EyeDiseases: an integrated resource for dedicating to genetic variants, gene expression and epigenetic factors of human eye diseases

Jian Yuan [1,2,3,†], **Fukun Chen**[1,2,3,†], **Dandan Fan**[1,2,3,†], **Qi Jiang**[1,2,3], **Zhengbo Xue**[1,2,3], **Ji Zhang**[1,2,3], **Xiangyi Yu**[1,2,3], **Kai Li**[4], **Jia Qu**[1,2,*] **and Jianzhong Su** [1,2,3,*]

[1]School of Ophthalmology & Optometry and Eye Hospital, Wenzhou Medical University, Wenzhou 325027, China, [2]National Clinical Research Center for Ocular Disease, Wenzhou 325027, China, [3]Institute of Biomedical Big Data, Wenzhou Medical University, Wenzhou 325027, China and [4]Wenzhou Institute, University of Chinese Academy of Sciences, Wenzhou 325011, Zhejiang, China

ABSTRACT

Eye diseases are remarkably common and encompass a large and diverse range of morbidities that affect different components of the visual system and visual function. With advances in omics technology of eye disorders, genome-scale datasets have been rapidly accumulated in genetics and epigenetics field. However, the efficient collection and comprehensive analysis of different kinds of omics data are lacking. Herein, we developed EyeDiseases (https://eyediseases.bio-data.cn/), the first database for multi-omics data integration and interpretation of human eyes diseases. It contains 1344 disease-associated genes with genetic variation, 1774 transcription files of bulk cell expression and single-cell RNA-seq, 105 epigenomics data across 185 kinds of human eye diseases. Using EyeDiseases, we investigated SARS-CoV-2 potential tropism in eye infection and found that the SARS-CoV-2 entry factors, ACE2 and TMPRSS2 are highly correlated with cornea and keratoconus, suggest that ocular surface cells are susceptible to infection by SARS-CoV-2. Additionally, integrating analysis of Age-related macular degeneration (AMD) GWAS loci and co-expression data revealed 9 associated genes involved in HIF-1 signaling pathway and voltage-gate potassium channel complex. The EyeDiseases provides a valuable resource for accelerating the discovery and validation of candidate loci and genes contributed to the molecular diagnosis and therapeutic vulnerabilities with various eyes diseases.

INTRODUCTION

Eye diseases encompass a large and diverse range of morbidities that affect different components of the visual system and visual function (1). In the medium- to long-term, they could cause vision impairment including blindness, such as diabetic retinopathy, glaucoma, age-related macular degeneration and cataract (2). Globally, it is estimated that at least 2.2 billion people have a vision impairment or blindness based on recently published epidemiological data (3,4). With behavioural and lifestyle changes, and urbanization, population growth and ageing will dramatically increase the number of people with eye conditions, vision impairment and blindness in the coming decades (5–9). Meanwhile, the burden of visual impairments and eye conditions tends to be greater in low and middle-income countries and underserved populations, such as women, indigenous peoples, persons with certain kinds of disability, and in rural communities.

Genetic variation also plays important roles in eye diseases in both children and adults. Over 60% of cases of childhood blindness are caused by genetic factors, such as congenital glaucoma, ocular malformations, atrophy of the optic nerve and retinitis pigmentosa (10). Although the triggering factors for many eye diseases remain unknown, genome-wide association studies (GWAS) have identified many nucleotide polymorphisms, with some of them likely to be causal, and associated risk loci (11–16). Additionally, evidences are accumulating that epigenetic changes resulting from environmental factors may play a significant role in eye disorder, and interventions to modify these factors hold increasing promise (17–19). Together, these results indicate that either hidden factors such as epigenetics or gene-environment interactions constitute the predominant factor for eye disorder risks.

