

RESEARCH

Open Access



Analysis of long-range chromatin contacts, compartments and looping between mouse embryonic stem cells, lens epithelium and lens fibers

Michael Camerino¹, William Chang² and Ales Cvekl^{1,2*}

Abstract

Background Nuclear organization of interphase chromosomes involves individual chromosome territories, “open” and “closed” chromatin compartments, topologically associated domains (TADs) and chromatin loops. The DNA- and RNA-binding transcription factor CTCF together with the cohesin complex serve as major organizers of chromatin architecture. Cellular differentiation is driven by temporally and spatially coordinated gene expression that requires chromatin changes of individual loci of various complexities. Lens differentiation represents an advantageous system to probe transcriptional mechanisms underlying tissue-specific gene expression including high transcriptional outputs of individual crystallin genes until the mature lens fiber cells degrade their nuclei.

Results Chromatin organization between mouse embryonic stem (ES) cells, newborn (P0.5) lens epithelium and fiber cells were analyzed using Hi-C. Localization of CTCF in both lens chromatins was determined by ChIP-seq and compared with ES cells. Quantitative analyses show major differences between number and size of TADs and chromatin loop size between these three cell types. In depth analyses show similarities between lens samples exemplified by overlaps between compartments A and B. Lens epithelium-specific CTCF peaks are found in mostly methylated genomic regions while lens fiber-specific and shared peaks occur mostly within unmethylated DNA regions. Major differences in TADs and loops are illustrated at the ~500 kb *Pax6* locus, encoding the critical lens regulatory transcription factor and within a larger ~15 Mb *WAGR* locus, containing *Pax6* and other loci linked to human congenital diseases. Lens and ES cell Hi-C data (TADs and loops) together with ATAC-seq, CTCF, H3K27ac, H3K27me3 and ENCODE *cis*-regulatory sites are shown in detail for the *Pax6*, *Sox1* and *Hif1a* loci, multiple crystallin genes and other important loci required for lens morphogenesis. The majority of crystallin loci are marked by unexpectedly high CTCF-binding across their transcribed regions.

Conclusions Our study has generated the first data on 3-dimensional (3D) nuclear organization in lens epithelium and lens fibers and directly compared these data with ES cells. These findings generate novel insights into lens-specific transcriptional gene control, open new research avenues to study transcriptional condensates in lens fiber cells, and enable studies of non-coding genetic variants linked to cataract and other lens and ocular abnormalities.

*Correspondence:

Ales Cvekl
ales.cvekl@einsteinmed.edu

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Chromatin, CTCF, Differentiation, DNA looping, ES cells, Hi-C, lens, Pax6, Topologically associated domains

Introduction

Transcriptional regulation of individual genes is primarily mediated by sequence-specific DNA-binding transcription factors bound to promoters and distal enhancers, local recruitment of specific chromatin remodeling enzymes/complexes, and generation of specific combinations of local histone posttranslational modifications (PTMs) that either facilitate or inhibit formation of molecular interactions required for active transcription. These regulatory mechanisms operate in the context of cell type-specific 3D-organization of the nucleus [1–3]. DNA organization involves individual chromosome territories, transcriptionally active and inactive chromatin domains, TADs, and chromatin loops of various sizes [2]. Thus, chromatin folding is critical for compacting DNA within the nuclear space and provides platform for transcription coupled with RNA splicing, DNA replication, DNA repair, and regulation of chromosome structure and maintenance [1]. Phase separation to generate nuclear condensates is now considered as the driving force of chromatin folding [1–3].

Interphase chromosomes have discrete territories within the nucleus [4, 5]. It has been shown that active gene expression and chromatin state is correlated with positioning of individual loci within the nucleus. Active transcription primarily occurs towards the center, whereas transcriptionally inactive heterochromatin is positioned toward the nuclear periphery [6–8]. The major features of higher-order genome organization are driven by the functional status of the genome including distribution of molecular complexes regulating transcription and other processes in a self-organizing system. In turn, architectural features of the genome modulate its function [2]. A series of studies have shown that the genome organization is cell-type specific and directly linked to tissue-specific transcription [3, 9–12]. Recent studies based on chromatin structural modeling suggest fast chromatin dynamics including promoter-enhancer loops [13]. Studies focused on individual cell types during embryonic development are thus important for our understanding how chromatin structure relates to the genome function.

Regulation of long-range chromatin interactions is mediated by structural maintenance complexes (SMCs) including cohesin and condensin [14–16]. The multifunctional sequence-specific DNA-binding transcription factor CTCF is the most prominent protein defining boundaries of the extruded loops [17–20]. CTCF is comprised from centrally-located 11 zinc-fingers [21] with five (ZF3–7) and two (ZF1 and ZF10) of them involved in DNA and RNA binding, respectively [22–24]. Both

the N- and C-terminal portions of CTCF contain intrinsically disordered regions (IDRs) [25]. Within the interphase chromosomes, TADs are formed at scale of several hundred kilo base pairs (kb) through loop-extrusion mechanism involving ATP-dependent molecular motor activity of the cohesin complex comprised of SMC1, SMC3, RAD21, and SCC3 subunits [26, 27]. Mapping of *in vivo* CTCF chromatin binding using ChIP-seq together with its gene loss-of-function studies revealed an unexpected complexity of downstream effects regarding promoter-enhancer interactions, chromatin looping and transcription, and variable dependence of these processes on both CTCF and cohesion [20, 28–31]. Super-resolution imaging revealed CTCF clusters of 4–8 molecules with approximately a quarter of them coupled with 3–15 cohesin molecules that are separated from RNA polymerase II clusters showing that cohesin and transcription have contrasting functions in CTCF clustering [32].

High-throughput chromosome conformation capture (Hi-C) employs a chromosome conformation capture (3C) technique coupled with next generation DNA sequencing and is currently used to generate organizational maps of chromatin interactions of individual tissues [33, 34]. This approach allows for high-throughput and unbiased data analysis of chromatin organization at a resolution under 10 kb [35]. Hi-C allows detection of compartmentalization status defined as “active compartments A” and “inactive compartments B”, TADs, and inter/intra chromosomal interactions [36, 37]. Combined with CTCF DNA-binding data via ChIP-seq, ATAC-seq data and relevant histone PTMs, Hi-C proves to be a powerful tool in identifying novel candidate promoter-enhancer interactions [38–40].

Embryonic development generates over 400 different basic cell types in the mammalian body from a single fertilized egg [41]. Ocular lens is a unique avascular tissue comprised from two types of cells of common origin from the anterior pre-placodal ectoderm called lens epithelium and lens fibers [42–45]. The anterior portion of the lens is comprised of an epithelial cell layer. Epithelial cells at the equatorial zone divide, migrate and differentiate into secondary fiber cells forming outer layers of the lens fiber cell compartment. Within the lens fiber cell compartment, lens transparency requires formation of organelle free zone (OFZ) to prevent light scattering with major consequences for gene expression control [46]. Between E16.5 to E18.5 of mouse embryonic development, lens fiber cell denucleation includes changes of the nuclear shape and size reduction, chromatin condensation [47], transfer of nuclear proteins into the cytoplasm,

up-regulation of lens-specific acidic DNase II β , phosphorylation of nuclear lamin A and C proteins by Cdk1, and culminating in abrupt disintegration of the individual nuclei within the centrally located “primary” lens fiber cell compartment [48–52]. Surprisingly, nascent crystallin gene expression remains at their maximal levels in these reorganizing nuclei [51]. At steady state levels, expression of crystallin genes ranks among the highest of any biological system found in nature, only comparable to globin genes in red blood cells [53]. Disrupted lens fiber cell denucleation results in both congenital and cortical cataracts [54–56].

A systematic analysis of the spatial and temporal organization of the genome is conducted by the four-dimensional nucleome (4DN) consortium [57]; however, no studies of the ocular cells are included. In the eye, earlier Hi-C studies analyzed mouse neural retina at different developmental stages [58], human adult neural retina [59] and human corneal limbal cells [60]. In the lens, temporal regulation of promoter-enhancer looping within the mouse α A-crystallin locus analyzed by 3 C also revealed two shadow enhancers [61]. To examine global 3D nuclear organization of the lens, we performed Hi-C and CTCF ChIP-seq using microdissected newborn mouse lenses and included ES cells for direct comparative analyses. Our earlier RNA-seq, ATAC-seq and whole genome bisulfite sequencing (WGBS) data from similar samples [62–64] serve for direct integration with nucleome mapping tools. ChIP-seq data including Pax6, RNA polymerase II and histone PTMs from newborn

lens chromatin are also available [53, 65] for extended analyses. The main findings show local functional differences in nuclei from the lens epithelium compared to lens fibers and marked differences compared to ES cells and localization of majority of crystallin loci outside of chromatin loops.

Results

Hi-C sequencing identifies chromatin reorganization in differentiating lens

To examine nucleome dynamics between the newborn mouse lens epithelial and fiber cells, we employed deep Hi-C sequencing. Parallel studies included mouse ES cells as a reference as it represents “ground” state of embryonic development and chromatin organization [9, 66]. A schematic of project workflow is shown in Fig. 1. To elucidate lens cell-type specific changes in chromosome organization, we initially investigated chromatin looping, TADs, compartments A and B, and chromatin state. With a total of 4.55 billion read pairs generated in ES cells, lens epithelium and lens fiber cells (1.48, 1.57, 1.50 billion read pairs, respectively), we detected 1.93 billion Hi-C contacts (Additional File 1: Table S1). Combining both biological replicates of ES cells, lens epithelium and lens fiber cells generated similar total number of 643.2, 588.0, and 699.1 million individual Hi-C contacts, respectively. The ES cells showed a far lower proportion of inter-chromosomal contacts (10.8%) compared to both lens epithelium (24.1%) and lens fiber cells (32.2%) (Additional File 1: Table S1). To determine the predicted resolution of the

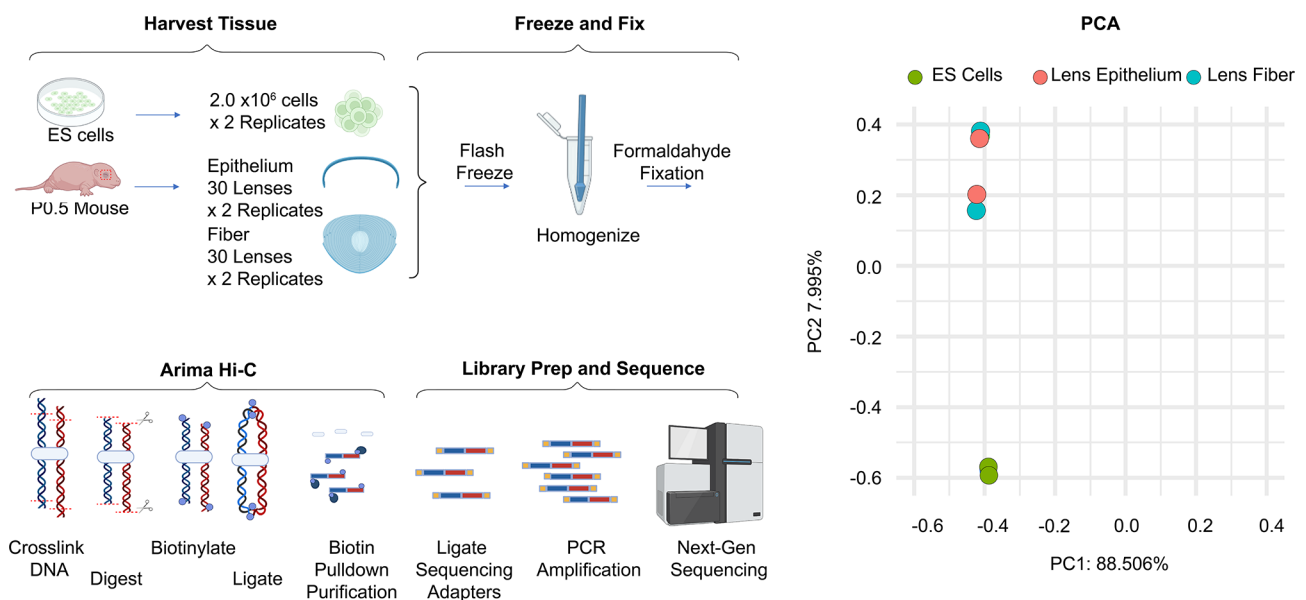


Fig. 1 Tissues and experimental design of the Hi-C study. Lenses were harvested from newborn (P0.5) mice and microdissected into epithelium and fiber cells (30 lenses x 2 replicates). Mouse ES cells were harvested near ~80% confluency (2.0 x 10⁶ cells x 2 replicates). Cells were crosslinked and processed according to Arima’s library preparation protocol. Chromatin contact maps were generated using the ENCODE pipeline. Differential compartment A/B analysis was performed using dHiC (see Materials and Methods). The principal component analysis (PCA) on compartmental eigenvectors shows that both lens cells are distinct from the ES cells

contact maps we used HiCRes [67]. When filtering out reads with $\text{MAPQ} < 30$ ESCs, lens epithelium, and lens fiber cells, all had a similar predicted resolution of 2.69–3.32 kb using 545–595 million read pairs (Additional File 2: Fig. S1). For A/B compartment analysis, dcHiC procedure was used [68]. As expected, principal component analysis (PCA) on A/B compartment status shows both lens cell types clustered together and separated from ES cells (Fig. 1).

Large scale changes in chromatin domains during lens cell differentiation

First, to demonstrate differences between ES cells and both lens cells at the chromosomal (250 kb resolution) and the locus level (10 kb resolution), interaction maps of chromosome 3 are displayed in Fig. 2a. For example, there are notable differences in contact density and long-range chromatin interactions (denoted by arrows) such as within the chr3: 67,000,000–75,000,000 regions between the ES and both lens cells (Fig. 2b).

Next, we performed comparative analyses of chromatin loops and TADs. While similar number of loops were found between ES cells ($n=8,459$), lens epithelium ($n=8,005$) and lens fiber cells ($n=7,682$), loop sizes and their distribution varied greatly. ES cells had a significantly higher mean loop size (989 kb) compared to the lens epithelium (582 kb) and lens fiber cells (581 kb) but all shared similar median sizes (220, 250 and 269 kb, respectively), indicating a higher proportion of long-range loop formations in ES cells (Fig. 2c). This is consistent with previous studies, showing that these interactions are reorganized after individual cell-fate specific differentiation programs [69–71]. ES cells had the largest number of TADs ($n=3,604$) compared to both lens epithelium ($n=1,796$) and lens fiber cells ($n=2,641$). Interestingly, ES and lens fiber cells shared similar median TAD sizes (145 and 135 kb, respectively), whereas lens epithelium median TAD size was more than 2-fold larger (335 kb), indicating reorganization of the nucleome during lens epithelial to fiber cell transition (Fig. 2d). Cell type-specific chromatin loop and TAD size trends are mostly consistent across each individual 20 chromosomes (Additional File 2: Fig. S2). Next, chromosome 1 is shown to illustrate the difference in long-range loop formations in ESCs (Fig. 3a) compared to both lens cells (Fig. 3b-c). Genome wide loop calling shows a near 4-fold difference in chromatin loop size > 3 Mb in ES cells ($n=315$) compared to both lens epithelium ($n=84$) and lens fiber cells ($n=82$). For statistical analysis of genome wide loop and TADs for every chromosome, see Additional File 7: Table S2. Though loop size distributions vary between ES and lens cells, genome wide loop anchor annotations show similar proportions of contacts

mapped to exons, intergenic regions, introns and promoters (Additional File 2: Fig. S3).

Given the most prominent role of DNA-binding transcription factor Pax6 in lens progenitor lens cell formation, separation of the lens vesicle from the surface ectoderm and formation of lens epithelium and lens fibers [72, 73], we used the ~ 500 kb *Pax6* locus as the model to visualize individual chromatin loops. Importantly, human PAX6 locus is located within a larger 7 Mb genomic WAGR (Wilms tumor, aniridia, genitourinary malformation and mental retardation syndromes) region, where large deletions and mutations including *WT1*, *RCN1*, *PAX6*, *PAX6OS1* and *ELP4* loci cause interrelated human diseases affecting kidney, eye, brain and genitals [74, 75]. Here we identified one notable long-range interactions between the mouse Pax6 promoter region with *Wt1* (~ 540 kb, Fig. 4a) and *Meis2* (~ 10.3 Mb, data not shown) within ES cells that was not found in lens cells (Fig. 4b-c). Interestingly, *Meis2* encodes another transcription factor directly regulating Pax6 expression via multiple distal enhancers [76]. These interactions are in agreement with other ES Hi-C data sets [17, 66]. As expected, the *Pax6* locus shows dramatic changes in chromatin reorganization in both lens epithelium (Fig. 4b) and lens fiber cells (Fig. 4c), making contacts with both up- and down-stream proximal and distal enhancers when compared to ES cells. Given the complexity of DNA loops found in these three cell types, expression levels of WAGR genes as well as Pax6os1, Paupar and Wt1os lncRNAs in lens epithelium and fibers are shown in Additional File 2: Fig. S4. Thus, analyses of individual model loci reveal new insights into long-range chromatin loops that can be linked to cell-specific transcription (see below).

Inter/intra-loops, TADs, type of chromatin loops, transcription, and compartment A/B analyses

Next, we examined the relationship between loop anchors and TADs using Juicer tools HiCCUPS and Arrowhead, respectively [77]. We grouped the possible arrangements into four categories: Inter- and intra-loops, loops outside of TAD, and one loop anchor in TAD as schematically shown in Fig. 5a. ES cells showed a higher number of Inter-TAD chromatin loops ($n=560$) compared to both lens epithelium ($n=94$) and fiber cells ($n=240$). Lens epithelial cells show the highest number of loops where anchors were not within a TAD boundary ($n=4,463$) (Fig. 5b). Taken together, these data show major chromatin remodeling in the pathway from ES to lens cells.

