Comparative transcriptome analysis of epithelial and fiber cells in newborn mouse lenses with RNA sequencing

Thanh V. Hoang,[1] Praveen Kumar Raj Kumar,[1] Sreeskandarajan Sutharzan,[1] Panagiotis A. Tsonis,[2] Chun Liang,[1] Michael L. Robinson[1]

(The first two authors contributed equally to the work)

[1]Department of Biology, Miami University, Oxford, OH; [2]Department of Biology and Center for Tissue Regeneration and Engineering, University of Dayton, Dayton, OH

Purpose: The ocular lens contains only two cell types: epithelial cells and fiber cells. The epithelial cells lining the anterior hemisphere have the capacity to continuously proliferate and differentiate into lens fiber cells that make up the large proportion of the lens mass. To understand the transcriptional changes that take place during the differentiation process, high-throughput RNA-Seq of newborn mouse lens epithelial cells and lens fiber cells was conducted to comprehensively compare the transcriptomes of these two cell types.

Methods: RNA from three biologic replicate samples of epithelial and fiber cells from newborn *FVB/N* mouse lenses was isolated and sequenced to yield more than 24 million reads per sample. Sequence reads that passed quality filtering were mapped to the reference genome using Genomic Short-read Nucleotide Alignment Program (GSNAP). Transcript abundance and differential gene expression were estimated using the Cufflinks and DESeq packages, respectively. Gene Ontology enrichment was analyzed using GOseq. RNA-Seq results were compared with previously published microarray data. The differential expression of several biologically important genes was confirmed using reverse transcription (RT)-quantitative PCR (qPCR).

Results: Here, we present the first application of RNA-Seq to understand the transcriptional changes underlying the differentiation of epithelial cells into fiber cells in the newborn mouse lens. In total, 6,022 protein-coding genes exhibited differential expression between lens epithelial cells and lens fiber cells. To our knowledge, this is the first study identifying the expression of 254 long intergenic non-coding RNAs (lincRNAs) in the lens, of which 86 lincRNAs displayed differential expression between the two cell types. We found that RNA-Seq identified more differentially expressed genes and correlated with RT-qPCR quantification better than previously published microarray data. Gene Ontology analysis showed that genes upregulated in the epithelial cells were enriched for extracellular matrix production, cell division, migration, protein kinase activity, growth factor binding, and calcium ion binding. Genes upregulated in the fiber cells were enriched for proteosome complexes, unfolded protein responses, phosphatase activity, and ubiquitin binding. Differentially expressed genes involved in several important signaling pathways, lens structural components, organelle loss, and denucleation were also highlighted to provide insights into lens development and lens fiber differentiation.

Conclusions: RNA-Seq analysis provided a comprehensive view of the relative abundance and differential expression of protein-coding and non-coding transcripts from lens epithelial cells and lens fiber cells. This information provides a valuable resource for studying lens development, nuclear degradation, and organelle loss during fiber differentiation, and associated diseases.

BACKGROUND

The ocular lens is an excellent model for studying development, physiology, and disease [1]. The mammalian lens is made up of only two cell types: epithelial cells, which comprise a monolayer of cells that line the anterior hemisphere of the lens, and fiber cells, which make up the remainder of the lens mass. The primary lens fiber cells result from differentiation of the cells in the posterior half of the lens vesicle while secondary fiber cells differentiate from lens epithelial cells displaced toward the equator by lens epithelial cell proliferation. During differentiation, lens epithelial cells undergo cell cycle arrest, elongate, and begin expressing genes characteristic of lens fiber cells [2]. Eventually, the differentiating fiber cells lose their nuclei and other intracellular organelles, such that the most mature lens fiber cells in the center of the lens exist in an organelle-free zone [3]. Lens growth, through epithelial cell proliferation and secondary fiber cell differentiation, occurs throughout the vertebrate lifespan.

Lens fiber cell differentiation is a highly coordinated process involving specific changes in gene expression between two different cell types. For example, several genes,

Correspondence to: Michael L. Robinson, Department of Biology, Miami University, Oxford, Ohio 45056; Phone: (513) 529-2353; FAX: (513) 529-6900; email: Robinsm5@miamioh.edu

including *Pax6, Foxe3*, and E-cadherin (*Cdh1*), are highly expressed in lens epithelial cells but downregulated in lens fiber cells [4-6]. In contrast, many other genes, including Aquaporin0 (*Aqp0 or Mip*), β- and γ-crystallins, CP49 (*Bfsp2*), and filensin (*Bfsp1*), are expressed at low levels in lens epithelial cells but are strongly upregulated during lens fiber cell differentiation [7-9]. However, a comprehensive understanding of gene expression changes during lens fiber differentiation remains incomplete. In particular, the expression and role of long non-coding RNAs during lens development largely await discovery.

Increasing evidence suggests an important role for long intergenic non-coding RNAs (lincRNAs) in regulating gene transcription and protein synthesis [10,11]. LincRNAs are non-coding transcripts more than 200 nucleotides long that are intergenically transcribed. LincRNAs can regulate gene expression via *cis* and *trans* mechanisms. LincRNAs potentially function in many different ways, including cotranscriptional regulation, bridging proteins to chromatin, and scaffolding of nuclear and cytoplasmic complexes [11]. Little information currently exists about the specific expression pattern or function of lincRNAs during lens development.

Microarrays provide a comprehensive approach for gene-expression studies [12]. Several previous investigations applied microarray technology to the lens, where transcriptional profiling was typically restricted to whole lenses [13,14], fiber cells [15], or lens epithelial explants [16-18]. However, microarrays have several limitations, including probe cross-hybridization, the selection of specific probes, and low detection thresholds that may reduce the ability to accurately estimate low-level transcripts. Additionally, novel transcripts and splice isoforms of annotated genes are often missed because microarray design often limits information to previously identified transcripts.

The application of next-generation sequencing (NGS) technology creates enormous potential to increase the sensitivity and resolution of genomic and comprehensive transcriptome analyses without many of the limitations of microarrays [19]. Visualization of mapped sequence reads spanning splice junctions can also reveal novel isoforms of previously annotated genes, which was not possible with microarrays [20,21]. Deep sequencing of RNA with NGS (RNA-Seq) permits comprehensive evaluation and quantification of all types of RNA molecules expressed in a cell or tissue, including mRNA and non-coding RNAs [22]. RNA-Seq can detect rare, novel, and alternatively spliced transcript isoforms. RNA-Seq also identifies rare mutations and RNA editing, and can elucidate non-model organism transcriptomes [23-25]. The declining costs and increased availability of RNA-Seq are fueling the emergence of this technology as the most powerful current method for comprehensive transcriptome profiling.

To date, two studies have used RNA-Seq in mouse lenses. However, these studies were restricted to mRNA expression profiling in whole lenses [26] or in lens epithelial explants [27]. Here, we present the first application of RNA-Seq to comprehensively investigate the transcriptomes of epithelial cells and fiber cells of newborn *FVB/N* mouse lenses. RNA-Seq results were then compared with those obtained with previously published microarray methods [28]. Furthermore, RT-qPCR confirmed the expression of several biologically important genes and provided a benchmark with which to compare the RNA-Seq and microarray analysis.

We found that RNA-Seq provided a more complete and accurate approach for comprehensive, comparative analysis of the lens transcriptome than previously described methods. RNA-Seq analysis of lens epithelial cells and lens fiber cells discovered 1,368 more differentially expressed transcripts than found in a similar microarray analysis of P13 *C57BL/6J* mouse lenses [28]. Surprisingly, only 22% of the genes (1,009) differentially expressed in this microarray study were differentially expressed in the current RNA-Seq analysis. The comprehensive and quantitatively continuous nature of RNA-Seq will facilitate better understanding of lens fiber differentiation and organelle destruction inherent in lens development.

METHODS

Animals: Newborn, inbred *FVB/N* mice, or P13 *C57BL/6J* mice, were euthanized with CO_2 inhalation before the lenses were removed. All procedures were approved by the Miami University Institutional Animal Care and Use Committee and complied with the ARVO Statement for Use of Animals in Research and consistent with those published by the Institute for Laboratory Animal Research (Guide for the Care and Use of Laboratory Animals).

RNA-Seq library preparation and HiSeq sequencing: Lenses were removed from the eye and carefully isolated by manual dissection from the blood vessels, retina, and cornea. The isolated lenses were dissected into two fractions. The epithelial fraction contained the lens capsules with adherent cells that include the entire central and equatorial epithelium and cells in the transitional zone. This fraction also likely contained some early fiber cells firmly attached to the capsule, although obvious elongated fiber cells in this fraction were manually removed. Although every attempt was made to physically remove the tunica vasculosa lentis, the lens epithelial fraction most likely contained cells from this vasculature. The fiber cell fraction included the bulk

of the lens mass that was non-adherent to the lens capsule. Epithelial and fiber cell fractions were each pooled into three biologic replicate samples for each cell type. Each sample contained tissue from eight lenses from four newborn mice. Total RNA, including mRNA and lincRNA, was isolated using the mirVana isolation kit (Ambion, Life Technologies, Grand Island, NY). Total RNA samples with the RNA integrity number (RIN, Agilent 2100 Bioanalyzer) ≥8.0 were used to prepare a library of template molecules suitable for subsequent sequencing on an Illumina (St. Louis, MO) HiSeq platform. Briefly, polyadenylated RNA was purified from the total RNA samples using Oligo dT conjugated magnetic beads and prepared for single-end sequencing according to the Illumina TruSeq RNA Sample Preparation Kit v3. The library was sequenced for 50 cycles using the TruSeq SBS kit on an Illumina HiSeq 2000 system at the Genomics and Sequencing Core Laboratory at the University of Cincinnati. Approximately 29 million sequences that were 51 bp long per sample were generated. Bases were called and translated to generate FASTQ sequence files. All raw sequencing data were deposited at the NCBI's Sequence Read Archive (SRA) and are accessible under the SRA accession number SRP040480.

Bioinformatic analysis:

Data preprocessing—The six Illumina HiSeq data sets, three pooled biologic replicates from each tissue (designated E1, E2, and E3 for the epithelial samples and F1, F2, and F3 for the fiber cell samples), were individually assessed for quality using prinseq [29] and FastQC. Raw Illumina sequence reads were trimmed for low-quality reads (phred <28), ambiguous bases (N), sequencing adapters, primers, and poly(A)/(T) tails.

Reference mapping—The trimmed reads were mapped to the *Mus musculus* reference genome (build GRC38.72, ENSEMBL/MGI annotations from *C57BL/6J* [30]) using the Genomic Short-read Nucleotide Alignment Program (GSNAP version gmap-gsnap-2013–09–30.v2) [31]. The GSNAP feature of single-nucleotide polymorphism (SNP) tolerant mapping was used with the available SNP information downloaded from ENSEMBL as Variant call format. The SNPs and the known splice sites were obtained from ENSEMBL genome annotation. Sequencing reads were mapped to the reference genome with two mismatches allowed. Unique and perfect mappings were identified using the Sequence Alignment Format (SAM) flag NH:i:1 and NM:i:0, respectively. For high confidence in calling gene expression, only uniquely and perfectly mapped reads (with the inclusion of SNPs) were used.

Estimation of gene expression and identification of differentially expressed genes—Transcript assembly and abundance estimation were performed using Cufflinks 2.1.1 [32]. Gene expression levels were expressed as reads per kilobase per million (RPKM) mapped reads. The minimum threshold of 0.5 RPKM was used to define gene expression. The DESeq package was used to study differential gene expression [33], where the input count table was obtained using HTseq. The significance of differentially expressed genes was identified with an adjusted p value (in DESeq, the adjusted p value considers multiple testing using the Benjamini-Hochberg method [33]) less than 0.05 (at 95% confidence) and with a fold change greater than 1.5. Protein-coding genes and non-coding genes were then sorted using Perl scripts. A Venn diagram was created using VENNY.

Functional annotation: The GOseq package was used to find Gene Ontology (GO) categories, including biochemical process, cellular components, and molecular function, to identify particular functions enriched among differentially expressed genes. GOseq has the ability to account for gene length and read count biases [34].

Validation of the gene expression with RT-qPCR: To validate the RNA-Seq data, the expression level of selected genes was analyzed with RT-qPCR. Genes that have low and high expression levels were chosen. The same epithelial and fiber cell RNA samples used for RNA-Seq were reverse transcribed into cDNA, using random primers and Superscript II reverse transcriptase (Invitrogen), according to the manufacturer's instructions. After incubation at 42 °C for 50 min, the reaction was stopped by heating to 75 °C for 15 min. The qPCR assays were performed on the cDNA using GoTaq Green Master Mix (Promega, Madison, WI) following the manufacturer's instructions and read using a CFX connect instrument (Bio-Rad, Hercules, CA). Intron-spanning primers were designed to specifically quantify targeted mRNA transcripts (Appendix 1). Although we did not perform an RNA-Seq analysis of *C57BL/6J* RNA, we compared our RNA-Seq results to a previous microarray study of lens epithelial cells and lens fiber cells from P13 *C57BL/6J* mice. Therefore, RNA from P13 *C57BL/6J* lens epithelial cell and lens fiber cell fractions was analyzed with RT-qPCR to allow for more direct comparison with a previously performed differential gene expression analysis with microarray [28]. Each biologic sample was analyzed in triplicate with qPCR. Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) expression was used as an endogenous control. The cycling conditions consisted of 1 cycle at 95 °C for 100 s for denaturation, followed by 40 three-step cycles for amplification (each cycle consisted of 95 °C incubation for 20 s, an appropriate annealing temperature for

10 s, and product elongation at 70 °C incubation for 20 s). The melting curve cycle was generated after PCR amplification. The reaction specificity was monitored by determining the product melting temperature and by checking for the presence of a single DNA band on agarose gels from the RT-qPCR products. To calculate the amplification efficiencies (E) of the qPCR primers, standard curves were obtained by performing qPCR reactions on serially diluted samples; the PCR efficiencies (E) were calculated based on the slopes of the standard curves (E=10 $^{(-1/slope)}$ −1). The quantification cycle (Cq, also commonly called the Ct) was obtained, and the ΔCq value was calculated with Cq $_{(gene)}$ − Cq $_{(GAPDH)}$. For fold-change (FC) expression, ΔΔCq was first calculated by subtracting the mean of the ΔCq values of the fiber samples from the meanΔCq values of epithelial samples and then converted to 2 $^{(-\Delta\Delta Cq)}$. The data are expressed as mean ± standard error of the mean (SEM).

RESULTS AND DISCUSSION

Transcriptome analysis overview of newborn lenses: Previously, gene expression profiling in the lens used expressed sequence tag analysis, microarrays, and RNA-Seq analysis; however, these studies focused on examining the whole lens [13,14,26,35-38], lens fiber cells [15], or lens epithelial explants [16-18,27]. Unlike prior studies, the current work examined changes in the expression of protein-coding genes and non-coding genes from epithelial cells and fiber cells in vivo. To our knowledge, these studies represent the first application of RNA-Seq to identify the transcriptional changes underlying epithelial cell differentiation into fiber cells during mouse lens development.

