Bidirectional Analysis of Cryba4-Crybb1 Nascent Transcription and Nuclear Accumulation of Crybb3 mRNAs in Lens Fibers

Saima Limi,[1] Yilin Zhao,[1] Peng Guo,[2] Melissa Lopez-Jones,[2] Deyou Zheng,[1,3,4] Robert H. Singer,[2] Arthur I. Skoultchi,[5] and Ales Cvekl[1,6]

[1]Departments of Genetics, Albert Einstein College of Medicine, Bronx, New York, United States
[2]Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, New York, United States
[3]Neurology, Albert Einstein College of Medicine, Bronx, New York, United States
[4]Neuroscience, Albert Einstein College of Medicine, Bronx, New York, United States
[5]Cell Biology, Albert Einstein College of Medicine, Bronx, New York, United States
[6]Ophthalmology and Visual Sciences, Albert Einstein College of Medicine, Bronx, New York, United States

Correspondence: Ales Cvekl, Ophthalmology and Visual Sciences, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Ullmann 123, Bronx, New York, NY 10461, USA;
ales.cvekl@einstein.yu.edu.

Submitted: October 9, 2018
Accepted: November 26, 2018

Citation: Limi S, Zhao Y, Guo P, et al. Bidirectional analysis of Cryba4-Crybb1 nascent transcription and nuclear accumulation of Crybb3 mRNAs in lens fibers. *Invest Ophthalmol Vis Sci*. 2019;60:234–244. https://doi.org/10.1167/iovs.18-25921

Purpose. Crystallin gene expression during lens fiber cell differentiation is tightly spatially and temporally regulated. A significant fraction of mammalian genes is transcribed from adjacent promoters in opposite directions ("bidirectional" promoters). It is not known whether two proximal genes located on the same allele are simultaneously transcribed.

Methods. Mouse lens transcriptome was analyzed for paired genes whose transcriptional start sites are separated by less than 5 kbp to identify coexpressed bidirectional promoter gene pairs. To probe these transcriptional mechanisms, nascent transcription of *Cryba4*, *Crybb1*, and *Crybb3* genes from gene-rich part of chromosome 5 was visualized by RNA fluorescent in situ hybridizations (RNA FISH) in individual lens fiber cell nuclei.

Results. Genome-wide lens transcriptome analysis by RNA-seq revealed that the *Cryba4-Crybb1* pair has the highest Pearson correlation coefficient between their steady-state mRNA levels. Analysis of *Cryba4* and *Crybb1* nascent transcription revealed frequent simultaneous expression of both genes from the same allele. Nascent Crybb3 transcript visualization in "early" but not "late" differentiating lens fibers show nuclear accumulation of the spliced Crybb3 transcripts that was not affected in abnormal lens fiber cell nuclei depleted of chromatin remodeling enzyme Snf2h (Smarca5).

Conclusions. The current study shows for the first time that two highly expressed lens crystallin genes, *Cryba4* and *Crybb1*, can be simultaneously transcribed from adjacent bidirectional promoters and do not show nuclear accumulation. In contrast, spliced Crybb3 mRNAs transiently accumulate in early lens fiber cell nuclei. The gene pairs coexpressed during lens development showed significant enrichment in human "cataract" phenotype.

Keywords: crystallin, differentiation, denucleation, head-to-head genes, lens, nucleus, RNA FISH, splicing

Cellular differentiation during embryonic development is marked by temporally and spatially regulated gene expression. Genes are located at different densities across the chromosomes, including gene-rich regions and gene deserts. Thus, multiple transcriptional events can occur in parallel, including transcriptions in opposite and parallel directions. In the human genome, bidirectional promoter gene pairs have been found to make up approximately 11% of the genome. There are potentially shared *cis*-elements between the promoters of these bidirectional promoter gene pairs and these shared elements are thought to regulate the gene expression by acting on their promoters.[1] Furthermore, recordings of transcriptional burst signals by live cell imaging has shown that regulation of bidirectional promoters are likely controlled by transcriptional hubs or cases where a single enhancer can control two promoters simultaneously.[2]

The ocular lens is an advantageous system to study transcription and cellular differentiation. Lens is a transparent tissue with two cellular compartments. The majority of the lens is composed of lens fiber cells with an overlying monolayer of cuboidal lens epithelial cells. Lens fiber cells differentiate from postmitotic lens cell precursors through a series of steps, including cell elongation, cell cycle exit by expression of cyclin-dependent kinase inhibitors Cdkn1b/p27 and Cdkn1c/p57, synthesis and accumulation of crystallin proteins, degradation of subcellular organelles, and other tissue morphogenetic and remodeling processes.[3–6]

Temporally and spatially regulated crystallin gene expression is the central pathway of the lens differentiation cascade. Lens crystallins represent approximately 90% of lens water-soluble proteins and are divided into evolutionarily conserved vertebrate α-, β-, and γ-crystallin families. A pair of αA- and αB-crystallins evolved from small heat shock proteins[7] and exhibit

protein chaperone-like and antiapoptotic activities in lens[8,9] and nonlens tissues.[10,11] The β/γ-crystallins evolved from an ancestral calcium binding protein,[12] through a common βγ-crystallin precursor,[13] and further separated into the β- and γ-crystallin subfamilies.[6,14,15] During lens fiber cell differentiation, there is a temporal and spatial regulation in the expression of individual α-, β-, and γ-crystallin genes; however, it is not precisely known how these processes are regulated at the transcriptional, posttranscriptional, translational, and posttranslational levels to achieve the desired individual crystallin protein composition and distribution within the mature lens fibers.

The β-crystallin subfamily is divided into two groups, including βA1/A3-, βA2-, and βA4-, and βB1-, βB2-, and βB3-crystallins.[14,15] Mutations in β-crystallin genes cause congenital cataracts[16] and other visual disorders.[17–19] The majority of these mutations include missense[20–24] and nonsense mutations.[25,26] These six mouse β-crystallin genes are located on chromosomes 1, 5, and 11. On chromosome 5, four β-crystallin genes are organized as two pairs, *Cryba4-Crybb1* and *Crybb2-Crybb3*, separated by over 0.8-Mbp region containing at least 12 coding and noncoding genes. The Cryba4, Crybb1, and Crybb3 mRNAs are expressed in the embryonic lens while Crybb2 mRNAs are abundantly expressed in the postnatal lenses.[27] Its initial "low" levels are detected by RNA-seq in E18.5 embryonic lenses.[28]

The β-crystallin gene cluster can be used to probe multiple mechanistic details of nascent transcription as the Cryba4 and Crybb1 transcriptional start sites are separated by only 3301 bp of DNA and are transcribed in the opposite directions as a "bidirectional" system.[1] Several concepts have been proposed for the existence of bidirectional promoter gene pairs, such as coordinate expression of histone genes to maintain stoichiometric relationship, coexpression of genes functioning in the same biological pathway to control expression of cell cycle genes through different time points, and synchronized heat shock responses.[1]

In differentiating lens fibers, transcription is restricted by the denucleation process imposing additional limitations on the efficiency of the transcription prior to nuclear degradation.[29] Recent studies of nascent transcription in multiple model cells and tissues revealed that the transcriptional process occurs in bursts[29–32] and is executed by convoys of 20 to 30 molecules of RNA polymerase.[33] Indeed, ChIP-seq data identified RNA polymerase II all across the active crystallin loci in lens chromatin, including their 3′-UTRs.[34] Concomitant with transcription, the mRNAs are spliced through a series of events at or near the site of transcription. Spliced mRNAs are transported out of the nucleus through the nuclear pore complexes into the cytoplasm for translation. Retention of transcripts within the nucleus for prolonged periods of time is rare.[35] Very little is known about the entire life course of mRNAs both in the nucleus and cytoplasm in any region of the lens.

Using bioinformatics and RNA-seq lens data, we first analyzed mouse lens transcriptomes for bidirectional promoter gene pairs and focused our experiments on head-to-head *Cryba4-Crybb1* gene pair to examine simultaneous expression of these genes from a single allele by RNA fluorescent in situ hybridizations (RNA FISH). In contrast, in the *Crybb2-Crybb3* pair, only the Crybb3 mRNAs are abundant in early embryonic stages. Our studies show nuclear accumulation of its spliced mRNAs and this accumulation is temporally and spatially regulated in differentiating lens fibers. Depletion of ATP-dependent chromatin remodeling enzyme Snf2h (Smarca5) from the lens does not affect the temporary nuclear accumulation of Crybb3 mRNAs. In summary, these studies provide novel insights into the complex dynamics of crystallin gene transcription and splicing in the differentiating lens fibers.

