

Multiple genome-wide association studies of type 2 diabetes implicate several genes are associated with diabetic retinopathy based on UK Biobank

Tengda Cai¹, Qi Pan¹, Yiwen, Tao¹, Charvi Nangia², Aravind Lathika Rajendrakumar³, Tania Dottorini⁴, Mainul Haque⁵, Colin Palmer², Weihua Meng^{1,2*}

1 Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute, University of Nottingham Ningbo China, Ningbo, China, 315100

2 Division of Population Health and Genomics, School of Medicine, University of Dundee, Dundee, UK, DD2 4BF

3 Duke University, 2080 Duke University Road, Durham, USA, NC 27708

4 School of Veterinary Medicine and Science, University of Nottingham, Nottingham, UK, LE12 5RD

5 School of Mathematical Sciences, University of Nottingham Ningbo China, Ningbo, China, 315100

Corresponding Author: Dr Weihua Meng

Address: Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK, DD2 4BF; Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute, University of Nottingham Ningbo China, Ningbo, China, 315100. Tel: +44 1382 383419; +86 574 88188632. Email: w.meng@dundee.ac.uk; weihua.meng@nottingham.edu.cn. ORCID: 0000-0001-5388-8494

ORCID codes

Tengda Cai: 0000-0003-1617-506X

Qi Pan: 0000-0003-0135-9122

Yiwen Tao: 0009-0000-0102-4205

Charvi Nangia: 0000-0003-3984-2603

Aravind Lathika Rajendrakumar: 0000-0002-4286-7265

Tania Dottorini: 0000-0001-7733-906X

Mainul Haque: 0000-0003-1836-8076

Colin Palmer: 0000-0002-6415-6560

Weihua Meng: 0000-0001-5388-8494

Abstract

Purpose

To identify the genetic variants associated with diabetic retinopathy in type 2 patients from the UK Biobank cohort ($n = 17,015$) and supporting replication cohorts GODARTS ($n = 5,013$), GOSHARE ($n = 1,754$), Caucasian Australians ($n = 518$), FinnGen ($n = 206,664$) and Chinese ($n = 1,007$).

Methods

Totally eleven genome-wide association studies were applied to search for significant genetic variants.

Results

We found 5 different loci associated with type 2 diabetic retinopathy in or nearest gene *EYA2*, *MPDZ*, *NTNG1*, *CTAGE14P* and *MREGP1*. In the primary GWAS, a significant SNP rs6066146 located in gene *EYA2* showed a p value of 4.21×10^{-8} and may play a role in the development of the disease, with “spleen” reaching a significant level produced by tissue expression analysis. Corresponding heritability of DR was estimated to be 26.73% by SumHer. Among five genes, we found that genes *EYA2*, *MPDZ*, *NTNG1* had genetic interactions and may affect the complex development of retinal blood vessels.

Conclusion

Diabetic retinopathy is a complication of diabetes that affects the eyes. It is highly likely to occur when high blood sugar damages the retinal blood vessels. There is limited awareness regarding the pathogenesis of DR. Our study identified multiple loci associated with diabetic retinopathy, which may lead to personalized treatments to reduce the burden of the disease.

Keywords: Diabetic retinopathy; UK Biobank; genome-wide association study; tissue expression analysis; heritability.

Introduction

Diabetic retinopathy (DR), a disease influenced by multiple genes, stands as one of the most frequent ocular disorders in individuals suffering from diabetes. Around 35% of the global diabetic population is affected by DR, which is a leading cause of blindness in the workforce-aged demographic, despite the implementation of screening and treatments (Yau et al. 2012; Bunce et al. 2010). If left untreated, half of all patients diagnosed with proliferative diabetic retinopathy (PDR) are projected to become blind within five years post-diagnosis (Williams et al. 2004). As the global population ages and diabetes becomes more prevalent, it is anticipated that instances of DR will rise.

In diabetic patients, diabetic retinopathy is a common microvascular complication and is associated with an increased risk of life-threatening systemic vascular complications (Su et al. 2023). Hyperglycemia is known to lead to changes in the retina, causing vascular permeability, inflammation, fluid leakage and ischemia, and that certain growth factors may play a key role in diabetic microvascular complications (Pathak et al. 2012). Hyperglycemia and dyslipidemia are thought to disturb the homeostasis of the retina by inducing inflammatory responses in retinal tissue, including oxidative stress (Heesterbeek 2012). Epidemiologic studies have suggested multiple risk factors associated with the development and progression of diabetic retinopathy from longitudinal studies, including higher blood glucose, higher blood pressure, nonsmoking, male gender, higher HbA1c, longer duration of diabetes, lower BMI, and higher blood urea concentration (Meng et al. 2018). Moreover, many risk factors for DR have been reported recently, including hypertension, hyperlipidemia, poor blood glucose control, and albuminuria, although the underlying genetic mechanisms responsible for DR have not been identified (Looker et al. 2003; Leske et al. 2005; Cikamatana et al. 2007; Singh et al. 2008). Tentative risk factors included poor glycaemic control and a longer duration of diabetes (Stephen et al. 2018).

Studies have shown that patients with 20 years of diabetes, almost all type 1 diabetes (T1D) patients and 58% of type 2 diabetes (T2D) patients have signs of diabetic retinopathy (Hietala et al. 2010). Both twin studies and family studies have substantiated that the susceptibility to DR is genetically inherited in patients with type 1 and type 2 diabetes (Monti et al. 2007; Cho

et al. 2014). In individuals, the retinal disorder progresses from non-proliferative diabetes to vision-threatening proliferative diabetes, which is characterized by the growth of abnormal new blood vessels in the retina (Liu et al. 2019). The new blood vessels and the ensuing contraction of fibrous tissue can deform the retina, leading to traction retinal detachment and severe, often irreversible, vision loss (Liu et al. 2019). Notably, the genetic influence is more pronounced in the severe forms of DR, with the heritability factor rising from 18% in non-PDR stages as sibling models, to a substantial 50% in PDR cases (Looker et al. 2012; Hallman et al. 2005; Hietala et al. 2008). Studies have shown that proliferative DR has a 52% heritability rate in families, suggesting that family genetic factors are also associated with DR. (Looker et al. 2007; Arar et al. 2008).

Lately, various genome-wide association studies (GWASs) have suggested numerous potential genetic loci linked to DR, yet none of them have attained genome-wide significance or been confirmed by replication studies (Fu et al. 2010; Grassi et al. 2011; Awata et al. 2014; Burdon et al. 2015; Meng et al. 2018; Pollack et al. 2019). Furthermore, none of the linkage analyses, candidate gene association studies, and genome-wide association studies have identified DR risk loci that can be consistently reproduced. Insufficient grasp of DR's genetic aspects impedes the discovery of new biological routes for intervention, consequently escalating the healthcare expenses related to DR.

To better understand the genetic mechanisms related to diabetic retinopathy with type 2 mellitus, we carried out multiple GWASs using the UK Biobank (UKB) cohort. So far, this is a novel GWAS study conducted on the UKB cohort specifically for diabetic retinopathy. This project is designed to explore the genetic loci associated with DR based on the UK Biobank cohort and will give new ideas to those researchers studying DR.

Materials and Methods

Information of Cohorts and Patients

The UK Biobank serves as a large-scale biomedical database and research resource, offering comprehensive genetic and phenotypic data on roughly half a million UK participants aged between 40 and 69 years across England, Scotland, and Wales. For additional details regarding the UK Biobank cohort, refer to the website www.ukbiobank.ac.uk. This study was conducted

in accordance with the Tenets of the Declaration of Helsinki. Ethical approval for this research was appropriately secured from the National Health Service National Research Ethics Service (reference 11/NW/0382). Current studies were conducted based on UKB cohorts.

In this research, the gene data came from the UK Biobank cohort. They used a standardized process for DNA extraction and quality control (QC), in which detailed methods can be found at https://biobank.ctsu.ox.ac.uk/crystal/ukb/docs/genotyping_sample_workflow.pdf.

Type 2 diabetic patients were identified as individuals who responded with “No” to a specific query about “Started insulin within one year diagnosis of diabetes” in UK Biobank questionnaire (Refer to UK Biobank Questionnaire field ID: 2986). That is, individuals classified as type 2 diabetic patients are those who reported not commencing insulin therapy within a year of diagnosis. Patients with DR are diagnosed and collective according to the ICD 10 - WHO International Classification of Diseases (Refer to UK Biobank Questionnaire field ID: 41202). Moreover, we focused our study on participants of White British (Refer to UK Biobank Questionnaire field ID: 21000).

Upon extracting data from the UKB cohort, the numbers and clinical characteristics of cases and controls would be shown in Table S1 with significant GWASs. The case-control definition compared any DR to no DR. Cases were identified by participants who had T2D as well as DR while control subjects were distinguished by participants who had T2D but not DR.

