Integrated bioinformatics analysis of aberrantly methylated-differentially expressed genes and pathways in age-related macular degeneration

Yinchen Shen
Shanghai General Hospital

Mo Li
Shanghai Jiao Tong University

Kun Liu
Shanghai General Hospital

Xiaoyin Xu
Shanghai General Hospital

Shaopin Zhu
Shanghai General Hospital

Ning Wang
Shanghai General Hospital

Wenke Guo
NHC Key Lab. of Reproduction Regulation

Qianqian Zhao
NHC Key Lab. of Reproduction Regulation

Ping Lu
NHC Key Lab. of Reproduction Regulation

Fudong Yu
NHC Key Lab. of Reproduction Regulation

Xun Xu

Research article

Keywords: Methylation, Gene expression, Enrichment analysis, Network analysis, Age-related macular degeneration

Abstract

Background: Age-related macular degeneration (AMD) represents the leading cause of visual impairment in the aging population. The goal of this study was to identify aberrantly methylated-differentially expressed genes (MDEGs) in AMD and explore the involved pathways by integrated bioinformatic analysis.

Methods: Data of expression profiling GSE29801 and methylation profiling GSE102952 were obtained from the Gene Expression Omnibus database. We analyzed differentially methylated genes and differentially expressed genes in R software. Functional enrichment and protein−protein interaction (PPI) network analysis were performed using R package and Search Tool for the Retrieval of Interacting Genes online database. Hub genes were identified using Cytoscape.

Results: 827 and 592 genes showed high and low expression, respectively, in GSE29801; 4117 hyper-methylated genes and 511 hypo-methylated genes were detected in GSE102952. After overlapping, we categorized 153 genes as hyper-methylated, low-expression genes (Hyper-LGs) and 24 genes as hypo-methylated, high-expression genes (Hypo-HGs). Four Hyper-LGs (CKB , PPP3CA , TGFB2 , SOCS2) overlapped with AMD risk genes in Public Health Genomics and Precision Health Knowledge Base. KEGG pathway enrichment analysis indicated Hypo-HGs were enriched in the calcium signaling pathway, whereas Hyper-LGs were enriched in sphingolipid metabolism. In GO analysis, Hypo-HGs were enriched in fibroblast migration, membrane raft, coenzyme binding, etc. Hyper-LGs were enriched in mRNA transport, nuclear speck, DNA binding, etc. In PPI networks analysis, 23 nodes and 2 edges were established from Hypo-HGs, and 151 nodes and 73 edges were established from Hyper-LGs. Hub genes (DHX9 , MAPT , PAX6) showed the greatest overlap.

Conclusion: This study revealed potentially aberrantly MDEGs and pathways in AMD, which may improve the understanding of this disease.

Background

Age-related macular degeneration (AMD) is the leading cause of adult blindness in developed countries, particularly in those over 55 years of age. Worldwide, this condition accounts for 7% of cases of blindness [1, 2] and is expected to affect 288 million people by the year 2040 [3]. AMD first appears as drusen (dry AMD), and advances to late AMD (wet AMD) characteristic by choroidal neovascularization (CNV) [4]. Drusen is composed of lipoproteinaceous deposits and acellular debris [5]. CNV involves the growth of abnormal new blood vessels originating from the choroid through a break in the Bruch's membrane, which then invade the subretinal pigment epithelium or sub-retinal space, resulting in severe vision loss [6, 7]. AMD affects central fine vision, significantly impairs a patient's ability to drive, read, and recognize faces, and greatly affects quality of life [8]. Previous studies indicated that many environmental factors are associated with an increased risk of AMD, such as age, race, smoking, obesity, and hypertension [9, 10]. Additionally, genetic factors are regarded important in the initiation and progression of AMD [11, 12].

No completely effective cure is available for this disease. Current treatment for wet AMD (i.e. advanced stages) requires repeated intravitreal anti-vascular endothelial growth factor injections, which represent a significant risk of infection and treatment burden for both the patients and ophthalmologists [5]. Other than tumor tissue samples, ocular fundus tissue samples (i.e. retina and choroid) of patients with AMD are quite difficult to obtain in practice, which has restricted our understanding of this blinding disease. Over the past few years, most genetic studies of AMD were case−control genome-wide association studies (GWASs) of single-nucleotide polymorphisms in the patients' peripheral blood [13, 14, 15], which were valuable but provided limited information.

With the rapid development of gene assay technology, studies of disease pathogenesis are no longer limited to gene deletion, gene mutation, gene insertion, etc. Microarrays based on high-throughput platforms are useful and efficient tools for searching for meaningful genes and epigenetic alternations in the identification of biomarkers for diagnosis or prognosis [16]. To better explore the molecular mechanism of AMD, it is necessary to conduct full transcriptome analysis at the tissue level; however, as mentioned above, obtaining ocular tissue samples is difficult. Compared to retina or choroid tissue, blood samples are relatively easy to obtain from patients with AMD. Therefore, it would be helpful to evaluate gene expression results from the ocular tissue of patients with AMD as biomarkers in the blood. In the present study, data from gene expression profiling microarrays of human retinal and choroidal samples from the Iowa and Oregon cohort AMD and control donors (GSE29801: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29801) [17] as well as gene methylation profiling microarrays of the peripheral blood of subjects with AMD (GSE102952: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102952) were integrated and analyzed using a series of bioinformatics tools. More precise screening results were obtained by overlapping these two AMD data sets. Few studies have attempted to combine gene expression profile microarrays and gene methylation profile microarrays to understand the development of AMD. The introduction of DNA methylation characteristics in the blood is useful for understanding the characteristics of AMD disease at the tissue level.

In the present study, data from gene expression profiling microarrays and gene methylation profiling microarrays were integrated and analyzed. Functional enrichment and protein−protein interaction (PPI) network analyses of screened genes were performed using the R package "clusterProfiler" and Search Tool for the Retrieval of Interacting Genes (STRING) online database. We identified methylated genes in the peripheral blood, which may be useful as biomarkers for the precise diagnosis and treatment of AMD.

Methods

The need for ethics approval was waived by the ethics committee of Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China (The document was attached as an additional file).

Microarray data information

We identified methylated-differentially expressed genes (MDEGs) between AMD and control samples by analyzing mRNA microarray and methylation profiling datasets. In this study, the gene expression profiling dataset (GSE29801: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29801) and gene

methylation dataset (GSE102952: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102952) were downloaded from the Gene Expression Omnibus (GEO, https://www.ncbi. nlm.nih.gov/geo/) of the National Center for Biotechnology Information (NCBI). The gene expression profiling data was based on the mRNA from the macular region of human donor eye of retina and retinal pigment epithelium (RPE)-choroids. The gene methylation microarray data was performed using genome-wide DNA methylation profiling of peripheral blood.