The lens epithelial and fiber samples in this RNA-Seq analysis consisted of cells physically separated based on adherence to the lens capsule. This method of separation raises important considerations for interpreting gene expression data. Lens epithelial cells at the anterior surface of the lens can be grouped into central and peripheral regions [39]. Epithelial cells in the central region are considered mitotically quiescent, while epithelial cells in the peripheral region can be divided into the germinative zone and the transitional zone. The germinative zone epithelium is mitotically active, while the transitional zone epithelial cells closest to the equator undergo cell cycle arrest and align into meridional rows before differentiating into fiber cells [40,41]. These earliest differentiating transitional zone cells remain adherent to the lens capsule and represent a fraction of the transcripts in the "lens epithelial" fraction in the current RNA-Seq analysis. This, coupled with the possible adherence of a few early differentiating fiber cells, may explain the

appearance of "fiber cell-specific" transcripts such as those for *Aqp0* and *Gja3* in the lens epithelial RNA-Seq data (see below). Likewise, the lens fiber fraction is heterogeneous, containing young, cortical fiber cells in the outer layer and fully mature, organelle-free fiber cells in the center of the lens [42]. Therefore, despite dividing the lens transcripts into epithelial and fiber cell fractions, the result of the RNA-Seq analysis reflected the transcriptional changes in somewhat heterogenous cell populations resulting from the isolation method.

RNA-Seq obtained approximately 25 to 33 million raw sequence reads for each of the six cDNA libraries (three from capsular adherent lens epithelial cells and three from fiber cells) obtained from pooled newborn *FVB/N* mouse lenses. Trimmed reads that passed the quality filter were mapped to the mouse *C57BL/6J* reference genome (GRC38.72) using the software GSNAP with the SNP tolerant mapping with SNPs known to differ between *C57BL/6J* and *FVB/N* mice. On average, 90% of the total reads were successfully mapped to the reference genome (Table 1). Using Cufflinks for estimating transcript abundance, with an expression-level threshold of RPKM ≥0.5, the lens epithelial fraction contained transcripts from 13,732 genes and the fiber cells expressed 10,850 genes for a total of 14,060 genes expressed in the lens. The epithelial and fiber cell fractions coexpressed 10,522 (74.8%) of these genes. Using DESeq (p value ≤0.05 and fold change ≥1.5, Figure 1, Appendix 2) for examining differential gene expression, the lens epithelial cells and the lens fiber cells differentially expressed 6,022 protein-coding genes. Of these genes, the lens epithelial cells upregulated 3,233 genes, and the lens fiber cells upregulated 2,789 genes. Among 1,746 annotated mouse lincRNA genes, the lens expressed 254 lincRNAs. Of these, 86 lincRNA genes were differentially expressed between the two lens cell types, with epithelial cells upregulating 32 genes and fiber cells upregulating 54 genes (Appendix 3).

Genes with the highest expression level in the epithelial cells (ranked by RPKM values) included several crystallins, clusterin (*Clu*), collagens (*Col4a1* and *Col4a2*), osteonectin (*Sparc*), heat shock protein 90-beta (*Hsp90ab1*), glucose transporter-1 (*Slc2a1*), and heparan sulfate proteoglycan-2 (*Hspg2*; Appendix 4). Furthermore, lens epithelial cells highly expressed many mitochondrial-encoded genes, including NADH dehydrogenases (*mt-Nd1*, *mt-Nd2*, *mt-Nd4*, and *mt-Nd5),* cytochrome b (*mt-Cytb*), tRNA leucine 1 (*mt-Tl1*), and leucyl-tRNA synthetase-2 (*Lars2*). Interestingly, although not differentially expressed, the epithelial and fiber cells expressed abundant transcripts for the novel lincRNA, *RP23–81C12.3.*

Sequence reads	E1	E2	E3	F1	F2	F3
Total raw reads	33,174,286	29,919,226	29,965,660	28,652,759	30,661,663	24,833,352
Reads passed quality filtering	30,524,674	27,392,629	28,203,944	27,019,907	27,760,247	22,735,990
	(92%)	(91.5%)	(94.1%)	(94.3%)	(90.5%)	(91.6%)
Mapped reads	29,414,088	27,286,119	27,299,912	25,814,802	27,495,080	22,630,822
	(89%)	(91%)	(91.1%)	(90%)	(89.6%)	(91.1%)
Uniquely and perfectly mapped reads	12,925,906	18,954,231	18,408,241	18,686,534	18,354,154	15,975,465
	(39%)*	(63.3%)	(61.4%)	(65.2%)	(59.8%)	(64.3%)

TABLE 1. SUMMARY OF BASE CALLING AND ALIGNMENTS.

* Although all 6 samples have similar percentages of mapped reads, we found a higher number of reads that mapped ambiguously to reference genome in E1 samples than the other 5 samples. We used other mapping tools to verify these results with similar outcome. We do not know why the E1 sample contains fewer unique reads, and at this point any explanation for this discrepancy would be speculation on our part.

The fiber cell RNA libraries contained numerous transcripts for genes encoding several different crystallins, as well as several structural proteins (*Lim2*, *Bfsp1*, *Sparc*, and *Vim*), CD24a antigen (*Cd24a*), Tudor domain containing 7 (*Tdrd7*), connexins (*Gja3 and Gja8*), and cyclin-dependent kinase inhibitor 1C (*Cdkn1c*, also known as *p57KIP2*; Appendix 5). The fiber cells also expressed several lincRNAs, including *RP23–81C12.3*, metastasis-associated lung adenocarcinoma transcript 1 (*Malat1*), growth arrest specific 5 (*Gas5*), and maternally expressed gene 3 (*Meg3*).

Genes differentially expressed (DEGs) in the epithelial cell and fiber cell fractions were first filtered based on whether they were upregulated in the epithelial cells or fiber cells. Genes were then ranked by the lowest adjusted p value rather than the greatest fold change to avoid large fold-change differences that failed to achieve sufficient significance as a result of low expression levels. The most significantly
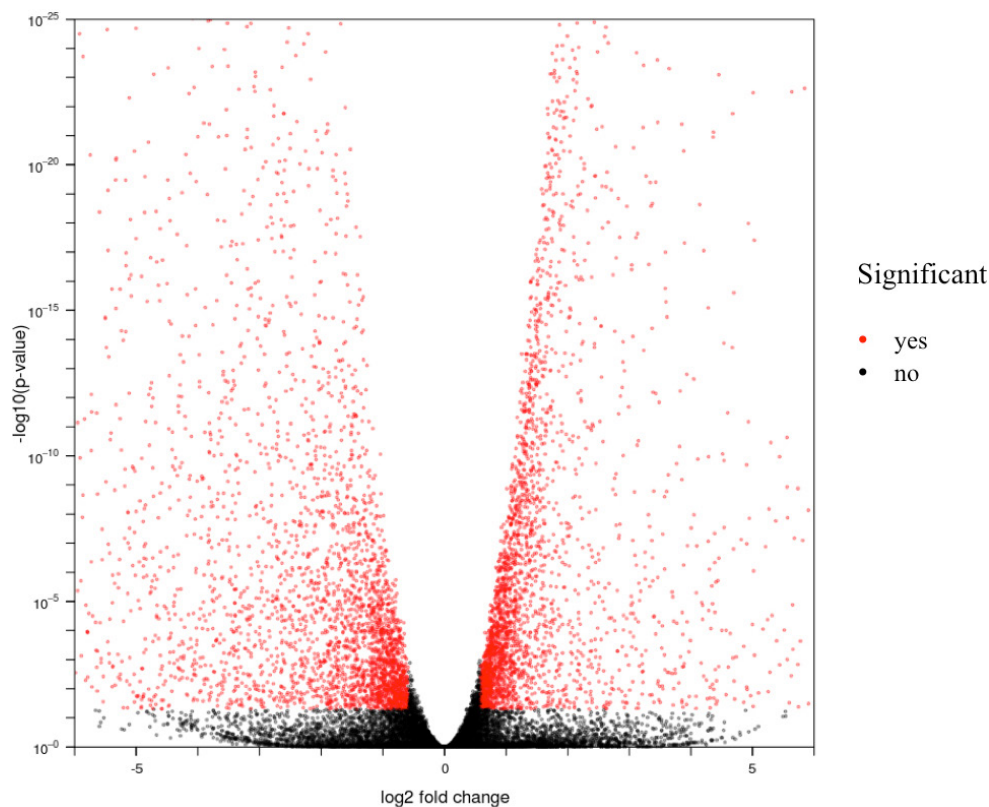

Significant

• yes
• no

Figure 1. Volcano plot showing differentially expressed genes between three replicates of lens epithelial cells and three replicates of lens fiber cells. The x-axis corresponds to the $\log_2$ fold-change value, and the y-axis displays the $\log_{10}$ p value. The red dots represent the significantly differentially expressed transcripts (p≤0.05 and fold change ≥1.5) between the lens epithelial cells and the lens fiber cells, while the black dots are not statistically significant (p>0.05).

differentially expressed genes, ranked by adjusted p value, in the epithelial cells included arylsulfatase family member I (*Arsi*), immunoglobulin superfamily containing leucine-rich repeat protein (*Islr*), sulfatase-1 (*Sulf1*), folate receptor-1 (*Folr1*), B-cell translocation gene-1 (*Btg1*), nepronectin (*Npnt*), vascular endothelial growth factor receptor-1 (*Flt1*), and cyclin dependent kinase-1 (*Cdk1*; Figure 2A). The most significantly DEGs in the fiber cells tended to be those that encode lens structural proteins and membrane proteins, including several different crystallins, filensin (*Bfsp1*), gap junction protein epsilon-1 (*Gje1*), aquaporin0 (*Aqp0* also known as *Mip*), tropomodulin1 (*Tmod1*), CD24 molecule (*Cd24a*), solute carrier family 24 (*Slc24a4*), and lens fiber membrane intrinsic protein (*Lim2*; Figure 2B). The structure and transparency of the lens require a high concentration (about 350 mg/ml) of crystallins, which represent the largest class of lens structural proteins. Mutations in several different crystallin genes result in cataract in mice and humans [43]. Len cell membranes contain high concentrations of the lens-specific water channel, Aqp0, as well as gap junctions and ion channels, which cooperate to allow the lens to transport nutrients and waste products despite the lack of vasculature in the mature lens [44].

Lens fiber cells upregulated the expression of *Dnase2b*, *Tdrd7*, *Hsf4*, and *Tmod1*, consistent with the importance of these genes in lens development and fiber cell differentiation. DNase2b activity is required for nuclear DNA degradation in the cortical fiber cells, and loss of *Dnase2b* leads to nuclear cataracts in mice [45]. Tdrd7, a member of a family of proteins that form RNA granules, has been shown to regulate gene expression in the lens by controlling the fate of mRNA. Loss-of-function mutations in *TDRD7* in humans and *Tdrd7* nullizygosity in mice cause cataracts as well as glaucoma [46]. *Hsf4* mutations are associated with autosomal dominant lamellar and Marner cataract [47]. *Hsf4* null mice have cataracts with abnormal fiber cells [48], due to downregulation of gamma-crystallin, *Bfsp1*, and *Dnase2b* expression [49,50]. Previous studies have shown that Tmod1 is required for coordinating fiber cell shapes and geometry during lens fiber elongation and maturation by stabilizing the spectrin-actin network in the lens fiber cell membrane. Deletion of Tmod1
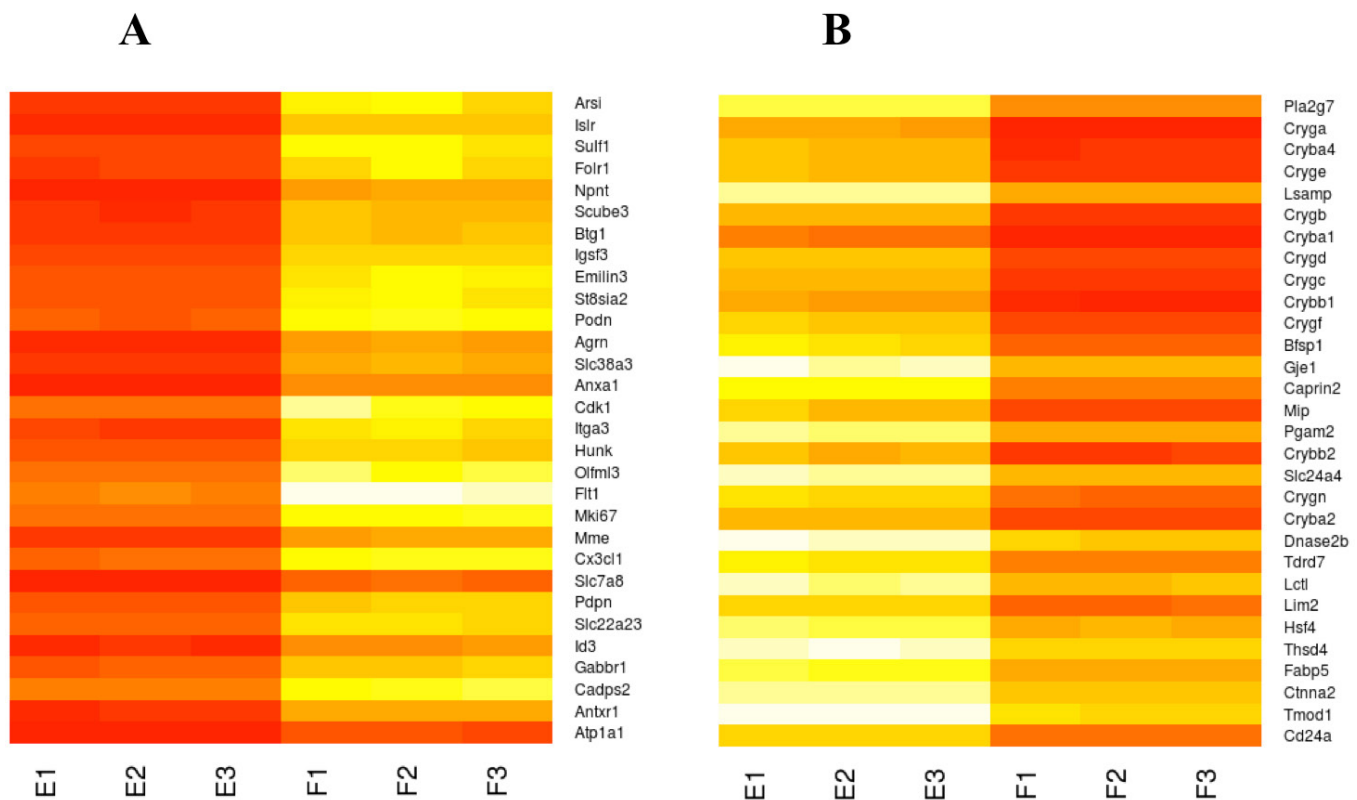

Figure 2. Top 30 most significant differentially expressed protein-coding genes between epithelial and fiber cells ranked by adjusted p value. **A**: Genes with upregulated expression in epithelial cells. **B**: Genes with upregulated expression in fiber cells. In the heatmap, white represents low levels of expression while red represents high levels. Sample abbreviations: E1, E2, and E3: epithelial replicates; F1, F2, and F3: fiber replicates.

leads to fiber cell disorder and impairs lens transparency [51,52].

