High-throughput sequencing reveals novel lincRNA in age-related cataract

NA ZHANG, CHUNMEI ZHANG, XU WANG and YANHUA QI

Department of Ophthalmology, The Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang 150081, P.R. China

Received January 6, 2017; Accepted October 2, 2017

DOI: 10.3892/ijmm.2017.3185

Abstract. Age-related cataract (ARC) is a major cause of blindness. Long non-coding RNAs (lncRNAs) are a heterogeneous class of RNAs that are non-protein-coding transcripts >200 nucleotides in length. LncRNAs are involved in various critical biological processes, such as chromatin remodeling, gene transcription, and protein transport and trafficking. Furthermore, the dysregulation of lncRNAs causes a number of complex human diseases, including coronary artery diseases, autoimmune diseases, neurological disorders and various cancers. However, the role of lncRNA in cataract remains unclear. Therefore, in the present study, lens anterior capsular membrane was collected from normal subjects and patients with ARC and total RNA was extracted. High-throughput sequencing was applied to detect differentially expressed lncRNAs and mRNAs. The analysis identified a total of 42,556 candidate differentially expressed mRNAs (27,447 +15,109) and a total of 7,041 candidate differentially expressed lncRNAs (4,146 + 2,895). Through bioinformatics analysis, the significant differential expression of novel lincRNA was observed and its possible molecular mechanism was explored. Reverse transcription-quantitative polymerase chain reaction was used to validate the different expression levels of selected lncRNAs. These findings may lead to the development of novel strategies for genetic diagnosis and gene therapy.

Introduction

Age-related cataract (ARC) is one of the most common chronic diseases and a main cause of blindness worldwide (1). At present, surgery is the most effective method of treating cataracts. However, surgery is an economic burden to society, and has associated risks and complications (2). On the basis of the location of the opacity in the lens, ARC is classified as cortical, nuclear or posterior sub-capsular cataract (3,4). In the present study, the main focus was on nuclear cataract (S2) and posterior capsule cataract (S3). ARC is associated with a variety of factors, including age, sex, radiation (visible light, ultraviolet and X-ray), oxidation, trauma, diet and drugs (5). ARC may also be associated with the immune response and growth factors (6,7).

With the emergence of high-throughput sequencing technology, long non-coding RNA (lncRNA) has emerged as an alternative to microRNA (miRNA) as a research topic of great interest. LncRNAs are transcripts that are >200 nucleotides in length, and are similar in structure to mRNA but have little or no protein-coding potential (8,9). According to their genomic locations, lncRNAs may be divided into several types, one of which is intronic lncRNAs (lincRNA), which refers to lncRNA located within an intergenic region of the genome (10). lncRNAs regulate gene expression at the transcriptional, epigenetic or translational level, and thereby alter cellular responses to various stresses (11). Aberrant lncRNA expression is implicated in several human diseases, including tumorigenesis, neurological diseases and cardiovascular diseases (12-14). Certain lncRNAs have been demonstrated to serve important roles in the development of the eye and ocular diseases. Ventral anterior homeobox 2, opposite strand (Vax2os), retinal non-coding RNA 2 (RNCR2), six3 opposite strand (Six3OS) and taurine upregulated 1 (Tug1) have been indicated to have an association with eye development: Vax2os1 controls the cell cycle progression of photoreceptor progenitors in the mouse retina, whereas RNCR2, Six3OS and Tug1 play key roles in the management of retinal cell-specific differentiation (9,15). In addition, metastasis associated lung adenocarcinoma transcript 1 (MALAT1) has been identified to have significantly increased expression in the retinal tissue and aqueous humor of diabetic mice, and in the vascular fiber membranes of diabetic patients (16,17). However, the role of lncRNA in the lens is unclear. Therefore, the present study was designed to explore the possible regulation mechanism of lncRNAs in ARC.

Materials and methods

Clinical sample collection. The lens samples for the lncRNA microarray were collected from postmortem donors (free from ocular diseases) (S1) and ARC patients (9 patients, free of other

Correspondence to: Professor Yanhua Qi, Department of Ophthalmology, The Second Affiliated Hospital of Harbin Medical University, 246 Baojian Road, Nangang, Harbin 150081, P.R. China
E-mail: qyh86605643@126.com

Key words: long non-coding RNA, novel lincRNA, age-related cataract, reverse transcription-quantitative polymerase chain reaction, protein-protein interaction network

ocular diseases; these were divided into 2 groups: the S2 group included 3 patients with nuclear cataract, and the S3 group included 6 patients with posterior capsule cataract) at the Second Affiliated Hospital of Harbin Medical University (Harbin, China) from November, 2014 to January, 2015. All samples were obtained from male donors. The average age of the normal control group was 36, and the average age of the cataract patients was 65. The present study was approved by the Human Ethics Review Board of the Second Affiliated Hospital of Harbin Medical University (Harbin, China) and informed consent was provided by all 9 cataract patients. All lens samples were obtained by intact continuous curvilinear capsulorhexis, without vascular contacting or damage to the iris or any other intraocular structures. Since the RNA of a single lens was not sufficient for microarray analysis, 3 cataract lenses were pooled together as a biological repeat to obtain enough RNA. The degree of lenticular opacification was determined using the Lens Opacities Classification System II (18).

Tissue collection and RNA extraction. The collected samples were stored in liquid nitrogen. Total RNAs were extracted using TRIzol reagent (Invitrogen; Thermo Fisher Scientific, Waltham, MA, USA). The yield of RNA was determined using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Inc.). LncRNA high throughput sequencing analysis, including labelling, hybridization, scanning, normalization and data analysis, was performed by Annoroad Geomics (Beijing, China) (16,17).

RNA sequencing. A quality check of the input total RNA was performed to confirm its integrity by running an aliquot on an Agilent Bioanalyzer (Agilent Technologies, Inc., Santa Clara, CA, USA), and the concentration of the RNA was measured using an ultra-micro spectrophotometer. A 3-μg quantity of RNA was selected as the initial amount of each sample from which to construct a lncRNA library, following removal of the ribosomal RNA (rRNA) in the sample using Ribo-Zero™ Gold kits (human/mouse/rat) (Epicentre, Madison, WI, USA). Different index tags were selected according to the NEB Next Ultra database direct operating protocol (NEB Next Ultra Directional RNA Library Prep kit for Illumina;NEB Ipswich, MA, USA; all reagents used here were provided in the kit). The specific steps of library construction were as follows: Firstly, the Ribo-Zero™ Gold kit was used to remove rRNA. Then, fragmentation buffer was added to the reaction system to fragment RNA into short segments, and the short fragments were used as templates for first cDNA chain synthesis using six base random primers (random hexamers). This was followed by second-strand cDNA synthesis using DNA polymerase I and RNaseH. The cDNA fragments were then subjected to an end-repair process, the addition of a single 'A' base, and ligation of the Illumina multiplexing adapters. The products were purified and enriched using polymerase chain reaction (PCR) to create the final cDNA sequencing library, as previously described (19-21). For PCR, the thermocycling conditions were as follows: step 1, 65˚C for 15 min; step 2, 30˚C for 10 min, 42-50˚C for 45 min and 95˚C for 5 min.

Data processing. The raw data was filtered to provide high-quality, cleaned reads, and a follow-up analysis was then conducted. The data processing steps were as follows: i) Removal of contamination reads, ii) removal of low-quality reads and iii) removal of the reads for bases having a nitrogen content >5% (5,22).

LncRNA analysis. Only lincRNA was analyzed in the present study. According to the characteristics of the lincRNA, a series of strict filters were used. The filter conditions were a length of ≥200 bp, exon number of ≥2 and minimum coverage of ≥3 transcription reads. Known mRNA transcripts, known non-coding RNA transcripts, homologous transcripts and transcripts with protein-coding capability were removed.