Microarray data from GSE29801 included 177 samples from the macular or extramacular region of human donor eye RPE-choroids and 118 samples from the macular or extramacular region of human donor retina with no reported ocular disease, with possible preclinical AMD or AMD [17]. RPE-choroid and retinal samples were isolated from human donor eyes obtained from the University of Iowa (GSH) and Lions Eye Bank of Oregon. The Iowa eyes were selected from a well-characterized repository derived from over 3900 donors. The Oregon eyes were generally classified as AMD based on medical histories confirmed by ophthalmological records [17, 18]. Global transcriptome profiling was carried out using the Agilent Whole Human Genome 4 × 44 K in situ oligonucleotide array platform (G4112F, Agilent Technologies, Santa Clara, CA, USA). After removing redundant data, the microarray data of 41 patients with AMD and 42 normal controls were included in our analysis [17]. Firstly, the authors removed redundant data and selected the subjects who had records of both Macular Retina and Macular RPE-choroid. Such selection criteria was due to the fact that AMD mainly affected central vision acuity, i.e., the central macular area. Secondly, we only analyzed samples that had stronger expression level out of Macular Retina and Macular RPE-choroid for one subject. The objective of such approach was to maximize the differentially expressed genes identified.

For the gene methylation microarray data, Oliver et al. previously performed genome-wide DNA methylation profiling of blood from 9 patients with AMD and 9 controls based on the platform GPL13534 Illumina HumanMethylation 450 BeadChip (HumanMethylation450_15017482, San Diego, CA, USA), which covers approximately 450,000 CpG sites in different gene regions including the transcription start sites 1500 and 200, 5' untranslated region, 1st exon, body, and 3' untranslated region. The authors generously shared their original data online for public use (GSE102952: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102952).The methylation profiles data of all 9 patients with AMD and 9 normal controls were included in this analysis.

Data processing to identify differentially expressed genes (DEGs) and differentially methylated genes (DMGs)

We used the Limma package in R software (version 3.4.2; Bell Laboratories, formerly AT&T, now Lucent Technologies, Murray Hill, NJ, USA) to analyze GSE29801 and GSE102952 to identify DEGs and DMGs. A P < 0.05 was regarded as statistically significant. For DEGs, we set the cut-off standard as P < 0.05 and absolute value of log (fold-change) > median (fold-change). For DMGs, we set the cut-off criteria as P < 0.05 and absolute value of log (fold-change) > 3/4 sort [summary log (fold-change)].

For further analysis, hypomethylation-high expression genes (Hypo-HGs) were obtained by overlapping hypomethylated and up-regulated genes; hypermethylation-low expression genes (Hyper-LGs) were obtained by overlapping hypermethylated and down-regulated genes.

Functional enrichment analysis

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed for the selected genes (Hypo-HGs and Hyper-LGs), and the enrichment results were illustrated using R package "clusterProfiler". This package can be used to extract biological meaning from large gene lists. We performed GO term enrichment analysis under the following three sub-ontologies: biological process (BP), molecular function (MF), and cellular component (CC). The cut-off criterion of significantly enriched KEGG pathways was $P < 0.05$.

Comprehensive PPI network

Determination of the comprehensive PPI is important for detecting the molecular mechanisms of AMD. In this study, we used the online STRING (version 11.0) to construct the network of Hypo-HGs and Hyper-LGs. STRING is an online database used to predict PPIs, which are essential for interpreting the molecular mechanisms of key cellular activities in AMD. The cut-off standard was defined as an interaction score of 0.4. The results were visualized in Cytoscape software (version 3.5.1). Hub genes were defined as the top 3 genes that appeared most frequently according to all cytoHubba ranking methods using Cytoscape software. Subsequently, the Molecular Complex Detection (MCODE) algorithm in Cytoscape software was used to screen the modules. An MCODE score > 3 and node number > 3 were used as the criteria to define a module.

Results

Identification of DEGs and DMGs in AMD

To identify DEGs or DMGs, we used the expression profile from GSE29801 (containing RPE-choroid and retina tissue samples from 41 patients with AMD and 42 normal samples) and methylation profile from GSE102952 (containing peripheral blood samples of 9 AMD patients and 9 normal controls) after data preprocessing and quality assessment using R software. We identified 827 high-expression genes and 592 low-expression genes and 4117 hypermethylated genes and 511 hypomethylated genes. The top 100 most significantly DEGs of GSE29801 are shown in Table S1, and the top 100 most significantly DMGs of GSE102952 are shown in Table S2.

Identification of aberrantly methylated-DEGs in AMD

To further explore the aberrantly methylated DEGs, Hypo-HGs were obtained by overlapping hypomethylated and up-regulated genes; Hyper-LGs were obtained by overlapping hypermethylated and down-regulated genes. We identified 24 Hypo-HGs and 153 Hyper-LGs. The flowchart of this study is presented in Fig. 1. All genes are shown in Table S3. To confirm the reliability of the results, we compared aberrantly MDEGs from the GSE29801 and GSE102952 datasets with genes in the Public Health Genomics and Precision Health Knowledge Base (PHGKB) (version. 5.8) (https://phgkb.cdc.gov/PHGKB/startPagePhenoPedia.action) studied in the category of "macular degeneration"; the results are shown in Fig. 2. We screened four Hyper-LGs (CKB, PPP3CA, TGFB2, and SOCS2) which overlapped with potential AMD risk genes in PHGKB. However, no Hypo-HG overlapped in PHGKB.

Functional enrichment analysis

Functional enrichment analysis was conducted using the R package "clusterProfiler", and all significantly enriched KEGG pathways with the 24 Hypo-HGs and 153 Hyper-LGs are shown in Table 1. The results of KEGG pathway enrichment analysis indicated that Hypo-HGs were significantly enriched in the phosphatidylinositol signaling system and calcium signaling pathway, whereas Hyper-LGs were significantly enriched in amphetamine addiction, morphine addiction, and sphingolipid metabolism. The top 5 GO terms in each category in which the 24 Hypo-HGs and 153 Hyper-LGs were significantly involved are shown in Table 2 and Table 3, respectively. Functional enrichment analysis suggested that the 24 Hypo-HGs were enriched in the BP of fibroblast migration and positive regulation of neurological system process. The GO CC category revealed enrichment in membrane raft and membrane microdomain. The MF category showed enrichment for factors involved in neuropeptide receptor activity and coenzyme binding (Table 2). The 153 Hyper-LGs were enriched in the BP category of mRNA transport, etc. The CC category revealed enrichment in nuclear speck, neuronal cell body, etc. The MF category indicated enrichment in DNA binding and phosphoric ester hydrolase activity (Table 3).

