Original Article

A Computational Approach to Discriminate AMD/ non-AMD Patients Based on Retina Tissue Gene Expression Profile

Fazel Amirvhaedi [1], MS; Mazaher Maghsoudloo [2,3,*], PhD; Farhad Adhami-Moghadam [4], MD

1. Department of Biotechnology, University of Tehran, Tehran, Iran.
2. The Center of Research and Training for Occupational Technical Safety and Health, Tehran, Iran.
3. Department of Genetics, Faculty of Advanced Science and Technology, Tehran Medical Sciences, Islamic Azad University, Tehran, Iran.
4. Department of ophthalmology, Faculty of Medicine, Tehran Medical Sciences Islamic Azad University, Tehran, Iran.

***Corresponding Author:** Mazaher Maghsoudloo

E-mail: m.maghsoudloo@ut.ac.ir

Abstract

Background: Age-related macular degeneration (AMD) is the progressive degenerative disease of the macula and the main cause of blindness in older adults. Various risk factors have been associated with disease progression among different individuals. As a multifactorial disease, AMD is affected by different risk factors such as aging, genetic susceptibility, environmental risk factors and lifestyle. Since the etiology of AMD is not fully known, it would be essential to identify disease risk factors and novel predictive risk factors to detect AMD at an early stage.

Material and Methods: The expression data were obtained from the Gene Expression Omnibus database. Samples were quantile normalized, and log2 transformed. Furthermore, outlier samples were removed by hierarchical clustering. R limma was used to run a linear model and identify differentially expressed genes (DEGs). As a result, 33 genes were discovered with a q-value less than 0.05 and a $|\log (FC)| \geq 0.7$. With a machine learning (ML) approach, DEGs were applied to discriminate between the case and control samples. Furthermore, FeatureSelect is used to extract the most effective separator genes. Nine genes were identified as the best disease discriminator genes through 11 feature selection algorithms.

Results: The gene set found in the study distinguishes healthy samples from patient samples with an accuracy of 87.5 %. We found DEF119B, UBD, and GRP to be three novel potential AMD candidate biomarkers using ML models and feature selection.

Conclusion: Machine learning can be beneficial in diagnosing, preventing and treating diseases, especially in diseases such as AMD that do not have a clear etiology.

Keywords: Gene Expression; Machine Learning; AMD; Gene Selection; Feature Selection.

Article Notes: Received: Aug. 29, 2020; Received in revised form: Sep 22, 2020; Accepted: Nov. 11, 2020; Available Online: Apr. 4, 2021.

How to cite this article: Amirvhaedi F, Maghsoudloo M, Adhami-Moghadam F. A computational approach to discriminate AMD/non-AMD patients based on retina tissue gene expression profile. Journal of Ophthalmic and Optometric Sciences . 2021;5(2): 6-20.

Introduction

AMD is the leading cause of blindness in people over 55 in industrialized countries, accounting for 6 to 9 % of the world's blindness [1]. By 2020, AMD is expected to affect 196 million people, and by 2040, it will affect 288 million [2]. Drusen are deposits of insoluble extracellular debris found between the retinal pigment epithelium (RPE) and Bruch's membrane (BM) [3, 4], causing progressive degeneration of photoreceptors and RPE, ultimately resulting in loss of central vision. The macula, which is the central part of the retina and has the highest density of photoreceptor cells, is significantly altered pathologically by AMD [5].

There are two types of AMD, wet and dry, based on whether blood vessels have disrupted the retina or not. In dry cases, there is often some degree of vision impairment and sometimes severe blindness. Approximately 15-30 percent of AMD patients have the wet form, which presents abruptly and rapidly progresses to blindness if untreated [4]. The late stage of dry AMD, known as geographic atrophy (GA), is the most damaging endpoint of the drusen(lipid) life cycle [5]. At the end stages of wet AMD, also known as choroidal neovascularization (CNV), immature blood vessels emerge from the underlying choroid and grow toward the outer retina [4]. GA currently does not have a treatment; however, complement cascades are expected to offer a potential therapeutic option [2], while the exudative late wet form of the disease is treatable via intravitreal injections of VEGF inhibitors, although it is a labour-intensive process [6].

AMD is known as a multifactorial disease in which many factors such as age, genetics, oxidative stress, lipid metabolism, and environmental and lifestyle risk factors such as diet and smoking are involved, and due to this, the underlying disease mechanism has not yet been established been identified clearly [5,7]. Due to all of this, early detection of AMD, identification of genes involved in it, and AMD pathogenesis pathways have become increasingly important. Recently, transcriptome analysis[8-11] was performed on RPE-choroid and retina samples to identify AMD global biomarkers [12]. In addition, differentially expressed genes have been identified by the co-expression network analysis approach in some studies [13]. Also, another study has analyzed the aqueous humor proteome in patients with AMD [14]. These analyses were carried out to identify potential therapeutic drugs in some cases [15].

Machine learning (ML) is a set of computational methods and models that can be used for various purposes, such as classifying or predicting the results of experimental data (transcriptomics or proteomics data) and selecting the best features[16-18]. These features are used to build a model that is more accurate than when all the features are used for prediction or classification [19].

AMD can be precisely diagnosed and staged with fundus imaging or optical coherence tomography (OCT) images of the retina. So far, machine learning classifiers are mainly used to detect and diagnose AMD through these images [20]. Furthermore, it has the ability to predict the longitudinal progression status of disease [21].

In this study, we aim to find candidate biomarkers for AMD by studying and analyzing transcriptome data using machine learning. Firstly, we analyzed gene expression, identified DEGs, and then used a machine learning approach to select the best disease-predicting genes. Chosen genes were evaluated and screened from the functional enrichment perspective. Finally, the best subset of genes

was selected that could accurately predict the status of the case.

Material and Methods

Gene Expression Profile and Study Population Demographic Information

The expression data were retrieved from GEO (Gene Expression Omnibus, http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE29801 [22]. The data were composed of 293 samples; 177 belonged to the human retinal pigmented epithelium (RPE)-choroid complex, and 118 belonged to retina tissue. Only two samples of choroidal tissue were replicated among the whole samples. The Series Matrix File was downloaded and used for subsequent analysis. The experiment type was Expression profiling by array, and the Agilent-014850 Whole Human Genome Microarray 4x44K G4112F was used for transcriptome analysis. The platform accession number is GPL4133. We performed our analyzes focusing on retinal tissue samples.