Although *Pla2g7*, *Cd24a*, and *Caprin2* were highly expressed in the fiber cells, previous studies demonstrated that normal lens development occurs in the absence of these genes. Pla2g7 is a secreted enzyme that catalyzes the degradation of platelet-activating factor into inactive products [53] and has been shown to be involved in arthrosclerosis and coronary diseases [54,55]. Pla2g7 null mice grow to adulthood without a reported eye phenotype [56]. Cd24a, a cell-surface glycoprotein, is highly expressed in fiber cells in our data, consistent with a previous study [57]. However, no eye phenotype was reported for Cd24a-knockout mice [58]. Caprin2 was highly expressed in the lens primary fiber cells and the secondary fiber cells in mouse and chick lens, and that expression was induced by fibroblast growth factor (FGF) signaling [59].

RNA-Seq data also demonstrated strong expression of *Clu*, *Hsp90ab1*, *Fabp5*, *Thsd4*, and *Ctnna2* in the fiber cells, suggesting that these genes may be biologically important for fiber differentiation or function. *Clu* and *Hsp90ab1* encode chaperone proteins and may play an important role in preventing protein misfolding and maintaining lens transparency in the fiber cells. Fabp5 is involved in fatty acid uptake and metabolism [60], but no lens phenotype was reported in mice with targeted mutations in *Fabp5* [61]. Pgam2 is an important enzyme in glycolytic pathway [62], but its expression and function in the lens has not been previously examined. Thsd4, also known as ADAMTSL4, is a secreted glycoprotein and is found in differentiating fiber cells. Thsd4 interacts with fibrillin-1 and promotes fibril formation [63]. Mutations in *Thsd4* cause ectopia lentis, which is malposition of the lens from its normal position [64]. However, the function of Thsd4 in the lens fiber cells remains unknown. The *Ctnna2* gene encoding cytoplasmic alpha-N-catenin is strongly expressed in the lens fibers [65]. It functions by interacting with actin cytoskeleton with beta-catenin at cell junctions [66]. It is not clear whether the effect of the loss of αN-catenin on the lens was examined in mice [67].

The lens also upregulates the expression of *Lctl* during differentiation. *Lctl* encodes a novel Klotho-lactase-phlorizin hydrolase-related protein [68]. Klotho and βKlotho act as coreceptors for the FGF15/19 subfamily of endocrine FGFs. Although the function of Lctl in vivo remains unknown, expression of Lctl potentiated FGF19-induced ERK1/2 phosphorylation in HEK293 cells [69].

Comparison of RNA-Seq result with microarray analysis: Nakahara et al. previously compared gene expression between mouse lens epithelial cells and lens fiber cells using a microarray [28]. This microarray analysis examined P13 lenses from the *C57BL/6J* mouse inbred strain and, like our RNA-Seq analysis, used a ±1.5 fold-change cutoff for differential expression. Lens epithelial and lens fiber cell fractions in the microarray analysis were isolated identically as described for the current RNA-Seq analysis. However, although our RNA-Seq analysis used newborn *FVB/N* lens RNA with three biologic replicates per fraction the previously published microarray lacked biologic replicates. The lack of biologic replicates prevented the assessment of p values for the microarray analysis.

When comparing the results of the RNA-Seq data with that obtained in this microarray study, the microarray detected 4,654 DEGs, much lower than that detected by RNA-Seq (6,022 DEGs). Of these DEGs discovered with the microarray, the lens epithelium upregulated 1,812 genes, and the fiber cells upregulated 2,842 genes. In contrast to RNA-Seq, microarray analysis requires previous knowledge of gene transcripts to construct hybridization probes, restricting analyses to transcripts built on the array. Therefore, the specific microarray used prevented the analysis of several genes (for example, *Cdkn1b*) due to a lack of specific probes for hybridization. In total, 1,009 DEGs were commonly identified with RNA-Seq and microarray analysis (Figure 3). This figure represents only 22% of the differentially expressed genes identified previously by microarray.

Validation of RNA-Seq and microarray results with RT-qPCR: To validate the RNA-Seq analysis, transcripts from 19 genes that have a wide range of expression (from no change to different levels of differential expression) were chosen for the RT–qPCR analysis. These genes included those involved in cell cycle regulation, receptor tyrosine kinase signaling, DNA methylation, and cytosolic enzymes. RT–qPCR facilitated the analysis of gene expression levels, from P0 *FVB/N* and P13 *C57BL/6J* lenses, to compare the RNA-Seq data obtained from P0 *FVB/N* RNA with the previous microarray data obtained from P13 *C57BL/6J* RNA. As in the RNA-Seq and microarray, we maintained a 1.5 FC as the threshold for differential gene expression with negative FC values indicating lower expression in the fiber cell fraction compared with the epithelial cell fraction. For genes that exhibit a positive FC, our RNA-Seq data correlated well with the RT-qPCR results from P0 *FVB/N* samples (Figure 4A), except *Mapk3* and *Atg3*, which were not differentially expressed according to RT-qPCR (FC=1.11 and 1.36, respectively). Meanwhile, the correlation of the microarray data with the RT-qPCR results for P13 *C57BL/6J* samples was weaker, notably in the cases of *Fgfr3*, *Evt5*, and *Prox1* expression. For genes more highly expressed in the epithelium (Figure 4B) and genes with large
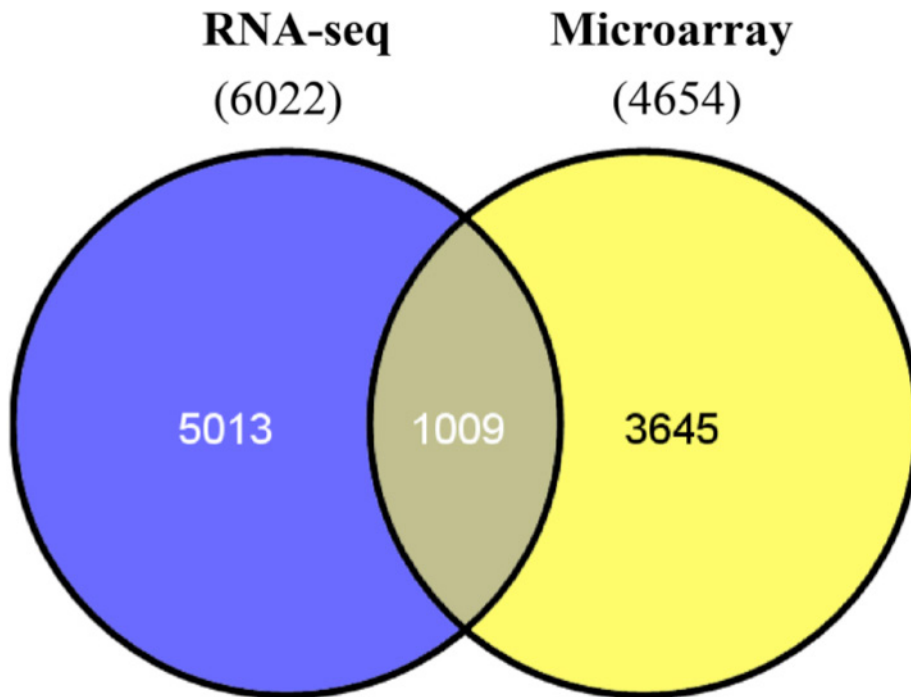
Figure 3. Number of differentially expressed genes shared between our RNA-Seq results and microarray analysis. In total, 1,009 differentially expressed genes are common among these two technologies, even though our RNA-Seq revealed 6,022 differentially expressed genes while the microarray found 4,654 genes.

differential expression (Figure 4C), the RNA-Seq data also exhibited better agreement with the RT-qPCR results than the microarray analysis. The lack of a specific microarray probe for *Cdkn1b* prevented microarray data analysis for this gene. Overall, the RNA-Seq data and the microarray data trended in the same direction, regarding differential gene expression, as the RT-qPCR analysis. However, the RNA-Seq data correlated more closely with the RT-qPCR data than with the microarray data.

The use of different inbred mouse strains and different aged lenses (P0 *FVB/N* in the RNA-Seq and P13 *C57BL/6J* in the microarray) likely contributed to some specific differences in the conclusions between these two analyses (for example, see Figure 4A for *Atg3*, and *Lgmn*) where the RT-qPCR of P13 C57BL/6 RNA correlated most closely with the microarray. However, RT-qPCR validation using P0 *FVB/N* and P13 *C57BL/6* RNA confirmed that RNA-Seq often more closely predicted the gene expression level than the microarray (for example, see Figure 4A for *Mapk1*, *Etv5*, *Spry2*, and *Prox1*).

Functional enrichment of DEGs: Genes upregulated in epithelial cells and genes upregulated in fiber cells were analyzed for GO enrichment using the GOseq software package, which corrects for gene length and read count biases during RNA-Seq [34]. A comparison of the top GO terms (based on p value) for cellular component, molecular function, and biologic process for genes more highly expressed in

epithelial cells or fiber cells is shown in Figure 5. For cellular component terms, genes upregulated in epithelial cells are enriched for components of the extracellular matrix, plasma membrane, cell surface, chromosomes, and intracellular organelles, whereas genes upregulated in fiber cells show enrichment for proteosome complexes, the cytoskeleton, and respiratory chain (Figure 5A).

For molecular function terms, genes upregulated in epithelial cells were enriched for transcripts that encode protein kinase and protein dimerization activities. Transcripts encoding binding activities, include calcium binding, ion binding, growth factor binding, protein kinase binding, receptor binding, integrin binding, and glycosaminoglycan binding, were abundantly expressed in the lens epithelial cell fraction. Not surprisingly, the transcripts upregulated in fiber cells were enriched for structural constituents of the eye lens, cytoskeletal protein binding, hydrolase activity, unfolded protein binding, and ubiquitin binding (Figure 5B). Biologic process terms enriched in genes upregulated in epithelial cells include development, cell cycle, cell division, and cell migration. This is consistent with the fact that epithelial cells undergo active proliferation and migration before fiber cell differentiation (Figure 5C). In contrast, genes upregulated in fiber cells were enriched for transcripts encoding transport, localization, and metabolic processes.

In addition to heterogeneity within the lens cell populations, potential non-lens transcript contamination should be
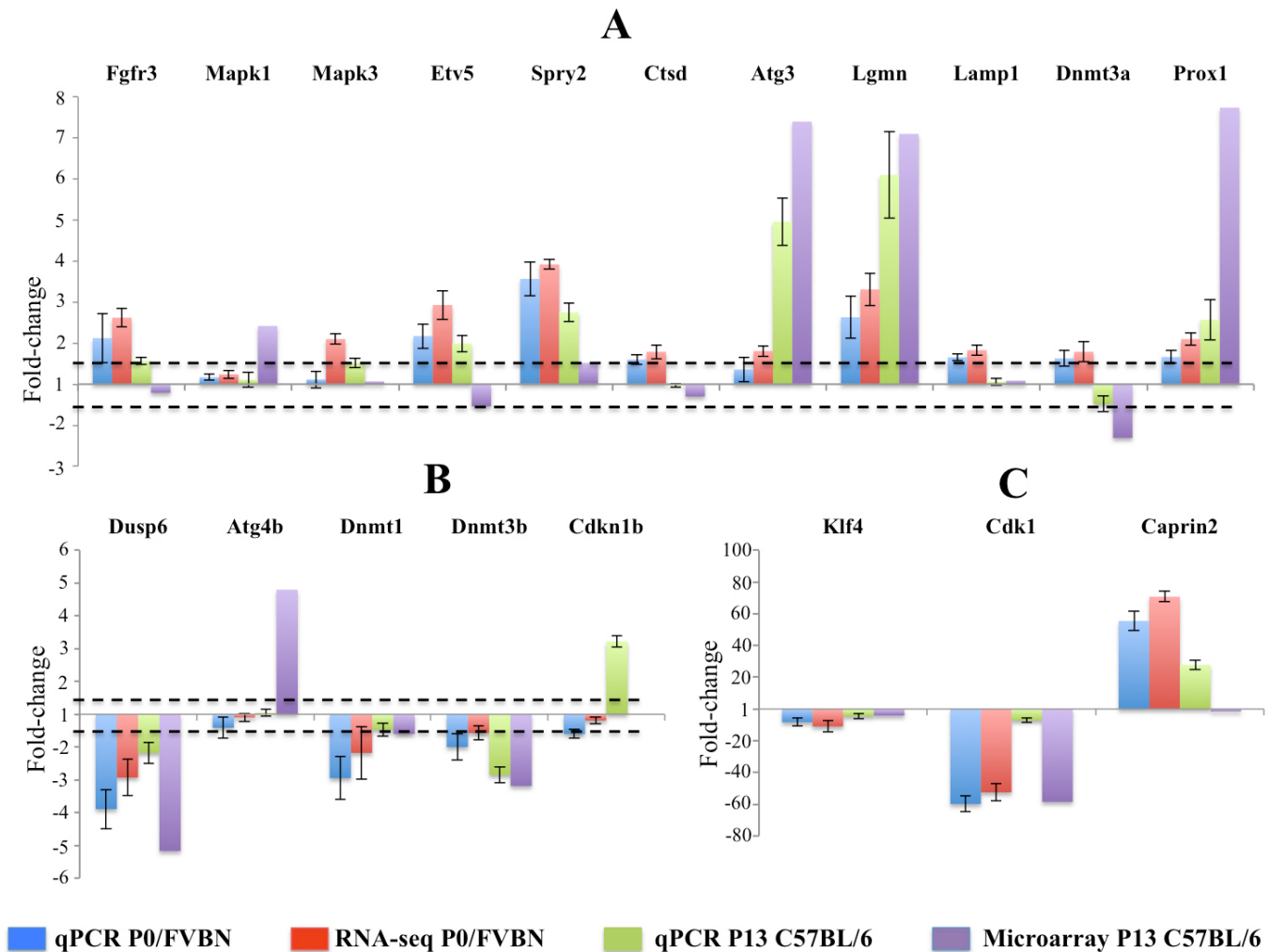
Figure 4. Comparison of RNA-Seq and microarray analysis with RT-qPCR validation assays. Quantitative measurement of gene expression was determined with quantitive reverse transcription (qRT)–PCR for newborn FVB lenses (blue) and for 13 day-old C57/BL6 lenses (green) compared with RNA-Seq for newborn FVB lenses (red) and microarray for 13-day-old C57/BL6 (purple) for 19 genes with a wide range of expression levels. **A**: Genes that were upregulated in fiber cells. **B**: Genes that were upregulated in epithelial cells. **C**: Genes with large fold changes in the expression levels. Fold change was calculated based on gene expression in the fiber cells relative to the epithelial cells. Negative values indicate lower expression in the fiber cells compared with the epithelial cells. Genes that have more than 1.5 fold changes (dashed line) are differentially expressed genes.

considered. Although the lenses were isolated from the eyes and carefully dissected from the blood vessels, retina, and cornea, the biologic process term enrichment of the RNA-Seq data showed that vascular development was enriched in epithelial-upregulated genes (Figure 5C). When looking at the expression of endothelial markers in the RNA-Seq data, *Pecam-1* and *Flt1* (also known as *Vegfr1)* were highly expressed in epithelial fractions, indicating the possibility of contaminating blood vessels. However, we cannot rule out that the gene expression in epithelial fractions is indeed bona fide. Opn1sw, a retinal marker, and Rpe65, an RPE cell marker, are not expressed in the lens [70]. Consistent with

this study, we did not detect the expression of those genes in either lens epithelial or fiber cells, excluding the possibility of the contaminating retina and RPE.

