Genome Organization with CUT and Tag (GO-CaT) identifies substructure and maturation of lamina-associated domains in neurons of the developing human brain

Chujing Zhang[1,3], Eugene Gil[1,3], Sajad Hamid Ahanger[1,3], Lynn Mingcong Li[1,3], Li Wang[2,3], Jingjing Li[2,3], Arnold R. Kriegstein[2,3], Daniel A. Lim[1,3,4]*

Affiliations:
[1]Department of Neurological Surgery, University of California San Francisco; San Francisco, CA 94143, USA.
[2]Department of Neurology; University of California San Francisco; San Francisco, CA 94143, USA.
[3]Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research; University of California San Francisco; San Francisco, CA 94143, USA.
[4]San Francisco Veterans Affairs Medical Center, San Francisco, CA 94143, USA.
* Corresponding author and lead contact: Daniel.Lim@ucsf.edu

Abstract

Approximately 30-40% of the human genome is anchored to the nuclear lamina through variably sized (10 kb–10 Mb) lamina-associated domains (LADs). Previous work in cultured murine and human cell lines indicate that LADs are not homogenous but can be classified into two subtypes (T1 and T2) based on levels of lamina-association. Due in part to technical limitations, the LAD substructure of neurons maturing *in vivo* is poorly understood. Here, we developed Genome Organization with CUT and Tag (GO-CaT) to map LAD substructure in neurons isolated from the midgestational human cortex and adult human brain. GO-CaT with LaminB1 antibodies mapped LADs with high efficiency as compared to other methods, and in prenatal neurons, we distinguished T1- and T2-LADs based on levels of LaminB1 enrichment. While T1-LADs had the transcriptional and epigenomic characteristics of cell type-invariant LADs including strong transcriptional repression, T2-LADs had a distinct epigenomic state that included enrichment of promoter-enhancer DNA interactions and intermediate levels of gene expression. In prenatal neurons, T2-LADs defined genomic regions that play key roles in neuronal development and were enriched for neurological GWAS phenotypes including those of human cognitive disorders. In a population of adult brain neurons, most prenatal T2-LADs were repositioned to either inter-LAD regions or further sequestered into T1-LADs, perhaps representing the completed, life-long spatial genome architecture of human neurons. These studies highlight the utility of GO-CaT for LAD subdomain mapping in cells isolated from human tissues and illustrate how the substructure of LADs may contribute to neural development, neuronal maturation, and human brain disease.

Introduction

In the interphase nucleus of mammalian cells, chromosomes are radially organized, with specific genomic regions interacting with the lamina of the inner nuclear membrane [1, 2]. At the nuclear periphery, approximately 30-40% of the genome are segmented into variably sized (10 kb to 10 Mb) lamina-associated domains (LADs), and genes within LADs have low transcriptional activity. In studies of cell differentiation in culture, LADs that "detach" from the lamina exhibit increased gene expression [3-6], and conversely, experimental tethering of genes to the lamina suppresses transcription [7-10]. Thus, the lamina constitutes a large, transcriptionally repressive nuclear subcompartment. Although LADs have been widely studied in cells cultured *in vitro*, relatively little is known about LADs in cells acutely isolated from primary tissues, partly due to limitations of methods used for LAD mapping. In particular, the LAD architecture of post-mitotic cells – such as neurons in the developing and adult brain – is poorly understood.

LADs have substructure based on their degree of lamina-association. Using LaminB1 chromatin immunoprecipitation-sequencing (ChIP-seq) of twelve cultured cell lines, Shah *et al*. classified LADs into two subtypes: T1-LADs, which are highly enriched in LaminB1 and transcriptionally repressed, and T2-LADs, which have moderate levels of LaminB1 enrichment and low levels of gene expression [11]. The potential functional implications of such LAD substructure have not been reported for neurons.

During mammalian brain development, proliferative neural precursor cells give rise to young neurons that at early to mid-gestational ages collect in a region called the cortical plate. Neurons in the cortical plate are postmitotic but still immature, acquiring additional functions and becoming incorporated into complex circuits later in development. While most cell types are periodically replaced through a process that involves cell division, post-mitotic neurons do not have such an opportunity to globally "reset" their genome architecture through DNA replication and mitosis. Thus, in neurons, the association of chromatin with the lamina may be of particular importance, providing a physical and regulatory scaffold that needs to last a lifetime.

There are several different methods for LAD mapping, each with distinct advantages and disadvantages. Most LAD maps have been generated by DNA adenine methyltransferase identification (Dam-ID) based methods wherein genomic DNA proximal to the lamina undergoes adenine methylation by a transgenic LaminB1-Dam fusion protein and is subsequently identified by sequencing. LaminB1-DamID is not well suited for studying human primary cells due to the need for transgene expression. Other approaches like LaminB1 ChIP-seq [11] and Tyramide signal amplification (TSA)-seq [12, 13] require tens of millions of cells, limiting their utility for clinical specimens that are usually limited in quantity.

We previously developed Genome Organization with CUT&RUN technology (GO-CaRT), which localizes protein-A micrococcal nuclease (pA-MNase) to LaminB1 antibodies in the nucleus, to map LADs from cells acutely isolated from mouse, macaque, and human brain cells [14]. In this study, we adapted this technique to utilize pA-Tn5 transposase, providing a streamlined approach with higher sensitivity termed Genome Organization with CUT and Tag (GO-CaT). Unlike GO-

CaRT, GO-CaT uses Tn5-mediated "tagmentation" to insert sequencing adaptors during DNA fragmentation, which greatly improved LAD mapping efficiency and simplified the process. The enhanced LAD mapping efficiency of GO-CaT allowed for precise dissection of LADs into T1 and T2 subdomains from small numbers of neurons isolated from prenatal and adult human brain specimens. Together, these data from human neurons *in vivo* provide novel insights into the substructure of 3D genome architecture and implicate dynamics of T1- and T2-LADs in neuronal differentiation, maturation and human disease.