Table 1

Kyoto Encyclopedia of Genes and Genomes pathway analysis of aberrantly methylated differentially expressed genes in age-related macular degeneration (AMD). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed for the selected genes. All significantly enriched KEGG pathways with the 24 hypo-methylated, high-expression genes (Hypo-HGs) and 153 hyper-methylated, low-expression genes (Hyper-LGs) are shown. The cut-off criterion was P < 0.05. The results indicated that Hypo-HGs were significantly enriched in the phosphatidylinositol signaling system and calcium signaling pathway, whereas Hyper-LGs were significantly enriched in amphetamine addiction, morphine addiction, and sphingolipid metabolism.

Category	ID	Description	GeneRatio	BgRatio	P value	Q value	geneSymbol	Count
Hypomethylated with high expression								
KEGG_PATHWAY	hsa04070	Phosphatidylinositol signaling system	2/13	99/7866	0.01117	0.28231	PIP5K1A/PIP4P2	2
KEGG_PATHWAY	hsa04020	Calcium signaling pathway	2/13	193/7866	0.03911	0.31445	NTSR1/TACR1	2
Hypermethylated with low expression								
KEGG_PATHWAY	hsa05031	Amphetamine addiction	3/59	68/7866	0.01423	0.73750	GRIN2C/PPP3CA/STX1A	3
KEGG_PATHWAY	hsa05032	Morphine addiction	3/59	91/7866	0.03057	0.73750	GRK4/PDE10A/PDE4D	3
KEGG_PATHWAY	hsa00600	Sphingolipid metabolism	2/59	47/7866	0.04820	0.73750	PLPP2/SGPL1	2

Table 2. Gene Ontology analysis of hypomethylated with high expression genes in age-related macular degeneration (AMD)

Category	Term	Description	GeneRatio	BgRatio	P value	Q value	GeneSymbol	Count
Hypomethylated with high expression								
GOTERM_BP	GO:0010761	fibroblast migration	2/22	41/18493	0.00108	0.11666	DDR2/PIP5K1A	2
GOTERM_BP	GO:0031646	positive regulation of neurological system process	2/22	58/18493	0.00214	0.11666	NTSR1/TNR	2
GOTERM_BP	GO:1901616	organic hydroxy compound catabolic process	2/22	72/18493	0.00328	0.11666	LDHD/NTSR1	2
GOTERM_BP	GO:0003333	amino acid transmembrane transport	2/22	86/18493	0.00465	0.11666	NTSR1/SLC7A4	2
GOTERM_BP	GO:0031644	regulation of neurological system process	2/22	127/18493	0.00988	0.11666	NTSR1/TNR	2
GOTERM_CC	GO:0045121	membrane raft	3/23	304/19659	0.00516	0.13697	NTSR1/SKAP1/TNR	3
GOTERM_CC	GO:0098857	membrane microdomain	3/23	305/19659	0.00520	0.13697	NTSR1/SKAP1/TNR	3
GOTERM_CC	GO:0098589	membrane region	3/23	316/19659	0.00574	0.13697	NTSR1/SKAP1/TNR	3
GOTERM_CC	GO:0032280	symmetric synapse	1/23	11/19659	0.01280	0.16553	NTSR1	1
GOTERM_CC	GO:0061827	sperm head	1/23	11/19659	0.01280	0.16553	TACR1	1
GOTERM_MF	GO:0008188	neuropeptide receptor activity	2/20	50/17632	0.00145	0.07933	NTSR1/TACR1	2
GOTERM_MF	GO:0050662	coenzyme binding	3/20	285/17632	0.00389	0.09513	GCAT/LDHD/UXS1	3
GOTERM_MF	GO:0034596	phosphatidylinositol phosphate 4-phosphatase activity	1/20	10/17632	0.01129	0.09513	PIP4P2	1
GOTERM_MF	GO:0008528	G protein-coupled peptide receptor activity	2/20	145/17632	0.01158	0.09513	NTSR1/TACR1	2
GOTERM_MF	GO:0015174	basic amino acid transmembrane transporter activity	1/20	11/17632	0.01241	0.09513	SLC7A4	1

Table 2. Gene ontology (GO) pathway enrichment analyses were performed for the selected genes. The cut-off criterion was $P < 0.05$. The top 5 GO terms in each category in which the 24 hypo-methylated, high-expression genes were significantly involved are shown. They were enriched in the biological process of fibroblast migration and positive regulation of neurological system process. The cellular component category revealed enrichment in membrane raft and membrane microdomain. The molecular function category showed enrichment for factors involved in neuropeptide receptor activity and coenzyme binding.

Table 3. Gene Ontology analysis of hypermethylated with low expression in age-related macular degeneration (AMD)

Category	Term	Description	GeneRatio	BgRatio	P value	Q value	GeneSymbol
Hypermethylated with low expression							
GOTERM_BP	GO:0051028	mRNA transport	7/141	151/18493	0.00016	0.11773	DHX9/HNRNPA3/NUP58/SLU7/SMG1/SRSF11/YTH
GOTERM_BP	GO:0042698	ovulation cycle	5/141	68/18493	0.00017	0.11773	ADAMTS1/ADNP/PAM/SGPL1/TGFB2
GOTERM_BP	GO:0022602	ovulation cycle process	4/141	47/18493	0.00045	0.11773	ADAMTS1/PAM/SGPL1/TGFB2
GOTERM_BP	GO:0060021	roof of mouth development	5/141	89/18493	0.00060	0.11773	CHD7/FOXE1/MEF2C/SGPL1/TGFB2
GOTERM_BP	GO:0021513	spinal cord dorsal/ventral patterning	3/141	22/18493	0.00060	0.11773	INTU/PAX6/SOX1
GOTERM_CC	GO:0016607	nuclear speck	12/146	382/19659	0.00003	0.00790	BAZ2A/DPP3/FAM76B/MAPT/MEF2C/PNISR/SLU7
GOTERM_CC	GO:0043025	neuronal cell body	12/146	483/19659	0.00027	0.03616	ADNP/CCK/CKB/GRK4/KCND3/MAPT/MPL/PAM/P
GOTERM_CC	GO:0044441	ciliary part	9/146	441/19659	0.00575	0.36035	AHI1/CEP126/CEP131/DYNC2H1/GRK4/INTU/MLF1
GOTERM_CC	GO:0001669	acrosomal vesicle	4/146	106/19659	0.00808	0.36035	CEP131/FLOT2/SPAG8/SV2B
GOTERM_CC	GO:1990351	transporter complex	7/146	333/19659	0.01232	0.36035	ABCD4/GRIK1/GRIN2C/KCND3/PDE4D/PEX13/STX
GOTERM_MF	GO:0003680	AT DNA binding	2/141	10/17632	0.00274	0.26873	MAPT/MEF2C
GOTERM_MF	GO:0042578	phosphoric ester hydrolase activity	9/141	369/17632	0.00290	0.26873	CTDSP1/PDE10A/PDE4D/PDPR/PFKFB2/PLPP2/Pl
GOTERM_MF	GO:0044325	ion channel binding	5/141	125/17632	0.00332	0.26873	KCND3/PDE4D/SLC8A1/STX1A/YWHAZ
GOTERM_MF	GO:0001046	core promoter sequence-specific DNA binding	3/141	38/17632	0.00344	0.26873	BAZ2A/H3F3A/PAX6
GOTERM_MF	GO:0004721	phosphoprotein phosphatase activity	6/141	182/17632	0.00346	0.2687	CTDSP1/PDPR/PLPP2/PPP3CA/PTPDC1/PTPRF