Analysis of important signaling pathways in lens development: The RNA-Seq data were analyzed for differentially expressed genes involved in selective signaling pathways that play important roles during lens development and fiber cell differentiation. Since others have discussed transcription factor expression in the developing lens [71], we will omit specific discussion of transcription factors.

Receptor tyrosine kinases: Among more than 40 receptor tyrosine kinases (RTKs) expressed in the lens, FGF

Up-regulated genes in Epithelial cells

A. Cellular component
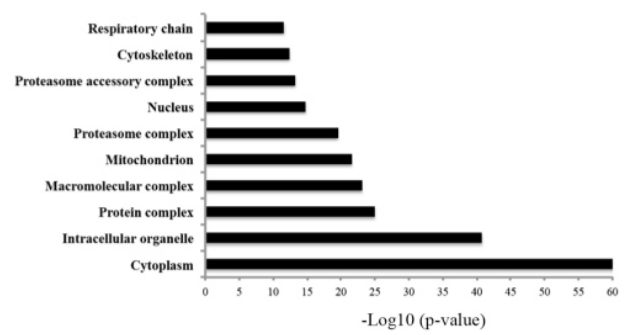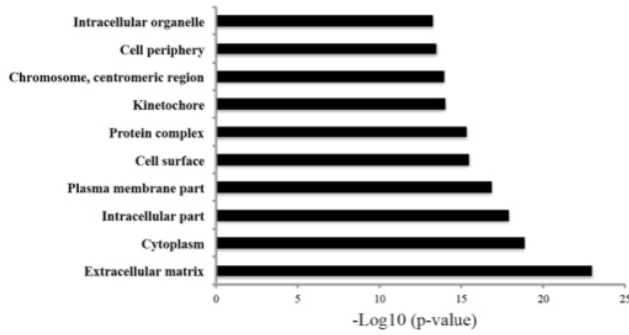
B. Molecular function

C. Biological process

Up-regulated in Fiber cells
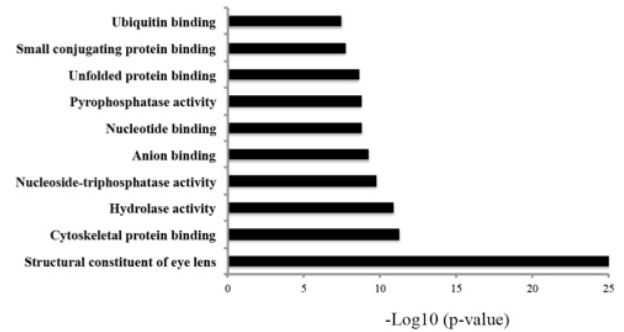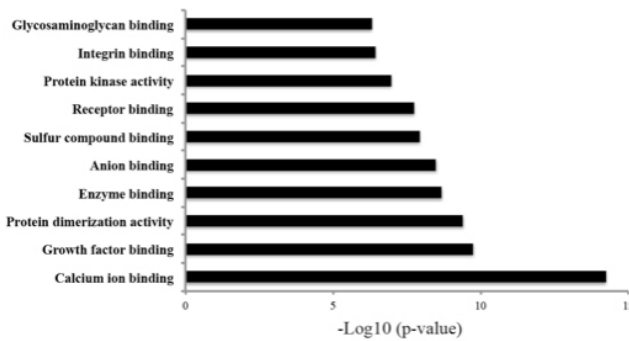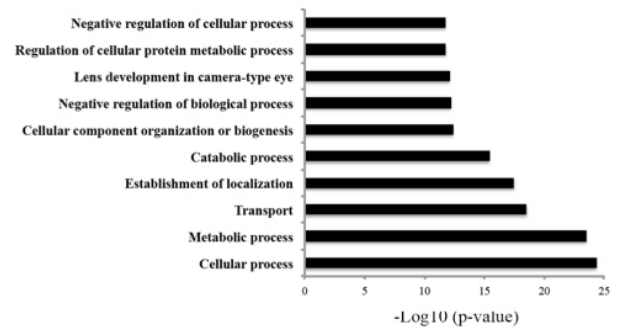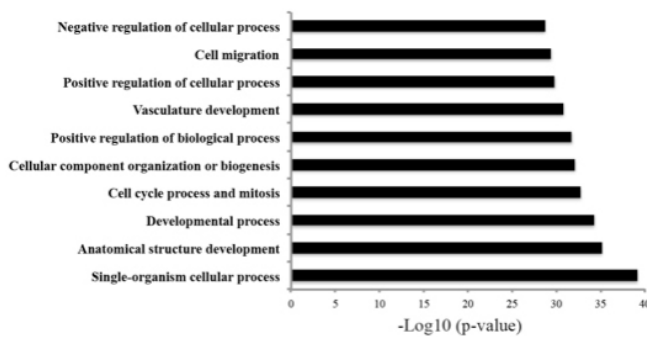

Figure 5. Gene Ontology enrichment analysis for genes upregulated in epithelial and genes upregulated in fiber cells. The top ten most enriched Gene Ontology (GO) terms are shown for the cellular component (**A**), molecular function (**B**), and biologic function (**C**).

receptors (Fgfrs) showed the highest gene expression levels in epithelial cells and fiber cells (Table 2). The abundant expression of Fgfrs is consistent with the known essential role for Fgfr signaling in lens development [72]. Of the four genes encoding Fgfrs capable of active tyrosine phosphorylation (*Fgfr1–4*), *Fgfr4* expression was barely detected in the lens RNA-Seq data. *Fgfr3* expression (the most highly

expressed of the *Fgfrs* in the lens) was upregulated in fiber cells as confirmed with our RT-qPCR (Figure 4A) and was shown in previous studies [73,74]. There was no significant difference in *Fgfr1* and *Fgfr2* expression levels between the epithelial and fiber cell fractions. Eph receptor A2 and the Met proto-oncogene were also upregulated in fiber cells. Most of the other tyrosine kinase receptors were expressed

more abundantly in the epithelial cells than in the fiber cells. These include discoidin domain receptor family member 1 (*Ddr1*), platelet derived growth factor receptors (*Pdgfra* and *Pdgfrb*), AXL receptor tyrosine kinase (*Axl*), Eph receptors (*Ephb2*, *Ephb4*, and *Ephb6*), and kinase insert domain protein receptor (*Kdr*).

Notch signaling: Notch signaling has been shown to regulate cell growth and differentiation in the mammalian lens [75,76]. We found that the majority of genes involved in Notch signaling pathways, including the Notch receptors (*Notch1, Notch2, Notch3* and *Notch4*), ligands (*Dll1, Dll4*, and *Jag2*), and effector genes (*Hes1* and *Hes5*) of this pathway, were expressed more abundantly in the epithelial cells, indicating that Notch signaling is more active in epithelial cells than in fiber cells (Table 3), consistent with previous functional studies [77,78].

Wnt signaling: Wnt signaling plays an important role in regulating eye development [79,80], including lens epithelial and fiber cell differentiation [81-83]. Our data showed that among ten Wnt receptors, five Wnt receptors (*Fzd1, Fzd2, Fzd4, Fzd7,* and *Fzd 8*) were significantly upregulated in the epithelial cells, while only *Fzd3* was slightly upregulated in the fiber cells (Table 4). *Fzd6* expression was not significantly different between the two cell types. Lens epithelial cells preferentially expressed the Wnt ligands, *Wnt5a* and *Wnt11*, more abundantly in epithelial cells, while fiber cells upregulated the expression of *Wnt7a* and *Wnt7b*. The lens epithelial cells and the fiber cells expressed the Wnt signaling effectors, *Dvl1, Dvl2,* and *Dvl3*, but only *Dvl2* exhibited differential expression with 2.47-fold more transcripts in the lens epithelium (Table 4).

Tgfβ superfamily receptors and their ligands: Bone morphogenetic protein (BMP) signals also play an essential role in lens development. In fact, lens induction requires BMP signaling [84,85], and BMP signaling interacts with FGF signaling to regulate cell cycle exit in primary lens fiber differentiation [86-88] and secondary lens fiber differentiation [89]. Of 21 transforming growth factor-β (Tgfβ) superfamily receptors, the fiber cells expressed bone morphogenetic protein receptor 1A (*Bmpr1a*) most abundantly and upregulated *Bmpr1b* (Appendix 6). The epithelial cells and the fiber cells expressed *Bmpr2* as well as activin A receptor IB and IIB (*Acrv2b* and *Acrv1b*). Most other Tgfβ receptors, including transforming growth factor-β receptor I (*Tgfbr1*), transforming growth factor-β receptor II (*Tgfbr2*), neural cell adhesion molecule 1 (*Ncam1*), activin A receptor type 1 (*Acvr1*), and endogen (*Eng*), were expressed more abundantly in epithelial cells than in fiber cells.

RNA-Seq of the newborn lenses detected a limited number of TGFβ superfamily ligand transcripts, and almost all of the differentially regulated transcripts were preferentially expressed in the epithelial cells (Appendix 6). Ligands showing the highest expression level are bone morphogenetic protein-1 (*Bmp1*), transforming growth factor-β2 (*Tgfb2*), bone morphogenetic protein-7 (*Bmp7*), and inhibin α (*Inha*). Other ligands, including *Bmp3, Bmp5, Bmp6, Bmp8a, Bmp8b, Bmp10,* and *Bmp15*, were not detected in either cell type. Overall, the upregulated expression of the Tgfβ superfamily receptors and ligands in the lens epithelial cells may indicate that Tgfβ signaling is more active in the epithelial cells than in the fiber cells.

Genes encoding DNA degradation and DNA repair enzymes: The final stages of lens fiber cell differentiation require nuclear degradation to promote lens transparency. The destruction of nuclear DNA requires DNases (DNases). Lens epithelial cells and lens fiber cells express transcripts for *DnaseI, Dnase2a,* and *Dnase2b*. Transcripts for *Dnase2b* outnumber the other two DNase transcripts in the lens. In particular, fiber cells upregulate DNase2b transcripts 100 fold relative to the lens epithelium (Table 5). The induction of DNase2b transcripts in fiber cells is consistent with the known importance of this enzyme for DNA destruction during fiber cell denucleation [28,45].

The failure to repair DNA double-stranded breaks (DSBs) can lead to programmed cell death [90]. Previous studies showed that this type of DNA damage was generated during normal fiber cell denucleation [91,92]. The possibility exists that reduced DNA repair activity plays a mechanistic role during lens fiber cell denucleation. Consistent with this view, transcripts of many key components of the DNA repair pathway were downregulated in the fiber cells (Table 5).

Genes involved in cell death and proteolysis: The induction of proteolytic pathways that trigger organelle loss during lens fiber cell differentiation bears some resemblance to the initiation of programmed cell death (apoptosis) [93,94]. The p53 family members (p53, p63, and p73) are tumor suppressor genes and regulate apoptosis and/or differentiation [95]. RNA-Seq analysis found p53 transcripts were more abundant in the lens than p63 or p73 transcripts. The number of p53 transcripts in the epithelial cells also exceeds the number in fiber cells (Appendix 7), consistent with previous evidence showing that p53 is expressed in epithelial cells and the lens fiber bow region [96]. The RNA-Seq data also revealed high expression of genes including the p53 regulator *Mdm2*, as well as several members of the Bcl family of apoptosis regulators in epithelial and fiber cells.

TABLE 2. EXPRESSION OF TYROSINE KINASE RECEPTORS IN EPITHELIAL CELLS AND FIBER CELLS.

