Phenotypic and Genetic Characteristics of Retinal Vascular Parameters and their Association with Diseases

Sofía Ortín Vela* [*1,2], Michael Johannes Beyeler* [†1,2], Olga Trofimova [1,2], Mattia Tomasoni [3], Ilaria Iuliani [1,2], David Presby [1,2], Florence Hoogewoud [3], and Sven Bergmann [‡1,2,4]

[1]Dept. of Computational Biology, University of Lausanne, Lausanne, Switzerland
[2]Swiss Institute of Bioinformatics, Lausanne, Switzerland
[3]Jules-Gonin Eye Hospital, FAA, University of Lausanne, Lausanne, Switzerland
[4]Dept. of Integrative Biomedical Sciences, University of Cape Town, Cape Town, South Africa

Abstract

Fundus images allow for non-invasive assessment of the retinal vasculature whose features provide important information on health. Using a fully automated image processing pipeline, we extracted 17 different morphological vascular phenotypes, including median vessels diameter, diameter variability, main temporal angles, vascular density, central retinal equivalents, the number of bifurcations, and tortuosity, from over 130k fundus images of close to 72k UK BioBank subjects. We performed genome-wide association studies (GWAS) of these parameters. From this, we estimated their individual heritabilities, ranging between 5 and 25%, and genetic cross-trait correlations, which mostly mirrored the corresponding phenotypic correlations, but tended to be slightly larger. Projecting our genetic association signals onto genes and pathways revealed remarkably low overlap suggesting largely decoupled mechanisms modulating the different traits. Our disease trait associations confirmed some previously known findings and revealed many novel connections. Notably, diameter variability, especially for the veins, seems to have many new and interesting associations with multiple disease traits, including age of death, pulmonary embolism, and heart attack. Our analyses provide evidence that large-scale analysis of image-derived vascular traits has sufficient power for obtaining functional, as well as some initial causal insights into the processes modulating the retinal vasculature.

Keywords: GWAS, retina, CFI, IDPs, vasculature, genes, vascular diseases, heritability, MR.

Nonstandard Abbreviations and Acronyms

IDP: Image-derived phenotype, PWASI: Pulse wave Arterial stiffness index, HDL: High-Density Lipoprotein, HbA1c: Glycated haemoglobin, BMI: Body mass index, DVT: Deep Vein Thrombosis, PE: Pulmonary Embolism, ED: Eye-diseases, DBP: Diastolic blood pressure, SBP: Systolic blood pressure, PR: Pulse rate, PAD: Peripheral artery disease, LDSR: Linkage Disequilibrium Score Regression, OD: Optic disc.

**: equal first authors in random order
*sofia.ortinvela@unil.ch
†michael.beyeler@unil.ch
‡sven.bergmann@unil.ch

Contents

Introduction **3**

Results **6**

Automated pipeline for phenotyping of the retinal vasculature 6

Correlation structure and heritability of retinal vascular IDPs 6

Gene and pathway level analysis of retinal vascular IDPs 8

Phenotypic association with diseases and risk factors 10

Genetic association with risk factors and causality analysis 12

Discussion **15**

Methods **18**

UK Biobank data 18

Image segmentation and quality control 19

Phenotyping 19

Correction for covariates 20

Disease association 20

Genome-wide analyses 20

Genetic association with diseases 21

Mendelian Randomisation 21

Data and code availability 22

Author contributions **22**

Funding **23**

References **30**

Introduction

The retina provides a unique opportunity for imaging human vasculature. In particular, retinal fundus photos allow for noninvasive in-vivo assessment of the vascular system of the superficial inner layer of the retina. Such images have been acquired in several cohorts and there is a large body of research on automatic extraction of vascular properties and their associations with medically relevant information.

It is well established that vascular properties obtained from retinal imaging not only enable the monitoring of ocular diseases such as diabetic retinopathy, macular degeneration, and glaucoma, but can also serve as a powerful screening tool for early detection of systemic diseases, including stroke [1–4], coronary heart disease [5, 6], peripheral artery diseases (PADs) [7], hypertension [3, 4, 8–18], atherosclerosis [2, 3, 19], and myocardial infarction [20, 21]. Retinal abnormalities have also been associated with common comorbidities such as diabetes [8, 22–24] and obesity [25].