*To whom correspondence should be addressed. Tel: +86 577 88068270; Email: sujz@wmu.edu.cn
Correspondence may also be addressed to Jia Qu. Tel: +86 88068272; Email: qujia@eye.ac.cn
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.
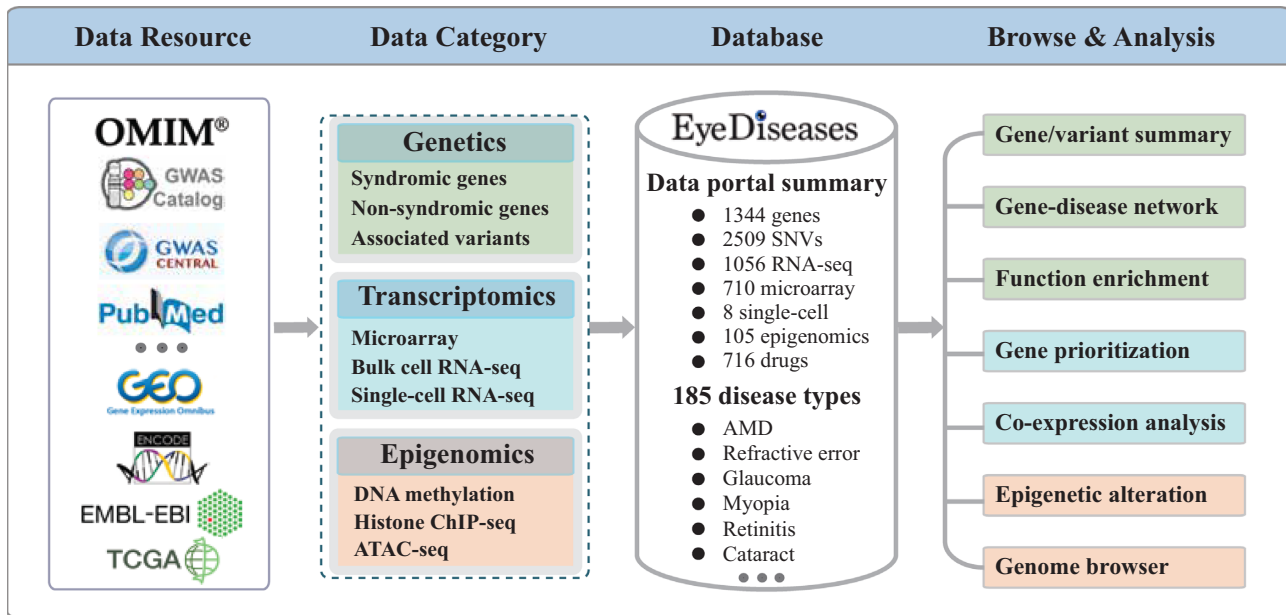
Figure 1. Data processing workflow and overall architecture of EyeDiseases. EyeDiseases is an omics database which integrates gene, mutation, RNA expression, Single-Cell RNAseq, DNA methylation, chromatin accessibility, histone modification and drug data. The database provides three functions, 'Genetics', 'Transcriptomics', and 'Epigenomics', to help researchers visualize the relationships between eye disorders and candidate genes. Data in this database were performed extended functional annotation, such as gene–disease networks, gene ontology, pathway analysis, gene prioritization, co-expression and epigenetic alteration.

Disease-specific databases are useful for researchers to find and retrieve the appropriate information. However, a gene–disease association database dedicated to eye diseases is still unavailable. With the next-generation sequencing technologies providing even more power for gene discoveries (20), it has become more and more difficult for an ophthalmologist to understand with confidence how many genes have been associated with eye diseases. In this study, we presented, Eyediseases, a manually curated database, aimed to enhance our understanding of eye diseases. Eye-Diseases is a multi-omics database which integrates gene, mutation, RNA expression, Single-Cell RNA-seq, DNA methylation, chromatin accessibility, histone modification and drug datasets (Figure 1). The database provides three functions, 'Genetics', 'Expression' and 'Epigenomics', to help researchers visualize the relationships between eye disease and associated genes. Data in this database were performed extended functional annotation, such as gene-disease networks, gene ontology, pathway analysis, co-expression and epigenetic alteration. Therefore, EyeDiseases uses uniform processing pipelines to create high-quality, consistent, and reproducible data, which provides the scientific community with freely available information on eye disease-related gene, RNA expression and epigenetic regulation.

MATERIALS AND METHODS

Database structure and organization

Eyediseases integrates multi-omics data from 1138 projects. The main diseases include Refractive error, Myopia, Retinitis, Glaucoma, Cataract, Age-related macular degener-

Table 1. Data content in Eyediseases

	Data type	Total No.
Genetics	Genes	1344
	Variants	2509
Expression	Microarray	710
	RNA-seq	1056
Single cell	RNA-seq	8
Epigenomics	ATAC-seq	7
	H3K4me1	2
	H3K4me3	4
	H3K9me3	1
	H3K27ac	5
	H3K27me3	1
	Bisulfite-Seq	2
	RRBS	3
	methylation 450	80

ation. There are 2509 records of single nucleotide variants (SNVs) and 1344 records of genes, and the corresponding annotation are based on hg19 and hg38. There are 710 and 1056 records in the microarray and bulk cell RNA-seq expression data, respectively. The single-cell RNAseq data includes six normal eye tissues and two disease samples. There are 105 records in the epigenomics which contain DNA methylation, ATAC-seq, H3K27ac, H3K27me3, H3K4me1, H3K4me3, H3K9me3 data from 15 normal tissues and 12 disease tissues (Table 1). There are 716 records regarding variant/gene-related drug information collected from the DrugBank.

The EyeDiseases comprises seven modules: Search, Browse, Analysis, Statistics, Download, Document and Contact. The search module provides three modes by dis-

ease, genes or variants. The browse module contains four perspectives: disease, gene expression, single-cell transcriptomes and epigenomics. The analysis comprises five analytical perspectives: gene-disease network, functional annotations, gene prioritization, co-expression analysis and epigenetic alteration.

Data sources and preprocessing

The disease list was obtained from the Medical Subject Headings (MeSH) (https://meshb.nlm.nih.gov/), cross-referenced with the list of diseases included in NCBI. Consequently, total 185 eye diseases were retained, which gene associations can be found by our data collection process (see below).