Lens cells are characterized by robust expression of genes encoding α -, β - and γ -crystallins [53, 64]. Thus, we next analyzed gene expression profiles of newborn lens epithelium and lens fiber cells for comparative

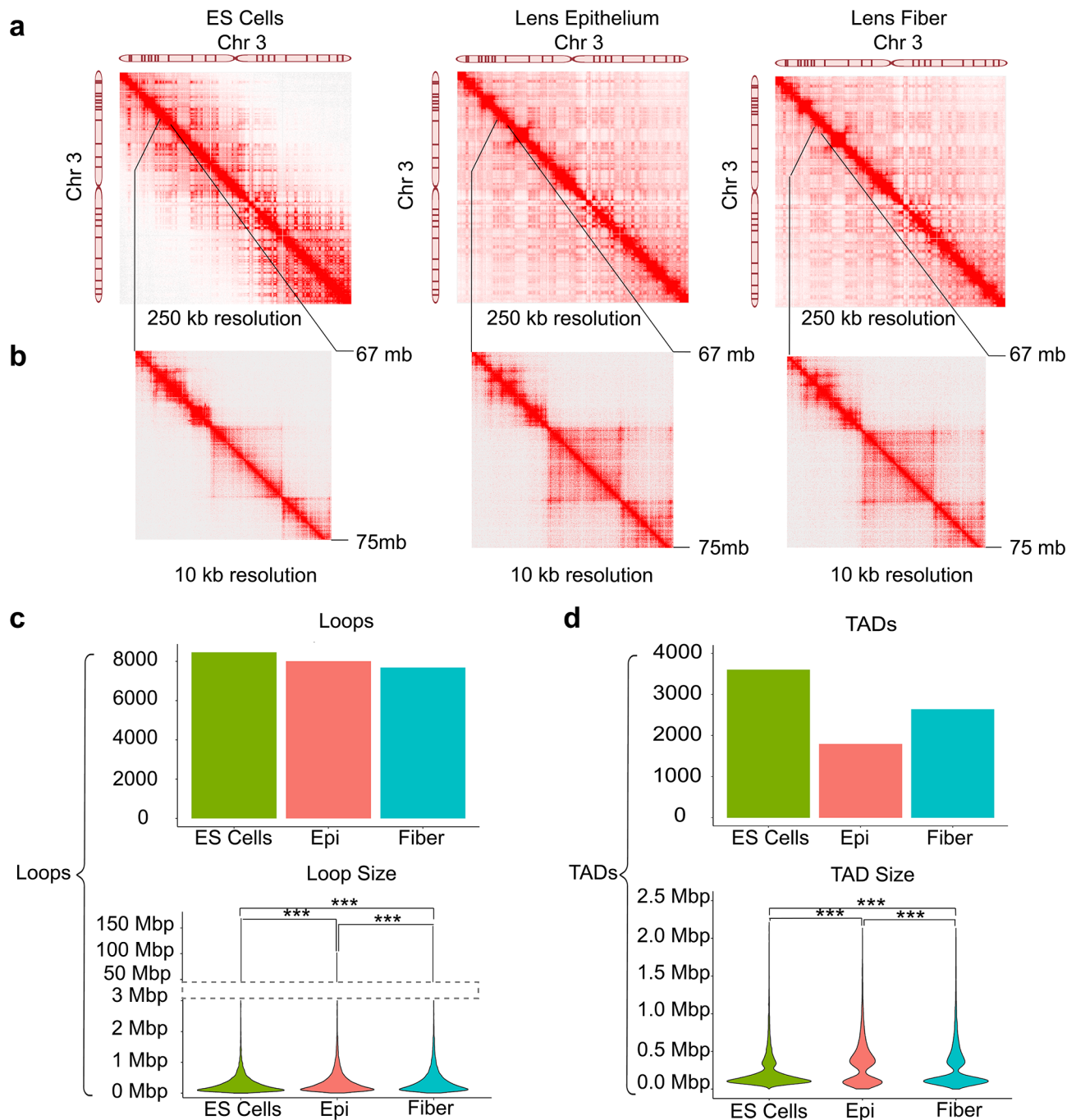


Fig. 2 Mapping of chromatin loops and TADs in ES and lens cells. **(a)** Interaction maps of chromosome 3 generated from ES, lens epithelium and lens fiber cells. Interactions shown are over the span of ~160 Mb and are binned at 250 kb resolution. **(b)** Magnified view of a representative segment of 67–75 Mb on chromosome 3 showing marked differences in chromosome interactions between ESCs and lens cells. Interactions shown are binned at 10 kb resolution. **(c)** Distribution of loop sizes called in ES, lens epithelium and lens fiber cells, respectively. Dashed lines connect scale ranges to show all data points. **(d)** Number of TADs called in ES, lens epithelium and lens fiber cells including their median TAD sizes: 145, 135 and 335 kb, respectively. Significance bars denoted with asterisks: $p \leq 0.05$, $p \leq 0.01$ and $p \leq 0.001$ denoted as *, **, *** respectively

analysis with A/B compartment status. Transcriptome profiling reveals that lens epithelium has enrichment for *Cryaa* and *Cryga* crystallin genes and a battery of small-nuclear non-coding RNAs. Lens fiber cells show higher expression ($\text{Log}_2(\text{FPKM}) > 10.0$) of lens structural genes

including five members of the γ -crystallin gene family (*Cryga*, *Crygc*, *Cryge*, *Crygd* and *Crygb*), as well as members the β -crystallin family (*Cryba1*, *Crybb1* and *Crybb3*) (Fig. 5c). Expression of all genes in lens epithelium and

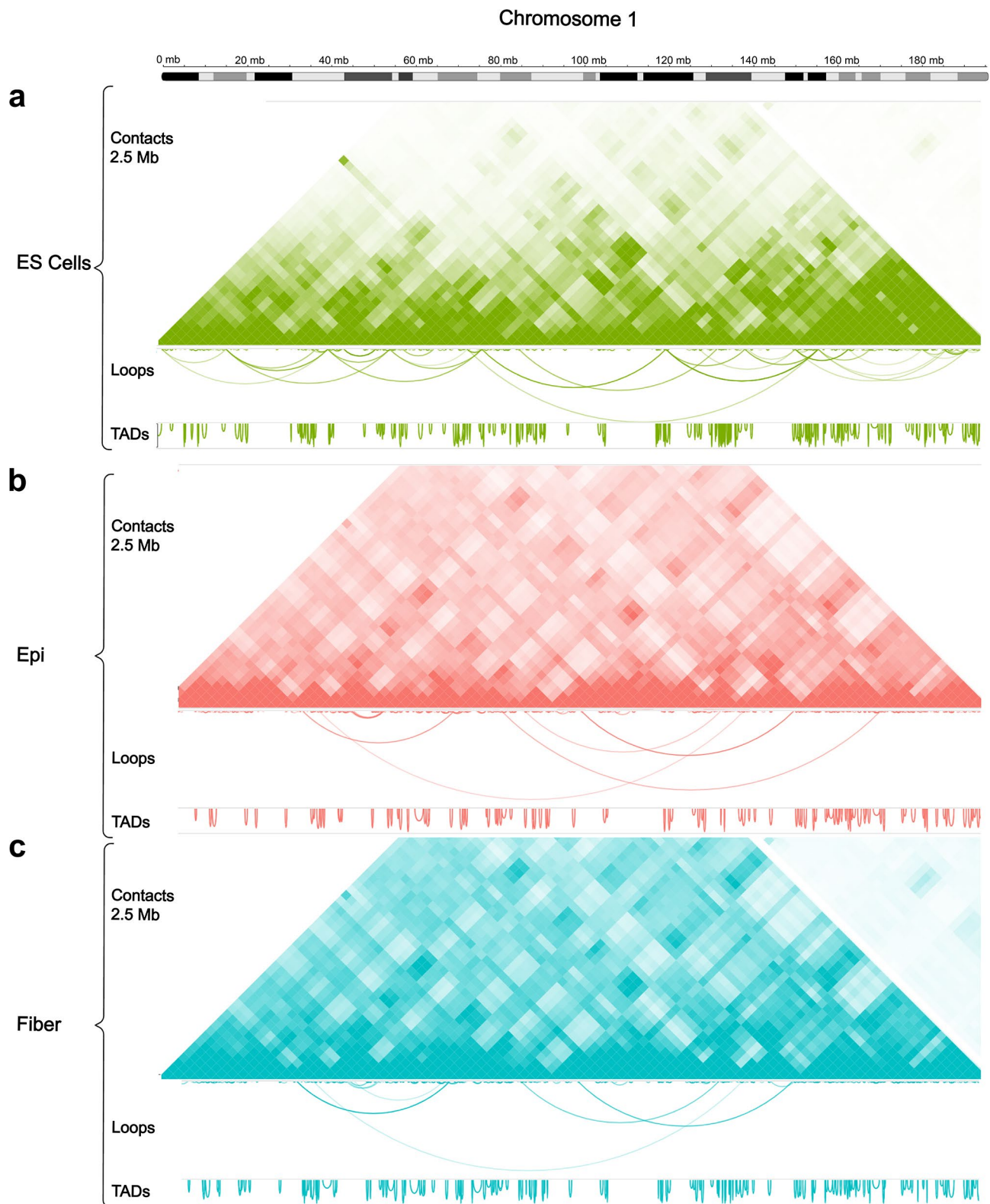


Fig. 3 Differences between the number and distribution of long-range chromatin loops pertinent to entire chromosome 1 in ES and lens cells. Chromatin contact map showing TAD and loop size distribution along chromosome 1 in **(a)** ES cells (572 loops and 232 TADs), **(b)** Lens epithelium (569 loops and 127 TADs), and **(c)** Lens fiber cells (533 loops and 187 TADs). Contact interactions are binned at 2.5 Mb resolution. Genome wide loop calling shows larger number of loops > 3 Mb in ES ($n=315$) compared to both lens epithelium ($n=84$) and lens fiber cells ($n=82$)

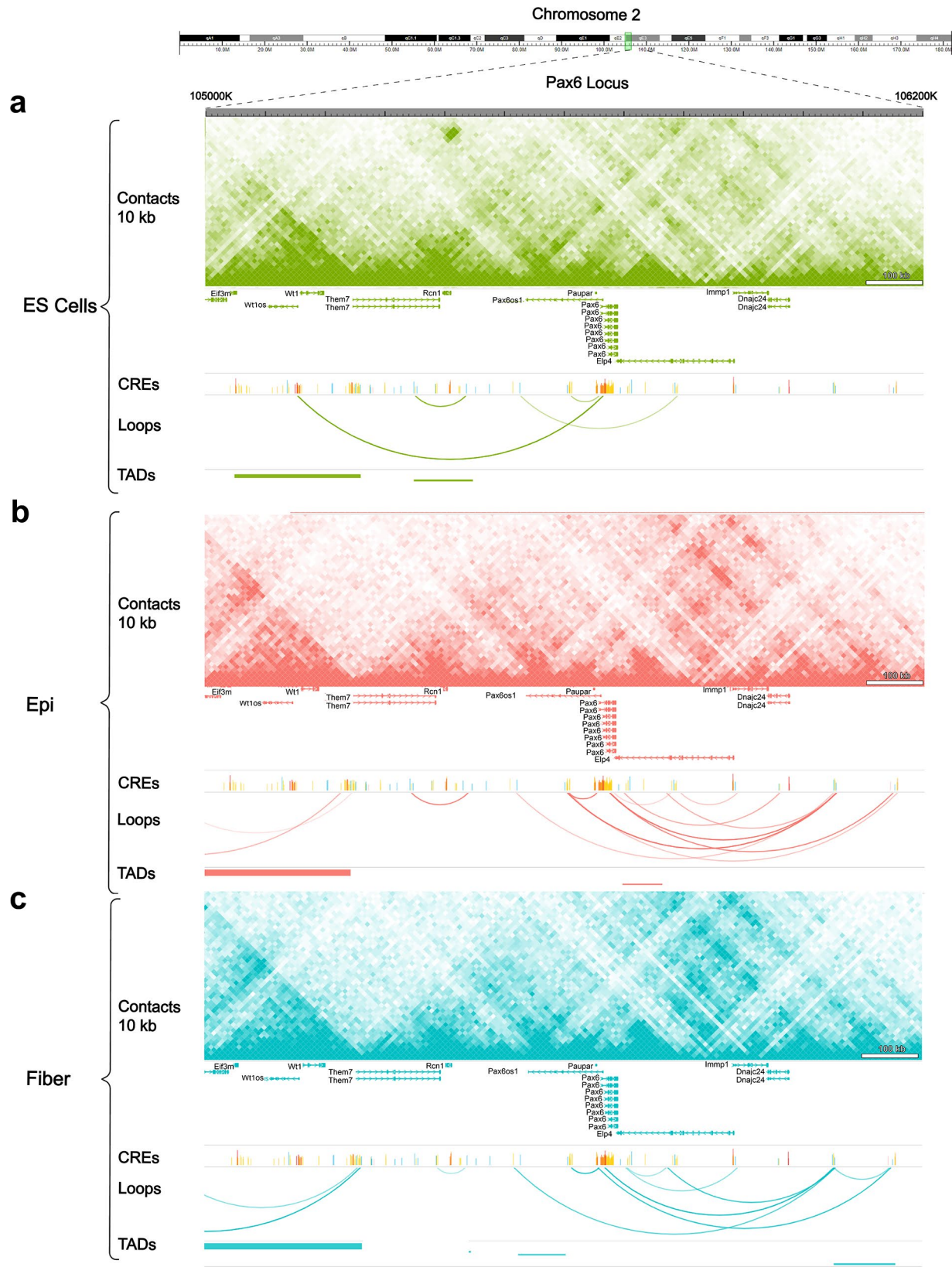


Fig. 4 Identification of chromatin loops and TADs at the representative *Pax6* locus. Loops and TADs at the *Pax6* locus (1.4 Mb DNA region of the chromosome 2) show markedly different chromatin interactions in ES compared to both lens cells. **(a)** The *Pax6* locus within the larger WAGR region in ES cells shows long-range chromatin looping between the *Pax6* and *Wt1* loci while very limited activity within the *Pax6* locus. Chromatin looping patterns within the *Pax6* locus are mostly shared between **(b)** lens epithelium and **(c)** lens fiber cells. Contact interactions are binned at 10 kb resolution

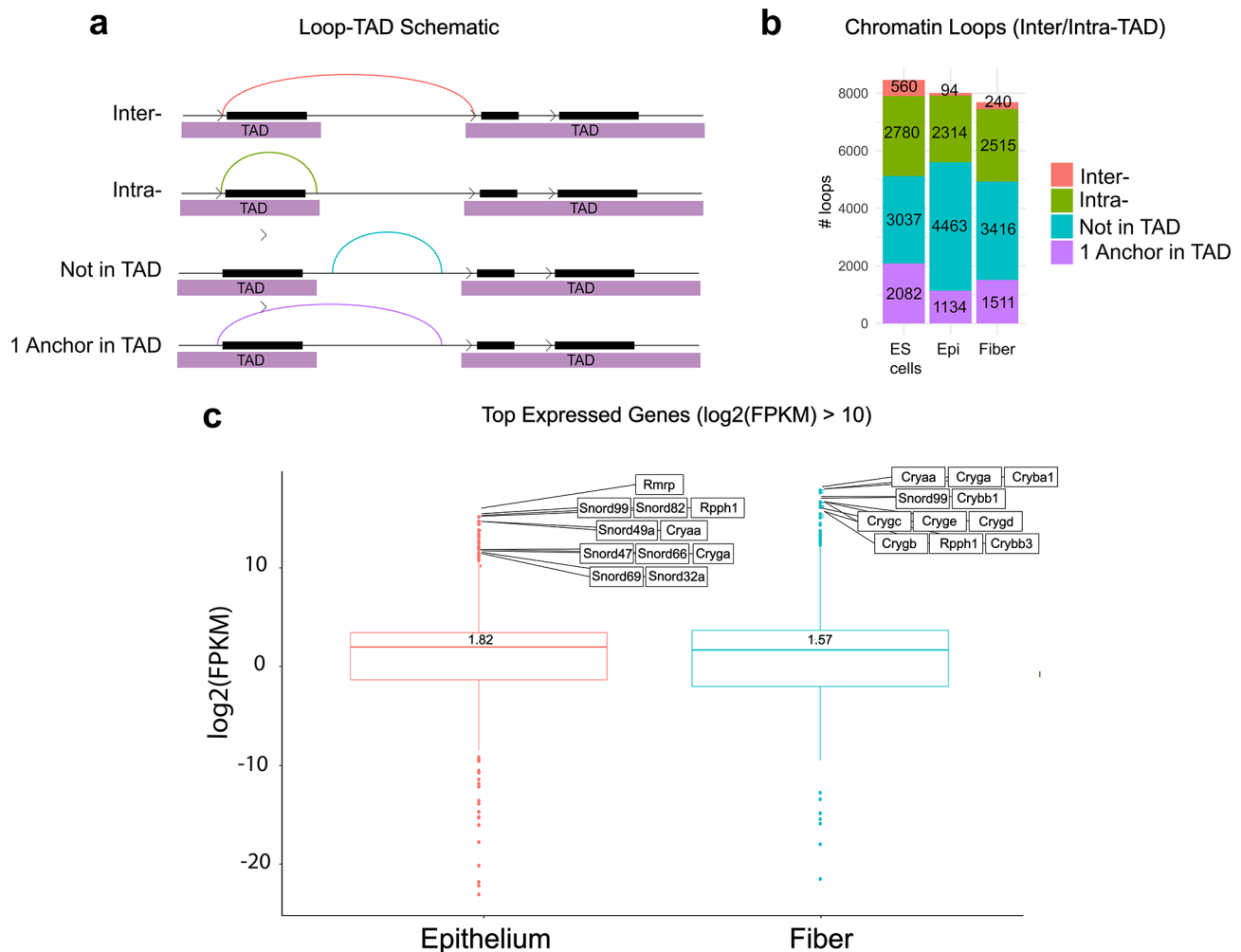


Fig. 5 Proportion of inter-/intra-TAD chromatin loops and distribution of RNA expression of individual genes in lens cells. **(a)** Schematic diagram showing four possible loop-TAD anchor arrangements. Inter- is defined where both loop anchors connect two separate TADs. Intra- is defined where both loop anchors are within a single TAD. Loops with anchors that were found between TADs and the other lies outside of the TAD were defined as “1 anchor in TAD”. **(b)** Quantitative analysis of four types of chromatin loops. ES cells show a higher number of Inter-TAD chromatin loops ($n=560$) than lens epithelium ($n=94$) and fiber cells ($n=240$). Lens epithelial cells show the highest number of loops where anchors are not within a TAD boundary ($n=4,463$). **(c)** Boxplot showing top differentially expressed coding and non-coding transcripts with fragments per kilobase of transcript per million mapped reads (FPKM). The top expressed genes in the lens epithelium mostly include small nuclear RNAs and α A-crystallin (*Cryaa*) gene whereas lens fiber cells show high expression of multiple crystallin genes

lens fiber cells can be found in Additional File 8: Table S3 and Additional File 9: Table S4, respectively.

Additional global analyses of steady-levels of individual RNAs found within compartments A and B are shown in Fig. 6. When cross-referencing transcriptional data with compartment status, we found that both lens epithelium and fiber cells had higher median gene expression levels in compartment A (1.26, 0.854) than in compartment B (0.332, 0.054), respectively (Fig. 6a). Overall, genome wide compartment analysis shows minor differences between lens cell types but marked differences between lens cells and ESCs. The significant differences are denoted by Log₁₀P adjusted values, notably at individual chromosomes 1, 2, 4, 7, 11, 13, 14, 18, and 19 (Fig. 6b).