Materials and Methods

Mice and Tissue

Animal husbandry and experiments were conducted in accordance with the approved protocol of the Albert Einstein College of Medicine Animal Institute Committee and the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research. Noon of the day vaginal plug was examined and was considered as E0.5 of embryogenesis. Individual lenses were harvested from E12.5, E14.5, and newborn (P0.5) FVB wild type strain mice. Animals were euthanized by CO_2 and mouse embryos were dissected from pregnant females. In some cases, whole eyeballs were removed from the postnatal animals. Tissues were then fixed in 4% paraformaldehyde at 4°C, submerged in 30% sucrose overnight at 4°C, and embedded in optimal cutting temperature (OCT). Serial sections were cut in 7-μm thickness through the midsection of the lens then used for hybridizations. A procedure to generate floxed Snf2h model and conditional inactivation in the lens are described elsewhere.[36–38] For all embryonic mouse studies the embryos were removed from the womb, flushed, and fixed in 4% paraformaldehyde overnight at 4°C, incubated overnight with 30% sucrose at 4°C, and then embedded in OCT medium before cryosectioning. All animals were euthanized at approximately the same time of day to remove any circadian rhythm influence on transcriptional activity.

Computational and Bioinformatics Studies

The Refseq gene annotation for the mm10 version of the mouse genome was downloaded from the UCSC genome browser in June 2018. Only known coding transcripts ("NM") were used for the current study. In total, 20,591 adjacent gene pairs on the same chromosome were identified. The shortest distances between the transcriptional start sites (TSSs) of the gene pairs were then calculated and 1692 pairs had distances within 5 kb. Among them, 1643 gene pairs were detected by lens epithelium and fiber RNA-seq data from embryonic E14.5 to newborn P0.5 lenses.[28] The Pearson's correlation coefficients (r) of the mean fragments per kilobase of transcript per million mapped (FPKMs) reads from E14.5 to P0.5 epithelium and fiber of each gene pair were calculated, with P values corrected by Benjamini-Hochberg method for multiple testing (Supplementary Table S1). The significantly coexpressed gene pairs were selected by absolute value of $r > 0.9$ and adjusted P value < 0.05. The functional analysis of the coexpressed and not coexpressed genes was conducted with the ToppGene.[39]

RNA FISH Hybridizations and Imaging

A procedure to generate probes for detection of nascent transcripts by RNA FISH and for RNA FISH and hybridization and imaging in the lens is described elsewhere.[29] Briefly, a probe library consisting of 12 to 48 probes of 20-bp length was constructed against the exon regions of the genes of interest. For the Crybb3 intronic probes, probes recognizing the first intron were designed and used. Hybridization was conducted overnight with fluorescently labelled probes using Quasar 570 and Quasar 670 fluorophores. Tissue sections were placed in a 1× decloaking buffer (10× Reveal Decloaking buffer; Biocare Medical, Pacheco, CA, USA) used as an antigen retrieval agent and underwent a series of heat and pressure treatments in a decloaker. The sections were then washed and processed through successive treatments with ammonia, sodium borohy-

Investigative Ophthalmology & Visual Science

dride, and magnesium chloride to reduce autofluorescence before undergoing overnight hybridization. After a brief wash with PBS, the slides were mounted with Antifade (Molecular Probes, Eugene, OR, USA) and nuclei were counterstained with 4′,6-diamidino-2-phenylindole (DAPI). Images were taken with Carl Zeiss (Pleasanton, CA, USA) inverted fluorescence microscope equipped with ×100 and ×63 oil-immersion objectives. Quantification was done on stacks of 41 optical sections with Z spacing of 0.2 μm.[40]

Imaging and Image Analysis

A procedure for imaging, image processing, and image analysis of RNA FISH data in the lens tissue is described elsewhere.[29] Briefly, three-dimensional image data were acquired using the Zeiss Axio Observer CLEM microscope (Carl Zeiss). Slides were evaluated at ×60 for the distribution of the fluorescence signal (nascent transcription sites). All 41-image z-stacks were compressed into one image using Maximum Projection in ImageJ software (http://imagej.nih.gov/ij/; provided in the public domain by the National Institutes of Health, Bethesda, MD, USA). Only sites within nuclei were counted. The number of nuclei in each field examined were also counted, including those within designated regions where transcription sites were not detected. Transcription sites were detected with Volocity software (PerkinElmer, University of Warwick Science Park, Coventry, CV4 7HS) that analyzed merged images with the specific probes against the gene of interest and DAPI for detection of nuclei. Transcription sites were automatically detected only inside the nuclei using Volocity software. Nuclear segmentation was carried out manually on a maximal projection of the DAPI channel. Images were then deconvoluted in Volocity prior to any measurements. Thresholding based on signal intensity, which uses Otsu's method on the histogram of intensities in the image, was used to separate signal from background. The 'mean pixel intensity' was used for quantifying and graphing the signal intensity values. The data collected in the DAPI channel provided the boundaries of nuclei in the specimen. *P* value of burst fraction was calculated using Student's *t*-test and burst intensity was calculated using Wilcoxon rank-sum test.

Image Analysis in the Lens: Four Regions of Progressive Lens Fiber Cell Differentiation

In order to differentiate between the progressive stages of lens fiber cell differentiation in the newborn lens, we divided the lens tissue into four geometrically equal sections from the periphery to the center of the lens starting from both the left and right sides of the lens, labeled a, b, c, and d, respectively.[29] The left and right sides of the lens tissue were treated symmetrically. Lens tissue area a represents early differentiating lens fiber cells, area b represents intermediate lens fiber cells, area c represents advanced lens fiber cells, and area d represents terminally pre- and denucleated organelle free zone (for details, see Fig. 3B).[29]

Results

Analysis of Coexpression Patterns of Paired Genes During Lens Development

To investigate the coexpression patterns of adjacent coding genes during mouse lens development, we first collected 20,591 pairs of adjacent coding genes, including 4943 "head-to-head," 4944 "tail-to-tail," and 10,704 "head-to-tail" pairs (Fig. 1A). As expected, the shortest distances between the TSSs of the gene pairs showed a bimodal distribution (Fig. 1B) and among the three types of gene pairs, "head-to-head" pairs had significantly shortest distances (P values $< 2.2^{e-16}$ by Mann-Whitney U tests, Fig. 1C). We further examined the 1692 gene pairs with distances less than 5 kb based on the first quantile of "head-to-head" gene pairs (Fig. 1D and Supplementary Table S2). The Pearson correlation coefficients were then calculated based on mean FPKMs of the gene pairs in embryonic lens development for the 1643 gene pairs detected in the RNA-seq data.[28] Among them, 296 gene pairs were found significantly coexpressed during lens development, by r values more than 0.9 or less than −0.9 and adjusted $P < 0.05$ (256 positively coexpressed and 40 negatively coexpressed pairs; Supplementary Table S3). This analysis revealed that the *Cryba4-Crybb1* pair was the top coexpressed gene pair followed by *Col4a4-Col4a3*, *Cpeb3-March5*, *Atg2b-Gskip*, and *Col4a6-Col4a5* (Fig. 1E).

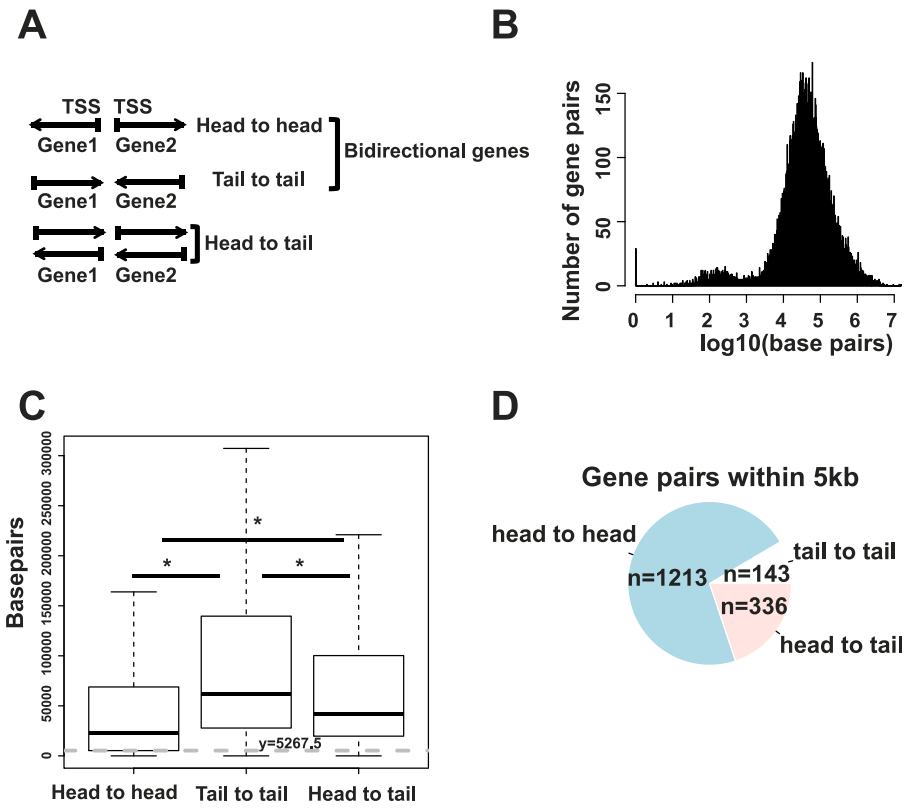
Exploration of Underlying Contribution Factors and Functional Classification of Coexpressed Genes Pairs During Lens Development