Results

GO-CaT maps LADs with high efficiency

To evaluate the performance of GO-CaT, we first mapped LADs in HEK293T cells, that had been previously mapped by LaminB1 GO-CaRT[14] and LaminB1-DamID-seq[15]. For GO-CaT, we used LaminB1 antibodies to tether pA-Tn5 to the nuclear lamina and identified the tagmented DNA fragments with paired-end sequencing (**Fig.1A**). LADs mapped by GO-CaT were highly similar (Jaccard index 0.68 - 0.84) to those of GO-CaRT and DamID-seq (**Fig.1B, C**). With GO-CaT, under the same sequencing depth, the unique fragments mapped to LADs were 11.3% and 25.1% greater than those of GO-CaRT and DamID-seq, respectively (**Fig. 1D**). Furthermore, at LAD boundaries, there was a larger difference of normalized LaminB1 signal within LADs vs. inter-LAD regions as compared to data from GO-CaRT and DamID-seq (**Fig. 1E**). These data indicate that the pA-Tn5 activity of GO-CaT efficiently produces accurate LAD maps with discrete LAD boundaries.

GO-CaT efficiently maps discrete LAD subtypes

Hidden Markov Model (HMM) algorithms have been used to identify LAD subtypes that differ in their levels of LaminB1 enrichment[11]. Using LaminB1 GO-CaT data from HEK293T cells, we partitioned the entire genome into 100kb bins, computed the normalized LaminB1 occupancy for each unit, and used HMM analysis to determine the number of distinct LaminB1 occupancy states that best fit the data (**Fig. 2A**). Receiver Operating Characteristic (ROC) analysis indicated that the four-state model outperformed the three-, five-, six- and seven-state models (**Fig. 2B, Supplementary Fig. 1A-C**). The four-state model segmented the genome into two LADs and two inter-LADs classifications (**Fig. 2C**). Notably, the two LAD states in our analysis corresponded to the previously described T1- and T2-LAD subtypes (**Supplementary Fig. 1D**)[11]. Consistent with prior studies of T1- and T2-LADs across multiple cell types, genes in HEK293T T1-LADs were strongly repressed, whereas genes in T2-LADs displayed moderate expression levels (**Fig. 2D**).

To investigate the efficiency of GO-CaT for distinguishing T1- and T2-LAD, we compared GO-CaT and GO-CaRT data, down-sampling the total number of uniquely mapped reads from both datasets to 20, 16, 12, 8 and 4 million and using each subsampled dataset for LAD subtypes identification. With GO-CaT data, LAD subtypes identification was highly consistent across the

down-sampled data, having a minimal decrease with fewer reads (**Fig. 2E**). In contrast, with GO-CaRT data, the consistency of T1 and T2-LAD discrimination was decreased as compared to GO-CaT with greater down-sampling (**Fig. 2E**). Additionally, we calculated the genome-wide transition rates at the boundaries of T1- and T2-LADs by the probability gradient computed from the HMM. In GO-CaT data, the transition rates between the two LAD subtypes were very high (transition rate: 0.12) near the boundary even when only 4 million reads were used (**Fig. 2F**). Although the T1- and T2-LAD boundaries were distinctly defined in GO-CaRT when using 20 million reads as an input (transition rate: 0.15), the rate significantly declined after down-sampling, dropping to background levels (transition rate: 0.05) when 4M reads were used (**Fig. 2F**). Taken together, these data highlight the efficiency of GO-CaT in identifying LADs, particularly concerning the distinction of T1- and T2-LAD subtypes.

GO-CaT identifies LAD subtypes in cells of the human cortical plate

An important advantage of GO-CaT is the highly efficient nature of the technique, enabling the study of relatively small numbers of cells. We therefore employed LaminB1 GO-CaT to identify T1- and T2-LADs of human prenatal cortical development, where only small tissue specimens can usually be obtained. From prenatal brain tissues at gestational weeks (GW) 17 and 22, we dissected the cortical plate regions that are enriched for immature excitatory neurons as previously described[16]. From the cortical plate at GW17 and GW22 we isolated approximately 20,000 cells for LaminB1 GO-CaT (**Fig. 3A**) and utilized the four-state HMM for LAD calling (**Fig. 3B**). The four LaminB1 occupancy states between the two cortical plate samples were highly concordant (**Supplementary Fig. 2**). Similar to our observations in HEK293T cells, two of these states corresponded to T1- and T2- LADs, exhibiting robust and moderate LaminB1 occupancy, respectively (**Fig. 3B**).

In cortical plate cells, T1- and T2-LADs exhibited similar genome coverage (741Mb and 757Mb, respectively), with each LAD subtype representing ~24% of the whole genome. Across the genome, T1-LADs were composed of broad domains marked by contiguous LaminB1 occupancy, whereas T2-LADs were characterized by narrower segments interspersed with LaminB1-depleted regions (**Fig. 3B**). Accordingly, the median domain size of T1-LADs (1.2MB) was significantly larger than that of T2-LADs (0.7MB) (**Fig. 3C**). Consistent with previous studies[11], T1-LADs had higher levels of LaminB1 enrichment as compared to T2-LADs (**Fig. 3D**). T2-LADs of the human prenatal cortex had higher gene density (7.5 vs 5.3 genes per MB) and, as expected, greater levels of gene expression as compared to T1-LADs (**Fig. 3E**). Also consistent with prior descriptions of LAD subtypes[11], T2-LADs in cortical plate cells exhibited greater enrichment of SINE but not LINE transposable elements when compared to T1-LADs (**Fig. 3F**). Together, these data indicate that GO-CaT can accurately map T1- and T2-LAD subtypes in relatively small numbers of cells from human prenatal brain tissues.