The data were obtained from a common cohort of 31 normal, 26 AMD, and 11 potential pre-AMD human eyes derived from the University of Iowa. Two retina specialists evaluated all Iowa donors and grouped them according to morphological and pathological features. From human donor eyes obtained from the University of Iowa, RPE-choroid and retinal samples were isolated. AMD cases are divided into five different subtypes based on their phenotypic features like drusen size: MD1 (Pre-AMD), MD2 (Sub-clinical pre-AMD), Dry AMD (non-GA Dry AMD), GA (Geographic atrophy), and CNV (Wet AMD) [22 23, 24].

Data Pre-Processing and Differential Expression Analysis

Data was read in R and then quantile normalized and $\log_2$ adjusted. Then, a box plot and density plot were drawn to visualize the distribution of modified data. Moreover, the hierarchical clustering method was used to determine noisy and outlier data and remove them. Finally, the probe IDs whit unknown gene names were excluded.

After data preprocessing, all AMD/pre-AMD phenotypes were combined and tested for significant differential expression against normal donor samples. A linear model was used to identify differentially expressed genes using the R package Limma 3.5.1 [25] as used in other gene expression analysis studies [23, 26, 27]. The P values adjusted by the False Discovery Rate method and those with FDR < 0.05 were considered statistically significant to identify significant DEG between patients and normal subjects.

Feature selection and Gene correlation

In order to select the best genes for predicting the state of a sample, a smaller number was chosen from the identified DEGs found using FeatureSelect software. FeatureSelect is a feature or gene selection software that offers the best features using 11 different algorithms based on wrapper methods [2829]. To reduce the possibility of overfitting, we conducted 10-fold cross-validation to our feature selection and model construction . This process was repeated for all 11 algorithms; moreover, the Max Feature was equivalent to 10, and the parameters of each algorithm were set to default.

Functional enrichment

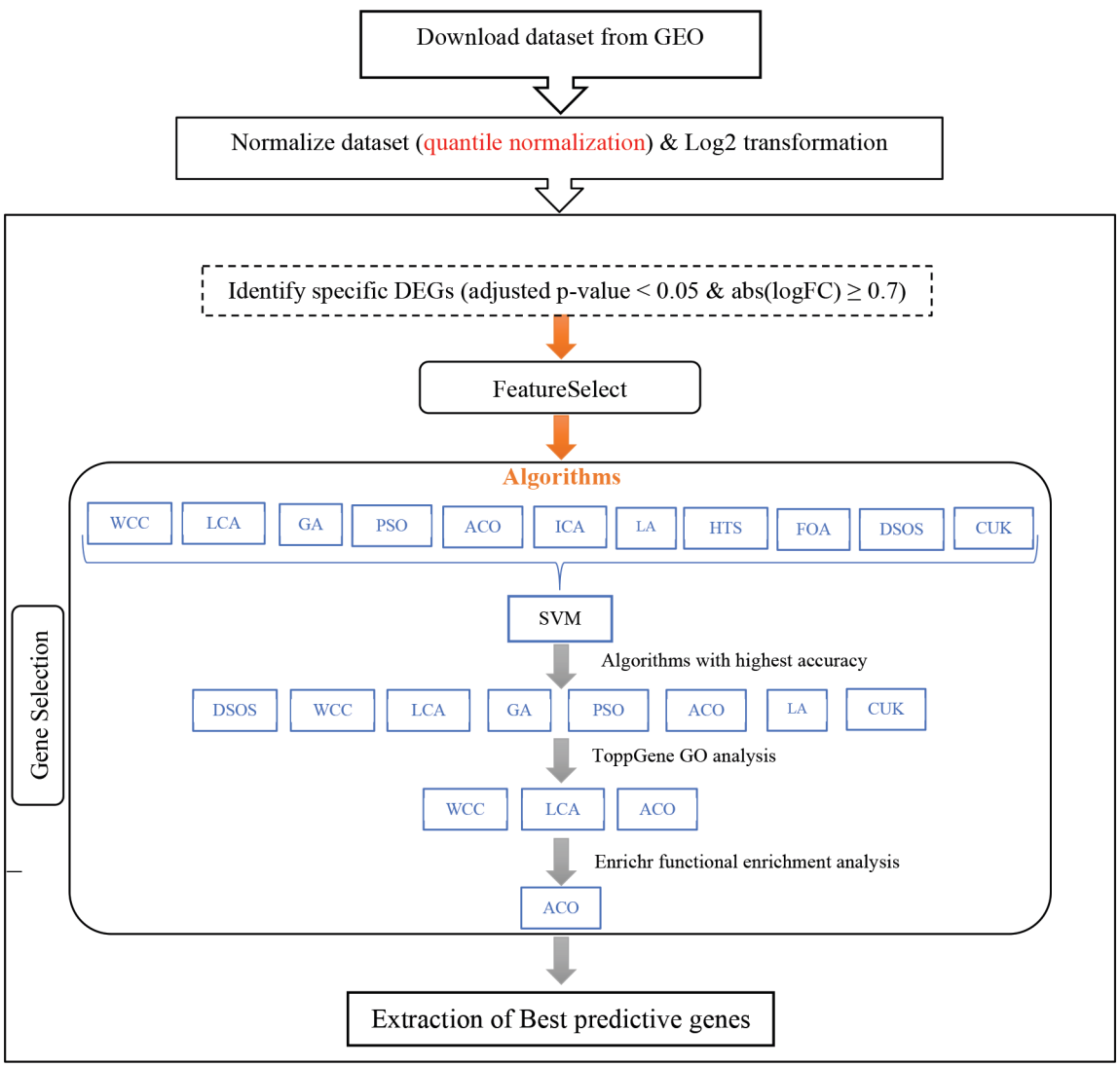
To select the best gene subset from the 11 proposed subsets, eight groups with the highest accuracy were selected. In the next stage, we used a gene ontology approach based on three categories, including Biological Process,

Pathway, and Disease, to select the best subset among the eight remaining groups.

Gene ontology and pathway terms are obtained using the ToppGene web tool (http://toppgene.cchmc.org/) [30]. We identified significant terms using the adjusted P value, which was less than 0.05. As a result of enrichment, we reached to top three gene subsets. In the next step, we used the Enrichr web tool (https://maayanlab.cloud/Enrichr/) [31] and KEGG 2021 Human module to select the most specific subsets. Finally, a gene subset was selected as the best feature class. Significant terms were identified using the adjusted P value, which was less than 0.05. The gene subset with the most similar biological pathways to AMD was selected. Fig1 shows the workflow of the proposed method discussed in this and subsequent sections.