As expected, the percentage of coexpressed genes for gene pairs with distances less than 5 kb is much higher than the gene pairs with distances more than 5 kb (18% vs. 11%). However, we found that the coexpression of gene pairs within 5 kb was not correlated to distances between TSSs or expression levels of genes (Figs. 2A, 2B, all P values > 0.05 and abs(r) < 0.1). We further examined the 296 significantly coexpressed gene pairs. As expected, 30 of 56 genes of histone cluster 1 family members located on the mouse chromosome 13 were coexpressed throughout lens development. Using ToppGene, we analyzed the functional enrichment of the coexpressed and non-coexpressed gene pairs after excluding the histone genes. Comparing the coexpressed genes versus non-coexpressed ones, we found that "collagen–activated tyrosine kinase receptor signaling pathway," "fatty acid beta–oxidation using acyl–CoA dehydrogenase," and "homophilic cell adhesion via plasma membrane adhesion molecules" gene ontology (GO) terms were uniquely enriched in the coexpressed genes. Interestingly, lens-associated human phenotypes, including "anterior lenticonus," "abnormality of lens," and "cataract," were only enriched in coexpressed gene pairs. Thirty-six of 569 genes (unique genes collected from 296 gene pairs) were characterized as cataract genes, accounting for 6% of all coexpressed genes (Supplementary Table S4). Given high abundance of both Cryba4 and Crybb1 mRNAs in embryonic lens fiber cells and opposite direction of their transcription, analysis of this system offers an opportunity to study nascent transcription from two adjacent alleles on the same chromosome.

Nascent Transcription of *Cryba4* and *Crybb1* Genes

Two mouse β-crystallin gene pairs and their transcriptional orientations are schematically shown in Figure 3A. Transcription of *Cryba4* and *Crybb1* genes occurs in opposite directions. In contrast, transcription of *Crybb2* and *Crybb3* is in the same orientation (Fig. 4). Different regions of lens fiber cell compartments are shown in E14.5 and P0.5 lenses in Figure 3B as established elsewhere.[29] Nascent Cryba4 and Crybb1 mRNA transcripts were thus simultaneously detected by two sets of Quasar 570 (red) and Quasar 670 (green) conjugated oligonucleotide probes to detect exons in individual lens fiber cell nuclei as described in Materials and Methods. A single Cryba4-Crybb1 allele generates two proximal hybrid-

Investigative Ophthalmology & Visual Science



E

Gene1	Gene2	Distance	Type	cc	padj
Cryba4	Crybb1	3301	head to head	0.99	1.01E-06
Col4a4	Col4a3	71	head to head	0.99	3.06E-06
Cpeb3	March5	73	head to head	0.99	1.22E-05
Atg2b	Gskip	110	head to head	0.99	1.91E-05
Col4a6	Col4a5	1342	head to head	0.99	1.91E-05

Figure 1. Analysis of expression of paired genes in mouse lens. (**A**) Organization of head-to-head, tail-to-tail, and head-to-tail genes. (**B**) Distribution of distances between paired genes. (**C**) Distribution of distances of head-to-tail, tail-to-tail, and head-to-tail gene pairs. (**D**) Total numbers of three types of paired genes within 5 kb of genomic DNA. (**E**) Top five head-to-head genes with the highest correlation of their mRNAs.

ization signals ("two foci") if both adjacent genes are expressed at the same time. Analysis of serial sections revealed five types of foci distributions within each individual nucleus; no nascent mRNA expression, expression of Cryba4 alone, expression of Crybb1 alone, adjacent expression of both Cryba4 and Crybb1, and "distal" expression of both Cryba4 and Crybb1 mRNAs originating from different alleles of chromosome 5 (Fig. 3C). Quantification of 479 nuclei revealed that approximately 40% and 56% of the nuclei within the whole newborn mouse lens generates nascent Cryba4 and Crybb1 mRNAs, respectively (Fig. 3D). We found nascent transcription of both *Cryba4* and *Crybb1* genes, no transcription of both genes, transcription of only Cryba4, and only Crybb1 in 28.8%, 33.8%, 11.6%, and 27.5% of nuclei, respectively (Fig. 3D). In the group of nascent coexpressed Cryba4 and Crybb1 mRNAs, approximately 14% of the nuclei from the whole lens show Crybb1 and Cryba4 co-localization with overlapping signal either from one or both alleles indicating simultaneous transcription in the opposite directions (Fig. 3D). There is an almost equal percentage of cells within newborn lens that transcribe both Cryba4 and Crybb1 (28.8%) versus those that transcribe neither Cryba4 or Crybb1 (33.8%). These data suggest there is no strong preference toward cells either transcribing both Cryba4 and Crybb1 or toward cells transcribing neither Cryba4 or Crybb1. However, the number of cells transcribing *Crybb1* solely is much higher (27.5%) than those transcribing only the *Cryba4* gene (11.5%) indicating that the lens fiber nuclei generate more Crybb1 than Cryba4 mRNAs. On average, Crybb1 appears to be "stronger" in recruitment of the transcription machinery compared with Cryba4. From these data, we conclude that that this head-to-head gene pair does not possess any limitation to prevent simultaneous nascent transcription in the opposite directions.

Nascent Transcription of the Mouse *Crybb3* Gene

The gene-rich region of mouse chromosome 5 contains another *Crybb2-Crybb3* gene pair (Figs. 3A, 4). However, transcriptions of these genes occur in the same direction (Fig. 3A). Expression of Crybb2 is detectable in lens fibers from E18.5 while expression of Crybb3 is found in both lens compartments from E12.5 (Fig. 5). Originally, we intended to perform simultaneous visualization of four nascent Cryba4,