The classical processing of retinal images typically can be divided into three steps. First, the image is processed at the level of pixels to identify which of them represents blood vessels (possibly distinguishing between arteries and veins), or other structures, like the optic disc (OD). In the second step this pixel-wise information is used to annotate the retina in terms of objects, such as vessel segments represented as a list of points along their midline, as well as the vessel widths along these points. Finally, the information from these objects is used to measure different vascular properties such as vessel diameter or tortuosity, number of bifurcations, as well as certain angles between major vessels. Some simple vascular parameters, such as vascular density or fractal dimension, can also be computed directly from the pixel-wise information.

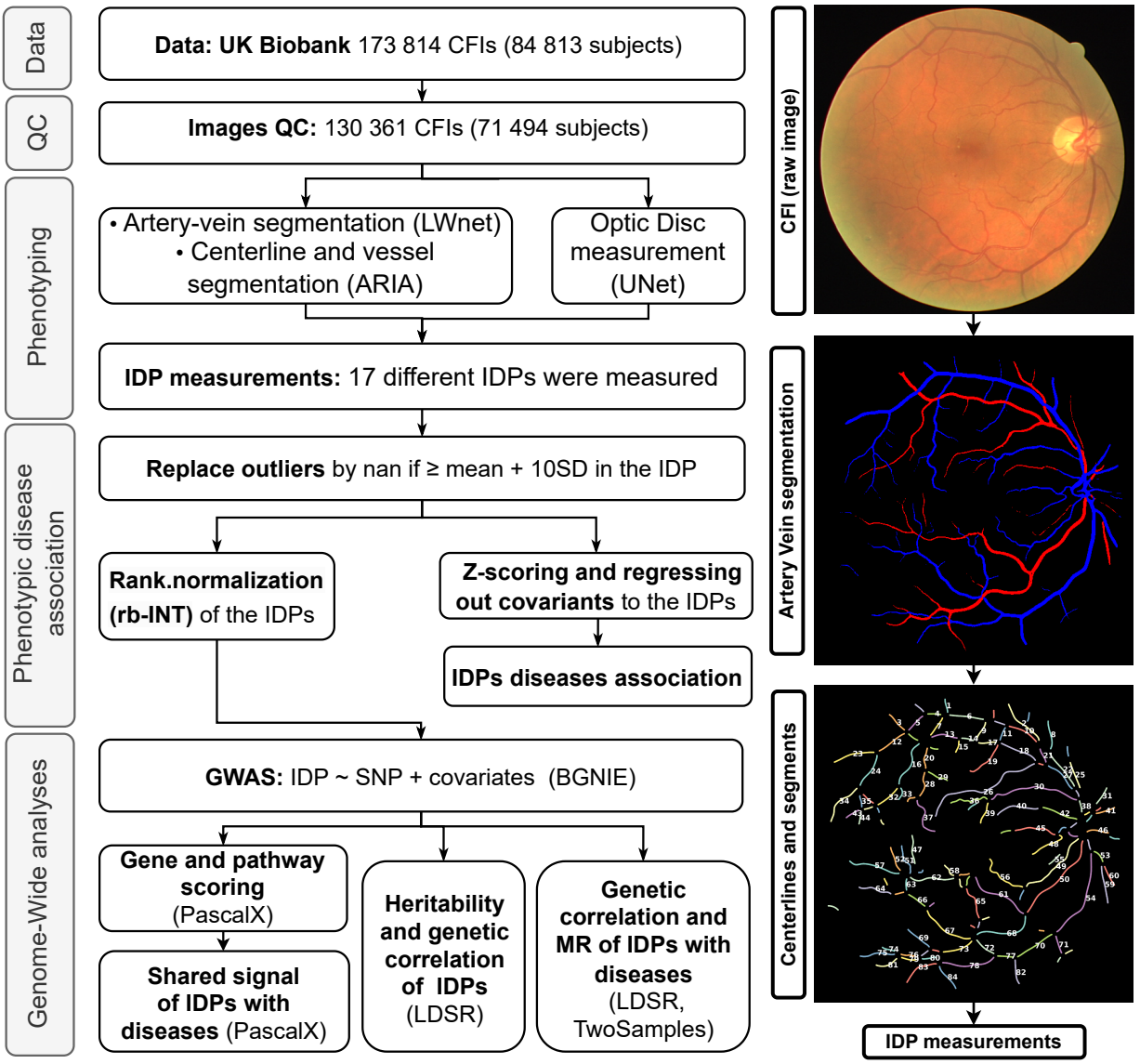
While there have been several studies analysing retinal vascular parameters [26–31], most of them focused on measuring just one or few retinal traits, often in small image sets, and some required expert input rather than being fully automated [32–34]. Furthermore, the software used for vascular phenotyping is usually not openly accessible, with one very recent exception [35]. Together this precludes the establishment of a comprehensive and reproducible characterisation of large retinal image collections.

