• Basic Research •

Screening of methylation genes in age-related cataract

Li Wang[1,2], Peng Li[3], Xiong Guo[2]

[1]Department of Medicine Technology Optometry, Xi'an Medical University, Xi'an 710021, Shaanxi Province, China
[2]Public Health School, Health Science Center, Xi'an Jiaotong University, Xi'an 710061, Shaanxi Province, China
[3]Department of Ophthalmology, the 451st Hospital of the PLA, Xi'an 710054, Shaanxi Province, China
Correspondence to: Li Wang. Department of Medicine Technology Optometry, Xi'an Medical University, Xi'an 710021, Shaanxi Province, China. teawangli@126.com
Received: 2018-02-12 Accepted: 2018-04-19

Abstract

• AIM: To analyze and screen the methylation status of whole-genome in age-related cataract samples.

• METHODS: Anterior lens capsule samples were collected from age-related cortical cataract patients over 50 years of age with LOCS III score of nuclear color ≥4 along with control subjects. DNAs were extracted and subjected to methylation microarray for the identification of methylated genes employing the high-throughput sequencing approach.

• RESULTS: Compared with the control group, 843 sites were found methylated, including 802 hypermethylation sites with 542 corresponding genes, 41 demethylation sites with 29 corresponding sites. COL4A1, GJA3, SIPA1L3 were confirmed by mass spectrometry, the results were consistent with high-throughput sequencing.

• CONCLUSION: DNA methylation microarrays is an efficient way for screening the aberrantly methylated genes. In this study, we are able to screen a few age-related cataract genes such as COL4A1, GJA3, and SIPA1L3 for their aberrant methylation patterns in cataract patients however further work is warranted to understand the significance of these findings.

• KEYWORDS: age-related cataract; methylation; genes; methylation microarrays

DOI:10.18240/ijo.2018.07.05

Citation: Wang L, Li P, Guo X. Screening of methylation genes in age-related cataract. *Int J Ophthalmol* 2018;11(7):1102-1107

INTRODUCTION

Age-related cataract (ARC), also known as senile cataract, is a clouding of the lens in the eye which leads to a decrease in vision less than 0.05, and is the most common cause of blindness and visual impairment globally. The World Health Organization data shows that cataract accounts for approximately 50% of all blindness[1]. The morbidity rate of ARC in Shaanxi Province is 36.66%. By far, surgical excision is the only effective therapeutic method for ARC, but surgical risk, complications and enormous economic burden on society have drawn extensive concerns. Epigenetics provides a new explication for ARC. As one of the important features of epigenetics, DNA methylation plays a key regulative role in ocular surface gene expression, and aberrant methylation has already been linked to cataract and other ophthalmological diseases. At present, the pathogenesis of cataract has not been determined. Generally, it is considered the result of multiple factors, like heredity, epigenetics, environment *etc*. This study selected the anterior lens capsule in ARC patients and the anterior lens capsule without cataract as the research objects, using high-throughput sequencing of Illumina Methylation Array to extract DNA, and using DNA methylation microarrays of Illumina 850K to screen the abnormally methylated genes.

SUBJECTS AND METHODS

Subjects The study was randomly divided into 3 experimental groups and 3 control groups. The inclusion criteria of experimental groups were: ARC patients over 50 years of age, cortical ARC with LOCS III score of nuclear color (NC) ≥4, regardless of diabetes, uveitis, glaucoma and other factors. The control groups were selected age matched anterior lens capsule without cataract in Eye Bank of Shaanxi Province, China. Patient's sex, age, classification of cortical, eye examination, and systemic disease were recorded, and the pictures of anterior segment were captured from each patients. The informed consent was obtained from all subjects, and the study was approved by the Ethics Committee of Xi'an Jiaotong University.

Methods

Whole-genome DNA methylation The anterior lens capsule of experimental and control groups were processed using Illumina Infinium Human Methylation850 BeadChip[2]. The procedure went in eight stages: 1) sample quality inspection: using appropriate method to extract genomic DNA, then using spectrophotometer to analyze qualitatively and quantitatively, the samples were diluted to 50 ng/uL in a final volume of 20 μL. Then 0.8% agarose gel electrophoresis (AGE) was used to separate DNA, the major band was clearly that its usually no less than 10 kb with no significant degradation, above size of 5 μg was processed DNA methylation downstream procedures; 2)

