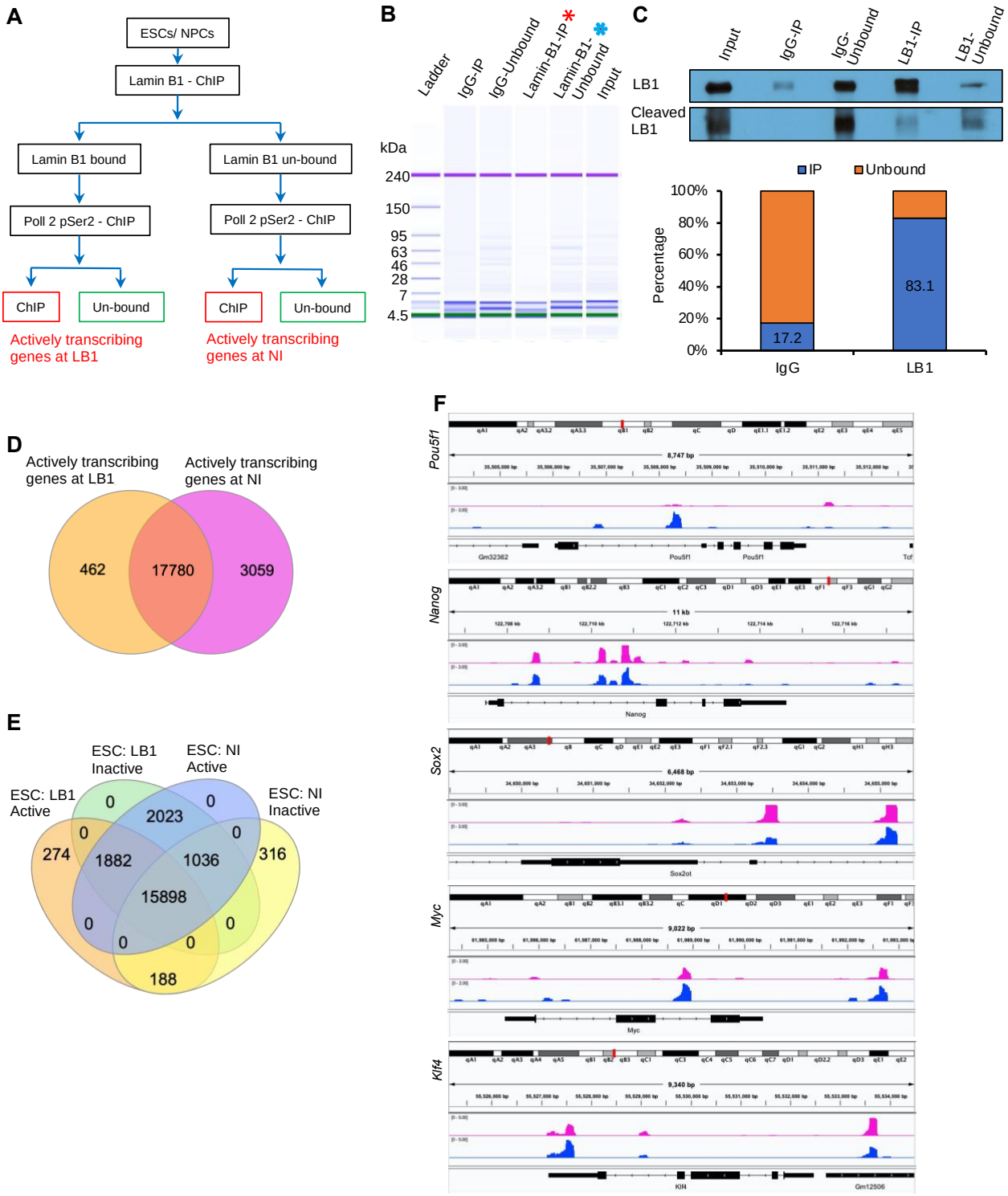


Figure 2. Sequential ChIP assay revealed genes that are actively transcribed at the LB1 nuclear region

A, Schematic presentation of sequential ChIP workflow. In the first step of the pull-down assay, LB1 bound and unbound fractions were separated and used as the input material in the second step using RNA Pol II pSer2 antibody. DNA was extracted from the chromatin fraction pull-down by pSer2 (red rectangles in the diagram) and the pSer2 unbound (green rectangles in the diagram) chromatin. (LB1: Lamin_B1 compartment, NI: nuclear interior). **B**, Protein content in different fractions in the first step (pull-down by LB1) of the sequential chip assay in ESCs. Only the LB1 bound (red star) and LB1 unbound (blue star) protein-chromatin fractions were used for the second step of the pull-down assay. The protein contents in the LB1 pull-down fractions are very low compared to the LB1 unbound and IgG control fractions. **C**, Immunoblot showing the specificity of the LB1 pull down in the first step of the sequential ChIP assay. LB1 unbound fraction shows that LB1 protein was not completely pulled down and remains in the LB1 unbound fraction. The cleaved LB1 protein fraction is higher in the LB1 unbound fraction compared to the LB1 and IgG-bound fractions. Histogram showing the percentages of the quantitated LB1 signal intensities of the bands in bound and unbound fractions. Total intensities were calculated by summing the signal intensities of the LB1 band and the cleaved LB1 band in each lane. **D**, Venn diagram showing the actively transcribing genes in LB1 compartment and NI compartment of mESCs assessed by ChIP-seq analysis. 18,242 and 20,839 actively transcribing genes were detected in LB1 and NI compartments, and 17,780 genes are common to both gene categories. 462 genes are exclusively expressed in the LB1 compartment, while 3,059 genes are exclusively expressed at the NI. **E**, Venn diagram demonstrating the monoallelically (188) and biallelically (274) expressed genes in the LB1 compartment in mESCs. **F**, ChIP-seq peak tracks for *Pou5f1*, *Nanog*, *Sox2*, *Myc*, and *Klf4* genes illustrated in integrative Genomic Viewer. Data summarized in each peak as the mean of all the pixels overlapped into a single genomic location. Magenta-colored tracks show the transcribing genes in the LB1 compartment, and the blue-colored tracks show the genes transcribed at the NI. The *Pou5f1* gene is mainly transcribed in the NI. *Nanog*, *Sox2*, *Myc*, and *Klf4* genes could be transcribed either in the LB1 compartment or NI compartment. Scales for both tracks were set manually to the same value.

We have previously shown that many monoallelic genes are preferentially expressed at Sun1-enriched territories at the NP[35]. Therefore, we asked whether the monoallelic genes expressed at the Sun1 territories and the LB1 region associated monoallelic genes in mESCs were common. We found that only ~5% of the transcriptionally active genes from the LB1 region and the Sun1-territories overlapped (Extended Data Fig. 2C), suggesting that the LB1 region and Sun1 territories may differentially scaffold the monoallelic genes at the NP. However, given that we could only pull-down ~83% of the LB1 protein at the first step of the sequential ChIP assay in mESCs, some of the active monoallelic genes associated with LB1 may not have been captured. However, the ChIP-seq peak association for biallelic and monoallelic genes tested (e.g., Biallelic: *Rpl7, Wnt10a, Marco*. Monoallelic: *Rpl37a, Pomt1, Spint1*) confirmed that their ChIP-seq peaks are enriched only at the LB1 region (Extended Data Fig. 2D).