Gene-disease association analysis

To investigate the gene-disease relationship, the DisGeNET web tool was used (https://www.disgenet.org/home/) [32]. DisGeNET contains one of the largest collections of genes and variants associated with human diseases, which can be used for several research purposes, including investigating the molecular basis of diseases and comorbidities, analyzing the properties of disease genes, and evaluating computationally predicted disease genes [33].



Results

Data preprocessing and DEGs Identification

Using the clustering method to analyze data and dendrograms to draw conclusions, three samples, consisting of two normal data and one patient data, were identified as outliers. They were removed from all future samples and analyses. Detailed information on retina samples has been shown in table 1. We excluded probes without annotated genes or NaN values from the experimental study. The total number of probes falls from 45220 to 19749. We identified differentially expressed genes using the general linear model. Genes with an adjusted P value < 0.05 were filtered and assigned as DEGs. As a result, 403 DEGs were identified (Additional File 1). Then, in order to further reduce the number of genes and reach more significant ones, the genes with ($|logFC|$) ≥ 0.7 were selected, and as a result, 33 DEGs were identified (Additional File 2) between normal and AMD (including pre-AMD). figure 1 shows identified DEGs. Three samples were removed due to using a hierarchical clustering method to determine noisy and outlier data (figure S1 in Supplementary file).

Feature selection using different wrapper methods

First, all FeatureSelect algorithms were applied to the data to select the genes/features with the utmost accuracy from the 33 DEGs for classification. According to the results, eight algorithms with high accuracy, including the WCC, LCA, GA, PSO, ACO, LA, DSOS, and CUK, were selected to choose the best gene set for classification and potential biomarkers. Furthermore, the gene correlation matrix of all 33 genes was calculated using the Pearson correlation coefficient (figure S2 in Supplementary file) to understand the relationship between genes better. table 3 shows the information for each feature selection algorithm. The highest accuracy was obtained using the DSOS algorithm (95.83 %). figure 1 shows the performance of each algorithm per iteration and the stability

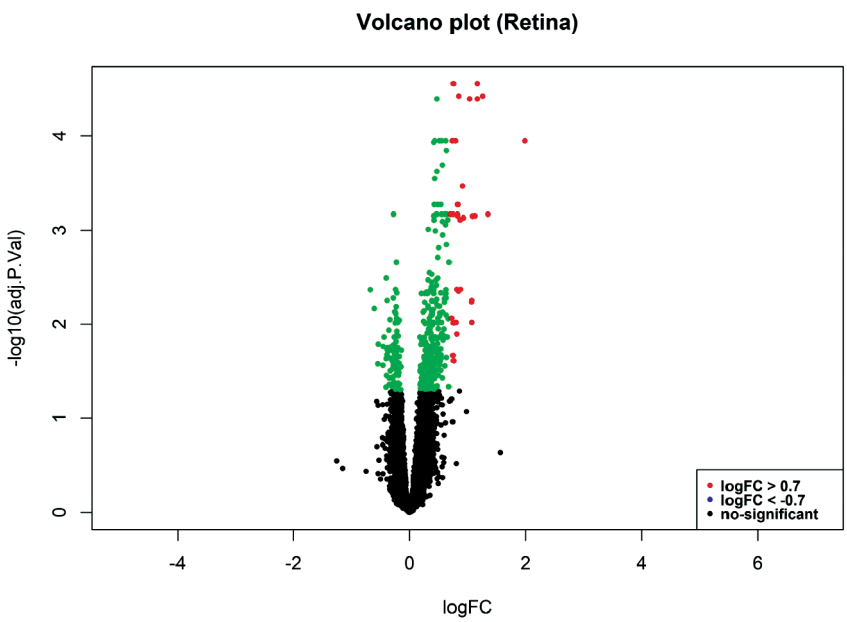


Figure 1: The volcano plot demonstrates identified DEGs. Red dots show DEGs with logFC > 0.7 and adjusted P value less than 0.05

Table 1: Demographic features of retina samples in the study

AMD classification	Age	Sex (male/female)	Retina samples
Normal	61.65 (± 25.04)	30 (54.55 %) 25 (45.45 %)	55
MD1	78.77 (± 11.60)	5 (55.55 %) 4 (44.45 %)	9
MD2	87 (± 7.35)	3 (60 %) 2 (40 %)	5
Dry AMD	79.1 (± 13.24)	8 (25.81 %) 23 (74.19 %)	31
GA	82 (± 9.24)	2 (50 %) 2 (50 %)	4
CNV	82 (± 3.93)	6 (75 %) 2 (25 %)	8
GA/CNV	85.67 (± 5.82)	2 (33.33 %) 4 (66.67 %)	6

of accuracy and error. In addition, receiver operating characteristic (ROC) curve was plotted to better compare algorithms and their discriminative power, figure 2.

The most accurate algorithms and features proposed

Algorithms with the highest accuracy were selected for feature selection. Finally, 7 to 10 features were proposed by different algorithms. These algorithms and the features/genes which have been proposed are demonstrated in table 3.

In the following, all of the eight algorithms and selected features were enriched for AMD based on three gene ontology combinations, including Biological Process, Pathway, and Disease in Toppgene web tool (Additional File 3). For DSOS, no common Biological Process was found for the selected genes and for several other algorithms, there were not enough common Pathways with AMD or not enough significantly relevant diseases found. Eventually, three algorithms with the highest similarity to AMD in terms of three gene ontologies were selected: WCC, LCA, and ACO.