Read comparison. In this study, we used the TopHat software version 2.0.12 and selected the default parameters, and compared the reference sequence to hg19 (31).

Differential expression analysis. The cataract samples and normal samples in each group were subjected to differential expression analysis using DESeq2. Differential expression analysis was then carried out following standardization and variance estimation. The standards for screening the expression differences of mRNA and lncRNA were $P<0.001$, a false discovery rate (FDR) of <0.001 and a $|\log_2 \text{ratio}| >2$ (21).

Gene set enrichment analysis. Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using the DAVID functional annotation web server (http://david.abcc.ncifcrf.gov) using default parameters. Enriched GO terms (FDR <0.05) and KEGG pathways ($P<0.1$) were obtained (21-24).

Screening target genes. Cis target prediction was based on the distance between the lncRNA its linked protein-coding gene and to forecast the targets. Trans target prediction was according to the sequence of lncRNA and mRNA sequence, and the evaluation of the combination of free energy, which is a combination of the stability (25).

During the construction of all of the networks, Cytoscape software was used. Cytoscape is an open source software platform for visualizing molecular interaction networks and biological pathways and integrating these networks with annotations, gene expression profiles and other state data (http://www.cytoscape.org/).

Conservation analysis. The conservation of the corresponding sequence and site of the identified novel lincRNA was analyzed in order to evaluate the extent of its variation. Conservation scores (mammalian phastCons scores) were used to carry out the conservation analysis, as previously described (26).

Reverse transcription-quantitative PCR (RT-qPCR) validation of differentially expressed lncRNAs. Total RNA was extracted for use in microarray experiments from cataract patients and postmortem eyes as described above. Agarose gel electrophoresis was used to determine the integrity of the RNA, and its concentration was determined using a UV spectrophotometer. Quantification was performed with a

two-step reaction process: RT and qPCR. Each RT reaction consisted of 0.5 mg RNA, 2 μl 5X PrimeScript buffer, 0.5 μl oligo(dT), 0.5 μl random 6 mers and 0.5 μl PrimeScript RT Enzyme Mix I (Takara Bio, Inc., Otsu, Japan), with RNase Free dH_2O added to a total volume of 10 μl. Reactions were performed using a GeneAmp PCR system 9700 (Applied Biosystems; Thermo Fisher Scientific, Inc.) for 15 min at 37˚C, followed by heat inactivation of RT for 5 sec at 85˚C. The 10-μl RT reaction mix was then diluted 10-fold with nuclease-free water and held at -20˚C. qPCR was performed using a Stratagene Mx3000P Real-time PCR instrument (Agilent Technologies, Inc., Santa Clara, CA, USA) with 20 μl PCR reaction mixture comprising 100 ng template DNA, 10 μl 2X GreenStar Master mix, 1 μl forward primer, 1 μl reverse primer and PCR grade water to a volume of 20 μl (Bioneer Corporation, Daejeon, Korea). Reactions were incubated in a 384-well optical plate at 95˚C for 10 min, followed by 40 cycles of 95˚C for 10 sec, 60˚C for 30 sec. Each sample was run and analyzed in triplicate. Data quantification was carried out using the 2-ΔΔCq method (27).

Results

Sequencing data analysis. The base distribution was analyzed, in order to investigate the presence of the AT and GC separation phenomenon; this phenomenon may arise during the sequencing or building of the library, and will affect the subsequent quantitative analysis. It was found that the frequencies of AT and CG was almost the same, which confirmed the reliability of the obtained data (Fig. 1A and B).

Read comparison. TopHat software was used to compare the filtered samples with the reference genome, and to locate the sequence within the genome. When the reference genome selection is appropriate and assembly is complete, in samples without exogenous species contamination, the matching rate of the sample is usually >80% (28-30) (Table I). The total number of reads in S2 was 126,905,134, including 92% mapped reads and 4% multiMap reads. The total number of reads in S3 was 102,287,578, including 96% mapped reads and 7% multiMap reads. The number and proportion of matching sequences of three functional components (exon, intron and intergenic) were confirmed. The majority of the matched sequences in the three samples were in the exon (Fig. 1C). In general, if the species of annotation information comprehensive, the majority of the sequence should match with the exon region (31,32).

Estimation of mRNA expression. Gene expression levels are generally measured on the basis of the mRNA transcription number of the gene. The units reads per kilobase million mapped reads (RPKM) were used as a measure of gene expression. It was observed that the number of differentially expressed genes accounted for only a small proportion of all genes, and so these genes are likely to have little impact on the overall distribution of genes in the samples, and all samples should have a similar distribution of expression (Fig. 1D) (33).

Characteristics of the novel lincRNA. In order to better define the characteristics of the novel lincRNA, its length (Fig. 2A), the number of exons (Fig. 2B) and expression quantity (Fig. 2C) were analyzed. The distribution of lincRNA expression was calculated according to the gene expression in all groups, and it was observed that all samples were similar in expression. The differentially expressed genes account for only a small fraction of the overall gene (34).

Table I. Comparison of samples [n (%)].

Reads	S2	S3
Total reads	126,905,134	102,287,578
Mapped reads	117,262,466 (92)	98,284,753 (96)
MultiMap reads	4,738,594 (4)	6,924,023 (7)

S2, nuclear cataract; S3, posterior capsule cataract; total reads, total number of filtered sequences; mapped reads, sequences matching the genome; MultiMap reads, sequences matching with multiple locations in the genome.

Table II. Differential expression of mRNAs.

Expression	S2 vs. S1	S3 vs. S1
Upregulated	24,947	10,478
Downregulated	2,500	4,613
Total	27,447	15,109

S1, normal eyes; S2, nuclear cataract; S3, posterior capsule cataract.

Differential expression and functional analysis of mRNA. DEseq was used to conduct a differential expression analysis between S2 and S1, and between S3 and S1 using $q<0.05$ (35). The analysis identified 27,447 and 15,019 candidate differentially expressed mRNAs, respectively, in these two comparisons. Hierarchical cluster analysis was conducted for the classification of candidate mRNAs in each group. Then, fold change values were used to classify the candidate mRNAs into upregulated mRNAs (fold change ≥2) and downregulated mRNAs (fold change ≤0.05). For S2 vs. S1 and S3 vs. S1, 24,947 and 10,478 upregulated mRNAs and 2,500 and 4,631 downregulated mRNAs, respectively, were identified (Table II). Further analysis of the differential expression is presented in Fig. 3. Intersections of the differentially expressed mRNAs were detected, and it was found that there were 11,348 differentially expressed mRNAs, 9,261 upregulated mRNAs and 2,087 downregulated mRNAs (Fig. 3C). Furthermore, an enrichment analysis was performed to investigate the functions of these upregulated and downregulated mRNAs. These mRNAs, whether upregulated or downregulated, were enriched in several basic biological or nerve-related terms and pathways, such as the phenylalanine metabolism pathway (Fig. 3A and B). Notable GO terms included cytoskeletal protein binding and actin binding. The phenylalanine metabolism pathway and cytoskeletal protein binding and actin binding are the biological processes of the marked enrichment of all data after processing (data not shown). In particular, the mRNAs were significantly rich in GO terms associated with