sulfite transformation: using the Illumina official recommendation of Zymo EZ DNA Methylation Kit to process the whole-genome DNA sulfite transformation; 3) DNA amplification: making MSA3 plate by adding 0.1 mol/L NaOH to denature DNA into single strands, then added the emulsion whole-genome amplification after neutralization for incubating at a constant temperature of 37℃; 4) DNA fragmentation, precipitation, resuspension: amplification of DNA, followed by the enzymatic digestion of the DNA fragment obtained, was used to precipitate DNA by adding isopropanol at 4℃, after drying in air, re-dissolve in the re-dissolved DNA was precipitated by adding the hybrid buffer; 5) DNA microarray hybridization: placing the resuspended DNA samples in hybridization oven for overnight. DNA fragments underwent denaturation with 50-base site-specific onto BeadChips during hybrid process; 6) wash BeadChip, single-base extension and staining: wash unhybridized and non-specifically hybridized DNA sample from the BeadChips to obtain the DNA plate. Add labeled nucleotides to extend the primers hybridized to the DNA. Stain the primers, disassemble the flow-through chambers, and coat the BeadChips for protection. Place the chip in XC4 reagent tube to ensure its surface with viscous transparent liquid, then dry in vacuum for 1h, the chip was wrapped so as to protect the signal for a long time; 7) chip scanning and data extraction: the iScan Reader used a laser to excite the fluorescence of the single-base extension product on the beads of the BeadChip sections. Light emissions from this fluorescence were then recorded in high-resolution images of the BeadChip sections. Data from these images were analyzed using Illumina's GenomeStudio Genotyping Module. The cytosine methylation of the CpG site expressed as β value, ranging from 0 (completely unmethylated) to 1 (completely methylated); 8) quality control of experimental: Illumina methylation microarrays introduced a series of quality control probes, which monitored the whole experimental process. It involved non-sample dependence controls (staining controls, extension controls, target removal controls, hybridization controls) and sample dependence controls [bisulfite conversion controls, stringency controls, and non-specific binding controls, non-polymorphic (NP) conrols][3].

Statistical analysis Color bias correction and quantile normalization were performed using the R package lumi 2.22.1[4], while probe bias was done using beta-mixture quantile normalization (BMIQv1.3)[5] to correct beta-values for the two assay-types on methylation level. Differential methylation analysis was performed using R package IMA 3.1.2[6] for calculating *P* value of each CpG. Empirical Bayes statistics of limma[7] was done using hypothesis testing method. Identification of significant CpG site was $P \leq 0.05$ with $|\beta| \geq 0.20$. All data were processed quality control, standards of quality control was $P \leq 0.05$, data that did not conform to the criterion was eliminated. For screening of significant CpG site was processed Cluster analysis using cluster3.0[8]. Using Kobas the genes' functional analysis for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) was done to reflect the difference in CpGs.

Functional annotation analysis of gene ontology and pathway GO analysis is a common bioinformatics method. The genes in each GO category were used as a functional gene set to measure genetic differences[9]. The GO database uses three separate ontologies, biological process (BP), molecular function (MF) and cellular component (CC). Pathway annotation analysis is an effectively enrichment analysis method, is used to identify related proteins within a pathway or building pathway de novo from the proteins of interest, hence considering its joint effect of multiple genes, pathway annotation analysis has important significance for research in complex diseases[10]. In our study, pathway enrichment analysis included KEGG, Reactome, Panther and BioCyc. The criterion of GO and pathway analysis was $P \leq 0.05$. Moreover, GO and pathway analysis were processed using DAVID (http://david.abcc.ncifcrf.gov/). DAVID could provide functional interpretation of large lists of genes derived from genomic studies.

Mass spectrometry analysis Mass spectrometry (MS) is an analytical technique that ionizes chemical species and sorts the ions based on their mass-to-charge ratios, and it has been applied successfully in many fields. The new samples were processed by MS analysis, PCR primer design was performed by Methprimer, methylated data of anterior lens capsule was generated by Epityper software version 1.0 (Sequenom, CA, USA).

RESULTS

Differential Methylation Sites Using GenomeStudio Software to screen the original data, and the original signal intensity value of each site was obtained according to the Methylation Analysis Algorithms of Illumina. Then color bias correction and quantile normalization were caused by different types of fluorescence and probe bias that obtained the significant CpG sites (β value; Table 1). According to above-mentioned analytical β value of CpG sites, Pearson correlation coefficient was calculated (Table 2, Figure 1).