We next explored the similarity of LAD subtypes between cultured cells and those of the developing human cortical plate. Comparing an atlas of LADs in differentiated cultured human cell types representing 12 different tissues[11], we found that T1-LADs of the cortical plate were more similar across different tissues (Jaccard index: 0.27-0.49) as compared to the T2-LADs,

which were more dissimilar (Jaccard index: 0.18-0.25) (**Fig. 3G**). Of note, T2-LADs of the developing cortical plate were largely distinct from that of cultured brain cells (Jaccard index: 0.20) (**Fig. 3G**), highlighting the unique LAD architecture of neurons *in vivo* as compared to models of neural differentiation in vitro. Together, these data suggest that T1-LADs are more conserved across different cell types while T2-LADs are more tissue type-specific.

T1- and T2-LADs of cells in the human cortical plate have distinct epigenomic characteristics

LADs are generally known to consist of heterochromatin enriched with the repressive histone modification H3K9me2, but the epigenomic differences between the T1- and T2-LAD subtypes are less well understood. We performed CUT and Tag analysis for H3K9me2 as well as H3K27me3 in the same cortical plate samples used for GO-CaT LAD mapping. Consistent with studies of LAD subtype in cultured cells [11], T1-LADs of prenatal cortical plate cells had greater levels of H3K9me2 enrichment as compared to T2-LADs (**Fig. 4A**). Interestingly, T2-LADs were characterized by regions enriched for H3K27me3, which were at lower levels in T1-LADs (**Fig. 4B**). Thus, T1- and T2-LADs have distinct histone modification states in the developing human brain.

DNA-DNA interactions are another important aspect of local chromatin structure[17], and analysis of DNA-DNA interaction data can broadly separate the genome into two major states called compartment A (more likely to be transcribed) and compartment B (less likely to be transcribed) [18-20]. We first analyzed high-throughput chromosome conformation capture (Hi-C) data [21, 22] from prenatal excitatory neurons in conjunction with our T1- and T2-LAD maps. Across the entire genome, 85.7% and 74.1% T1- and T2-LADs, respectively, were designated as compartment B. However, the magnitude of the compartment scores between T1- and T2-LADs was notably different, with T1-LADs being associated with high compartment B scores, whereas T2-LADs had compartment scores that were intermediate to A and B[23] (**Fig. 4C**). This observation demonstrates the differential chromosomal conformation of genomic regions across the T1- and T2-LAD subtypes.

To study DNA-DNA looping interactions between promoters and enhancers, we analyzed anti-H3K4me3 PLAC-seq data in excitatory neurons isolated from GW17 to GW22 prenatal brain tissues [24]. To examine this 3D aspect of local genome organization across the LAD subtypes, we counted the numbers of interactions within each individual T1-, T2- and inter-LAD (*i.e.,* intra-domain DNA-DNA interactions). Notably, we observed a significant enrichment of the intra-domain interactions in T2-LADs as compared to T1-LADs (Fold Change: 10.3, 4.59 per MB vs 0.44 per MB) in the developing cortex (**Fig. 4D, 4E**). Within the T2-LADs, the intra-domain interactions were more than twice as frequent as long-range interactions across the boundaries of T2-LADs (**Fig. 4D, 4E**). Thus, the pronounced transcriptional repression of T1-LADs aligns with high levels of H3K9me2, strong association with compartment B, and relative depletion of promoter-enhancer interactions. By contrast, T2-LADs are characterized by lower H3K9me2 levels, representing atypical heterochromatin with higher levels of H3K27me3, intermediate compartment A/B scores and greater levels of promoter-enhancer interactions. Taken together,

these data illustrate the unique epigenomic features that distinguish T1- and T2-LADs and emphasize the importance of considering LAD substructures when studying 3D genome organization in developing brains.

T2-LADs in human cortical plate cells contain genes important for the development of excitatory neurons

The relative enrichment of promoter-enhancer interactions within T2-LADs suggests a transitional state wherein genes possess 3D genomic interactions that may influence their transcriptional activity as cells differentiate. To explore this notion, we first hypothesized that T2-LADs are enriched for genes that are more likely to be neuronal lineage-specific. By analyzing snRNA-seq data in developing human prefrontal cortex [25], we mapped gene expression variance across terminally differentiated cell types (upper layer excitatory neuron [EN], deep layer EN, astrocyte and microglia) to T1-, T2-LADs, and inter-LADs identified from cortical plate samples. Genes located in T2-LADs (n=3,986) were nearly twice as variable in terms of expression across these differentiated brain cell types as compared to genes in T1-LADs (n=2,878) and inter-LADs regions (n=29,310) (**Fig. 5A**). These data indicate that genes within the T2-LADs of the prenatal cortical plate are more likely to become expressed in a lineage-type-specific manner as compared to T1-LADs or inter-LAD regions.