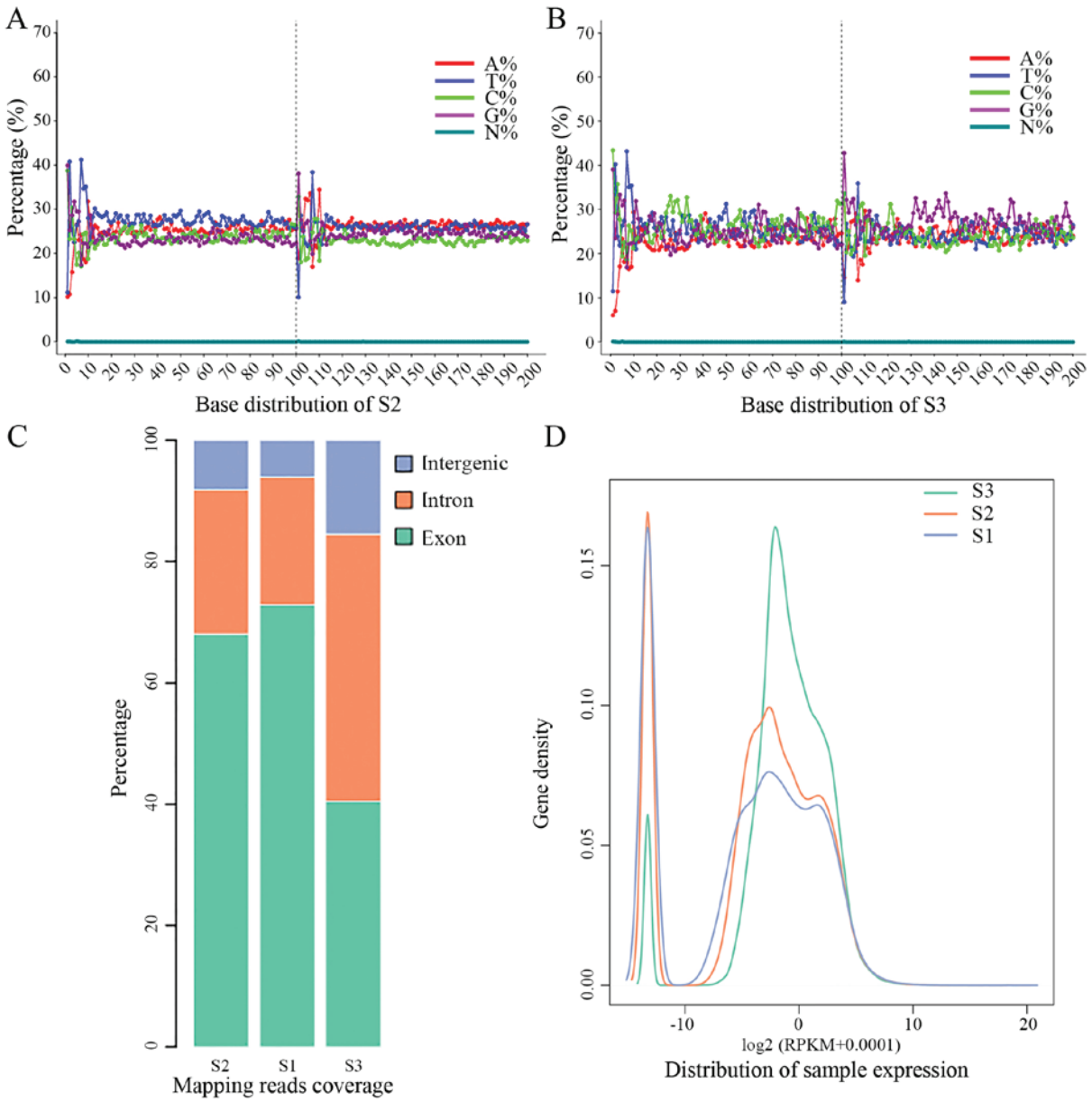

Figure 1. (A and B) Sample base content distribution map. The base position of the filtered sequence is used as the horizontal coordinate, and the ratio of the A, T, G and C base content in each position as the longitudinal coordinate. (C) Distribution of the unique sequence in the reference genome. The histogram presents the proportions of matched exons, introns and intergenic sequences. (D) Distribution of sample expression. S1, normal eyes; S2, nuclear cataract; S3, posterior capsule cataract; RPKM, reads per kilobase million mapped reads.

structural constituents of the eye lens, which was related to the function and nerves of eyes. Some researchers considered ARC to be a structural disease (36). The lens is composed of >90% crystalline protein, with the remainder comprising skeletal, membrane, connective and water channel proteins (37). A large number of studies have indicated that the occurrence of cataract is directly associated with the apoptosis of lens epithelial cells induced by the modification of lens proteins, such as changes in the ratio of proteins or protein structure (38-40). The results of the present study revealed various differentially expressed mRNAs that were strongly associated with the function of the eyes. In the KEGG pathway analysis (data not shown), downregulated mRNA expression was particularly enriched in the tyrosine metabolism pathway, which is closely associated with cataract. The abnormal metabolism of tyrosine in the lens results in the production of quinones, which may be oxidized and thereby damage the protein of the lens, causing lens opacity and ultimately the occurrence of cataract (40-42). In addition, tyrosine is able to bind to lens proteins through phosphorylation, which may also have an association with cataract (43).

Differential expression and functional analysis of lncRNA. Due to the low expression feature of lncRNAs, q<1.1 was used to identify the differentially expressed lncRNAs from the novel lincRNAs between S2 and S1, and between S3 and S1. This revealed that there were 4,146 and 2,895 differentially expressed lncRNAs, respectively (Table III). Fold-change values were used to classify the candidate lncRNAs into upregulated lncRNAs (fold change ≥2) and downregulated