One important feature of EyeDiseases is comprehensive collection of data from major genetic studies for most eye disorders. For gene linked with syndromic eye disorder, we searched Online Mendelian Inheritance in Man databases (OMIM, https://omim.org/) to get the official disease names and linked all the disorders to OMIM, and searched PubMed for additional citations using the query 'disease name'. All citations were double-checked manually. For association, we first searched and collected data with MeSH and OMIM name from the established GWAS catalog (https://www.ebi.ac.uk/gwas/) and GWAS Central (https://www.gwascentral.org/). To obtain a complete resource of multi-omics relevant to eye disorders, comprehensive searches were performed for eye studies.

We searched the GEO (https://www.ncbi.nlm.nih.gov/geo/) and EMBL-EBI (https://www.ebi.ac.uk/) database with the following query terms: '(eye OR disease name) AND (microarray OR transcriptome OR RNAseq)' for expression profiling studies, '(eye OR disease name) AND (single cell RNA-seq OR scRNA-seq)' for single cell profiling studies and '(eye OR disease name) AND (DNA methylation OR histone OR ATAC-seq)' for epigenomic datasets. Each type of data includes normal tissues and disease tissues. We also achieved the datasets from ENCODE (https://www.encodeproject.org/), and TCGA (https://portal.gdc.cancer.gov/) database.

For the expression profile generated with the next-generation RNA-seq platform, we aligned RNA-seq reads in BAM format using a two-pass method with STAR. Following alignment, gene-level expression quantifications based on a collapsed version of a reference transcript annotation, provided as read counts and TPM. For the microarray profile, Illumina microarrays were log2 transformed and quantile normalized using the lumi package in R. Affymetrix microarrays were RMA normalized (background correction, log2 transformation, quantile normalization, and probe summarization) using the affy package in R. Agilent microarrays were normalized using the limma package in R (read.maimages, backgroundCorrect, normalizeWithinArrays, normalizeBetweenArrays, getEAWP functions). To provide a systematic nomenclature for assessment of gene expression across platforms, microarray probes were re-mapped to Ensembl gene IDs and gene symbol using the biomaRt package in R, taking the maximum mean signal across all probes available for each gene, using the collapseRows function. The collapseRows MaxMean function was explicitly developed to perform cross-platform microarray meta-analysis and has been extensively validated in its ability to increase between-study consistency and enhance reproducibility (21). Strict care was taken to ensure data integrity.

For single cell RNA-seq, we following the Seurat Clustering Tutorial (https://satijalab.org/seurat) to cluster the cells and find cell markers that define clusters via differential expression (22). DNA methylation, histone ChIP-seq and ATAC-seq analysis were performed by the ENCODE and TCGA data coordinating center uniform processing pipelines to create high-quality, consistent, and reproducible data (23). We also developed a user-friendly and configurable genome browser through which multiple genomic and epigenomic resources can be visualized simultaneously.

Data analysis

In order to better understanding the phenotypic and genetic complexity in eye diseases, gene-disease network developed by D3.js was constructed to demonstrate the overlapping features between diseases. Next, enriched gene ontology, including biological process, cellular component, molecular function and functional pathways, were identified by DAVID. For each gene, we computed and displayed its co-expressed genes by weighted gene co-expression network analysis (WGCNA) in disease and normal samples, respectively. We used the soft threshold method for pearson correlation analysis of the expression profiles to determine the connection strengths between two genes. To select and prioritize candidate genes for complex eye diseases, we calculated the gene significance (GS) of a gene as the Pearson correlation coefficient for testing the association of the gene expression to the disease type. Generally, if GS were highly associated, it implied that genes were most significantly correlated with the disease. Furthermore, to systematic characterize candidate causal genes for AMD, we observed the assay for transposase-accessible chromatin with sequencing (ATAC-Seq) peak in AMD associated genes and find the extent to which epigenetic changes regulate AMD.

Implementation

EyeDiseases codes were developed using an integrated development environment, the server is implemented based on the Flask (https://flask.palletsprojects.com) framework. PostgreSQL (12–12.1–2) is used to store and manage the data information of the database. EyeDiseases uses Vue (2.6.11) for most of its front-end components and adopts. The UI that is compatible with PC and mobile terminals is completed through Bulma (https://bulma.zcopy.site). All drawing functions in the EyeDiseases are realized by using D3 (5.16.0) and Plotly (http://chart-studio.plotly.com), The tables of the database are visualized using ag-gird. The entire web application is deployed on CentOS (one of the most popular Linux distribution; https://www.centos.org) operating system. All the data and code used is publicly available at https://github.com/sulab-wmu/eyediseases.