Genes may be organized in different nuclear regions depending upon the state of the cell[36,37]. Therefore, we asked whether any known genes involved in pluripotency were actively transcribed at the LB1 compartment in mESCs. However, we did not find pluripotency-associated genes (e.g., *Pouf51*, *Nanog*, *Sox2*, *Klf4*, and *Myc)* exclusively expressed in either compartment in mESC nuclei (Fig. 2F). Further, to understand whether the genes localized at the LB1 region play a significant role in ESCs specific functions, we searched for the gene ontologies associated with the actively transcribing genes (monoallelic and biallelic) at the LB1 region. However, we did not observe any significant gene ontology enrichment specific for stem cells, instead these genes consisted of housekeeping genes.

Actively transcribed genes proportionally increase their occupancy at LB1 in ESC-derived OPCs

It has been previously shown by RNA-seq analysis that when mESCs differentiate into NPCs *in vitro*, the total number of expressed genes declines ~50%[38]. Therefore, to test whether we could utilize our sequential ChIP assay to recapitulate the reported gene expression changes in NPCs, we pulled-down the RNA Pol II pSer2-associated chromatin and assessed the total number of pSer2-associated genes in mESCs versus NPCs (Extended Data Fig. 3A – D). Using our assay, we observed that compared to mESCs, the total number of genes expressed in NPCs had declined by nearly 28% (Fig. 3A – Blue columns), showing the same trend as previously reported[38]. Further, it has been shown that when cells differentiate, chromatin organization in the nucleus changes[39-43]. Moreover, a study using the LB1 Dam-ID method has shown that the number of heterochromatic genes increases at LADs in differentiated cells vs undifferentiated naïve cells[44]. However, in our immunofluorescence assay, we observed that the percentage of the RNA Pol II pSer2 signals at the LB1 region was increased in NPCs compared to mESC (Fig.1 B, E & G). When we assessed our ChIP-seq data for gene enrichment in the LB1 region, aligned with the immunofluorescence observation, we found that the number of actively transcribing genes at the LB1 region was also increased by ~20 fold in NPCs (Fig. 3A- orange columns & B).

When mESCs differentiate into NPCs, we observed 8000 genes (423 genes in mESCs) established their transcription at the LB1 region. Therefore, to understand whether the increased gene activity at the LB1 compartment in NPCs is linked to the specific onset of lineage commitment of these cells, we assessed the gene ontologies of the newly established actively transcribing genes at the LB1 region and the total newly established actively transcribing genes (Fig. 3C - 13 and 851 genes, respectively) in NPCs. For the newly established actively transcribing genes at the LB1 region in

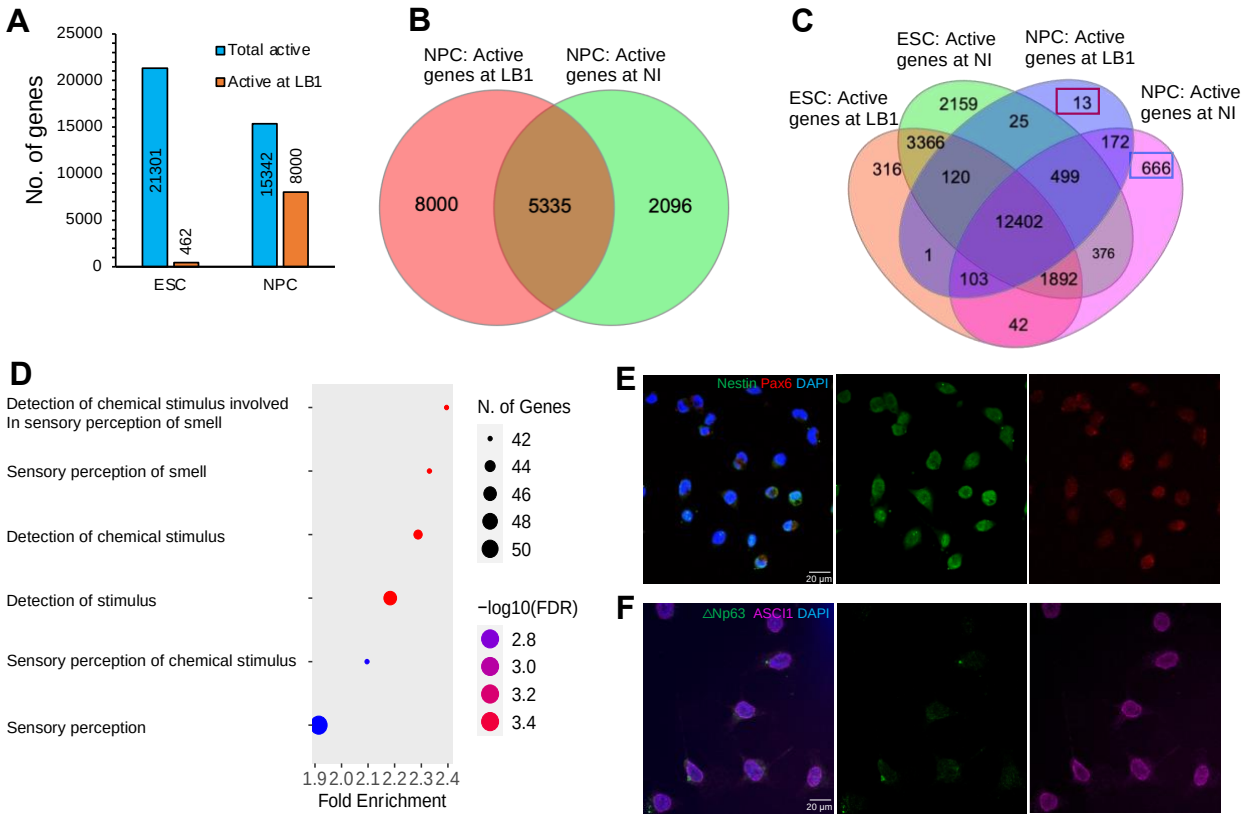


Figure 3. Actively transcribed genes proportionally increase their occupancy at LB1 in mESC-derived OPCs

A, The histogram shows the total number of actively transcribing genes (Blue-coloured) and the actively transcribing genes at the LB1 compartment and NI in mESCs and NPCs assessed by ChIP-seq assay. The total number of actively transcribing genes declines by ~28% in NPCs compared to ESCs. The actively transcribed genes at the LB1 compartment in NPCs have increased by ~20% from the total actively transcribed genes in ESCs. **B**, Venn diagram showing the intersects between actively transcribed genes at the LB1 compartment and NI in NPCs. The number of genes expressed from both compartments is less than in ESCs. **C**, Venn diagram shows the intersects of actively transcribed genes compartmentalization in ESCs and NPCs. Genes that are uniquely expressed at the LB1 (red rectangle) and NI (blue rectangle) are indicated in the diagram. **D**, Gene Ontology analysis of the genes that are exclusively actively transcribed at NI in NPCs. The gene set is exclusively enriched in the stimuli-sensing gene ontology. The size of the dots indicates the number of genes in each GO category, and the color of each dot indicates the -Log10 of

FDR. **E**, IF images showing Nestin and Pax6 expression in the mESC-derived NPCs. The NPCs are positive for both Nestin and Pax6. Scale bar: 20 µm**. F**, IF images showing the expression of △Np63 and Ascl1in NPCs. The NPCs are only positive for Ascl1; therefore, they are Globose basal (olfactory progenitor in early developing embryo's olfactory epithelium) cells. △Np63-green signals show background staining. Scale bar: 20 µm.

NPCs (13 genes), we did not observe any neural-related or cell lineage-specific gene enrichments; instead, we found that those genes exhibit "housekeeping" molecular functions. However, for the total newly established actively transcribing genes in NPCs (851 genes), we found that KEGG pathways are only enriched for olfactory transduction (fold enrichment: 1.9, gene enrichment FDR: 7.2E-03) and proteasome (fold enrichment: 6.5, gene enrichment FDR: 3.4E-02) pathways (Extended Data Fig. 3E). However, testing the genes actively transcribing only at NI regions (666 genes), we found that those genes were solely enriched in the olfactory transduction pathway (fold enrichment: 2.4, gene enrichment FDR: 6.0E-05) (Fig.3 D) as reported previously[45,46].