Gene enrichment analysis

In the next step to select the best algorithm,

Table 2: Detailed information for each algorithm

Algorithm name	Accuracy 10-fold cross-validation										
	WCC	LCA	GA	PSO	ACO	ICA	LA	HTS	FOA	DSOS	CUK
NOF[a]	10	7	9	9	9	9	8	6	8	7	9
Accuracy	87.5	87.5	87.5	87.5	87.5	83.33	87.5	83.33	83.33	95.83	87.5
Precision	87.78	87.06	87.78	87.06	87.06	82.86	87.78	57.14	83.33	96.67	87.78
Sensitivity	86.43	87.86	86.43	87.86	87.86	82.86	86.43	57.14	84.29	95	86.43

a. Number of features

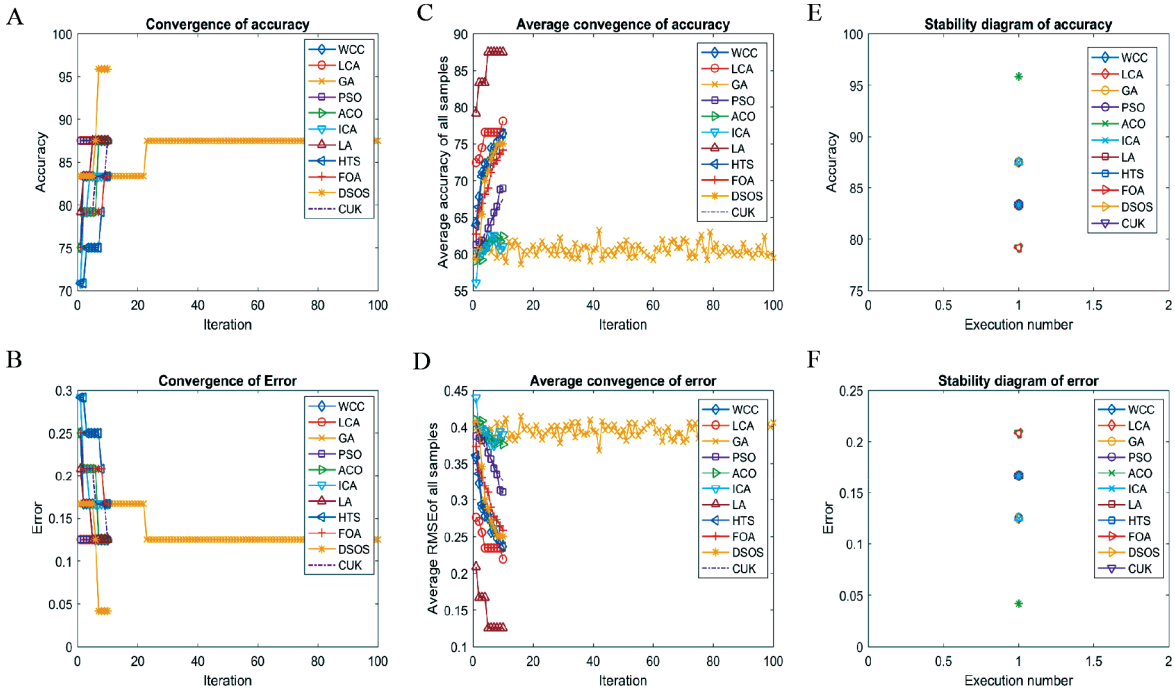


Figure 2: (A) demonstrates the Convergence of accuracy per iteration for each algorithm. As can be seen, the DSOS algorithm has converged rapidly and achieved maximum accuracy (96.675 %). GA converged at a slower rate than other algorithms. (B) shows the Convergence of Error per iteration for each algorithm. DSOS has reached the minimum Error of slightly less than 0.05. (C) & (D) demonstrate the Average convergence of accuracy and error, respectively. (E) & (F) show the accuracy and error Stability diagram, respectively. In terms of accuracy stability, ACO has reached the best accuracy, slightly over 95 %; also, it has the minimum error value in the stability diagram, below 0.05

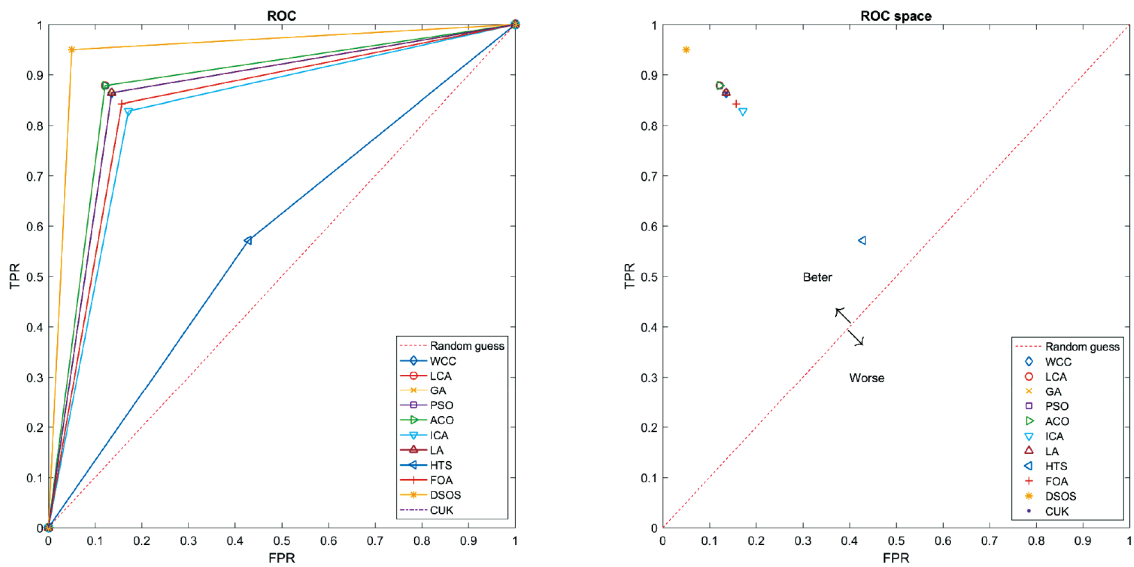


Figure 3: ROC curve for each model shows how much the model can distinguish between classes, in this case, healthy and disease people. DSOS has the highest area under the curve by about 97 %, and the lowest belongs to HTS

Table 3: Highest accuracy algorithm and their selected features

Algorithm name	Accuracy	NOF[a]	Features/genes
DSOS	95.8	7	C4B, SCRG1, MOXD1, NPVF, ANO3, DEFB119, ICAM1
WCC	87.5	10	C4B, SCRG1, ANO3, SERPINA3, ANXA1, CFB, GSTT2, SLC47A2, LGALS3, CD163
LCA	87.5	7	SCRG1, NPVF, C1S, SQRDL, CFB, ICAM1, LGALS3
GA	87.5	9	SCRG1, NPVF, C1S, PTGS1, SQRDL, FZD7, ANXA1, ICAM1, LGALS3
PSO	87.5	9	SCRG1, C1S, ABCC3, FCGBP, IGFBP2, SQRDL, SERPINA3, ICAM1, LOC100128008
ACO	87.5	9	C4B, HLA-DRA, ANO3, FAM181A, GRP, CFB, DEFB119, UBD, SLC47A2
LA	87.5	8	CP, ABCC3, ANO3, FAM181A, ANXA1, SLC47A2, LOC100128008, LGALS3
CUK	87.5	9	C4B, HLA-DRA, C1S, PTGS1, ABCC3, FAM181A, SQRDL, GRP, SLC47A2

a. Number of features

the features brought to the Enrichr web tool and examined from the functional enrichment perspective; the KEGG 2021 Human module was used for this mean, figure 4. Significant terms with an adjusted P value less than 0.05 were used for enrichment analysis. ACO chosen gene subset with the highest pathway similarity to AMD was selected for further analysis. Using the ACO-SVM algorithm, the following genes were identified for AMD: C4B, HLA-DRA, ANO3, FAM181A, GRP, CFB, DEFB119, UBD, and SLC47A2.