Recently, Deep Learning (DL) approaches have gained popularity in retinal image analysis. The first contribution is at the level of pixel-wise annotation, where state-of-the-art segmentation can be achieved with Convolutional Neural Network (CNN) architectures. For example, the little W-Net (LWNET) vessel-annotates pixels as being part of an artery or vein, outperforming classical segmentation approaches [36]. Such networks can also be used to annotate pixels belonging to the OD [37], vessel bifurcations [38], or other structures. The second contribution of deep CNN is to learn latent variables providing efficient low-dimensional representations of retinal images [39]. Such self-supervised, image-based phenotyping can generate novel traits complementing explicit retinal features. Finally, DL approaches have been used to directly predict health-relevant traits from retinal images [40, 41].

Beyond their value for assessing ocular or systemic health, retinal phenotypes have also been used in Genome-Wide Associations Studies (GWAS) to identify genetic variants modulating these traits. However, previous

GWAS have focused on a limited set of vascular properties, such as vessel diameter [42–45], tortuosity [46, 47], vascular density, fractal dimension [48], and certain deep latent variables [39] (see Suppl. text "Comparison with previous GWAS").

In this study, we present the first joint analysis of 17 retinal vascular parameters, including novel features like temporal angles, number of bifurcations, or diameter variability, which have not been analyzed previously at a large scale. Leveraging data from the UK Biobank (UKBB), and employing our open-source fully automated analysis platform, we provide retinal vascular phenotyping for close to 72k subjects, after quality control (QC). We performed GWAS for all traits, heritability estimates, gene and pathway analyses, as well as associations with a broad set of systemic and ocular diseases. This allowed us to compare cross-trait correlations, both at the phenotypic and genotypic levels, study in detail which genes affect individual or multiple traits, and identify potential causal relationships with diseases (see Fig. 1 depicting the overall methodology of our study).



(a) Pipeline (b) CFI processing

Figure 1: **a)** Methods pipeline. Subjects' basic and medical information, genotypes, and CFIs were collected from the UKBB. Applying the Image QC method of [48] removed $\sim$ 25% of all CFIs. Pixel-wise vessel segmentation and classification were performed using LWNET [36]. ARIA [49] was used to identify vessel segment objects. A DL network was used to measure the position of the OD [37]. Based on this primary information our bespoke algorithms measured vascular IDPs. IDPs (z-scored and corrected for covariates) were associated with diseases through linear and logistic regressions, and Cox models. GWAS was performed on all IDPs after rank-based inverse normal transformation (rb-INT) and correction for covariates, and the resulting summary statistics were used to estimate heritabilities and genetic correlations, to identify relevant genes and pathways, and to study the genetic association and potential causal relationships between the IDPs and some of the disease traits. **b)** IDP measuring process summary on DRIVE image. Top: original CFI. Middle: Pixel-wise segmented vasculature with artery-vein classification using LWNET [36]. Bottom: Vessel segment objects in terms of centerlines and diameters were identified using ARIA [49], providing the starting point for measuring vascular IDPs (see Suppl. text "Methods for phenotype extraction" for details).

Results

Automated pipeline for phenotyping of the retinal vasculature

To analyze the retinal vasculature, we developed an automated pipeline building on our previous work on retinal vessel tortuosity [46]. This pipeline enabled us to segment and annotate the retinal vasculature and OD from 130 361 CFIs of 71 494 subjects after QC. Applying a range of quality control thresholds, we observed that trait heritabilities tended to be higher, while disease incident rates were lower when applying more stringent thresholds, possibly because some diseases impact the vascular density which is a good proxy for image quality. We thus settled for an intermediate QC threshold removing $\sim 25\%$ of all CFIs (see Suppl. text "Quality control threshold effect on results" for details).