Although it is generally accepted in most mammalian cells that LB1 regions predominantly harbor heterochromatin, chromatin organization in the nuclei of olfactory neuron's has been shown to have an inverse heterochromatin and euchromatin organization[23]. Therefore, based on the gene ontologies of the newly established actively transcribing genes in NPCs, we speculated that the cells in NPC cultures may represent primed precursor olfactory neuron cells of which euchromatin and heterochromatin chromatin arrangements were predetermined. Although Pax6 and Nestin are known neural progenitor markers, it has also been shown that Nestin is expressed in radial glia-like progenitors that differentiate to zonally restricted olfactory and vomeronasal neurons[47] and that Pax6 is critical in neuronal incorporation into the adult olfactory bulb[48]. Interestingly, the cells

in our NPC cultures are positive for both markers (Fig. 3E). Therefore, we assessed the OPC-specific signature in our NPC cultures using the olfactory progenitor cell markers Ascl1 (Globose basal cells-GBCs) and △Np63 (Horizontal basal cells-HBCs). GBCs originate in early development, and the HBCs appear in the olfactory epithelium at a later developmental stage[49]. In our immunofluorescence assay for Ascl1 and △Np63, we observed that the NPC cells are only positive for Ascl1, confirming that our *in vitro* derived NPCs mimic the cellular identity of olfactory progenitors of the embryonic placode[50] (Fig. 3F).

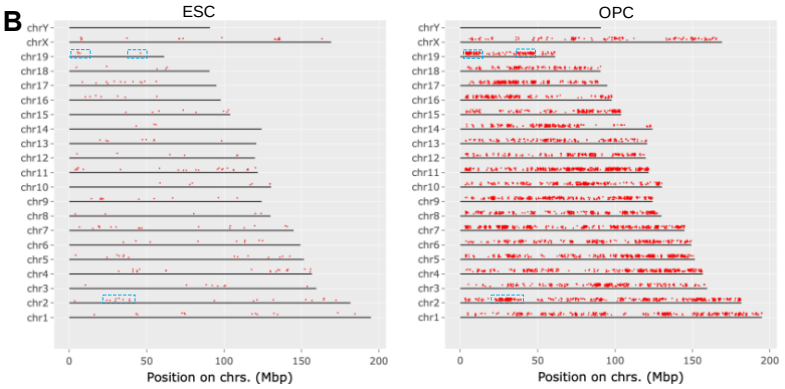
Actively transcribing genes at the LB1 region exhibit unique molecular characteristics

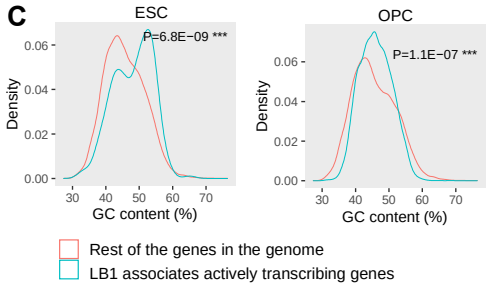
RNA Pol II is known to transcribe protein-coding and non-coding RNA genes[51]. Therefore, to investigate whether there are preferentially transcribing gene classes at the LB1 region, we next assessed the enrichment of gene classes in the LB1 associated actively transcribing genes (mESCs-423 genes and NPCs-8000 genes). We found that the input gene sets represent genes from various gene classes (Fig. 4A), and the enrichment of some classes was seemingly increased when mESCs differentiated into OPCs. Furthermore, those LB1-associated actively transcribing genes were spread across the genome (Fig. 4B) and we found that more genes were accumulated around those "preset loci" in ESCs when differentiated to OPCs (Fig. 4B – e.g., dotted blue rectangles on Chr2 and Chr19.

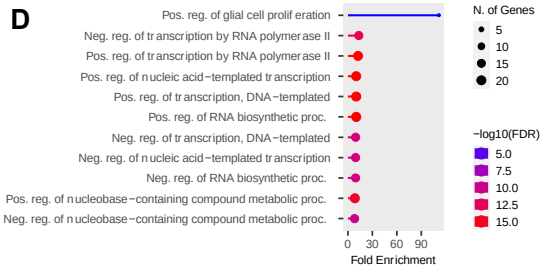
Furthermore, to understand the molecular features of the actively transcribing genes at the LB1 region, we next investigated the molecular characteristics of those genes, focusing on the coding sequence length, transcript length, genome span, 5' UTR length, 3' UTR length, and guanine and cytosine (GC) content (Fig. 4C, Extended Data Fig. 4A). Compared to all other genes in the

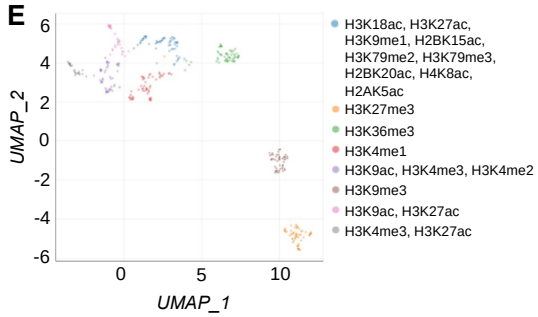
A

Gene classification	ESCs	OPCs
protein_coding	213	5868
lncRNA	145	1450
polymorphic_pseudogene	1	7
processed_pseudogene	34	135
snoRNA	2	18
snRNA	4	21
TEC	42	296
transcribed_processed_pseudogene	5	32
transcribed_unprocessed_pseudogene	4	22
unprocessed_pseudogene	3	8
miRNA	0	12
misc_RNA	0	2
TR_C_gene	0	1
transcribed_unitary_pseudogene	0	5
translated_unprocessed_pseudogene	0	1
unitary_pseudogene	0	9
scaRNA	2	1
Not mapped	7	92









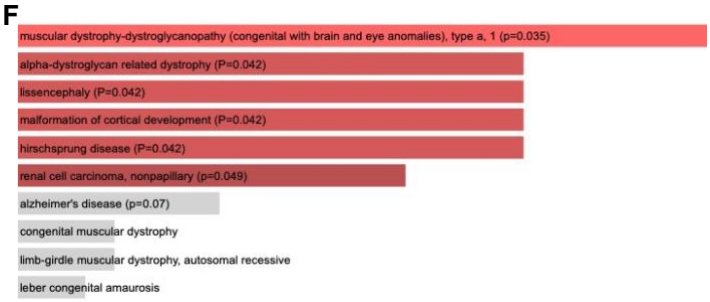


Figure 4. Actively transcribed genes at the LB1 compartment exhibit unique molecular characteristics