RESULTS

Database interface

We developed a user-friendly web interface for Eyediseases. The user may access all data and conduct analysis through the web interface (https://eyediseases.bio-data.cn/). To support facilitating access to genetic and epigenetic data in Eyediseases, five main search and explore approaches have been provided in home page: search by disease, search by gene, search by variant, explore by transcriptomics and explore by epigenomics (Figure 1). For disease query, we provide widgets that allow users to search for diseases by user supplied keywords (e.g. age-related macular degeneration). After clicking on a disease name, users will be redirected to a new page to choose data types. Once the user clicks genetics module, a detailed gene and variant page is shown, includes a list of associated genes and evidence supporting the association. For genetics query, users can find a quick search box on the home page to search by gene symbol (e.g. CFH) or variant (e.g. rs10490924). The result page for single gene search has five major views: (i) summary view provides an overview of general information of the queried gene; (ii) disease and phenotype view provides an associated disease and phenotype information from HPO; (iii) external reference view provides web link to the related databases when available; (iv) drug views provides a suite of a drug interacts with queried gene powered by the DrugBank platform and (v) expression view displays the expression patterns of a specific gene in the archived transcriptome datasets. The 'variant search interface' allows the user to input a single variant site (e.g. rs10490924), and return results for the collected mutation information list and the population frequency of the mutation. In disease and genetics query, all data sets are relationally linked. Once the user clicks a gene symbol or rs ID, a detail gene or variant page is shown.

In the transcriptomics module, the user may browse expression data by (i) clicking one of the three data categories ('Multi Gene Query', 'Eye Development' and 'Single-Cell RNAseq') on the home page; or (ii) selecting expression or single-cell RNAseq of 'Browse' on the function bar on the top of the web page. The expression search section, multi-gene query visualization allows users to select platform, species and datasets, and enter a list of genes of interest.

Next, users can click submit to view the expression of those genes in a table and a heat map that shows expression level across all normal or disease tissues. By default, the genes and tissues are clustered using hierarchical clustering. You can download the expression matrix and figure for further analysis. In the single-cell RNA-seq section, the results are showed by project. The view shows t-SNE plots of all cell type as well as a computed matrix table of cell markers with columns (i) cluster ID, (ii) cell type, (iii) cell marker, (iv) P-value, (v) average $\log_2$ fold change; (v) PC1 and (vi) PC2. In the Epigenomics view, Eyediseases provides a user-friendly and configurable genome browser through which multiple genomic and epigenomic resources can be visualized simultaneously. The genomic viewer connecting to a PostgreSQL backend is used to show the landscape for normal and disease tissue epigenomic information. Features of the viewer include the ability to zoom through the given regions, to enter a region by specifying the genomic coordinates, to change the order of tracks and to show and hide certain feature tracks and configure the appearance of the displayed information.

Example usage 1: EyeDiseases is used to find susceptible tissues in SARS-CoV-2 infection

COVID-19 runs rampant around the world (24). Ocular signs and symptoms are observed in a subset of patients with COVID-19, including conjunctival hyperemia, chemosis, epiphora, increased secretions, ocular pain, photophobia and dry eye (25). Recently, immunohistochemical studies show the ocular surface localisation of SARS-CoV-2 entry factors, ACE receptor and TMPRSS2 protease in human eyes (26,27). We performed association analysis of individual genes with the eye tissues and diseases by WGCNA co-expression network analysis. Using 'gene significance' module in 'analysis' page of database, we searched seven susceptibility gene, including ACE and TMPRSS2, to view the gene-tissue/disease association score in a table. We found that ACE2 and TMPRSS2 expression level were highly correlated with corneas and keratoconus (Table 2). This result indicated that corneal cells are susceptible to infection by SARS-CoV-2 and eye protection is advisable to prevent contamination from contact lenses, external droplets and hand-eye contact.

Example usage 2: integrating GWAS and co-expression network identifies AMD associated module in Eyediseases

Age-related macular degeneration (AMD) is the leading cause of blindness in the elderly with limited therapeutic options (28). Genetic variants can help uncover disease mechanisms and provide entry points into therapy. Since the first AMD GWAS in 2005 (29), genome-wide association studies for AMD have identified dozens of associations; yet, the genes responsible for most associations remain elusive. Genes underlying a disease are often functionally related and functionally related genes are often co-expressed (30). Therefore, we can build a co-expression network of eye tissues to predict associated genes at available GWAS loci of eye diseases. We began by generating a list of 190 genes located within age-related macular degeneration (AMD) GWAS loci from 19 projects in 'disease-genetics' function. Next, we constructed co-expression network in R which partitioned into 12 network modules using 107 retina expression profiles from 'co-expression' function (Figure 2A). Of the 12 network modules, two (modules 7 and 9) were significantly enriched for AMD associated genes (OR = 4.2, Fisher's P-value = 0.01 and OR = 5.1, Fisher's P-value = 0.004, respectively) (Figure 2B). To more formally evaluate the relationship between the two modules, we created a network based on the eigengene of each of the 12 modules. The module eigengene corresponds to the first principal component of a given module. In this network, the module 7 and 9 eigengenes were not correlated, which imply AMD have different pathogenesis (Figure 2C).