Gene symbol	Description	F_RPKM	E_RPKM	F_counts	E_counts	FC*	P value
Fgfr3	Fibroblast growth factor receptor 3	121.86	135.8	9683.7	3695.2	2.62	3.85E-16
Fgfr2	Fibroblast growth factor receptor 2	30.09	74.3	1764.3	1624.6	1.09	1
Fgfr1	Fibroblast growth factor receptor 1	27.77	64.9	2228.2	1894.7	1.18	0.424884574
Ddr1	Discoidin domain receptor family, member 1	26.72	114.4	2436.2	3987.0	-1.64	4.74E-05
Epha2	Eph receptor A2	14.89	16.7	1223.4	496.3	2.46	0.000111118
Ephb3	Eph receptor B3	8.65	18.4	414.7	356.8	1.16	0.717922501
Ryk	Receptor-like tyrosine kinase	7.95	21.1	619.7	626.3	-1.01	1
Igf1r	Insulin-like growth factor 1 receptor	6.79	24.37	862.5	1108	-0.77	0.17
Insr	Insulin receptor	3.99	13.6	932.0	1144.6	-1.23	0.283411755
Pdgfra	Platelet derived growth factor receptor, alpha	3.91	96.0	720.3	6218.4	-8.63	1.22E-19
Met	Met proto-oncogene	2.92	0.6	162.8	27.3	5.95	3.26E-11
Erbb4	Epidermal growth factor receptor 4	2.23	6.8	663.2	821.7	-1.24	0.290412165
Epha5	Eph receptor A5	1.74	5.1	204.3	237.0	-1.16	0.993890715
Ror1	Receptor tyrosine kinase-like orphan receptor 1	1.53	14.7	246.2	885.9	-3.60	0.000237418
Ephb2	Eph receptor B2	1.47	7.7	204.0	391.8	-1.92	0.00033882
Ephb4	Eph receptor B4	1.07	21.6	119.3	867.7	-7.28	1.82E-35
Erbb2	Epidermal growth factor receptor 2	1.03	10.5	137.3	431.7	-3.15	5.82E-10
Ptk7	PTK7 protein tyrosine kinase 7	0.85	19.3	100.7	845.8	-8.40	6.54E-24
Tek	Endothelial-specific receptor tyrosine kinase	0.81	15.1	78.6	705.3	-8.97	1.06E-36
Ntrk2	Neurotrophic tyrosine kinase, receptor, type 2	0.38	5.1	75.5	398.2	-5.27	0.120798265
Ret	Ret proto-oncogene	0.36	0.2	73.3	15.9	4.61	6.03E-05
Pdgfrb	Platelet derived growth factor receptor, beta	0.30	6.8	45.1	377.4	-8.38	9.25E-24
Ephb1	Eph receptor B1	0.20	0.6	23.9	33.2	-1.39	1
Epha4	Eph receptor A4	0.16	0.5	27.8	33.6	-1.21	0.992147532
Axl	AXL receptor tyrosine kinase	0.13	10.7	6.1	356.0	-58.32	1.93E-24
Ephb6	Eph receptor B6	0.13	1.8	6.0	66.2	-11.04	6.75E-07
Erbb3	Epidermal growth factor receptor 3	0.05	0.3	9.2	14.8	-1.60	0.762104733
Ntrk1	Neurotrophic tyrosine kinase, receptor, type 1	0.04	0.4	6.6	9.2	-1.38	0.972770678
Epha7	Eph receptor A7	0.04	1.8	2.1	59.0	-28.45	1.29E-06
Epha8	Eph receptor A8	0.04	0.0	5.5	1.3	4.12	0.554412732
Musk	Muscle, skeletal, receptor tyrosine kinase	0.03	0.0	3.1	0.2	20.15	0.595515715
Epha10	Eph receptor A10	0.03	0.5	1.1	21.1	-19.10	0.004189458
Kdr	Kinase insert domain protein receptor	0.03	11.5	4.4	708.8	-161.25	1.15E-29

Gene symbol	Description	F_RPKM	E_RPKM	F_counts	E_counts	FC*	P value
Flt4	FMS-like tyrosine kinase 4	0.03	2.3	5.0	145.7	−29.20	4.33E-19
Egfr	Epidermal growth factor receptor	0.02	0.6	1.7	37.2	−22.34	2.53E-05
Tie1	Tyrosine kinase receptor 1	0.02	8.7	1.6	384.3	−237.81	5.23E-10
Ltk	Leukocyte tyrosine kinase	0.02	0.0	1.0	0.0	inf	0.992147532
Ror2	Receptor tyrosine kinase-like orphan receptor 2	0.02	2.3	1.6	91.5	−56.61	0.005034625
Epha3	Eph receptor A3	0.01	0.2	1.1	12.7	−11.52	0.087755426
Fgfr4	Fibroblast growth factor receptor 4	0.01	0.3	1.2	9.1	−7.76	0.232438017
Epha1	Eph receptor A1	0.01	0.7	1.5	24.3	−16.72	0.003364677
Flt1	FMS-like tyrosine kinase 1	0.00	8.5	0.6	566.8	−962.19	6.66E-83
Ntrk3	Neurotrophic tyrosine kinase, receptor, type 3	0.00	0.1	0.0	1.7	0.00	0.981954195
Epha6	Eph receptor A6	0.00	0.2	0.0	4.6	0.00	0.393262883
Ddr2	Discoidin domain receptor family, member 2	0.00	0.6	0.0	59.0	0.00	0.005585263

Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the fiber cells (F) with F counts and E counts representing normalized read counts from DESeq software from the fiber and epithelial cell samples, respectively. *Fold-change (FC) was based on the expression level in the fiber cells relative to the epithelial cells in normalized counts produced by DESeq. Negative values indicate expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less than 0.05 and more than a 1.5 fold-change in normalized read counts were considered as differential expression.

TABLE 3. EXPRESSIONS OF GENES INVOLVED IN NOTCH SIGNALING IN EPITHELIAL CELLS AND FIBER CELLS.

Gene function	Gene symbol	Description	F_RPKM	E_RPKM	F_counts	E_counts	FC*	P value
	Notch1	Notch gene homolog 1	2.84	12.25	333.87	910.85	−2.73	3.83E-12
Receptors	Notch3	Notch gene homolog 3	1.28	26.82	276.06	2084.18	−7.55	1.78E-52
	Notch2	Notch gene homolog 2	0.83	30.55	243.54	3164.57	−12.99	1.48E-42
	Notch4	Notch gene homolog 4	0.13	6.98	17.67	187.21	−10.60	3.66E-13
	Jag1	Jagged 1	77.73	109.40	11,686.66	5515.88	2.12	1.07E-09
	Jag2	Jagged 2	0.78	6.39	92.41	271.70	−2.94	0.00000087
Ligands	Dll3	Delta-like 3	0.20	0.45	8.33	9.42	−1.13	1
	Dll4	Delta-like 4	0.14	2.95	11.29	113.43	−10.04	0.035633554
	Dll1	Delta-like 1	0.06	1.72	1.59	43.01	−27.09	0.00000467
	Herpud1	Herpud family member 1	21.57	19.77	992.83	363.88	2.73	2.84E-10
	Herpud2	HERPUD family member 2	6.09	14.73	476.03	409.29	1.16	0.844553336
	Hes6	Hairy and enhancer of split 6	4.45	25.84	141.54	358.37	−2.53	0.0000121
	Hes1	Hairy and enhancer of split 1	2.18	40.05	50.54	559.74	−11.08	6.13E-18
Media-tors/ Effectors	Rbpj	Recombination signal binding protein for immunoglobulin kappa J region	0.35	1.36	56.67	117.19	−2.07	0.029792975
	Hes5	Hairy and enhancer of split 5	0.18	10.49	5.50	113.62	−20.66	0.0000279
	Hes7	Hairy and enhancer of split 7	0.01	0.55	0.49	10.79	−22.22	0.087807157
	Hes2	Hairy and enhancer of split 2	0.00	0.11	0.00	0.62	NA	1
	Hes3	Hairy and enhancer of split 3	0.00	0.04	0.00	0.66	NA	1
Inhibitor	Numb	Numb gene homolog	0.67	4.38	56.67	117.19	−2.07	0.029792975

Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the fiber cells (F) with F counts and E counts representing normalized read counts from DESeq software from the fiber and epithelial cell samples, respectively. *Fold-change (FC) was based on the expression level in the fiber cells relative to the epithelial cells in normalized counts produced by DESeq. Negative values indicate expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less than 0.05 and more than a 1.5 fold-change in normalized read counts were considered as differential expression.

TABLE 4. EXPRESSION OF GENES INVOLVED IN WNT SIGNALING IN EPITHELIAL CELLS AND FIBER CELLS.

Gene function	Gene symbol	Description	F_RPKM	E_RPKM	F_counts	E_counts	FC*	P value
	Fzd6	Frizzled homolog 6	9.34	24.33	1056.07	1006.92	1.05	0.959331326
	Fzd3	Frizzled homolog 3	3.57	6.24	411.68	260.93	1.58	0.020293911
	Fzd1	Frizzled homolog 1	2.46	64.81	302.20	2726.49	−9.02	0.000000128
	Fzd7	Frizzled homolog 7	0.74	17.61	98.76	834.31	−8.45	1.52E-27
	Fzd5	Frizzled homolog 5	0.50	1.18	93.53	88.04	1.06	0.992147532
Receptors	Fzd2	Frizzled homolog 2	0.29	21.30	29.30	761.85	−26.00	5.02E-42
	Fzd4	Frizzled homolog 4	0.28	2.54	28.52	97.89	−3.43	0.000469262
	Fzd8	Frizzled homolog 8	0.06	1.62	5.35	56.68	−10.59	0.0000111
	Fzd10	Frizzled homolog 10	0.03	0.16	2.74	4.82	−1.76	0.993884113
	Fzd9	Frizzled homolog 9	0.02	0.38	1.03	8.27	−8.06	0.307196403

Gene function	Gene symbol	Description	F_RPKM	E_RPKM	F_counts	E_counts	FC*	P value
	Wnt7b	Wingless-related MMTV integration site 7B	104.79	114.65	9978.36	3926.80	2.54	6.59E-12
	Wnt7a	Wingless-related MMTV integration site 7A	34.72	7.78	2233.87	219.48	10.18	4.96E-62
	Wnt5b	Wingless-related MMTV integration site 5B	15.96	29.92	1002.85	681.68	1.47	0.018690685
	Wnt9a	Wingless-type MMTV integration site 9A	0.55	1.22	53.18	42.70	1.25	0.841380211
	Wnt3	Wingless-related MMTV integration site 3	0.28	0.24	24.03	6.99	3.44	0.168242537
	Wnt16	Wingless-related MMTV integration site 16	0.23	0.17	8.83	3.36	2.63	0.504663541
	Wnt4	Wingless-related MMTV integration site 4	0.08	0.02	8.83	1.02	8.67	0.105123216
	Wnt5a	Wingless-related MMTV integration site 5A	0.07	5.81	10.48	251.72	−24.02	1.89E-28
	Wnt2b	Wingless related MMTV integration site 2b	0.06	0.73	6.51	27.07	−4.16	0.057763171
	Wnt9b	Wingless-type MMTV integration site 9B	0.04	0.05	4.91	2.32	2.12	0.89964175
Ligands	Wnt2	Wingless-related MMTV integration site 2	0.01	0.00	0.49	0.00	inf	1
	Wnt1	Wingless-related MMTV integration site 1	0.00	0.00	0.00	0.00	inf	NA
	Wnt3a	Wingless-related MMTV integration site 3A	0.00	0.00	0.59	0.00	inf	1
	Wnt6	Wingless-related MMTV integration site 6	0.00	0.04	0.00	0.66	0.00	1
	Wnt8a	Wingless-related MMTV integration site 8A	0.00	0.03	0.00	0.68	0.00	1
	Wnt8b	Wingless-related MMTV integration site 8b	0.00	0.00	0.00	0.00	inf	NA
	Wnt10a	Wingless related MMTV integration site 10a	0.00	0.07	0.00	2.38	0.00	0.886460331
	Wnt10b	Wingless related MMTV integration site 10b	0.00	0.00	0.00	0.00	inf	NA
	Wnt11	Wingless-related MMTV integration site 11	0.00	0.81	0.00	11.45	0.00	0.028073099

Gene function	Gene symbol	Description	F_RPKM	E_RPKM	F_counts	E_counts	FC*	P value
	Dvl3	Dishevelled 3, dsh homolog	9.05	23.07	651.84	616.48	1.06	0.934788241
Mediators	Dvl1	Dishevelled 1, dsh homolog	7.05	18.86	489.97	413.01	1.19	0.630996947
	Dvl2	Dishevelled 2, dsh homolog	2.94	18.57	199.59	492.67	−2.47	0.0000132

Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the fiber cells (F) with F counts and E counts representing normalized read counts from DESeq software from the fiber and epithelial cell samples, respectively. *Fold-change (FC) is based on expressions in the fiber cells relative to the epithelial cells in normalized counts. Negative values indicated expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less than 0.05 and more than a 1.5 fold-change in normalized read count were considered as differential expression.

TABLE 5. EXPRESSION OF GENES ENCODING DNA DEGRADATION ENZYMES AND GENES INVOLVED IN DNA REPAIR.

Gene function	Gene symbol	Description	F_RPKM	E_RPKM	F_counts	E_counts	FC*	P value
DNases	Dnase2b	DNase II beta	34.89	1.00	2539.30	25.21	100.73	1.42E-176
	Dnase2a	DNase II alpha	2.27	10.00	39.87	64.31	−1.61	0.297074881
	Dnase1	DNase I	1.94	2.81	43.39	43.67	−1.01	1
	Lig3	Ligase III, DNA	6.27	17.17	540.75	492.46	1.10	1
	H2afx	H2A histone family, member X	5.64	44.44	217.55	601.89	−2.77	2.35E-10
	Rad52	RAD52 homolog	3.69	18.37	98.93	197.52	−2.00	0.006096111
	Mre11a	Meiotic recombination 11 homolog A	3.26	7.88	255.29	205.85	1.24	0.495213491
	Lig1	Ligase I, DNA	2.18	34.39	59.61	392.69	−6.59	0.001185211
	Lig4	Ligase IV, DNA	1.41	1.70	160.55	66.73	2.41	0.003428077
DNA repairs	Rad50	RAD50 homolog	1.34	4.20	98.11	208.59	−2.13	0.03211083
	Nbn	Nibrin	1.09	9.82	77.74	255.87	−3.29	0.000000271
	Mdc1	Mediator of DNA damage checkpoint 1	1.00	5.60	104.95	346.84	−3.30	0.000375043
	Atm	Ataxia telangiectasia mutated homolog	0.99	6.17	176.91	272.31	−1.54	0.047464437
	Chek1	Checkpoint kinase 1 homolog	0.64	10.24	10.16	102.10	−10.05	1.92E-09
	Brca1	Breast cancer 1	0.42	1.87	65.72	104.73	−1.59	0.462151787
	Rad51	RAD51 homolog	0.27	4.07	10.52	85.09	−8.09	0.000000624
	Brca2	Breast cancer 2	0.21	1.38	63.81	142.65	−2.24	0.174088223

Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the fiber cells (F) with F counts and E counts representing normalized read counts from DESeq software from the fiber and epithelial cell samples, respectively. *Fold-change (FC) was based on the expression level in the fiber cells relative to the epithelial cells in normalized counts produced by DESeq. Negative values indicated expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less than 0.05 and more than a 1.5 fold-change in normalized read counts were considered as differential expression.

The lens also expresses genes involved in proteolytic pathways, including caspases, caspase inhibitors, autophagy-related genes, capthepsins, and calpains (Appendix 7). Among all caspases, the lens most highly expresses caspase-7, where fiber cells possessed increased caspase 7 transcripts relative to epithelial cells. Meanwhile, transcripts for caspase-2 and -9 were more abundant in the epithelial cells. Fiber cells also upregulated caspase inhibitors (IAPs), *Birc7* and *Birc2*, whereas the lens epithelium upregulated *Xiap* and *Birc5* expression. Epithelial cells strongly expressed cathepsin-L and cathepsin-F, while fiber cells preferentially expressed cathepsin-D and cathepsin-Z. Fiber cells also strongly upregulated transcripts of calpain-1, calpain-2, and calpain-3, suggesting that calpain-mediated protein degradation may play an important role in organelle loss during fiber cell differentiation.