A, Table summarizes the classification of the actively transcribed genes in the LB1 region and NI. **B**, Illustration of RNA Pol II pSer2 pull-down gene distribution through the genome. Compared to ESCs, there is an accumulation of actively transcribing genes at the LB1 compartment in NPCs. Also, in NPCs, more genes become accumulated around the gene loci, which were already established at the LB1 region in ESCs (e.g., Chr2 and Chr19 – blue rectangles). **C**, Density plot showing the GC content of the actively transcribing genes at LB1. GC content is significantly enriched at actively transcribing genes at the LB1 compartment compared to the rest of the genes in the genome in both ESCs and NPCs. **D**, Gene Ontologies

of the transcription factors known to bind to the promoters of the actively transcribing genes at the LB1 region. GO terms exhibit that those TFs are associated with RNA Pol II. The size of the dots shows the number of genes, and the color of the dots shows the -Log10 of the FDR of each GO term. **E**, In silico analysis of histone signature association with the actively transcribing genes at the LB1 compartment. The dots in the UMAP plot indicate the known histone signatures associated with the assessed genes. "Blue" dots indicate the known histone signature of active gene bodies, as well as active enhancers (H3K27ac, H2BK20ac, HeK9me2), active promoters (H4K8ac, H2AK5ac, H3K9ac), and TSS of the transcribing genes (H3K4me3, H3K4me2). **F**, Bar plot showing the gene set association in Clinvar data. The X-axis represents the -log10 of the p value, and the actual p values are indicated in each column. The red color shows the enriched terms for the input gene set from mESCs and OPCs.

genome (tested by Chi-square and Student's t-tests), we found that the genome span, transcript length (Extended Data Fig. 4A) and the GC content (Fig. 4C) of the actively transcribing genes (mESCs: 462 and OPCs: 8000) at the LB1 region in mESCs and OPCs were significantly higher (ESC: p=6.8E-09 and OPCs: p=1.1E-07) compared to the rest of the genes in the genome. Moreover, we found that actively transcribing genes at the LB1 region transcribe a single or two isoforms, while the actively transcribing genes at NI were generally found to transcribe more than two isoforms (Extended Data Fig. 4B).

It has also been shown that the nuclear envelope is enriched with transcription factors (TFs) and that they play a role in localizing TF-bound genes at the nuclear pore complexes[52]. Therefore, taking this as a lead, we assessed the molecular features of the promoters of LB1-associated genes and the known TFs bound to those promoters. We found that the genes exclusively transcribed at the LB1 region were enriched with G-rich promoters (Extended Data Table 4A). Further, these

promoters are known to be predominantly regulated by zinc finger TF family proteins (Extended Data Table 4A) and these TFs are known to regulate RNA Pol II-dependent gene transcription of both protein coding and non-coding genes (Fig. 4D).

It has recently been shown that within the LB1 region, small "pocket-like" structures lacking LB1 accommodate active enhancer elements marked by H3K9me2[17]. Because the actively transcribing genes at the NP that we found in this study were associated with LB1, we investigated the known histone signatures enriched for those genes using the ENCODE database embedded in Enricher - a data analysis web tool[53-55]. We found that the genes we tested exhibit associations with known histone signatures (H3K36me3, H2BK15ac, H3K79me2, H3K79me3, H4K8ac, H3K4me3), which mark actively transcribing gene bodies (Fig. 4E).

Discussion

Although LADs are known to predominantly harbor heterochromatin, growing evidence strongly suggests the LADs are not completely heterochromatic territories but do harbor active genes[16,20,56,57]. Our immunofluorescence data in mESCs and OPCs strongly suggests that actively transcribing genes (marked by RNA Pol II pSer2) are present at the NP associated with LB1. Sequential ChIP assays have been developed to identify the physical interaction of multiple proteins, for example chromatin and transcription factors, simultaneous protein modifications on the same molecule, and to validate interactions in protein complexes[58,59]. We carried out a second-line approach utilizing a sequential ChIP-seq assay to validate our immunofluorescence observation and to investigate the enriched active genes associated with the LB1 region. Using

two distinct experimental approaches, we demonstrated that LB1 harbors actively transcribing genes in mESCs and OPCs. It has been shown that upon cell differentiation into more specialized cells, more genes tend to be repositioned at the inner NE and gain heterochromatic status[44]. However, in line with the fact that euchromatin and heterochromatin organization is inverted in olfactory cell nuclei, our findings support the emerging idea that the LB1 compartment is not predominantly heterochromatic but can also accommodate transcriptionally active genes based on the cellular context. Therefore, we suggest that because different cell types can have different chromatin architectures, the organization of transcriptionally active and silent genes needs to be assessed within the context of a specific cell type being studied.

The RNA Pol II pSer2 bound, therefore actively transcribing, genes we found at the LB1 region contain unique molecular characteristics, for example high GC enrichment, higher genomic span, and long transcript length. Genes that transcribe longer transcripts have been shown to gain fewer SNPs than shorter genes through evolution[60-64]. Therefore, since the LB1-associated genes are housekeeping genes and these genes share identical molecular functions among all cell types, it is reasonable to argue that these genes may be less prone to gain changes during gene evolution. Although longer genes are associated with embryonic development[65] and neuronal development[66], we found no cell type specificity relating to the LB1-associated actively transcribing genes.

Furthermore, the relative GC enrichment of a gene is positively correlated with the transition of double-stranded DNA from its B form to Z form, increasing the bendability and reducing the thermostability of the DNA[67]. In addition, increased GC content is consistent with an open chromatin structure, consistent with stable and continuous transcription, which would be beneficial for housekeeping genes. In addition, open chromatin provides an advantage of minimizing the

energy cost to the cell by reducing frequent histone modifications, which requires energy usage. Therefore, the evolutionary advantage of high GC content in these genes may be used as a mechanism that could result in continuous open chromatin with minimal involvement of histone modifications, a hypothesis worth examining in detail in future studies.

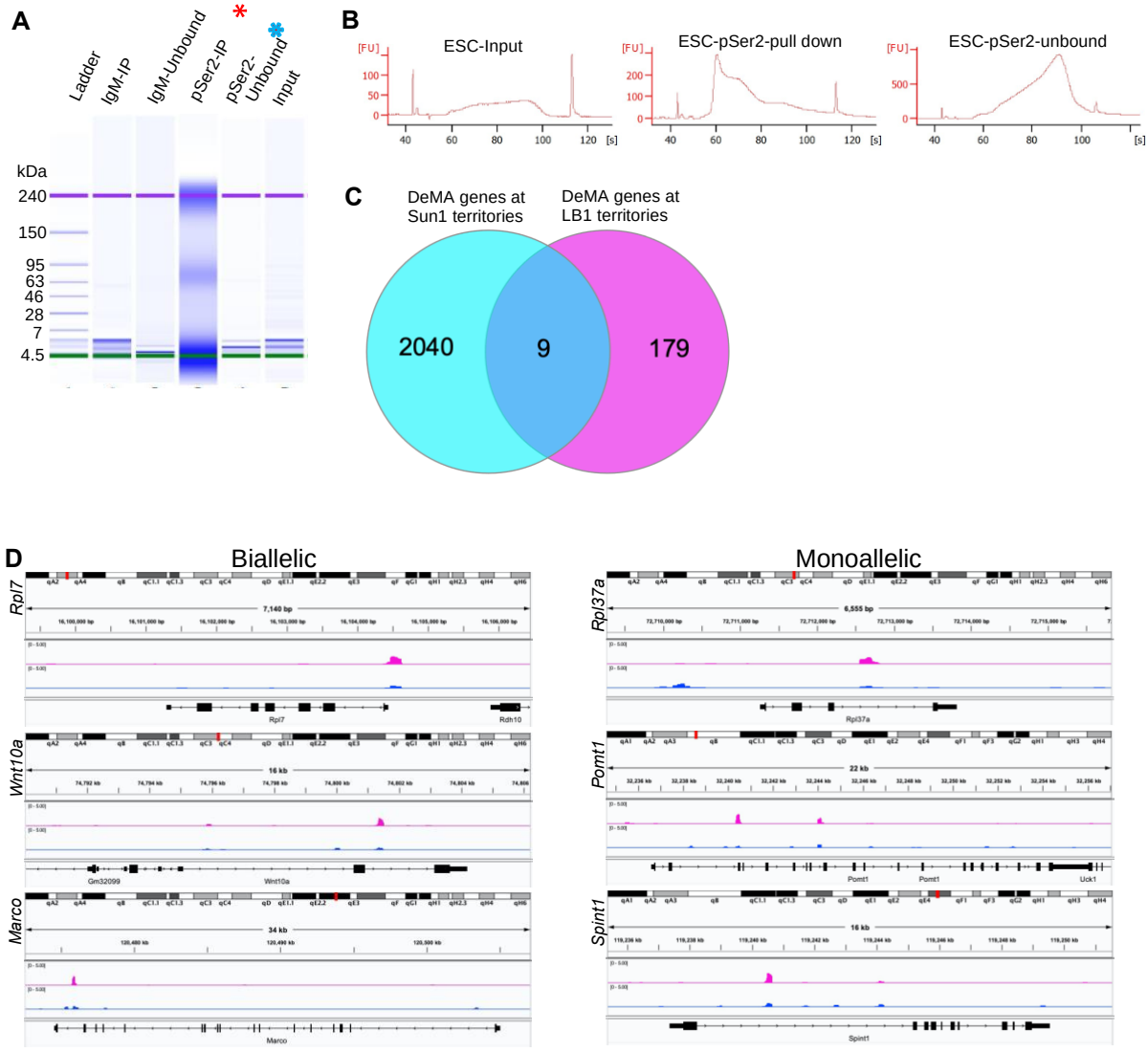
By compiling the LB1 associated genes with ENCODE histone data, we observed that histone signatures known to mark actively transcribing gene bodies are associated with genes we pulled-down with RNA Pol II pSer2 from the LB1 region. Along with H3K27ac and H2BK20ac, we observed that H3K9me2, an active enhancer mark, is also associated with the LB1 region-associated genes (Fig. 4E). H3K9me2 reportedly marks the primed enhancers that are shown to be localized in the LB1-lacking regions at the inner NE. Therefore, it is reasonable to argue that, because of the proximity of the LB1-associated genes and the H3K9me2 primed enhancers at the inner NE, genes at the LB1 region could be regulated by *cis-elements* localized at those LB1 lacking NE areas. This potential mechanism is worth assessing in future studies by developing a sequential ChIP assay combined with Hi-C techniques and enhancer editing approaches.