We also noted locus specific compartmental changes. For representative loci, the *Pax6* locus introduced above exhibits dramatic changes from compartment B to A downstream of the gene body between ES to lens cells. *Pax6* regulates α B- and γ F-crystallin gene expression via synergistic action with RAR β /RXR β heterodimeric transcription factors [78, 79]. The *Rarb* locus (novel cataract risk locus) [80], chromosome 14) also shows compartmental changes within the gene body (Fig. 6c-d). Finally, GO analysis of significant A/B compartment changes (500 unique genes) showed enrichment for terms related to cell morphogenesis, positive regulation of Ras protein signal transduction, homophilic cell adhesion, and

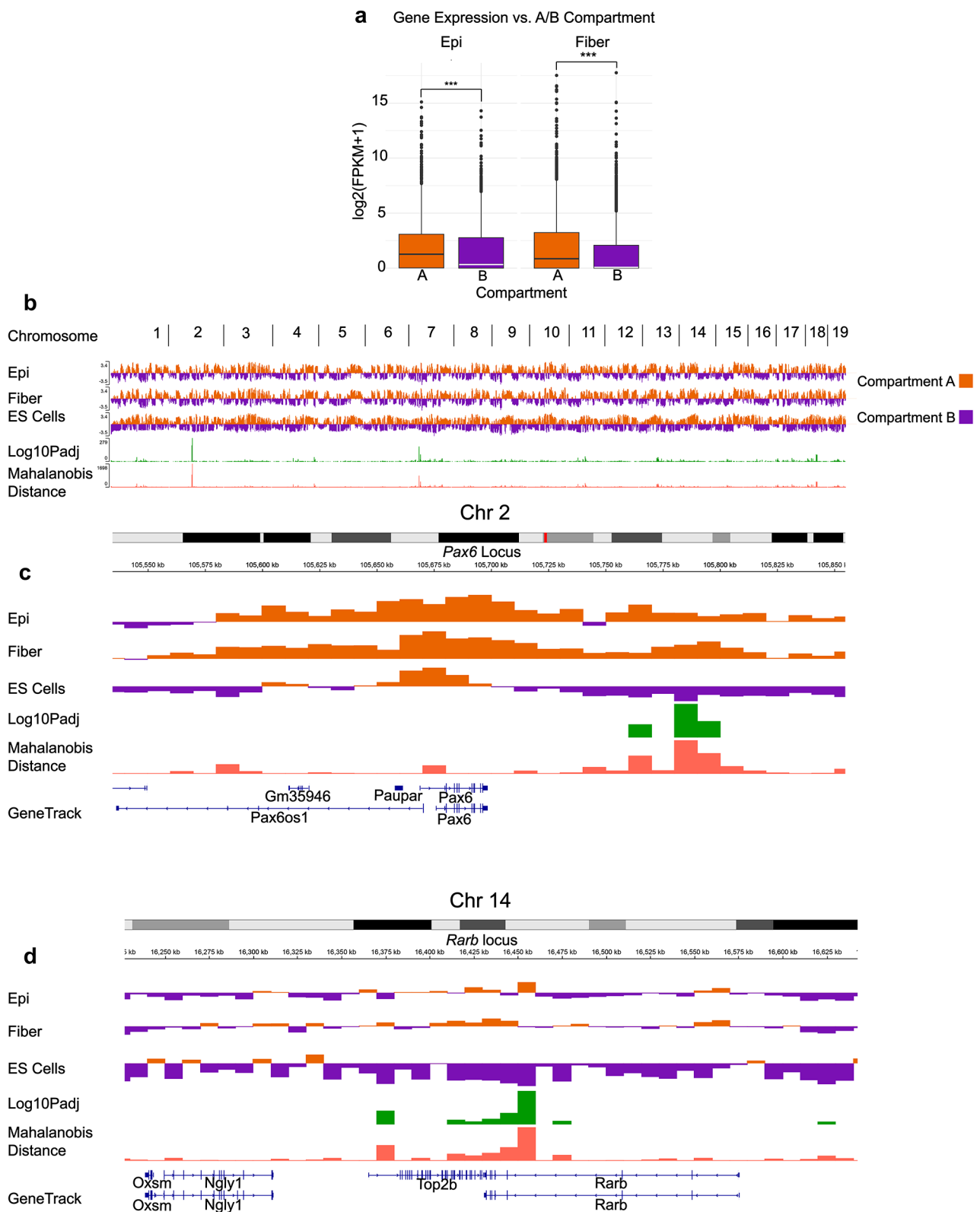


Fig. 6 Comparative analysis of compartments A/B between both lens and ES cells. **(a)** Median gene expression of genes in A/B compartments in of lens epithelial (median: A=1.26, B=0.332) and lens fiber cells (median: A=0.854, B=0.054). **(b)** Genome wide compartment analysis. Two *Pax6* **(c)** and *Rarb* **(d)** loci are used to illustrate regional compartmental changes. Positive Eigen values shown in orange are denoted as "Compartment A", whereas negative Eigen values shown in purple are denoted as "Compartment B". Green peaks are show significant differences in A/B compartmental changes between cell types (Log 10 P adjusted of significant Mahalanobis peaks)

response to leukemia inhibitory factor (Additional File 10: Table S5).

To further understand functional roles of genes involved in chromatin looping, we performed GO analysis on loop anchors found within 1 kb upstream and 100 bp downstream of the transcriptional start site (TSS). Promoter contacts from ES cells show high enrichment for biological processes related to embryonic development. Notable terms (Fig. 7a) include pattern specification process ($p=3.27\times 10^{-21}$), embryonic organ morphogenesis ($p=4.45\times 10^{-18}$), and cell fate

commitment ($p=1.92\times 10^{-17}$). Notable genes with promoter looping include structural development genes *Runx2*, *Tbx1* and *Lhx1* (Additional File 11: Table S6); however, no data exists on their roles in lens. There was also promoter looping in important genes related to cell signaling and patterning, including *Wnt5a* [81, 82] and *Shh* [83, 84] (Additional File 11: Table S6), both genes involved in lens development. In lens epithelium, notable GO enrichments were found for terms related to epithelial development including epithelial tube morphogenesis ($p=8.61\times 10^{-9}$), morphogenesis of a branching

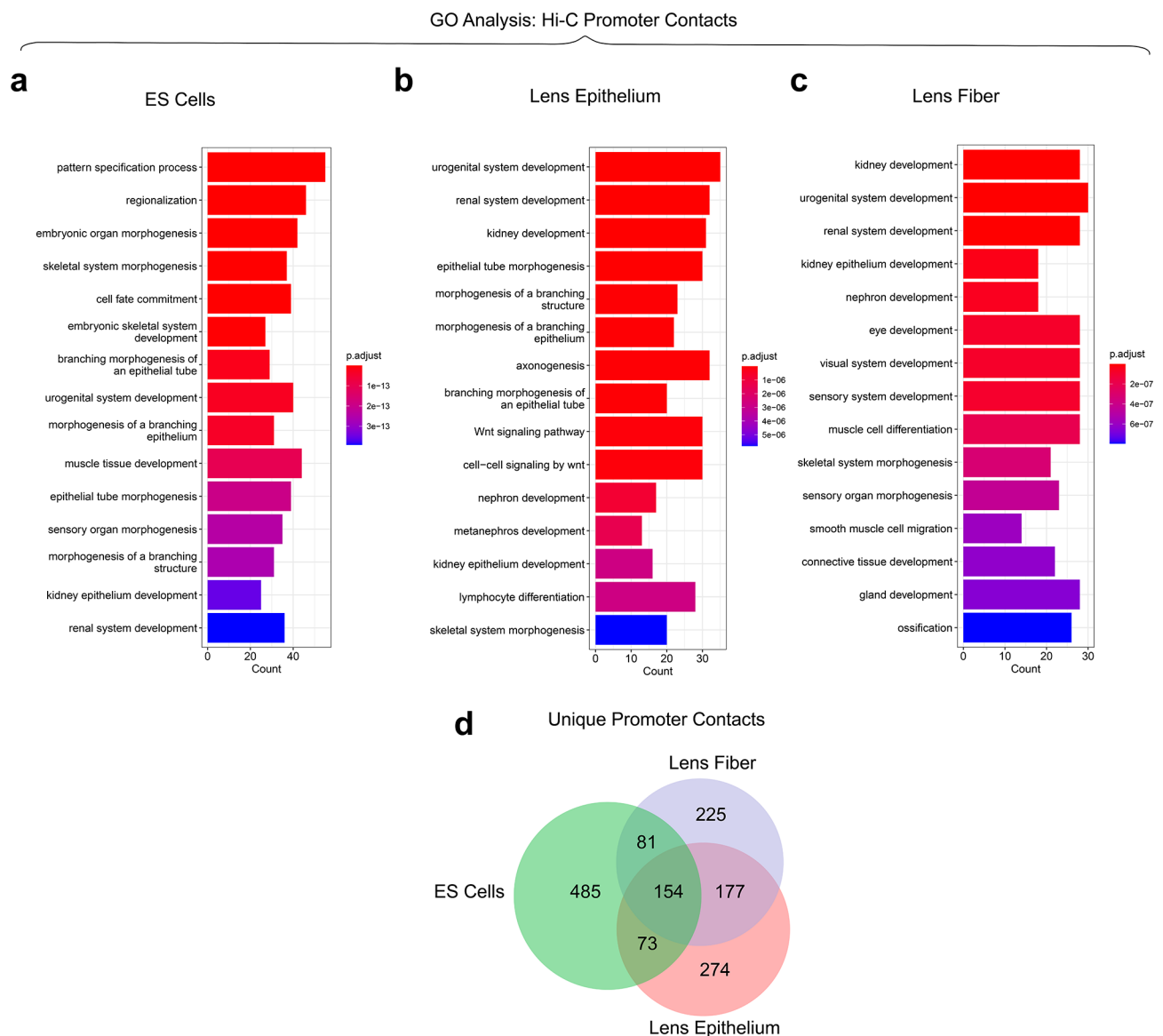


Fig. 7 Gene Ontology (GO) analysis of Hi-C promoter chromatin contacts. GO analysis of chromatin loop contacts at or near promoters (2 kb upstream and 0.1 kb downstream of transcriptional start sites). **(a)** GO analysis of promoter contacts in ES cells. Note enriched terms related to embryonic development and cell fate commitment. **(b)** GO analysis of promoter contacts in lens epithelium. Note enriched terms related to epithelial morphogenesis, differentiation and proliferation. **(c)** GO analysis of promoter contacts in lens fiber cell promoter contacts. Note enriched terms related to development of the visual system. **(d)** Lens epithelium and fiber cells share more promoter contacts ($n=331$) than with ES cells ($n=227$ and $n=235$, respectively). ES cells had the highest number of unique promoter contacts ($n=485$) when compared with lens epithelium ($n=274$) and lens fiber cells ($n=225$)

epithelium ($p=5.21 \times 10^{-11}$), and Wnt based cell-cell signaling ($p=1.98 \times 10^{-10}$) (Fig. 7b). Some prominent genes related to epithelial development with promoter looping include *Bmp4*, *Rara*, and *Pax2* [85–87] (Additional File 11: Table S6). There were also genes related to epithelial cell migration ($p=5.36 \times 10^{-4}$) including *Pxn* and *Hdac7* [88, 89] with promoter looping. Eye development was also significant GO category ($p=4.73 \times 10^{-4}$), listing 21 genes with promoter looping (Additional File 11: Table S6). Finally, lens fiber cell promoter contacts had notable GO enrichment for eye related terms including eye development ($p=8.08 \times 10^{-8}$), visual system development ($p=8.14 \times 10^{-8}$), and sensory system development ($p=8.82 \times 10^{-8}$) (Fig. 7c).

Notable genes encoding transcription factors regulating lens morphogenesis found here with promoter looping include *Pax6* [76], *Meis2* [76] and *Sox1* [90]. Lens fiber cell promoter contacts were also found in *Fgfr2* and *Rara* loci encoding proteins involved in FGF [91–93] and retinoic acid [94] signaling during lens development. Genes encoding anti-apoptotic protein *Bcl2* [95] and chromosome organizational protein *Nipbl* involved in cohesion activities [96] were also noted for their lens fiber-specific promoter looping. Lens epithelium and lens fiber both share a similar number of unique promoter contacts ($n=274$ and $n=225$, respectively), while ES cells had the highest number ($n=485$). Lens epithelium and fiber cells share more promoter contacts ($n=331$) and with ES cells ($n=227$ and $n=235$, respectively) (Fig. 7d).

Lens epithelium and fiber cells show different CTCF-binding associated with distinct CTCF-anchored looping distributions, but common DNA methylation patterns

To further examine chromatin organization in both lens cells, we determined binding of CTCF by ChIP-seq. A total number of individual 16,424 and 23,859 CTCF peaks were found in lens epithelium and lens fiber cells, respectively (Fig. 8a). These numbers are within the similar range as results obtained in mouse retina and brain [97, 98]. Between the two lens cell types, 11,947 (42.2%) of CTCF peaks are shared (Fig. 8a). To compare lens CTCF peaks with ES cells we compared the present data with an earlier study that by Casellas lab [99]. We then aggregated all lens CTCF peaks and found that lens cells have 8,090 (18.4%) unique and 20,221 (45.9%) shared peaks with ES cells (Fig. 8a). DNA *cis*-motif analysis of the individual peaks revealed that both lens epithelium and lens fiber cells shared the same top ranked motifs aligned with CTCF and BORIS aka CTCFL consensus binding sites (Fig. 8b). Note that BORIS is a germline-specific paralogue of CTCF [19, 100].

Earlier studies have shown that DNA methylation regulates binding of CTCF to DNA, impacting 3D genome structure and gene regulation [24, 101, 102]. Next, we

used our recent WGBS data for newborn lenses and ES cells [63] to examine methylation patterns within the individual lens CTCF peaks. Overall, demethylation was observed in both lens epithelium and fiber at loci corresponding to epithelium-specific, fiber-specific, and shared CTCF peaks, though this demethylation was less pronounced at epithelium-specific peaks. Intriguingly, differential CTCF binding between lens epithelium and fiber was not associated with differential demethylation; instead, all DNA methylation patterns were similar between epithelium, fiber, and ES cells at epithelium-specific, fiber-specific, and shared CTCF peaks. This suggests that, while DNA demethylation may be necessary for CTCF-binding, it is not sufficient for predicting cell-specific CTCF (Fig. 8c).

Finally, to understand loop organization in the context of CTCF binding we analyzed proportion of loops containing 0, 1, or 2 anchors bound by CTCF (Fig. 8d). We found that lens fiber cells had nearly a ~2.5-fold higher proportion of loops (17.2%) where both loop anchors were bound by CTCF when compared to lens epithelium (7.15%). In contrast, lens epithelial cells have the highest proportion of loops (57.2%) where either anchor was not bound by CTCF when compared to lens fiber (38.1%).

Subnuclear localization and changes in CTCF in differentiating lens fibers

To evaluate potential changes in CTCF subnuclear localization during lens development we performed immunofluorescence analyses of E14.5 and P0.5 mouse lenses. By E14.5, the lens epithelium and fiber cell compartments are fully formed and primary lens fibers execute their terminal differentiation. Lens epithelial cells located at the lens equator exit the cell cycle and differentiate into the secondary lens fibers [72] (Fig. 9a).

At E14.5 central and peripheral lens epithelium, distinct puncta of CTCF are localized in both nucleoplasm and nucleoli (Fig. 9b). Likewise, similar CTCF staining patterns are found in both differentiating primary and secondary lens fibers (Fig. 9c). At P0.5, primary lens fibers undergo their denucleation and the OFZ is formed (Fig. 9d). Both central and peripheral lens epithelium showed similar patterns of CTCF localization within both nucleoplasm and nucleoli (Fig. 9e). In contrast, the P0.5 fiber cells showed localization of CTCF to the nucleoplasm with much weaker CTCF staining within the nucleoli (Fig. 9f). Taken together, these findings demonstrate changes in CTCF subnuclear localization during lens fiber cell differentiation.

Looping structures, CTCF binding and other chromatin features at individual model loci

To fully harness power of the current data, several representative loci are shown together with additional

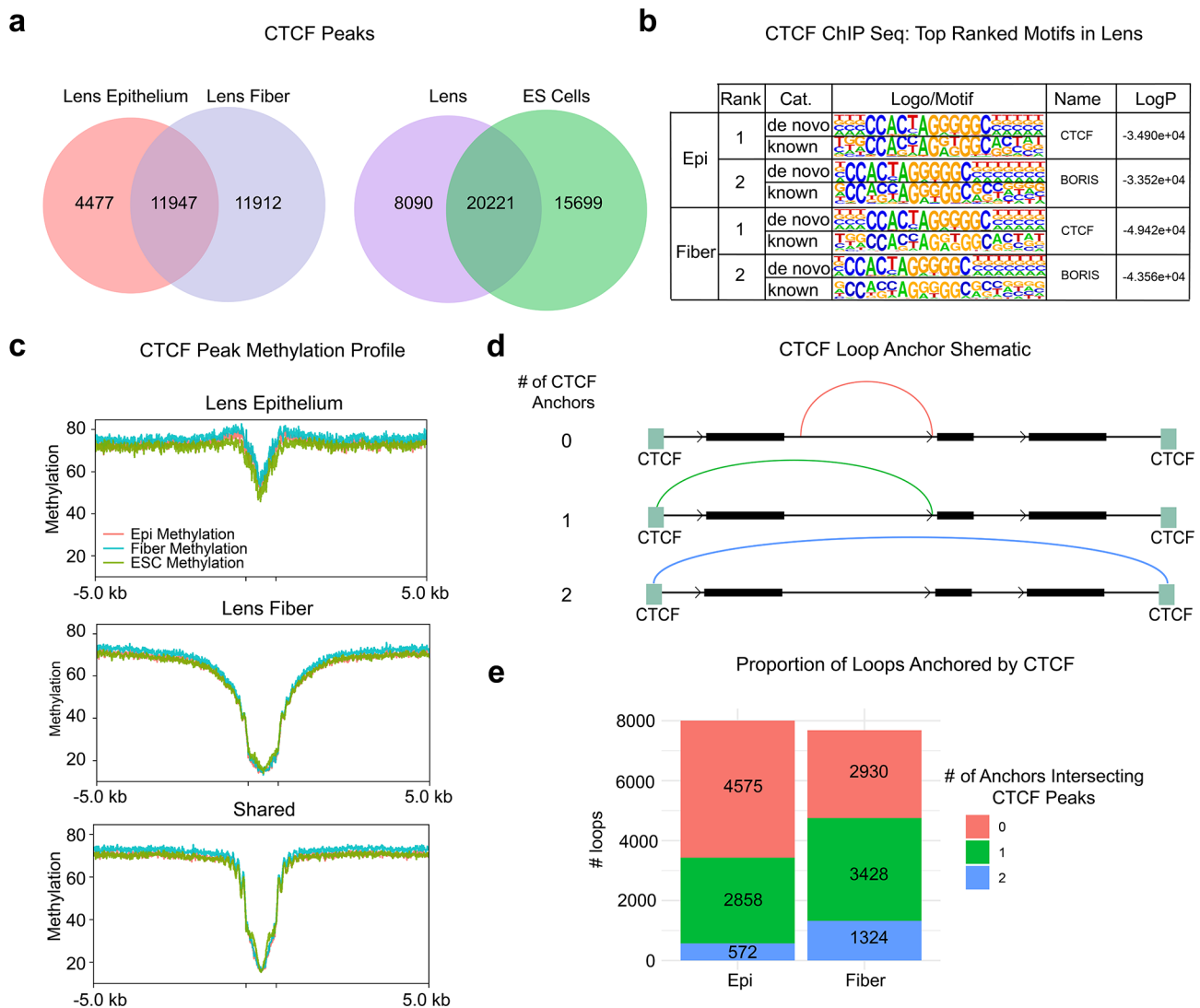


Fig. 8 CTCF binding in lens epithelium and fibers and ES cells, role of DNA methylation and different CTCF-anchored looping distributions. **(a)** Two comparisons between total number of peaks identified in lens epithelium and fiber cells with 11,947 of sheared CTCF peaks. **(b)** Identification of the top enriched motifs with the peaks and their direct comparison with known CTCF binding motifs. **(c)** CTCF peaks and DNA methylation (epi, fiber, and ES cells) in lens epithelium-specific, fiber-specific and shared peaks. Shown are centers of the peaks and their 5 kb flanking regions. **(d)** Proportions of loops anchored by CTCF. Three possible loop-CTCF anchor arrangements are shown. Note gain of total numbers of peaks with one or two anchors in lens fiber cells

tracks, including CTCF binding (ES cells, epithelium and fibers), our earlier “open” chromatin peaks determined by ATAC-seq [62] and ENCODE *cis*-regulatory elements [103]. In addition, histone PTMs and RNA polymerase II tracks are shown using our other data obtained from newborn lenses [53, 65] as well as H3K27ac data in ES cells [104]. Representative loci encoding major lens regulatory and structural proteins are shown in Figs. 10, 11, 12, 13, 14 and 15 and in Additional File 3. As a result of the initial analyses, a set of three contact maps in ES cells, lens epithelium and lens fibers for individual *Pax6*, *Sox1*, *Hif1a* and *Cryaa* loci as well as for the *Cryga-Crygb-Crygc-Crygd-Cryge* and *Crybb2-Crybb3* clustered crystallin loci are shown in Additional File 3: Figure S6-11.