Expression of genes encoding aquaporins, connexins, and intermediate filaments:

Aquaporins—Aquaporins are membrane water transport proteins that are found in many cell types [97]. Aquaporin-0, also known as the major intrinsic protein (*Mip*), is the most abundant protein in lens fiber cell membranes. The RNA-Seq data showed that transcripts for this lens-specific aquaporin outnumbered all other aquaporins in the lens with fiber cells expressing 60-fold more Aqp0 transcripts than epithelial cells (Appendix 8). The expression of aquaporin-5 (*Apq5*) and aquaporin-11 (*Aqp11*) was low in both cell types, while the lens epithelium preferentially expressed transcripts for aquaporin 1 (*Aqp1*), aquaporin 4 (*Aqp4*), and aquaporin-8 (*Aqp8*). Other aquaporin genes, including aquaporin-7 (*Aqp7*), aquaporin-2 (*Aqp2*), aquaporin-6 (*Aqp6*), aquaporin-9 (*Aqp9*), and aquaporin-12 (*Aqp12*), were minimally expressed or absent in the newborn mouse lens. Numerous mutations in *Aqp0* give rise to cataracts in humans and mice [98], and although *Aqp1* mutant mice display increased sensitivity to stress induced cataract [99], humans with *Aqp1* mutations have normal lenses. *Aqp5* null mice also fail to develop cataracts [100].

Gap junction proteins—Gap junction-mediated coupling in lens cells plays a particularly important role in lens homeostasis. In fact, lens transparency depends on the activity of sodium pumps, gap junctions, and aquaporins to maintain fluid transport in the lens [101]. Most gap junction genes exhibited higher expression in the fiber cells. The most abundantly expressed gap junction transcripts in the fiber cells included gap junction protein-α3 (*Gja3*), also known as connexin 46, gap junction protein-α8 (*Gja8*), also known as connexin 50, and gap junction protein-ε1 (*Gje1*), also known as connexin 23(Appendix 8). Transcripts for gap junction protein-α1 (*Gja1*), also known as connexin 43, gap junction protein gamma-1 (*Gjc1*), also known as connexin 45, and gap junction protein-α4 (*Gja4*), also known as connexin 37, were upregulated in the epithelial cells. Mutation of connexins, including Gje1, Gja3, and Gja8, causes microphthalmia and cataracts [102,103]. Lim2 is an integral protein found in cell gap junctions. A missense mutation of LIM2 in the human lens [104] and targeted disruption of *Lim2* in the mouse lens [105] also cause cataracts.

Intermediate filaments—Intermediate filaments (IFs) are a key component of the cytoskeleton of all vertebrate cells. The lens exclusively expresses a subset of specialized intermediate filament proteins known as bead filament proteins. The beaded filament proteins in the lens are commonly called filensin (*Bfsp1*) and phakinin (*Bfsp2*) [44]. The RNA-Seq data showed the highest expression of transcripts for *Bfsp1*, *Bfsp2*, vimentin (*Vim*), and synemin (*Synm*) in the fiber cells (Appendix 8). Previous studies reported that mouse *FVB/N* and 129 inbred strains harbor a natural 6-kb deletion mutation in *Bfsp2* gene. This deletion causes exon 1 to inappropriately splice to exon 3. This incorrect mRNA splicing changes the reading frame of the Bfsp2 transcript. As a result, a stop codon at position 2 of exon 3 in the Bfsp2 transcript prevents the translation of functional BFSP2 protein in the lens [106,107]. Consistent with these studies, we did not find any reads mapping to the exon 2 of *Bfsp2* gene in our *FVB/N* lens samples (data not shown). Other IFs were expressed at moderate levels, including lamin A (*Lmna*), lamin B1 (*Lmnb1*), lamin B2 (*Lmnb2*), and nestin (*Nes*), with *Lmnb1*, *Lmnb2*, and *Nes* expressed more in epithelial cells. There was low expression of several keratin genes in the lens, including keratin 10 (*Krt10*), keratin 18 (*Krt18*), keratin 19, (*Krt19*), and keratin 40 (*Krt40*).

Expression of lincRNA genes: LincRNAs are a group of newly identified non-coding RNAs that play an important role in gene transcription and protein translation [11]. The expression pattern and functional role of lincRNAs in the lens are largely unknown. Like mRNAs, lincRNA transcripts undergo polyadenylation. Thus, lincRNAs are subjected to RNA-Seq via polyA selection [11,108,109]. Among 1,746 annotated lincRNA genes, we detected the expression of 254 lincRNAs in the lens, of which 86 lincRNA genes were differentially expressed (32 genes upregulated in epithelial cells and 54 genes upregulated in fiber cells). In the top 30 most differentially expressed lincRNAs (ranked by p values, Figure 6), most lincRNA genes were upregulated in fiber cells, including *RP23–237H8.2*, *AC135859.1*, *AL663030.1*, *AC128663.1*, and *AC100730.1*. Little information exists concerning the specific functions of these lincRNAs.
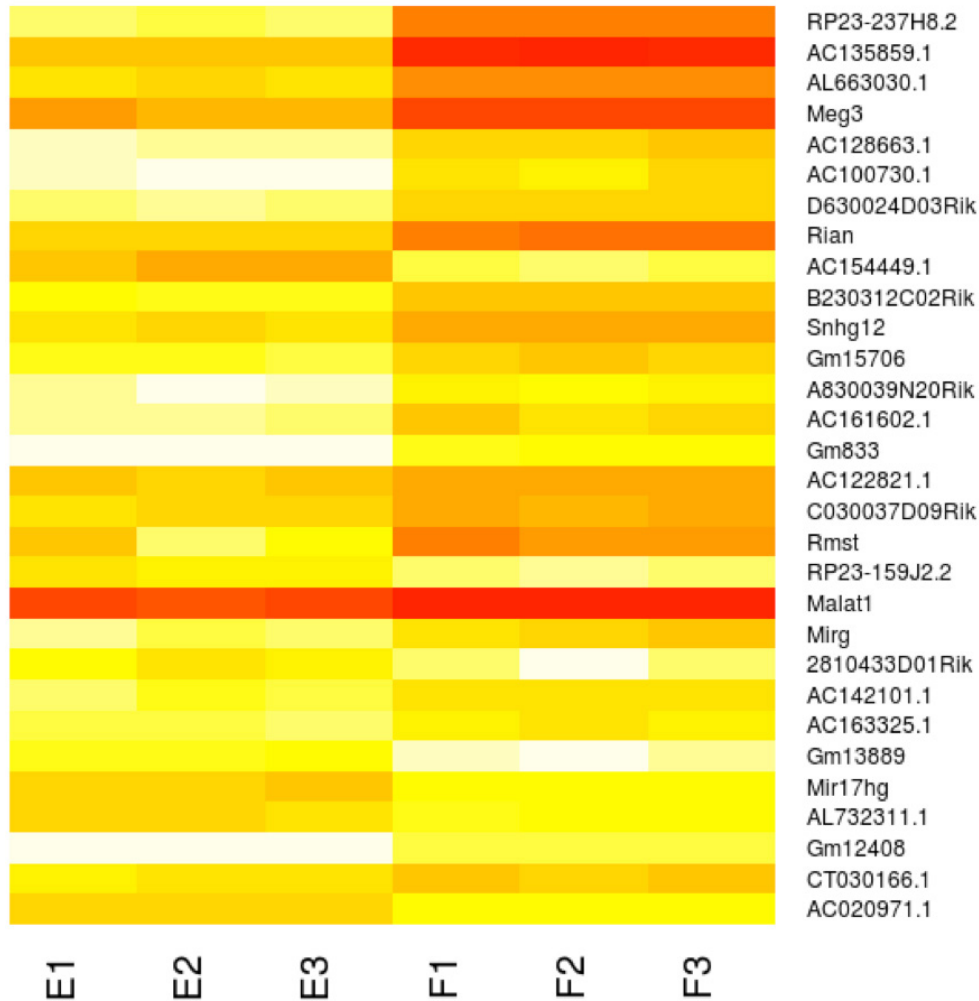
Figure 6. Top 30 most significantly differentially expressed lincRNAs between the epithelial and fiber cells ranked by adjusted p value. LincRNA genes with a p-adjusted value less than 0.05 and more than a 1.5-fold change in normalized read counts, with at least five unique mapped reads in either epithelial or fiber samples were considered differential expression. Sample abbreviations: E1, E2, and E3: epithelial replicates; F1, F2, and F3: fiber replicates.

LincRNAs upregulated in epithelial cells include *AC154449.1, RP23–159J2.2, 2810433D01Rik, Gm13889, Mir17hg, AL732311.1,* and *AC020971.1.*

Maternally expressed gene 3 (*Meg3*) and maternally expressed gene 8 (*Meg8* or *Rian*) are overlapping but non-identical transcripts found in the Dlk1-Dio3 imprinted region. *Meg3* and *Rian* were upregulated in fiber cells. *Meg3* is expressed in many normal tissues, including the eyes [110]. Previous studies suggested that *Meg3* acts as a tumor suppressor and regulates vascularization as well as pattern specification and cell differentiation during ear development [111,112]. Mice carrying a maternal *Meg3* deletion die before birth, but no eye phenotype was described [113]. Rian is highly expressed in various tissues in mouse embryos, including the brain, tongue, and liver [114], but the function of Rian is unknown, and, until now, the expression of Rian in the eye had not been reported.

Malat1 is one of the most abundant and conserved long non-coding RNAs. In situ hybridization showed that Malat1 is expressed in many tissues including the lens and in cancers [115,116]. RNA-Seq detected high expression of *Malat1* in epithelial cells and fiber cells. A recent study demonstrated that microRNA-9 targets Malat1 transcripts for degradation in the nucleus [117]. Surprisingly, three independent research groups that created targeted mutations in Malat1 failed to detect any obvious phenotypes [117-120].

RNA-Seq detected *Mirh17hg* preferentially in the lens epithelium. Previous studies showed that *Mir17hg* is the host gene for a polycistronic transcript containing the MIR17–92 cluster, a group of at least six miRNAs that may be involved in cell proliferation, differentiation, and cell survival [121-123]. However, the functional role of *Mir17hg* in the lens is unknown.

Conclusion: This first comprehensive transcriptome analysis with RNA-Seq in the ocular lens provides a valuable resource

for studying lens development, fiber differentiation, and lens pathogenesis. In addition to providing a platform for comparing lens epithelial cell and fiber cell protein coding gene expression, for the first time, the expression patterns of lincRNAs in the lens were characterized. However, the functional significance of these lincRNAs in lens development or physiology remains unknown. Further studies are needed to elucidate their functional significances in lens development and fiber cell differentiation.

APPENDIX 1: LIST OF SPECIFIC PRIMERS FOR GENES USED IN RT-QPCR.

To access the data, click or select the words "Appendix 1." This is a PDF file containing specific primers sequences for 19 genes used in RT-qPCR quantification.

APPENDIX 2. DIFFERENTIAL EXPRESSION OF ALL PROTEIN-CODING GENES.

To access the data, click or select the words "Appendix 2." This is an EXCEL file that contains the DESeq result of differential expression of total 22,601 known protein-coding genes. The table includes the means of normalized read counts of epithelial samples and fiber samples, fold change and p value.

APPENDIX 3. DIFFERENTIAL EXPRESSION OF LINCRNAS.

To access the data, click or select the words "Appendix 3." This is an EXCEL file that contains the DESeq result of differential expression of total 1746 known lincRNA genes. The table includes the means of normalized read counts of epithelial samples and fiber samples, fold change and p value.

APPENDIX 4. TOP 50 HIGHEST EXPRESSED GENES IN EPITHELIAL CELLS BASED ON RPKM VALUES.

To access the data, click or select the words "Appendix 4." Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the epithelial cells (E) with E counts and F counts representing normalized read counts from DESeq software from the epithelial and fiber cell samples, respectively. *Fold-change (FC) calculation was based on the expression level in the fiber cells relative to the epithelial cells in normalized read counts produced by DESeq. Negative values indicate expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less than 0.05 and more than a 1.5

fold-change in normalized read counts were considered as differential expression.

APPENDIX 5. TOP 50 HIGHEST EXPRESSED GENES IN FIBER CELLS BASED ON THE FIBER RPKM VALUES

To access the data, click or select the words "Appendix 5." Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the fiber cells (F) with F counts and E counts representing normalized read counts from DESeq software from the fiber and epithelial cell samples, respectively. *Fold-change (FC) was based on the expression level in the fiber cells relative to the epithelial cells in normalized counts produced by DESeq. Negative values indicate expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less than 0.05 and more than a 1.5 fold-change in normalized read counts were considered as differential expression.

APPENDIX 6. EXPRESSION OF GENES INVOLVED IN TGFB SIGNALING IN LENS EPITHELIAL CELLS AND FIBER CELLS

To access the data, click or select the words "Appendix 6." Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the fiber cells (F) with F counts and E counts representing normalized read counts from DESeq software from the fiber and epithelial cell samples, respectively. *Fold-change (FC) was based on the expression level in the fiber cells relative to the epithelial cells in normalized counts produced by DESeq. Negative values indicated expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less than 0.05 and more than a 1.5 fold-change in normalized read counts were considered as differential expression.

APPENDIX 7. EXPRESSION OF GENES INVOLVED IN APOPTOSIS AND PROTEIN DEGRADATION PATHWAYS

To access the data, click or select the words "Appendix 7." Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the fiber cells (F) with F counts and E counts representing normalized read counts from DESeq software from the fiber and epithelial cell samples, respectively. *Fold-change (FC) was based on the expression level in the fiber cells relative to the epithelial cells in normalized counts produced by DESeq. Negative values indicated expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less

than 0.05 and more than a 1.5 fold-change in normalized read counts were considered as differential expression.

APPENDIX 8. EXPRESSION OF GENES ENCODING FOR AQUAPORINS, CONNEXINS AND INTERMEDIATE FILAMENTS

To access the data, click or select the words "Appendix 8." Genes were ranked based on the RPKM (Reads Per Kilobase per Million mapped reads) values in the fiber cells (F) with F counts and E counts representing normalized read counts from DESeq software from the fiber and epithelial cell samples, respectively. *Fold-change (FC) was based on the expression level in the fiber cells relative to the epithelial cells in normalized counts produced by DESeq. Negative values indicated expression lower in the fiber cells compared with the epithelial cells. Genes with a p-adjusted value less than 0.05 and more than a 1.5 fold-change in normalized read counts were considered as differential expression.

REFERENCES

1. Wormstone IM, Wride MA. The ocular lens: a classic model for development, physiology and disease. Philos Trans R Soc Lond B Biol Sci 2011; 366:1190-2. [PMID: 21402579].

2. Lovicu FJ, Robinson ML. Lens fiber differentiation. In: Lovicu FJ, Robinson ML, editors. Development of the Ocular lens. New York, NY: Cambridge University Press; 2004. p. 214–44.