Interestingly, we found that the G-rich promoters associated with actively transcribing genes at the LB1 region were predominantly regulated by zinc finger TFs. G-rich DNA motifs are likely to form G-quadruplexes in different physiological conditions and the promoters containing G-quadruplexes were shown to interact with zinc finger family TFs[68]. It is possible that the promoters associated with LB1 regions may contain G-quadruplexes and thus recruit zinc finger TFs to regulate gene expression[69] and this would also be an interesting mechanism to explore in future studies.

Nuclear periphery-localized genes are known to misregulate their expression in laminopathy diseases[70]. In addition, using mouse models, it has been shown that misregulation of LB1 in the nuclear periphery of striatal medium-sized spiny and CA1 hippocampal neurons is linked to Huntington-mediated neurodegeneration and therefore, the LB1-associated actively transcribing genes may bear a pathological relevance in neurons[71]. In our study, by compiling the LB1-associated transcriptionally active genes in OPCs with Clinvar data, we found that several candidates (e.g., Fkrp, Dcx) from our gene sets are associated with known neuronal disease-related dystrophies in humans (Fig.4 F), suggesting that LB1-associated genes may have clinical relevance in some laminopathy diseases.

In summary, we present evidence that the LB1 region at the inner nuclear membrane of mESCs and OPCs contains a significant number of actively transcribing genes that exhibit unique molecular features. Further, we showed that the transcriptionally active gene bodies at the LB1 region are enriched for G and C, and the promoters of those genes are G-rich and are known to interact mainly with zinc finger family TFs'. Our data support a model whereby the complexity and activity of genes associated with LB1 is cell-type specific and can change upon differentiation.

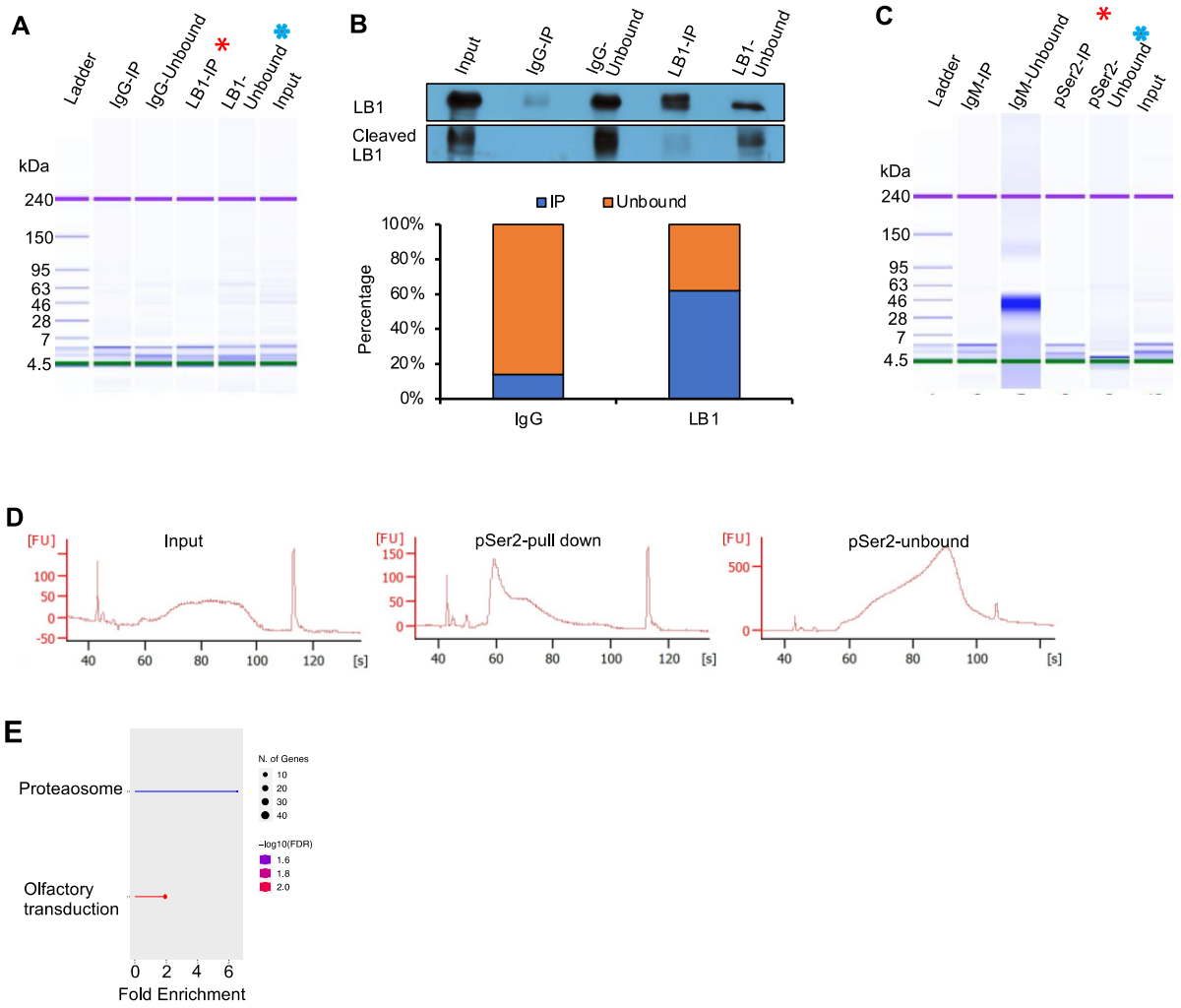
Extended Data Figures



Extended Data Figure 2. Sequential ChIP assay revealed that genes are actively transcribed at the LB1 nuclear region