During the replication phase, it is challenging to obtain summary statistics of DR patients with type 2 mellitus, as most provide only small datasets of the top loci (e.g., $p < 1 \times 10^{-5}$) and limited researchs. We attempted to expand our collection to different datasets for DR research which may not match exactly in definition but favored exploring the replication of loci. we matched the significant loci of four GWAS studies on diabetic retinopathy of mainly type 2 patients with our results. The studies are using the following cohorts: 1. GoDARTS (evolved from the Diabetes Audit and Research in Tayside Scotland setup in 1996 jointly by the University of Dundee etc.) and GoSHARE (GODRTS recruitment began in 1998 intending to obtain genotypic information on individuals with diabetes) (Rajendrakumar unpublished). 2. Caucasian Australians (funded by the National Health and Medical Research Council of Australia (GNT 595918) (Graham et al. 2018). 3. FinnGen cohorts (based on a public-private partnership between Finnish universities, biobanks, hospital districts, and several international pharmaceutical companies)

(FinnGen. GWAS summary results: diabetic retinopathy). 4. Chinese cohorts (from Taiwan-US Diabetic Retinopathy study) (Sheu et al. 2013). More information about these cohorts and GWASs was shown in Table 1.

Table 1. Information of four GWASs (five cohorts) in replication phase.

Cohorts	Studied Trait	Background Disease	Definition of Cases and Controls	Discovery Sample Size (N)		
				Cases	Controls	Total
GoDARTS	DR	Primarily Type 2 Mellitus, Few Type 1 Mellitus	Case: samples with any DR in T2D or T1D Control: samples without DR in T2D or T1D	1,791	3,222	5,013
GoSHARE	DR	Primarily Type 2 Mellitus, Few Type 1 Mellitus	Case: samples with any DR in T2D or T1D Control: samples without DR in T2D or T1D	527	1,227	1,754
Caucasian Australians	PDR	Type 2 Mellitus	Case: samples with PDR in T2D Control: samples without DR in T2D	176	435	518
FinnGen	DR	Type 1 & 2 Mellitus	Case: samples with DR Control: samples without DR	3,646	203,018	206,664
Chinese	PDR	Type 2 Mellitus	Case: samples with PDR in T2D Control: samples without DR in T2D	437	570	1,007

Definitions of multiple GWASs

When performing primary GWAS, the tool we used is genome-wide association methods based on generalized linear mixed model association tool (fastGWA-GLMM) (Yang et al. 2021), which in model covariates were age, sex, bmi, duration of diabetes (DoD), and top 8 genetic principal components. Based on the model in primary GWAS, there were two further GWASs generated from models adding other important covariates HbA1c and urea respectively as they have a high likelihood of impact in DR according to epidemiological studies (Roychoudhury et al. 2021; Meng et al. 2018). In most general sense, "covariates" typically refer to X variables that are included in a model to refine the precision of estimated treatment effects. In observational study designs, covariates might be added to a model to 1) to enhance predictive power, 2) to investigate specific conditional effects, or 3) to control for confounding (Walker 2018). Overall, the inclusion of covariates HbA1c and Urea respectively aims to increase the statistical power and validity of the model, ensuring the reliability of the research findings.

Considering the impact of data distribution in specific covariates, there are 8 secondary GWASs generated from different datasets. As we are interested in the datasets that have high age, HbA1c, Urea, and prolonged DoD which may help us to find out significant genetic loci with DR, datasets with age, DoD, HbA1c and Urea larger than corresponding Median (Q2) & Upper quartile (Q3) are mined out to perform GWAS respectively since the magnitude of these covariates is not clearly delineated in terms of the stages. Totally 11 different GWASs (3 GWASs and 8 Secondary GWASs) designed were shown in Table 2. The first row of the table represents three basic GWASs, where GWAS(+HbA1c) and GWAS(+Urea) indicate the addition of

covariates HbA1c and Urea, respectively, to the linear model of the Primary GWAS. Second row shows secondary GWASs studies conducted using sub-datasets based on the model of the Primary GWASs. For example, in DoD>Q2, it means that the chosen sub-dataset is where the variable (duration of diabetes) is greater than its own median, based on the model used in the Primary GWAS.

Table 2. *Upper part* shows 3 GWASs using different models in fastGWA-GLMM. *Lower part* totally shows 8 secondary GWASs based on different dataset requirements in each grey part.

GWASs	GWAS(+HbA1c)	Primary GWAS	GWAS(+Urea)
Secondary GWASs	HbA1c > Q2 HbA1c > Q3	Age > Q2, Age > Q3 DoD > Q2, DoD > Q3	Urea > Q2 Urea > Q3

Statistical analysis

Genome-Wide Complex Trait Analysis (GCTA v1.93.3), a package used in this research under Linux, is designed to estimate the proportion of phenotypic variation explained by all genome-wide SNPs for complex traits (refer to <https://yanglab.westlake.edu.cn/software/gcta/#Overview>). Our study performed fastGWA-GLMM for binary traits in UK Biobank data scale. In QC step, SNPs with imputation INFO scores < 0.3, minor allele frequency < 0.5% were removed by using GCTA, as well as SNPs that failed Hardy-Weinberg < 0.001 and samples with missing call rates < 0.1 would be filter out by using Plink v1.90. SNPs situated on the X and Y chromosomes, as well as mitochondrial SNPs, were eliminated from consideration. Gender difference between cases and controls was compared using chi-square testing, whereas age and BMI were compared by conducting independent t-testing with R v4.0.3. SNPs were considered reaching genome-wide association significance if they had a p value less than 5×10^{-8} . Moreover, heritability was calculated by utilizing “BLD-LDAK Model” in SumHer (Speed & Balding 2018; Speed et al. 2020).

GWAS analysis by FUMA and LocusZoom

GRCh37 genome assembly was followed for SNP annotation. SNP functional annotations, Manhattan plot and Q-Q plot were applied and produced by the FUMA v1.5.6, a web platform for annotating, prioritizing, visualizing and interpreting GWAS results (Watanabe et al. 2017).

Regional visualization was provided by LocusZoom (<http://locuszoom.org/>). The gene analysis, gene-set analysis and tissue expression analysis were performed with MAGMA v1.08, which was integrated in FUMA. In the gene analysis, the summary statistics of SNPs are consolidated to the level of entire genes, evaluating the combined association of all SNPs in that gene with the phenotype. Gene-set analysis groups individual genes based on shared biological, functional, or other traits. This helps in understanding the role of particular biological pathways or cell functions in the genetic origins of a phenotype. To test the relationship between highly expressed genes in a specific tissue and genetic associations with DR, tissue expression analysis, based on GTEx (<https://www.gtexportal.org/home/>), was performed separately for 30 general tissue types and 53 specific tissue types.

Results

Primary GWAS results

Through data extraction from questionnaire and quality control on gene data, we identified 1,824 cases and 15,164 controls. Clinical characteristics of the case and control groups were compiled in Table S1, Sheet 1. Age, sex, body mass index (BMI), duration of diabetes, HbA1c and Urea were all found to be significantly different ($p < 0.005$) between cases and controls. There was a cluster appearing of top SNPs in the Manhattan plot, where a locus reaches significant level (Fig. 1). Corresponding Q-Q plot is shown in Figure S1. The top SNP in this region was rs6066146 in the *EYA2* gene with a p value of 4.21×10^{-8} in Chromosome 20. Following suggestive SNPs are rs117659116 in chromosome 12 with a p value of 5.77×10^{-8} in gene *MREGP1*, rs17277029 in chromosome 4 with a p value of 6.01×10^{-7} in gene *FBXW7* and rs112773829 in chromosome 2 with a p value of 8.95×10^{-7} in gene *RNU6-546P*. Altogether 11,169,041 imputed SNPs passed from routine quality control checking and imputation INFO quality score larger than 0.3. The heritability of DR was estimated to be 26.73% in this type 2 diabetic population based on the combination of the Baseline LD model (Finucane et al. 2015; Gazal et al. 2019) and the LDAK model (Speed et al. 2020) by SumHer.

Gene, gene-set and tissue expression analysis by FUMA

The parameters in the FUMA platform are set as follows. N = 16988, exMHC = 1, ensembl =

v102, genotype = protein_coding, leadP = 5e-8, gwasP = 0.05, r2 = 0.6, 2ndr2 = 0.1, refpanel = UKB/release2b, pop = WBrits_10k, MAF = 0, mergeDist = 250, magma = 1, magma_window = 42, magma_exp = GTEEx/v8/gtex_v8_ts_avg_log2TPM, GTEEx/v8/gtex_v8_ts_general_avg_log2TPM, posMap = 1, posMapWindowSize = 10.

In gene analysis, all the SNPs that are located within genes were mapped to 19,238 protein-coding genes. No gene demonstrated the strongest association, see Figure S2.