3. Bassnett S. On the mechanism of organelle degradation in the vertebrate lens. Exp Eye Res 2009; 88:133-9. [PMID: 18840431].

4. Pontoriero GF, Smith AN, Miller LA, 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. [PMID: 18996109].

5. Medina-Martinez O, Brownell I, Amaya-Manzanares F, Hu Q, Behringer RR, Jamrich M. Severe defects in proliferation and differentiation of lens cells in Foxe3 null mice. Mol Cell Biol 2005; 25:8854-63. [PMID: 16199865].

6. Shaham O, Smith AN, Robinson ML, Taketo MM, Lang RA, Ashery-Padan R. Pax6 is essential for lens fiber cell differentiation. Development 2009; 136:2567-78. [PMID: 19570848].

7. FitzGerald PG. Lens intermediate filaments. Exp Eye Res 2009; 88:165-72. [PMID: 19071112].

8. Graw J. Genetics of crystallins: cataract and beyond. Exp Eye Res 2009; 88:173-89. [PMID: 19007775].

9. Cvekl A, Duncan MK. Genetic and epigenetic mechanisms of gene regulation during lens development. Prog Retin Eye Res 2007; 26:555-97. [PMID: 17905638].

10. Hung T, Chang HY. Long noncoding RNA in genome regulation Prospects and mechanisms. RNA Biol 2010; 7:582-5. [PMID: 20930520].

11. Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. Cell 2013; 154:26-46. [PMID: 23827673].

12. Hoheisel JD. Microarray technology: beyond transcript profiling and genotype analysis. Nat Rev Genet 2006; 7:200-10. [PMID: 16485019].

13. Mansergh FC, Wride MA, Walker VE, Adams S, Hunter SM, Evans MJ. Gene expression changes during cataract progression in Sparc null mice: differential regulation of mouse globins in the lens. Mol Vis 2004; 10:490-511. [PMID: 15303089].

14. Wride MA, Mansergh FC, Adams S, Everitt R, Minnema SE, Rancourt DE, Evans MJ. Expression profiling and gene discovery in the mouse lens. Mol Vis 2003; 9:360-96. [PMID: 12942050].

15. Ivanov D, Dvoriantchikova G, Pestova A, Nathanson L, Shestopalov VI. Microarray analysis of fiber cell maturation in the lens. FEBS Lett 2005; 579:1213-9. [PMID: 15710416].

16. Segev F, Mor O, Segev A, Belkin M, Assia EI. Downregulation of gene expression in the ageing lens: a possible contributory factor in senile cataract. Eye (Lond) 2005; 19:80-5. [PMID: 15105821].

17. Hawse JR, Hejtmancik JF, Horwitz J, Kantorow M. Identification and functional clustering of global gene expression differences between age-related cataract and clear human lenses and aged human lenses. Exp Eye Res 2004; 79:935-40. [PMID: 15642332].

18. Ruotolo R, Grassi F, Percudani R, Rivetti C, Martorana D, Maraini G, Ottonello S. Gene expression profiling in human age-related nuclear cataract. Mol Vis 2003; 9:538-48. [PMID: 14551529].

19. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. Cell 2013; 155:27-38. [PMID: 24074859].

20. Draghici S, Khatri P, Eklund AC, Szallasi Z. Reliability and reproducibility issues in DNA microarray measurements. Trends Genet 2006; 22:101-9. [PMID: 16380191].

21. Forster T, Roy D, Ghazal P. Experiments using microarray technology: limitations and standard operating procedures. J Endocrinol 2003; 178:195-204. [PMID: 12904167].

22. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 2009; 10:57-63. [PMID: 19015660].

23. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, Chen ZH, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 2011; 29:644-52. [PMID: 21572440].

24. Martin JA, Wang Z. Next-generation transcriptome assembly. Nat Rev Genet 2011; 12:671-82. [PMID: 21897427].

25. Ramaswami G, Zhang R, Piskol R, Keegan LP, Deng P, O'Connell MA, Li JB. Identifying RNA editing sites using RNA sequencing data alone. Nat Methods 2013; 10:128-32. [PMID: 23291724].

26. Manthey AL, Lachke SA, Fitzgerald PG, Mason RW, Scheiblin DA, McDonald JH, Duncan MK. Loss of Sip1 leads to migration defects and retention of ectodermal markers during lens development. Mech Dev 2014; 131:86-110. [PMID: 24161570].

27. WolfLGaoCSGuetaKXieQChevallierTPodduturiNRSunJConteIZelenkaPSAshery-PadanRZavadilJCveklA. Identification and characterization of FGF2-dependent mRNA: microRNA networks during lens fiber cell differentiation. G3 (Bethesda). 2013;3:2239-55

28. Nakahara M, Nagasaka A, Koike M, Uchida K, Kawane K, Uchiyama Y, Nagata S. Degradation of nuclear DNA by DNase II-like acid DNase in cortical fiber cells of mouse eye lens. FEBS J 2007; 274:3055-64. [PMID: 17509075].

29. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics 2011; 27:863-4. [PMID: 21278185].

30. Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, Gil L, Garcia-Giron C, Gordon L, Hourlier T, Hunt S, Juettemann T, Kahari AK, Keenan S, Komorowska M, Kulesha E, Longden I, Maurel T, McLaren WM, Muffato M, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS, Ritchie GR, Ruffier M, Schuster M, Sheppard D, Sobral D, Taylor K, Thormann A, Trevanion S, White S, Wilder SP, Aken BL, Birney E, Cunningham F, Dunham I, Harrow J, Herrero J, Hubbard TJ, Johnson N, Kinsella R, Parker A, Spudich G, Yates A, Zadissa A, Searle SM. Ensembl 2013. Nucleic Acids Res 2013; 41:Database issueD48-55. [PMID: 23203987].

31. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 2010; 26:873-81. [PMID: 20147302].

32. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 2010; 28:511-5. [PMID: 20436464].

33. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol 2010; 11:R106-[PMID: 20979621].

34. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Gene ontology analysis for RNA-Seq: accounting for selection bias. Genome Biol 2010; 11:R14-[PMID: 20132535].

35. Wistow G, Bernstein SL, Wyatt MK, Behal A, Touchman JW, Bouffard G, Smith D, Peterson K. Expressed sequence tag analysis of adult human lens for the NEIBank Project: over 2000 non-redundant transcripts, novel genes and splice variants. Mol Vis 2002; 8:171-84. [PMID: 12107413].

36. Tang Y, Crowley TE, Kumar NM. Global gene expression analysis of lenses from different mouse strains and in the alpha3Cx46 knockout mouse. Mol Vis 2010; 16:113-21. [PMID: 20104256].

37. Wu F, Lee S, Schumacher M, Jun A, Chakravarti S. Differential gene expression patterns of the developing and adult mouse cornea compared to the lens and tendon. Exp Eye Res 2008; 87:214-25. [PMID: 18582462].

38. Lachke SA, Ho JW, Kryukov GV, O'Connell DJ, Aboukhalil A, Bulyk ML, Park PJ, Maas RL. iSyTE: integrated Systems Tool for Eye gene discovery. Invest Ophthalmol Vis Sci 2012; 53:1617-27. [PMID: 22323457].

39. Beebe DC. The control of lens growth: relationship to secondary cataract. Acta Ophthalmol Suppl 1992; 20553-7. [PMID: 1332414].

40. Zhou M, Leiberman J, Xu J, Lavker RM. A hierarchy of proliferative cells exists in mouse lens epithelium: Implications for lens maintenance. Invest Ophthalmol Vis Sci 2006; 47:2997-3003. [PMID: 16799045].

41. Zelenka PS, Gao CY, Rampalli A, Arora J, Chauthaiwale V, He HY. Cell cycle regulation in the lens: Proliferation, quiescence, apoptosis and differentiation. Prog Retin Eye Res 1997; 16:303-22. .

42. Wride MA. Lens fibre cell differentiation and organelle loss: many paths lead to clarity. Philos Trans R Soc Lond B Biol Sci 2011; 366:1219-33. [PMID: 21402582].

43. Churchill A, Graw J. Clinical and experimental advances in congenital and paediatric cataracts. Philos Trans R Soc Lond B Biol Sci 2011; 366:1234-49. [PMID: 21402583].

44. Song S, Landsbury A, Dahm R, Liu Y, Zhang Q, Quinlan RA. Functions of the intermediate filament cytoskeleton in the eye lens. J Clin Invest 2009; 119:1837-48. [PMID: 19587458].

45. Nishimoto S, Kawane K, Watanabe-Fukunaga R, Fukuyama H, Ohsawa Y, Uchiyama Y, Hashida N, Ohguro N, Tano Y, Morimoto T, Fukuda Y, Nagata S. Nuclear cataract caused by a lack of DNA degradation in the mouse eye lens. Nature 2003; 424:1071-4. [PMID: 12944971].

46. Lachke SA, Alkuraya FS, Kneeland SC, Ohn T, Aboukhalil A, Howell GR, Saadi I, Cavallesco R, Yue Y, Tsai AC, Nair KS, Cosma MI, Smith RS, Hodges E, Alfadhli SM, Al-Hajeri A, Shamseldin HE, Behbehani A, Hannon GJ, Bulyk ML, Drack AV, Anderson PJ, John SW, Maas RL. Mutations in the RNA

granule component TDRD7 cause cataract and glaucoma. Science 2011; 331:1571-6. [PMID: 21436445].

47. Bu L, Jin Y, Shi Y, Chu R, Ban A, Eiberg H, Andres L, Jiang H, Zheng G, Qian M, Cui B, Xia Y, Liu J, Hu L, Zhao G, Hayden MR, Kong X. Mutant DNA-binding domain of HSF4 is associated with autosomal dominant lamellar and Marner cataract. Nat Genet 2002; 31:276-8. [PMID: 12089525].

48. Fujimoto M, Izu H, Seki K, Fukuda K, Nishida T, Yamada S, Kato K, Yonemura S, Inouye S, Nakai A. HSF4 is required for normal cell growth and differentiation during mouse lens development. EMBO J 2004; 23:4297-306. [PMID: 15483628].

49. Shi X, Cui B, Wang Z, Weng L, Xu Z, Ma J, Xu G, Kong X, Hu L. Removal of Hsf4 leads to cataract development in mice through down-regulation of gamma S-crystallin and Bfsp expression. BMC Mol Biol 2009; 10:10-[PMID: 19224648].

50. Cui X, Wang L, Zhang J, Du R, Liao S, Li D, Li C, Ke T, Li DW, Huang H, Yin Z, Tang Z, Liu M. HSF4 regulates DLAD expression and promotes lens de-nucleation. Biochim Biophys Acta 2013; 1832:1167-72. [PMID: 23507146].

51. Nowak RB, Fowler VM. Tropomodulin 1 constrains fiber cell geometry during elongation and maturation in the lens cortex. J Histochem Cytochem 2012; 60:414-27. [PMID: 22473940].

52. Gokhin DS, Nowak RB, Kim NE, Arnett EE, Chen AC, Sah RL, Clark JI, Fowler VM. Tmod1 and CP49 synergize to control the fiber cell geometry, transparency, and mechanical stiffness of the mouse lens. PLoS ONE 2012; 7:e48734-[PMID: 23144950].

53. Tjoelker LW, Wilder C, Eberhardt C, Stafforini DM, Dietsch G, Schimpf B, Hooper S, Le Trong H, Cousens LS, Zimmerman GA, Yamada Y, McIntyre TM, Prescott SM, Gray PW. Anti-inflammatory properties of a platelet-activating factor acetylhydrolase. Nature 1995; 374:549-53. [PMID: 7700381].

54. Lp PLASC, Thompson A, Gao P, Orfei L, Watson S, Di Angelantonio E, Kaptoge S, Ballantyne C, Cannon CP, Criqui M, Cushman M, Hofman A, Packard C, Thompson SG, Collins R, Danesh J. Lipoprotein-associated phospholipase A(2) and risk of coronary disease, stroke, and mortality: collaborative analysis of 32 prospective studies. Lancet 2010; 375:1536-44. [PMID: 20435228].

55. Zalewski A, Macphee C. Role of lipoprotein-associated phospholipase A2 in atherosclerosis: biology, epidemiology, and possible therapeutic target. Arterioscler Thromb Vasc Biol 2005; 25:923-31. [PMID: 15731492].

56. Jiang Z, Fehrenbach ML, Ravaioli G, Kokalari B, Redai IG, Sheardown SA, Wilson S, Macphee C, Haczku A. The effect of lipoprotein-associated phospholipase A2 deficiency on pulmonary allergic responses in Aspergillus fumigatus sensitized mice. Respir Res 2012; 13:100-[PMID: 23140447].

57. Landgren H, Blixt A, Carlsson P. Persistent FoxE3 expression blocks cytoskeletal remodeling and organelle degradation during lens fiber differentiation. Invest Ophthalmol Vis Sci 2008; 49:4269-77. [PMID: 18539941].

58. Nielsen PJ, Lorenz B, Muller AM, Wenger RH, Brombacher F, Simon M, von der Weid T, Langhorne WJ, Mossmann H, Kohler G. Altered erythrocytes and a leaky block in B-cell development in CD24/HSA-deficient mice. Blood 1997; 89:1058-67. [PMID: 9028339].

59. Lorén CE, Schrader JW, Ahlgren U, Gunhaga L. FGF signals induce Caprin2 expression in the vertebrate lens. Differentiation 2009; 77:386-94. [PMID: 19275872].

60. Zimmerman AW, Veerkamp JH. Fatty-acid-binding proteins do not protect against induced cytotoxicity in a kidney cell model. Biochem J 2001; 360:159-65. [PMID: 11696003].

61. Owada Y, Takano H, Yamanaka H, Kobayashi H, Sugitani Y, Tomioka Y, Suzuki I, Suzuki R, Terui T, Mizugaki M, Tagami H, Noda T, Kondo H. Altered water barrier function in epidermal-type fatty acid binding protein-deficient mice. J Invest Dermatol 2002; 118:430-5. [PMID: 11874481].

62. DiMauro S, Miranda AF, Khan S, Gitlin K, Friedman R. Human muscle phosphoglycerate mutase deficiency: newly discovered metabolic myopathy. Science 1981; 212:1277-9. [PMID: 6262916].

63. Gabriel LA, Wang LW, Bader H, Ho JC, Majors AK, Hollyfield JG, Traboulsi EI, Apte SS. ADAMTSL4, a secreted glycoprotein widely distributed in the eye, binds fibrillin-1 microfibrils and accelerates microfibril biogenesis. Invest Ophthalmol Vis Sci 2012; 53:461-9. [PMID: 21989719].

64. Ahram D, Sato TS, Kohilan A, Tayeh M, Chen S, Leal S, Al-Salem M, El-Shanti H. A homozygous mutation in ADAMTSL4 causes autosomal-recessive isolated ectopia lentis. Am J Hum Genet 2009; 84:274-8. [PMID: 19200529].