Next, we evaluated gene ontology (GO) enrichments to determine function of the genes between module 7
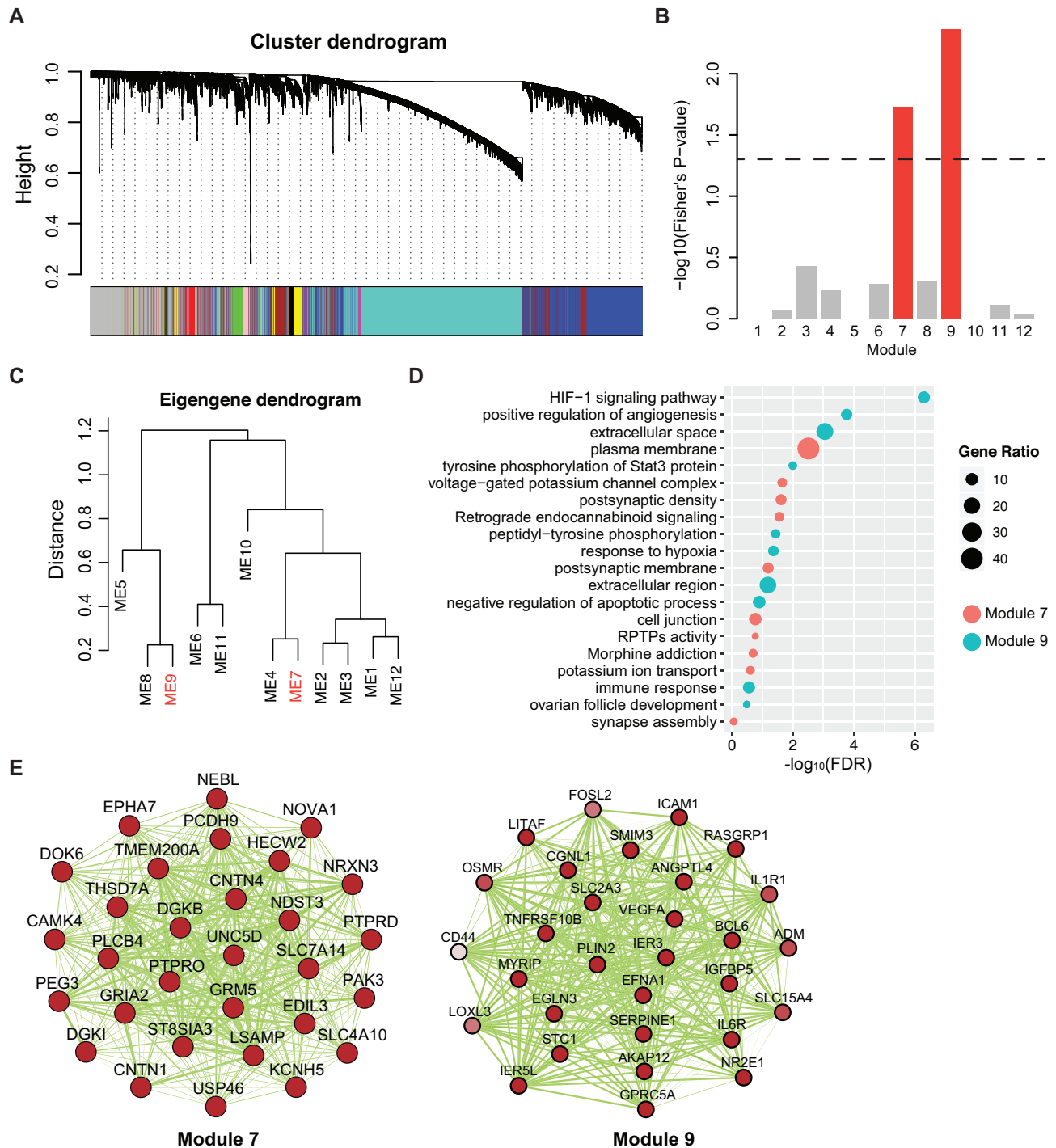
Figure 2. Identification of the AMD Functional Module. (**A**) Weighted gene co-expression network analysis (WGCNA) in retina, showing a hierarchical clustering tree of co-expression modules. Each module corresponds to a branch, which is labeled by a distinct color shown underneath. (**B**) Enrichment of genes located in AMD GWAS regions in network modules 7 and 9. (**C**) Eigengenes for modules 7 and 9 cluster. (**D**) GO enrichment analysis of the module 7 and 9. The rich factor refers to the proportion of enriched genes for each term. (**E**) Top hub genes along with edges supported by co-expression are shown. Hub genes are defined by being the top 30 most connected genes based on kME intermodular connectivity. Increasing edge width indicates increasing topological overlap measure (tom). Increasing color gradient indicates increasing connectivity degree.

Table 2. Gene significance (GS) score of SARS-CoV-2 entry factors and susceptibility gene in disease and normal tissues