Gene-disease association investigation

We compared the genes found in AMD with those in the Summary of Gene-Disease Associations database in DisGeNET. It was found that six genes or their close family members were reported in the database and three genes were not. These three genes were: UBD, DEFB119, and GRP.

Discussion

By using machine learning, candidate biomarkers of AMD can be obtained more accurately and quickly. Each gene as a feature/ candidate biomarkers can predict disease and its progression; furthermore, the features selection can improve the disease classification accuracy. In this study, we proposed an efficient method of combining gene expression analysis with a machine learning approach and high-performance feature selection algorithms to gain the best disease predicting candidate biomarkers [34, 35] . The study probably contributed to the discovery of more reliable genes involved in the pathogenesis of diseases, as well as some new candidate biomarkers .

It may seem like a large number of features or genes in microarray data is advantageous, but many of these are redundant or irrelevant, resulting in the loss of classification accuracy

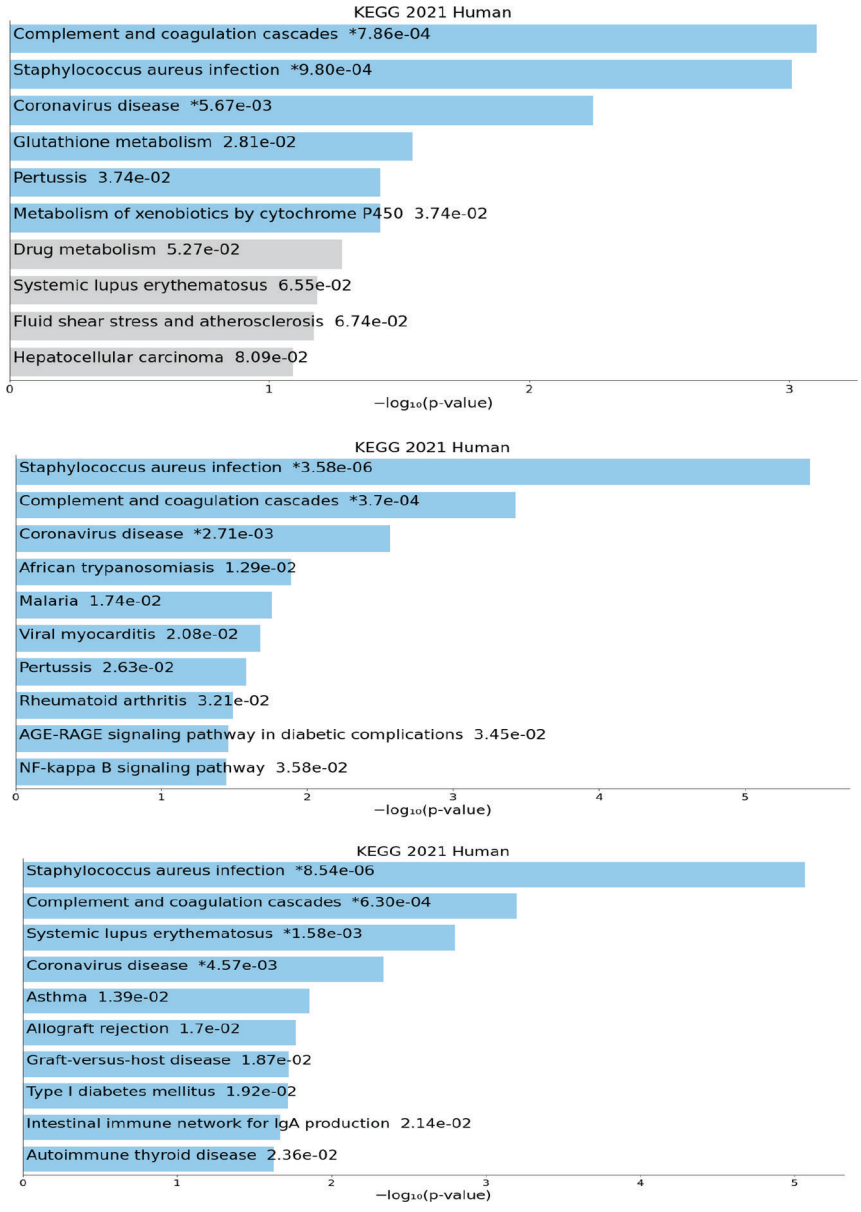


Figure 4: A pathway analysis was conducted using the KEGG 2021 Human module in Enrichr for three gene subsets. The bar chart shows the P values associated with the top 10 enriched terms in the chosen library. The colored bars correspond to terms with significant P values (< 0.05). (a) shows the bar chart that belongs to WCC, (b) to LCA, and (c) to ACO. The asterisk (*) beside a P value indicates it also has a significant adjusted P value (< 0.05)

and causing overfitting. A microarray analyzes the expression levels of thousands of genes simultaneously and can be used as a detective tool. Nevertheless, the accuracy is impacted by a large number of features and the limited number of samples. To overcome this deterioration inaccuracy, dimensionality reduction techniques, mainly feature selection approaches, are employed [36]. Feature selection techniques maximize prediction accuracy, minimize training time, and counteract the threat of model overfitting by handling the dimension and better interpreting the feature space [37].