65. Uchida N, Shimamura K, Miyatani S, Copeland NG, Gilbert DJ, Jenkins NA, Takeichi M. Mouse alpha N-catenin: two isoforms, specific expression in the nervous system, and chromosomal localization of the gene. Dev Biol 1994; 163:75-85. [PMID: 8174789].

66. Pokutta S, Drees F, Yamada S, Nelson WJ, Weis WI. Biochemical and structural analysis of alpha-catenin in cell-cell contacts. Biochem Soc Trans 2008; 36:141-7. [PMID: 18363554].

67. Park C, Falls W, Finger JH, Longo-Guess CM, Ackerman SL. Deletion in Catna2, encoding alpha N-catenin, causes cerebellar and hippocampal lamination defects and impaired startle modulation. Nat Genet 2002; 31:279-84. [PMID: 12089526].

68. Ito S, Fujimori T, Hayashizaki Y, Nabeshima Y. Identification of a novel mouse membrane-bound family 1 glycosidase-like protein, which carries an atypical active site structure. Biochim Biophys Acta 2002; 1576:341-5. [PMID: 12084582].

69. Fon Tacer K, Bookout AL, Ding X, Kurosu H, John GB, Wang L, Goetz R, Mohammadi M, Kuro-o M, Mangelsdorf DJ, Kliewer SA. Research resource: Comprehensive expression atlas of the fibroblast growth factor system in adult mouse. Mol Endocrinol 2010; 24:2050-64. [PMID: 20667984].

70. Mustafi D, Kevany BM, Bai X, Maeda T, Sears JE, Khalil AM, Palczewski K. Evolutionarily conserved long intergenic

non-coding RNAs in the eye. Hum Mol Genet 2013; 22:2992-3002. [PMID: 23562822].

71. Ogino H, Ochi H, Reza HM, Yasuda K. Transcription factors involved in lens development from the preplacodal ectoderm. Dev Biol 2012; 363:333-47. [PMID: 22269169].

72. Robinson ML. An essential role for FGF receptor signaling in lens development. Semin Cell Dev Biol 2006; 17:726-40. [PMID: 17116415].

73. Zhao H, Yang TY, Madakashira BP, Thiels CA, Bechtle CA, Garcia CM, Zhang H, Yu K, Ornitz DM, Beebe DC, Robinson ML. Fibroblast growth factor receptor signaling is essential for lens fiber cell differentiation. Dev Biol 2008; 318:276-88. [PMID: 18455718].

74. Govindarajan V, Overbeek PA. Secreted FGFR3, but not FGFR1, inhibits lens fiber differentiation. Development 2001; 128:1617-27. [PMID: 11290300].

75. Rowan S, Conley KW, Le TT, Donner AL, Maas RL, Brown NL. Notch signaling regulates growth and differentiation in the mammalian lens. Dev Biol 2008; 321:111-22. [PMID: 18588871].

76. Jia J, Lin M, Zhang L, York JP, Zhang P. The Notch signaling pathway controls the size of the ocular lens by directly suppressing p57Kip2 expression. Mol Cell Biol 2007; 27:7236-47. [PMID: 17709399].

77. Le TT, Conley KW, Mead TJ, Rowan S, Yutzey KE, Brown NL. Requirements for Jag1-Rbpj mediated Notch signaling during early mouse lens development. Dev Dyn 2012; 241:493-504. [PMID: 22275127].

78. Saravanamuthu SS, Le TT, Gao CY, Cojocaru RI, Pandiyan P, Liu C, Zhang J, Zelenka PS, Brown NL. Conditional ablation of the Notch2 receptor in the ocular lens. Dev Biol 2012; 362:219-29. [PMID: 22173065].

79. Grocott T, Johnson S, Bailey AP, Streit A. Neural crest cells organize the eye via TGF-beta and canonical Wnt signalling. Nat Commun 2011; 2:265-[PMID: 21468017].

80. Fuhrmann S. Wnt signaling in eye organogenesis. Organogenesis 2008; 4:60-7. [PMID: 19122781].

81. Stump RJ, Ang S, Chen Y, von Bahr T, Lovicu FJ, Pinson K, de Iongh RU, Yamaguchi TP, Sassoon DA, McAvoy JW. A role for Wnt/beta-catenin signaling in lens epithelial differentiation. Dev Biol 2003; 259:48-61. [PMID: 12812787].

82. Chen Y, Stump RJW, Lovicu FJ, Shimono A, McAvoy JW. Wnt signaling is required for organization of the lens fiber cell cytoskeleton and development of lens three-dimensional architecture. Dev Biol 2008; 324:161-76. [PMID: 18824165].

83. Antosova B, Smolikova J, Borkovcova R, Strnad H, Lachova J, Machon O, Kozmik Z. Ectopic Activation of Wnt/beta-Catenin Signaling in Lens Fiber Cells Results in Cataract Formation and Aberrant Fiber Cell Differentiation. PLoS ONE 2013; 8:e78279-[PMID: 24205179].

84. Furuta Y, Hogan BL. BMP4 is essential for lens induction in the mouse embryo. Genes Dev 1998; 12:3764-75. [PMID: 9851982].

85. Sjödal M, Edlund T, Gunhaga L. Time of exposure to BMP signals plays a key role in the specification of the olfactory and lens placodes ex vivo. Dev Cell 2007; 13:141-9. [PMID: 17609116].

86. Jarrin M, Pandit T, Gunhaga L. A balance of FGF and BMP signals regulates cell cycle exit and Equarin expression in lens cells. Mol Biol Cell 2012; 23:3266-74. [PMID: 22718906].

87. Faber SC, Robinson ML, Makarenkova HP, Lang RA. Bmp signaling is required for development of primary lens fiber cells. Development 2002; 129:3727-37. [PMID: 12117821].

88. Rajagopal R, Huang J, Dattilo LK, Kaartinen V, Mishina Y, Deng CX, Umans L, Zwijsen A, Roberts AB, Beebe DC. The type I BMP receptors, Bmpr1a and Acvr1, activate multiple signaling pathways to regulate lens formation. Dev Biol 2009; 335:305-16. [PMID: 19733164].

89. Boswell BA, Overbeek PA, Musil LS. Essential role of BMPs in FGF-induced secondary lens fiber differentiation. Dev Biol 2008; 324:202-12. [PMID: 18848538].

90. Shrivastav M, De Haro LP, Nickoloff JA. Regulation of DNA double-strand break repair pathway choice. Cell Res 2008; 18:134-47. [PMID: 18157161].

91. Wang WL, Li Q, Xu J, Cvekl A. 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. Mol Biol Cell 2010; 21:2453-68. [PMID: 20484573].

92. Yang YG, Frappart PO, Frappart L, Wang ZQ, Tong WM. A novel function of DNA repair molecule Nbs1 in terminal differentiation of the lens fibre cells and cataractogenesis. DNA Repair (Amst) 2006; 5:885-93. [PMID: 16790366].

93. Dahm R. Lens fibre cell differentiation - A link with apoptosis? Ophthalmic Res 1999; 31:163-83. [PMID: 10224500].

94. Wride MA. Lens fibre cell differentiation and organelle loss: many paths lead to clarity. Philos Trans R Soc Lond B Biol Sci 2011; 366:1219-33. [PMID: 21402582].

95. Levrero M, De Laurenzi V, Costanzo A, Gong J, Wang JY, Melino G. The p53/p63/p73 family of transcription factors: overlapping and distinct functions. J Cell Sci 2000; 113:1661-70. [PMID: 10769197].

96. Ayala M, Strid H, Jacobsson U, Soderberg PG. p53 expression and apoptosis in the lens after ultraviolet radiation exposure. Invest Ophthalmol Vis Sci 2007; 48:4187-91. [PMID: 17724205].

97. Verkman AS, Mitra AK. Structure and function of aquaporin water channels. Am J Physiol Renal Physiol 2000; 278:F13-28. [PMID: 10644652].

98. Schey KL, Wang Z. J LW, Qi Y. Aquaporins in the eye: expression, function, and roles in ocular disease. Biochim Biophys Acta 2014; 1840:1513-23. [PMID: 24184915].

99. Ruiz-Ederra J, Verkman AS. Accelerated cataract formation and reduced lens epithelial water permeability in aquaporin-1-deficient mice. Invest Ophthalmol Vis Sci 2006; 47:3960-7. [PMID: 16936111].

100. Kumari SS, Varadaraj M, Yerramilli VS, Menon AG, Varadaraj K. Spatial expression of aquaporin 5 in mammalian cornea and lens, and regulation of its localization by phosphokinase A. Mol Vis 2012; 18:957-67. [PMID: 22550388].

101. Donaldson PJ, Musil LS, Mathias RT. Point: A critical appraisal of the lens circulation model–an experimental paradigm for understanding the maintenance of lens transparency? Invest Ophthalmol Vis Sci 2010; 51:2303-6. [PMID: 20435604].

102. Li L, Cheng C, Xia CH, White TW, Fletcher DA, Gong X. Connexin mediated cataract prevention in mice. PLoS ONE 2010; 5:e12624-[PMID: 20844585].

103. Rong P, Wang X, Niesman I, Wu Y, Benedetti LE, Dunia I, Levy E, Gong X. Disruption of Gja8 (alpha8 connexin) in mice leads to microphthalmia associated with retardation of lens growth and lens fiber maturation. Development 2002; 129:167-74. [PMID: 11782410].

104. Ponnam SP, Ramesha K, Tejwani S, Matalia J, Kannabiran C. A missense mutation in LIM2 causes autosomal recessive congenital cataract. Mol Vis 2008; 14:1204-8. [PMID: 18596884].

105. Shiels A, King JM, Mackay DS, Bassnett S. Refractive defects and cataracts in mice lacking lens intrinsic membrane protein-2. Invest Ophthalmol Vis Sci 2007; 48:500-8. [PMID: 17251442].

106. Simirskii VN, Lee RS, Wawrousek EF, Duncan MK. Inbred FVB/N mice are mutant at the cp49/Bfsp2 locus and lack beaded filament proteins in the lens. Invest Ophthalmol Vis Sci 2006; 47:4931-4. [PMID: 17065509].

107. Sandilands A, Wang X, Hutcheson AM, James J, Prescott AR, Wegener A, Pekny M, Gong X, Quinlan RA. Bfsp2 mutation found in mouse 129 strains causes the loss of CP49 and induces vimentin-dependent changes in the lens fibre cell cytoskeleton. Exp Eye Res 2004; 78:109-23. [PMID: 14667833].

108. Hangauer MJ, Vaughn IW, McManus MT. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. PLoS Genet 2013; 9:e1003569-[PMID: 23818866].

109. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011; 25:1915-27. [PMID: 21890647].

110. Manji SS, Sorensen BS, Klockars T, Lam T, Hutchison W, Dahl HH. Molecular characterization and expression of maternally expressed gene 3 (Meg3/Gtl2) RNA in the mouse inner ear. J Neurosci Res 2006; 83:181-90. [PMID: 16342203].

111. Zhou Y, Zhang X, Klibanski A. MEG3 noncoding RNA: a tumor suppressor. J Mol Endocrinol 2012; 48:R45-53. [PMID: 22393162].

112. Zhou Y, Zhong Y, Wang Y, Zhang X, Batista DL, Gejman R, Ansell PJ, Zhao J, Weng C, Klibanski A. Activation of p53 by MEG3 non-coding RNA. J Biol Chem 2007; 282:24731-42. [PMID: 17569660].

113. Zhou Y, Cheunsuchon P, Nakayama Y, Lawlor MW, Zhong Y, Rice KA, Zhang L, Zhang X, Gordon FE, Lidov HG, Bronson RT, Klibanski A. Activation of paternally expressed genes and perinatal death caused by deletion of the Gtl2 gene. Development 2010; 137:2643-52. [PMID: 20610486].

114. Gu T, He H, Han Z, Zeng T, Huang Z, Liu Q, Gu N, Chen Y, Sugimoto K, Jiang H, Wu Q. Expression of macro non-coding RNAs Meg8 and Irm in mouse embryonic development. Acta Histochem 2012; 114:392-9. [PMID: 21855964].

115. Nakagawa S, Ip JY, Shioi G, Tripathi V, Zong X, Hirose T, Prasanth KV. Malat1 is not an essential component of nuclear speckles in mice. RNA 2012; 18:1487-99. [PMID: 22718948].

116. Tripathi V, Shen Z, Chakraborty A, Giri S, Freier SM, Wu X, Zhang Y, Gorospe M, Prasanth SG, Lal A, Prasanth KV. Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. PLoS Genet 2013; 9:e1003368-[PMID: 23555285].

117. Leucci E, Patella F, Waage J, Holmstrom K, Lindow M, Porse B, Kauppinen S, Lund AH. microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus. Sci Rep 2013; 3:2535-[PMID: 23985560].

118. Zhang B, Arun G, Mao YS, Lazar Z, Hung GN, Bhattacharjee G, Xiao XK, Booth CJ, Wu J, Zhang CL, Spector DL. The lncRNA Malat1 Is Dispensable for Mouse Development but Its Transcription Plays a cis-Regulatory Role in the Adult. Cell Reports 2012; 2:111-23. [PMID: 22840402].

119. Eißmann M, Gutschner T, Hammerle M, Gunther S, Caudron-Herger M, Gross M, Schirmacher P, Rippe K, Braun T, Zornig M, Diederichs S. Loss of the abundant nuclear non-coding RNA MALAT1 is compatible with life and development. RNA Biol 2012; 9:1076-87. [PMID: 22858678].

120. Nakagawa S, Ip JY, Shioi G, Tripathi V, Zong X, Hirose T, Prasanth KV. Malat1 is not an essential component of nuclear speckles in mice. RNA 2012; 18:1487-99. [PMID: 22718948].

121. de Pontual L, Yao E, Callier P, Faivre L, Drouin V, Cariou S, Van Haeringen A, Genevieve D, Goldenberg A, Oufadem M, Manouvrier S, Munnich A, Vidigal JA, Vekemans M, Lyonnet S, Henrion-Caude A, Ventura A, Amiel J. Germline deletion of the miR-17 approximately 92 cluster causes skeletal and growth defects in humans. Nat Genet 2011; 43:1026-30. [PMID: 21892160].

122. Patel V, Williams D, Hajarnis S, Hunter R, Pontoglio M, Somlo S, Igarashi P. miR-17~92 miRNA cluster promotes kidney cyst growth in polycystic kidney disease. Proc Natl Acad Sci USA 2013; 110:10765-70. [PMID: 23759744].

123. He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, Powers S, Cordon-Cardo C, Lowe SW, Hannon GJ, Hammond SM. A microRNA polycistron as a potential human oncogene. Nature 2005; 435:828-33. [PMID: 15944707].