We selected 17 representative IDPs from a broader set of 36 parameters to characterize each image, including vascular densities, median tortuosities, central retinal equivalents, median diameters, diameter variabilities, and main temporal angles for arteries and veins. We also calculated the ratios between artery and vein values for the first four parameters, and we estimated the total number of vessel bifurcations. These IDPs were chosen based on their associations with diseases and the reliability of measurement. The distributions of these main IDPs are presented in Suppl. text "Distribution of retinal vascular IDPs", and their specific sample sizes in Suppl. table "Sample size trait". The broader set of 36 IDPs can be found in Suppl. text "Methods for phenotype extraction".

Correlation structure and heritabilities of retinal vascular IDPs

To explore the relationships between our main IDPs, we calculated pair-wise phenotypic Pearson correlations, $r_{ij}^{(p)}$ (upper right triangle in Fig. 2a). We observed that the same IDPs measured for arteries and veins tended to cluster together, particularly the temporal angles, tortuosities, and vascular densities. Vascular densities were highly correlated with the number of bifurcations, $r^{(p)}$=0.91, the strongest cross-IDP correlation observed. Vascular densities and number of bifurcations were significantly correlated with each other and with venous, but not arterial, tortuosity. They were also strongly anti-correlated with median diameters. The temporal angles formed their own branch, with low correlations to the other traits. Finally, most artery-vein ratios were highly correlated with their corresponding arterial, and highly anti-correlated with their corresponding venous measures.

To assess the genetic relationships between the IDPs, we estimated pair-wise genetic correlations, $r_{ij}^{(g)}$, using cross-trait LDSR [50] (lower left triangle in Fig. 2a). On average, genetic correlations were slightly higher than phenotypic correlations (standardized mean difference across the 136 pairs (ij) is $d = 0.34$, t(135)=6.53, $p = 6.2 \times 10^{-10}$). The largest difference occurred between temporal angles. Overall there was a strong correspondence between genetic and phenotypic correlations, $corr(r^{(g)}, r^{(p)}) = 0.86$, permutation $p < 1 \times 10^{-4}$ (see Suppl. text "Correlation structure and heritabilities of extended list of retinal vascular

IDPs" for the corresponding analysis for our broad set of 36 IDPs).

To gain deeper insights into the genetic control of retinal vascular morphology, we estimated the single nucleotide polymorphism (SNP) heritability, h^2, for each IDP (Fig. 2b). Our analysis revealed varying levels of heritability across different traits. Arterial tortuosity exhibited the highest heritability, h^2=0.25±0.02, followed by the tortuosity ratio, h^2=0.20±0.02, and venous diameter variability, h^2=0.18±0.02. In contrast, the median diameter measures, particularly arterial median diameter, showed the lowest heritability, h^2=0.05±0.01 (see Suppl. text "Manhattan and Quantile-Quantile plots" for Manhattan and Quantile-Quantile plots from the GWAS of each IDP).