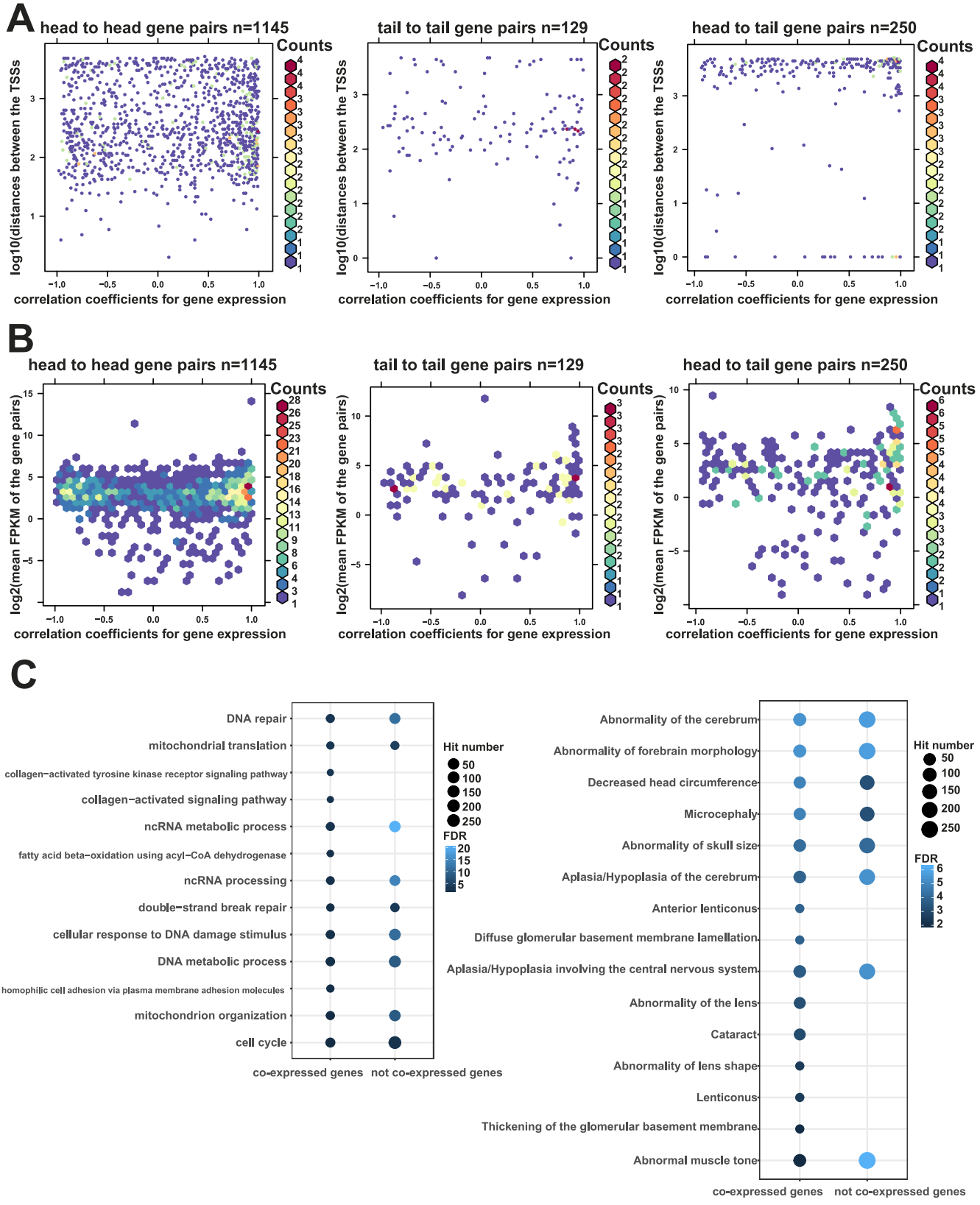


Figure 2. Distances between TSSs, expression levels of genes, or transcription direction patterns do not correlate with gene coexpression. (**A**) The hexbin plots show log10(distances between the TSSs) together with Pearson correlation coefficients (*x* axis) for gene expression for three groups of gene pairs. (**B**) The hexbin plots show log2(mean FPKM of the gene pairs) together with Pearson correlation coefficients (*x* axis) for gene expression for three groups of gene pairs. (**C**) The dot plots show GO term and human phenotype enriched at coexpressed and non-coexpressed gene pairs.

Investigative Ophthalmology & Visual Science

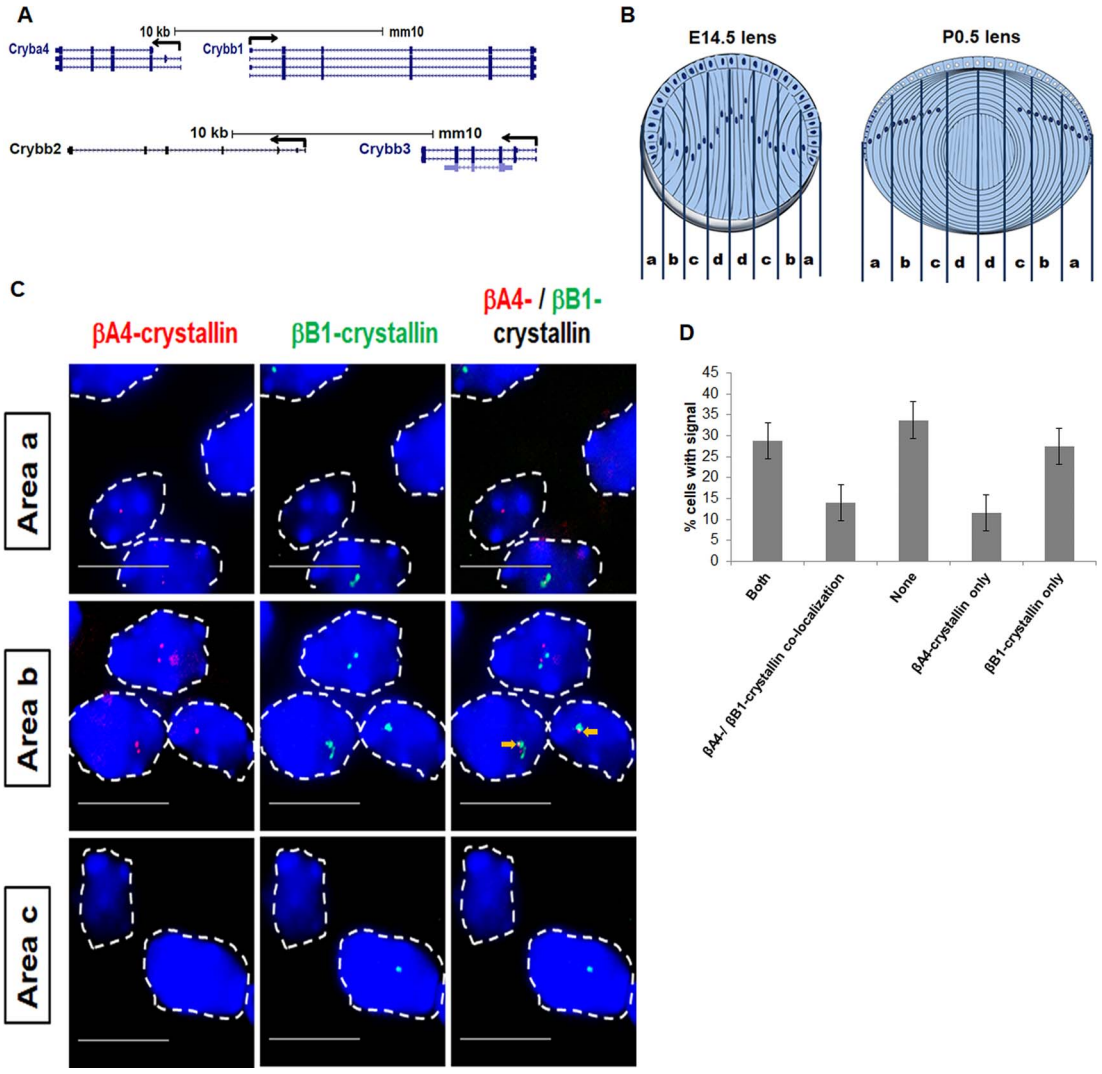


Figure 3. Nascent transcription of Cryba4 and Crybb1 in differentiating lens fibers of newborn lens. (**A**) Schematic diagram of two mouse β-crystallin gene pairs on chromosome 5 showing bidirectional promoter arrangement of the *Cryba4* and *Crybb1* gene pair and same transcriptional direction of the *Crybb2* and *Crybb3* gene pair. (**B**) Schematic diagram of E14.5 and P0.5 lens to indicate different regions of the lens fiber cells: areas a, b, c, d (see Materials and Methods, Limi et al.[29]). (**C**) Visualization of Cryba4 and Crybb1 nascent mRNA coexpression in individual lens fiber cell nuclei. Regions of colocalization of the Cryba4 and Crybb1 nascent transcripts originating from the same allele are indicated by *yellow arrows*. (**D**) Quantitative analysis of nascent transcription of *Cryba4* and *Crybb1* genes showing percentage of nuclei expressing both genes, both genes from the same allele, *Cryba4* alone, *Crybb1* alone, or none of the two transcripts.

Crybb1, Crybb2, and Crybb3 mRNAs to further probe their coexpression dynamics; however, we found additional abundant Crybb3 mRNA nuclear foci (Fig. 5). To resolve these unexpected findings, we designed a pair of individual RNA FISH probe libraries recognizing either the introns or exons of the Crybb3 mRNAs conjugated to two spectrally resolvable fluorophores, Quasar 570 (red) and Quasar 670 (green), respectively. Previous studies in liver and red blood cells have shown that co-localization of signals recognized by these two distinct probe sets corresponds to the nascent transcripts.[30,31] The earliest nascent βB3-crystallin mRNA transcription in the lens was examined in mouse E12.5 embryos in early differentiating primary lens fibers; however, multiple hybridization signals ($n > 2$) specific for the exons were also detected in individual nuclei (Fig. 5). Co-localization of both intron and exon signals generated one or two yellow "foci" per nuclei representing one or two actively transcribed alleles while nuclei without intron/exon co-localization did not produce any double signals observed as yellow dots/foci (Fig. 5). Additional exon-specific signals (Quasar 670, green) are detected within the cluster of nuclei in early differentiating lens fibers suggesting that spliced primary transcripts are accumulated inside of the nuclei at multiple locations. Thus, we achieved visualization of nascent transcription of the βB3-crystallin gene using intronic probes and demonstrate nuclear retention/accumulation using exonic probes recognizing the spliced mRNA transcripts in the lens fiber cells but not in the lens epithelium.