Table 3. Gene ontology (GO) pathway enrichment analyses were performed for the selected genes. The cut-off criterion was $P < 0.05$. The top 5 GO terms in each category in which the 153 hyper-methylated, low-expression genes were significantly involved are shown. They were enriched in the biological process of mRNA transport, etc. The cellular component category revealed enrichment in nuclear speck, neuronal cell body, etc. The molecular function category indicated enrichment in DNA binding and phosphoric ester hydrolase activity.

Comprehensive gene regulation network

The STRING database was used to construct PPI networks. Ultimately, 23 nodes and 2 edges were established from the Hypo-HGs and 151 nodes and 73 edges were established from the Hyper-LGs. The results are shown in Fig. 3 and Fig. 4. We use Cytoscape software to determine the largest subgroup interaction network of Hyper-LGs genes (Fig. 5). We identified hub genes using cytoHubba and validated that DHX9, MAPT, and PAX6 showed the greatest overlap. Two core module analysis of this subgroup network for the Hyper-LGs genes were structured, including module1: HNRNPA3, DHX9, SRSF11, and SLU7 and module2: SOX1, PAX6, and DLX2.

Discussion

AMD is a disease with complex inheritance and epigenetic changes [5]. Identification of the underlying genes has been difficult. Both genomic screen (locational) and candidate gene (functional) approaches have been used. Based on numerous genetic studies of AMD, approximately 50% of the heritability of AMD can be explained by two major loci harboring coding and non-coding variations at chromosomes 1q (CFH) and 10q (ARMS2/HTRA1) [19, 20, 21, 22]. Recently, a large GWAS highlighted new genes and pathways involved in the development of AMD, including complement activation, collagen synthesis, lipid metabolism/cholesterol transport, receptor-mediated endocytosis, endodermal cell differentiation, and extracellular matrix organization, indicating that many unknown genetic changes remain to be identified in the initiation and development of AMD [14]. In this study, we screened for novel biomarkers by combining microarray information RPE-choroid and retina tissue samples from patients with AMD as well as peripheral blood samples by overlapping relevant datasets (GSE29801 and GSE10295) using integrated bioinformatics analysis for available microarray data. This is the first study to employ this approach.

The Hyper-LGs identified are potential biomarkers of AMD methylation microarrays for pre-clinical detection in peripheral blood. Among them, four Hyper-LGs (CKB, PPP3CA, TGFB2, and SOCS2) overlapped with risk genes in the category of "macular degeneration" in PHGKB. One study revealed that CKB was unlikely

to explain the significant fraction of the risk of developing AMD in a family-based association dataset including 162 families and an independent case-control dataset of 399 cases and 159 fully evaluated controls [23]. PPP3CA is a druggable molecule that inactivates MAP3K5, but has not been widely investigated for its role in AMD. One previous study revealed AMD-related sequence variants in genes encoding PPP3CA, underlying its relationship with AMD [24]. TGFB2 induces RPE cell and collagen gel contraction. Subretinal fibrosis contributes to the loss of vision associated with AMD, and RPE cells play a key role in the fibrotic reaction [25]. Under hypoxia conditions, RPE cells can increase the secretion of TGFB2 and induce epithelial-mesenchymal transition, resulting in the formation of scar-like fibrous tissue in AMD [26]. Targeted inhibition of TGFB signaling may be an effective approach for retarding AMD progression [27]. SOCS proteins are modulators of cytokine and growth factor signaling whose aberrant regulation has been linked to a variety of inflammatory and neoplastic diseases [28]. In a GWAS study of 919 patients with exudative AMD treated with intravitreal ranibizumab, SOCS2 was a candidate gene whose levels were associated with visual loss at month three [29]. These results provide insight into AMD pathogenesis, but must be confirmed by in vivo and in vitro experiments. The methylation patterns of PPP3CA, TGFB2, and SOCS2 in AMD have not been previously described. We found that these genes were hypermethylated and low regulated, suggesting that aberrant methylation of these genes affects the pathogenesis of AMD. No Hypo-HGs overlapped in PHGKB, likely because of the limited number of genes identified.

Among the top 5 pathways identified in KEGG and GO, calcium signaling [30, 31], sphingolipid metabolism [32, 33], fibroblast migration [34, 35], membrane [36], coenzyme [37, 38, 39], and DNA binding [40] have been investigated in AMD. Calcium signaling, sphingolipid metabolism, and coenzyme showed strong relationships with AMD, whereas the other categories require further evaluation. Calcium signaling play a vital role in RPE cell function. Intracellular calcium mobilization activates gene expression and the secretion of inflammatory cytokines such as interleukin-8 in human RPE cells [30]. Complement attack on RPE cells leading to cell death was also modulated by extracellular calcium and intracellular signaling mechanisms [31]. Sphingosine 1-phosphate is a potent lipid mediator that modulates inflammatory responses and proangiogenic factors, and it has been suggested that this protein upregulates CNV and is deeply involved in the pathogenesis of exudative AMD [33]. Free radicals play a pathogenic role in AMD, whereas coenzyme Q10 has a protective effect [39]. A combination of acetyl-L-carnitine, n-3 fatty acids, and coenzyme Q10 benefited visual functions in early AMD [38]. However, drug metabolism pathways such as amphetamine addiction and morphine addiction may have been identified by chance and may not be related to AMD. The specific manner in which the other pathways affect AMD development and progression must be further investigated.

In the PPI network, 23 nodes and 2 edges were established from the Hypo-HGs and 151 nodes and 73 edges were established from the Hyper-LGs. DHX9, MAPT, and PAX6 were identified as hub genes. Two core modules for Hyper-LGs were structured, including module1: HNRNPA3, DHX9, SRSF11, and SLU7 and module2: SOX1, PAX6, and DLX2. Among the hub genes and core modules above, PAX6 is expressed in retinal progenitor cells throughout retinogenesis [41]. PAX6 is a novel regulatory gene between RPE transcription factors that controls the timing of RPE differentiation and adjacent choroid maturation, suggesting that PAX6 is involved in choroid development in the pathogenesis of AMD [42]. Other genes have not been previously investigated in AMD.