In the gene-set analysis, a total of 10,678 gene sets were tested. There is no gene set achieve statistically significant association ($p < 0.05/10,678 = 4.68 \times 10^{-6}$). The top first gene set are GOBP_NEGATIVE_REGULATION_OF_CELL_JUNCTION_ASSEMBLY demonstrated p value of 1.39×10^{-5} . The top ten gene sets from this analysis are shown in Table S2.

In the tissue expression analysis, tissue "Spleen" demonstrated statistically significant associations in the expression analysis of 30 general tissue types from multiple organs and the 53 specific tissue types within some of these organs. See Fig. 2.

Additional GWASs results

All GWASs which had significant results are concluded in Table 3. Fig. 3 shows circular Manhattan plot for GWAS(+HbA1c), GWAS(+Urea), primary GWAS(DoD>Q2), GWAS(HbA1c>Q3). Gene *EYA2* was shown to be a significant gene in several experiments. A Marker of 1: 107,782,413 with the nearest gene *NTNG1* is also shown to be a significant gene in several GWASs while there is a SNP cluster in both gene *CTAGE14P* and *MPDZ* respectively. All regional plots for 4 loci are shown in Figure S3, while Figure S4 shows the Q-Q plot of GWAS(+HbA1c), GWAS(+Urea), primary GWAS(DoD>Q2) and GWAS(HbA1c>Q3) respectively.

Replication Stage

In the GoDARTS and GoSHARE cohorts, rs6066146 was not significantly replicated, while rs17273542 had a p-value consistently less than 0.1. In the Caucasian Australians cohort, no loci were significantly replicated. In the FinnGen cohort's study on VII Diseases of the eye and adnexa (H7_), both rs6066146 and rs71183224 had p-values less than 0.01, indicating suggestive replication. In Chinese cohorts, rs6066146 showed p value less than 0.1. Detailed

p values of loci in these studies are shown in the Table S3.

Table 3. Significant and suggestive loci of all GWASs using UK Biobank

GWASs	Marker	Chromosome	Position	Nearest gene	p value
Primary GWAS	rs6066146	20	45,599,553	<i>EYA2</i>	4.21 x10 ⁻⁸ *
	rs117659116	12	31,418,552	<i>MREGP1</i>	5.77 x10 ⁻⁸
	1: 107,782,413	1	107,782,413	<i>NTNG1</i>	2.22 x10 ⁻⁷
	rs17277029	4	153,428,649	<i>FBXW7</i>	6.01 x10 ⁻⁷
	rs112773829	2	156,268,365	<i>RNU6-546P</i>	8.95 x10 ⁻⁷
GWAS (+HbA1c)	rs117659116	12	31,418,552	<i>MREGP1</i>	1.75 x10 ⁻⁸ *
	1: 107,782,413	1	107,782,413	<i>NTNG1</i>	1.96 x10 ⁻⁸ *
	rs6066146	20	45,599,553	<i>EYA2</i>	3.77 x10 ⁻⁸ *
	rs72721645	4	153,494,209	<i>AC023424.2</i>	3.78 x10 ⁻⁷
	rs552254054	16	549,951	<i>RAB11FIP3</i>	8.99 x10 ⁻⁷
	11: 9,199,268	11	9,199,268	<i>DENND5A</i>	9.66 x10 ⁻⁷
GWAS (+Urea)	1: 107,782,413	1	107,782,413	<i>NTNG1</i>	4.08 x10 ⁻⁸ *
	rs71183224	20	45,603,386	<i>EYA2</i>	5.96 x10 ⁻⁸
	rs6435333	2	155,560,333	<i>KCNJ3</i>	4.04 x10 ⁻⁷
	rs34759185	1	217,293,468	<i>ESRRG</i>	6.28 x10 ⁻⁷
	rs552254054	16	549,951	<i>RAB11FIP3</i>	7.17 x10 ⁻⁷
	rs142927559	6	97,386,148	<i>KLHL32</i>	7.21 x10 ⁻⁷
	rs137934083	22	49,479,088	<i>AL078622.1</i>	8.02 x10 ⁻⁷
Primary GWAS (DoD > Q2)	rs13013265	2	168,556,935	<i>CTAGE14P</i>	2.69 x10 ⁻⁸ *
	rs79202341	9	115,618,322	<i>SNX30</i>	5.18 x10 ⁻⁷
	rs147264322	2	78,314,433	<i>AC012494.1</i>	8.09 x10 ⁻⁷
	rs190386742	2	170,057,766	<i>LRP2</i>	9.08 x10 ⁻⁷
GWAS (HbA1c > Q3)	rs62532872	9	13,226,669	<i>MPDZ</i>	6.90 x10 ⁻⁹ *
	rs17273542	9	13,224,491	<i>MPDZ</i>	2.32 x10 ⁻⁸ *
	rs186759284	2	116,694,780	<i>AC016721.1</i>	1.19 x10 ⁻⁷
	rs550645235	4	26,779,514	<i>TBC1D19</i>	1.63 x10 ⁻⁷
	rs147617022	5	164,192,487	<i>AC109466.1</i>	2.89 x10 ⁻⁷
	17: 57,849,382	17	57,849,382	<i>VMP1</i>	3.94 x10 ⁻⁷
	rs145145808	2	116,905,368	<i>AC016721.1</i>	4.23 x10 ⁻⁷
	rs189997796	14	58,510,393	<i>ARMH4</i>	4.70 x10 ⁻⁷
	rs117849333	18	70,439,751	<i>NETO1</i>	5.51 x10 ⁻⁷
	rs77360543	15	76,160,587	<i>UBE2Q2</i>	5.62 x10 ⁻⁷
	rs548908893	2	69,970,492	<i>ANXA4</i>	7.56 x10 ⁻⁷

* indicates significant loci reach significant p value 5x10⁻⁸ , while others are suggestive loci.

Discussion

In these GWAS studies of type 2 diabetic retinopathy using the UKB dataset, our research group identified noteworthy genetic variants at five different loci. These findings highlight the complex genetic basis associated with DR. Notably, in this study we used the UKB's generic questionnaire as a valuable screening tool, providing researchers with the advantage of mitigating the potential problem of reduced study power due to heterogeneity. Results showed that in primary GWAS on diabetic retinopathy, we identified variants in *EYA2* have reached genome-wide significance ($p < 5 \times 10^{-8}$). In addition, other GWASs further revealed that variants in or near gene *MREGP1*, *MPDZ*, *NTNG1* and *CTAGE14P* were identified to be involved in the

development of DR.

Leading locus rs6066146 in gene *EYA2* in chromosome 20 for DR was identified from primary GWAS. The *EYA2* gene, which stands for "Eyes Absent Homolog 2", encodes a member of the eyes absent family of proteins, which encodes proteins that may be post-translationally modified and may play a role in eye development (Wang et al. 2016). *EYA2* is involved in DNA repair and has, besides type 2 diabetes, also been reported to associate with triglyceride levels and waist-hip ratio (Jonsson et al. 2021). Zhang et al. had found that *EYA2* is also expressed in other sensory and developmental systems (eye) that have not been reported to date. The research showed that in the retina, X-gal staining of *EYA2^{lacZ}* at E12.5 and E14.5 showed β -Gal activity in *EYA2^{lacZ/+}* and *EYA2^{lacZ/lacZ}*, as well as β -Gal activity was also observed in the sclera, extraocular muscles, and nerve fibers within the optic stalk, which will develop into optic nerve (Zhang et al. 2021). Studies above supported a possible role of *EYA2* in retinal neuronal differentiation. To expand on that, the four human *EYA* paralogs (*EYA1-4*) are initially elucidated within the framework of *Drosophila* eye development as the retinal determination gene network (Bonini et al. 1993). Research findings indicate that the pharmacological suppression of *EYA* protein tyrosine phosphatase (PTP) activity leads to a slowdown in vascular growth within the postnatal mouse retina (Roychoudhury et al. 2021). An experiment in a mouse model of oxygen-induced retinopathy demonstrated that *EYA* inhibitors reduced pathological neovascularisation in an in vivo model for proliferative retinopathy (Hegde & OH 2017). Observations above sheds light on the potential impact of *EYA2* on human retinal vasculature. Additionally, the *EYA* tyrosine phosphatase activity plays a crucial role in deciding between repair and apoptosis following DNA double-stranded breaks, which activity might promote hypoxia-induced angiogenesis by dephosphorylating H2AX, which implies that *EYA* is likely involved in both developmental and disease-related retinal angiogenesis (Wang et al. 2016). In summary, many studies showcase the pro-angiogenic nature of *EYA* tyrosine phosphatase activity (Roychoudhury et al. 2021; Wang et al. 2016; Tadjuidje et al. 2012). Both Benzbromarone and Benzarone emerge as promising contenders for repurposing as potential drugs to address cancer metastasis, tumor-driven angiogenesis, and vascular disorders (Tadjuidje et al. 2012).