The results of differential methylation analysis showed that compared with the controls, 843 sites were methylated, including 802 hypermethylation sites with 542 corresponding genes, 41 demethylation sites with 29 corresponding genes sites (Table 3).

Gene Ontology and Pathway Functional Enrichment Analysis GO analysis identified 1478 meaningful GO terms, of which biological process for 1138, cellular component for 160, molecular function for 180. The top 30 of enrichment GO terms, shown in Table 1 and Figure 2, were found by GO analysis. In addition, pathway annotation analysis identified

Table 1 List of the top 30 of enrichment GO terms

Term	Database	Corrected *P*
Multicellular organismal development	Biological process	4.11×10^{32}
Anatomical structure development	Biological process	5.51×10^{32}
Single-organism developmental process	Biological process	5.54×10^{32}
Developmental process	Biological process	7.77×10^{31}
System development	Biological process	7.77×10^{31}
Anatomical structure morphogenesis	Biological process	7.11×10^{29}
Cell periphery	Cellular component	4.9×10^{28}
Plasma membrane part	Cellular component	8.76×10^{28}
Regulation of multicellular organismal process	Biological process	2.11×10^{27}
Plasma membrane	Cellular component	1.12×10^{26}
Single-organism process	Biological process	1.16×10^{26}
Regulation of developmental process	Biological process	4.7×10^{26}
Single-multicellular organism process	Biological process	5.94×10^{26}
Nervous system development	Biological process	7.27×10^{26}
Multicellular organismal process	Biological process	1.62×10^{24}
Generation of neurons	Biological process	3.12×10^{23}
Cell development	Biological process	3.86×10^{23}
Regulation of multicellular organismal development	Biological process	3.86×10^{23}
Signaling	Biological process	1.31×10^{22}
Single organism signaling	Biological process	1.52×10^{22}
Neuron differentiation	Biological process	4.6×10^{22}
Cell junction	Cellular component	4.6×10^{22}
Neuron development	Biological process	8.07×10^{22}
Regulation of cell differentiation	Biological process	2.42×10^{21}
Positive regulation of biological process	Biological process	2.76×10^{21}
Regulation of cell communication	Biological process	2.9×10^{21}
Membrane region	Cellular component	2.9×10^{21}
Regulation of signaling	Biological process	3.4×10^{21}
Cell differentiation	Biological process	3.74×10^{21}
Neuron part	Cellular component	5.4×10^{21}

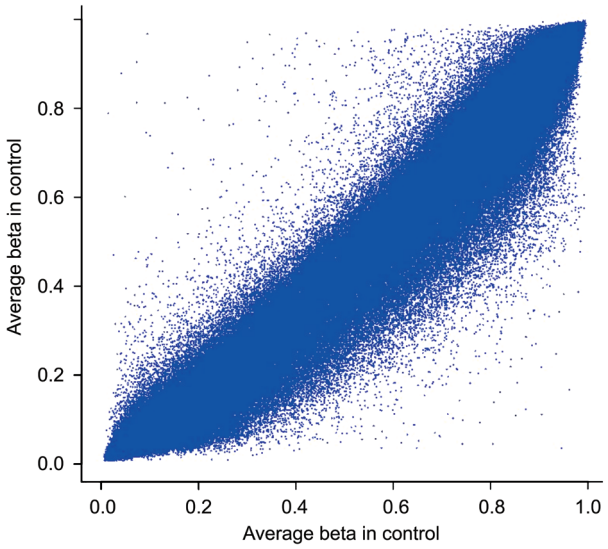

Figure 1 Pearson correlation coefficient.

215 pathways, of which BioCyc for 0, KEGG for 60, Reactome for 138, Panther for 17. The top 30 of significant pathways and distribution diagram of KEGG pathway were shown in Table 2, Figures 3 and 4.

Mass Spectrometry Analysis Sequenom methylation was used to measure the methylation status of COL4A, SIPA1L3 and GJA3 promoter in ARC and control groups. And cg18792803, cg07724793, cg08036955 of COL4A, cg25690202, cg17671383 of SIPA1L3, cg26133081, cg03129014, cg07039404 of SIPA1L3 were assessed according to the characteristic of primers design. Meanwhile, the study made comparative analysis of eight CpG sites in ARC and NC groups by *t*-test and Wilcox test. There had significant difference in methylated level between two groups (*t*-test, $P<0.05$; Wilcox test, $P<0.05$; Table 3), this result was consistent with the findings of methylation microarrays. Moreover, the methylation level of eight CpG sites in ARC group was higher than that in control group.