Notably, genes within T2-LADs of the prenatal cortical plate eventually become expressed and are highly specific to two terminal EN cell types (upper and deep layer) but not to other terminally differentiated states including oligodendrocytes, astrocytes, or microglia (**Fig. 5B**). Because the prenatal cortical plate is enriched for immature ENs [25], these observations support the notion that T2-LADs in immature cell types represent a transitional epigenomic state, containing information necessary for subsequent cellular differentiation and function, such as genes that drive cellular lineage specification, identity, and cell type-specific functions. To explore this notion further, we performed pseudotime and lineage analysis using the snRNA-seq data to identify putative lineage drivers of EN differentiation[26]. We found T2-LADs to be significantly enriched for EN lineage drivers as compared to T1-LADs and inter-LAD regions (**Fig. 5C**). SATB2 is an EN lineage driver gene with key roles in neuronal development (**Supplementary Fig. 3A**), and DNA fluorescent *in situ* hybridization analysis demonstrated that this T2-LAD gene is located further from the nuclear periphery than T1-LAD gene *LINC01677* (**Fig. 5D, E**). Taken together, these findings suggest that T2-LADs represent a transitional epigenomic state that contains developmental information important for subsequent neuronal maturation.

Brain-specific genetic susceptibility variants are enriched in T2-LADs

We next sought to utilize the identified LAD subtypes to interpret genetic variants associated with complex diseases and multigenic phenotypes. We accessed the NHGRI-EBI GWAS catalog [27] and selected 176 well-powered GWAS datasets with over 20,000 cases. Many GWAS primarily report only index SNPs with extreme significance, which often fail to capture causal variants as a substantial portion of trait heritability resides in SNPs with associations below the genome-wide significance threshold. To investigate how this heritability is distributed across LAD subtypes, we

examined the enrichment of GWAS-associated SNPs within gene loci located in T1-, T2-, and inter-LAD regions via stratified linkage disequilibrium score regression (LDSC) [28, 29], a method that quantifies functional enrichment from GWAS summary statistics by leveraging genome-wide SNP data while explicitly accounting for linkage disequilibrium (LD) patterns.

Our analysis revealed a significant difference in per-SNP heritability enrichment across LAD subtypes for multigenic phenotypes and complex diseases (**Table 1**). As expected, risk variants within inter-LADs displayed significant heritability enrichment for 150 complex traits and diseases, but only 15% (26/176) of these traits were neurological phenotypes. Notably, T2-LAD gene loci demonstrated significant genetic enrichment for phenotypes specific to neuronal development (42.1%, 24/57), whereas no enrichment was observed in T1-LAD loci (**Fig. 6A-D**). For example, risk variants associated with autism spectrum disorder (ASD), putamen volume (Mean Putamen), and attention-deficit/hyperactivity disorder (ADHD) were strongly enriched exclusively within T2-LADs, but not within T1-LADs, inter-LADs, or LADs as a whole (**Fig. 6E**). Additionally, nominal enrichment for non-neurological diseases, such as Type 1 diabetes, was detected only in inter-LADs but not in T2-LADs, supporting the hypothesis that T2-LADs are genetically predisposed to neuronal development disorders and may harbor neuronal-specific regulatory elements.

T2-LADs are reduced in mature neurons in adult brain

Given that T2-LADs in prenatal neurons appeared to represent a transitional epigenomic state related to synapse development and function, we hypothesized that the LAD architecture of adult brain neurons would be substantially different, depleted for T2-LADs and perhaps enriched for T1-LADs. Although the cortical plate of prenatal cortex is composed of mostly ENs, there are also a considerable number of inhibitory neurons (INs). To reduce potential heterogeneity of the cell population, we purified EN from the GW20 prenatal cortical plate by fluorescent activated nuclei sorting (FANS) for SATB2-expression and performed LaminB1 GO-CaT (**Fig. 7A**). The LAD substructures identified from purified prenatal ENs recapitulated those of the developing cortical plate (**Supplementary Fig. 4A**), where T2-LADs account for ~46.8% of total LADs.

To investigate the LAD architecture of adult brain neurons, we performed FANS to purify SATB2+ EN from postmortem specimen of adult frontal lobe cortex (49-year-old male). We sorted for LaminB1 GO-CaT on the isolated nuclei that were also positive for the NeuN nuclear antigen (**Fig. 7A**). Leveraging our four-state HMM, we mapped both the T1- and T2-LAD subtypes in adult EN (**Fig. 7B**). Strikingly, in adult EN, the proportion of the genome found in T2-LADs (423 Mb, 15.1% of the whole genome) was about half of that found in prenatal EN (837Mb,27.1% of the whole genome) (**Fig. 7C**). This apparent loss of T2-LADs in adult EN was mostly accounted for by two groups of changes: First, ~35% of T2-LADs in prenatal EN (303 Mb, 10.8% of the whole genome) were relocated to inter-LAD regions in adult EN (**Fig. 7C**). Second, ~43% of prenatal neuronal T2-LADs (372 Mb, 13.3% of the whole genome) were sequestered to T1-LAD regions, increasing the overall proportion of T1-LADs in the LAD genome of adult brain neurons (**Fig. 7C**). In contrast to the highly dynamic nature of prenatal neuronal T2-LADs, ~72% of T1-LADs (478 Mb, 17.0% of the whole genome) remained as T1-LADs in neurons of the adult cortex (**Fig. 7C**).

Only a small proportion of prenatal inter-LAD regions were relocated to T1- and T2-LADs in adult EN (~14% and ~18%, respectively) (**Fig. 7C**). Overall, the broad maintenance of T1-LADs and additional shift of T2-LADs to T1-LADs identity resulted in a ~27.1% increase (201 Mb, 7.2% of the whole genome) in the proportion of T1-LADs in the adult EN (**Fig. 7C**).