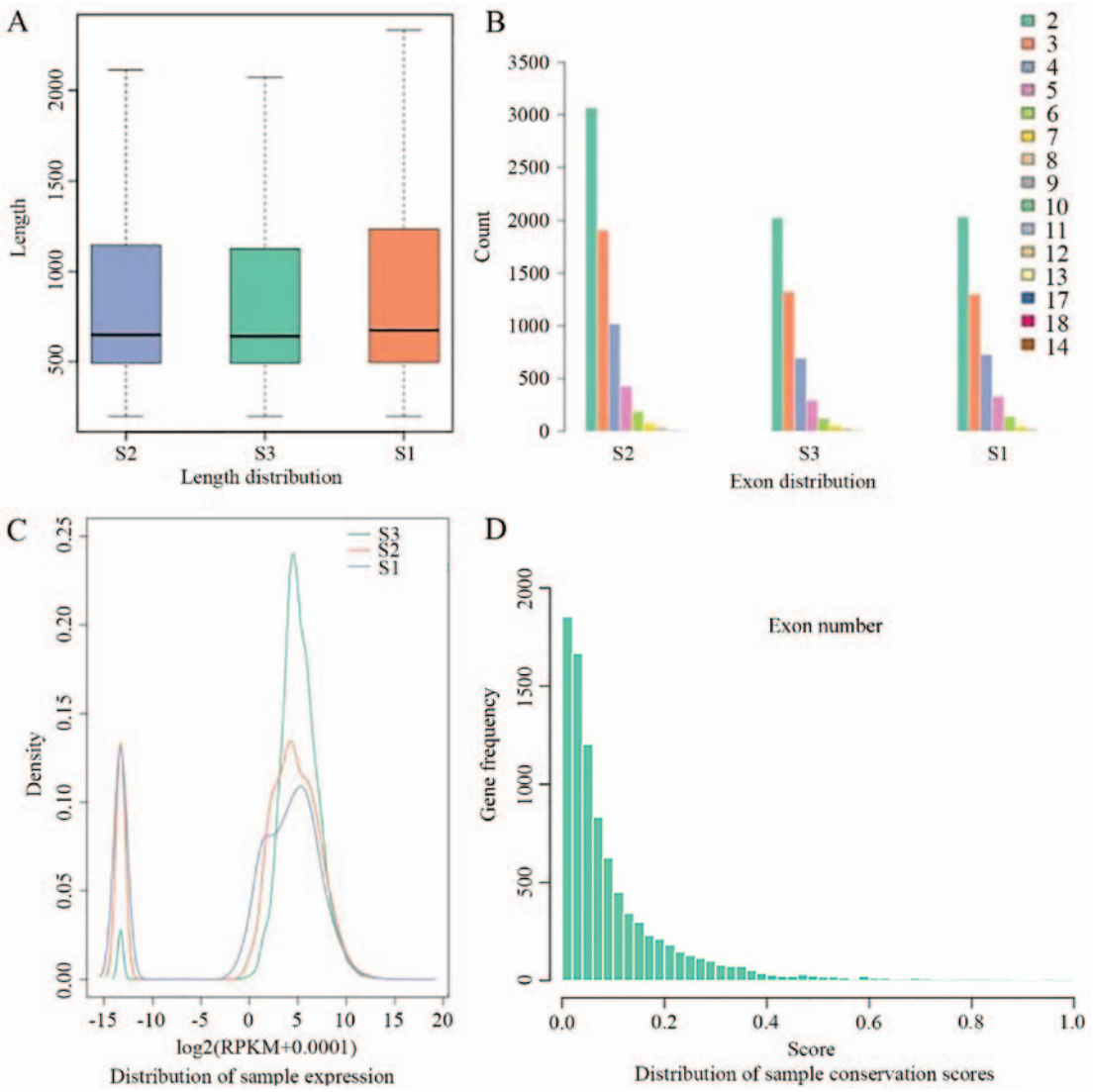

Figure 2. (A) Length distribution trend of the samples in the three groups. The average length of each sample is >500 bp. (B) Novel long non-coding RNA exon number analysis in the S1, S2 and S3 groups. Different colors represent different numbers of exons of each transcript. The numbers in the y axis indicate the number of transcripts. (C) Distribution of gene expression. (D) Distribution of sample conservation scores. S1, normal eyes; S2, nuclear cataract; S3, posterior capsule cataract; RPKM, reads per kilobase million mapped reads.

Table III. Differential expression of lncRNAs.

Expression	S2 vs. S1	S3 vs. S1
Upregulated	3,477	2,051
Downregulated	669	844
Total	4,146	2,895

lncRNA, long non-coding RNA; S1, normal eyes; S2, nuclear cataract; S3, posterior capsule cataract.

lncRNAs (fold change ≤0.05), and 3,477 and 2,051 upregulated lncRNAs and 669 and 844 downregulated lncRNAs were identified, respectively. Intersections of the lncRNAs were also obtained for each group. There were found to be 2,269 common differentially expressed lncRNAs, 1,701 upregulated lncRNAs and 568 downregulated lncRNAs (Fig. 3D). Hierarchical cluster analysis was also performed, to classify the candidate lncRNAs for each group.

LncRNA-mRNA network. The 12,097 and 2,332 differential mRNAs and lncRNAs were further analyzed by integrating all differential mRNAs and lncRNAs for each group using intersection sets. Firstly, comprehensive mRNA-lncRNA interactions were obtained by integrating cis-targets and trans-targets. The interactions were also classified into upregulated and downregulated interactions using the dysregulated mRNAs. The functions of the dysregulated lncRNAs were analyzed on the basis of their target mRNAs (Figs. 4 and 5). It was observed that these dysregulated lncRNAs were enriched in certain basic pathways and GO terms, including regulation of the actin cytoskeleton pathway and protein binding terms. The result was coincident with the dysregulated mRNAs. Genes such as paired box 6 and chaperone-like activity of αA-crystallin (CRYAA), have similar functions. The lncRNA target gene major intrinsic protein of lens

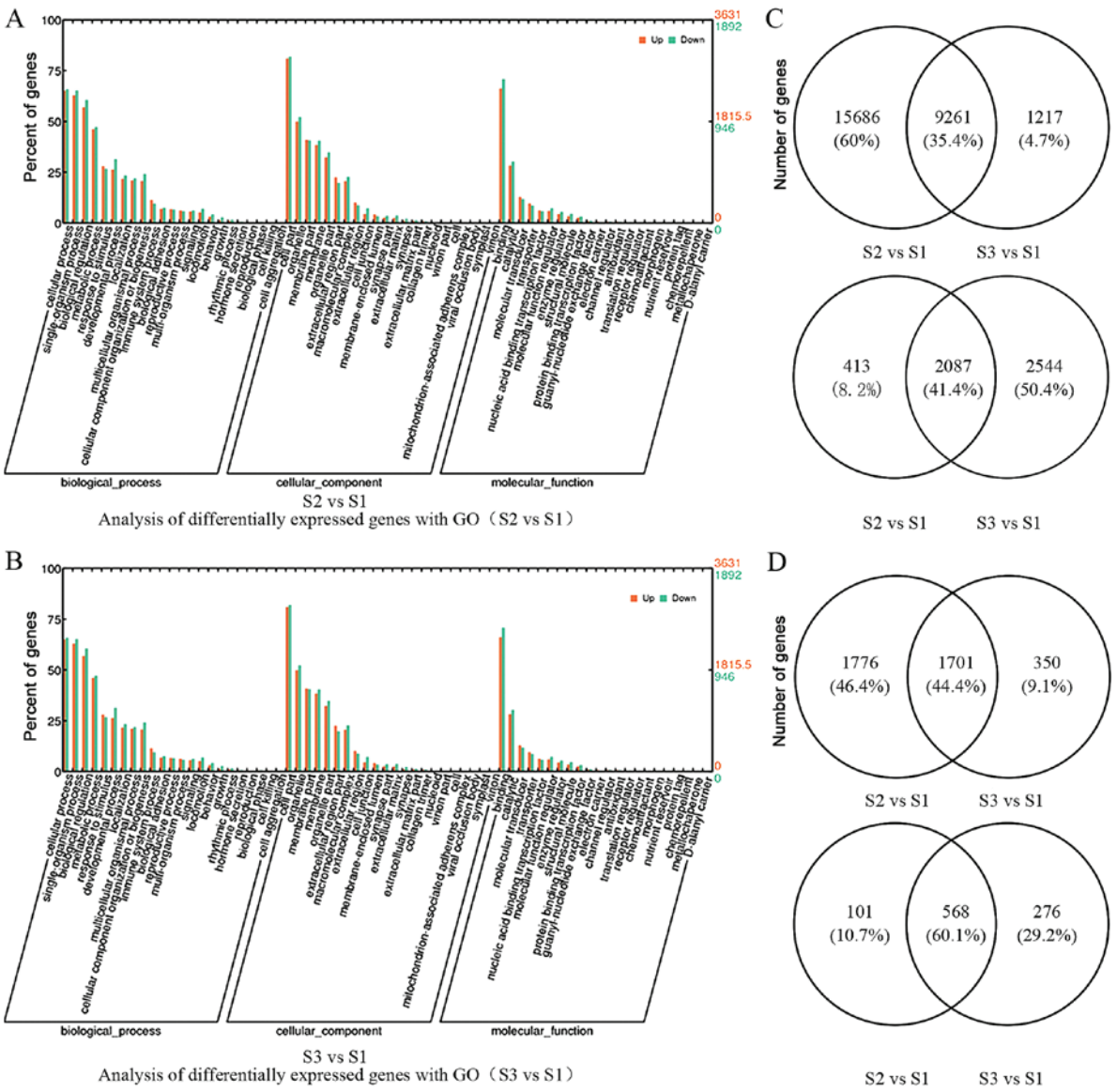

Figure 3. GO analysis of the differentially expressed genes for (A) S2 vs. S1 and (B) S3 vs. S1. The horizontal coordinates for the large classes under the GO annotation represent a variety of biological processes, cellular components and molecular functions. The x axis is a variety of biological processes, cell components, and molecular functions of GO. The left vertical coordinate is the proportion of the class, and the right vertical is the specific gene number of the class.Different colors represent different groups. Venn diagrams showing the intersection of the differentially expressed (C) mRNAs (upper diagrams, upregulated mRNAs; lower diagrams, downregulated mRNAs) and (D) long non-coding RNAs (upper diagrams, upregulated lncRNAs; lower diagrams, downregulated lncRNAs). GO, gene ontology; S1, normal eyes; S2, nuclear cataract; S3, posterior capsule cataract.