A, Protein content in different fractions in the sequential chip assay's second step (pull-down by RNA Pol II pSer2). Compared to IgM control, a significantly higher amount of protein is in the RNA Pol II pSer2 pull-down fraction. **B**, Analysis showing the quality and the concentrations of sequencing libraries constructed from DNA isolated from each fraction. **C**, Venn diagram showing the intersection of DeMA (deterministic monoallelic genes) genes at the Sun1 territories at the inner nuclear membrane and the LB1 region. The results show that a very low percentage of genes are common to both compartments. **D**,

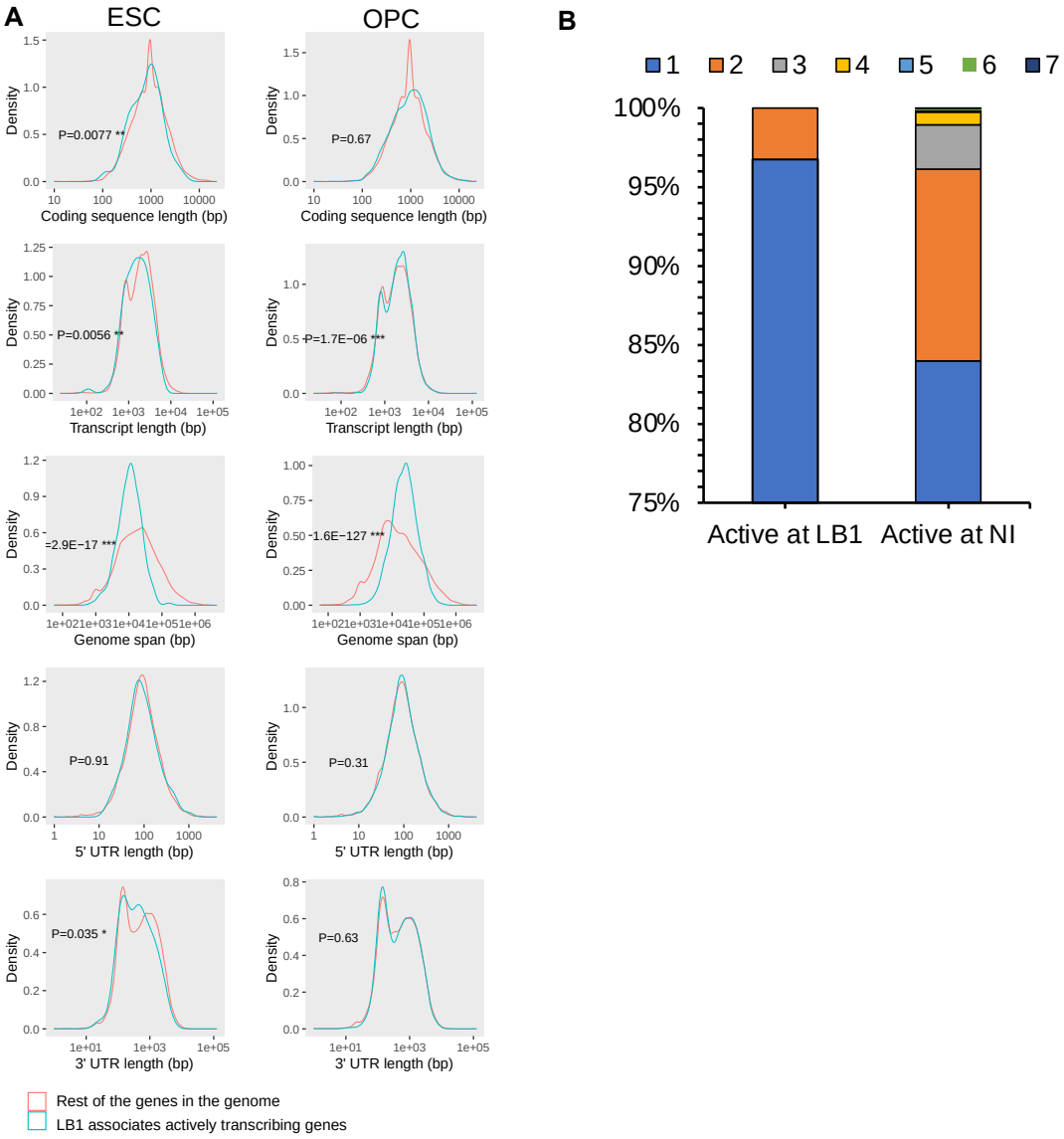
Representative ChIP-seq peak tracks to illustrate the peak distribution of several expressed biallelic and monoallelic genes associated with the LB1 region.



Extended Data Figure 3. Actively transcribed genes proportionally increase their occupancy at LB1 in mESC-derived OPCs

A, Protein chip showing the relative protein fractions pulled down by LB1 in the first step of the sequential ChIP assay using NPCs. Only the LB1 bound (red star) and LB1 unbound (blue star) protein-chromatin fractions were used for the second step of the pull-down. **B**, The representative immunoblot is showing the LB1 pull-down efficiency in NPCs. Compared to the unbound fractions of IgG and LB1, the signals relevant to cleaved LB1 are lower in IgG and LB1 fractions, similar to the observation made in mESCs. The

histogram shows the percentages of signal intensities of IgG and LB1 pull-down bands. Percentages were calculated over the total LB1 (in each IgG and LB1 fraction), the totals of uncleaved and cleaved signals were summed to have the total LB1 signals. In NPCs, ~62% of the total LB1 protein is pulled down. **C**, Protein contents in different fractions in the second step of the sequential ChIP assay in NPCs. **D**, Analysis showing the quality of the DNA-seq libraries constructed from RNA Pol II pSer2 pull-down chromatin-associated DNA in NPCs. **E**, IF images of *in vitro* mESC-derived NPCs showing the NPCs are dual positive for Pax6 and Nestin. Scale bar-20 µm.



Extended Data Figure 4. Actively transcribing genes at the LB1 region exhibit unique molecular characteristics

A, Density plots show the enrichment of the significant molecular characteristics of the actively transcribing genes at the LB1 compartments in mESCs and OPCs. The data indicate that those genes' transcript length and genome span are significantly different from the rest of the genes in the genome. **B**, Histogram shows the number of isoforms per gene in LB1-associated versus NI-associated genes. The data indicate that the LB1-associated genes exclusively have either one or two known isoforms, while actively transcribing genes at the NI could have more than two isoforms.

Data and code availability

The RNA Pol II pSer2 ChIP-seq data (raw and processed) supporting this study's findings have been deposited into the Gene Expression Omnibus (GEO) under accession code GSE246040. We have provided all other data supporting the findings of this study with the manuscript.

Acknowledgements

We would like to thank members of the Spector lab for critical discussions and advice throughout the course of this study. We thank Osama El Demerdash for assistance with setting up a computational pipeline. We would also like to thank Sara Goodwin (CSHL) for assisting with DNA sequencing. We acknowledge the CSHL Cancer Center Shared Resources (Microscopy, Next-Gen Sequencing) for services and technical expertise (NCI 2P3OCA45508). Sequencing analysis was performed using equipment purchased through NIH grant S10OD028632-01. This research was supported by R35GM131833 (D.L.S.) and grant 18-26 from the Charles H. Revson Foundation (G.I.B.).

Author Contribution

G.I.B. and D.L.S. conceived and designed the overall study. G.I.B. designed and performed all the experiments and analysis. G.I.B. and D.L.S. wrote and edited the manuscript. T.W. edited the images and the manuscript.