DISCUSSION

Cataract is a major cause of blindness around the world, as well as the leading cause of blindness in China. Cataracts become more common with age and increase obviously along with age, especially in those older than 50 years of age. Medium and large-sized cities in China have entered a stage of aging,

Table 2 List of the top 30 of significant pathways.

Term	Database	Corrected *P*
Axon guidance	Reactome	1.98×10^{12}
Developmental Biology	Reactome	3.27×10^{11}
Pathways in cancer	KEGG PATHWAY	1.75×10^{8}
Signal transduction	Reactome	5.28×10^{8}
Axon guidance	KEGG PATHWAY	9.88×10^{8}
Neuronal system	Reactome	1.67×10^{7}
Hemostasis	Reactome	2.62×10^{7}
Focal adhesion	KEGG PATHWAY	2.9×10^{7}
Wnt signaling pathway	PANTHER	3.52×10^{7}
Signalling by NGF	Reactome	3.21×10^{6}
Transmembrane transport of small molecules	Reactome	9.11×10^{6}
Hippo signaling pathway	KEGG PATHWAY	1.26×10^{5}
Muscle contraction	Reactome	3.47×10^{5}
Angiogenesis	PANTHER	3.7×10^{5}
Signaling by platelet derived growth factor	Reactome	3.85×10^{5}
Signaling by vascular endothelial growth factor	Reactome	3.85×10^{5}
VEGFA-VEGFR2 Pathway	Reactome	8.05×10^{5}
Signaling by epidermal growth factor receptor	Reactome	9.87×10^{5}
NGF signalling via TRKA from the plasma membrane	Reactome	1.16×10^{4}
Cardiac conduction	Reactome	1.4×10^{4}
DAP12 signaling	Reactome	1.85×10^{4}
Extracellular matrix organization	Reactome	1.85×10^{4}
Fc epsilon receptor (FCERI) signaling	Reactome	1.85×10^{4}
Rap1 signaling pathway	KEGG PATHWAY	1.85×10^{4}
Phospholipase D signaling pathway	KEGG PATHWAY	1.85×10^{4}
DAP12 interactions	Reactome	1.85×10^{4}
Downstream signaling of activated FGFR1	Reactome	2.03×10^{4}
Glutamatergic synapse	KEGG PATHWAY	2.03×10^{4}
Downstream signal transduction	Reactome	2.2×10^{4}
Circadian entrainment	KEGG PATHWAY	2.2×10^{4}

NGF: Nerve growth factor; VEGFA: Vascular endothelial growth factor A; VEGFR2: Vascular endothelial growth factor receptor 2; TRKA: Tropomyosin receptor kinase A; DAP12: DNAX-activation protein 12; FGFR1: Fibroblast growth factor receptor 1.

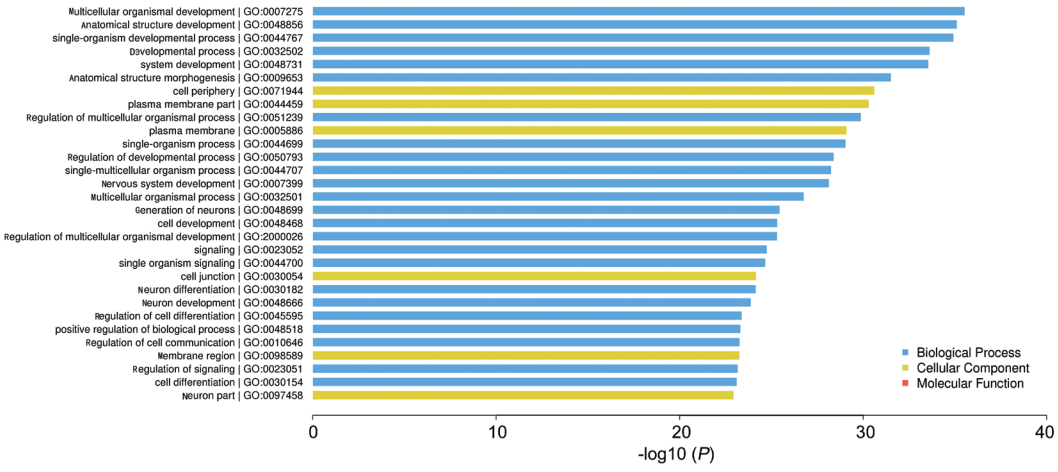

Figure 2 The top 30 of enrichment GO terms.