Normal	ACE2	TMPRSS2	CCR9	FYCO1	CXCR6	XCR1	SLC6A20	LZTFL1
cornea	0.364	0.300	0.032	− 0.425	0.022	− 0.017	− 0.090	− 0.258
retina	− 0.119	− 0.155	− 0.264	0.225	− 0.254	− 0.192	− 0.301	0.305
retina_macula	− 0.138	− 0.137	− 0.146	0.086	− 0.174	− 0.162	− 0.255	0.311
retina_non_macula	− 0.144	− 0.134	− 0.146	0.292	− 0.195	− 0.176	− 0.265	0.424
rpe_macula	0.305	− 0.016	0.078	− 0.073	0.458	0.222	0.319	− 0.410
rpe_non_macula	− 0.037	0.120	0.396	− 0.120	0.243	0.263	0.609	− 0.448
retinal_endothelial_cells	0.053	0.379	0.468	− 0.170	− 0.015	0.414	− 0.026	− 0.157
ipsc_derived_retinal_organoids	− 0.035	0.045	− 0.029	− 0.122	− 0.038	− 0.012	− 0.056	0.077
trabecular_meshwork_cells	− 0.061	− 0.046	− 0.083	− 0.308	− 0.118	− 0.056	− 0.088	− 0.230
Disease								
Keratoconus	0.858	0.394	0.293	0.390	0.548	− 0.538	− 0.389	0.461
Age-related macular degeneration	− 0.548	− 0.545	− 0.231	− 0.436	− 0.522	0.731	0.554	− 0.561
Diabetic retinopathy	− 0.015	0.167	0.077	0.080	− 0.032	− 0.308	− 0.303	0.176
Primary open-angle glaucoma	0.013	− 0.026	− 0.099	0.264	0.423	− 0.117	− 0.219	0.002
Retinitis pigmentosa	− 0.032	0.391	0.010	0.025	0.038	− 0.102	0.113	0.097
Retinoblastoma	− 0.053	0.201	0.076	0.064	− 0.027	− 0.354	− 0.264	0.310

and 9 (Figure 2D). Notably, the module 9 was significantly (FDR < 0.05) enriched for Hypoxia-Inducible Factor (HIF)-1 signaling pathway (FDR = 5.04×10^{-7}) and positive regulation of angiogenesis (FDR = 1.73×10^{-4}), which have been associated with choroidal neovascularization and the progression of AMD into the neovascular clinical phenotype (nAMD). Regard to nAMD, HIFs regulate the expression of multiple growth factors and cytokines involved in angiogenesis and inflammation, hallmarks of neovascular AMD. HIFs also increase RPE apoptosis and autophagy, similarly to the choriocapillaries, which stimulates the endothelial proliferation observed in nAMD. Additionally, we found that the module 7 was enriched for voltage-gate potassium channel complex (FDR = 0.022) and postsynaptic density (FDR = 0.024), which imply neuronal processes are essential elements in human retinal tissues in response to AMD. Another way to infer the function of the modules is based on the known function of highly connected genes with central positions within the modules ("hub" genes). We explored the strongest connections in module 6 and 9 using Cytoscape software (Figure 2E). These results indicated that integrating GWAS and co-expression data, it is possible to provide insight into the identity of associated GWAS genes involved in AMD-related pathways.

DISCUSSION

EyeDiseases is a first comprehensive database that has been developed for the integration of multiple omics and interactive analysis in eye diseases. The data in EyeDiseases are from GWAS catalog, GWAS Central, OMIM, the GEO database, the EMBL-EBI, ENCODE, TCGA and Drugbank database with thousands of samples included. Data type consisting of the genome, transcriptome, epigenome, single-cell and targeted drug. EyeDiseases is a time-saving, free, and intuitive tool for tapping the full potential of publicly available genomics big data, which enables biologists and clinicians without any programming experience to obtain ready-to-use multi-omics data and perform a diverse range of data analyses. It also has the potential to become a one-stop service for data query and analysis for the scientific and clinical community associated with the eye disorder field.

The majority of genetic variants potentially associated with eye diseases (e.g. AMD) that are identified by GWAS are located in non-coding sequences, suggesting that their functions, if any, may be exerted through regulating the expression of genes producing proteins or non-coding RNAs (31–34). Thus, integration of GWAS and gene expression network is essential for elucidating the underlying mechanisms for causal gene in diseases (35–37). Our Eyediseases database provides integrative data, from which users can explore the functional biological network in both disease tissues and normal eye tissues. As an effective application of Eyediseases to identify potential causal gene for eye disorders, we used an eye-relevant co-expression network to inform GWAS loci robustly associated with AMD and identify putative causal genes. We identified two network modules (7 and 9) that contained more genes implicated by GWAS than would be expected by chance. Modules 7 and 9 were enriched for processes the known mechanisms of AMD, such as 'Hypoxia-Inducible Factor (HIF)-1 signaling pathway' (38,39). Based on our results, we believe that significant insight into complex disease etiology can be gained using networks to inform GWAS.

To better serve the research community of eye diseases, we will not only continuously update multi-omics data from both eye disease and normal samples, but also develop new analytical features for further exploration of the available big genomic data. Our next plan is to obtain more public multi-omics data on eye diseases, such CNV, proteomics, ocular image and metagenomics, and build an enhanced database based on comprehensive genome-level data for the effective visualization and analysis of all human genes in the future. We believe Eyediseases could serve as a very useful public resource for both bench scientists and computational biologists, and contribute to clinical and translational studies.