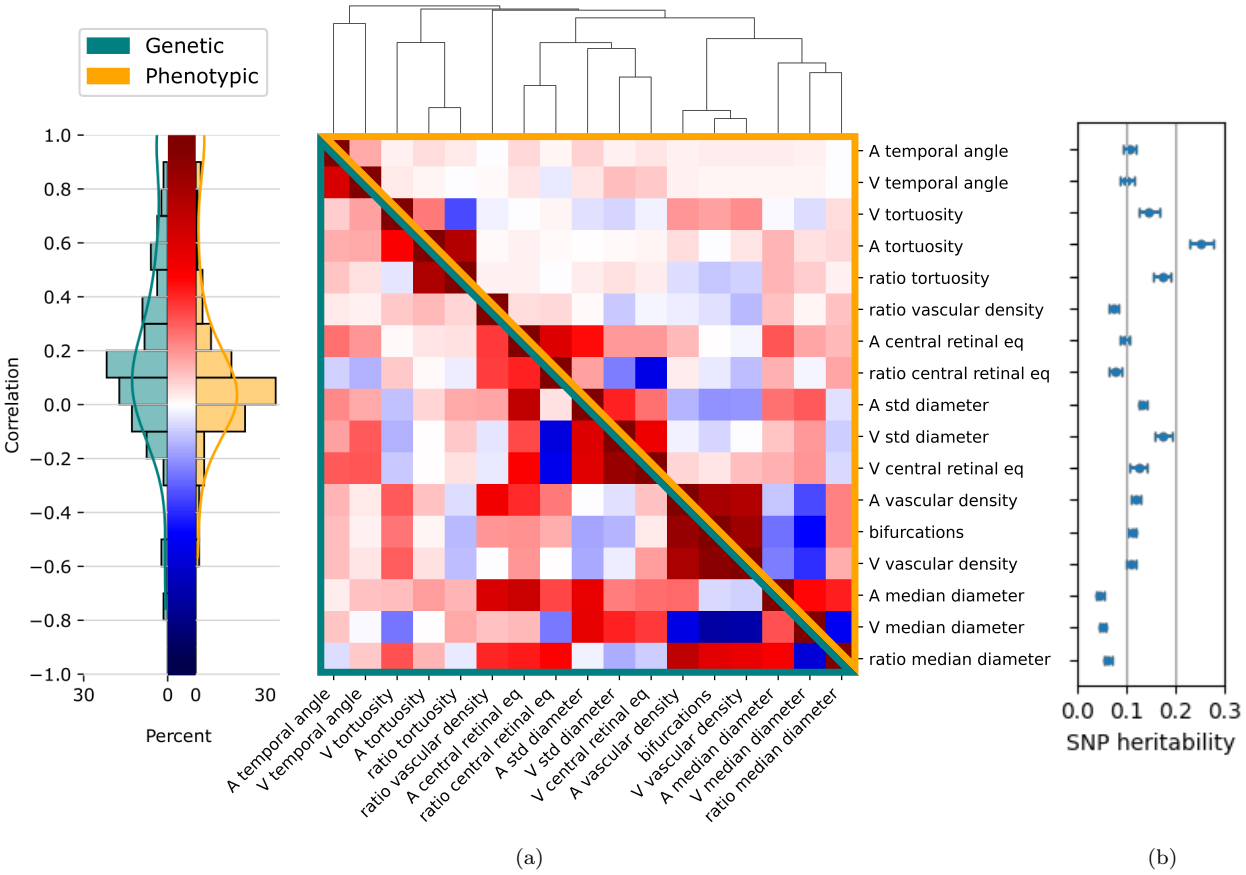


Figure 2: **a)** Phenotypic (upper-left orange triangle) and genetic (lower-right green triangle) correlations between retinal vascular traits, clustered by absolute phenotypic correlation distance, $1 - |corr|$. The 17 traits are (A: artery, V: vein): main temporal angles ('A/V temporal angle'), median tortuosities and their ratio ('A/V tortuosity' and 'ratio tortuosity'), central retinal equivalents and their ratio ('A/V central retinal eq' and 'ratio central retinal eq'), diameter variabilities ('A/V std diameter'), vascular densities and their ratio('A/V vascular density' and 'ratio vascular density'), median diameters and their ratio ('A/V median diameter' and 'ratio median diameter'), and the number of bifurcations ('bifurcations'). Traits were corrected for age, sex, eye geometry, batch effects, and ethnicity before phenotypic clustering and before GWAS (see Methods). **b)** Corresponding trait SNP heritabilities, h^2, and their standard error, estimated using LDSR [50].