as the number of the aged increases, the number of ARC also grows rapidly. From some reports, about 60 million people are blind due to cataracts, there are 5 million people with a vision less than 0.3, and the number is growing by 0.4-1.2 million people per year, while ARC have the maximum percentage. DNA methylation is a critical regulator of gene expression in the eye, some studies have been confirmed that aberrant methylation associates with cataract, age-related macular and eye diseases[11-12]. In order to identify sites of differential methylation and differential gene expression in ARC, we exploited the anterior lens capsule in ARC patients for whole-genome methylation analysis. We found that anterior lens

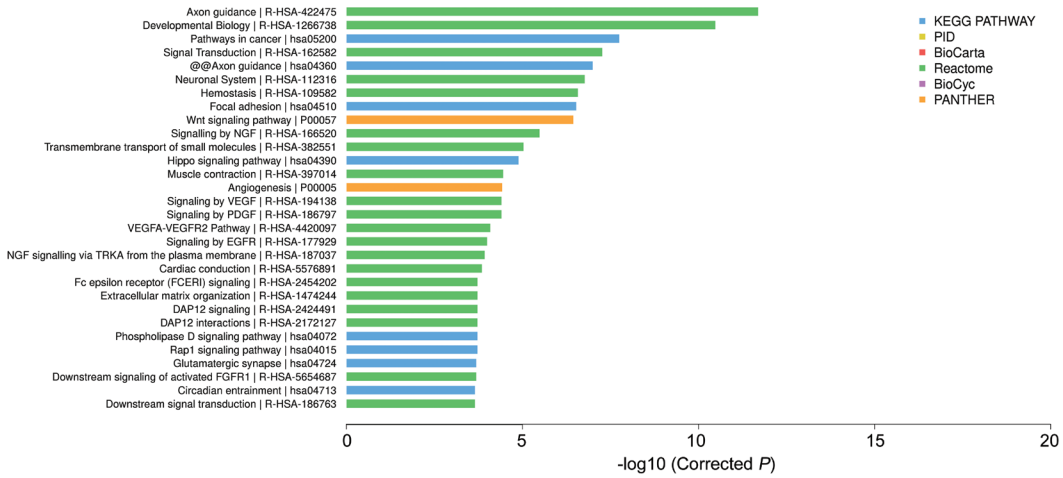

Figure 3 The top 30 of significant pathways.

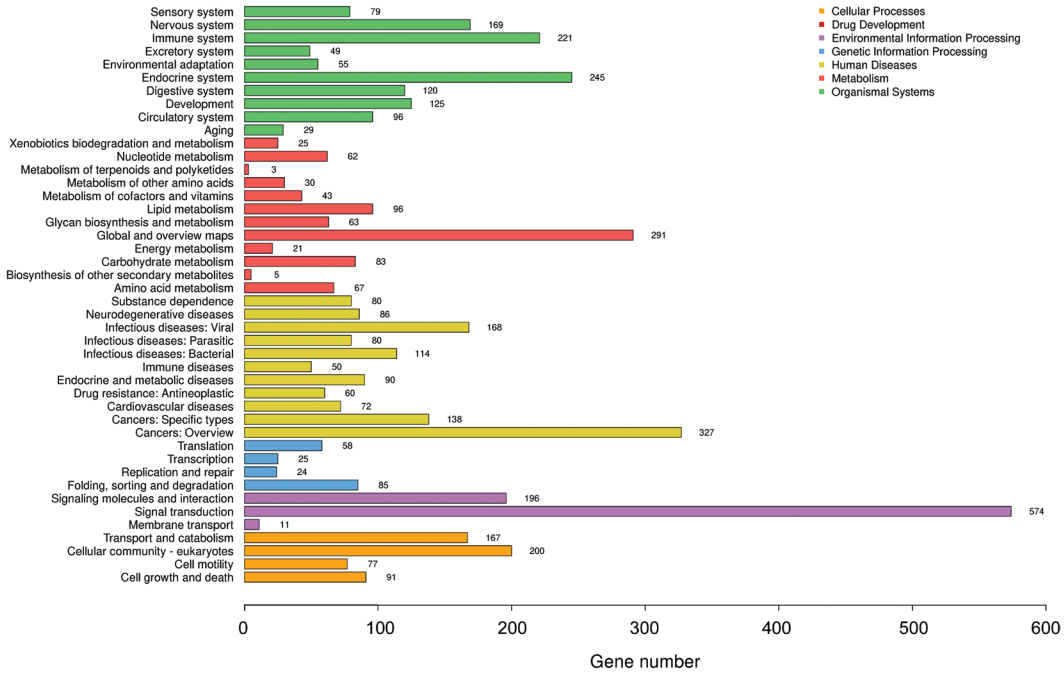

Figure 4 The top 30 of KEGG pathways.