To follow these findings, we next analyzed more advanced E14.5 and P0.5 differentiating lens fibers. In E14.5 lenses, the RNA FISH data resembled those identified at E12.5 (Fig. 6). Accumulations of spliced Crybb3 mRNAs were found both in early (area a) and advanced lens fibers (area c). In contrast, in newborn mouse lens, the advanced lens fibers (area c, P0.5, Fig. 6) that originally expressed Crybb3 (Fig. 5, E12.5 stage and Fig. 6, E14.5 stage) no longer show any signal indicating that the accumulated mRNAs were exported from the nuclei and nascent transcription was attenuated. We propose that

Investigative Ophthalmology & Visual Science

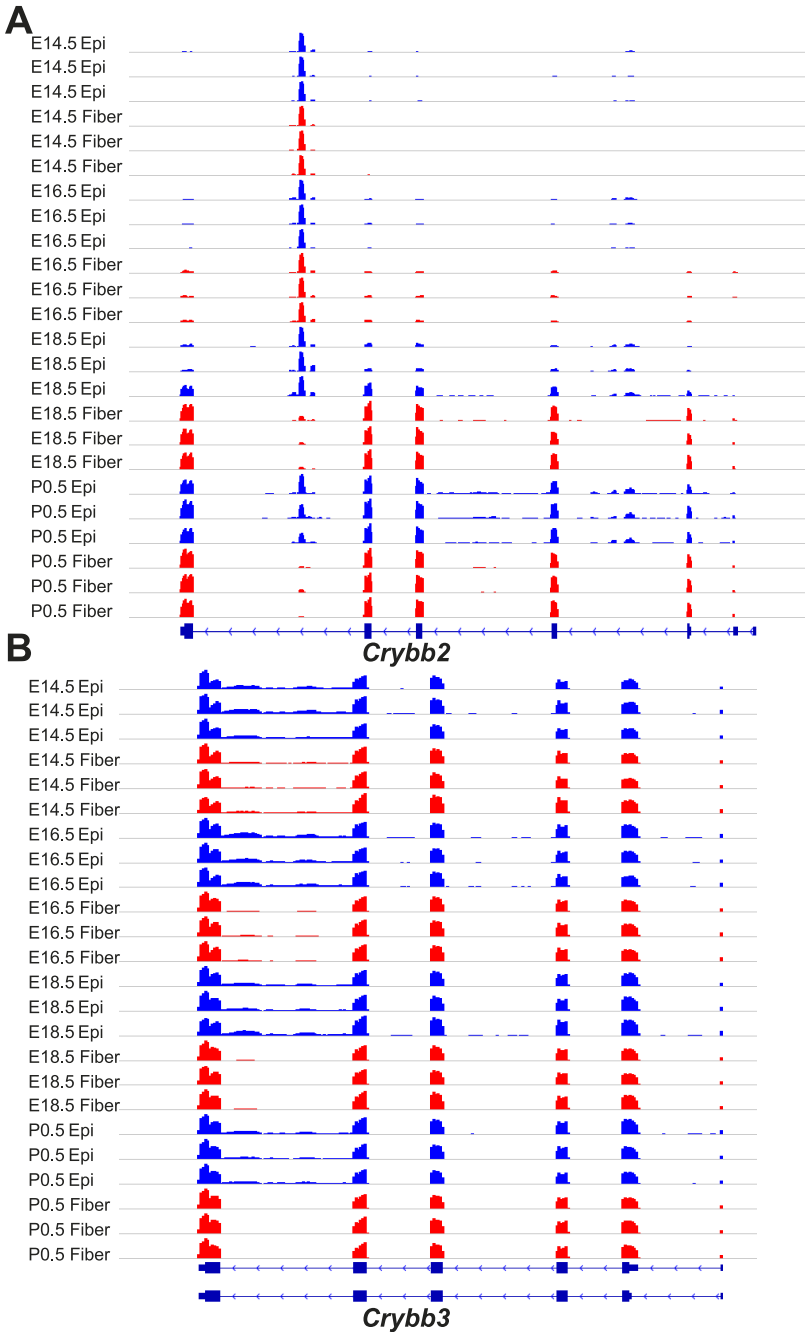


Figure 4. Distribution of RNA-seq reads at the (**A**) βB2- and (**B**) βB3-crystallin gene expression in lens epithelium (*blue*) and lens fibers (*red*). Results of triplicate experiments are shown in Zhao et al.[28]

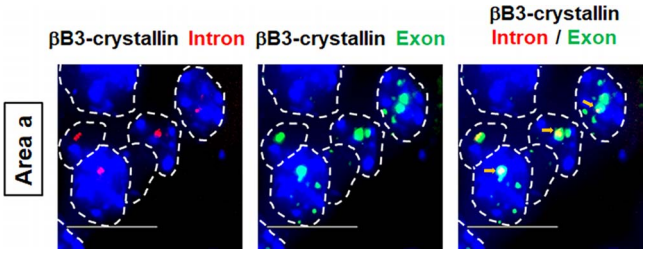


Figure 5. RNA FISH analysis of βB3-crystallin nascent transcription using exon- and intron-specific probes first detected in E12.5 embryonic mouse lens. Sites of exon/intron co-localizations of nascent primary Crybb3 transcripts are indicated by *yellow arrows*.

developmentally controlled mRNA nascent transcription of βB3-crystallin differs significantly from other crystallin genes examined here (i.e., Cryba4 and Crybb1) and elsewhere (i.e., Cryaa, Crybb1, and Cryga).[29]

To dissect the developmental difference in nascent transcription and nuclear retention, the transcription burst fraction and burst intensity of the βB3-crystallin intronic/unspliced and exonic/spliced mRNAs transcripts were measured between the two developmental stages, E14.5 and newborn P0.5 lens fiber cells (Fig. 7). The transcription burst fraction was measured as "percent cells showing signal," which refers to the percentage of nuclei within the entire tissue actively transcribing the gene (nascent transcripts) or retaining the spliced mRNA transcripts

Investigative Ophthalmology & Visual Science

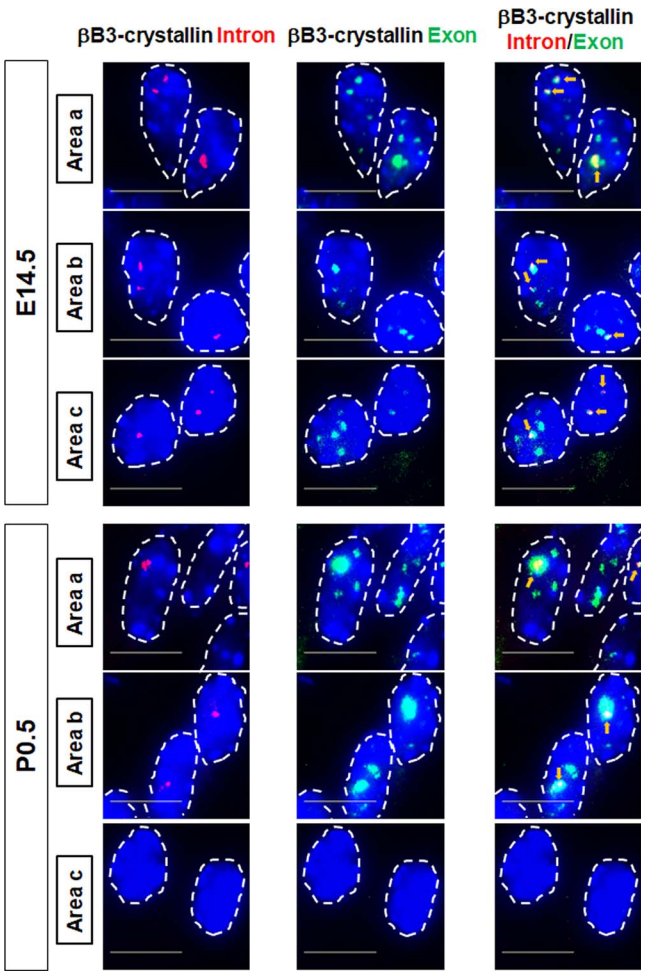


Figure 6. Transient accumulation of spliced βB3-crystallin mRNAs in the lens fiber cell nuclei in E14.5 lens and in P0.5 lens from periphery to the center of the lens tissue. The progressive stages of differentiation of the lens cells are shown by different regions of the tissue indicated by areas a–c, from periphery to the center of the lens, respectively (see Fig. 3B). *Yellow arrows* indicate areas of exon/intron co-localization.