To characterize the expression level of genes that relocated from T2-LADs to T1-LADs and inter-LAD regions, we analyzed snRNA-seq data from Herring et al., which profiled excitatory neurons across 24 developmental time-points from GW22 to 40 years of age [25]. To quantify the rate of gene expression change over development, we fit expression trends for each gene to a regularized regression model. Genes that transited from T2- to T1-LADs in adult EN exhibited lower levels of gene expression (**Fig. 7D**). In contrast, genes that transitioned from T2-LADs to inter-LAD regions showed increased expression levels as compared to genes that remained in T2-LADs (p=5.92e-7) or those that were within regions that transitioned from T2- to T1-LADs (T2-to-T1 genes, p=1.27e-6) (**Fig. 7D**). Of note, these expected gene expression trends were not observed in snRNA-seq data from IN or oligodendrocyte cells (**Fig. 7D**), confirming that our analysis of adult EN primarily represented the LAD architecture of EN.

Genes that relocated from prenatal EN T2-LADs to inter-LAD regions in adult EN (e.g., *CHRM3* and *DMD*) were enriched in gene ontology (GO) terms related to calcium signaling responses (GO: 1901385, FDR=0.024; GO: 0051592, FDR=0.031) (**Fig. 7E, Supplementary Fig. 4B**), supporting previous findings of increased functional demands of human neuronal maturation for rapid calcium influx to support synaptic transmission and complex processes like learning and memory [30-33]. While T1-LADs found in both prenatal and adult brain were not enriched for any GO terms, genes that relocated from T2- to T1-LADs in adult EN (e.g., *SEMA6D* and *LPAR1*) were enriched for GO terms related to neuronal migration and growth (GO:0014032, FDR=0.036; GO: 0001764, FDR=0.045) (**Fig. 7F, Supplementary Fig. 4B**), which are essential for early postmitotic neurons but are less involved after neural circuits are established[33-36]. Together, these findings suggest that LAD substructure matures across human neurodevelopment, underlying a shift away from developmental functions to those that are specialized to the functions and demands of fully differentiation neurons.

DISCUSSION

Although LADs have been studied in numerous different cell lines in culture[11], there are very few reports of this aspect of 3D genome architecture in cells *in vivo*. Leveraging the advantages of Tn5 and related CUT&Tag methods, LaminB1 GO-CaT exhibited greater efficiency for LAD mapping, enabling the study of relatively small numbers of cells acutely isolated from tissues. Using GO-CaT, we identified the LAD substructure of neuronal cell populations from the developing and adult human brain. Our studies of T1- and T2-LADs in the mid-gestational cortical plate revealed novel insights into the potential roles of LAD substructure in neurogenesis and phenotypes related to genetic variation. Mapping LAD substructure in a population of adult brain neurons provides foundational groundwork for understanding how spatial 3D genome architecture can mature in cells that remain postmitotic for decades.

For LAD mapping, LaminB1 GO-CaT has performance characteristics that make it well-suited for experimental scenarios in which samples are limited (*e.g.* clinical specimens and FACS-sorting of rare cell populations). In our benchmarking studies, GO-CaT mapped LADs with high confidence from as few as 10,000 cells. Subsampling analyses revealed LaminB1 GO-CaT to have the lowest marginal input cell requirement as compared to DamID- and GO-CaRT-based methods. Furthermore, with GO-CaT data, T1- and T2-LAD subdomains had clearer boundaries, resulting in higher resolution and lower background during LAD calling.

In addition to confirming previously described key characteristics of LAD subdomains, we discovered additional epigenomic characteristics that distinguish T1- and T2-LADs in prenatal cortical plate cells. Consistent with prior observations made in cultured human cell lines[11], T1-LADs in primary neuronal cells exhibited high levels of H3K9me2 and correspondingly low gene expression. Adding to the understanding of T2-LADs being a less repressed LAD state, we found that T2-LADs in prenatal brain cells have relatively higher levels of H3K27me3, greater association with compartment A, and more frequent DNA interactions between promoters and potential distal regulatory sequences. Collectively, this epigenomic signature of T2-LADs suggests that this LAD substructure represents a distinct chromatin state that may be poised for transcriptional activation.

The higher levels of DNA-DNA interactions with promoters in T2-LADs suggests a chromatin configuration that allows genes to be activated in a lineage-specific manner as they differentiate. Genes located in T2-LADs demonstrated higher variability in expression across differentiated brain cell types as compared to T1-LADs and inter-LADs regions, indicating that T2-LADs harbor lineage-specific genes essential for neural cell identity. These genes were enriched for neuronal-specific activities, such as neuronal projection development (GO:0010976, FDR=0.0014; GO:0048666, FDR=0.0031), while T1-LADs genes were related to more general cellular functions like cell-cell adhesion and cell cycle (GO:0098742, FDR=9.39E-04; GO:0090068, FDR=0.0049). Overall, genes in T2-LADs were predominantly activated in terminally differentiated ENs, emphasizing the role of T2-LADs in encoding information relevant for EN lineage specification.

In adult neurons, the landscape of LADs was substantially reorganized, as compared to prenatal cells. Most notably, adult neurons have a reduction in T2-LADs. This shift appears to reflect a maturation process wherein T2-LADs are repositioned to more distinct chromatin states, either transitioning to inter-LAD regions or T1-LADs. Genes that relocated from T2- to T1-LADs in adult neurons exhibited reduced gene expression, consistent with the role of T1-LADs as repressive chromatin domains that help maintain a transcriptionally restrained environment. The observed 27.1% increase in T1-LAD coverage may signify the establishment of a more differentiated genome architecture, supporting neuronal function while restricting the processes that are more relevant to earlier stages of neuronal maturation. The relocation of T2-LADs to inter-LAD regions in adult neurons corresponded to increased gene expression, suggesting a transition from a poised to active epigenomic state. More broadly, the comparison of prenatal and adult neuron LAD substructure provides insights into how the spatial genome architecture

may "mature" in postmitotic cells to achieve a long-lived, cell-type specific transcriptional regulatory scaffold.