Notable 3D-chromatin differences are revealed by the contact maps between the ES and lens cells as shown earlier (see Figs. 2b and 3).

We first show this set of comprehensive tracks at the ~1.2 Mb *Pax6* locus (Fig. 10, chromosome 2).

Expression of *Pax6* is higher in lens epithelium compared to lens fibers (Additional File 2: Fig. S5). The *Pax6* locus shows a shared complex distal looping network in lens epithelium and fiber cells that is markedly different compared to the ES cells. This loop network spans ~670 kb forming nested loop structures sharing multiple loop anchors bound by CTCF. We found three shared CTCF peaks between ES, lens epithelium, and fibers (boxes 1–3). Lens fibers also show three internal

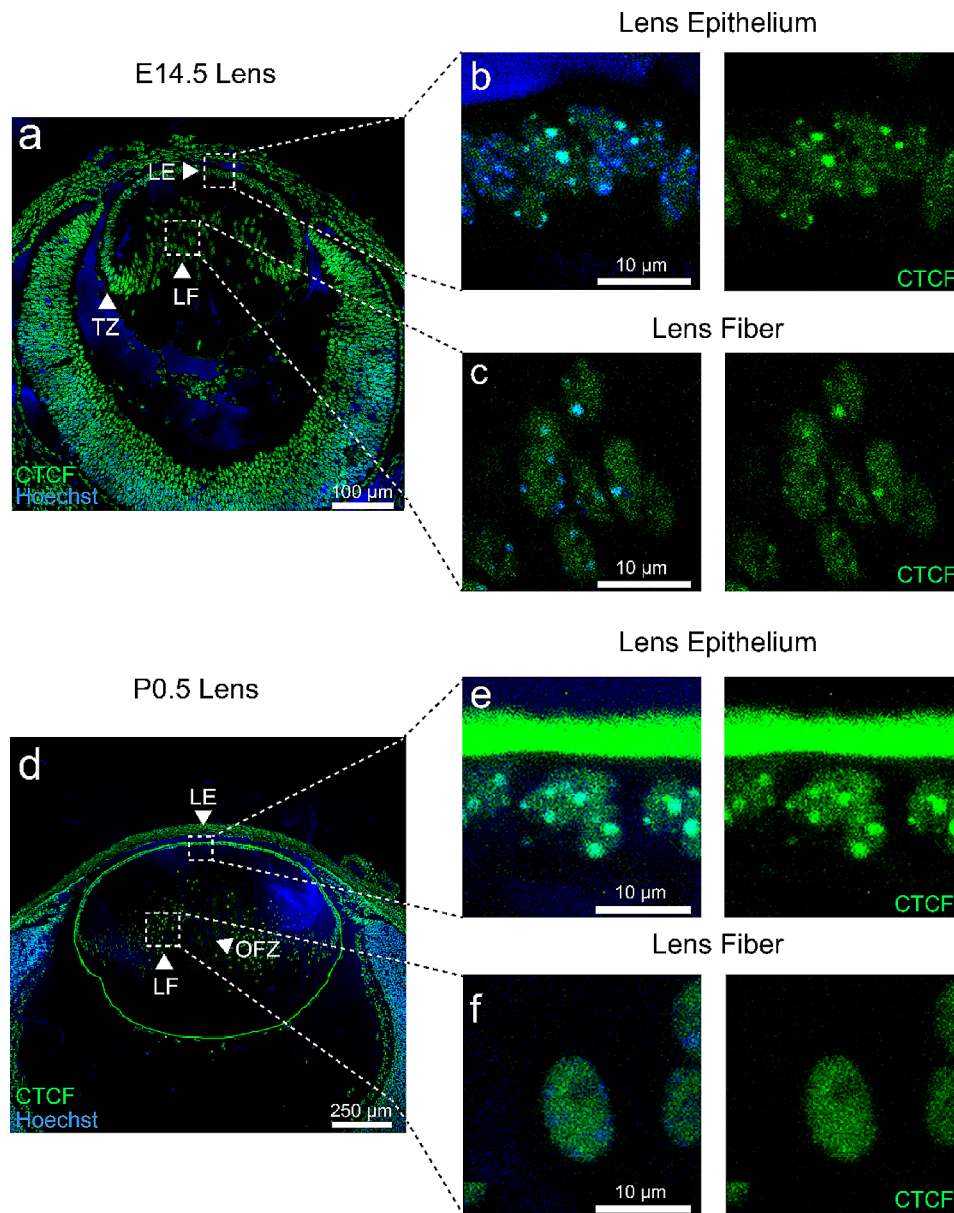


Fig. 9 Immunofluorescence analysis of CTCF in developing mouse lens. **(a)** E14.5 mouse lens labeled for CTCF and nuclear stain Hoechst showing lens epithelium (LE) and primary lens fiber (LF) elongation and secondary LF formation at the transition zone (TZ). **(b)** Magnification of lens epithelium nuclei showing localization of CTCF in both nucleoplasm and nucleoli. **(c)** Early primary lens fiber cells also show localization of CTCF in nucleoli, but more diffusion in the nucleoplasm. **(d)** P0.5 lens showing mature structure of the lens and formation of the organelle-free zone (OFZ) as primary lens fiber cells are already denucleated. **(e)** Lens epithelial cells show strong CTCF signal in nucleoli and staining across the nucleoplasm. **(f)** Primary and secondary fiber cells show translocation of CTCF to the lens nucleoplasm prior their denucleation

CTCF peaks (numbers 4–6) shared with ES cells [66] and peak 4 is shared with rod photoreceptors [58]. Both lens cell types share an upstream distal loop of ~50 kb in length bound by CTCF in lens fiber cell chromatin (loops L1-L2). In whole lens chromatin, a 40 kb domain of the *Pax6* locus shows H3K27ac activity (Fig. 10, horizontal bracket) spanning the gene, overlapping with “open” chromatin (see ATAC-seq tracks), and regions rich with candidate *cis*-regulatory elements (cCREs) that is absent in ES cells. Notably, we found two distinct internal loops

in the 3' region (loops L3 and L4) extending across the *Elp4* locus expressed in the opposite orientation only in lens epithelium and another loop extended even further into the *Dnajc24* gene (Fig. 10).

The *Sox1* locus (Fig. 11, chromosome 8) encodes another DNA-binding transcription factor that is more expressed in lens fibers compared to lens epithelium (Additional File 2: Fig. S5) and directly regulates γ -crystallin gene expression [90]. There is also an overlapping *Sox1* other transcript (*Sox1ot*) long noncoding RNA

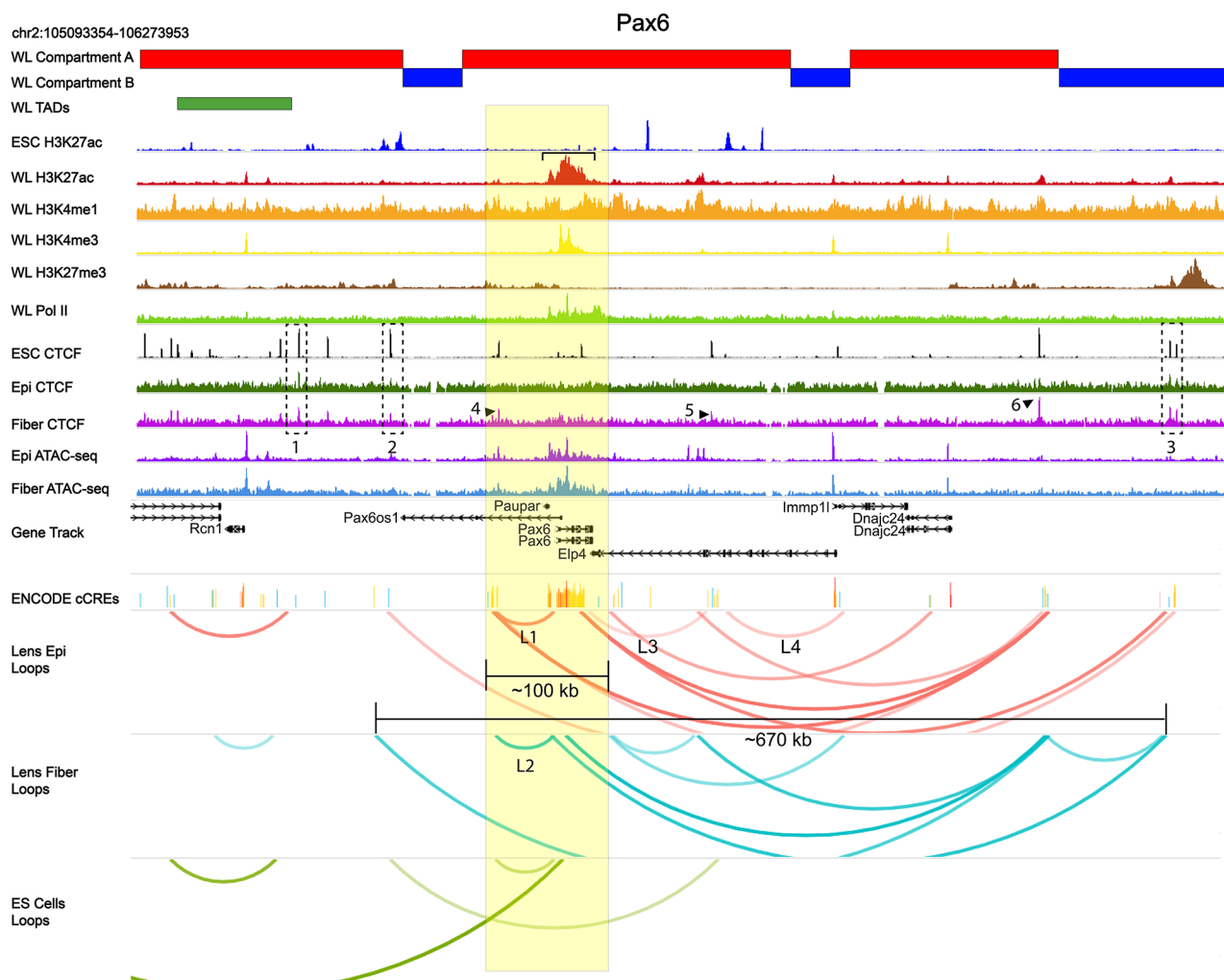


Fig. 10 Chromatin loops, CTCF binding and other features of the *Pax6* locus. Genome browser representation of the loops (shown as arches) found at the *Pax6* locus and its flanking regions in lens epithelium (red), lens fibers (blue) and ES (green) cells. *Pax6* coding regions including both proximal 5'- and 3'-flanking non-coding regions are highlighted in yellow. Compartments A and B are shown by red and blue horizontal bars, respectively. TAD regions identified by "Arrowhead" are shown in green. Whole lens ChIP-seq (H3K27ac, H3K4me1, H3K4me3 and H3K27me3 [59] and RNA polymerase II (Pol II) [47] tracks using whole lens (WL) chromatin are also shown. A broad region of H3K27ac is marked by a horizontal bracket. Lens epithelium (epi) and fiber cell CTCF (present study), ES cell CTCF [66], and ATAC-seq [56] are shown together with ENCODE candidate *cis*-regulatory elements (cCREs): Promoter-like signatures (red), proximal enhancer-like signature (orange), distal enhancer-like signature (yellow), DNase-H3K4me3 (pink) and CTCF (blue). Specific CTCF peaks and loops discussed in the text are marked by dotted boxes (1–3), numbered arrows (4–6) and (L1–L4), respectively. Yellow box is used to indicate locus of interest and scale.

(lncRNA) involved in neuronal differentiation [105]. *Sox1* is encoded by a single ~1.2 kb exon, marked by H3K27ac domain in lens chromatin, and flanked by strong CTCF binding (Fig. 11, arrow 1, also found in ES cells) in lens fibers proximal to the 5'-promoter (-3.9 kb) region where multiple similar loops (L3–L5) are found in all three chromatin samples (Fig. 11).

Lens epithelium and fiber cells mostly share similar loop anchors but have marked differences between ES cells. One notable difference between ES cells and lens cells is the shared loop anchor near the *Sox1* promoter making a distal contact ~50 kb with cCRE enhancer site (loops L1 and L2). In fiber cells, the *Sox1* promoter makes

a loop formation with an unbound CTCF site ~400 kb upstream (loop L3). Five additional CTCF downstream peaks (arrows 1–2, 5–6, dotted boxes 3–4) were found and peak 2 overlaps with loop anchors L5 and L6 in lens fiber and ES cells. A similar sized loop is formed with the *Sox1* promoter near a cCRE region only in lens epithelium (loop L4). Both lens epithelium and fiber share the same CTCF bound distal downstream contact (arrow 7) but have different upstream contacts, one being near the last exon of *Sox1* (cCRE track, arrow 8) in lens epithelium (~130 kb), and the other within an intronic element of *Sox1* (~160 kb) in lens fiber (cCRE track, arrow 9). Overall, the present data show that the *Sox1* gene body

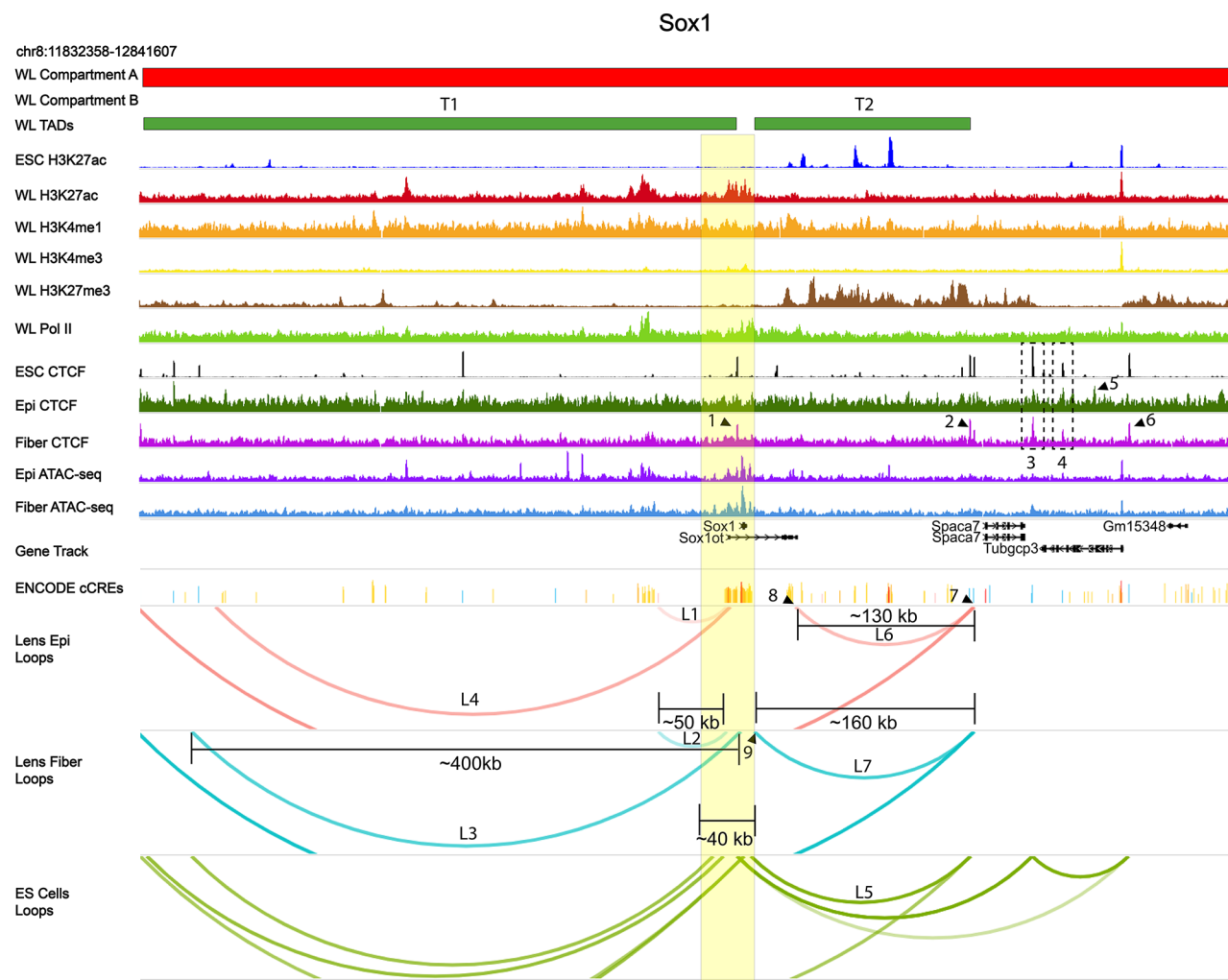


Fig. 11 Chromatin loops, CTCF binding and other features of the *Sox1* locus. *Sox1* coding regions and majority of overlapping *Sox1ot* are highlighted in yellow. The *Sox1* gene body is located between two large upstream 430 kb and downstream 160 kb TADs (T1 and T2), respectively. See Fig. 10 for individual track description. Specific CTCF peaks and loops discussed in the text are marked by numbered arrows and dotted boxes (1–6) and L1–L7, respectively. See Fig. 10 for individual track description.

is located between two large upstream and downstream TADs of 430 kb (T1) and 160 kb (T2) in size, respectively (Fig. 11).

It has been shown earlier that lens fiber cell differentiation occurs at hypoxia conditions [106]. The basic helix-loop-helix transcription factor *Hif1 α* is the main regulator of hypoxia-regulated transcription [107, 108]. Depletion of *Hif1 α* in mouse lens disrupts its growth and lead to lens degeneration [109]. *Hif1 α* locus (Fig. 12, chromosome 12) shows distinct chromatin looping signatures between lens epithelium and lens fiber cells. Four loops found in lens epithelium are shared by lens fibers (loops L1–L4); however, lens fibers display six additional unique chromatin loops, L5–L10. Two notable lens fiber cell loops (loops L6 and L7) make contacts with *Hif1 α* promoter and ~190 kb upstream intronic element (fiber cell loops, arrow 1) of the *Prkch* gene marked by CTCF

binding (dotted box 1) and a unique downstream looping contact with the *Snai1* promoter region (fiber cell loop L9, arrow 2). Both CTCF (dotted boxes 1 and 2) and ATAC-seq data (dotted box 3) do in lens epithelium and fiber show similar peak profiles across the locus. Whole lens H3K4me1 and H3K27ac data (dotted box 4, distinct from ES cells) predict two potential distal 5'-enhancers of the *Hif1 α* locus (loop L1) and overlap with open chromatin regions (dotted box 3). Most loops found here originate from CTCF binding in lens epithelium, fibers and ES cells (Fig. 12, dotted boxes 1 and 4). Thus, loops L6–L9 in lens fibers appear as potential regulatory mechanisms of *Hif1 α* expression in hypoxic (1.5–2% O_2) lens fiber cells [110].