Many studies related to AMD prognosis and diagnosis have utilized machine learning methods. For example, in an investigation, AMD GWAS studies combined with machine learning to predict AMD risk through four different approaches using genome-wide significant SNPs as features [38]. In another study, the researchers combined fundus images and feature SNPs in a deep learning-based approach for predicting late AMD progression using convolutional neural networks and reached an average area under curve (AUC) value of 0.85 [39]. Genetic variants have been identified at 34 loci associated with AMD by genome-wide association studies (GWAS[40, 41]). In a different investigation, using random walk (RW) and support vector machines (SVM), researchers identified AMD-related genes based on gene interaction relationships and gene significance to understand the disease pathogenesis better. Their model's area under the curve (AUC) was 0.927, resulting in 65 new genes associated with AMD [42]. First-line therapy for CNV in AMD patients consists of intravitreal injections of anti-vascular endothelial growth factor (VEGF) agents. The anti-VEGF treatment, however, does not always produce satisfactory results in all patients. Transcriptome analysis of peripheral blood mononuclear cells (PBMCs) from AMD patients before treatment to identify biomarkers of response to ranibizumab with the help of machine learning and correlation-based feature selection algorithms showed high accuracy with the area under curve of the value 0.968 [43].

Our proposed method identified C4B, HLA-DRA, ANO3, FAM181A, GRP, CFB, DEFB119, UBD, and SLC47A2 as potential candidate biomarkers for AMD. Gene-disease association investigation showed DEFB119, UBD and GRP as novel candidate biomarkers, and the remaining six genes were either directly related to the disease or belonged to a very close family related to AMD.

Ubiquitin D or FAT10 is a ubiquitin-like modifier (ULM) encoded in the major histocompatibility complex (MHC). It is induced in various tissues upon stimulation by inflammatory factors like tumour necrosis factor (TNF) and interferon γ (IFN- γ) in mammals. Through docking to the 26S proteasome, FAT10 cause the bulk of its conjugates to degrade with an overall half-life of one hour. It is tightly regulated by transcriptional as well as post-transcriptional mechanisms. Thus, FAT10 is mainly expressed in organs of the immune system, such as the lymph nodes and spleen. The FAT10 protein is implicated in various cellular functions, including apoptosis, spindle checkpoints during mitotic cell division, and NF-κB activation. Additionally, it interacts with a broad spectrum of proteins involved in tumorigenesis [44].

AIPL1, a retinal chaperone, is essential for maintaining retinal phototransduction in photoreceptors, and losing AIPL1 causes rapid rod and cone cell degeneration. It plays a vital role in phosphodiesterase 6 (PDE6) assembly, an essential effector enzyme in phototransduction which hydrolyzes cGMP. FAT mRNA is expressed in the human retina, and rod PDE6 is identified as a retina-specific substrate of FAT10 conjugation. AIPL1 stabilizes the PDE6-FAT10 conjugate. AIPL1 interacts with FAT10 in a non-covalent manner. PDE6 FAT10ylatoin leads to its proteasomal degradation and consequently inhibits PDE6 cGMP hydrolyzing activity. Thus, FAT10 may contribute to the loss of PDE6 and, as a result, the degeneration of retinal cells in diseases associated with inflammation and inherited blindness caused by mutations in AIPL1 [45].

In conjunction with aging, oxidative damage is accumulated, and oxidative damage is considered the primary cause of age-related degenerative diseases. In addition, oxidative stress has been shown to trigger inflammation during AMD pathology. Proteins, lipids and DNA are damaged in pathological oxidative damage, and mitochondria are dysfunctional, resulting in pro-inflammatory responses, macrophage infiltration, and polarization. It is believed that inflammation caused by tissue damage is necessary for the immune system to create a protective response. Drusogenesis, degeneration of the RPE/photoreceptor, BM disruption, and development of CNV are all caused by local inflammation. Thus inflammation is thought to play an integral role in the pathogenesis of both dry and wet AMD [2].

Defensins are cationic antimicrobial peptides that are part of a large family of antimicrobial peptides that share the same signature pattern that consists of six conserved cysteines. Defensins have a wide range of functions and have been implicated in a variety of human diseases. Natural antibacterial peptides, such as defensin, have emerged as a crucial form of innate immunity in animals and plants. Human β-defensin (DEFB) is produced by epithelial cells lining tissues exposed to the outside world. Antibacterial activity has been detected in vitro, suggesting that they may be involved in innate immunity. Aside from their role in innate immunity, human β-defensins also act as chemo-attractants for immature dendritic cells and memory T cells, which link the innate and adaptive immune systems. β-defensins also have essential roles in the male reproductive tract and perform reproduction-specific tasks. There is also evidence that defensins may contribute to HIV- 1 resistance in male and female genitourinary tract [46]. In a study on the effects of ageing and anatomic location on gene expression in the human retina, DEFB119 were expressed at higher levels in young macula compared to older macula [47]. Another study carried out transcriptomic analysis across different regions of the human retina and RPE/choroid by RNA-seq, DEFB119, was identified as one of the DEGs between temporal RPE/choroid vs temporal retina [48].

Gastrin-releasing peptide (GRP) is a neuropeptide comprised of 27 amino acids. It binds to the GRP receptor (GRPR, also known as BB2) on cell membranes. The GRP/GRPR system appears to influence a wide range of physiological processes: including exocrine and endocrine secretions, thermal regulation, blood sugar levels, feeding, inflammatory responses, gastrin and somatostatin secretion, gastric acid secretion, pancreatic secretion, gastrointestinal motility, lung development, pain perception, satiety, itchy responses, memory formation and expression, and cell proliferation in the immune system. Also, brain function may be controlled by GRP, a transmitter in the central nervous system (CNS) that activates GRPRs on postsynaptic membranes [49-51].

GRP acts as a mitogen, morphogen and pro-angiogenic factor in specific cancers. Normal kidney and renal cancer express GRP and its receptor. It is known that GRP may promote the growth of renal cancer cells. Studies have shown that chronic kidney disease (CKD) is associated with an increased risk of AMD [52]. Furthermore, it has been discovered that in male sexual function, GRPs and GRPRs play an essential role in spinal cord control of erection and ejaculations [50]. The evidence from animal and human studies indicates that abnormalities in GRPR-triggered cellular signalling are linked to Alzheimer's disease (AD). The results of mouse and human studies

suggest that abnormalities of the GRPR gene expression and GRPR signalling (cellular calcium homeostasis and oxidative stress) are likely associated with AD.

There are several clinical and pathologic similarities between AD and AMD. First and foremost, aging is a major risk factor for both multifactorial degenerative disorders. Secondly, the pathological hallmark of AD, the deposition of extracellular amyloid β-peptide, is also a component of AMD drusen [3]. Moreover, preclinical research has shown that GRP and its receptor are important for modulating synaptic plasticity and memory formation in the hippocampus and other areas of the brain [53]. Although the GRP gene was not explicitly mentioned in the DisGeNET database, however GAST (gastrin coding) gene was reported as one of the disease-associated genes with a low association score.