Competing Interest

The authors declare no competing interests.

Corresponding author

David L. Spector (spector@cshl.edu)

References

1 Solovei, I., Thanisch, K. & Feodorova, Y. How to rule the nucleus: divide et impera. *Curr Opin Cell Biol* **40**, 47-59, doi:10.1016/j.ceb.2016.02.014 (2016).

2 van Steensel, B. & Belmont, A. S. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* **169**, 780-791, doi:10.1016/j.cell.2017.04.022 (2017).

3 Mir, M., Bickmore, W., Furlong, E. E. M. & Narlikar, G. Chromatin topology, condensates and gene regulation: shifting paradigms or just a phase? *Development* **146**, doi:ARTN dev182766 10.1242/dev.182766 (2019).

4 Morris, K. J., Chotalia, M. & Pombo, A. Nuclear Architecture in Stem Cells. *Adv Exp Med Biol* **695**, 14-25, doi:Book_Doi 10.1007/978-1-4419-7037-4 (2010).

5 Rae, M. M. & Franke, W. W. The interphase distribution of satellite DNA-containing heterochromatin in mouse nuclei. *Chromosoma* **39**, 443-456, doi:10.1007/BF00326177 (1972).

6 Bonev, B. & Cavalli, G. Organization and function of the 3D genome. *Nat Rev Genet* **17**, 661-678, doi:10.1038/nrg.2016.112 (2016).

7 Dechat, T., Adam, S. A., Taimen, P., Shimi, T. & Goldman, R. D. Nuclear Lamins. *Csh Perspect Biol* **2**, doi:ARTN a000547 10.1101/cshperspect.a000547 (2010).

8 Melcer, S., Gruenbaum, Y. & Krohne, G. Invertebrate lamins. *Exp Cell Res* **313**, 2157-2166, doi:10.1016/j.yexcr.2007.03.004 (2007).

9 Wilson, K. L. & Foisner, R. Lamin-binding Proteins. *Csh Perspect Biol* **2**, doi:ARTN a000554 10.1101/cshperspect.a000554 (2010).

10 Dechat, T. *et al.* Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Gene Dev* **22**, 832-853, doi:10.1101/gad.1652708 (2008).

11 Kind, J. *et al.* Genome-wide Maps of Nuclear Lamina Interactions in Single Human Cells. *Cell* **163**, 134-147, doi:10.1016/j.cell.2015.08.040 (2015).

12 Scholz, B. A. *et al.* WNT signaling and AHCTF1 promote oncogenic MYC expression through super-enhancer-mediated gene gating. *Nat Genet* **51**, 1723-+, doi:10.1038/s41588-019-0535-3 (2019).

13 Brown, C. R. & Silver, P. A. Transcriptional regulation at the nuclear pore complex. *Curr Opin Genet Dev* **17**, 100-106, doi:10.1016/j.gde.2007.02.005 (2007).

14 Brickner, J. H. & Walter, P. Gene recruitment of the activated INO1 locus to the nuclear membrane. *Plos Biol* **2**, 1843-1853, doi:ARTN e342 10.1371/journal.pbio.0020342 (2004).

15 Casolari, J. M. *et al.* Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* **117**, 427-439, doi:Doi 10.1016/S0092-8674(04)00448-9 (2004).

16 Kumaran, R. I. & Spector, D. L. A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *J Cell Biol* **180**, 51-65, doi:DOI 10.1083/jcb.200706060 (2008).

17 Smith, C. L., Poleshko, A. & Epstein, J. A. The nuclear periphery is a scaffold for tissue-specific enhancers. *Nucleic Acids Res* **49**, 6181-6195, doi:10.1093/nar/gkab392 (2021).

18 Poleshko, A. *et al.* Genome-Nuclear Lamina Interactions Regulate Cardiac Stem Cell Lineage Restriction. *Cell* **171**, 573-+, doi:10.1016/j.cell.2017.09.018 (2017).

19 Gonzalez-Sandoval, A. *et al.* Perinuclear Anchoring of H3K9-Methylated Chromatin Stabilizes Induced Cell Fate in C. elegans Embryos. *Cell* **163**, 1333-1347, doi:10.1016/j.cell.2015.10.066 (2015).

20 Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G. & Cremer, T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* **8**, 104-115, doi:10.1038/nrg2041 (2007).

21 Falk, M. *et al.* Heterochromatin drives compartmentalization of inverted and conventional nuclei. *Nature* **570**, 395-+, doi:10.1038/s41586-019-1275-3 (2019).

22 Solovei, I. *et al.* LBR and Lamin A/C Sequentially Tether Peripheral Heterochromatin and Inversely Regulate Differentiation. *Cell* **152**, 584-598, doi:10.1016/j.cell.2013.01.009 (2013).

23 Bashkirova, E. & Lomvardas, S. Olfactory receptor genes make the case for inter-chromosomal interactions. *Curr Opin Genet Dev* **55**, 106-113, doi:10.1016/j.gde.2019.07.004 (2019).

24 Buchwalter, A., Kaneshiro, J. M. & Hetzer, M. W. Coaching from the sidelines: the nuclear periphery in genome regulation. *Nat Rev Genet* **20**, 39-50, doi:10.1038/s41576-018-0063-5 (2019).

25 Manzo, S. G., Dauban, L. & van Steensel, B. Lamina-associated domains: Tethers and looseners. *Curr Opin Cell Biol* **74**, 80-87, doi:10.1016/j.ceb.2022.01.004 (2022).

26 Mehrmohamadi, M., Sepehri, M. H., Nazer, N. & Norouzi, M. R. A Comparative Overview of Epigenomic Profiling Methods. *Front Cell Dev Biol* **9**, 714687, doi:10.3389/fcell.2021.714687 (2021).

27 Spracklin, G. *et al.* Diverse silent chromatin states modulate genome compartmentalization and loop extrusion barriers. *Nat Struct Mol Biol* **30**, 38-+, doi:10.1038/s41594-022-00892-7 (2023).

28 Nielsen, R. *et al.* Genome-wide profiling of PPAR gamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Gene Dev* **22**, 2953-2967, doi:10.1101/gad.501108 (2008).

29 Bender, T. P., Thompson, C. B. & Kuehl, W. M. Differential expression of c-myb mRNA in murine B lymphomas by a block to transcription elongation. *Science* **237**, 1473-1476, doi:10.1126/science.3498214 (1987).

30 Krumm, A., Meulia, T., Brunvand, M. & Groudine, M. The Block to Transcriptional Elongation within the Human C-Myc Gene Is Determined in the Promoter - Proximal Region. *Gene Dev* **6**, 2201-2213, doi:DOI 10.1101/gad.6.11.2201 (1992).

31 Chen, Y. B., Chen, Q., McEachin, R. C., Cavalcoli, J. D. & Yu, X. C. H2A.B facilitates transcription elongation at methylated CpG loci. *Genome Res* **24**, 570-579, doi:10.1101/gr.156877.113 (2014).

32 Wolf, E. *et al.* Activation and repression by oncogenic Myc shape tumor-specific gene expression profiles. *Febs J* **282**, 282-282 (2015).

33 Kinyamu, H. K., Bennett, B. D., Bushel, P. R. & Archer, T. K. Proteasome inhibition creates a chromatin landscape favorable to RNA Pol II processivity. *Journal of Biological Chemistry* **295**, 1271-1287, doi:10.1074/jbc.RA119.011174 (2020).

34 Bowman, E. A. & Kelly, W. G. RNA Polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation A tail of two kinases. *Nucleus-Phila* **5**, 224-236, doi:10.4161/nucl.29347 (2014).