Our integrative genome analyses highlight enriched genetic susceptibility of genome variations in T2-LADs to neurodevelopmental disorders like ASD and ADHD, suggesting their specialized roles in defining neuronal identity. In contrast, inter-LADs are broadly associated with human phenotypes but show limited relevance to neurological traits, further emphasizing the specificity of T2-LADs to brain-related genetic predispositions. These findings challenge the view of LADs as purely structural and propose T2-LADs are integral to the genetic regulation of neuronal development. The unique enrichment of T2-LADs for neurodevelopmental disorders suggests that these regions may represent a critical interface between genetic risk and cellular function in the brain. Future studies focused on how T2-LADs facilitate dynamic gene regulation during neurodevelopment may reveal new epigenomic principles that underlie neurodevelopmental diseases.

References

1.   Belmont AS. Nuclear Compartments: An Incomplete Primer to Nuclear Compartments, Bodies, and Genome Organization Relative to Nuclear Architecture. *Cold Spring Harb Perspect Biol* 2022 Jul 1; **14**(7).

2.   Bickmore WA, van Steensel B. Genome architecture: domain organization of interphase chromosomes. *Cell* 2013 Mar 14; **152**(6)**:** 1270-1284.

3.   Brueckner L, Zhao PA, van Schaik T, Leemans C, Sima J, Peric-Hupkes D*, et al.* Local rewiring of genome-nuclear lamina interactions by transcription. *EMBO J* 2020 Mar 16; **39**(6)**:** e103159.

4.   Guerreiro I, Rang FJ, Kawamura YK, Kroon-Veenboer C, Korving J, Groenveld FC*, et al.* Antagonism between H3K27me3 and genome-lamina association drives atypical spatial genome organization in the totipotent embryo. *Nat Genet* 2024 Oct; **56**(10)**:** 2228-2237.

5.   Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, Brugman W*, et al.* Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell* 2010 May 28; **38**(4)**:** 603-613.

6.   Poleshko A, Shah PP, Gupta M, Babu A, Morley MP, Manderfield LJ*, et al.* Genome-Nuclear Lamina Interactions Regulate Cardiac Stem Cell Lineage Restriction. *Cell* 2017 Oct 19; **171**(3)**:** 573-587 e514.

7.   Reddy KL, Zullo JM, Bertolino E, Singh H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 2008 Mar 13; **452**(7184)**:** 243-247.

8.   Finlan LE, Sproul D, Thomson I, Boyle S, Kerr E, Perry P*, et al.* Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genet* 2008 Mar 21; **4**(3)**:** e1000039.

9.   Wang H, Xu X, Nguyen CM, Liu Y, Gao Y, Lin X*, et al.* CRISPR-Mediated Programmable 3D Genome Positioning and Nuclear Organization. *Cell* 2018 Nov 15; **175**(5)**:** 1405-1417 e1414.

10.   van Steensel B, Belmont AS. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* 2017 May 18; **169**(5)**:** 780-791.

11.   Shah PP, Keough KC, Gjoni K, Santini GT, Abdill RJ, Wickramasinghe NM*, et al.* An atlas of lamina-associated chromatin across twelve human cell types reveals an intermediate chromatin subtype. *Genome Biol* 2023 Jan 23; **24**(1)**:** 16.

12.   Chen Y, Zhang Y, Wang Y, Zhang L, Brinkman EK, Adam SA*, et al.* Mapping 3D genome organization relative to nuclear compartments using TSA-Seq as a cytological ruler. *J Cell Biol* 2018 Nov 5; **217**(11)**:** 4025-4048.

13.   Kumar P, Gholamalamdari O, Zhang Y, Zhang L, Vertii A, van Schaik T*, et al.* Nucleolus and centromere TSA-Seq reveals variable localization of heterochromatin in different cell types. *bioRxiv* 2023 Nov 1.

14.   Ahanger SH, Delgado RN, Gil E, Cole MA, Zhao J, Hong SJ*, et al.* Distinct nuclear compartment-associated genome architecture in the developing mammalian brain. *Nat Neurosci* 2021 Sep; **24**(9)**:** 1235-1242.

15.   Altemose N, Maslan A, Rios-Martinez C, Lai A, White JA, Streets A. muDamID: A Microfluidic Approach for Joint Imaging and Sequencing of Protein-DNA Interactions in Single Cells. *Cell Syst* 2020 Oct 21; **11**(4)**:** 354-366 e359.

16.   Wang L, Wang C, Moriano JA, Chen S, Zuo G, Cebrian-Silla A*, et al.* Molecular and cellular dynamics of the developing human neocortex at single-cell resolution. *bioRxiv* 2024 Aug 4.

17.   Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT*, et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 2014 Dec 18; **159**(7)**:** 1665-1680.

18.   Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A*, et al.* Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 2009 Oct 9; **326**(5950)**:** 289-293.

19.   Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* 2013 Jun; **14**(6)**:** 390-403.

20.   Gibcus JH, Dekker J. The hierarchy of the 3D genome. *Mol Cell* 2013 Mar 7; **49**(5)**:** 773-782.

21.   Heffel MG, Zhou J, Zhang Y, Lee DS, Hou K, Pastor-Alonso O*, et al.* Temporally distinct 3D multi-omic dynamics in the developing human brain. *Nature* 2024 Nov; **635**(8038)**:** 481-489.