Gene and pathway level analysis of retinal vascular IDPs

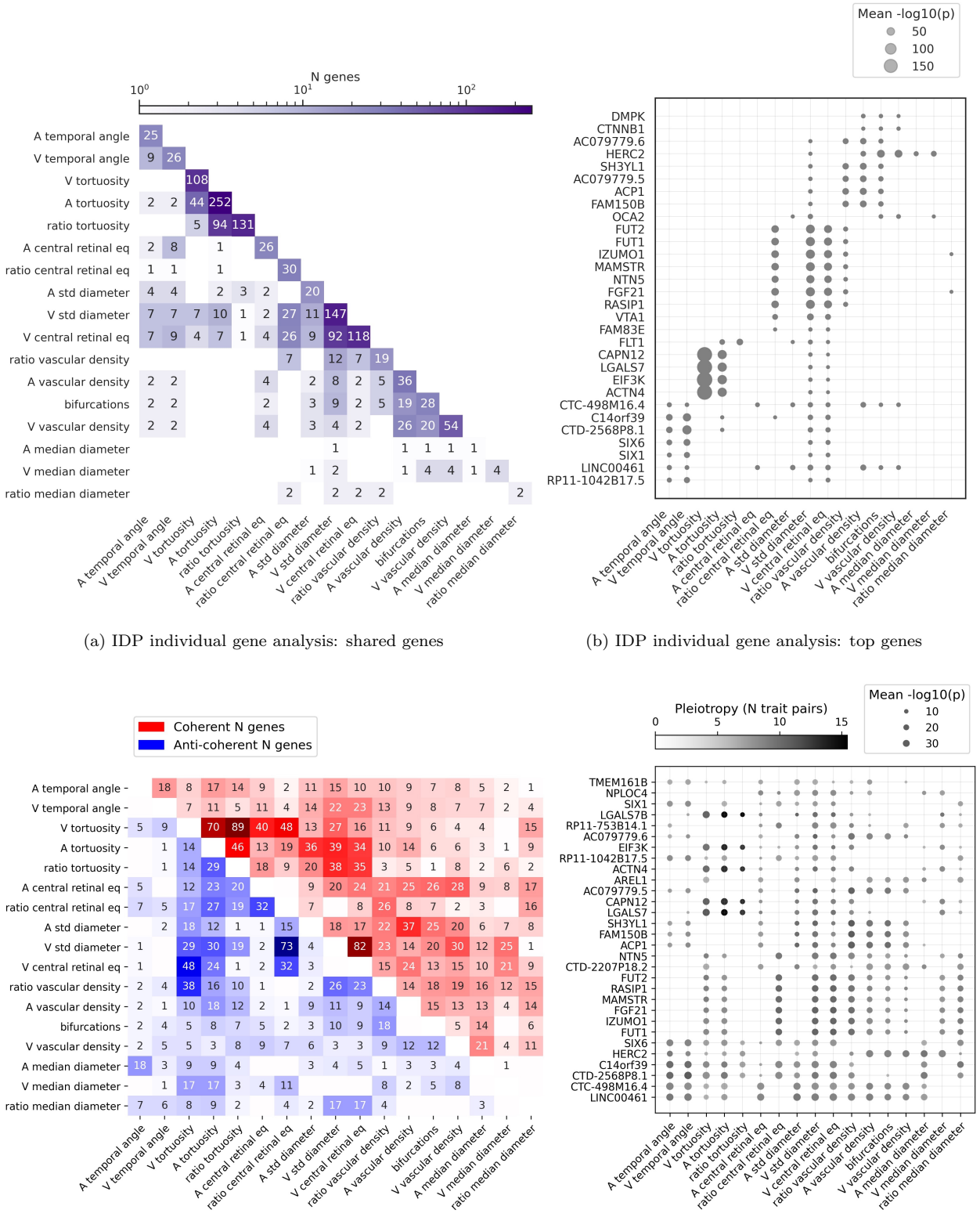
To identify the genes modulating specific vascular IDPs, we employed our *PascalX* analysis tool [51, 52]. This tool aggregates SNP-wise association signals within gene windows and generates gene scores.

The number of genes associated with IDPs (diagonal of Fig. 3a) ranged from 1 gene for the arterial median diameter to 252 genes for arterial tortuosity. We observed a higher number of genes for tortuosities, venous diameter variability, and venous central retinal equivalent. The off-diagonal elements in Fig. 3a show the number of common genes for each pair of IDPs. Generally, more correlated IDPs tended to share more genes. However, even among highly similar IDPs, a significant portion of genes were trait-specific, and some IDPs shared only a few or no genes.

While no single gene was shared among all IDPs, two genes, *LINC00461* and *CTC-498M16.4*, were associated with 9 of the 17 IDPs (Fig. 3b). Other genes associated with multiple IDPs included *SIX6*, *FLT1*, *FUT1*, *HERC2*, and *PDE6G*. The complete list of significant genes associated with each IDP is available as Suppl. data "Data Significant genes per IDP".

Furthermore, we conducted pathway analysis using *PascalX* to identify gene sets, or "pathways", that exhibited higher association signals for each IDP GWAS than expected by chance. Although no single pathway was shared among all IDPs, some of the most frequent pathways included 'Fetal retina fibroblast', and 'Abnormal retinal morphology' (Suppl. text "Pathway analyses"). The complete list of significant pathways per IDP is available as Suppl. data "Data Significant pathways per IDP".