35 Balasooriya, G. I. & Spector, D. L. Allele pairing at Sun1-enriched domains at the nuclear periphery via T1A3 tandem DNA repeats. *bioRxiv*, 2023.2004.2007.536031, doi:10.1101/2023.04.07.536031 (2023).

36 Thorpe, S. D. & Lee, D. A. Dynamic regulation of nuclear architecture and mechanics-a rheostatic role for the nucleus in tailoring cellular mechanosensitivity. *Nucleus-Phila* **8**, 287-300, doi:10.1080/19491034.2017.1285988 (2017).

37 Lanctot, C., Cheutin, T., Cremer, M., Cavalli, G. & Cremer, T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* **8**, 104-115, doi:10.1038/nrg2041 (2007).

38 Efroni, S. *et al.* Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell* **2**, 437-447, doi:10.1016/j.stem.2008.03.021 (2008).

39 Bodnar, M. S. & Spector, D. L. Chromatin Meets Its Organizers. *Cell* **153**, 1187-1189, doi:10.1016/j.cell.2013.05.030 (2013).

40 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

41 Hou, C. H., Li, L., Qin, Z. H. S. & Corces, V. G. Gene Density, Transcription, and Insulators Contribute to the Partition of the Drosophila Genome into Physical Domains. *Mol Cell* **48**, 471-484, doi:10.1016/j.molcel.2012.08.031 (2012).

42 Nora, E. P. *et al.* Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381-385, doi:10.1038/nature11049 (2012).

43 Sexton, T. *et al.* Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. *Cell* **148**, 458-472, doi:10.1016/j.cell.2012.01.010 (2012).

44 Peric-Hupkes, D. *et al.* 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 (2010).

45 Tan, L. Z., Li, Q. & Xie, X. S. Olfactory sensory neurons transiently express multiple olfactory receptors during development. *Mol Syst Biol* **11**, doi:ARTN 844 DOI 10.15252/msb.20156639 (2015).

46 Hanchate, N. K. *et al.* Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. *Science* **350**, 1251-1255, doi:10.1126/science.aad2456 (2015).

47 Murdoch, B. & Roskams, A. J. A novel embryonic nestin-expressing radial glia-like progenitor gives rise to zonally restricted olfactory and vomeronasal neurons. *J Neurosci* **28**, 4271-4282, doi:10.1523/Jneurosci.5566-07.2008 (2008).

48 Curto, G. G. *et al.* Pax6 Is Essential for the Maintenance and Multi-Lineage Differentiation of Neural Stem Cells, and for Neuronal Incorporation into the Adult Olfactory Bulb. *Stem Cells Dev* **23**, 2813-2830, doi:10.1089/scd.2014.0058 (2014).

49 Packard, A., Schnittke, N., Romano, R. A., Sinha, S. & Schwob, J. E. DeltaNp63 regulates stem cell dynamics in the mammalian olfactory epithelium. *J Neurosci* **31**, 8748-8759, doi:10.1523/JNEUROSCI.0681-11.2011 (2011).

50 Schwob, J. E. *et al.* Stem and progenitor cells of the mammalian olfactory epithelium: Taking poietic license. *J Comp Neurol* **525**, 1034-1054, doi:10.1002/cne.24105 (2017).

51 Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051-4060, doi:10.1038/sj.emboj.7600385 (2004).

52 Heessen, S. & Fornerod, M. The inner nuclear envelope as a transcription factor resting place. *Embo Rep* **8**, 914-919, doi:10.1038/sj.embor.7401075 (2007).

53 Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128, doi:10.1186/1471-2105-14-128 (2013).

54 Kuleshov, M. V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* **44**, W90-97, doi:10.1093/nar/gkw377 (2016).

55 Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Curr Protoc* **1**, e90, doi:10.1002/cpz1.90 (2021).

56 Brickner, J. H. Transcriptional memory at the nuclear periphery. *Curr Opin Cell Biol* **21**, 127-133, doi:10.1016/j.ceb.2009.01.007 (2009).

57 Brueckner, L. *et al.* Local rewiring of genome-nuclear lamina interactions by transcription. *Embo Journal* **39**, doi:ARTN e103159 10.15252/embj.2019103159 (2020).

58 de Medeiros, R. B. Sequential chromatin immunoprecipitation assay and analysis. *Methods Mol Biol* **791**, 225-237, doi:10.1007/978-1-61779-316-5_17 (2011).

59 Beischlag, T. V., Prefontaine, G. G. & Hankinson, O. ChIP-re-ChIP: Co-occupancy Analysis by Sequential Chromatin Immunoprecipitation. *Methods Mol Biol* **1689**, 103-112, doi:10.1007/978-1-4939-7380-4_9 (2018).

60 Lopes, I., Altab, G., Raina, P. & de Magalhaes, J. P. Gene Size Matters: An Analysis of Gene Length in the Human Genome. *Front Genet* **12**, 559998, doi:10.3389/fgene.2021.559998 (2021).

61 Wolf, Y. I., Novichkov, P. S., Karev, G. P., Koonin, E. V. & Lipman, D. J. The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes

of different apparent ages. *Proc Natl Acad Sci U S A* **106**, 7273-7280, doi:10.1073/pnas.0901808106 (2009).

62 Gorlova, O., Fedorov, A., Logothetis, C., Amos, C. & Gorlov, I. Genes with a large intronic burden show greater evolutionary conservation on the protein level. *Bmc Evol Biol* **14**, doi:Doi 10.1186/1471-2148-14-50 (2014).

63 Grishkevich, V. & Yanai, I. Gene length and expression level shape genomic novelties. *Genome Res* **24**, 1497-1503, doi:10.1101/gr.169722.113 (2014).

64 Vishnoi, A., Kryazhimskiy, S., Bazykin, G. A., Hannenhalli, S. & Plotkin, J. B. Young proteins experience more variable selection pressures than old proteins. *Genome Res* **20**, 1574-1581, doi:10.1101/gr.109595.110 (2010).

65 Yang, D. *et al.* A two-level model for the role of complex and young genes in the formation of organism complexity and new insights into the relationship between evolution and development. *Evodevo* **9**, doi:ARTN 22 10.1186/s13227-018-0111-4 (2018).

66 Sahakyan, A. B. & Balasubramanian, S. Long genes and genes with multiple splice variants are enriched in pathways linked to cancer and other multigenic diseases. *Bmc Genomics* **17**, doi:ARTN 225 10.1186/s12864-016-2582-9 (2016).

67 Vinogradov, A. E. DNA helix: the importance of being GC-rich. *Nucleic Acids Res* **31**, 1838-1844, doi:10.1093/nar/gkg296 (2003).

68 Kumar, P. *et al.* Zinc-finger transcription factors are associated with guanine quadruplex motifs in human, chimpanzee, mouse and rat promoters genome-wide. *Nucleic Acids Res* **39**, 8005-8016, doi:10.1093/nar/gkr536 (2011).

69 Shen, J. Z. *et al.* Promoter G-quadruplex folding precedes transcription and is controlled by chromatin. *Genome Biology* **22**, doi:ARTN 143 10.1186/s13059-021-02346-7 (2021).

70 Evangelisti, C. *et al.* The wide and growing range of lamin B-related diseases: from laminopathies to cancer. *Cell Mol Life Sci* **79**, doi:ARTN 126 10.1007/s00018-021-04084-2 (2022).

71 Alcala-Vida, R. *et al.* Neuron type-specific increase in lamin B1 contributes to nuclear dysfunction in Huntington's disease. *Embo Mol Med* **13**, doi:ARTN e12105 10.15252/emmm.202012105 (2021).