There were some limitations to this study. First, we focused on Hyper-LGs and Hypo-HGs without analyzing contra-regulated genes; thus, further analysis is required to evaluate these genes. Second, our study was limited to only 2 datasets, and we did not conduct validation in animals or patients' samples. Thus, the results are preliminary and larger sample sizes as well as further fundamental experiments are needed to confirm these results. Third, the clinical characteristics of AMD patients included were not analyzed because these data were not available, so the results should be conservatively interpreted.

Conclusions

In summary, data from gene expression profiling microarrays and gene methylation profiling microarrays of patients with AMD were integrated and analyzed using a series of bioinformatics tools. Our results indicated aberrantly MDEGs (PPP3CA, TGFB2, and SOCS2) and pathways (calcium signaling, sphingolipid metabolism, fibroblast migration, membrane, coenzyme, and DNA binding) in AMD. These genes may serve as biomarkers for the precise diagnosis and treatment of AMD. Further studies are needed to confirm the functional significance of the identified genes and pathways in AMD.

Abbreviations

AMD:age-related macular degeneration; BP:biological process; CC:cellular component; CNV:choroidal neovascularization; DEGs:differentially expressed genes; DMGs:differentially methylated genes; GEO:Gene Expression Omnibus; GO:Gene ontology; GWASs:genome-wide association studies; Hyper-LGs:hyper-methylated low-expression genes; Hypo-HGs:hypo-methylated high-expression genes; KEGG:Kyoto Encyclopedia of Genes and Genomes; MDEGs:methylated-differentially expressed genes; MF:molecular function; NCBI:National Center for Biotechnology Information; PHGKB:Public Health Genomics and Precision Health Knowledge Base; PPI:protein−protein interaction; RPE:retinal pigment epithelium; STRING:Search Tool for the Retrieval of Interacting Genes

Declarations

Ethics approval and consent to participate: All procedures performed in the study were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The written consent was not applicable.

Consent for publication: As the manuscript did not contain any individual details, images or videos, the written consent was not applicable.

Availability of data and materials: The datasets analysed during the current study are available in the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) of the National Center for Biotechnology Information (NCBI). (GSE29801: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29801; GSE102952: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE102952).

Competing interests: The authors declare that they have no competing interests.

Funding: This study was supported by National Key R&D Program of China (2016YFC0904800), National Natural Science Foundation of China (81800831, 81800799) and Shanghai Municipal Commission of Health and Family Planning (201740088). The sponsor or funding organization had no role in the design or conduct of this research.

Authors' contributions: YS conceived and coordinated the project, complied the figures and the tables and wrote the manuscript. ML performed statistical analysis of all data, compiled functional enrichment analysis and protein–protein interaction network analysis. KL, SPZ and XYX helped to coordinate data analysis and screened the hub genes. WG, QZ, PL and NW helped with data curation and searching information in different databases. FY sourced the database and performed statistical analysis of all data. XX conceived and directed the project and wrote the manuscript. All authors reviewed the manuscript.

Acknowledgements: Not applicable.

References

1. Bressler NM. Age-related macular degeneration is the leading cause of blindness… JAMA. 2004;291:1900-1.

2. Bourne RR, Stevens GA, White RA, Smith JL, Flaxman SR, Price H, et al. Causes of vision loss worldwide, 1990-2010: a systematic analysis. Lancet Glob Health. 2013;1: e339-49.

3. Wong WL, Su X, Li X, Cheung CM, Klein R, Cheng CY, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. Lancet Glob Health. 2014;2:e106-16.

4. Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study (AREDS): design implications. AREDS report no. 1. Control Clin Trials. 1999;20:573-600.

5. DeAngelis MM, Owen LA, Morrison MA, Morgan DJ, Li M, Shakoor A, et al. Genetics of age-related macular degeneration (AMD). Hum Mol Genet. 2017;26:R45-50.

6. Jian L, Panpan Y, Wen X. Current choroidal neovascularization treatment. Ophthalmologica. 2013;230:55-61.

7. Do DV. Detection of new-onset choroidal neovascularization. Curr Opin Ophthalmol. 2013;24:244-7.

8. Handa JT. How does the macula protect itself from oxidative stress? Mol Aspects Med. 2012;33:418-35.

9. Klein R, Klein BE, Knudtson MD, Meuer SM, Swift M, Gangnon RE. Fifteen-year cumulative incidence of age-related macular degeneration: the Beaver Dam Eye Study. Ophthalmology. 2007;114:253-62.

10. Evans JR., Fletcher AE, Wormald RP. 28,000 Cases of age related macular degeneration causing visual loss in people aged 75 years and above in the United Kingdom may be attributable to smoking. Br J Ophthalmol. 2005;89:550-3.

11. Hammond CJ, Webster AR, Snieder H, Bird AC, Gilbert CE, Spector TD. Genetic influence on early age-related maculopathy: a twin study. Ophthalmology. 2002;109:730-6.

12. Shahid H, Khan JC, Cipriani V, Sepp T, Matharu BK, Bunce C, et al. Age-related macular degeneration: the importance of family history as a risk factor. Br J Ophthalmol. 2012;96:427-31.

13. Fritsche LG, Chen W, Schu M, Yaspan BL, Yu Y, Thorleifsson G, et al. Seven new loci associated with age-related macular degeneration. Nat Genet. 2013;45:433-9, 439e1-2.

14. Fritsche LG, Igl W, Bailey JN, Grassmann F, Sengupta S, Bragg-Gresham JL, et al. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. Nat Genet. 2016;48:34-43.

15. Hong N, Shen Y, Yu CY, Wang SQ, Tong JP. Association of the polymorphism Y402H in the CFH gene with response to anti-VEGF treatment in age-related macular degeneration: a systematic review and meta-analysis. Acta Ophthalmol. 2016;94:334-45.

16. Kulasingam V, Diamandis EP. Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat Clin Pract Oncol. 2008;5:588-99.

17. Newman AM, Gallo NB, Hancox LS, Miller NJ, Radeke CM, Maloney MA, et al. Systems-level analysis of age-related macular degeneration reveals global biomarkers and phenotype-specific functional networks. Genome Med. 2012;4:16.

18. Radeke MJ, Peterson KE, Johnson LV, Anderson DH. Disease susceptibility of the human macula: differential gene transcription in the retinal pigmented epithelium/choroid. Exp Eye Res. 2007;85:366-80.

19. Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, et al. Complement factor H polymorphism in age-related macular degeneration. Science. 2005;308:385-9.