The most highly expressed gene in the lens encodes the α A-crystallin (*Cryaa*, chromosome 17) and is directly regulated by Pax6, c-Maf, CREB, c-Jun and Ets5

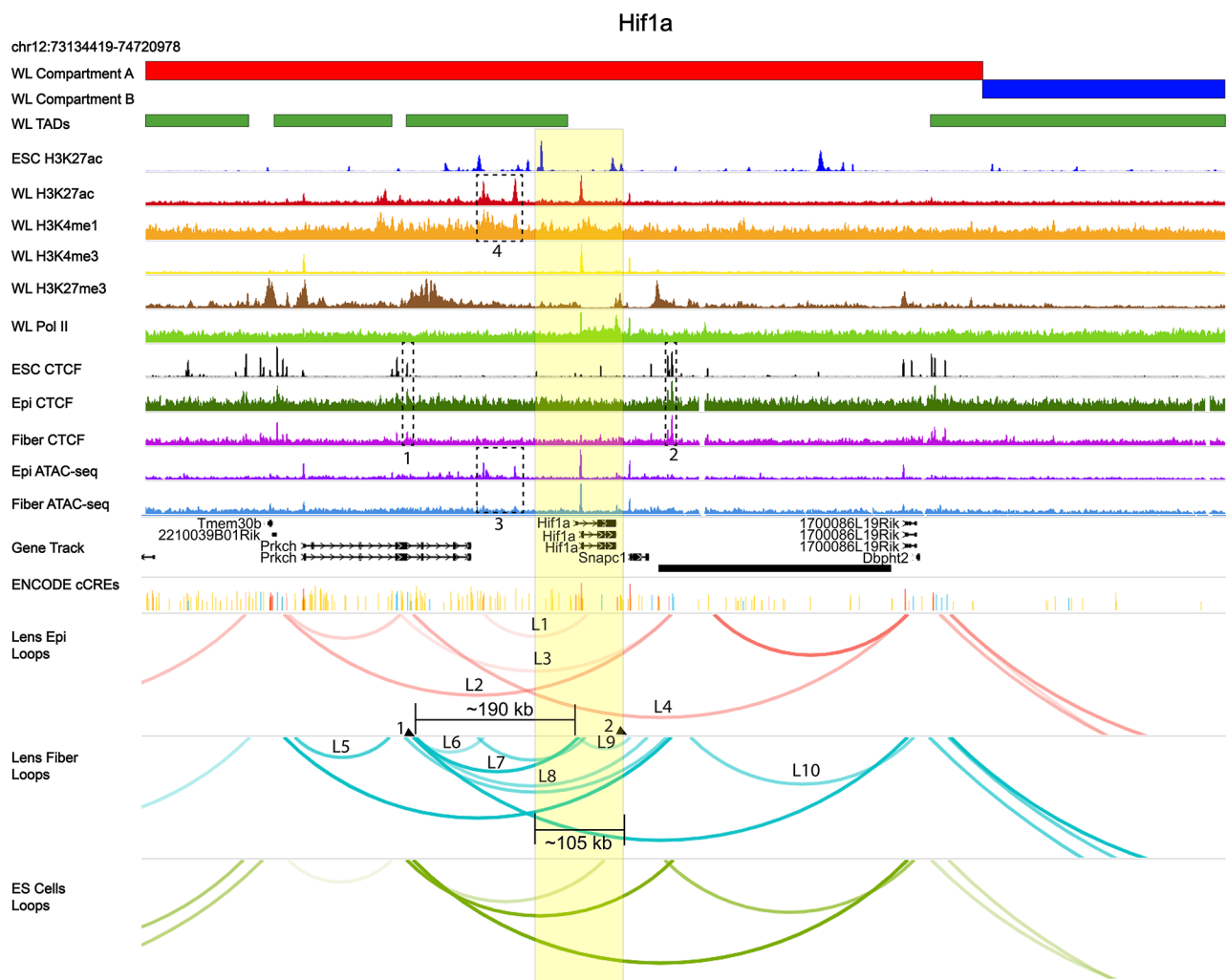


Fig. 12 *Hif1a* locus, CTCF binding and its extensive chromatin looping in lens fiber cells. *Hif1a* locus including its 5'-flanking genomic regions is highlighted in yellow. Lens fiber cell chromatin looping patterns shows marked differences compared to lens epithelium. Two candidate upstream enhancers are boxed in the H3K27ac, H3K4me1 and ATAC-seq tracks. Specific CTCF peaks and loops discussed in the text are marked by numbered arrows and dotted boxes (1–10) and L1–L10, respectively. See Fig. 10 for individual track description

transcription factors [111–113] and marked by abundant RNA polymerase II across over 12 kb of its coding region (Fig. 13). Two 3'-regions of CTCF-binding (dotted boxes 2 and 3) were found in all three chromatins while “proximal” CTCF-binding was only detected at the 5'-region of the *Cryaa* locus in lens epithelium (arrow 1). A striking difference between lens epithelium and fiber cells was increased presence of CTCF found in lens fiber cells across the entire *Cryaa* gene body and overlapping with RNA polymerase II. The internal “peak” of this CTCF domain corresponds to strong binding of CTCF (box 2) in ES cell chromatin.

A cluster of five γ -crystallin genes (*Cryga*, *Crygb*, *Crygc*, *Crygd* and *Cryge*) occupies over 70 kb of chromosome 1 in the absence of any major loops in both lens chromatins analyzed and is marked by CTCF binding at both flanking sides (Fig. 14, dotted boxes 1–2). Both lens cells

and ES cells share distal loop structures (loops L1–L3) of ~60–65 kb in length outside of this cluster of five crystallin genes. Both anchors of this loop structure are bound by CTCF (dotted box 3). In contrast, in ES cells, not expressing crystallin genes, a large distal ~130 kb loop (L4) spans the entire γ -crystallin cluster.

Inspection of chromatin organization of the *Crybb2*–*Crybb3* locus (Fig. 15, chromosome 5) also shows large loops originating just downstream of the 3'-UTR of the *Crybb2* gene that is expressed at much lower level compared to the adjacent *Crybb3* in newborn lens [64] (Additional File 2: Fig. S5).

Three CTCF peaks are found upstream of the *Crybb3* promoter in both lens chromatins as well as in ES cells (Fig. 15, dotted boxes 1–3) where other multiple outside loops originate. Most importantly, a marked overlap between RNA polymerase II and CTCF binding

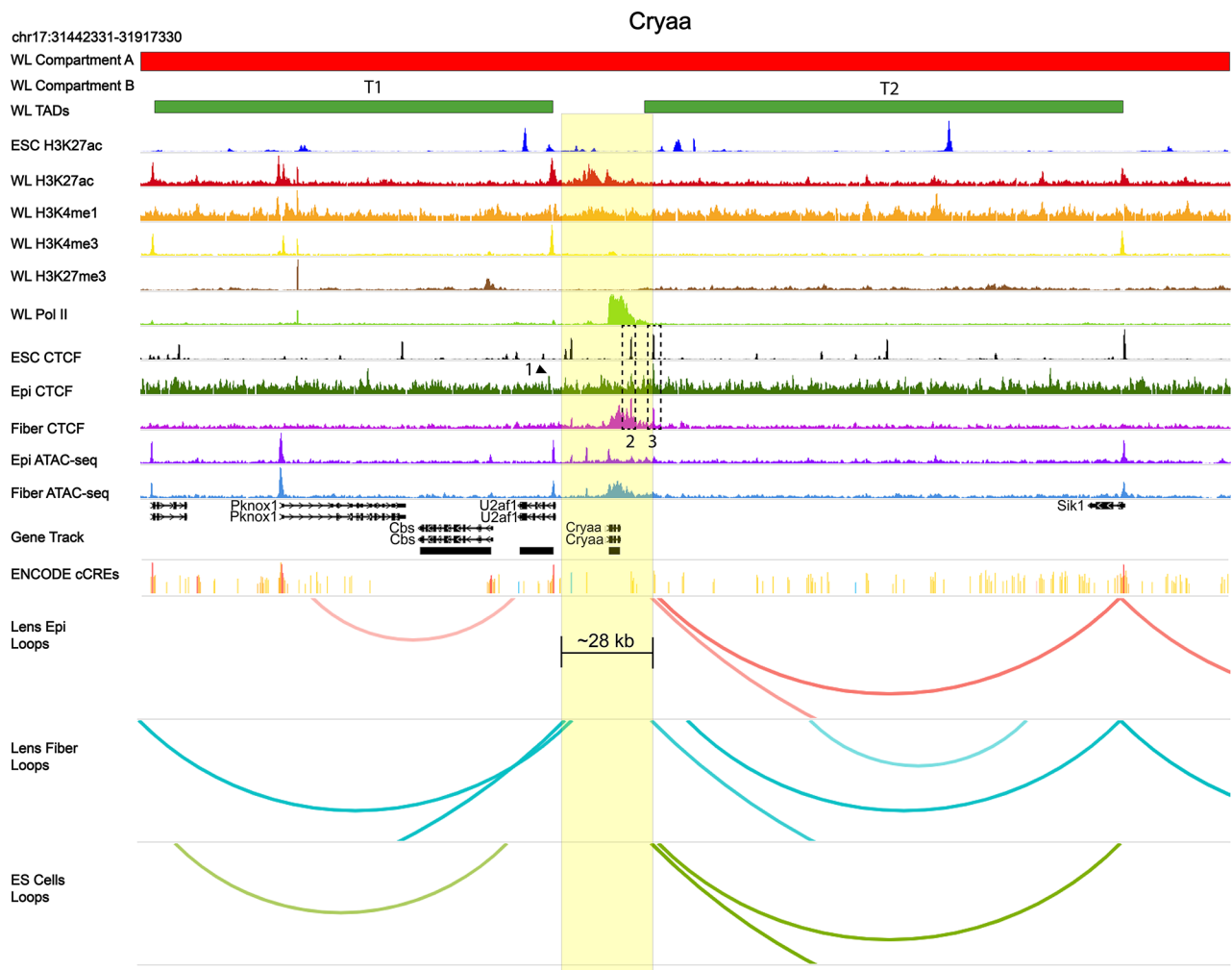


Fig. 13 CTCF-binding, RNA polymerase II and other features of the *Cryaa* locus. The *Cryaa* locus including its 5'- and 3'-flanking genomic regions is highlighted in yellow and resides outside of large looping systems. Note two flanking upstream and downstream ~140 kb (T1) and ~160 kb (T2) TADs, respectively. Specific CTCF peaks discussed in the text are marked by numbered arrows and dotted boxes (1–3). Note an overlap between RNA polymerase II (Pol II) signal across the gene body in lens ChIP-seq and CTCF in fiber cell chromatin. See Fig. 10 for individual track description

across the entire *Crybb3* locus and reduced amounts at the *Crybb2* locus are found (Fig. 15). For contact maps of each cell type, see Additional File 3: Figs. S6–S11.

Three additional groups of individual and/or clustered loci include genes encoding DNA-binding transcription factors regulating lens development (Foxe3, Gata3, Hsf4, Maf, Prox1, Pitx3, Rarb, Sox2 and Tfp2a), proteins involved in lens morphogenesis and differentiation (Bfsp2, Bmp4, Bmp7, Cryba4–Crybb1, Cryba1, Cryba2 and Rb1), and novel cataract genes identified by recent human genome-wide association studies (GWAS), such as *Casz1*, *Gstm2*, *Krtp2-Dpm3-Efna1* and *Sema4d* [80] as shown in Additional file 4, Figs. S12–20; file 5, Figs. S21–27; and file 6, Figs. S28–31; respectively. For example, the centrally located *Gstm2* [80] is a part of larger cluster of seven *Gstm* genes and within a single loop found only in lens fiber cells. The *Rb1* locus encodes the retinoblastoma

protein (pRb) highly expressed in lens fibers and both controlling their cell cycle exit-coupled terminal differentiation [114] and binding to Pax6 proteins [115, 116]. It shows a unique fiber-cell specific long-range loop marked by CTCF binding sites in all three chromatins. The *Sema4d* locus [80] encoding plasma membrane receptor protein semaphorin 4D is also located within a region lacking large loops described above. Taken together, detailed analyses of chromatin looping show unique organization of multiple crystallin loci as well as other loci encoding important lens regulatory and structural proteins. The data suggest an intriguing possibility that clustered CTCF proteins participate in RNA polymerase II convoys and/or formation of the condensates supported by our earlier finding of colocalization of nascent RNA transcription detected using single molecule RNA

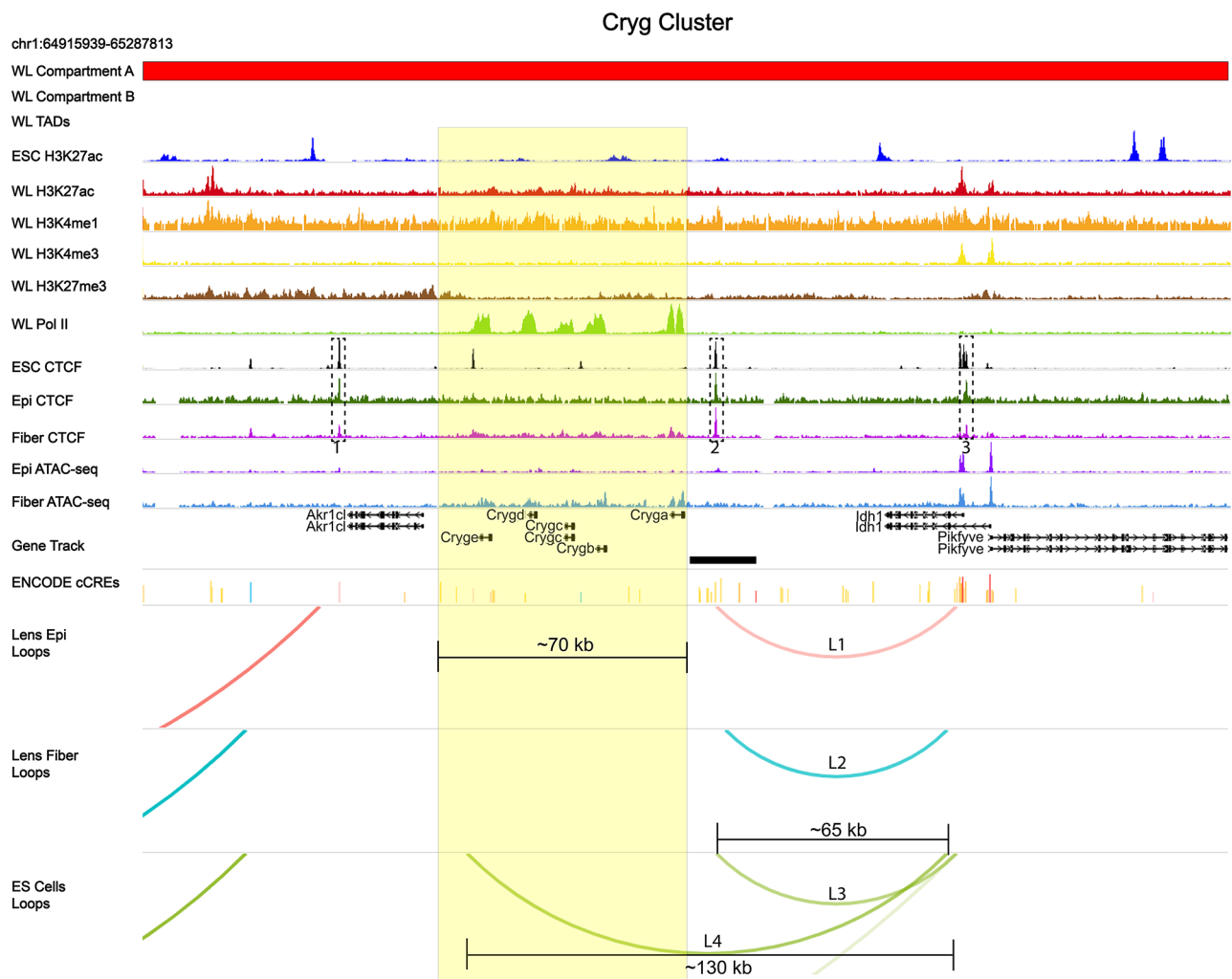


Fig. 14 Limited large loop patterns, CTCF binding and other features of the γ -crystallin cluster. The γ -crystallin cluster occupies over 70 kb of DNA (yellow box) and is marked by high RNA polymerase II (Pol II) consecutive domains at the *Cryga*, *Crygb*, *Crygc*, *Crygd*, and *Cryge* loci in lens chromatin. Three CTCF peaks shared by lens epithelium and fibers are marked dotted boxes (1–3). Note increased CTCF broad binding corresponding to RNA polymerase II presence in lens chromatin. No proximal large loops are found through the entire γ -crystallin cluster in both lens chromatin. See Fig. 10 for individual track description

FISH and immunofluorescent visualization of transcriptionally active RNA polymerase II [51].

Discussion

The present Hi-C data show for the first-time 3D-chromatin organization in mouse newborn lens chromatin separated into lens epithelial and lens fiber cells and direct comparisons with chromatin organization of ES cells. In addition, these data are analyzed in the context of CTCF binding determined by ChIP-seq as well as our earlier studies of chromatin landscape by ATAC-seq and ChIP-seq studies of histone PTMs and RNA polymerase II. These lens data can serve for comparative purposes with other mouse tissues and provide potential insights into distal non-coding variants associated with abnormal

lens development, cataracts, and other diseases involving ocular lenses.

Hi-C experiments have already been conducted using multiple mouse and human cells and tissues of different complexity. Multiple similarities exist between erythroid maturation and differentiating lens fiber cells. Mammalian erythrocytes are marked by high levels of α - and β -globin gene expression that are directly comparable to crystallin gene expression at the quantitative levels (comparing bulk RNA-seq data) with lens fiber cells [53]. In addition, chromatin condensation and transfer of nuclear proteins into the cytoplasm occurs in maturing red blood cells followed by nuclear extrusion [117, 118]. In contrast, lens fiber cell nuclei disintegrate within the individual lens fiber cells as described above (see Fig. 9). Our data using RNA FISH to detect nascent RNA expression

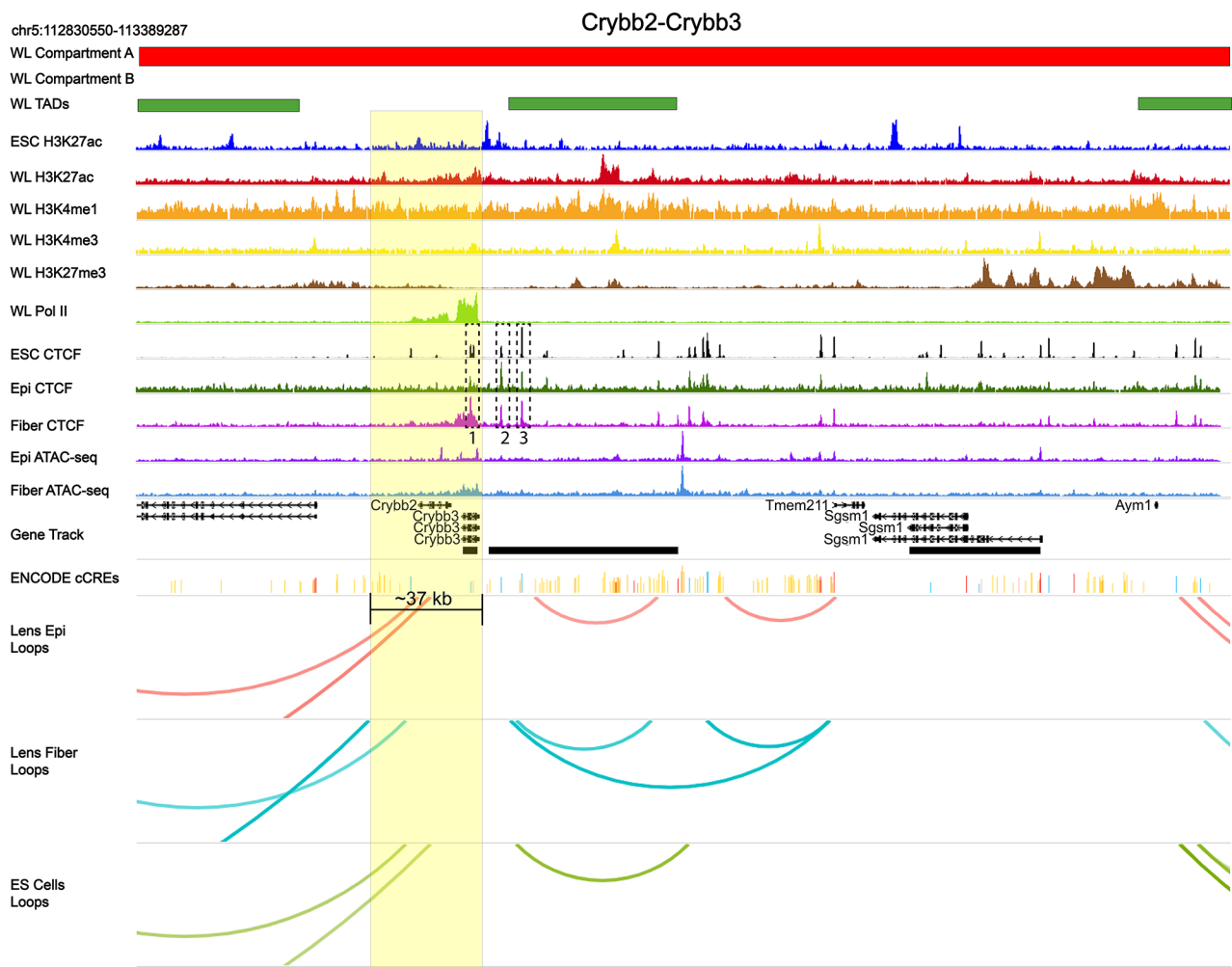


Fig. 15 Large loop patterns, CTCF binding, RNA polymerase II, and other features of the *Crybb2-Crybb3* locus. Both gene bodies are located between ~70 kb upstream and ~110 kb long downstream TADs. Similar to the γ -crystallin cluster, no proximal large loop interactions were detected within these loci. Consistent with RNA-seq data [58], the *Crybb3* shows higher Pol II signal compared to the adjacent *Crybb2* locus. Multiple CTCF peaks are found upstream of the *Crybb3* locus (boxes 1–3). A large domain of CTCF overlaps with Pol II at the *Crybb3* gene. See Fig. 10 for individual track description

show that *Cryba4* and *Crybb1* genes can be simultaneously transcribed from adjacent alleles [119]. It has been shown earlier that low insulation scores have been associated with upregulation of genes related to terminal differentiation [10, 66].