Conclusion

The results found in this paper show the power of ML and feature selection algorithms in better understanding multifactorial diseases like AMD. The use of feature selection algorithms, in addition to increasing the accuracy of the model, also prevents it from being overfit. Enrichment analysis approach was used to assess the biological validity of the found genes. The mentioned DEGs might be potential diagnostic or therapeutic candidate biomarkers. Therefore, further experimental evaluation of these DEGs in AMD is recommended.

Authors ORCIDs

Fazel Amirvhaedi:
https://orcid.org/0000-0002-1056-5983
Mazaher Maghsoudloo:
https://orcid.org/0000-0002-3218-1059

References

1. Fleckenstein M, Keenan TDL, Guymer RH, Chakravarthy U, Schmitz-Valckenberg S, Klaver CC, et al. Age-related macular degeneration. Nat Rev Dis Primers. 2021;7(1):31.

2. Tan W, Zou J, Yoshida S, Jiang B, Zhou Y. The Role of Inflammation in Age-Related Macular Degeneration. Int J Biol Sci. 2020;16(15):2989-3001.

3. Heesterbeek TJ, Lorés-Motta L, Hoyng CB, Lechanteur YTE, den Hollander AI. Risk factors for progression of age-related macular degeneration. Ophthalmic Physiol Opt. 2020;40(2):140-70.

4. Ambati J, Fowler BJ. Mechanisms of age-related macular degeneration. Neuron. 2012;75(1):26-39.

5. Somasundaran S, Constable IJ, Mellough CB, Carvalho LS. Retinal pigment epithelium and age-related macular degeneration: A review of major disease mechanisms. Clin Exp Ophthalmol. 2020;48(8):1043-56.

6. Stahl A. The Diagnosis and Treatment of Age-Related Macular Degeneration. Dtsch Arztebl Int. 2020;117(29-30):513-20.

7. Tsuchida M, Fukushima T, Nasuda S, Masoudi-Nejad A, Ishikawa G, Nakamura T, et al. Dissection of rye chromosome 1R in common wheat. Genes & Genetic Systems. 2008;83(1):43-53.

8. Kouhsar M, Azimzadeh Jamalkandi S,

Moeini A, Masoudi-Nejad A. Detection of novel biomarkers for early detection of Non-Muscle-Invasive Bladder Cancer using Competing Endogenous RNA network analysis. Scientific Reports. 2019;9(1):8434.
9. Najafi A, Bidkhori G, Bozorgmehr JH, Koch I, Masoudi-Nejad A. Genome scale modeling in systems biology: algorithms and resources. Curr Genomics. 2014;15(2):130-59.
10. Motieghader H, Kouhsar M, Najafi A, Sadeghi B, Masoudi-Nejad A. mRNA-miRNA bipartite network reconstruction to predict prognostic module biomarkers in colorectal cancer stage differentiation. Mol Biosyst. 2017;13(10):2168-80.
11. Alaei S, Sadeghi B, Najafi A, Masoudi-Nejad A. LncRNA and mRNA integration network reconstruction reveals novel key regulators in esophageal squamous-cell carcinoma. Genomics. 2019;111(1):76-89.
12. 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(2):16.
13. Liang G, Ma W, Luo Y, Yin J, Hao L, Zhong J. Identification of differentially expressed and methylated genes and construction of a co-expression network in age-related macular degeneration. Ann Transl Med. 2022;10(4):223.
14. Rinsky B, Beykin G, Grunin M, Amer R, Khateb S, Tiosano L, et al. Analysis of the Aqueous Humor Proteome in Patients With Age-Related Macular Degeneration. Invest Ophthalmol Vis Sci. 2021;62(10):18.
15. Abedi Z, MotieGhader H, Maghsoudloo M, Sheikh Beig Goharrizi MA, Shojaei A. Novel Potential Drugs for Therapy of Age-Related Macular Degeneration Using Protein-Protein Interaction Network (PPI) Analysis. Journal of Ophthalmic and Optometric Sciences. 2019;3(4):11-23.
16. Hooshmand SA, Zarei Ghobadi M, Hooshmand SE, Azimzadeh Jamalkandi S, Alavi SM, Masoudi-Nejad A. A multimodal deep learning-based drug repurposing approach for treatment of COVID-19. Mol Divers. 2021;25(3):1717-30.
17. Masoudi-Sobhanzadeh Y, Omidi Y, Amanlou M, Masoudi-Nejad A. DrugR+: A comprehensive relational database for drug repurposing, combination therapy, and replacement therapy. Comput Biol Med. 2019;109:254-62.
18. Abbasi K, Razzaghi P, Poso A, Ghanbari-Ara S, Masoudi-Nejad A. Deep Learning in Drug Target Interaction Prediction: Current and Future Perspectives. Curr Med Chem. 2021;28(11):2100-13.
19. Maghsoudloo M, Azimzadeh Jamalkandi S, Najafi A, Masoudi-Nejad A. An efficient hybrid feature selection method to identify potential biomarkers in common chronic lung inflammatory diseases. Genomics. 2020;112(5):3284-93.
20. Cheung R, Chun J, Sheidow T, Motolko M, Malvankar-Mehta MS. Diagnostic accuracy of current machine learning classifiers for age-related macular degeneration: a systematic review and meta-analysis. Eye. 2022;36(5):994-1004.
21. Ganjdanesh A, Zhang J, Chew EY, Ding Y, Huang H, Chen W. LONGL-Net: temporal correlation structure guided deep learning model to predict longitudinal age-related macular degeneration severity. PNAS Nexus. 2022;1(1):pgab003.
22. 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(2):16.
23. Maghsoudloo M, Azimzadeh Jamalkandi S, Najafi A, Masoudi-Nejad A. Identification of biomarkers in common chronic lung diseases by co-expression networks and drug-target interactions analysis. Molecular Medicine. 2020;26(1):9.
24. Mortezaei Z, Lanjanian H, Masoudi-Nejad A. Candidate novel long noncoding RNAs, MicroRNAs and putative drugs for Parkinson's disease using a robust and efficient genome-wide association study. Genomics. 2017;109(3-4):158-64.
25. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
26. Hosseini SS, Abedi Z, Maghsoudloo M, Sheikh Beig Goharrizi MA, Shojaei A. Investigation of genes associated with primary open-angle glaucoma (POAG) using expression profile analysis. Journal of Ophthalmic and Optometric Sciences. 2019;3(3):37-54.
27. Maghsoudloo, M. and M.H. Noroozizadeh, A gene selection approach for Diabetic retinopathy microarray data classification using Ant Colony Optimization. Journal of Ophthalmic and Optometric Sciences. Volume, 2019. 3(4).
28. Masoudi-Sobhanzadeh Y, Motieghader H, Masoudi-Nejad A. FeatureSelect: A software for feature selection based on machine learning approaches. BMC Bioinformatics. 2019;20(1). doi: 10.1186/s12859-019-2754-0.
29. Abbasi K, Poso A, Ghasemi J, Amanlou M, Masoudi-Nejad A. Deep Transferable Compound Representation across Domains and Tasks for Low Data Drug Discovery. Journal of Chemical Information and Modeling. 2019;59(11):4528-39.
30. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Research. 2009;37(suppl_2):W305-W11.
31. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14(1):128.
32. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, et al. The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Research. 2019;48(D1):D845-D55.
33. Piñero J, Bravo À, Queralt-Rosinach N, Gutiérrez-Sacristán A, Deu-Pons J, Centeno E, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. Nucleic Acids Research. 2016;45(D1):D833-D9.
34. Ahmadi H, Ahmadi A, Azimzadeh-Jamalkandi S, Shoorehdeli MA, Salehzadeh-Yazdi A, Bidkhori G, et al. HomoTarget: a new algorithm for prediction of microRNA targets in Homo sapiens. Genomics. 2013;101(2):94-100. E
35. Masoudi-Nejad A, Wang E. Cancer modeling and network biology: accelerating toward personalized medicine. Semin Cancer Biol. 2015;30:1-3.
36. Abd-Elnaby M, Alfonse M, Roushdy M. Classification of breast cancer using microarray gene expression data: A survey. Journal of Biomedical Informatics. 2021;117:103764.
37. Mahendran N, Durai Raj Vincent PM, Srinivasan K, Chang C-Y. Machine Learning Based Computational Gene Selection Models: A Survey, Performance Evaluation, Open Issues, and Future Research Directions. Frontiers in Genetics. 2020;11.
38. Yan Q, Jiang Y, Huang H, Swaroop A,