Tissue expression analysis facilitated by FUMA indicated an association between 'Spleen' and gene *EYA2*, where *EYA PTP* activity had been identified (Pandey et al. 2013; Tadjuidje et al. 2012), which is associated with angiogenesis (Wang et al. 2018) and proliferative retinopathy (Wang et al. 2016) and have been validated in animal disease models. Recent studies had elaborated that microglial spleen tyrosine kinase might serve as a potential therapeutic target for diabetic retinopathy (Liu et al. 2021; Su et al. 2020), which means that some physiological change in the spleen to some degree might contribute to the development of DR. Another experimental study showed that in diabetic patients, circulating pro-inflammatory monocytes of bone marrow origin were increased, whereas reparative circulating angiogenic cells were trapped in the bone marrow and spleen and had an impaired ability to be released into the circulatory system, leading to inadequate retinal vascular repair, that is, developing diabetic retinopathy (Chakravarthy et al. 2016). Moreover, in the type 2 diabetic individuals, the heritability of DR was calculated to be 26.73%. It is noteworthy that this heritability estimation falls within the range of heritability reported by most studies (Simó-Servat et al. 2013; Looker et al. 2012; Hietala et al. 2008; Hallman et al. 2005). This assessment does not encompass the contribution of gene-environment interactions, etc., implying that the genuine heritability of this trait might be higher. We have posited that DR exhibits hereditary characteristics in this GWAS, underscoring the need for additional genetic investigations.

Significant loci found in other GWASs were as follows. Variant rs117659116 in gene *MREGP1* in chromosome 12 and variants rs62532872, rs17273542 in *MPDZ* gene in chromosome 9. *MREGP1* have been pointed out twice in the GWAS studies, but there were no others supporting except that the gene was found out associated with type 2 diabetes mellitus, with rs117659116 could not found any information about DR. In addition, we have found two significant loci, which are the Marker 1: 107,782,413 in chromosome 1 near gene *NTNG1*, and the locus rs13013265 near gene *CTAGE14P* in chromosome 2. Gene *NTNG1* (Netrin G1), identified by GWAS(+HbA1c) and GWAS(+Urea), was a protein-coding gene. A study shows that the expression of *NTNG1* was high in the retina and kidney in GTEx (<https://www.gtexportal.org/home>), and in the brain and kidney in HPA-RNA (Ke et al. 2023). Abnormal methylation-modified differentially expressed genes of *NTNG1* may be involved in

inflammatory response and vascular remodeling in the process of Coronary artery ectasia (Yang et al. 2023), which may likely influence the retinal vasculature. Gene *CTAGE14P* was identified from GWAS(DoD>Q2), which was done to explore whether the case and control group would show significant differences in a group of people who have had diabetes for more than 5 years. Though gene *CTAGE14P* has been found to be associated with visual perception in previous studies (Zhu et al. 2020), there is not enough evidence that the gene is associated with DR. DR patients were somewhat insensitive to age according to the secondary GWASs in age. For all secondary GWASs(+Urea), we did not find any significant loci to support the involvement of high urea in the development of DR.

The *MPDZ* gene was obtained from a study GWAS(HbA1c>Q3), which aimed to explore whether there was a significant difference between the case and control group in DR patients with high levels of HbA1c. *MPDZ* (Multiple PDZ Domain Crumbs Cell Polarity Complex Component) is a protein-coding gene. Diseases associated with *MPDZ* include hydrocephalus, congenital, 2, with or without brain or eye anomalies and congenital communicating hydrocephalus (HYC2 2023). From the platform GeneCards, *MPDZ* mRNA express in normal expression pattern in human tissues from Serial Analysis of Gene Expression. Evidence shows that *MPDZ* is directly associated with Crumbs homolog 1 (Crb1), in which loss of Crb1 function leads to either recessively inherited retinitis pigmentosa or Leber congenital amaurosis in humans. The identification of human mutations suggests that *MPDZ* plays a role in human retinal disease, but the precise nature of this role remains to be determined (Ali et al. 2011). It has also been demonstrated that *MPDZ* enhances Notch signaling activity and that angiogenesis is coordinated by *VEGF* and Notch signaling (Tetzlaff et al. 2018). Inactivation of the *MPDZ* gene leads to impaired Notch signaling activity and increased blood vessel sprouting in cellular models (Tetzlaff et al. 2018), which may lead to increased vascularity in the DR. Since a recent study identified that *MPDZ* plays an essential role in the occurrence and maintenance of the macula (Zhang et al. 2022), a genetic correlation analysis of DR and macula can be performed in future studies.

It is worth mentioning that a suggestive locus rs17277029 in chromosome 4 with a *p* value of

6.01×10^{-7} in gene *FBXW7* had been found in primary GWAS. *FBXW7* was reported to be associated with DR, which in detail, in mouse experiments, silencing *FBXW7* enhanced angiogenesis and overexpressing *FBXW7* depressed angiogenesis in retinal tissue (Hu et al. 2020). Another study showed that the abundant mobilization of DNA damage repair mediated by *FBXW7* leads to the downregulation of poly-ADP-ribose-polymerase(PARP) expression and activity in both human endothelial cells and diabetic rat retinas, where RARP plays important roles in endothelial cell impairment in the pathogenesis of diabetic vascular diseases (Li et al. 2022). The GeneMANIA (<https://genemania.org/>) shows that *FBXW7* and *EYA2*, *NTNG1* had directly genetic interactions, with *MPDZ* had indirectly genetic interaction in Fig. 4. The remaining suggestive loci from other studies are not described in further detail here, but they hold certain value. For instance, primary GWAS(DoD>Q2) have shown that gene *LRP2* has been identified in relation to DR (Sheu et al. 2013), and *KCNJ3* from GWAS(+Urea) has been associated with the retina (Pattnaik et al. 2012). Moreover, Fig. 4 also shows in the gene linkage network, *RIMS2* serves as the intersection of *EYA2*, *MPDZ* and *NTNG1*, showing a certain association between them, which had reported potentially associated with DR. Data mining of human adult bulk and single-cell retinal transcriptional datasets revealed predominant expression in rod photoreceptors, and immunostaining demonstrated *RIMS2* localization in the human retinal outer plexiform layer, purkinje cells, and pancreatic islets (Mechaussier et al. 2020). A study shows that circular RNA is a potential target to control diabetic proliferative retinopathy (Shan et al. 2017), while circular *Rims2* (*circRims2*) is highly expressed and conserved in both the human and mouse brains and the high-throughput RNA-seq analysis reveals a high expression of *circRims2* in the retina (Sun et al. 2021). Furthermore, the current study finds that *circRims2* deficiency evokes retinal inflammation and activates the tumor necrosis factor signaling pathway (Sun et al. 2021), which is similar to that caused by inactivation of gene *MPDZ* (Tetzlaff et al. 2018). Although this is an alternative approach to using the identified loci to search for possible sources of mutations, there is no intuitive evidence that the mutation in the source gene *RIMS2* causes the development of DR in our GWAS studies.

Our study, we defined T2D as the immediate use of insulin by the patient within one year of the

detection of diabetes, a definition that was not present in UKB and carries some risk. We used multiple methods to define DR cases and controls datasets, so that doing different GWAS will not only improve data and study quality, but also help to discover new mechanisms of DR at the molecular level. For example, significant loci in primary GWAS showed more significant expression in GWAS (+HbA1c), suggesting that the inclusion of the covariate HbA1c helps to identify loci associated with DR, demonstrating the importance of this metric in studying DR. Our study, while providing initial insights, is not highly interpretive in designing different GWAS studies. The addition of covariates to the fastGWA-GLMM model is somewhat empirical, and the use of quartiles for dataset partitioning is somewhat subjective. This partitioning was necessitated in order to mine more information out of the limited dataset. Currently, most studies conduct GWASs using existing pipelines without considering the diversity of variable inputs into the model itself and the quality of model fitting. It is hoped that in the future, statistical knowledge can be utilized to explain the rationality, significance, and limitations of the fitted models more comprehensively. Furthermore, while we have highlighted the p-values of significant loci in four cohorts during the replication phase, it is essential to recognize that the statistical power of these results is relatively limited. In GoDARTS and GoSHARE, due to the limited number of T1D cases in the samples used, the disease background differs slightly from our study, which is entirely focused on T2D. In the study of Caucasian Australians cohort and Chinese cohort, the subject is PDR, which also differs from our research subject, DR. In FinnGen, the definition of controls applies to all participants, not just those restricted to having T2D.