REFERENCES

1. W.H.O. (2019) World report on vision.
2. W.H.O. (2012) *Global Data on Visual Impairments 2010*. WHO, Geneva, pp. 1–5.
3. Fricke,T.R., Tahhan,N., Resnikoff,S., Papas,E., Burnett,A., Ho,S.M., Naduvilath,T. and Naidoo,K.S. (2018) Global prevalence of presbyopia and vision impairment from uncorrected presbyopia: systematic review, meta-analysis, and modelling. *Ophthalmology*, **125**, 1492–1499.
4. Bourne,R.R., Flaxman,S.R., Braithwaite,T., Cicinelli,M.V., Das,A., Jonas,J.B., Keeffe,J., Kempen,J.H., Leasher,J. and Limburg,H. (2017) Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: a systematic review and meta-analysis. *Lancet Global Health*, **5**, e888–e897.
5. Chen,M., Yu,X., Xu,J., Ma,J., Chen,X., Chen,B., Gu,Y. and Wang,K. (2019) Association of gene polymorphisms with primary open angle glaucoma: a systematic review and meta-analysis. *Invest. Ophthalmol. Vis. Sci.*, **60**, 1105–1121.
6. Wong,T.Y., Sun,J., Kawasaki,R., Ruamviboonsuk,P., Gupta,N., Lansingh,V.C., Maia,M., Mathenge,W., Moreker,S. and Muqit,M.M. (2018) Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. *Ophthalmology*, **125**, 1608–1622.
7. Wong,W.L., Su,X., Li,X., Cheung,C.M.G., Klein,R., Cheng,C.-Y. and Wong,T.Y. (2014) Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Global Health*, **2**, e106–e116.
8. Azari,A.A. and Barney,N.P. (2013) Conjunctivitis: a systematic review of diagnosis and treatment. *JAMA*, **310**, 1721–1730.
9. Ye,J., He,J., Wang,C., Wu,H., Shi,X., Zhang,H., Xie,J. and Lee,S.Y. (2012) Smoking and risk of age-related cataract: a meta-analysis. *Invest. Ophthalmol. Vis. Sci.*, **53**, 3885–3895.
10. Foster,A. and Gilbert,C. (1992) Epidemiology of childhood blindness. *Eye*, **6**, 173–176.
11. Tedja,M., Wojciechowski,R., Hysi,P., Eriksson,N., Furlotte,N., Verhoeven,V., Iglesias,A., Meester-Smoor,M., Tompson,S., Fan,Q. *et al.* (2018) Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. *Nat. Genet.*, **50**, 834–848.
12. Khor,C., Do,T., Jia,H., Nakano,M., George,R., Abu-Amero,K., Duvesh,R., Chen,L., Li,Z., Nongpiur,M. *et al.* (2016) Genome-wide association study identifies five new susceptibility loci for primary angle closure glaucoma. *Nat. Genet.*, **48**, 556–562.
13. Hysi,P., Choquet,H., Khawaja,A., Wojciechowski,R., Tedja,M., Yin,J., Simcoe,M., Patasova,K., Mahroo,O., Thai,A. *et al.* (2020) Meta-analysis of 542,934 subjects of European ancestry identifies new genes and mechanisms predisposing to refractive error and myopia. *Nat. Genet.*, **52**, 401–407.
14. Fritsche,L., Igl,W., Bailey,J., Grassmann,F., Sengupta,S., Bragg-Gresham,J., Burdon,K., Hebbring,S., Wen,C., Gorski,M. *et al.* (2016) A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat. Genet.*, **48**, 134–143.
15. Bailey,J., Loomis,S., Kang,J., Allingham,R., Gharahkhani,P., Khor,C., Burdon,K., Aschard,H., Chasman,D., Igo,R. *et al.* (2016) Genome-wide association analysis identifies TXNRD2, ATXN2 and FOXC1 as susceptibility loci for primary open-angle glaucoma. *Nat. Genet.*, **48**, 189–194.
16. Craig,J., Han,X., Qassim,A., Hassall,M., Cooke Bailey,J., Kinzy,T., Khawaja,A., An,J., Marshall,H., Gharahkhani,P. *et al.* (2020) Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. *Nat. Genet.*, **52**, 160–166.
17. Hoang,T., Horowitz,E., Chaffee,B., Qi,P., Flake,R., Bruney,D., Rasor,B., Rosalez,S., Wagner,B. and Robinson,M. (2017) Lens development requires DNMT1 but takes place normally in the absence of both DNMT3A and DNMT3B activity. *Epigenetics*, **12**, 27–40.
18. Zhang,W., Song,M., Qu,J. and Liu,G. (2018) Epigenetic modifications in cardiovascular aging and diseases. *Circ. Res.*, **123**, 773–786.
19. Bonnin,N., Belville,C., Chiambaretta,F., Sapin,V. and Blanchon,L. (2014) DNA methyl transferases are differentially expressed in the human anterior eye segment. *Acta Ophthalmol. (Copenh)*, **92**, e366–e371.
20. Lohmann,K. and Klein,C. (2014) Next generation sequencing and the future of genetic diagnosis. *Neurotherapeutics*, **11**, 699–707.
21. Gandal,M., Haney,J., Parikshak,N., Leppa,V., Ramaswami,G., Hartl,C., Schork,A., Appadurai,V., Buil,A., Werge,T. *et al.* (2018) Shared molecular neuropathology across major psychiatric disorders parallels polygenic overlap. *Science*, **359**, 693–697.
22. Satija,R., Farrell,J., Gennert,D., Schier,A. and Regev,A. (2015) Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.*, **33**, 495–502.
23. Davis,C., Hitz,B., Sloan,C., Chan,E., Davidson,J., Gabdank,I., Hilton,J., Jain,K., Baymuradov,U., Narayanan,A. *et al.* (2018) The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.*, **46**, D794–D801.
24. Zhu,N., Zhang,D., Wang,W., Li,X., Yang,B., Song,J., Zhao,X., Huang,B., Shi,W. and Lu,R. (2020) A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.*, **382**, 727–733.
25. Wu,P., Duan,F., Luo,C., Liu,Q., Qu,X., Liang,L. and Wu,K. (2020) Characteristics of ocular findings of patients with coronavirus disease 2019 (COVID-19) in Hubei Province, China. *JAMA Ophthalmol.*, **138**, 575–578.
26. Collin,J., Queen,R., Zerti,D., Dorgau,B., Georgiou,M., Djidrovski,I., Hussain,R., Coxhead,J., Joseph,A., Rooney,P. *et al.* (2020) Co-expression of SARS-CoV-2 entry genes in the superficial adult human conjunctival, limbal and corneal epithelium suggests an additional route of entry via the ocular surface. *Ocular Surf.*, **19**, 190–200.
27. L,Z., Z,X., GM,C., US,S., CG,E. and EJ,D. (2020) ACE2 and TMPRSS2 are expressed on the human ocular surface, suggesting susceptibility to SARS-CoV-2 infection. *Ocular Surf.*, **18**, 537–544.
28. Mitchell,P., Liew,G., Gopinath,B. and Wong,T. (2018) Age-related macular degeneration. *Lancet (London, England)*, **392**, 1147–1159.
29. Klein,R., Zeiss,C., Chew,E., Tsai,J., Sackler,R., Haynes,C., Henning,A., SanGiovanni,J., Mane,S., Mayne,S. *et al.* (2005) Complement factor H polymorphism in age-related macular degeneration. *Science (New York, N. Y.)*, **308**, 385–389.
30. Calabrese,G.M., Mesner,L.D., Stains,J.P., Tommasini,S.M., Horowitz,M.C., Rosen,C.J. and Farber,C.R. (2017) Integrating GWAS and co-expression network data identifies bone mineral density genes SPTBN1 and MARK3 and an osteoblast functional module. *Cell Syst.*, **4**, 46–59.
31. Yan,Q., Ding,Y., Liu,Y., Sun,T., Fritsche,L.G., Clemons,T., Ratnapriya,R., Klein,M.L., Cook,R.J. and Liu,Y. (2018) Genome-wide analysis of disease progression in age-related macular degeneration. *Hum. Mol. Genet.*, **27**, 929–940.
32. Fritsche,L.G., Igl,W., Bailey,J.N.C., Grassmann,F., Sengupta,S., Bragg-Gresham,J.L., Burdon,K.P., Hebbring,S.J., Wen,C. and Gorski,M. (2016) A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat. Genet.*, **48**, 134–143.
33. AMD Gene Consortium (2013) Seven new loci associated with age-related macular degeneration. *Nat. Genet.*, **45**, 433–439.
34. Arakawa,S., Takahashi,A., Ashikawa,K., Hosono,N., Aoi,T., Yasuda,M., Oshima,Y., Yoshida,S., Enaida,H. and Tsuchihashi,T. (2011) Genome-wide association study identifies two susceptibility loci for exudative age-related macular degeneration in the Japanese population. *Nat. Genet.*, **43**, 1001–1004.