22.   Lee DS, Luo C, Zhou J, Chandran S, Rivkin A, Bartlett A*, et al.* Simultaneous profiling of 3D genome structure and DNA methylation in single human cells. *Nat Methods* 2019 Oct; **16**(10)**:** 999-1006.

23.   Harris HL, Gu H, Olshansky M, Wang A, Farabella I, Eliaz Y*, et al.* Chromatin alternates between A and B compartments at kilobase scale for subgenic organization. *Nat Commun* 2023 Jun 6; **14**(1)**:** 3303.

24.   Song M, Pebworth MP, Yang X, Abnousi A, Fan C, Wen J*, et al.* Cell-type-specific 3D epigenomes in the developing human cortex. *Nature* 2020 Nov; **587**(7835)**:** 644-649.

25.   Herring CA, Simmons RK, Freytag S, Poppe D, Moffet JJD, Pflueger J*, et al.* Human prefrontal cortex gene regulatory dynamics from gestation to adulthood at single-cell resolution. *Cell* 2022 Nov 10; **185**(23)**:** 4428-4447 e4428.

26.   Lange M, Bergen V, Klein M, Setty M, Reuter B, Bakhti M*, et al.* CellRank for directed single-cell fate mapping. *Nat Methods* 2022 Feb; **19**(2)**:** 159-170.

27. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C*, et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 2019 Jan 8; **47**(D1)**:** D1005-D1012.

28. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C*, et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015 Mar; **47**(3)**:** 291-295.

29. Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh PR*, et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* 2015 Nov; **47**(11)**:** 1228-1235.

30. Dolphin AC, Lee A. Presynaptic calcium channels: specialized control of synaptic neurotransmitter release. *Nat Rev Neurosci* 2020 Apr; **21**(4)**:** 213-229.

31. Higgins D, Graupner M, Brunel N. Memory maintenance in synapses with calcium-based plasticity in the presence of background activity. *PLoS Comput Biol* 2014 Oct; **10**(10)**:** e1003834.

32. Kennedy MB. Synaptic Signaling in Learning and Memory. *Cold Spring Harb Perspect Biol* 2013 Dec 30; **8**(2)**:** a016824.

33. Wallace JL, Pollen AA. Human neuronal maturation comes of age: cellular mechanisms and species differences. *Nat Rev Neurosci* 2024 Jan; **25**(1)**:** 7-29.

34. Bron R, Vermeren M, Kokot N, Andrews W, Little GE, Mitchell KJ*, et al.* Boundary cap cells constrain spinal motor neuron somal migration at motor exit points by a semaphorin-plexin mechanism. *Neural Dev* 2007 Oct 30; **2:** 21.

35. Nakanishi Y, Izumi M, Matsushita H, Koyama Y, Diez D, Takamatsu H*, et al.* Semaphorin 6D tunes amygdalar circuits for emotional, metabolic, and inflammatory outputs. *Neuron* 2024 Sep 4; **112**(17)**:** 2955-2972 e2959.

36. Yung YC, Stoddard NC, Mirendil H, Chun J. Lysophosphatidic Acid signaling in the nervous system. *Neuron* 2015 Feb 18; **85**(4)**:** 669-682.

Figure Legend

Figure 1. GO-CaT is an efficient method for profiling genome-wide LAD structure
(**A**) A working scheme of GO-CaT workflow. (**B**) Genome tracks comparing LaminB interaction profiles captured by LaminB DamID, GO-CaRT and GO-CaT in HEK293T cells at a representative region (chr1:63,888,968-109,339,020). The Y axis represents the log ratio of LaminB signals over negative controls for each experiment. Identified LADs were denoted by solid horizontal bars. (**C**) Correlation heatmap showing the Jaccard similarity of between LADs identified by LaminB1 DamID, GO-CaRT and GO-CaT in HEK293T cells. (**D**) Percentage of sequenced reads in LADs identified by DamID (orange), GO-CaRT (dark blue) and GO-CaT (light blue) in HEK293T. All experiments were sub-sampled to have equivalent number of reads and sequencing depth for comparison. (**E**) Metaplot showing normalized LaminB1 enrichment across all 0.5Mb flanking regions of the boundaries that distinguish LADs from inter-LAD in HEK293T.

Figure 2. GO-CaT efficiently maps d discrete LAD subtypes from a four-state-HMM. (**A**) A description of workflow LAD subtypes identification by Hidden Markov Model (HMM) on GO-CaT experiments. (**B**) Benchmarking HMM performance by Area Under Receiver Operating Characteristic curve (AUROC) with different numbers of chromatin states from three to seven. (**C**) Genome tracks highlighting a representative region containing all four LADs subtypes in HEK293T. (**D**) Bar plot showing genome-wide gene expression difference between T1- and T2-LADs in HEK293T cells. P value was determined by Wilcoxon rank sum test, and **** denotes a p value less than 0.0001. (**E**) Line plots demonstrating higher robustness in the T1-(left) and T2-LADs (right) identification in GO-CaT as compared to GO-CaRT when decreasing subsampling rate from 20M to 4M reads. (**F**) Metaplots comparing chromatin alternate rate across all T1- and T2-LAD boundaries between GO-CaT and GO-CaRT when subsampling reads from 20M to 4M. Boundary scores were defined as the averaged transitioning probability from one LAD subtype to another for each 10kb bin predicted by HMM.