Overall, this study suggests a role for the *EYA2* gene in retinal angiogenesis and proliferative retinopathy, with potential involvement of the spleen in retinal blood vessel repair. The *MPDZ* gene, possibly influencing Notch signaling, may promote retinal vascular sprouting in diabetic patients, while *FBXW7*'s overexpression depresses angiogenesis in retinal tissues. While limited evidence exists for other loci and neighboring genes, they offer reference points for future DR research. During the replication phase, no significant loci were replicated due to 1) the limited number of GWAS studies on DR, 2) the challenge of finding consistent subject definitions in GWAS studies targeting DR, and 3) the difficulty of accessing summary statistics for DR as many datasets remained unpublished. The replication cohorts that most closely

matched the definition of our study were GoDARTS and GoSHARE, but direct replication was not achieved. This study highlights the *EYA2*, *MPDZ*, *NTNG1*, *CTAGE14P* and *MREGP1* genes, along with *FBXW7* and *LRP2*, for their higher association with DR in type 2 patients, helping to understand the mechanisms of DR and identify potential targets for DR treatment.

Statements and Declarations

Funding

This study was mainly funded by the Pioneer and Leading Goose R&D Program of Zhejiang Province 2023 with reference number 2023C04049 and Ningbo International Collaboration Program 2023 with reference number 2023H025.

Data availability

This research complies with all ethical standards and data privacy measures set by the UK Biobank. Detailed summary statistics data pertaining to Diabetic Retinopathy within the UK Biobank are retrievable from <https://figshare.com/s/9d55aa774fb19307f7c8>. Should there be any additional data pertinent to this study that are not presented within this paper or its additional files, the authors can provide such data upon reasonable request. Data is available upon publication.

Author information

Authors and Affiliations

**Nottingham Ningbo China Beacons of Excellence Research and Innovation Institute,
University of Nottingham Ningbo China, Ningbo, China, 315100**

Tengda Cai, Qi Pan, Yiwen Tao, Weihua Meng

**Division of Population Health and Genomics, School of Medicine, University of
Dundee, Dundee, UK, DD2 4BF**

Charvi Nangia, Colin Palmer

Duke University, 2080 Duke University Road, Durham, USA, NC 27708

Aravind Lathika Rajendrakumar

**School of Veterinary Medicine and Science, University of Nottingham, Nottingham,
UK, LE12 5RD**

Tania Dottorini

**School of Mathematical Sciences, University of Nottingham Ningbo China, Ningbo,
China, 315100**

Mainul Haque

Authors' Contributions

TC drafted the paper and performed GWAS analysis. QP, YT contributed to data formatting and correction. AR, CN, CP, TD, MH provided comments to the paper. WM organized the project and provided comments.

Acknowledgment

The authors gratefully thank all the participants and professionals contributing to the UK Biobank. This research has been conducted using the UK Biobank Resource (50604).

Corresponding authors

Correspondence to Weihua Meng

Consent to Publish

All authors have consent for publication.

Ethics approval

This study was approved by the Ethics Committee of the University of Nottingham, Ningbo, China.

References

- Ali M, Hocking PM, McKibbin M, Finnegan S, Shires M, Poulter JA ... Inglehearn CF (2011): mpdznul allele in an avian model of retinal degeneration and mutations in human leber congenital amaurosis and retinitis pigmentosa; *Investigative Ophthalmology & amp. Visual Science* 52: 7432.
- Arar NH, Freedman BI, Adler SG, Iyengar SK, Chew EY, Davis MD ... Abboud HE(2008): Heritability of the severity of diabetic retinopathy: the FIND-Eye study; *Investigative ophthalmology & amp. Visual Science* 49: 3839–3845.
- Awata T, Yamashita H, Kurihara S, Morita-Ohkubo T, Miyashita Y, Katayama S ... Koike A (2014): A genome-wide association study for diabetic retinopathy in a Japanese population: Potential Association with a long intergenic non-coding RNA. *PLoS ONE* 9.
- Bain SC, Klufas MA, Ho A & Matthews DR (2018): Worsening of diabetic retinopathy with rapid improvement in Systemic Glucose Control: A Review; *Diabetes. Obesity and Metabolism* 21: 454–466.
- Bonini NM, Leiserson WM & Benzer S (1993): The eyes absent gene: Genetic control of cell survival and differentiation in the developing *Drosophila* Eye. *Cell* 72 : 379–395.
- Bunce C, Xing W & Wormald R (2010): Causes of blind and partial sight certifications in England and Wales: April 2007–March 2008. *Eye* 24 : 1692–1699.
- Burdon KP, Fogarty RD, Shen W, Abhary S, Kaidonis G, Appukuttan B ... Craig JE (2015): Genome-wide association study for sight-threatening diabetic retinopathy reveals association with genetic variation near the *grb2* gene. *Diabetologia* 58 : 2288–2297.
- Chakravarthy H, Beli E, Navitskaya S, O'Reilly S, Wang Q, Kady N ... Busik JV (2016): Imbalances in mobilization and activation of pro-inflammatory and vascular reparative bone marrow-derived cells in diabetic retinopathy. *PLOS ONE* 11.
- Cho H & Sobrin L (2014): Genetics of diabetic retinopathy. *Current Diabetes Reports* 14.
- Cikamatana L, Mitchell P, Rochtchina E, Foran S, Wang JJ (2007): Five-year incidence and progression of diabetic retinopathy in a defined older population: The Blue Mountains Eye Study. *Eye* 21 : 465–471.
- FinnGen. GWAS summary results: diabetic retinopathy. (2020) https://r5.finnngen.fi/pheno/H7_RETINOPATHYDIAB (accessed November 03, 2023).
- Finucane HK, Bulik-Sullivan B, Gusev A, Trynka G, Reshef Y, Loh P ... Price A (2015): Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nature Genetics* 47 : 1228–1235.
- Fu YP, Hallman DM, Gonzalez VH, Klein BE, Klein R, Hayes MG ... Hanis CL(2010): Identification of diabetic retinopathy genes through a genome-wide association study among Mexican-americans from Starr County, Texas. *Journal of Ophthalmology* 2010 : 1–9.
- Gazal S, Finucane HK, Furlotte NA, Loh P, Palamara PF, Liu X ... Price AL (2019): Author correction: Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nature Genetics* 51 : 1295–1295.
- Graham PS, Kaidonis G, Abhary S, Gillies MC, Daniell M, Essex RW ... Burdon KP (2018): Genome-wide association studies for diabetic macular edema and proliferative diabetic retinopathy. *BMC Medical Genetics* 19.
- Grassi MA, Tikhomirov A, Ramalingam S, Below JE, Cox NJ & Nicolae DL (2011): Genome-wide meta-analysis for severe diabetic retinopathy. *Human Molecular Genetics* 20 : 2472–2481.
- Hallman DM, Huber JC, Gonzalez VH, Klein BE, Klein R & Hanis CL (2005): Familial aggregation of severity of diabetic retinopathy in Mexican Americans from Starr County, Texas. *Diabetes Care* 28 : 1163–1168.
- Haller JA & Decroos FC (2012): Risk factors for age-related macular degeneration and choroidal neovascularization. *Age-Related Macular Degeneration* 68–79.
- Hegde R & OH C (2017): US9725430B2 - Use of small molecule inhibitors targeting EYA tyrosine phosphatase. Google Patents. <https://patents.google.com/patent/US9725430B2/en> (accessed September 20, 2023).
- Hietala K, Forsblom C, Summanen P & Groop PH (2008): Heritability of proliferative diabetic retinopathy. *Diabetes* 57 : 2176–2180.
- Hietala K, Harjutsalo V, Forsblom C, summanen P & Groop P (2010): Age at onset and the risk of proliferative retinopathy in type 1 diabetes. *Diabetes Care* 33 : 1315–1319.
- Hu L, Lv X, Li D, Zhang W, Ran G, Li Q & Hu J (2020): The anti-angiogenesis role of FBXW7

in diabetic retinopathy by facilitating the ubiquitination degradation of c-myc to orchestrate the HDAC2. *Journal of Cellular and Molecular Medicine* 25 : 2190–2202.

Hydrocephalus, Congenital, 2, with or Without Brain or Eye Anomalies (HYC2)- malacards. (2023)https://www.malacards.org/card/hydrocephalus_congenital_2_with_or_without_brain_or_eye_anomalies (accessed August 31, 2023).

Jonsson A, Stinson SE, Torekov SS, Clausen TD, Færch K, Kelstrup L & Hansen T (2021): Genome-wide association study of circulating levels of glucagon during an oral glucose tolerance test. *BMC Medical Genomics* 14.

Ke S, Guo J, Wang Q, Shao H, He M, Li T ... Guo J (2023): Netrin family genes as prognostic markers and therapeutic targets for clear cell renal cell carcinoma: Netrin-4 acts through the Wnt/ β -catenin signaling pathway. *Cancers* 15 : 2816.

Leske M, Wu S, Hennis A, Hyman L, Nemesure B, Yang L & Schachat A (2005): Hyperglycemia, blood pressure, and the 9-year incidence of Diabetic Retinopathythe Barbados Eye Studies. *Ophthalmology* 112 : 799–805.

Li S, Deng J, Sun D, Chen S, Yao X, Wang N ... Wei F (2022): FBXW7 alleviates hyperglycemia-induced endothelial oxidative stress injury via ROS and PARP inhibition. *Redox Biology* 58 : 102530.