20. Hageman GS, Anderson DH, Johnson LV, Hancox LS, Taiber AJ, Hardisty LI, et al. A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. Proc Natl Acad Sci U S A. 2005;102:7227-32.

21. Rivera A, Fisher SA, Fritsche LG, Keilhauer CN, Lichtner P, Meitingern T, et al. Hypothetical LOC387715 is a second major susceptibility gene for age-related macular degeneration, contributing independently of complement factor H to disease risk. Hum Mol Genet. 2005;14:3227-36.

22. Deangelis MM, Ji F, Adams S, Morrison MA, Harring AJ, Sweeney MO, et al. Alleles in the HtrA serine peptidase 1 gene alter the risk of neovascular age-related macular degeneration. Ophthalmology. 2008;115:1209-15.e7.

23. Haines JL, Schnetz-Boutaud N, Schmidt S, Scott WK, Agarwal A, Postel EA, et al. Functional candidate genes in age-related macular degeneration: significant association with VEGF, VLDLR, and LRP6. Invest Ophthalmol Vis Sci. 2006;47:329-35.

24. SanGiovanni JP, Lee PH. AMD-associated genes encoding stress-activated MAPK pathway constituents are identified by interval-based enrichment analysis. PLoS One. 2013;8:e71239.

25. Kimura K, Orita T, Liu Y, Yang Y, Tokuda K, Kurakazu T, et al. Attenuation of EMT in RPE cells and subretinal fibrosis by an RAR-γ agonist. J Mol Med (Berl). 2015;93:749-58.

26. Feng Z, Li R, Shi H, Bi W, Hou W, Zhang X. Combined silencing of TGF-β2 and Snail genes inhibit epithelial-mesenchymal transition of retinal pigment epithelial cells under hypoxia. Graefes Arch Clin Exp Ophthalmol. 2015;253:875-84.

27. Radeke MJ, Radeke CM, Shih YH, Hu J, Bok D, Johnson LV, et al. Restoration of mesenchymal retinal pigmented epithelial cells by TGFβ pathway inhibitors: implications for age-related macular degeneration. Genome Med. 2015;7:58.

28. Letellier E, Haan S. SOCS2: physiological and pathological functions. Front Biosci (Elite Ed). 2016;8:189-204.

29. Akiyama M, Takahashi A, Momozawa Y, Arakawa S, Miya F, Tsunoda T, et al. Genome-wide association study suggests four variants influencing outcomes with ranibizumab therapy in exudative age-related macular degeneration. J Hum Genet. 2018;63:1083-91.

30. Yang IH, Wong JH, Chang CM, Chen BK, Tsai YT, Chen WC, et al. Involvement of intracellular calcium mobilization in IL-8 activation in human retinal pigment epithelial cells. Invest Ophthalmol Vis Sci. 2015;56:761-9.

31. Yang P, Baciu P, Kerrigan BC, Etheridge M, Sung E, Toimil BA, et al. Retinal pigment epithelial cell death by the alternative complement cascade: role of membrane regulatory proteins, calcium, PKC, and oxidative stress. Invest Ophthalmol Vis Sci. 2014;55:3012-21.

32. Pujol-Lereis LM, Liebisch G, Schick T, Lin Y, Grassmann F, Uchida K, et al. Evaluation of serum sphingolipids and the influence of genetic risk factors in age-related macular degeneration. PLoS One. 2018;13:e0200739.

33. Terao R, Honjo M, Aihara M. Apolipoprotein M Inhibits Angiogenic and Inflammatory Response by Sphingosine 1-Phosphate on Retinal Pigment Epithelium Cells. Int J Mol Sci. 2017;19:pii: E112.

34. Theodoropoulou S, Copland DA, Liu J, Wu J, Gardner PJ, Ozaki E, et al. Interleukin-33 regulates tissue remodelling and inhibits angiogenesis in the eye. J Pathol. 2017;241:45-56.

35. Santulli RJ, Kinney WA, Ghosh S, Decorte BL, Liu L, Tuman RW, et al. Studies with an orally bioavailable alpha V integrin antagonist in animal models of ocular vasculopathy: retinal neovascularization in mice and retinal vascular permeability in diabetic rats. J Pharmacol Exp Ther. 2008;324:894-901.

36. Kumar-Singh R. The role of complement membrane attack complex in dry and wet AMD - From hypothesis to clinical trials. Exp Eye Res. 2019;184:266-77.

37. McGwin G Jr, Modjarrad K, Hall TA, Xie A, Owsley C. 3-hydroxy-3-methylglutaryl coenzyme a reductase inhibitors and the presence of age-related macular degeneration in the Cardiovascular Health Study. Arch Ophthalmol. 2006;124:33-7.

38. Feher J, Kovacs B, Kovacs I, Schveoller M, Papale A, Balacco Gabrieli C. Improvement of visual functions and fundus alterations in early age-related macular degeneration treated with a combination of acetyl-L-carnitine, n-3 fatty acids, and coenzyme Q10. Ophthalmologica. 2005;219:154-66.

39. Blasi MA, Bovina C, Carella G, Genova ML, Jansen AM, Lenaz G, et al. Does coenzyme Q10 play a role in opposing oxidative stress in patients with age-related macular degeneration? Ophthalmologica. 2001;215:51-4.

40. Paimela T, Ryhänen T, Mannermaa E, Ojala J, Kalesnykas G, Salminen A, et al. The effect of 17beta-estradiol on IL-6 secretion and NF-kappaB DNA-binding activity in human retinal pigment epithelial cells. Immunol Lett. 2007;110:139-44.

41. Remez LA, Onishi A, Menuchin-Lasowski Y, Biran A, Blackshaw S, Wahlin KJ, et al. Pax6 is essential for the generation of late-born retinal neurons and for inhibition of photoreceptor-fate during late stages of retinogenesis. Dev Biol. 2017;432:140-50.

42. Cohen-Tayar Y, Cohen H, Mitiagin Y, Abravanel Z, Levy C, Idelson M, et al. Pax6 regulation of Sox9 in the mouse retinal pigmented epithelium controls its timely differentiation and choroid vasculature development. Development. 2018;145:pii: dev163691.

Additional Files

Additional file 1-Table S1. The top 100 most significantly differentially expressed genes (DEGs) of GSE29801. To identify DEGs, we used the expression profile from GSE29801 (containing RPE-choroid and retina tissue samples from 41 patients with AMD and 42 normal samples). After data preprocessing and quality assessment using R software, we identified 827 high-expression genes and 592 low-expression genes. The top 100 most significantly DEGs of GSE29801 are shown in Table S1.

Additional file 2-Table S2. The top 100 most significantly differentially methylated genes (DMGs) of GSE29801. To identify DMGs, we used the methylation profile from GSE102952 (containing peripheral blood samples of 9 AMD patients and 9 normal controls). After data preprocessing and quality assessment using R software, we identified 4117 hypermethylated genes and 511 hypomethylated genes. The top 100 most significantly DMGs of GSE102952 are shown in Table S2.