The unexpected presence of CTCF and overlap with RNA polymerase II signals in *Cryaa* and *Crybb3-Crybb2*, *Cryba4-Crybb1* and *Cryga-Cryge* gene clusters represents potentially a new research avenue to understand composition of condensates formed at late stages of lens fiber cell differentiation, just prior their denucleation, originally visualized through co-localization studies of nascent crystallin mRNA transcription and transcriptionally active RNA polymerase II [51]. To assess the association of RNA Polymerase II and CTCF signals using an unbiased approach, we analyzed genome-wide both ChIP-seq signals. We specifically evaluated the

mean Pol II and CTCF signals across all gene bodies, focusing on protein-coding genes, and comparing them with gene expression data (see Materials and Methods for details). Our genome-wide analysis shows that the crystallin family of genes indeed both had the highest RNA Polymerase II-CTCF signal overlap and expression levels (Additional File 12: Fig. S32a). Outside of the crystallin genes, we found other highly expressed lens genes with high RNA Polymerase II-CTCF signal overlaps. Some of these genes include Mip/Aqaporin 0, Gja8, and Vim, all genes are involved in lens fiber cell differentiation and implicated in cataract formation (Additional File 12: Fig. S32b). The most highly expressed α A-crystallin locus [53, 64] show highly specific CTCF binding in lens fiber cells that overlaps with transcriptionally active RNA polymerase II detected in whole lens chromatin (Fig. 13). Similar patterns are also found at both *Crybb2-Crybb3*

and *Cryba4-Crybb1* clusters and individual *Cryba1* and *Cryba2* regions (Additional File 5 Figs. S24 and S25). This trend is even detectable at the γ -crystallin gene cluster (Fig. 14). Recent studies have shown presence of RNA polymerase II and CTCF via their intrinsically disordered domains in formation of phase-separated droplets [120]. It is thus possible that the presence of CTCF in these highly transcribed regions is indirect and related to the protein-protein interactions within the phase separated structures [121]. Our earlier studies of nascent expression of *Cryaa*, *Cryba4*, *Crybb1*, *Crybb3* and *Cryga* genes evaluated by single molecule RNA FISH showed correlation with the largest foci containing transcriptionally active RNA polymerases II [51, 119]. Thus, these findings may relate to a broad multifunctionality of CTCF outside of organizing TADs and forming clusters of 2–8 CTCF molecules [32] that may even at much larger quantities assist large RNA polymerase II convoys transcribing crystallin loci as found in lens chromatin [46, 53]. Another possibility is that the CTCF proteins, through their RNA binding zinc-finger domains [22], can also bind nascent crystallin mRNAs. This possibility is not mutually exclusive with the other mechanisms described above. Finally, recent studies using cancer cell line have shown presence of CTCF in phase-separated condensates and their integrity requires presence of CTCF suggesting instructive function of these proteins for condensate formation [122]. Thus, future studies of these phenomena in lens fiber cell chromatin are highly warranted.

In addition to the representative loci discussed above (Figs. 10, 11, 12, 13, 14 and 15), three groups of loci are included for comparative purposes. For example, *Sox2* (Additional file 4, Fig. S19) regulates pluripotency of ES cells [123] as well as early stages of lens placode morphogenesis [124, 125]. The large *Sox2-Sox2ot* locus includes multiple proximal and distal enhancers our recent studies demonstrated different DNA methylation patterns between ES and both types of lens cells [63]. *FoxE3* (Additional File 4, Fig. S13) is another DNA-binding transcription factor regulating early lens morphogenesis [126]. The second group includes *Cryba4-Crybb1*, *Cryba1* and *Cryba2* loci (Additional File 5, Figs. S24–26), *Bfsp2*, encoding lens-specific intermediate beaded filament protein [127], also shows co-localization of RNA polymerase II with CTCF region in lens fibers (Additional File 5, Fig. 21). Finally, four novel cataract loci, including *Casz1*, *Gstm2*, *Krt2-Dpm3-Efna1* and *Sema4d*, identified by recent human GWAS studies [80], are also shown in Additional file 6, Figs. S28–31, that might help in the interpretation of non-coding variants if found in evolutionarily conserved distal regions marked by looping anchors. Indeed, GWAS found that most variants in a wide spectrum of human diseases are located outside of protein-coding regions [128]. Identification

of truly causal GWAS variants is challenging given their repertoire and distal enhancers representing the most challenging tasks [128]. Thus, the present Hi-C studies will not only aid in interpretation of GWAS studies of cataract genes [80] and of the WAGR syndrome [74, 75] but also aid studies of the microphthalmia-anophthalmia-coloboma (MAC) syndrome caused by mutations in genes with multiple tissue-specific distal enhancers such as *PAX6*, *SOX2*, *ATOH7*, *OTX2*, *VSX2*, *FOXE3*, *BMP4*, *MAB21L1* and other loci [129–131] due to high evolutionary conservation of transcriptional control of these genes between human and mouse. For example, Hi-C maps of promoter-enhancer interactions in neural tissues [132], multiple sclerosis [133] and age-related macular degeneration [59] have already been used to analyze the GWAS data.

Multiple roles of CTCF in DNA- and RNA-binding can be also inferred from different subnuclear localization related to lens differentiation. We found marked difference of CTCF nuclear localization in lens epithelium and fiber cells that is both temporally and spatially regulated (Fig. 9). Outside of being an insulator protein regulating genome organization, previous studies have shown CTCF to be involved in epigenetic control of rDNA and enhancement of rRNA transcription catalyzed by RNA polymerase I within the nucleoli [134, 135]. Other studies also show CTCF regulates myeloid and erythroid differentiation in human cell lines [136, 137]. Taken together, high abundance of CTCF within nucleoli in differentiating lens epithelium and fiber cells may be a cellular mechanism to augment ribosomal biogenesis to meet high demands for translational output of crystallin proteins in maturing lens fiber cells [40].

Lens chromatin landscape is regulated by various chromatin remodeling complexes as show by lens-specific depletions of Brg1 (Smarca4) [138], Snf2h (Smarca5) [49], CBP and p300 [139], Nco6a [140] and Znhit1 [141] proteins in mouse models. We have shown localization of Brg1, Snf2h, p300 and CBP at the 16 kb *Cryaa* locus using qChIPs in mouse lens chromatin [112, 142] and that Pax6 forms complexes with BAF complexes in neurons [143], retina [144] and lens [112, 145]. Recent studies have shown co-localization of Snf2h/Smarca5 with CTCF in human cell lines [146]. Thus, Snf2h proteins found at the *Cryaa* locus [112] might be also involved in our findings of CTCF across the *Cryaa* locus in lens fiber cell chromatin. This another mechanism is not mutually exclusive with those described above.

Future studies will be aimed to pursue parallel opportunities to employ living cells to study cohesin and condensin molecular machines governing ATP-dependent loop extrusion [28], use of single cells combined with Hi-C [147], detailed analysis of the *Pax6* locus using 4 C-seq [148], mapping of enhancer RNAs via PRO-seq

[149], deletion of candidate enhancers in the *Pax6*, *Prox1*, and *Hif1a* loci together with transgenic reporter assays, and conditional inactivation of CTCF [150] using lens-specific Cre lines acting in more advanced stages of lens differentiation [151]. Interesting loci can be selected from the present studies and subjected to both 3 C studies to map looping patterns with much improved resolution and directly visualize these interactions using DNA FISH. Finally, chromatin structural modeling (HiP-HoP) based on a combinatorial use of ATAC-seq, H3K27ac, and CTCF data [13] just requires mapping of the H3K27ac landscape in microdissected P0.5 lenses.

In conclusion, the present study has expanded earlier transcriptomics and epigenomics data on mouse lens embryonic development and differentiation that now includes chromatin loops and paves the road for similar studies using human lens cells. Future studies will probe chromatin condensation within lens fiber cells undergoing early stages of their denucleation while preserving the maximal transcriptional output of β - and γ -crystallin genes just prior their abrupt disintegration. Super-resolution microscopy now allows quantification of transcriptionally active RNA polymerase II, CTCF and other proteins in parallel with direct visualization of nascent crystallin gene expression detected by single molecule RNA FISH within individual nuclei of differentiating lens fiber cells [51, 119].

Materials and methods

ES cells and lens tissues

To study chromatin interactions of differentiating lens cells, newborn (P0.5) CD-1 mice and mouse ES cells were used. Two biological replicates of lens epithelium, lens fiber and ES cells were used for statistical power. Lenses were dissected at P0.5 and then micro-dissected into lens epithelium and lens fiber under a dissection microscope. Each replicate of lens tissue was comprised of 30 lens epithelium and fiber samples. Samples were kept on dry ice for the duration of the dissection process. Samples were then homogenized using a disposable pestle tissue grinder. Homogenized tissues were fixed in 2.0% formaldehyde for 10 min at room temperature. Formaldehyde was quenched with 0.125 M glycine solution. Mouse ES cells v6.5 (mixed 129/B6, male) were provided by Dr. Meelad Dawlaty [13]. Cells were grown under feeder-free conditions on 0.2% gelatin and supplemented with LIF (24 ng/mL). Each replicate contained $\sim 2.0 \times 10^6$ cells and were harvested near $\sim 80\%$ confluency. The crosslinking of ES cells was performed as described above.

Generation of Hi-C library and sequencing

The Hi-C library was generated using the Arima-HiC kit according to the manufacturers protocols (A510008) and performed by the NYU Langone Health Genome

Technology Center (New York, NY). DNA libraries were sequenced on the Illumina NovaSeq 6000 with $\sim 700 \times 10^6$ reads per sample with mean quality score $Q > 36$.

Quantitative Hi-C analyses and statistics

Read alignment and computation of Hi-C contact maps were performed using the ENCODE Hi-C pipeline (code available at <https://github.com/ENCODE-DCC/hic-pipeline>). Alignment was performed within the pipeline using bwa-mem (Li 2013 <https://arxiv.org/abs/1303.3997>). Contact map computation was performed with maximum resolution=5 kb within the pipeline using Juicer [77, 152]; contact maps created from filtered reads aggregated across two biological replicates of two lanes each per cell type with alignment quality score $MAPQ \geq 30$ were used for subsequent per-cell type analyses. For whole-lens contact maps, reads from all replicates from both epithelium and fiber were pooled. Loop and TAD calling were done with Juicer Tools using HiCCUPS and Arrowhead, respectively [77]. Default parameters were used for all pipeline analysis stages. Whole-lens contact maps and subsequent analyses were created from combined lens epithelium and fiber contact maps.

For A/B compartment analysis, dcHiC [68] was used to calculate the first two principal components (PCs) of contact maps at 10 kb resolution, select the appropriate PC for downstream analysis, normalize, sign-correct, and comparatively analyze the selected PC between cell types. Association of genes with compartmentalization was performed as follows. Transcriptional start sites were obtained from the RefSeq database [153]. All transcriptional isoforms were kept for each gene ID. Promoter regions were defined as 2 kb upstream/500 bp downstream of transcriptional start sites. Intersections between promoter regions and A/B compartment score bins were obtained using bedtools intersect. For promoters that intersected with more than one compartment score bin, the bin with the largest overlap was kept, and its compartment label used to annotate the gene associated with the promoter. Visualizations of contact maps were created using WaSHU Epigenome Browser and Juicebox [152, 154].

CTCF ChIP-seq and motif analysis

P0.5 lenses ($n=200$) were obtained from CD-1 mice, micro-dissected into lens epithelium and lens fibers and stored in liquid nitrogen prior the use as we described earlier [53]. Preparation for ChIP-seq was provided by ActiveMotif (Carlsbad, CA, U.S.A.). Briefly, immunoprecipitation was performed on 12 μ g chromatin from microdissected lens cells with 5 μ l anti-CTCF antibody (ActiveMotif cat. # 61,311, lot #11,219,006), $n=2$ biological replicates. The 75-nt single-end (SE75) sequence reads generated by Illumina sequencing (using NextSeq 500)

were mapped to the genome using the BWA algorithm (“bwa aln/samse” with default settings) [155]. Uniquely mapped reads passing Illumina’s purity filter with ≤ 2 mismatches were retained for downstream analyses. Duplicate reads were removed. Peaks were called using MACS2 [156]. Methylation profiles within CTCF peaks were plotted using deeptools 3.5.1 [157].

Analysis of RNA expression in lens cells

Bulk mouse lens RNA-seq data were generated earlier [64]. Boxplot showing top differentially expressed genes with fragments per kilobase of transcript per million mapped reads (FPKM). FPKM values were Log₂ transformed for scaling purposes. Callout boxes were used to highlight the most upregulated genes.

Cross-analysis of ChIP-seq and mRNA expression

Genes were ranked by their expression levels according to the earlier RNA-seq data [64]. The top 100 most expressed genes in fiber cells were selected. The longest transcripts associated with each gene ID were obtained from the RefSeq All database. Mean epithelium-specific and fiber-specific CTCF ChIP-seq signals and whole-lens RNA Polymerase II ChIP-seq signals were calculated across each transcript using multiBigwigSummary from the deeptools package. The data were merged with expression data of the top 100 expressed genes. Non-coding genes were manually removed to obtain the final list ($n=63$) associating expression, CTCF, and RNA Polymerase II signals in top-expressed genes in lens fiber.

Immunofluorescence analysis of lens nuclei

Eyes were fixed in 4.0% paraformaldehyde for 2 h at room temperature, transferred into 30% sucrose for cryopreservation and embedded in OCT (Tissue Tek). Tissue was stored at -80°C until used. Cryostat sections were cut in the transverse plane at 10 microns and stored at -20°C until IF protocol. Slides were permeabilized with 0.5% Triton-X-100 in PBS (PBS-T) for 30 min at room temperature. Slides were washed 3x PBS for 5 min each. Following permeabilization, slides were blocked with 4.0% BSA for 1 h at room temperature. Slides were washed 3x PBS for 5 min each. Slides were incubated for 24 h at 4 C with primary antibodies diluted in 1.0% BSA and PBS-T in a humidity chamber. Slides were then washed 4x PBS for 10 min each. Slides were incubated with secondary antibodies diluted in 1.0% BSA in PBS-T for 2 h at room temperature in humidity chamber at room temperature then washed 3x PBS for 10 min each, then a final wash in PBS with Hoechst (1:2,000) for 10 min. Slides were imaged on Leica SP8 at 63x magnification. Antibodies and dyes: CTCF (Santa Cruz sc-271,514, 1:100), Alexa Flour 488 (Jackson 115-547-185, 1:250) and Hoechst 33,342 (Fisher H3570, 1:2,000).

Abbreviations

ATAC-seq	Assay for transposase-accessible chromatin with high-throughput sequencing
ChIP	Chromatin immunoprecipitation
ES	Embryonic stem
cCREs	ENCODE candidate <i>cis</i> -regulatory elements
FISH	Fluorescence in situ hybridization
GO	Gene ontology
GWAS	Genome-wide association studies
kb	Kilo base pairs
IDRs	Intrinsically disordered regions
LCR	Locus control region
OFZ	Organelle free zone
PTM	Posttranslational modification
PCA	Principal component analysis
SMCs	Structural maintenance complexes
TADs	Topologically associated domains
3D	3-dimensional
WGBS	Whole genome bisulfite sequencing

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13072-024-00533-x>.

Supplementary Material 1
Supplementary Material 2
Supplementary Material 3
Supplementary Material 4
Supplementary Material 5
Supplementary Material 6
Supplementary Material 7
Supplementary Material 8
Supplementary Material 9
Supplementary Material 10
Supplementary Material 11
Supplementary Material 12

Acknowledgements

We thank Dr. Sitharam Ramaswami at The New York University Langone Health Genome Technology Center for providing Hi-C services. We are grateful to Dr. Roy S. Chuck for support. We thank Dr. Meelad Dawlaty for mouse ES cells. We thank Drs. Daniel Chauss and Robert Coleman for early advice and discussions. We thank Analytical Imaging Facility of Albert Einstein College of Medicine for use of the Leica SP8 confocal microscope.

Author contributions

MC performed all experiments and partial data analysis and wrote major portions of the manuscript. WC analyzed the data and wrote specific individual subsection. AC conceived the study, contributed to the data analysis, and edited the manuscript.

Funding

NIH R01 EY012200, EY014237 and EY014237-19S1 (to A.C.). NCI Cancer Center Support Grant P30CA013330. Shared Instrument Grant 1S100D02359.

Availability of data and material

Hi-C and CTCF ChIP-seq data were deposited into Gene Expression Omnibus (GEO) with accession ID GSE243851. ATAC-seq (GSE124497; [62]), RNA-seq (GSE113887; [64]), WGBS (GSE213901; [63]), ChIP-seq (GSE66961; [53, 65]).

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Animal studies using mice were approved by the Institute of Animal Studies at the Albert Einstein College of Medicine, protocols #20181105 and 00001533.

Competing interests

The authors declare no competing interests.