Liu C, Chen G, Bentley AR, Doumatey A, Zhou J, Adeyemo A ... Rotimi C (2019): Genome-wide association study for proliferative diabetic retinopathy in Africans. *npj Genomic Medicine* 4.

Liu X, Xu B & Gao S (2021): Spleen tyrosine kinase mediates microglial activation in mice with diabetic retinopathy; *Translational Vision Science & Technology* 10 : 20.

Looker HC, Krakoff J, Knowler WC, Bennett PH, Klein R & Hanson RL (2003): Longitudinal studies of incidence and progression of diabetic retinopathy assessed by retinal photography in Pima Indians. *Diabetes Care* 26 : 320–326.

Looker HC, Nelson RG, Chew E, Klein R, Klein BE, Knowler WC & Hanson RL (2007): Genome-wide linkage analyses to identify loci for diabetic retinopathy. *Diabetes* 56 : 1160–1166.

Looker HC, Nyangoma SO, Cromie D, Olson JA, Leese GP, Black M ... Colhoun HM (2012): Diabetic retinopathy at diagnosis of type 2 diabetes in Scotland. *Diabetologia* 55 : 2335–2342.

Mechaussier S, Almoallem B, Zeitz C, Van Schil K, Jeddawi L, Van Dorpe J ... Perrault I (2020): Loss of function of rims2 causes a syndromic congenital cone-rod synaptic disease with neurodevelopmental and pancreatic involvement. *The American Journal of Human Genetics* 106 : 859–871.

Meng W, Shah KP, Pollack S, Toppila I, Hebert HL, McCarthy MI ... Palmer CN (2018): A genome-wide association study suggests new evidence for an association of the NADPH oxidase 4 (nox4) gene with severe diabetic retinopathy in type 2 diabetes. *Acta Ophthalmologica* 96.

Monti MC, Lonsdale JT, Montomoli C, Montross R, Schlag E & Greenberg DA (2007): Familial risk factors for microvascular complications and differential male-female risk in a large cohort of American families with type 1 diabetes; *The Journal of Clinical Endocrinology & Metabolism* 92 : 4650–4655.

Pandey RN, Wang TS, Tadjuidje E, McDonald MG, Rettie AE & Hegde RS (2013): Structure-activity relationships of Benzbromarone metabolites and derivatives as EYA inhibitory anti-angiogenic agents. *PLoS ONE* 8.

Pathak D, Gupta A, Kamble B, Kuppusamy G & Suresh B (2012): Oral targeting of protein kinase C receptor: Promising route for diabetic retinopathy? *Current Drug Delivery* 9 : 405–413.

Pattnaik BR, Asuma MP, Spott R & Pillers DM (2012): Genetic defects in the hotspot of inwardly rectifying K⁺ (KIR) channels and their metabolic consequences: A Review. *Molecular Genetics and Metabolism* 105 : 64–72.

Pollack S, Jr RP, Jensen RA, Christiansen M, Li X, Cheng C ... Sobrin L (2018): Multiethnic genome-wide association study of diabetic retinopathy using liability threshold modeling of duration of diabetes and glycemic control. *Diabetes* 68 : 441–456.

Roychoudhury K & Hegde RS (2021): The eyes absent proteins: Unusual had family tyrosine phosphatases. *International Journal of Molecular Sciences* 22 : 3925.

Sapieha P, Joyal J, Rivera JC, Kermorvant-Duchemin E, Sennlaub F, Hardy P ... Chemtob S (2010): Retinopathy of prematurity: Understanding ischemic retinal vasculopathies at an extreme of life. *Journal of Clinical Investigation* 120 : 3022–3032.

Shan K, Liu C, Liu B, Chen X, Dong R, Liu X ... Yan B (2017): Circular noncoding RNA HIPK3 mediates retinal vascular dysfunction in diabetes mellitus. *Circulation* 136 : 1629–1642.

Sheu WH, Kuo JZ, Lee I, Hung Y, Lee W, Tsai H ... Chen YI (2013): Genome-wide association study in a Chinese population with diabetic retinopathy. *Human Molecular Genetics* 22 : 3165–3173.

Simó-Servat O, Hernández C & Simó R (2013): Genetics in diabetic retinopathy: Current concepts and new insights. *Current Genomics* 14 : 289–299.

Gupta A, Singh R, Ramasamy K, Abraham C & Gupta V (2008): Diabetic retinopathy: An update, *Indian journal of ophthalmology* 56 : 179.

Speed D & Balding DJ (2018): Sumher better estimates the SNP heritability of complex traits from summary statistics. *Nature Genetics* 51 : 277–284.

Speed D, Holmes J & Balding DJ (2020): Evaluating and improving heritability models using summary statistics. *Nature Genetics* 52 : 458–462.

Su X, Sun Z, Ren Q, Liu J, Yin L, Liang N ... Sun R (2020): The effect of spleen tyrosine kinase inhibitor R406 on diabetic retinopathy in experimental diabetic rats. *International Ophthalmology* 40 : 2371–2383.

Su Z, Wu Z, Liang X, Xie M, Xie J, Li H ... Jiang F (2023): Diabetic retinopathy risk in patients with unhealthy lifestyle: A Mendelian randomization study. *Frontiers in Endocrinology* 13.

Sun L, Ma Y, Ji Y, Wu Z, Wang Y, Mou H & Jin Z (2021): Circular rims2 deficiency causes retinal degeneration. *Advanced Biology* 5.

Tadjuidje E, Wang TS, Pandey RN, Sumanas S, Lang RA & Hegde RS (2012): The EYA tyrosine phosphatase activity is pro-angiogenic and is inhibited by Benzbromarone. *PLoS ONE* 7.

Tetzlaff F, Adam MG, Feldner A, Moll I, Menuchin A, Rodriguez-Vita J ... Fischer A (2018): Author response: MPDZ promotes DLL4-induced notch signaling during angiogenesis [Preprint].

Walker JA (2022): Applied Statistics for Experimental Biology. https://www.middleprofessor.com/files/applied-biostatistics_bookdown/_book/ (accessed November 13, 2023).

Wang Y, Tadjuidje E, Pandey RN, Stefater JA, Smith LE, Lang RA & Hegde RS (2016): The eyes absent proteins in developmental and pathological angiogenesis. *The American Journal of Pathology* 186 : 568–578.

Wang Y, Pandey RN, Riffle S, Chintala H, Wikenheiser-Brokamp KA & Hegde RS (2018): The protein tyrosine phosphatase activity of eyes absent contributes to tumor angiogenesis and tumor growth. *Molecular Cancer Therapeutics* 17 : 1659–1669.

Watanabe K, Taskesen E, Arjen van B & Posthuma D (2017): Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* 8 : 1826.

Williams R, Airey M, Baxter H, Forrester J, Kennedy-Martin T & Girach A (2004): Epidemiology of Diabetic retinopathy and Macular Oedema: A systematic review. *Eye* 18 : 963–983.

Winkler TW, Grassmann F, Brandl C, Kiel C, Günther F, Strunz T ... Weber BH (2019): Genome-wide association meta-analysis for early age-related macular degeneration highlights novel loci and insights for advanced disease [Preprint].

Yang J, Jiang L & Zheng Z (2021): Fastgwa-GLMM: A generalized linear mixed model association tool for biobank-scale data [Preprint].

Yang X, Zong Y, Zhang Z, Zhao Y, Gao X, Zhang J ... Xiao B (2023): Identification of potential abnormal methylation-modified genes in coronary artery ectasia. *International Journal of Genomics* 2023 : 1–17.

Yau JWY, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T ... Wong TY (2012): Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care* 35 : 556–564.

Zhang S, Zhang F, Wang J, Yang S, Ren Y, Rui X ... Sheng X (2022): Novel compound heterozygous variations in MPDZ gene caused isolated bilateral macular coloboma in a Chinese family. *Cells* 11 : 3602.

Zhang T, Xu J & Xu P (2021): *eya2* expression during mouse embryonic development revealed by *eya2lacZ* knockin reporter and homozygous mice show mild hearing loss. *Developmental Dynamics* 250 : 1450–1462.

Zhu Z, Chen B, Na R, Fang W, Zhang W, Zhou Q ... Fang F (2020): A genome-wide association study reveals a substantial genetic basis underlying the ebbinghaus illusion. *Journal of Human Genetics* 66 : 261–271.

Figure legends

Fig. 1 The Manhattan plot of the primary GWAS analysis on diabetic retinopathy in type 2 patients (N = 16,988)

The dashed grey line indicates the cut-off p value of 5×10^{-8}

Fig. 2 Tissue expression results on 30 general tissue types and 53 specific tissue types by GTEx in the FUMA.