Table 3 The results of COL4A1, SIPA1L3, GJA3 methylation level

mean±SD

CpG	ARC group	Control group	[a]*P*	[b]*P*
COL4A1				
cg18792803	0.262±0.013	0.129±0.052	0.014	0.026
cg07724793	0.562±0.023	0.345±0.083	0.016	0.027
cg08036955	0.280±0.034	0.076±0.022	0.025	0.029
SIPA1L3				
cg25690202	0.576±0.029	0.245±0.093	0.048	0.037
cg17671383	0.448±0.054	0.202±0.023	0.032	0.032
GJA3				
cg26133081	0.854±0.063	0.586±0.027	0.037	0.043
cg03129014	0.497±0.053	0.245±0.039	0.039	0.042
cg07039404	0.514±0.037	0.258±0.026	0.042	0.037

CpG: Cytosine polyguanine; ARC: Age-related cataract; [a]*t*-test; [b]Wilcox test.

capsule of ARC was obviously changed, meanwhile COL4A1, GJA3, SIPA1L3 were examined by MS, the results of MS and high-throughput analysis revealed that COL4A1, GJA3, SIPA1L3 played a key role in ARC.

Our findings suggest that COL4A1 associates with ARC, and represents as hypermethylation target (*P*=0.017, β=0.241). COL4A1 encodes the α1 chain of type IV collagen, type IV collagen is the main collagenous constituent of basement membranes. COL4A1 is located in chromosome 13q34, mRNAs are 6549 bp long, including 52 exons. Pöschl *et al*[13] reported that the first intron of COL4A1 interact with CTC binding factor, suggesting that the mutation is related to transcription.

Many studies have shown that mutations in the gene have been linked to diseases of the kidney, eye, cardiovascular system *etc*. Van Agtamel *et al*[14] indicated that dominant mutant mice with COL4A1 gene mutations had the symptom of iris corneal adhesion, hydrophthalmos, and corneal opacification. Gould *et al*[15] demonstrated that mutation of COL4A1 may cause a spectrum of cerebrovascular phenotypes and that persons

with COL4A1 mutations may be predisposed to hemorrhage, especially after environmental stress. And some studies have found that there are 13 single mutation sites of COL4A1 that could cause Bubble cataract and lens abnormality[16].

Our study also found that gap-junction protein alpha3 (GJA3) associates with ARC, and represents as hypermethylation target as well (P=0.013, β=0.283). Jiang *et al*[17] reported that a mutation is located in the first transmembrane region of GJA3, and GJA3 is important in the maintenance of optical clarity. Studies by Yuan *et al*[18] and Liu *et al*[19] noted that intercellular communication through the extensive network of gap junctions including GJA3 facilitates intercellular exchange of molecules including antioxidants, it is vital for the maintenance of eye lens transparency. And the expression of GJA3 gene is lower than that in control group for anterior lens capsule in ARC patients[19]. Consequently, the results showed that the down-regulated expression of GJA3 gene plays an important role in the development of cataract.

SIPA1L3 associates with ARC, and represents as hypermethylation target too (P=0.03, β=0.223). Greenlees *et al*[20] reported that SIPA1L3 downregulation reveals morphogenetic and cell polarity abnormalities. And decreased expression of SIPA1L3 in zebrafish and mouse could cause severe lens and eye abnormalities[20]. But the relationship of SIPA1L3 and ARC needs further work towards its contribution to ARC molecular biology experiment.

In conclusion, we used high-throughput sequencing to screen the abnormal methylation genes of ARC by methylation microarrays, and the aberrant methylation genes were further biologically validated by MS. We found that aberrant methylation expression of COL4A1, GJA3, SIPA1L3 associates was associated with ARC. Induced methylation alteration is an important mechanism for ARC, the strategies of methylation for ARC in the future can provide a new research direction.