Author details

¹The Departments Genetics, Albert Einstein College of Medicine, NY10461 Bronx, USA

²Ophthalmology and Visual Sciences, Albert Einstein College of Medicine, NY10461 Bronx, USA

Received: 15 November 2023 / Accepted: 8 March 2024

Published online: 20 April 2024

References

- Sabari BR, Dall'Agnese A, Young RA. Biomolecular condensates in the Nucleus. *Trends Biochem Sci.* 2020;45:961–77. <https://doi.org/10.1016/j.tibs.2020.06.007>.
- Misteli T. The Self-Organizing genome: principles of Genome Architecture and function. *Cell.* 2020;183:28–45. <https://doi.org/10.1016/j.cell.2020.09.014>.
- Ling X, Liu X, Jiang S, Fan L, Ding J. The dynamics of three-dimensional chromatin organization and phase separation in cell fate transitions and diseases. *Cell Regen.* 2022;11:42. <https://doi.org/10.1186/s13619-022-00145-4>.
- Cremer T, Cremer C, Schneider T, Baumann H, Hens L, Kirsch-Volders M. Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments. *Hum Genet.* 1982;62:201–9. <https://doi.org/10.1007/BF00333519>.
- Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet.* 2001;2:292–301. <https://doi.org/10.1038/35066075>.
- Kosak ST, Skok JA, Medina KL, Riblet R, Le Beau MM, Fisher AG, Singh H. Subnuclear compartmentalization of immunoglobulin loci during Lymphocyte Development. *Science.* 2002;296:158–62. <https://doi.org/10.1126/science.1068768>.
- van Steensel B, Belmont AS. 2017 Lamina-Associated domains: links with chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* 169, 780–91. (<https://doi.org/10.1016/j.cell.2017.04.022>).
- Bonev B, Cavalli G. Organization and function of the 3D genome. *Nat Rev Genet.* 2016;17:661–78. <https://doi.org/10.1038/nrg.2016.112>.
- Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, et al. Chromatin architecture reorganization during stem cell differentiation. *Nature.* 2015;518:331–6. <https://doi.org/10.1038/nature14222>.
- Winick-Ng W, Kukalev A, Harabula I, Zea-Redondo L, Szabó D, Meijer M, Serebreni L, Zhang Y, Bianco S, Chiariello AM, et al. Cell-type specialization is encoded by specific chromatin topologies. *Nature.* 2021;599:684–91. <https://doi.org/10.1038/s41586-021-04081-2>.
- Long HS, Greenaway S, Powell G, Mallon A-M, Lindgren CM, Simon MM. Making sense of the linear genome, gene function and TADs. *Epigenetics Chromatin.* 2022;15(4). <https://doi.org/10.1186/s13072-022-00436-9>.
- Tan J, Shenker-Tauris N, Rodriguez-Hernaez J, Wang E, Sakellaropoulos T, Boccacatte F, Thandapani P, Skok J, Aifantis I, Fenyő D, et al. Cell-type-specific prediction of 3D chromatin organization enables high-throughput in silico genetic screening. *Nat Biotechnol.* 2023. <https://doi.org/10.1038/s41587-022-01612-8>.
- Forte G, Buckle A, Boyle S, Marenduzzo D, Gilbert N, Brackley CA. Transcription modulates chromatin dynamics and locus configuration sampling. *Nat Struct Mol Biol.* 2023;30:1275–85. <https://doi.org/10.1038/s41594-023-01059-8>.
- Merkenschlager M, Odom DT. 2013 CTCF and cohesin: linking gene regulatory elements with their targets. *Cell.* 152, 1285–1297. (10.1016/j.cell.2013.02.029).
- Rowley MJ, Corces VG. 2018 Organizational principles of 3D genome architecture. *Nat Rev Genet.* 19, 789–800. (10.1038/s41576-018-0060-8).
- van Ruiten MS, Rowland BD. 2018 SMC Complexes: Universal DNA Looping Machines with Distinct Regulators. *Trends Genet.* 34, 477–487. (10.1016/j.tig.2018.03.003).
- Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. 2017 Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell.* 169, 930–944.e922. (10.1016/j.cell.2017.05.004).
- Wutz G, Várnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, Schoenfelder S, Jessberger G, Muhar M, Hossain MJ. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* 2017;36:3573–99.
- Pugacheva EM, Kubo N, Loukinov D, Tajmull M, Kang S, Kovalchuk AL, Strunnikov AV, Zentner GE, Ren B, Lobanenko VV. 2020 CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. *Proceedings of the National Academy of Sciences.* 117, 2020–2031.
- Hsieh T-HS, Cattoglio C, Slobodyanyuk E, Hansen AS, Darzacq X, Tjian R. Enhancer–promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1. *Nat Genet.* 2022;54:1919–32. <https://doi.org/10.1038/s41588-022-01223-8>.
- Klenova EM, Nicolas RH, Paterson HF, Carne AF, Heath CM, Goodwin GH, Neiman PE, Lobanenko VV. 1993 CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-myc gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol Cell Biol.* 13, 7612–7624. (10.1128/mcb.13.12.7612-7624.1993).
- Saldaña-Meyer R, Rodriguez-Hernaez J, Escobar T, Nishana M, Jácome-López K, Nora EP, Bruneau BG, Tsigos A, Furlan-Magaril M, Skok J et al. 2019 RNA Interactions Are Essential for CTCF-Mediated Genome Organization. *Mol Cell.* 76, 412–422.e415. (10.1016/j.molcel.2019.08.015)
- Oh HJ, Aguilar R, Kesner B, Lee HG, Kriz AJ, Chu HP, Lee JT. 2021 Jpx RNA regulates CTCF anchor site selection and formation of chromosome loops. *Cell.* 184, 6157–6173.e6124. (10.1016/j.cell.2021.11.012).
- Hashimoto H, Wang D, Horton JR, Zhang X, Corces VG, Cheng X. 2017 structural basis for the versatile and methylation-dependent binding of CTCF to DNA. *Mol Cell.* 66, 711–20. e713.
- Zhou R, Tian K, Huang J, Duan W, Fu H, Feng Y, Wang H, Jiang Y, Li Y, Wang R. CTCF DNA-binding domain undergoes dynamic and selective protein–protein interactions. *Science.* 2022;25:105011.
- Vietri Rudan M, Hadjur S. Genetic tailors: CTCF and cohesin shape the genome during evolution. *Trends Genet.* 2015;31:651–60. <https://doi.org/10.1016/j.tig.2015.09.004>.
- Ghirlando R, Felsenfeld G. CTCF: making the right connections. *Genes Dev.* 2016;30:881–91. <https://doi.org/10.1101/gad.277863.116>.
- Gabriele M, Brandão HB, Grosse-Holz S, Jha A, Dailey GM, Cattoglio C, Hsieh T-HS, Mirny L, Zechner C, Hansen AS. Dynamics of CTCF-and cohesin-mediated chromatin looping revealed by live-cell imaging. *Science.* 2022;376:496–501.
- Aljahani A, Hua P, Karpinska MA, Quililan K, Davies JO, Oudelaar AM. 2022 analysis of sub-kilobase chromatin topology reveals nano-scale regulatory interactions with variable dependence on cohesin and CTCF. *Nat Commun.* 13, 2139.
- Chakraborty S, Kopitchinski N, Zuo Z, Eraso A, Awasthi P, Chari R, Mitra A, Tobias IC, Moorthy SD, Dale RK. 2023 enhancer–promoter interactions can bypass CTCF-mediated boundaries and contribute to phenotypic robustness. *Nat Genet.* 55, 280–90.
- Davidson IF, Barth R, Zaczek M, van der Torre J, Tang W, Nagasaka K, Janissen R, Kerssemakers J, Wutz G, Dekker C. 2023 CTCF is a DNA-tension-dependent barrier to cohesin-mediated loop extrusion. *Nature.* 1–6.
- Gu B, Comerici CJ, McCarthy DG, Saurabh S, Moerner WE, Wysocka J. 2020 Opposing Effects of Cohesin and Transcription on CTCF Organization Revealed by Super-resolution Imaging. *Mol Cell.* 80, 699–711.e697. (10.1016/j.molcel.2020.10.001).
- Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO et al. 2009 Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science.* 326, 289–293. (10.1126/science.1181369)
- Dekker J, Rippe K, Dekker M, Kleckner N. 2002 Capturing Chromosome Conformation. *Science.* 295, 1306–1311. (10.1126/science.1067799).
- Eagen KP. 2018 Principles of Chromosome Architecture Revealed by Hi-C. *Trends Biochem Sci.* 43, 469–478. (10.1016/j.tibs.2018.03.006).
- Pal K, Forcato M, Ferrari F. Hi-C analysis: from data generation to integration. *Biophys Rev.* 2019;11:67–78.
- Liu N, Low WY, Alinejad-Rokny H, Pederson S, Sadlon T, Barry S, Breen J. 2021 seeing the forest through the trees: prioritising potentially functional interactions from Hi-C. *Epigenetics Chromatin.* 14, 1–17.

38. Andrey G, Mundlos S. 2017 The three-dimensional genome: regulating gene expression during pluripotency and development. *Development*. 144, 3646–3658. (10.1242/dev.148304).
39. MacGregor IA, Adams IR, Gilbert N. Large-scale chromatin organisation in interphase, mitosis and meiosis. *Biochem J*. 2019;476:2141–56.
40. Collas P, Liyakat Ali TM, Brunet A, Germier T. Finding friends in the crowd: three-dimensional cliques of topological genomic domains. *Front Genet*. 2019;10:602.
41. Vickaryous MK, Hall BK. 2006 Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. *Biol Rev Camb Philos Soc*. 81, 425–455. (10.1017/s1464793106007068).
42. Lovicu F, McAvoy J. 2005 Growth factor regulation of lens development. *Developmental biology*. 280, 1–14. (10.1016/j.ydbio.2005.01.020).
43. Bassnett S, Shi Y, Vrensen GF. 2011 Biological glass: structural determinants of eye lens transparency. *Philos Trans R Soc Lond B Biol Sci*. 366, 1250–1264. (10.1098/rstb.2010.0302).
44. Gunhaga L. 2011 The lens: a classical model of embryonic induction providing new insights into cell determination in early development. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 366, 1193–1203. (10.1098/rstb.2010.0175).
45. Cvekl A, Zhang X. Signaling and Gene Regulatory Networks in mammalian Lens Development. *Trends Genet*. 2017;33:677–702. <https://doi.org/10.1016/j.tig.2017.08.001>.
46. Cvekl A, Eliscovich C. 2021 Crystallin gene expression: Insights from studies of transcriptional bursting. *Experimental Eye Research*. 207, 108564. (10.1016/j.exer.2021.108564).
47. Vrensen GF, Graw J, De Wolf A. 1991 Nuclear breakdown during terminal differentiation of primary lens fibres in mice: a transmission electron microscopic study. *Exp Eye Res*. 52, 647–659. (10.1016/0014-4835(91)90017-9).
48. Chaffee BR, Shang F, Chang M-L, Clement TM, Eddy EM, Wagner BD, Nakahara M, Nagata S, Robinson ML, Taylor A. 2014 Nuclear removal during terminal lens fiber cell differentiation requires CDK1 activity: appropriating mitosis-related nuclear disassembly. *Development*. 141, 3388–3398. (10.1242/dev.106005).
49. He S, Limi S, McGreal RS, Xie Q, Brennan LA, Kantorow WL, Kokavec J, Majumdar R, Hou H Jr, Edelmann W et al. 2016 Chromatin remodeling enzyme Snf2h regulates embryonic lens differentiation and denucleation. *Development*. 143, 1937–1947. (10.1242/dev.135285).
50. Lyu L, Whitcomb EA, Jiang S, Chang ML, Gu Y, Duncan MK, Cvekl A, Wang WL, Limi S, Reneker LW et al. 2016 Unfolded-protein response-associated stabilization of p27(Cdkn1b) interferes with lens fiber cell denucleation, leading to cataract. *Faseb j*. 30, 1087–1095. (10.1096/fj.15-278036).
51. Limi S, Senecal A, Coleman R, Lopez-Jones M, Guo P, Polumbo C, Singer RH, Skoultschi AI, Cvekl A. 2018 Transcriptional burst fraction and size dynamics during lens fiber cell differentiation and detailed insights into the denucleation process. *Journal of Biological Chemistry*. 293, 13176–13190. (10.1074/jbc.RA118.001927).
52. Martynova E, Zhao Y, Xie Q, Zheng D, Cvekl A. Transcriptomic analysis and novel insights into lens fibre cell differentiation regulated by Gata3. *Open Biology*. 2019;9:190220.
53. Sun J, Rockowitz S, Chauss D, Wang P, Kantorow M, Zheng D, Cvekl A. Chromatin features, RNA polymerase II and the comparative expression of lens genes encoding crystallins, transcription factors, and autophagy mediators. *Mol Vis*. 2015;21:955–73.
54. Hamai Y, Fukui HN, Kuwabara T. 1974 Morphology of hereditary mouse cataract. *Exp Eye Res*. 18, 537–546. (10.1016/0014-4835(74)90060-8).
55. Pendergrass W, Penn P, Possin D, Wolf N. 2005 Accumulation of DNA, nuclear and mitochondrial debris, and ROS at sites of age-related cortical cataract in mice. *Invest Ophthalmol Vis Sci*. 46, 4661–4670. (10.1167/iovs.05-0808).
56. Pendergrass WR, Penn PE, Possin DE, Wolf NS. 2006 Cellular debris and ROS in age-related cortical cataract are caused by inappropriate involution of the surface epithelial cells into the lens cortex. *Mol Vis*. 12, 712–724.
57. Dekker J, Alber F, Aufmkolk S, Beliveau BJ, Bruneau BG, Belmont AS, Bintu L, Boettiger A, Calandrelli R, Disteche CM et al. . 2023 *Spatial and temporal organization of the genome: Current state and future aims of the 4D nucleome project*. *Molecular Cell*. 83, 2624–2640. (10.1016/j.molcel.2023.06.018)
58. Norrie JL, Lupo MS, Xu B, Al Dirri I, Valentine M, Putnam D, Griffiths L, Zhang J, Johnson D, Easton J et al. . 2019 *Nucleome Dynamics during Retinal Development*. *Neuron*. 104, 512–528.e511. (10.1016/j.neuron.2019.08.002)
59. Marchal C, Singh N, Batz Z, Advani J, Jaeger C, Corso-Díaz X, Swaroop A. 2022 High-resolution genome topology of human retina uncovers super enhancer-promoter interactions at tissue-specific and multifactorial disease loci. *Nat Commun*. 13, 5827. (10.1038/s41467-022-33427-1).
60. Li M, Huang H, Wang B, Jiang S, Guo H, Zhu L, Wu S, Liu J, Wang L, Lan X. Comprehensive 3D epigenomic maps define limbal stem/progenitor cell function and identity. *Nat Commun*. 2022;13:1293.
61. McGreal-Estrada RS, Wolf LV, Cvekl A. 2018 Promoter-enhancer looping and shadow enhancers of the mouse αA-crystallin locus. *Biol Open*. 7, (10.1242/bio.036897).
62. Zhao Y, Zheng D, Cvekl A. 2019 Profiling of chromatin accessibility and identification of general cis-regulatory mechanisms that control two ocular lens differentiation pathways. *Epigenetics & Chromatin*. 12, 27. (10.1186/s13072-019-0272-y).
63. Chang W, Zhao Y, Rayêe D, Xie Q, Suzuki M, Zheng D, Cvekl A. 2023 Dynamic changes in whole genome DNA methylation, chromatin and gene expression during mouse lens differentiation. *Epigenetics & Chromatin*. 16, 4. (10.1186/s13072-023-00478-7).
64. Zhao Y, Zheng D, Cvekl A. A comprehensive spatial-temporal transcriptomic analysis of differentiating nascent mouse lens epithelial and fiber cells. *Exp Eye Res*. 2018;175:56–72. <https://doi.org/10.1016/j.exer.2018.06.004>.
65. Sun J, Rockowitz S, Xie Q, Ashery-Padan R, Zheng D, Cvekl A. 2015 Identification of in vivo DNA-binding mechanisms of Pax6 and reconstruction of Pax6-dependent gene regulatory networks during forebrain and lens development. *Nucleic Acids Research*. 43, 6827–6846. (10.1093/nar/gkv589).
66. Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot JP, Tanay A et al. . 2017 *Multiscale 3D Genome Rewiring during Mouse Neural Development*. *Cell*. 171, 557–572.e524. (10.1016/j.cell.2017.09.043)
67. Marchal C, Singh N, Corso-Díaz X, Swaroop A. 2021 HiCRes: a computational method to estimate and predict the genomic resolution of Hi-C libraries. *Nucleic Acids Research*. 50, e35–e35. (10.1093/nar/gkab1235).
68. Chakraborty A, Wang JG, Ay F. 2022 dChIC detects differential compartments across multiple Hi-C datasets. *Nature Communications*. 13, 6827. (10.1038/s41467-022-34626-6).
69. Novo CL, Javierre BM, Cairns J, Segonds-Pichon A, Wingett SW, Freire-Pritchett P, Furlan-Magaril M, Schoenfelder S, Fraser P, Rugg-Gunn PJ. 2018 Long-Range Enhancer Interactions Are Prevalent in Mouse Embryonic Stem Cells and Are Reorganized upon Pluripotent State Transition. *Cell Rep*. 22, 2615–2627. (10.1016/j.celrep.2018.02.040).
70. Fang F, Xu Y, Chew K-K, Chen X, Ng H-H, Matsudaira P. 2014 Coactivators p300 and CBP Maintain the Identity of Mouse Embryonic Stem Cells by Mediating Long-Range Chromatin Structure. *Stem Cells*. 32, 1805–1816. (10.1002/stem.1705).
71. Pękowska A, Klaus B, Xiang W, Severino J, Daigle N, Klein FA, Oleś M, Casellas R, Ellenberg J, Steinmetz LM. 2018 Gain of CTCF-anchored chromatin loops marks the exit from naive pluripotency. *Cell Syst*. 7, 482–95. e410.
72. Cvekl A, Ashery-Padan R. 2014 The cellular and molecular mechanisms of vertebrate lens development. *Development*. 141, 4432–4447. (10.1242/dev.107953).
73. Cvekl A, Callaerts P. 25th anniversary and more to learn. *Exp Eye Res*. 2017;PAX6:156, 10–21. <https://doi.org/10.1016/j.exer.2016.04.017>.
74. Marakhonov AV, Vasilyeva TA, Voskresenskaya AA, Sukhanova NV, Kadyshchev VV, Kutsev SI, Zinchenko RA. 2019 LMO2 gene deletions significantly worsen the prognosis of Wilms' tumor development in patients with WAGR syndrome. *Human Molecular Genetics*. 28, 3323–3326. (10.1093/hmg/ddz168).
75. Fischbach BV, Trout KL, Lewis J, Luis CA, Sika M. 2005 WAGR Syndrome: a clinical review of 54 cases. *Pediatrics*. 116, 984–8. (10.1542/peds.2004-0467).
76. Antosova B, Smolkova J, Klimova L, Lachova J, Bendova M, Kozmikova I, Machon O, Kozmik Z. 2016 The Gene Regulatory Network of Lens Induction Is Wired through Meis-Dependent Shadow Enhancers of Pax6. *PLOS Genetics*. 12, e1006441. (10.1371/journal.pgen.1006441).
77. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, Aiden EL. 2016 Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst*. 3, 95–98. (10.1016/j.cels.2016.07.002).
78. Gopal-Srivastava R, Cvekl A, Piatigorsky J. 1998 Involvement of retinoic acid/retinoid receptors in the regulation of murine αB-crystallin/small heat shock protein gene expression in the lens. *Journal of Biological Chemistry*. 273, 17954–17961. (10.1074/jbc.273.28.17954).
79. Yang Y, Chauhan BK, Cveklava K, Cvekl A. Transcriptional regulation of mouse αB- and γF-crystallin genes in lens: opposite promoter-specific interactions between Pax6 and large maf transcription factors. *J Mol Biol*. 2004;344:351–68.