The dashed line shows the cut-off p value for significance with Bonferroni adjustment for multiple hypothesis testing

Fig. 3 The Circular Manhattan plots of additional GWASs analysis on diabetic retinopathy in type 2 patients

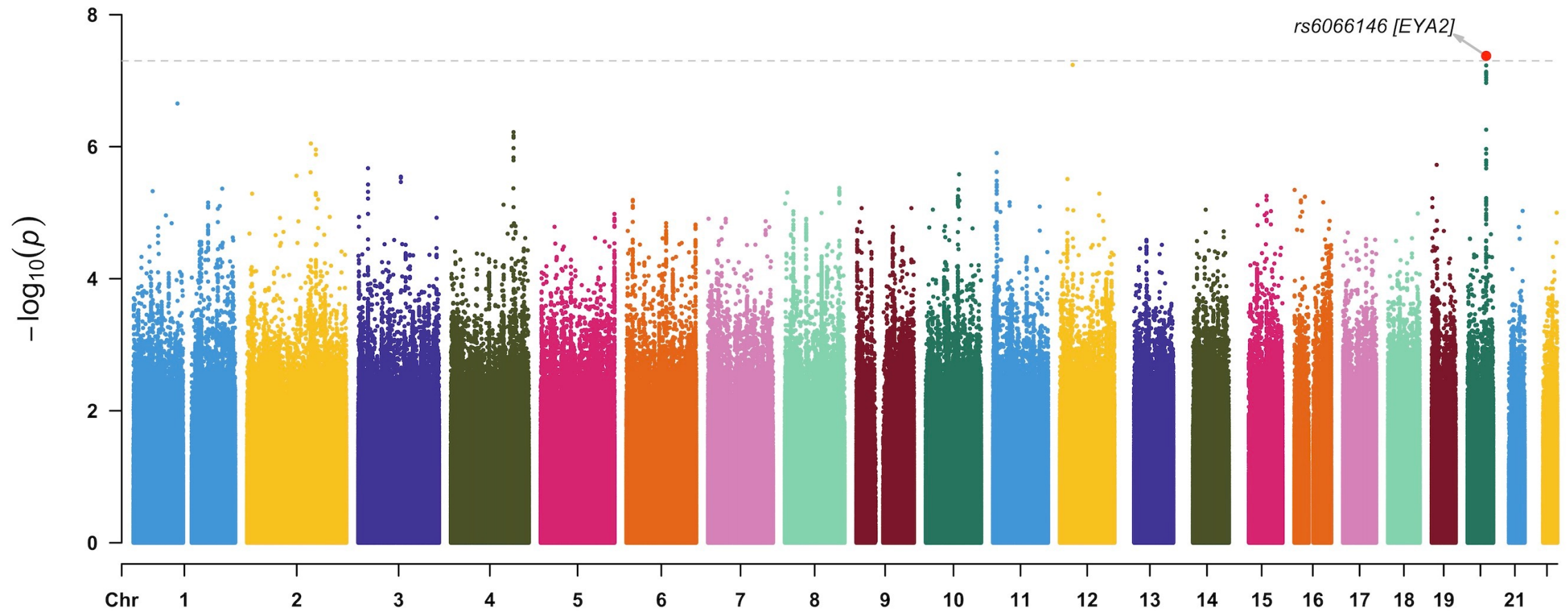
Fig. 4 The genetic interaction network of the key genes from GeneMANIA

Figure S1 The QQ plot of primary GWAS

Figure S2 Manhattan plot of gene analysis in primary GWAS

Figure S3 The regional plots of significant loci of five significant GWASs

Figure S4 The QQ plots of additional GWASs



GTEx v8 30 general tissue types

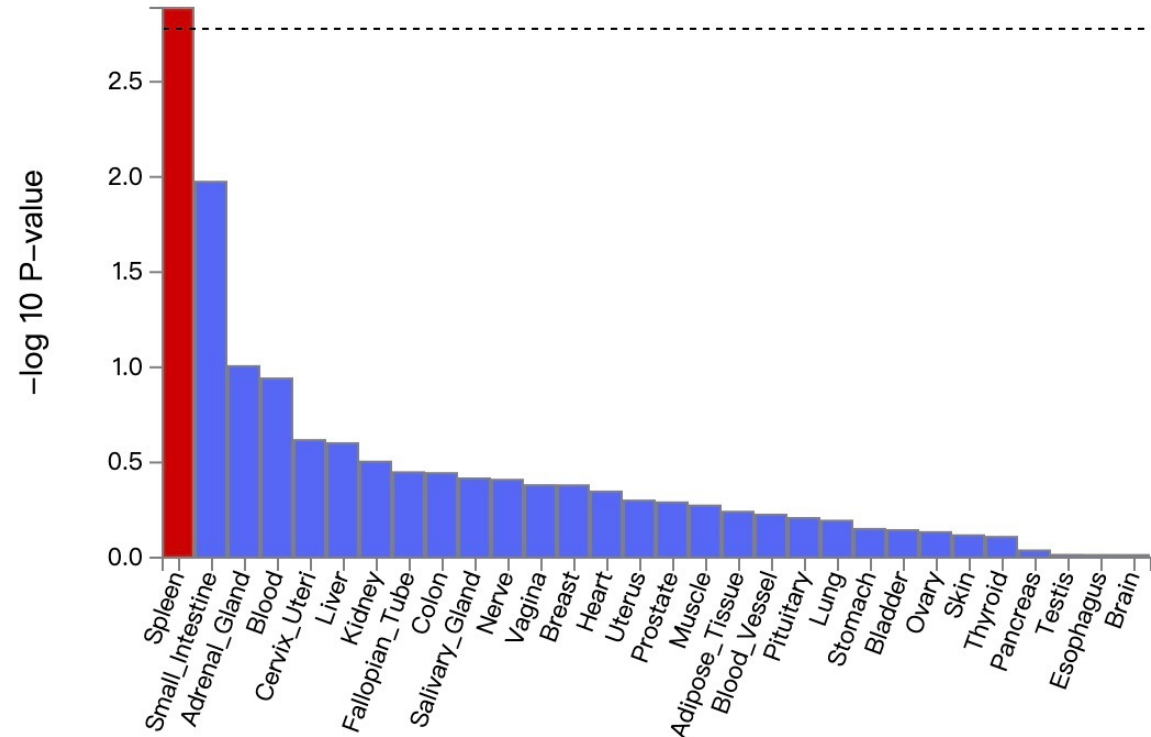
Download the plot as

[PNG](#)

[JPG](#)

[SVG](#)

[PDF](#)