Chew EY, Weeks DE, et al. Genome-Wide Association Studies-Based Machine Learning for Prediction of Age-Related Macular Degeneration Risk. Transl Vis Sci Technol. 2021;10(2):29.
39. Yan Q, Weeks DE, Xin H, Swaroop A, Chew EY, Huang H, et al. Deep-learning-based Prediction of Late Age-Related Macular Degeneration Progression. Nat Mach Intell. 2020;2(2):141-50.
40. Masoudi-Nejad A, Goto S, Jauregui R, Ito M, Kawashima S, Moriya Y, et al. EGENES: Transcriptome-Based Plant Database of Genes with Metabolic Pathway Information and Expressed Sequence Tag Indices in KEGG. Plant Physiology. 2007;144(2):857-66.
41. Darabi M, Izadi-Darbandi A, Masoudi-Nejad A, Naghavi MR, Nemat-Zadeh G. Bioinformatics study of the 3-hydroxy-3-methylglotaryl-coenzyme A reductase (HMGR) gene in Gramineae. Mol Biol Rep. 2012;39(9):8925-35.
42. Du Y, Kong N, Zhang J. Genetic Mechanism Revealed of Age-Related Macular Degeneration Based on Fusion of Statistics and Machine Learning Method. Frontiers in Genetics. 2021;12.
43. Oca AI, Pérez-Sala Á, Pariente A, Ochoa R, Velilla S, Peláez R, et al. Predictive Biomarkers of Age-Related Macular Degeneration Response to Anti-VEGF Treatment. J Pers Med. 2021;11(12). Epub 20211208.
44. Mah MM, Roverato N, Groettrup M. Regulation of Interferon Induction by the Ubiquitin-Like Modifier FAT10. Biomolecules. 2020;10(6).
45. Boehm AN, Bialas J, Catone N, Sacristan-Reviriego A, van der Spuy J, Groettrup M, et al. The ubiquitin-like modifier FAT10 inhibits retinal PDE6 activity and mediates its proteasomal degradation. Journal of Biological Chemistry. 2020;295(42):14402-18.
46. Radhakrishnan Y, Hamil KG, Yenugu S, Young SL, French FS, Hall SH. Identification, characterization, and evolution of a primate beta-defensin gene cluster. Genes Immun. 2005;6(3):203-10.
47. Cai H, Fields M, Hoshino R, Del Priore L. Effects of Aging and Anatomic Location on Gene Expression in Human Retina. Frontiers in Aging Neuroscience. 2012;4.
48. Whitmore SS, Wagner AH, DeLuca AP, Drack AV, Stone EM, Tucker BA, et al. Transcriptomic analysis across nasal, temporal, and macular regions of human neural retina and RPE/choroid by RNA-Seq. Exp Eye Res. 2014;129:93-106.
49. Roesler R, Kent P, Luft T, Schwartsmann G, Merali Z. Gastrin-releasing peptide receptor signaling in the integration of stress and memory. Neurobiol Learn Mem. 2014;112:44-52.
50. Ischia J, Patel O, Bolton D, Shulkes A, Baldwin GS. Expression and function of gastrin-releasing peptide (GRP) in normal and cancerous urological tissues. BJU Int. 2014;113 Suppl 2:40-7.
51. Petronilho F, Danielski LG, Roesler R, Schwartsmann G, Dal-Pizzol F. Gastrin-releasing peptide as a molecular target for inflammatory diseases: an update. Inflamm Allergy Drug Targets. 2013;12(3):172-7.
52. Chen YJ, Yeung L, Sun CC, Huang CC, Chen KS, Lu YH. Age-Related Macular Degeneration in Chronic Kidney Disease: A Meta-Analysis of Observational Studies. Am J Nephrol. 2018;48(4):278-91.
53. Roesler R, Luft T, Schwartsmann G. Targeting the gastrin-releasing peptide receptor pathway to treat cognitive dysfunctionassociated with Alzheimer's Disease. Dement Neuropsychol. 2007;1(2):118-23.

Footnotes and Financial Disclosures

Conflict of interest:

The authors have no conflict of interest with the subject matter of the present manuscript.