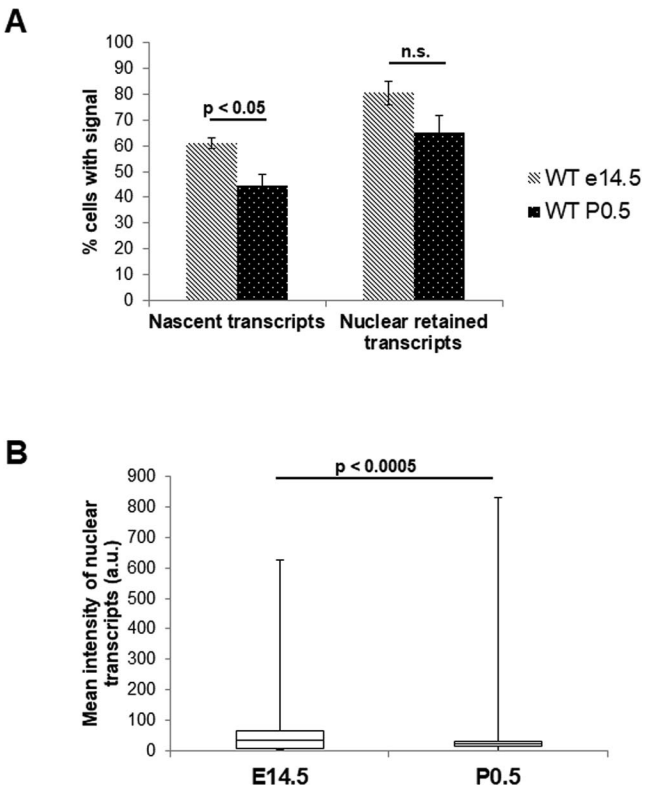


Figure 7. Quantification of transcriptional burst fraction and intensity of the βB3-crystallin gene in E14.5 versus P0.5 lens. (**A**) Burst fraction measured as percent cells within the entire lens tissue showing either intronic signal (nascent transcript) or exonic signal (nuclear retained transcripts). (**B**) Burst intensity measured as signal intensity in arbitrary units (a.u.) of the exonic signals (nuclear retained transcripts) within the nuclei of the entire lens tissue. Significance value shown by *P* value and nonsignificance is indicated by n.s., ($n = 3$ biological replicates).

(nuclear retained transcripts). There was a significant difference in the percentage of cells actively transcribing the transcripts between the two developmental stages (Fig. 7A). A significantly higher percentage of cells generate the βB3-crystallin nascent mRNAs in the embryonic E14.5 stage compared with the newborn P0.5 stage (Fig. 7A). However, the percentage of cells that retain the spliced βB3-crystallin mRNAs within the nucleus (nuclear retained transcripts) showed no significant difference between the E14.5 and P0.5 stages of development (Fig. 7A). In both the E14.5 and P0.5 stages of development examined, although there is slightly higher percentage of cells with nuclear retention of the βB3-crystallin transcripts than those actively transcribing the transcript, the differences were not found significant (Fig. 7A).

The difference in signal intensity of the primary spliced/nuclear retained βB3-crystallin mRNAs was significant between E14.5 and P0.5 lenses, with a higher signal intensity in the earlier E14.5 stage compared with the newborn stage (Fig. 7B). Embryonic E14.5 lens thus shows a higher amount of nuclear primary spliced transcripts that is retained compared with P0.5. The βB3-crystallin is transcribed by cells all across the lens fiber cell compartment in the early E14.5 developmental stage and the number of βB3-crystallin transcripts retained within the nuclei is higher than that of in P0.5 differentiated lens. These findings support a model proposing that at earlier stages βB3-crystallin is being produced for storage rather than utilization, and at later stages Crybb3 mRNAs are fully used by the translational systems.

Nascent transcription of three crystallin genes, including *Cryaa*, *Crybb1*, and *Cryga*, is disrupted in lenses that are depleted of the multifunctional ATP-dependent chromatin remodeling enzyme Snf2h (Smarca5).[29] This enzyme regulates both transcription and DNA repair.[41] In addition, Snf2h-null lenses retain their nuclei.[37] Thus, it is important to determine if these abnormal nuclei are capable of expressing Crybb3 and whether there are any abnormalities related to the accumulation and delayed export of spliced Crybb3 mRNAs. The βB3-crystallin intronic (nascent transcription) and exonic (nuclear retention) transcripts are shown in Snf2h-null newborn lens fiber cells (Fig. 8) and the control lens fiber nuclei are shown in Figure 6 from wild-type (WT) newborn P0.5 lens. The mutated nuclei show similar patterns of foci that illuminate nascent sites of transcription as well as nuclear accumulation of spliced Crybb3 mRNAs. There are also nuclei devoid of any signal indicating that depletion of Snf2h does not directly or indirectly influence either process.

To further examine a mechanistic link between chromatin remodeling on nascent transcription and nuclear retention the transcription, burst fraction and burst intensity of the βB3-crystallin intronic (nascent transcripts) and exonic transcripts (nuclear retained transcripts) were measured between WT and Sn2h-null newborn (P0.5) lens fiber cells (Figs. 8, 9).

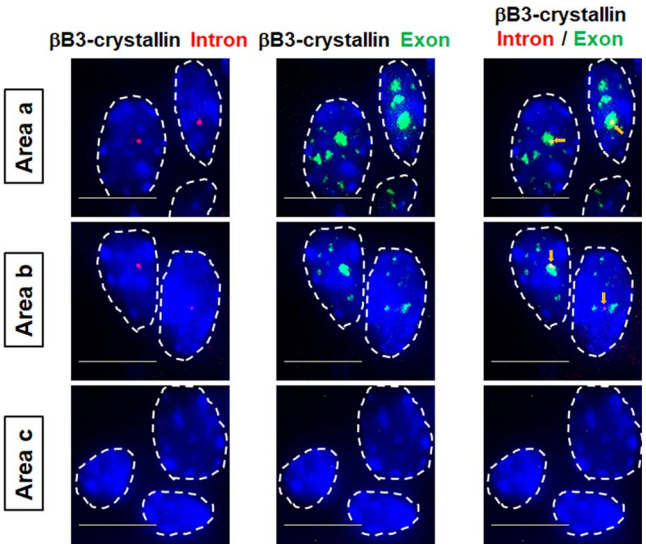


FIGURE 8. Transient accumulation of spliced βB3-crystallin mRNAs in newborn lens fiber cell nuclei following depletion of Snf2h. RNA-FISH images of βB3-crystallin nascent transcripts within different areas of the Snf2h-null newborn lens tissue from periphery to the center of the lens indicated by areas a–c, respectively. *Yellow arrows* indicate co-localization of the βB3-crystallin intronic and exonic signals.

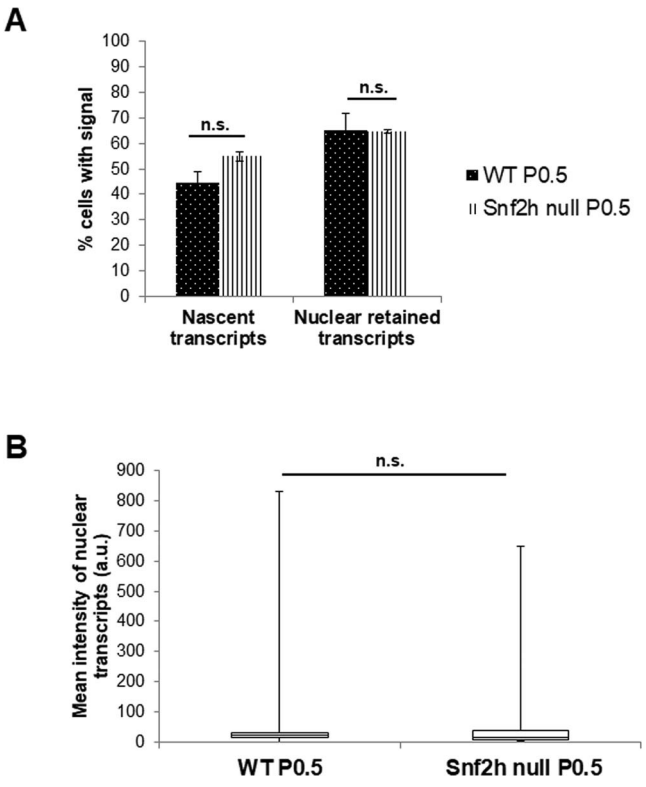


FIGURE 9. Analysis of transcriptional bursting parameters of βB3-crystallin gene in WT versus Snf2h-null newborn (P0.5) mouse lenses. (**A**) Transcriptional burst fraction of βB3-crystallin transcripts measured as percent cells within the entire lens tissue showing either intronic signal (nascent transcript) or exonic signal (nuclear retained transcripts). (**B**) Transcriptional burst intensity of βB3-crystallin transcripts measured as signal intensity of exonic signal (nuclear retained transcripts) in arbitrary units (a.u.) in WT versus Snf2h-null lens. Significance value of n.s. indicates not significant ($n = 3$ biological replicates).