GTEx v8 53 tissue types

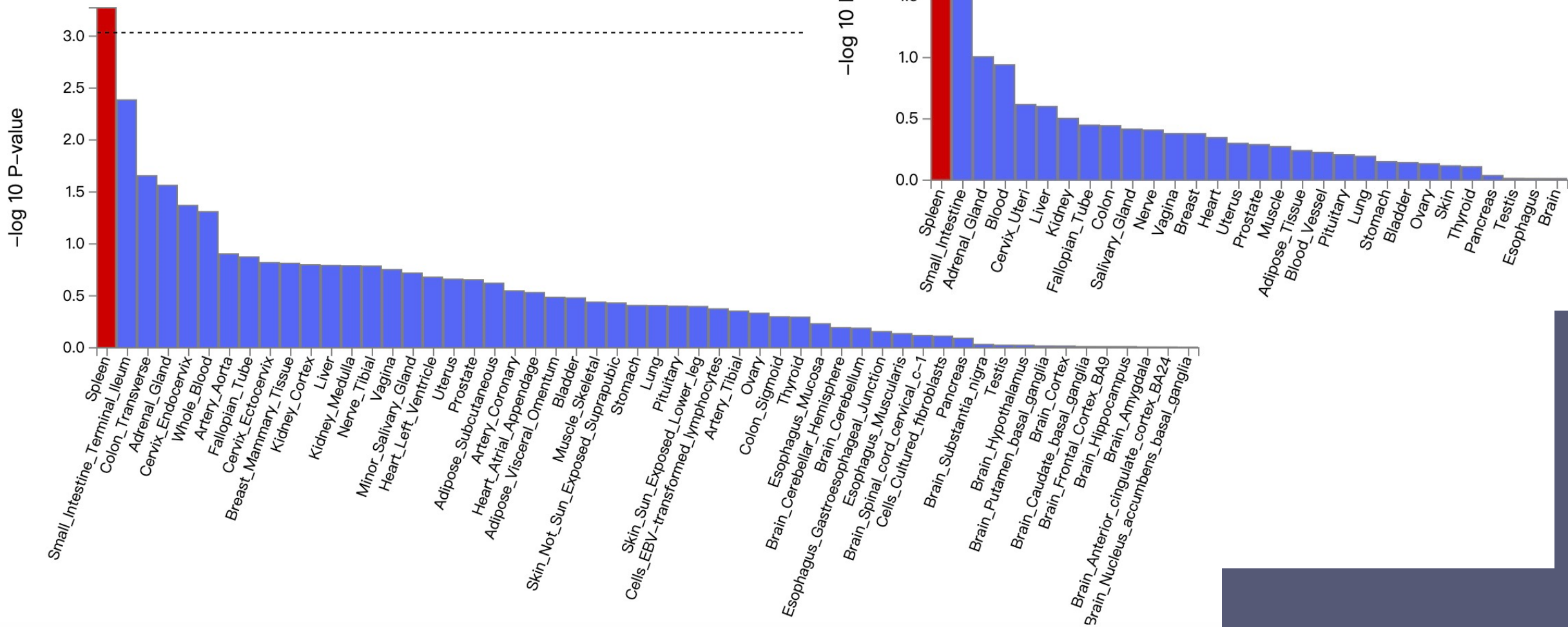
Download the plot as

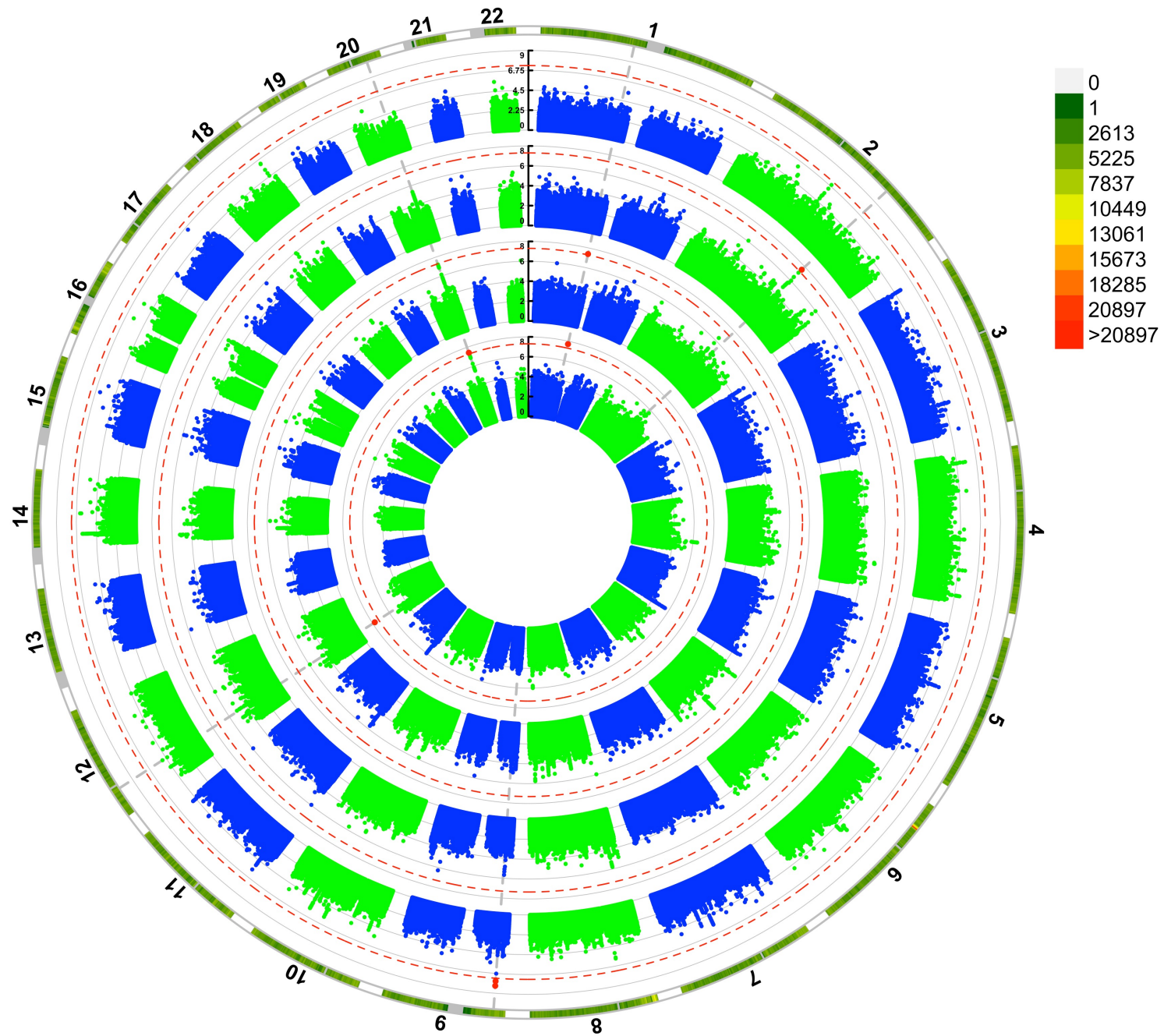
[PNG](#)

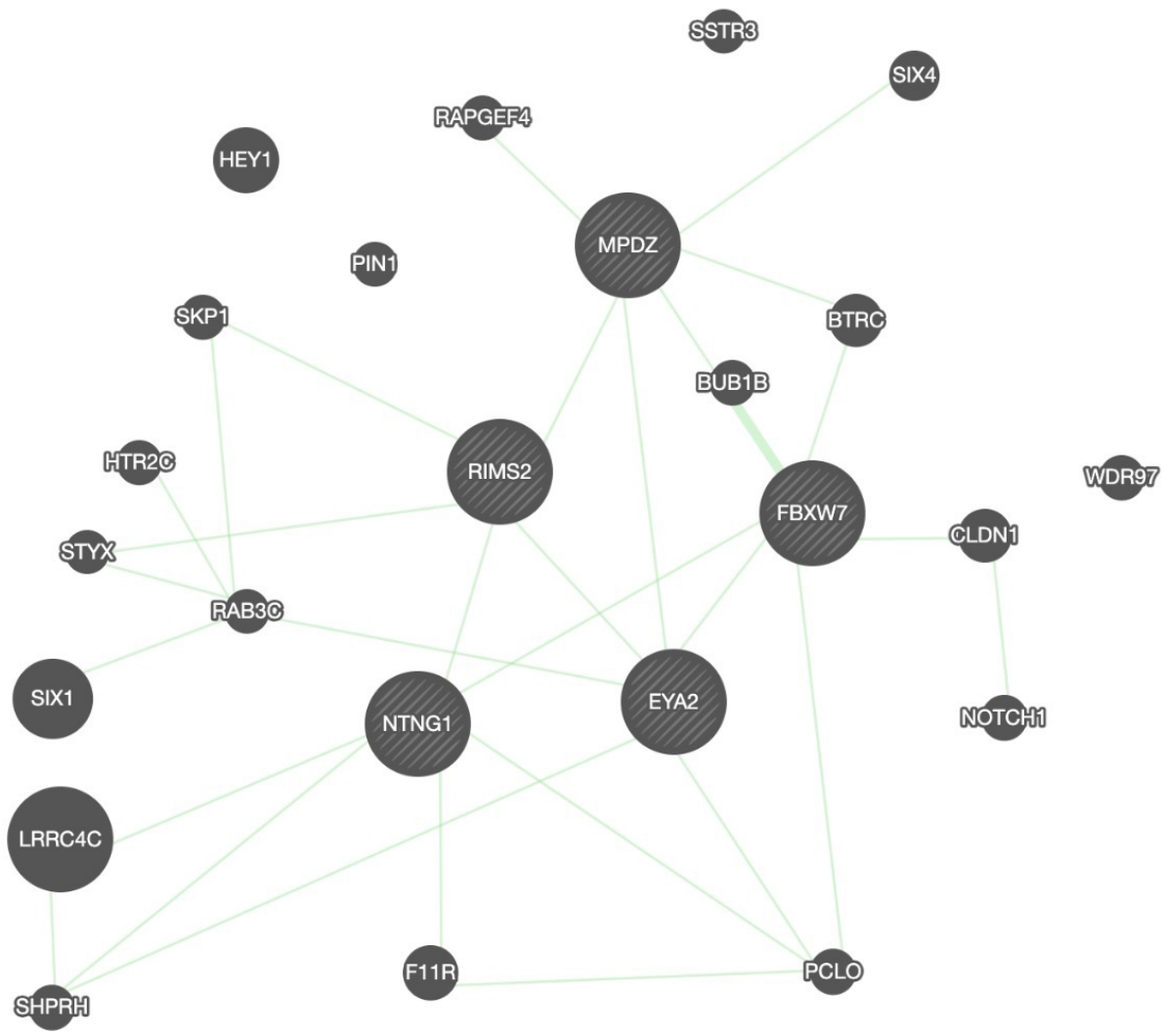
[JPG](#)

[SVG](#)

[PDF](#)







Networks

- Physical Interactions 77.64%
- Co-expression 8.01%
- Predicted 5.37%
- Co-localization 3.63%
- Genetic Interactions 2.87%
- Pathway 1.88%
- Shared protein domains 0.60%