80. Choquet H, Melles RB, Anand D, Yin J, Cuellar-Partida G, Wang W, Team aR, Hoffmann TJ, Nair KS, Hysi PG. 2021 a large multiethnic GWAS meta-analysis of cataract identifies new risk loci and sex-specific effects. *Nat Commun.* 12, 3595.
81. Stump RJ, Ang S, Chen Y, von Bahr T, Lovicu FJ, Pinson K, de longh RU, Yamaguchi TP, Sassoan DA, McAvoy JW. 2003 A role for Wnt/beta-catenin signaling in lens epithelial differentiation. *Dev Biol.* 259, 48–61. (10.1016/s0012-1606(03)00179-9).
82. Dawes LJ, Sugiyama Y, Lovicu FJ, Harris CG, Shelley EJ, McAvoy JW. 2014 interactions between lens epithelial and fiber cells reveal an intrinsic self-assembly mechanism. *Dev Biol.* 385, 291–303. (<https://doi.org/10.1016/j.ydbio.2013.10.030>).
83. Kerr CL, Huang J, Williams T, West-Mays JA. 2012 Activation of the hedgehog signaling pathway in the developing lens stimulates ectopic FoxE3 expression and disruption in fiber cell differentiation. *Investigative ophthalmology & visual science.* 53, 3316–3330. (10.1167/iops.12-9595).
84. Brown NL, Cheema S, Torre L, A. Diverse eye defects occur by varying the developmental timing of hedgehog signaling. *Investig Ophthalmol Vis Sci.* 2019;60:4307–4307.
85. Furuta Y, Hogan BL. 1998 BMP4 is essential for lens induction in the mouse embryo. *Genes Dev.* 12, 3764–3775. (10.1101/gad.12.23.3764).
86. Matt N, Ghyselinck NB, Pellerin I, Dupé V. Impairing retinoic acid signalling in the neural crest cells is sufficient to alter entire eye morphogenesis. *Dev Biol.* 2008;320:140–8.
87. Bosze B, Suarez-Navarro J, Soofi A, Lauderdale JD, Dressler GR, Brown NL. 2021 multiple roles for Pax2 in the embryonic mouse eye. *Dev Biol.* 472, 18–29.
88. Beebe DC, Vasiliev O, Guo J, Shui Y-B, Bassnett S. Changes in adhesion complexes define stages in the differentiation of lens fiber cells. *Investig Ophthalmol Vis Sci.* 2001;42:727–34.
89. Pontoriero GF, Smith AN, Miller L-AD, Radice GL, West-Mays JA, Lang RA. Co-operative roles for E-cadherin and N-cadherin during lens vesicle separation and lens epithelial cell survival. *Dev Biol.* 2009;326:403–17.
90. Nishiguchi S, Wood H, Kondoh H, Lovell-Badge R, Episkopou V. Sox1 directly regulates the γ -crystallin genes and is essential for lens development in mice. *Genes Dev.* 1998;12:776–81.
91. Robinson ML. An essential role for FGF receptor signaling in lens development. *Semin Cell Dev Biol.* 2006;17:726–40.
92. Lovicu F, McAvoy J, De longh R. 2011 Understanding the role of growth factors in embryonic development: insights from the lens. *Philosophical Transactions of the Royal Society B: Biological Sciences.* 366, 1204–1218. (10.1098/rstb.2010.0339).
93. Makrides N, Wang Q, Tao C, Schwartz S, Zhang X. Jack of all trades, master of each: the diversity of fibroblast growth factor signalling in eye development. *Open Biology.* 2022;12:210265.
94. Cvekl A, Wang W-L. Retinoic acid signaling in mammalian eye development. *Exp Eye Res.* 2009;89:280–91. (<https://doi.org/10.1016/j.exer.2009.04.012>).
95. Fromm L, Overbeek PA. 1997 Inhibition of cell death by lens-specific over-expression of bcl-2 in transgenic mice. *Dev Genet.* 20, 276–287. ([https://doi.org/10.1002/\(sici\)1520-6408\(1997\)20:3](https://doi.org/10.1002/(sici)1520-6408(1997)20:3)).
96. Alonso-Gil D, Losada A. 2023 NIPBL and cohesin: new take on a classic tale. *Trends in Cell Biology.* 33, 860–871. (10.1016/j.tcb.2023.03.006).
97. Aldiri I, Xu B, Wang L, Chen X, Hiler D, Griffiths L, Valentine M, Shirinifard A, Thiagarajan S, Sablauer A et al. . 2017 *The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis.* *Neuron.* 94, 550–568.e510. (10.1016/j.neuron.2017.04.022)
98. Prickett AR, Barkas N, McCole RB, Hughes S, Amante SM, Schulz R, Oakey RJ. Genome-wide and parental allele-specific analysis of CTCF and cohesin DNA binding in mouse brain reveals a tissue-specific binding pattern and an association with imprinted differentially methylated regions. *Genome Res.* 2013;23:1624–35.
99. Vian L, Pękowska A, Rao SSP, Kieffer-Kwon KR, Jung S, Baranello L, Huang SC, El Khattabi L, Dose M, Pruett N et al. . 2018 *The Energetics and Physiological Impact of Cohesin Extrusion.* *Cell.* 173, 1165–1178.e1120. (10.1016/j.cell.2018.03.072)
100. Loukinov DI, Pugacheva E, Vatolin S, Pack SD, Moon H, Chernukhin I, Mannan P, Larsson E, Kanduri C, Vostrov AA et al. . 2002 BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc Natl Acad Sci U S A.* 99, 6806–6811. (10.1073/pnas.092123699).
101. Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, Lee K, Canfield T, Weaver M, Sandstrom R. 2012 widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* 22, 1680–8.
102. Renaud S, Loukinov D, Abdullaev Z, Guilleret I, Bosman F, Lobanenko V, Benhattar J. Dual role of DNA methylation inside and outside of CTCF-binding regions in the transcriptional regulation of the telomerase hTERT gene. *Nucleic Acids Res.* 2007;35:1245–56.
103. Snyder MP, Gingeras TR, Moore JE, Weng Z, Gerstein MB, Ren B, Hardison RC, Stamatoyannopoulos JA, Graveley BR. 2020 perspectives on ENCODE. *Nature* 583, 693–8.
104. Luo Z, Lin C, Woodfin AR, Bartom ET, Gao X, Smith ER, Shilatfard A. 2016 Regulation of the imprinted Dlk1-Dio3 locus by allele-specific enhancer activity. *Genes Dev.* 30, 92–101. (10.1101/gad.270413.115).
105. Xi J, Xu Y, Guo Z, Li J, Wu Y, Sun Q, Wang Y, Chen M, Zhu S, Bian S. 2022 LncRNA SOX1-OT V1 acts as a decoy of HDAC10 to promote SOX1-dependent hESC neuronal differentiation. *EMBO Rep.* 23, e53015.
106. McNulty R, Wang H, Mathias RT, Ortwerth BJ, Truscott RJW, Bassnett S. 2004 Regulation of tissue oxygen levels in the mammalian lens. *The Journal of Physiology.* 559, 883–898. (10.1113/jphysiol.2004.068619).
107. Kaelin WG Jr. Proline hydroxylation and gene expression. *Annu Rev Biochem.* 2005;74:115.
108. Semenza GL. 2012 Hypoxia-inducible factors in physiology and medicine. *Cell.* 148, 399–408. (10.1016/j.cell.2012.01.021).
109. Shui Y-B, Beebe DC. 2008 Age-dependent control of lens growth by hypoxia. *Investigative ophthalmology & visual science.* 49, 1023–1029. (10.1167/iops.07-1164).
110. Beebe DC. Maintaining transparency: a review of the developmental physiology and pathophysiology of two avascular tissues. *Semin Cell Dev Biol.* 2008;19:125–33. (<https://doi.org/10.1016/j.semcdb.2007.08.014>).
111. Yang Y, Cvekl A. Tissue-specific regulation of the mouse α A-crystallin gene in Lens via recruitment of Pax6 and c-Maf to its promoter. *J Mol Biol.* 2005;351:453–69. (<https://doi.org/10.1016/j.jmb.2005.05.072>).
112. Yang Y, Stopka T, Golestaneh N, Wang Y, Wu K, Li A, Chauhan BK, Gao CY, Cveklóvá K, Duncan MK et al. 2006 Regulation of alphaA-crystallin via Pax6, c-Maf, CREB and a broad domain of lens-specific chromatin. *Embo j.* 25, 2107–2118. (10.1038/sj.emboj.7601114).
113. Xie Q, McGreal R, Harris R, Gao CY, Liu W, Reneker LW, Musil LS, Cvekl A. 2016 Regulation of c-Maf and α A-Crystallin in Ocular Lens by Fibroblast Growth Factor Signaling. *Journal of Biological Chemistry.* 291, 3947–3958. (10.1074/jbc.M115.705103).
114. Morgenbesser SD, Williams BO, Jacks T, DePino RA. 1994 p53-dependent apoptosis produced by Rb-deficiency in the developing mouse lens. *Nature.* 371, 72–74. (10.1038/371072a0).
115. Cvekl A, Kashanchi F, Brady JN, Piatigorsky J. Pax-6 interactions with TATA-box-binding protein and retinoblastoma protein. *Investig Ophthalmol Vis Sci.* 1999;40:1343–50.
116. Cvekl A, Yang Y, Chauhan BK, Cveklóva K. 2004 Regulation of gene expression by Pax6 in ocular cells: a case of tissue-preferred expression of crystallins in lens. *The International journal of developmental biology.* 48, 829. (10.1387/ijdb.041866ac).
117. Zhao B, Mei Y, Schipma MJ, Roth EW, Bleher R, Rappoport JZ, Wickrema A, Yang J, Ji P. 2016 Nuclear condensation during mouse erythropoiesis requires caspase-3-mediated nuclear opening. *Developmental cell.* 36, 498–510. (10.1016/j.devcel.2016.02.001).
118. Yoshida H, Kawane K, Koike M, Mori Y, Uchiyama Y, Nagata S. Phosphatidylserine-dependent engulfment by macrophages of nuclei from erythroid precursor cells. *Nature.* 2005;437:754–8.
119. Limi S, Zhao Y, Guo P, Lopez-Jones M, Zheng D, Singer RH, Skoultchi AI, Cvekl A. Bidirectional analysis of Cryba4-Crybb1 nascent transcription and nuclear accumulation of Crybb3 mRNAs in lens fibers. *Investig Ophthalmol Vis Sci.* 2019;60:234–44.
120. Wang H, Zhou R, Ji X. Droplet formation assay for investigating phase-separation mechanisms of RNA Pol II transcription and CTCF functioning. *STAR Protocols.* 2023;4:102202. (<https://doi.org/10.1016/j.xpro.2023.102202>).
121. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. 2017 a phase separation model for transcriptional control. *Cell* 169, 13–23.
122. Lee R, Kang M-K, Kim Y-J, Yang B, Shim H, Kim S, Kim K, Yang CM, Min B-g, Jung W-J. 2022 CTCF-mediated chromatin looping provides a topological framework for the formation of phase-separated transcriptional condensates. *Nucleic Acids Res.* 50, 207–26.
123. Jaenisch R, Young R. 2008 Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell.* 132, 567–582. (10.1016/j.cell.2008.01.015).

124. Kamachi Y, Uchikawa M, Tanouchi A, Sekido R, Kondoh H. 2001 Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. *Genes Dev.* 15, 1272–1286. (10.1101/gad.887101).
125. Uchikawa M, Ishida Y, Takemoto T, Kamachi Y, Kondoh H. Functional analysis of chicken Sox2 enhancers highlights an array of diverse regulatory elements that are conserved in mammals. *Dev Cell.* 2003;4:509–19. (10.1016/s1534-5807(03)00088–1).
126. Blixt Å, Mahlapuu M, Aitola M, Pelto-Huikko M, Enerbäck S, Carlsson P. A forkhead gene, FoxE3, is essential for lens epithelial proliferation and closure of the lens vesicle. *Genes Dev.* 2000;14:245–54.
127. Song S, Landsbury A, Dahm R, Liu Y, Zhang Q, Quinlan RA. 2009 Functions of the intermediate filament cytoskeleton in the eye lens. *The Journal of clinical investigation.* 119, 1837–1848. (10.1172/JCI38277).
128. Zhu Y, Tazearslan C, Suh Y. Challenges and progress in interpretation of non-coding genetic variants associated with human disease. *Experimental Biology Med.* 2017;242:1325–34.
129. Hingorani M, Hanson I, van Heyningen V. 2012 Aniridia. *European Journal of Human Genetics.* 20, 1011–1017. (10.1038/ejhg.2012.100).
130. Williamson KA, FitzPatrick DR. The genetic architecture of microphthalmia, anophthalmia and coloboma. *Eur J Med Genet.* 2014;57:369–80.
131. Lima Cunha D, Arno G, Corton M, Moosajee M. 2019 The Spectrum of PAX6 Mutations and Genotype-Phenotype Correlations in the Eye. *Genes (Basel).* 10, 1050. (10.3390/genes10121050).
132. Lu L, Liu X, Huang WK, Giusti-Rodríguez P, Cui J, Zhang S, Xu W, Wen Z, Ma S, Rosen JD et al. 2020 Robust Hi-C Maps of Enhancer-Promoter Interactions Reveal the Function of Non-coding Genome in Neural Development and Diseases. *Mol Cell.* 79, 521–534.e515. (10.1016/j.molcel.2020.06.007)
133. Martin P, McGovern A, Massey J, Schoenfelder S, Duffus K, Yarwood A, Barton A, Worthington J, Fraser P, Eyre S et al. 2016 Identifying Causal Genes at the Multiple Sclerosis Associated Region 6q23 Using Capture Hi-C. *PLOS ONE.* 11, e0166923. (10.1371/journal.pone.0166923).
134. van de Nobelen S, Rosa-Garrido M, Leers J, Heath H, Soochit W, Joosen L, Jonkers I, Demmers J, van der Reijden M, Torrano V et al. 2010 CTCF regulates the local epigenetic state of ribosomal DNA repeats. *Epigenetics Chromatin.* 3, 19. (10.1186/1756-8935-3-19).
135. Huang K, Jia J, Wu C, Yao M, Li M, Jin J, Jiang C, Cai Y, Pei D, Pan G. Ribosomal RNA gene transcription mediated by the master genome regulator protein CCCTC-binding factor (CTCF) is negatively regulated by the condensin complex. *J Biol Chem.* 2013;288:26067–77.
136. Torrano V, Chernukhin I, Docquier F, D'Arcy V, León J, Klenova E, Delgado MD. 2005 CTCF regulates growth and erythroid differentiation of human myeloid leukemia cells. *J Biol Chem.* 280, 28152–61.
137. Ouboussad L, Kreuz S, Lefevre PF. CTCF depletion alters chromatin structure and transcription of myeloid-specific factors. *J Mol Cell Biol.* 2013;5:308–22.
138. He S, Pirity MK, Wang W-L, Wolf L, Chauhan BK, Cveklava K, Tamm ER, Ashery-Padan R, Metzger D, Nakai A. 2010 chromatin remodeling enzyme Brg1 is required for mouse lens fiber cell terminal differentiation and its denucleation. *Epigenetics Chromatin.* 3, 1–20. (<https://doi.org/10.1186/1756-8935-3-21>).
139. Wolf L, Harrison W, Huang J, Xie Q, Xiao N, Sun J, Kong L, Lachke SA, Kuracha MR, Govindarajan V et al. 2013 Histone posttranslational modifications and cell fate determination: lens induction requires the lysine acetyltransferases CBP and p300. *Nucleic Acids Research.* 41, 10199–10214. (10.1093/nar/gkt824).
140. Wang W-L, Li Q, Xu J, Cvekl A. 2010 Lens fiber cell differentiation and denucleation are disrupted through expression of the N-terminal nuclear receptor box of NCOA6 and result in p53-dependent and p53-independent apoptosis. *Molecular biology of the cell.* 21, 2453–2468. (10.1091/mbc.E09-12-1031).
141. Lu J, An J, Wang J, Cao X, Cao Y, Huang C, Jiao S, Yan D, Lin X, Zhou X. 2022 Znhit1 Regulates p21Cip1 to Control Mouse Lens Differentiation. *Investigative ophthalmology & visual science.* 63, 18–18. (10.1167/iovs.63.4.18).
142. Yang Y, Wolf LV, Cvekl A. Distinct embryonic expression and localization of CBP and p300 histone acetyltransferases at the mouse α -crystallin locus in lens. *J Mol Biol.* 2007;369:917–26.
143. Bachmann C, Nguyen H, Rosenbusch J, Pham L, Rabe T, Patwa M, Sokpor G, Seong RH, Ashery-Padan R, Mansouri A. 2016 mSWI/SNF (BAF) complexes are indispensable for the neurogenesis and development of embryonic olfactory epithelium. *PLoS Genet.* 12, e1006274.
144. Ovadia S, Cui G, Elkon R, Cohen-Gulker M, Zuk-Bar N, Tuoc T, Jing N, Ashery-Padan R. 2023 SWI/SNF complexes are required for retinal pigmented epithelium differentiation and for the inhibition of cell proliferation and neural differentiation programs. *Development.* 150.
145. Sun J, Zhao Y, McGreal R, Cohen-Tayar Y, Rockowitz S, Wilczek C, Ashery-Padan R, Shechter D, Zheng D, Cvekl A. 2016 Pax6 associates with H3K4-specific histone methyltransferases Mll1, Mll2, and Set1a and regulates H3K4 methylation at promoters and enhancers. *Epigenetics & Chromatin.* 9, 37. (10.1186/s13072-016-0087-z).
146. Bomber ML, Wang J, Liu Q, Barnett KR, Layden HM, Hodges E, Stengel KR, Hiebert SW. 2023 human SMARCA5 is continuously required to maintain nucleosome spacing. *Mol Cell.* 83, 507–22. e506.
147. Ramani V, Deng X, Qiu R, Gunderson KL, Steemers FJ, Disteche CM, Noble WS, Duan Z, Shendure J. 2017 massively multiplex single-cell Hi-C. *Nat Methods.* 14, 263–6.
148. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet.* 2013;14:390–403.
149. Mahat DB, Kwak H, Booth GT, Jonkers IH, Danko CG, Patel RK, Waters CT, Munson K, Core LJ, Lis JT. 2016 base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). *Nat Protoc.* 11, 1455–76.
150. Heath H, De Almeida CR, Sleutels F, Dingjan G, Van De Nobelen S, Jonkers I, Ling KW, Gribnau J, Renkawitz R, Grosveld F. 2008 CTCF regulates cell cycle progression of $\alpha\beta$ T cells in the thymus. *EMBO J.* 27, 2839–50.
151. Zhao H, Yang Y, Rizo CM, Overbeek PA, Robinson ML. 2004 insertion of a Pax6 consensus binding site into the α -crystallin promoter acts as a lens epithelial cell enhancer in transgenic mice. *Investig Ophthalmol Vis Sci.* 45, 1930–9.
152. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. 2016 Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst.* 3, 99–101. (10.1016/j.cels.2015.07.012).
153. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D et al. 2016 Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733-745. (10.1093/nar/gkv1189).
154. Li D, Hsu S, Purushotham D, Sears RL, Wang T. 2019 WashU Epigenome Browser update 2019. *Nucleic Acids Res.* 47, W158-w165. (10.1093/nar/gkz348).
155. Li H. 2013 Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.
156. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W. 2008 Model-based analysis of ChIP-Seq (MACS). *Genome biology.* 9, 1–9.
157. Ramirez F, Dündar F, Diehl S, Grüning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 2014;42:W187–91.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.