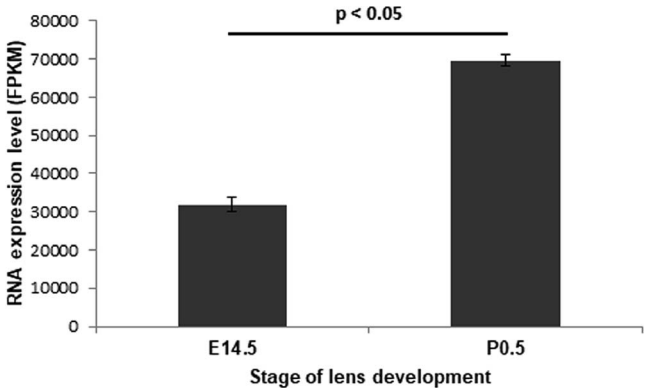


FIGURE 10. Comparative RNA-seq analysis of the βB3-crystallin gene expression in E14.5 versus newborn P0.5 lens fiber cells. Significant upregulation of βB3-crystallin gene expression was found as lens development progresses using primary data we published elsewhere.[28]

Transcription burst fraction of nascent mRNA transcripts encoding the βB3-crystallin showed no significant difference between WT and Snf2h-null newborn lens fiber cells. Likewise, there was no significant difference between WT and Snf2h-null newborn lens fibers in the percentage of cell nuclei that accumulated the spliced exonic βB3-crystallin transcripts (Fig. 9A). When comparing the signal intensity of the nuclear retained βB3-crystallin spliced exonic transcripts between WT and Snf2h-null newborn lens fibers, there was also no significant difference between these lenses (Fig. 9B). This suggests that chromatin remodeling by Snf2h does not affect the number of cell nuclei actively transcribing the βB3-crystallin gene nor does this enzyme play a role in nuclear retention of this gene. Similarly, our previous study has shown that Snf2h also does not affect transcriptional bursting parameters of the other gene from the βB3-crystallin group member, the βB1-crystallin.[29] We conclude that although Snf2h has different specific roles for different genes during transcription depending on the gene being studied it does not play a role in the βB3-crystallin gene burst fraction or nuclear retention.

Finally, the overall steady-state level of βB3-crystallin gene expression was assessed using RNA-seq in E14.5 and P0.5 lens (Fig. 10). The RNA-seq data do not show any sign of intronic transcript expression within the newborn P0.5 lens; the majority of the analyzed transcripts comprise spliced exons (Fig. 4). In E14.5 lens fiber cells, the RNA-seq data show very low levels of primary transcripts, including the last intron (Fig. 4). The steady-state RNA expression level by RNA-seq shows that the overall gene expression level of βB3-crystallin mRNA is significantly higher in the newborn P0.5 lens compared with the E14.5 embryonic lens (Fig. 10). This indicates that the total gene expression level of βB3-crystallin increases as lens development progresses.

DISCUSSION

The goal of the present studies was to explore lens transcriptome to identify individual genes to study general regulatory mechanisms of transcriptional control. In the first part, we identified a pair of *Cryba4-Crybb1* crystallin genes as the most highly correlated head-to-head pair at the steady-state mRNA levels. We then demonstrate that both of these genes can be simultaneously transcribed from the same allele in the opposite directions. In the second part, we found an unexpected transient accumulation of spliced Crybb3 mRNAs in differentiating lens fiber cell nuclei that differs from other

Investigative Ophthalmology & Visual Science

crystallin mRNAs, such as those encoding αA-, βA4-, βB1-, and γA-crystallins.

The bidirectional genes represent a significant portion of coding capacity of the human and mouse genomes.[1] It has been shown previously that in humans, a majority of bidirectional promoter gene pairs are coexpressed to a higher degree than random gene pairs but a few of the pairs can also compete with each other for gene expression. In the case where only one gene is expressed from the bidirectional promoter gene pair, the promoter of one gene may suppress the expression of the other gene due to competition for polymerases, competition for chromatin modifiers, absence of tissue-specific regulatory regions in one gene, or tissue-specific transcription factors.[42] These models of action are not mutually exclusive. In other cases, it has been shown that linked reporter genes with an enhancer sequence inserted between the two genes display coordinated transcription profile by that shared enhancer.[2] Our data show, that at least in case of the *Cryba4-Crybb1* pair, there are no restrictions that prevent parallel formation of transcriptional machineries and their actions as RNA polymerase II convoys in opposite orientations. Although there are differences in the spatial distribution of nascent transcription between the two genes as lens fibers differentiate and the transcriptional burst changes of Crybb1 during lens differentiation has been previously addressed[29] a proportion of the nuclei with overlapping Cryba4-Crybb1 nascent transcription sites still show co-localization.

Transient nuclear accumulation of spliced mRNAs has been reported in a number of systems, including liver, intestine, and pancreas.[35] It has been proposed that nuclear retention of transcripts may function to reduce gene expression noise caused by cell-to-cell variability in gene expression[35] or serve as backup storage in case of physiologic stress in the cellular environment.[43] Using RNA FISH, our earlier studies of αA-, βB1-, and γA-crystallins did not reveal any abundant nuclear accumulation of their mRNAs.[29] In contrast, the present studies show that in early developmental stages of lens fiber differentiation (E12.5 and E14.5), lens fiber cell nuclei show both nascent βB3-crystallin transcripts as well as nuclear accumulation of their spliced transcripts. In lens fiber cell nuclei approaching their physical destruction in the newborn lens, the Crybb3 foci are no longer present, indicating that the previously accumulated mRNAs were completely transferred into the cytoplasm as βB3-crystallin is highly abundant crystallin in lens fibers.[44]

The present data show that depletion of Snf2h and persistence of abnormal nuclei in mutated lens fibers does not disrupt the Crybb3 mRNA synthesis and its transient nuclear localization. Because coordinated crystallin gene expression is critical for lens development it is expected that a core set of regulatory factors, including the chromatin remodeler Snf2h may regulate transcriptional bursting at crystallin genes. We found that while Snf2h regulates transcriptional bursting of αA- and γA-crystallin it has no effect on βB1-crystallin[29] or βB3-crystallin (Fig. 9). Therefore, expression of βB3-crystallin appears to be regulated by other factors and mechanisms. This is evidenced by nuclear accumulation of Crybb3-spliced mRNAs at distinct developmental stages that has not been seen with other crystallin mRNAs. Because most mRNAs are rapidly transcribed and processed at their sites of transcription then transported out of the nucleus into the cytoplasm for translation, the nuclear retention of this particular crystallin mRNAs raises an interesting possibility that this mechanism evolved to delay translation of the corresponding mRNAs to optimize the cytoplasmic production line of different crystallin proteins.

The RNA-binding proteins that control mRNA cycles in the lens remain elusive. Novel insights were also obtained regarding the cytoplasmic control of multiple crystallin genes by the translational initiation factor eIF3h.[45] Recent studies have also shown that posttranscriptional regulation of lens fiber cell RNAs can be disrupted by mutations in multiple genes, including cytoplasmic RNA granule protein Tdrd7,[46] RNA-binding protein caprin 2,[47] and RNA-binding protein Celf1.[48] Our identification of RNA-binding proteins with different abundance in lens epithelium and lens fiber cells provides ample opportunities to initiate research to better understand the transcriptional-translational machineries in the lens.[44]

In conclusion, the present studies shed new light into lens fiber cell differentiation, mechanisms of crystallin gene transcription, and specific nuclear retention of spliced Crybb3 mRNAs. The gene pairs coexpressed during embryonic lens development show significant enrichment in human "cataract" phenotype and can serve as additional models to study coordinated gene expression in lens differentiation and pave the roads for follow-up studies of RNA biology in the lens and formation of cataracts.

Acknowledgments

The authors thank Robert Coleman for critical reading of the manuscript and Adrien Senecal for numerous advices. We also thank Jie Zhao for her help with mouse colonies.

Supported by grants from NIH R01 EY014237 (AC; NIH, Bethesda, MD, USA), U01 EB021236 (RHS; NIH, Bethesda, MD, USA), R01 DK096266 (AIS; Bethesda, MD, USA), and T32 GM007491 (SL; NIH, Bethesda, MD, USA). The core services were partially funded by NCI Cancer Center Support Grant (P30 CA013330; Bethesda, MD, USA).

Disclosure: **S. Li**, None; **Y. Zhao**, None; **P. Guo**, None; **M. Lopez-Jones**, None; **D. Zheng**, None; **R.H. Singer**, None; **A.I. Skoultchi**, None; **A. Cvekl**, None

References

1. Trinklein ND, Aldred SF, Hartman SJ, Schroeder DI, Otillar RP, Myers RM. An abundance of bidirectional promoters in the human genome. *Genome Res*. 2004;14:62–66.
2. Fukaya T, Lim B, Levine M. Enhancer control of transcriptional bursting. *Cell*. 2016;166:358–368.
3. Piatigorsky J. Lens differentiation in vertebrates. A review of cellular and molecular features. *Differentiation*. 1981;19:134–153.
4. Bassnett S. On the mechanism of organelle degradation in the vertebrate lens. *Exp Eye Res*. 2009;88:133–139.
5. Cvekl A, Ashery-Padan R. The cellular and molecular mechanisms of vertebrate lens development. *Development*. 2014;141:4432–4447.
6. Cvekl A, Zhang X. Signaling and gene regulatory networks in mammalian lens development. *Trends Genet*. 2017;33:677–702.
7. Cvekl A, Zhao Y, McGreal R, Xie Q, Gu X, Zheng D. Evolutionary origins of Pax6 control of crystallin genes. *Genome Biol Evol*. 2017;9:2075–2092.
8. Andley UP, Patel HC, Xi JH. The R116C mutation in αA-crystallin diminishes its protective ability against stress-induced lens epithelial cell apoptosis. *J Biol Chem*. 2002;277:10178–10186.
9. Morozov V, Wawrousek EF. Caspase-dependent secondary lens fiber cell disintegration in αA-/αB-crystallin double-knockout mice. *Development*. 2006;133:813–821.