Additional file 3-Table S3. Identification of aberrantly methylated-differentially expressed genes (DEGs) in age-related macular degeneration (AMD). To further explore the aberrantly methylated DEGs, hypo-methylated, high-expression genes (Hypo-HGs) were obtained by overlapping hypomethylated and up-regulated genes; hyper-methylated, low-expression genes (Hyper-LGs) were obtained by overlapping hypermethylated and down-regulated genes. We identified 24 Hypo-HGs and 153 Hyper-LGs, and all genes are shown in Table S3.

Additional file 4-The statement provided by the ethics committee of Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China.

Figures
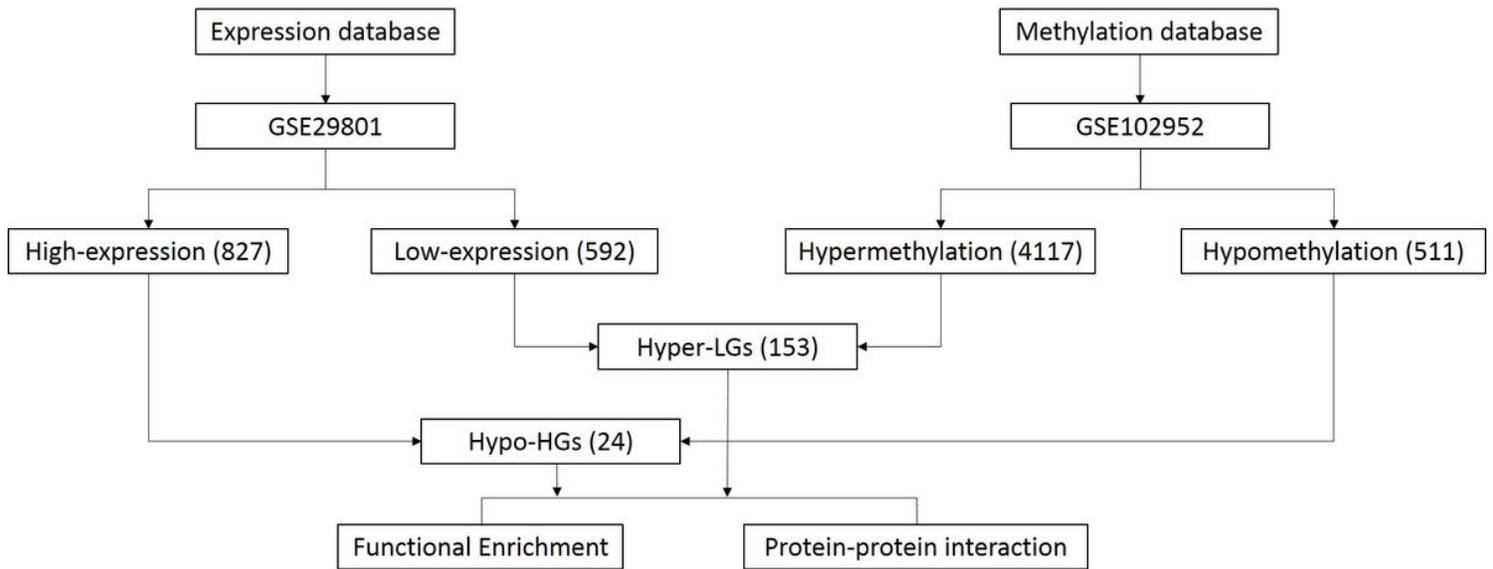
Figure 1

The flowchart of this study. The data of expression profiling GSE29801 and methylation profiling GSE102952 were obtained from Gene Expression Omnibus (GEO) database. As a result, we identified 827 high-expression genes and 592 low expression genes; 4117 hypermethylated genes and 511 hypomethylated genes. To further explore the aberrantly methylated differentially expressed genes, Hypo-HGs were obtained by overlapping hypomethylation and up-regulated genes; Hyper-LGs were obtained by overlapping hypermethylation and down-regulated genes. In total, we identified 24 Hypo-HGs and 153 Hyper-LGs. Then, functional enrichment analysis and protein−protein interaction (PPI) network analysis of screened genes were performed.
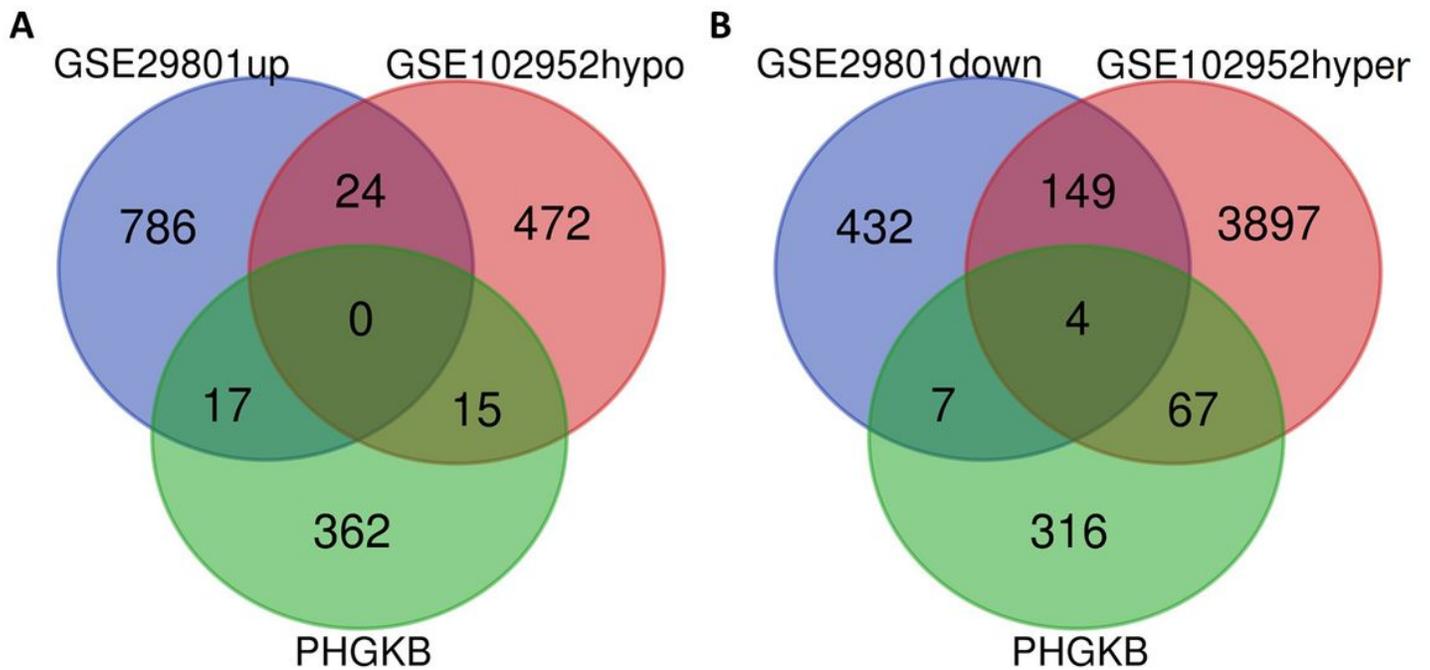
Figure 2