fiber (MIP, ENSG00000135517) encodes the most abundant junctional membrane protein in the mature lens and this protein serves a critical role in the maintenance of the normal structure and internal circulation of the lens (44,45). In ARC, CRYAA is considered to be critical for the maintenance of lens transparency (46,47). In the GO analysis, it was found that ENSG00000160202 and CRYAA were linked nodes (Fig. 4C), indicating that they have similar biological functions (48). ENSG00000166147 and fibrillin-1 (FBN1) were also linked nodes (Fig. 4B). FBN1 is the most common pathogenic gene for Marfan syndrome, and encodes the fibrillin-1 protein. Eye diseases including lens dislocation or subluxation, and cataract, are observed in typical Marfan syndrome patients. Studies have suggested that fibrillin-1 is not a component of the normal lens, but may serve as a connecter between latent-transforming growth factor β-binding protein 1 and the extracellular matrix (ECM) in capsular opacification. In addition, FBN1 is an important gene in eye development (49,50).

Furthermore, a human protein-protein interaction (PPI) network based on the HPRD data in the present study was constructed. The largest sub-network of the upregulated and downregulated mRNAs was then obtained from the PPI network. During the construction of all of the networks, the software Cytoscape was used. The upregulated network included 359 nodes and 395 edges, and the downregulated network included 305 nodes and 346 edges. The two networks approximated the scale-free network topology of dysregulated networks. The connectivity and betweenness of nodes were also analyzed. Some other networks also have similar topological features (51,52) (data not shown).

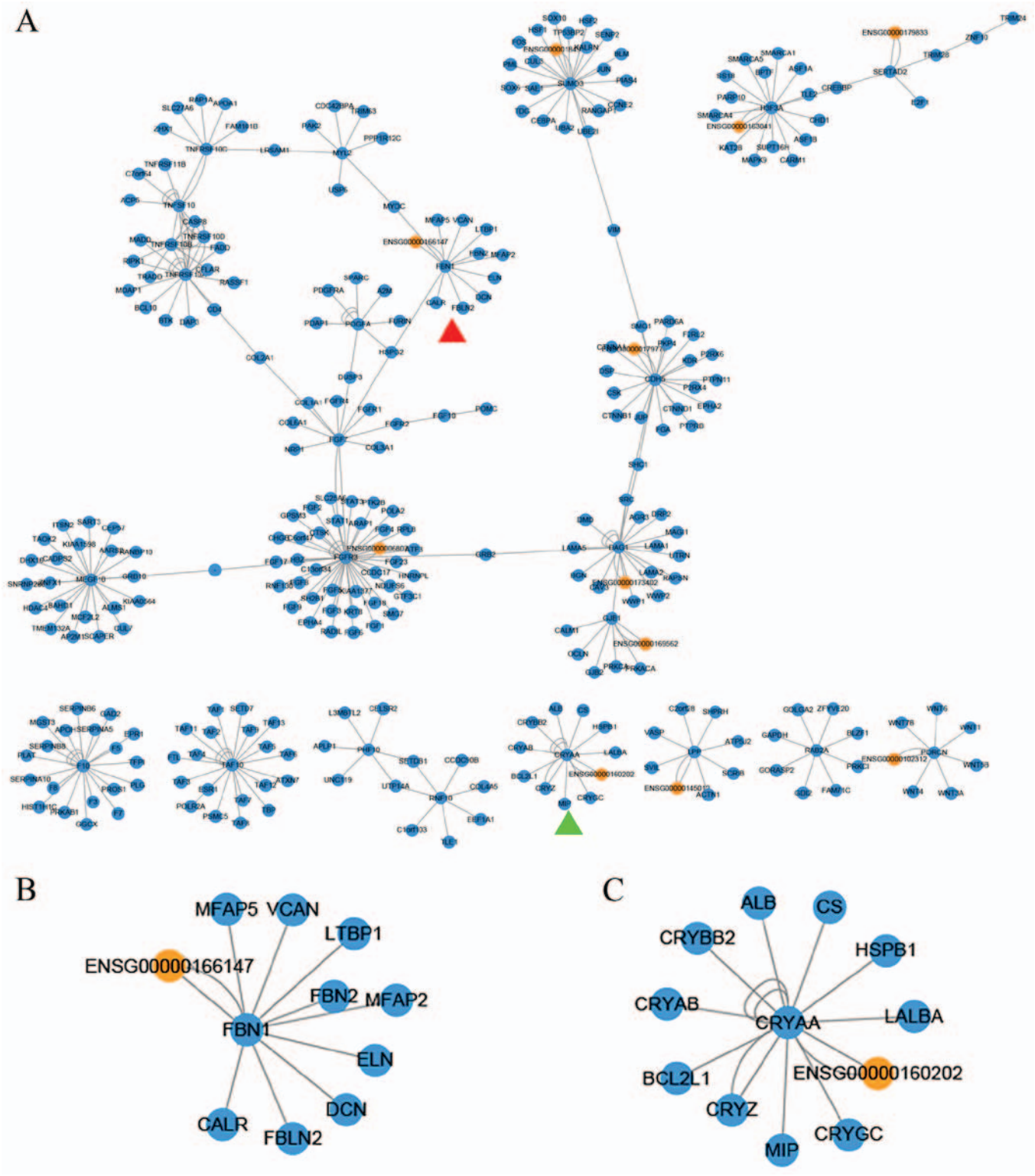

Figure 4. (A) Upregulated-lncRNA network and sub-networks obtained using the HPRD network. The red and green triangles indicate the position of the networks shown in (B and C). (B) An example of a sub-network of the HPRD network. (C) The network of genes associated with age-related cataract extracted from (A). lncRNA, long non-coding RNA; HPRD, Human Protein Reference Database.

Conservation analysis. The PhastCons score of each chromosome was downloaded, which contains all loci on each gene from the University of California, Santa Cruz database, in order to analyze the conservative of lncRNA. The conservation scores of the majority of the lncRNAs in the present study (Fig. 2D) were at a low level (53).

RT-qPCR validation. To validate the results of sequencing, 30 differentially expressed lncRNA transcripts were selected for validation using RT-qPCR. The RT-qPCR results were consistent with the sequencing results for 28 of the lncRNA transcripts, showing the same trends of differential expression ($P<0.05$). Of note, due to individual differences or other reasons, two were