ACKNOWLEDGEMENTS

Foundation: Supported by the Shaanxi Provincial Department of Science and Technology Agency Project (No.2017SF-288).

Conflicts of Interest: Wang L, None; **Li P,** None; **Guo X,** None.

REFERENCES

1 Wu QY, Luan J. Cataract and oxidative damage. *Mordern Medical Journal* 2010;38(3):306-310.

2 Mark MW, Smith AK, Wolf EJ, Maniates H, Stone A, Schichman SA, McGlinchey RE, Milberg W, Miller MW. The correlation of methylation levels measured using Illumina 450K and EPIC BeadChips in blood samples. *Epigenomics* 2017;9(11):1363-1371.

3 Lorés-Motta L, Paun CC, Corominas J, Pauper M, Geerlings MJ, Altay L, Schick T, Daha MR, Fauser S, Hoyng CB, den Hollander AI, de Jong EK. Genome-wide association study reveals variants in CFH and CFHR4 associated with systemic complement activation: implications in age-related macular degeneration. *Ophthalmology* 2018;1(17):32520-32524.

4 Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 2008;24(13):1547-1548.

5 Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* 2013;29(2):189-196.

6 Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB, Johnson CS, Smiraglia DJ, Liu S. IMA: an R package for high-throughput analysis of Illumina's 450K Infinium methylation data. *Bioinformatics* 2012;28(5):729-730.

7 Smyth GK. Limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor* 2005;397-420.

8 de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. *Bioinformatics* 2004;20(9):1453-1454.

9 Gene Ontology Consortium. The Gene Ontology project in 2008. *Nucleic Acids Res* 2008;36(Database issue):D440-D444.

10 Wang K, Li M, Bucan M. Pathway-based approaches for analysis of genomewide association studies. *Am J Hum Genet* 2007;81(6):1278-1283.

11 Liu MM, Chan CC, Tuo J. Epigenetics in ocular diseases. *Curr Genomics* 2013;14(3):166-172.

12 Zhang B, Bai J, Dong L, Liu P, Zheng Y. Advances in vitamin D deficiency and DNA methylation in age-related macular degeneration. *Guoji Yanke Zazhi (Int Eye Sci)* 2016;16(12):2225-2228

13 Pöschl E, Pollner R, Kühn K. The genes for the alpha 1(IV) and alpha 2(IV) chains of human basement membrane collagen type IV are arranged head-to-head and separated by a bidirectional promoter of unique structure. *EMBO J* 1988;7(9):2687-2695.

14 Van Agtmael T, Bailey MA, Schlötzer-Schrehardt U, Craigie E, Jackson IJ, Brownstein DG, Megson IL, Mullins JJ. Col4a1 mutation in mice causes defects in vascular function and low blood pressure associated with reduced red blood cell volume. *Hum Mol Genet* 2010;19(6):1119-1128.

15 Gould DB, Phalan FC, van Mil SE, Sundberg JP, Vahedi K, Massin P, Bousser MG, Heutink P, Miner JH, Tournier-Lasserve E, John SW. Role of COL4A1 in small-vessel disease and hemorrhagic stroke. *N Engl J Med* 2006;354(14):1489-1496.

16 Gould DB, Phalan FC, Breedveld GJ, van Mil SE, Smith RS, Schimenti JC, Aguglia U, van der Knaap MS, Heutink P, John SW. Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. *Science* 2005;308(5725):1167-1171.

17 Jiang H, Jin Y, Bu L, Zhang W, Liu J, Cui B, Kong X, Hu L. A novel mutation in GJA3 (connexin46) for autosomal dominant congenital nuclear pulverulent cataract. *Mol Vis* 2003;9:579-583.

18 Yuan L, Guo Y, Yi J, Xiao J, Yuan J, Xiong W, Xu H, Yang Z, Zhang J, Deng H. Identification of a novel GJA3 mutation in congenital nuclear cataract. *Optom Vis Sci* 2015;92(3):337-342.

19 Liu YY, Wan WC, Yang G, Pang XN, Yang GG, Jin XM. Analysis of GJA3 mutation associated with a Chinese family with autosomal dominant congenital cataract by whole-exome sequencing. *Chin J Exp Ophthalmol* 2016;34(10):916-919.

20 Greenlees R, Mihelec M, Yousoof S, *et al*. Mutations in SIPA1L3 cause eye defects through disruption of cell polarity and cytoskeleton organization. *Hum Mol Genet* 2015;24(20):5789-5804.