35. Jia,P. and Zhao,Z. (2014) Network-assisted analysis to prioritize GWAS results: principles, methods and perspectives. *Hum. Genet.*, **133**, 125–138.

36. Huan,T., Meng,Q., Saleh,M.A., Norlander,A.E., Joehanes,R., Zhu,J., Chen,B.H., Zhang,B., Johnson,A.D. and Ying,S. (2015) Integrative network analysis reveals molecular mechanisms of blood pressure regulation. *Mol. Syst. Biol.*, **11**, 799.

37. Gustafsson,M., Gawel,D.R., Alfredsson,L., Baranzini,S., Björkander,J., Blomgran,R., Hellberg,S., Eklund,D., Ernerudh,J. and Kockum,I. (2015) A validated gene regulatory network and GWAS identifies early regulators of T cell–associated diseases. *Sci. Transl. Med.*, **7**, 313ra178.

38. Mammadzada,P., Corredoira,P.M. and André,H. (2020) The role of hypoxia-inducible factors in neovascular age-related macular degeneration: a gene therapy perspective. *Cell. Mol. Life Sci.*, **77**, 819–833.

39. Krishna Vadlapatla,R., Dutt Vadlapudi,A. and Mitra,A.K. (2013) Hypoxia-inducible factor-1 (HIF-1): a potential target for intervention in ocular neovascular diseases. *Curr. Drug Targets*, **14**, 919–935.