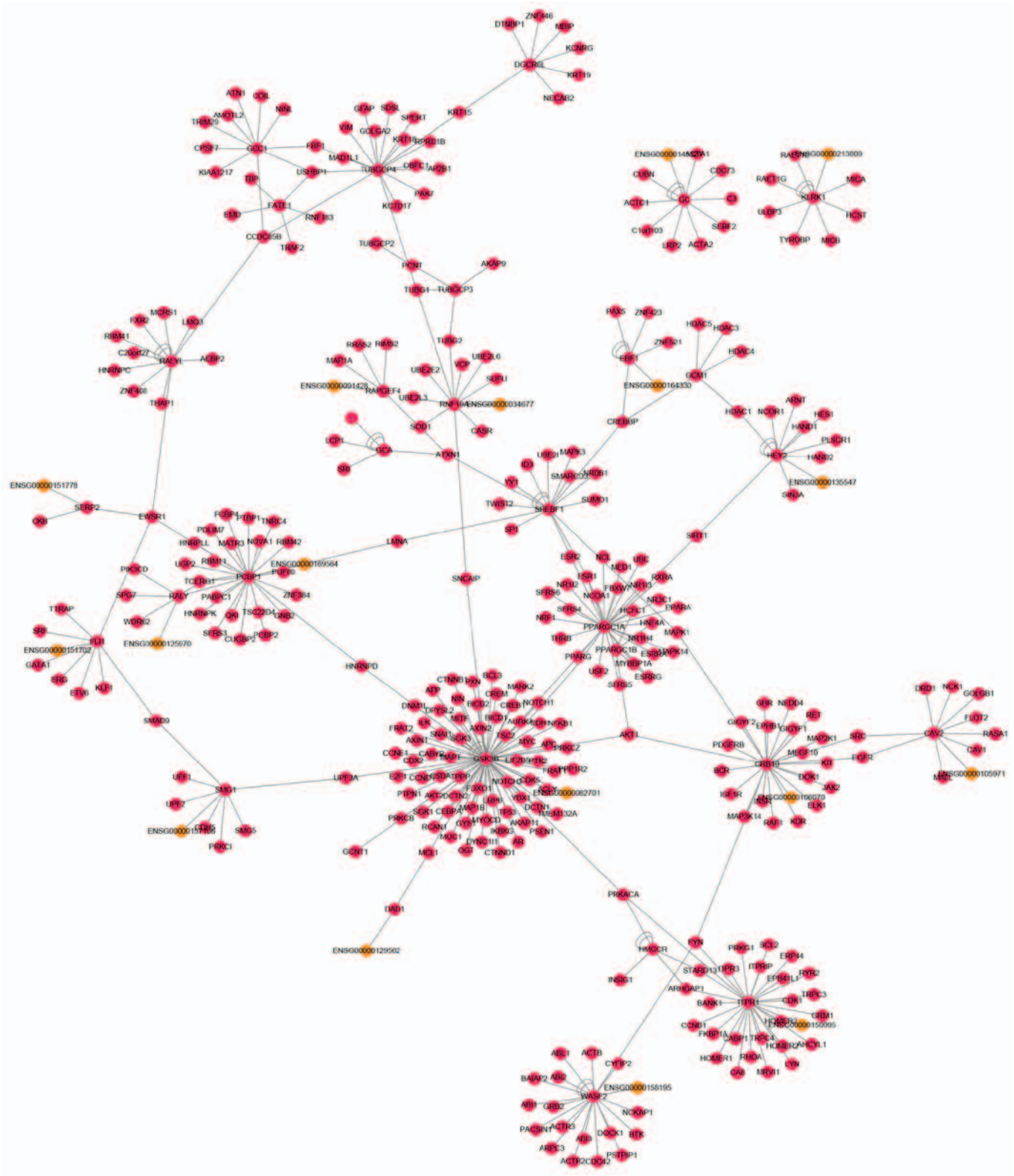

Figure 5. Downregulated mRNA and corresponding long non-coding RNA networks obtained using the Human Protein Reference Database.

inconsistent with the sequencing results (Table IV).The specific mechanisms need to be further explored.

Discussion

Cataract is the leading cause of blindness globally, and its incidence increases yearly (1). To date, the pathogenesis of cataract has mainly been investigated at the cellular and molecular levels, including the post-translational modification, physical and chemical factors of lens proteins. miRNA has tissue and cell specificity in the eye. It serves an important regulatory role in the regulation of cell proliferation, differentiation and apoptosis processes (54).

With the rapid development of high-throughput sequencing and bioinformatics technology, lncRNAs have been identi-

Table IV. Validation of sequencing results by RT-qPCR.

Target of lncRNA	High-throughput sequencing	RT-qPCR
ENSG00000188536	Upregulated	Upregulated
ENSG00000206172	Upregulated	Upregulated
ENSG00000225282	Upregulated	Upregulated
ENSG00000175147	Upregulated	Upregulated
ENSG00000237988	Upregulated	Upregulated
ENSG00002225282	Upregulated	Upregulated
ENSG00000008517	Upregulated	Upregulated
ENSG00000077080	Upregulated	Downregulated
ENSG00000244734	Upregulated	Upregulated
ENSG00000196361	Upregulated	Upregulated
ENSG00000197414	Upregulated	Upregulated
ENSG00000171346	Upregulated	Upregulated
ENSG00000253418	Upregulated	Upregulated
ENSG00000262516	Upregulated	Upregulated
ENSG00000118231	Upregulated	Upregulated
ENSG00000146166	Downregulated	Downregulated
ENSG00000100122	Downregulated	Downregulated
ENSG00000135517	Downregulated	Downregulated
ENSG00000196431	Downregulated	Downregulated
ENSG00000122986	Downregulated	Downregulated
ENSG00000233797	Downregulated	Upregulated
ENSG00000106868	Downregulated	Downregulated
ENSG00000082515	Downregulated	Downregulated
ENSG00000138376	Downregulated	Downregulated
ENSG00000105325	Downregulated	Downregulated
ENSG00000136158	Downregulated	Downregulated
ENSG00000186314	Downregulated	Downregulated
ENSG00000134258	Downregulated	Downregulated
ENSG00000213949	Downregulated	Downregulated
ENSG00000182218	Downregulated	Downregulated

fied as a class of non-encoding RNAs abundantly expressed in mammalian genomes; however, their functions are mostly unknown. The study of lncRNA has revealed that although their sequences are less conservative than those of mRNAs, the promoter and functional components are highly conservative. This indicates that lncRNA may have a specific function (55). There is evidence to support the hypothesis that lncRNAs are able to regulate the expression of protein-encoding genes, including the stability and subcellular localization of proteins. This is involved in a wide variety of biological processes, such as imprinting control, cell differentiation and immune response (56). The role of lncRNA in the eye has attracted attention. Several lncRNAs including Vax2os, RNCR2, Six3OS and Tug1, have been confirmed to be associated with the development of the eye. Vax2os controls the cell cycle progression of receptors in mouse retina; RNCR2, Six3OS and Tug1 serve key roles in the differentiation of retinal cells (57,58). Another lncRNA, MIAT, has an effect on the proliferation and differentiation of lens epithelial cells (59). However, the function of lncRNA in cataract is not yet clear.

In the present study, high-throughput sequencing revealed a large number of differentially expressed lncRNAs and mRNAs in ARC compared with ocular disease-free eyes (from postmortem donors). Construction of an lncRNA-mRNA network was used to analyze the functions of the lncRNA. The reliability of the differential expression of the lncRNAs was evaluated using RT-qPCR. The RT-qPCR results for 28 of the 30 selected differentially expressed lncRNAs were consistent with the sequencing results. LncRNA is RNA characterized by abundant expression but low conservation, low cell expression specificity and low transcription characteristics (60). As a result, it is difficult to directly identified the function of lncRNA via the nucleic acid sequences with current technologies. Therefore, the present study focused on their biological functions and the regulation pathways of differentially expressed genes using bioinformatics, to provide a theoretical basis for further study of the roles of lncRNAs in cataract. The lncRNA-mRNA expression network was constructed and GO analysis of the differentially expressed mRNA was conducted. In order to promote the accuracy of analysis, the intersections of the differentially expressed mRNAs were determined. It was found there were 12,097 differentially expressed mRNAs, 9,261 upregulated mRNAs and 2,087 downregulated mRNAs. These mRNAs were enriched in several basic biological or nerve-related terms and pathways, such as sodium symporter activity, neurotransmitter transporter activity and cytoskeletal protein binding. Notably, some were significantly enriched in GO terms relating to structural constituents of the eye lens. This demonstrated the complexity of the regulatory mechanism. KEGG pathway analysis of differentially expressed mRNAs in the lncRNA-mRNA expression networks was also conducted. The results reveal that genes involved in biological pathways including tyrosine metabolism, phenylalanine metabolism and regulation of the actin cytoskeleton. These signaling pathways are associated with numerous pathological processes. Thus, it appears that lncRNAs and the differentially expressed gene network are involved in the regulation of multiple cellular processes. Therefore, they may be a novel target for the treatment of cataract and associated diseases, which may have potential clinical significance.