10. Nagaraj RH, Nahomi RB, Mueller NH, Raghavan CT, Ammar DA, Petrash JM. Therapeutic potential of α-crystallin. *Biochim Biophys Acta*. 2016;1860(1 Pt B):252–257.
11. Arrigo AP, Simon S, Gibert B, et al. Hsp27 (HspB1) and αB-crystallin (HspB5) as therapeutic targets. *FEBS Lett*. 2007;581:3665–3674.
12. Srivastava SS, Mishra A, Krishnan B, Sharma Y. Ca2+-binding motif of βγ-crystallins. *J Biol Chem*. 2014;289:10958–10966.
13. Shimeld SM, Purkiss AG, Dirks RP, Bateman OA, Slingsby C, Lubsen NH. Urochordate βγ-crystallin and the evolutionary origin of the vertebrate eye lens. *Curr Biol*. 2005;15:1684–1689.
14. Slingsby C, Wistow GJ, Clark AR. Evolution of crystallins for a role in the vertebrate eye lens. *Protein Sci*. 2013;22:367–380.
15. Slingsby C, Wistow GJ. Functions of crystallins in and out of lens: roles in elongated and post-mitotic cells. *Prog Biophys Mol Biol*. 2014;115:52–67.
16. Shiels A, Bennett TM, Hejtmancik JF. Cat-Map: putting cataract on the map. *Mol Vis*. 2010;16:2007–2015.
17. Sinha D, Klise A, Sergeev Y, et al. βA3/A1-crystallin in astroglial cells regulates retinal vascular remodeling during development. *Mol Cell Neurosci*. 2008;37:85–95.
18. Valapala M, Edwards M, Hose S, et al. βA3/A1-crystallin is a critical mediator of STAT3 signaling in optic nerve astrocytes. *Sci Rep*. 2015;5:8755.
19. Zigler JS Jr, Valapala M, Shang P, Hose S, Goldberg MF, Sinha D. βA3/A1-crystallin and persistent fetal vasculature (PFV) disease of the eye. *Biochim Biophys Acta*. 2016;1860(1 Pt B):287–298.
20. Wang S, Zhao WJ, Liu H, Gong H, Yan YB. Increasing βB1-crystallin sensitivity to proteolysis caused by the congenital cataract-microcornea syndrome mutation S129R. *Biochim Biophys Acta*. 2013;1832:302–311.
21. Riazuddin SA, Yasmeen A, Yao W, et al. Mutations in βB3-crystallin associated with autosomal recessive cataract in two Pakistani families. *Invest Ophthalmol Vis Sci*. 2005;46:2100–2106.
22. Weisschuh N, Aisenbrey S, Wissinger B, Riess A. Identification of a novel CRYBB2 missense mutation causing congenital autosomal dominant cataract. *Mol Vis*. 2012;18:174–180.
23. Wang KJ, Wang S, Cao NQ, Yan YB, Zhu SQ. A novel mutation in CRYBB1 associated with congenital cataract-microcornea syndrome: the p.Ser129Arg mutation destabilizes the betaB1/betaA3-crystallin heteromer but not the betaB1-crystallin homomer. *Hum Mutat*. 2011;32:E2050–E2060.
24. Jiao X, Kabir F, Irum B, et al. A common ancestral mutation in CRYBB3 identified in multiple consanguineous families with congenital cataracts. *PLoS One*. 2016;11:e0157005.
25. Mackay DS, Boskovska OB, Knopf HL, Lampi KJ, Shiels A. A nonsense mutation in CRYBB1 associated with autosomal dominant cataract linked to human chromosome 22q. *Am J Hum Genet*. 2002;71:1216–1221.
26. Devi RR, Yao W, Vijayalakshmi P, Sergeev YV, Sundaresan P, Hejtmancik JF. Crystallin gene mutations in Indian families with inherited pediatric cataract. *Mol Vis*. 2008;14:1157–1170.
27. van Leen RW, van Roozendaal KE, Lubsen NH, Schoenmakers JG. Differential expression of crystallin genes during development of the rat eye lens. *Dev Biol*. 1987;120:457–464.
28. 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.
29. Limi S, Senecal A, Coleman RA, et al. Transcriptional burst fraction and size dynamics during lens fiber cell differentiation and detailed insights into the denucleation process. *J Biol Chem*. 2018;293:13176–13190.
30. Bartman CR, Hsu SC, Hsiung CC, Raj A, Blobel GA. Enhancer regulation of transcriptional bursting parameters revealed by forced chromatin looping. *Mol Cell*. 2016;62:237–247.
31. Bahar Halpern K, Tanami S, Landen S, et al. Bursty gene expression in the intact mammalian liver. *Mol Cell*. 2015;58:147–156.
32. Corrigan AM, Tunnacliffe E, Cannon D, Chubb JR. A continuum model of transcriptional bursting. *Elife*. 2016;5:e13051.
33. Tantale K, Mueller F, Kozulic-Pirher A, et al. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun*. 2016;7:12248.
34. Sun J, Rockowitz S, Chauss D, et al. Chromatin features, RNA polymerase II and the comparative expression of lens genes encoding crystallins, transcription factors, and autophagy mediators. *Mol Vis*. 2015;21:955–973.
35. Bahar Halpern K, Caspi I, Lemze D, et al. Nuclear retention of mRNA in mammalian tissues. *Cell Rep*. 2015;13:2653–2662.
36. Alvarez-Saavedra M, De Repentigny Y, Lagali PS, et al. Snf2h-mediated chromatin organization and histone H1 dynamics govern cerebellar morphogenesis and neural maturation. *Nat Commun*. 2014;5:4181.
37. He S, Limi S, McGreal RS, et al. Chromatin remodeling enzyme Snf2h regulates embryonic lens differentiation and denucleation. *Development*. 2016;143:1937–1947.
38. Kokavec J, Zikmund T, Savvulidi F, et al. The ISWI ATPase Smarca5 (Snf2h) is required for proliferation and differentiation of hematopoietic stem and progenitor cells. *Stem Cells*. 2017;35:1614–1623.
39. Chen J, Xu H, Aronow BJ, Jegga AG. Improved human disease candidate gene prioritization using mouse phenotype. *BMC Bioinform*. 2007;8:392.
40. Lionnet T, Czaplinski K, Darzacq X, et al. A transgenic mouse for in vivo detection of endogenous labeled mRNA. *Nat Methods*. 2011;8:165–170.
41. Toiber D, Erdel F, Bouazoune K, et al. SIRT6 recruits SNF2H to sites of DNA breaks, preventing genomic instability through chromatin remodeling. *Mol Cell*. 2013;51:454–468.
42. Wei W, Pelechano V, Jarvelin AI, Steinmetz LM. Functional consequences of bidirectional promoters. *Trends Genet*. 2011;27:267–276.
43. Prasanth KV, Prasanth SG, Xuan Z, et al. Regulating gene expression through RNA nuclear retention. *Cell*. 2005;123:249–263.
44. Zhao Y, Wilmarth PA, Cheng C, et al. Proteome-transcriptome analysis and proteome remodeling in mouse lens epithelium and fibers. *Exp Eye Res*. 2019;179:32–46.
45. Choudhuri A, Maitra U, Evans T. Translation initiation factor eIF3h targets specific transcripts to polysomes during embryogenesis. *Proc Natl Acad Sci U S A*. 2013;110:9818–9823.
46. Lachke SA, Alkuraya FS, Kneeland SC, et al. Mutations in the RNA granule component TDRD7 cause cataract and glaucoma. *Science*. 2011;331:1571–1576.
47. Dash S, Dang CA, Beebe DC, Lachke SA. Deficiency of the RNA binding protein caprin2 causes lens defects and features of Peters anomaly. *Dev Dyn*. 2015;244:1313–1327.
48. Siddam AD, Gautier-Courteille C, Perez-Campos L, et al. The RNA-binding protein Celf1 post-transcriptionally regulates p27Kip1 and Dnase2b to control fiber cell nuclear degradation in lens development. *PLoS Genet*. 2018;14:e1007278.

Investigative Ophthalmology & Visual Science