The results of overlapping GSE29801 and GSE102952 datasets with the genes reported to be related to AMD in Public Health Genomics and Precision Health Knowledge Base (PHGKB). (A) Overlapping of hypomethylation-high expression genes (Hypo-HGs) with the genes reported to be related to AMD in PHGKB. (B) Overlapping of hypermethylation-low expression genes (Hyper-LGs) with the genes reported to be related to AMD in PHGKB.
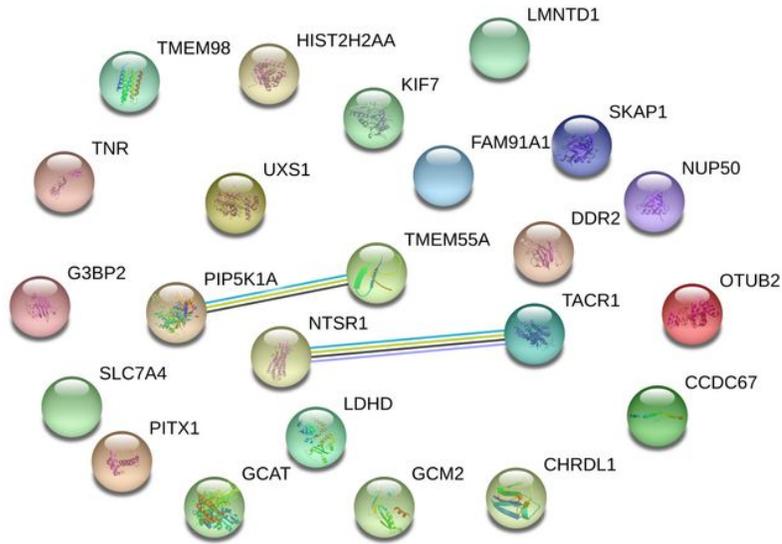
Figure 3

The protein–protein interaction (PPI) networks of the 24 hypomethylation-high expression genes (Hypo-HGs). 23 nodes and 2 edges were established from the Hypo-HGs.
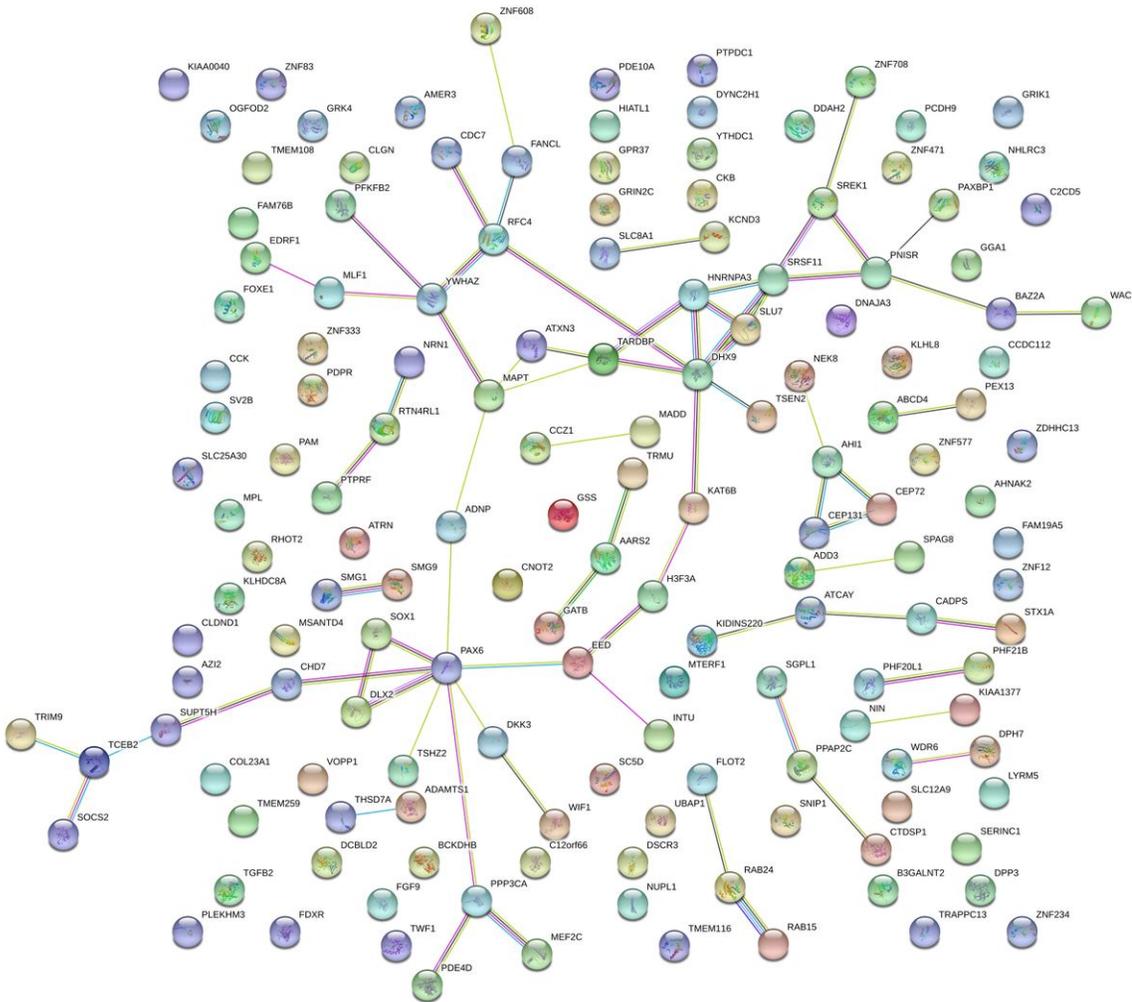
Figure 4

The protein–protein interaction (PPI) networks of the 153 hypermethylation-low expression genes (Hyper-LGs). 151 nodes and 73 edges were established from the Hyper-LGs.

Figure 5

The subgroup protein–protein interaction (PPI) analysis of the hypermethylation-low expression genes (Hyper-LGs).