It is worthy of note that the calcium signaling pathway, and regulation of the actin cytoskeleton, are closely associated with the underlying pathological processes of a number of diseases. The calcium ion is the most common signal transduction factor in the human body, and plays an important role in cell division, growth and death. Calcium ions act as the second messenger in the regulation of cell apoptosis, which is known to be one of the main causes of ARC (61,62). It may be speculated that the involvement of the lncRNAs in the pathogenesis of ARC involves regulation of the calcium signaling pathway, which provides new opportunities for the diagnosis and treatment of ARC. The lens is a fascinating and unique transparent tissue that grows continuously throughout life. During the process of differentiation into fiber cells, lens epithelial cells undergo marked morphological changes, membrane remodeling, polarization, transcriptional activation and the elimination of cellular organelles, including nuclei, concomitant with migration towards the lens interior (63,64). The majority of these events are considered to be influenced by dynamic reorganization of the cellular actin cytoskeleton

and by intercellular and cell (65). These observations suggest a number of new directions for exploration with regard to the pathogenic effects of lncRNA.

Acknowledgements

The authors are grateful to all patients and normal volunteers for their participation in this study. This study was supported by the National Natural Science Foundation of China, Harbin, Y.Q. (grant no. 81170833).

References

1. Resnikoff S and Keys TU: Future trends in global blindness. Indian J Ophthalmol 60: 387-395, 2012.
2. Pescosolido N, Barbato A, Giannotti R, Komaiha C and Lenarduzzi F: Age-related changes in the kinetics of human lenses: Prevention of the cataract. Int J Ophthalmol 9: 1506-1517, 2016.
3. Hammond CJ, Duncan DD, Snieder H, de Lange M, West SK, Spector TD and Gilbert CE: The heritability of age-related cortical cataract: The twin eye study. Invest Ophthalmol Vis Sci 42: 601-605, 2001.
4. Tsai SY, Hsu WM, Cheng CY, Liu JH and Chou P: Epidemiologic study of age-related cataracts among an elderly Chinese population in Shih-Pai, Taiwan. Ophthalmology 110: 1089-1095, 2003.
5. Sheng Y, He F, Lin JF, Shen W and Qiu YW: Tea and risk of age-related cataracts: A cross-sectional study in Zhejiang Province, China. J Epidemiol 26: 587-592, 2016.
6. Chen W, Lin H, Zhong X, Liu Z, Geng Y, Xie C and Chen W: Discrepant expression of cytokines in inflammation- and age-related cataract patients. PloS One 9: e109647, 2014.
7. Hwang HB, Yim HB, Cho YK and Choi JA: The Association Between Aqueous Connective Tissue Growth Factor and the Severity of Age-related Cataracts as Graded by the Lens Opacities Classification System III. Curr Eye Res 41: 350-356, 2016.
8. Wilusz JE, Sunwoo H and Spector DL: Long noncoding RNAs: Functional surprises from the RNA world. Genes Dev 23: 1494-1504, 2009.
9. Yan B, Tao ZF, Li XM, Zhang H, Yao J and Jiang Q: Aberrant expression of long noncoding RNAs in early diabetic retinopathy. Invest Ophthalmol Vis Sci 55: 941-951, 2014.
10. Kung JT, Colognori D and Lee JT: Long noncoding RNAs: Past, present, and future. Genetics 193: 651-669, 2013.
11. Li J, Xuan Z and Liu C: Long non-coding RNAs and complex human diseases. Int J Mol Sci 14: 18790-18808, 2013.
12. Uchida S and Dimmeler S: Long noncoding RNAs in cardiovascular diseases. Circ Res 116: 737-750, 2015.
13. Zhao Q, Li T, Qi J, Liu J and Qin C: The miR-545/374a cluster encoded in the Ftx lncRNA is overexpressed in HBV-related hepatocellular carcinoma and promotes tumorigenesis and tumor progression. PloS One 9: e109782, 2014.
14. Roberts TC, Morris KV and Wood MJ: The role of long non-coding RNAs in neurodevelopment, brain function and neurological disease. Philos Trans R Soc Lond B Biol Sci 369: pii: 20130507, 2014.
15. Meola N, Pizzo M, Alfano G, Surace EM and Banfi S: The long noncoding RNA Vax2os1 controls the cell cycle progression of photoreceptor progenitors in the mouse retina. RNA 18: 111-123, 2012.
16. Shen Y, Dong LF, Zhou RM, Yao J, Song YC, Yang H, Jiang Q and Yan B: Role of long non-coding RNA MIAT in proliferation, apoptosis and migration of lens epithelial cells: A clinical and in vitro study. J Cell Mol Med 20: 537-548, 2016.
17. Yao J, Wang XQ, Li YJ, Shan K, Yang H, Wang YN, Yao MD, Liu C, Li XM, Shen Y, *et al*: Long non-coding RNA MALAT1 regulates retinal neurodegeneration through CREB signaling. EMBO Mol Med 8: 346-362, 2016.
18. Chylack LT, Jr., Leske MC, McCarthy D, Khu P, Kashiwagi T and Sperduto R: Lens opacities classification system II (LOCS II). Arch Ophthalmol 107: 991-997, 1989.
19. Kissopoulou A, Jonasson J, Lindahl TL and Osman A: Next generation sequencing analysis of human platelet PolyA$^+$ mRNAs and rRNA-depleted total RNA. PLoS One 8: e81809, 2013.
20. Lin Y, Golovnina K, Chen ZX, Lee HN, Negron YL, Sultana H, Oliver B and Harbison ST: Comparison of normalization and differential expression analyses using RNA-Seq data from 726 individual *Drosophila* melanogaster. BMC Genomics 17: 28, 2016.
21. Zhang ZH, Jhaveri DJ, Marshall VM, Bauer DC, Edson J, Narayanan RK, Robinson GJ, Lundberg AE, Bartlett PF, Wray NR, *et al*: A comparative study of techniques for differential expression analysis on RNA-Seq data. PLoS One 9: e103207, 2014.
22. Wu Z, Fu Y, Cao J, Yu M, Tang X and Zhao S: Identification of differentially expressed miRNAs between white and black hair follicles by RNA-sequencing in the goat (*Capra hircus*). Int J Mol Sci 15: 9531-9545, 2014.
23. Shah DH: RNA sequencing reveals differences between the global transcriptomes of *Salmonella enterica* serovar enteritidis strains with high and low pathogenicities. Appl Environ Microbiol 80: 896-906, 2014.
24. Huang S, Feng C, Chen L, Huang Z, Zhou X, Li B, Wang LL, Chen W, Lv FQ and Li TS: Identification of potential key long non-coding RNAs and target genes associated with pneumonia using long non-coding RNA sequencing (lncRNA-Seq): A preliminary study. Med Sci Monit 22: 3394-3408, 2016.
25. Wainberg M, Alipanahi B and Frey B: Does conservation account for splicing patterns? BMC Genomics 17: 787, 2016.
26. Huang CH, Wang YT, Tsai CF, Chen YJ, Lee JS and Chiou SH: Phosphoproteomics characterization of novel phosphorylated sites of lens proteins from normal and cataractous human eye lenses. Mol Vis 17: 186-198, 2011.
27. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408, 2001.
28. Finotello F and Di Camillo B: Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. Brief Funct Genomics 14: 130-142, 2015.
29. Kumar R, Ichihashi Y, Kimura S, Chitwood DH, Headland LR, Peng J, Maloof JN and Sinha NR: A High-Throughput Method for Illumina RNA-Seq Library Preparation. Front Plant Sci 3: 202, 2012.
30. Yu J, He K, Ren T, Lou Y and Zhao A: High-throughput sequencing reveals differential expression of miRNAs in prehierarchal follicles of laying and brooding geese. Physiol Genomics 48: 455-463, 2016.
31. Trapnell C, Pachter L and Salzberg SL: TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105-1111, 2009.
32. Langmead B, Trapnell C, Pop M and Salzberg SL: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25, 2009.
33. Wagner GP, Kin K and Lynch VJ: Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci 131: 281-285, 2012.
34. Chen X and Yan GY: Novel human lncRNA-disease association inference based on lncRNA expression profiles. Bioinformatics 29: 2617-2624, 2013.
35. Wang L, Feng Z, Wang X, Wang X and Zhang X: DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26: 136-138, 2010.
36. Panda AK, Nandi SK, Chakraborty A, Nagaraj RH and Biswas A: Differential role of arginine mutations on the structure and functions of alpha-crystallin. Biochim Biophys Acta 1860: 199-210, 2016.
37. Antosova B, Smolikova J, Borkovcova R, Strnad H, Lachova J, Machon O and Kozmik Z: Ectopic activation of Wnt/beta-catenin signaling in lens fiber cells results in cataract formation and aberrant fiber cell differentiation. PloS One 8: e78279, 2013.
38. Li Y, Jia Y, Zhou J and Huang K: Effect of methionine sulfoxide reductase B1 silencing on high-glucose-induced apoptosis of human lens epithelial cells. Life Sci 92: 193-201, 2013.
39. Linetsky M, Raghavan CT, Johar K, Fan X, Monnier VM, Vasavada AR and Nagaraj RH: UVA light-excited kynurenines oxidize ascorbate and modify lens proteins through the formation of advanced glycation end products: implications for human lens aging and cataract formation. J Biol Chem 289: 17111-17123, 2014.
40. Staniszewska MM and Nagaraj RH: 3-hydroxykynurenine-mediated modification of human lens proteins: structure determination of a major modification using a monoclonal antibody. J Biol Chem 280: 22154-22164, 2005.
41. Ipson BR and Fisher AL: Roles of the tyrosine isomers meta-tyrosine and ortho-tyrosine in oxidative stress. Ageing Res Rev 27: 93-107, 2016.
42. Srivastava SK and Beutler E: Cataract produced by tyrosinase and tyrosine systems in rabbitens in vitro. Biochem J 112: 421-425, 1969.