Figure 3. Characteristics of LAD subtypes identified by GO-CaT in the human prenatal brain. (**A**) Schematic workflow of GO-CaT experiments performed on cells acutely isolated from dissected cortical plates of 2nd trimester human brain. (**B**) Representative genome tracks of genetic locus of four distinct LADs subtypes identified in developing brains of GW22 (upper panel) and GW17 (bottom panel). (**C**) Violin plot depicting the size of T1-LADs and T2-LADs identified in cortical plates (in Mb, million base pairs). (**D**) Metaplot showing the normalized LMNB1 occupancy enrichment over T1-(dark blue) and T2-LADs (light blue). All LADs were scaled to the same size as represented by the solid horizontal bars. (**E**) Bar plots depicting gene density (left panel) and average gene expression (right panel) in T1- and T2-LADs as determined by RNA-seq in purified EN. (**F**) Bar plots comparing the percentage of the regions covered by SINE (left) and LINE (right) retrotransposons between T1- and T2-LADs. (**G**) Correlation heatmap illustrating the Jaccard similarity of identified T1- (left) and T2-LADs (right) across 13 different human tissue/cell types including prenatal cortical plate. *: p value < 0.05; **: p value < 0.01; *** : p value < 0.001; ****: p value <0.0001; All p values were determined by Wilcoxon rank-sum test.

Figure 4. T2-LADs are associated with distinct epigenomic characteristics in the prenatal human cortical plate. (**A, B**) Metaplots showing normalized H3K9me2 (**A**) and H3K27me3 (**B**) enrichment signals across T1-LADs (dark blue) and T2-LADs (light blue) of prenatal cortical plates. All LADs were scaled into a size of 0.5Mb. (**C**) Metaplots showing compartment A/B scores across 100KB flanking regions of all LAD-interLAD boundaries grouped by LAD subtypes (T1: dark blue, T2: light blue). Compartment scores were inferred from HiC experiment in neurons, and a positive value indicates the state of type A compartment. (**D**) Comparison of promoter-enhancer interaction frequency within the same LAD subtypes (intra-subdomain) and crossing the boundaries (inter-subdomain) for T1-, T2- and inter-LADs. All P values were

from Wilcoxon rank-sums test and corrected by Benjamini-Hochberg method. **** denotes adjusted P values less than 0.0001. (**E**) Genome tracks showing different types of chromatin interactions within the annotated LAD subtypes at one representative genome locus (chr4: 72,995,600-99,124,210). The conceptual scheme that illustrates intra- and inter-subdomain chromatin interactions.

Figure 5. T2-LADs in human cortical plate contain information important for excitatory neuronal development. (**A**) Box plot showing the standardized expression variability across the genes located within T1-, T2-, and inter-LADs. (**B**) Box plots showing the significantly higher gene expression specificity in EN lineage cell types (Upper Layer and Deep Layer EN) for genes located in T2-LADs comparing to those genes from T1-LADs or genome background. Cell type-specific expression was quantified as the log2-fold changes over the mean expression of all other cell types. All P values were from Wilcoxon rank-sum tests and corrected by Benjamini-Hochberg method. **** denotes adjusted P value less than 0.0001. (**C**) Stack barplots demonstrating overrepresentation of neuronal lineage drivers in T2-LADs over T1-LADs and inter-LADs. The proportion of the bar in black denotes the percentage of lineage driver genes inferred from scRNA-seq analysis. All P values were from Fisher's exact test and corrected by Benjamini-Hochberg method. **** denotes adjusted P value less than 0.0001. (**D**) Representative DNA-FISH images in the nucleus of two representative locus (*LINC01677*, and *SATB2*) (**E**) Box plot quantifying the mean distance of each locus to nuclear lamina.

Figure 6. Genetic analysis of LAD subtypes with complex neuropsychiatric disorder-associated variants. (**A-D**) Scatter plots showing the enriched per-SNP heritability from partitioned LDSC analyses for SNPs associated with genes from (**A**) inter-LADs, (**B**) LADs as a whole, (**C**) T2-LADs, and (**D**) T1-LADs. X-axis: P values in minus log10 scale; Y-axis: Enrichment of heritability for each SNP. Each dot represents a GWAS phenotype tested and is color-coded by their significance of enrichment. (**E**) Comparisons of per-SNP heritability enrichment between genetic variants associated with different LAD subtypes on selected complex traits and multigenic diseases (e.g. autism, mean volume of putamen, ADHD, and type I diabetes). ***: P values < 0.001; **: P value < 0.01; ns: not significant.

Figure 7. LAD subtypes are extensively remodeled during human neuronal maturation. (**A**) Schematic of the sorting strategy of prenatal and adult EN. (**B**)Genome track of representative regions (chr2:140,307,830-241,801,469) of four different LAD subtypes (T1-LAD, T2-LAD, T3-inter-LAD, and T4-inter-LAD) identified by GO-CaT in prenatal and adult EN. (**C**) Alluvial plot illustrating the genome wide transitions of T1-LADs, T2-LADs and inter-LADs between prenatal and adult EN. (**D**) Boxplots showing genome-wide gene expression changes within excitatory neurons (ENs), inhibitory neurons (IN) and oligodendrocytes (Oligo) during neuronal maturation. Genes located within regions transited from T2-LAD to inter-LAD, from T2-LAD to T2-LAD, and from T2-LAD to T1-LAD are color-coded. Dashed red line represents genome background. Statistical significance indicated as **** (p < 0.0001) or ns (not significant). (**E,F**) Dot plots of gene ontology analysis for genes transitioning from T2-LADs in prenatal EN to inter-LADs in adult EN (**E**), from T2-LADs in prenatal EN to T1-LADs in adult EN (**F**). FDR (adjusted P values) are color-coded, and size of the dots represents the enrichment odds ratio.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7