43. Liu X, Zhou P, Fan F, Li D, Wu J, Lu Y and Luo Y: CpG site methylation in CRYAA promoter affect transcription factor Sp1 binding in human lens epithelial cells. BMC Ophthalmol 16: 141, 2016.
44. Yang J, Zhou S, Guo M, Li Y and Gu J: Different alpha crystallin expression in human age-related and congenital cataract lens epithelium. BMC Ophthalmol 16: 67, 2016.
45. Zhou Z, Wang B, Luo Y, Zhou G, Hu S, Zhang H, Ma X and Qi Y: Major intrinsic protein (MIP) polymorphism is associated with age-related cataract in Chinese. Mol Vis 17: 2292-2296, 2011.
46. Takemoto L and Sorensen CM: Protein-protein interactions and lens transparency. Exp Eye Res 87: 496-501, 2008.
47. Ma X, Jiao X, Ma Z and Hejtmancik JF: Polymorphism rs7278468 is associated with Age-related cataract through decreasing transcriptional activity of the CRYAA promoter. Sci Rep 6: 23206, 2016.
48. Bhagyalaxmi SG, Srinivas P, Barton KA, Kumar KR, Vidyavathi M, Petrash JM, Reddy GB and Padma T: A novel mutation (F71L) in αA-crystallin associated with age-related cataract results in defective chaperone-like function despite unaltered structure. Biochim Biophys Acta 1792: 974-981, 2009.
49. Hubmacher D, Reinhardt DP, Plesec T, Schenke-Layland K and Apte SS: Human eye development is characterized by coordinated expression of fibrillin isoforms. Invest Ophthalmol Vis Sci 55: 7934-7944, 2014.
50. Saika S, Miyamoto T, Tanaka T, Ishida I, Ohnishi Y and Ooshima A: Latent TGFbeta binding protein-1 and fibrillin-1 in human capsular opacification and in cultured lens epithelial cells. Br J Ophthalmol 85: 1362-1366, 2001.
51. Ning S, Gao Y, Wang P, Li X, Zhi H, Zhang Y, Liu Y, Zhang J, Guo M, Han D, *et al*: Construction of a lncRNA-mediated feed-forward loop network reveals global topological features and prognostic motifs in human cancers. Oncotarget 7: 45937-45947, 2016.
52. Xu J, Feng L, Han Z, Li Y, Wu A, Shao T, Ding N, Li L, Deng W, Di X, *et al*: Extensive ceRNA-ceRNA interaction networks mediated by miRNAs regulate development in multiple rhesus tissues. Nucleic Acids Res 44: 9438-9451, 2016.
53. Marques AC and Ponting CP: Catalogues of mammalian long noncoding RNAs: Modest conservation and incompleteness. Genome Biol 10: R124, 2009.
54. Wu C, Lin H, Wang Q, Chen W, Luo H, Chen W and Zhang H: Discrepant expression of microRNAs in transparent and cataractous human lenses. Invest Ophthalmol Vis Sci 53: 3906-3912, 2012.
55. McLaughlin PA, Nirmalan GP, Tam KH and Robinson GA: Release of 131I from the ovary of the laying Japanese quail after injection of perchlorate, thiocyanate or iodide. J Endocrinol 63: 337-342, 1974.
56. McIllmurray MB, Price MR and Langman MJ: Inhibition of leucocyte migration in patients with large intestinal cancer by extracts prepared from large intestinal tumours and from normal colonic mucosa. Br J Cancer 29: 305-311, 1974.
57. Young TL, Matsuda T and Cepko CL: The noncoding RNA taurine upregulated gene 1 is required for differentiation of the murine retina. Curr Biol 15: 501-512, 2005.
58. Cumming GR, Dufresne C, Kich L and Samm J: Exercise electrocardiogram patterns in normal women. Br Heart J 35: 1055-1061, 1973.
59. Gosak M, Markovič R, Fajmut A, Marhl M, Hawlina M and Andjelić S: The analysis of intracellular and intercellular calcium signaling in human anterior lens capsule epithelial cells with regard to different types and stages of the cataract. PLoS One 10: e0143781, 2015.
60. Schmitz SU, Grote P and Herrmann BG: Mechanisms of long noncoding RNA function in development and disease. Cell Mol Life Sci 73: 2491-2509, 2016.
61. Sundararajan M, Thomas PA, Teresa PA, Anbukkarasi M and Geraldine P: Regulatory effect of chrysin on expression of lenticular calcium transporters, calpains, and apoptotic-cascade components in selenite-induced cataract. Mol Vision 22: 401-423, 2016.
62. Cekic O: Copper, lead, cadmium and calcium in cataractous lenses. Ophthalmic Res 30: 49-53, 1998.
63. Fischer RS, Lee A and Fowler VM: Tropomodulin and tropomyosin mediate lens cell actin cytoskeleton reorganization in vitro. Invest Ophthalmol Vis Sci 41: 166-174, 2000.
64. Zhou CJ and Lo WK: Association of clathrin, AP-2 adaptor and actin cytoskeleton with developing interlocking membrane domains of lens fibre cells. Exp Eye Res 77: 423-432, 2003.
65. Gostishchev VK and Khokhlov AM: Pathogenesis of trophic ulcers in varicose veins of the lower extremities. Khirurgiia (Mosk) 10: 100-105, 1991.