To further investigate pleiotropic genes, we employed *PascalX* cross-GWAS analysis [52], a method that examines coherent effects across the SNPs associated with two traits within a gene window. This approach has more power than just intersecting the gene sets of two individual traits and also allows for distinguishing between coherent and anti-coherent effects. However, even in this more sensitive analysis, no single gene was shared between all pairs of traits. The most pleiotropic genes largely overlapped with those identified by simple gene-scoring, yet they tended to be shared among more IDPs (Fig. 3c and Fig. 3d). See Suppl. data "Data PascalX retina-retina cross gene scores" for a comprehensive list of genes. Generally, IDP pairs with positive LDSR genetic correlation shared more coherent gene signals, while those with negative LDSR genetic correlation had predominantly anti-coherent signals (see Suppl. text "LDSR genetic correlation against PascalX"). We also performed pathway analyses for some of the trait pairs for which *PascalX* identified large gene sets: For example, the coherence test between the variability in vein diameter and central retinal venular equivalent yielded 82 genes that were enriched gene sets annotated as related to abnormal morphology of the retina, choroid and uvea, as well as pigmentation.



(a) IDP individual gene analysis: shared genes

(b) IDP individual gene analysis: top genes

(c) IDP pairs gene analysis: shared genes

(d) IDP pairs gene analysis: top genes

Figure 3: **a)** Number of genes associated with the different IDPs, obtained using *PascalX* gene-scoring [52]. The diagonal shows the number of genes significantly associated with each IDP. The lower triangle shows the number of genes in the intersection between pairs of IDPs. **b)** 30 genes most frequently associated with the IDPs. Dot sizes are inversely proportional to p-values. **c)** Number of genes showing coherent (top right) or anti-coherent (bottom left) signal between pairs of IDPs, obtained using *PascalX* cross-scoring [52]. **d)** 30 genes most frequently found in the cross-trait analysis. Dot color represents pleiotropy, i.e. the number of trait pairs showing (anti-)coherent signal for a given gene. Dot sizes are inversely proportional to p-values.

Phenotypic association with diseases and risk factors

To evaluate the clinical relevance of retinal vascular morphology, we examined the phenotypic associations between the retinal vascular IDPs and various eye-related diseases, vascular diseases, and their associated risk factors. We employed linear regression for continuous traits (Fig. 4a), logistic regression for binary traits (Fig. 4b), and Cox models for diagnoses with age-of-onset (Fig. 4c). All IDPs were standardized (z-scored) and adjusted for potential confounders such as age, sex, eye geometry, batch, and ethnicity.

Amongst the continuous traits, diastolic blood pressure (DBP) and systolic blood pressure (SBP) displayed similar associations with most IDPs. Smoking exhibited the strongest positive association with venous diameter variability, while strong negative correlations were observed between blood pressure (BP) and arterial-related IDPs, particularly the central retinal equivalent, vascular density, and median diameter (Fig. 4a).

For binary traits, hypertension was positively associated with venous tortuosity and slightly with arterial tortuosity, while negatively associated with central retinal artery equivalent, and the ratios of central retinal equivalents and vascular densities, among others. Eye diseases such as amblyopia were highly negatively associated with the number of bifurcations and venous vascular density, $\beta \in$ [-0.161, -0.158]. Others like presbyopia, hypermetropia, and myopia showed smaller effects, $\beta \in$ [-0.075 to 0.057], but were still statistically significant. Retinal diabetes was highly negatively associated with the vascular densities and the number of bifurcations, $\beta \in$ [-0.298 to -0.248], while atherosclerosis was highly positively associated with venous diameter variability, $\beta = 0.344$ (Fig. 4b).

The Cox analysis for age-of-onset traits included severe eye- and cardiovascular diseases, diabetes, and age at death. Both vascular densities displayed strong and consistent associations with the age of diagnosis for all eye-related diseases. Diabetes age-at-diagnosis shared similar associations with our vascular IDPs as diabetes in the eye, but was also associated with venous tortuosity and central retinal arterial equivalent, among others. Heart attack was associated with higher median arterial diameter and central retinal arterial equivalent, as well as diameter variability in both vessel types. Importantly, earlier deaths were most strongly associated with increased venous diameter variability and less strongly with increased central venular retinal equivalent, increased venous temporal angles, and conversely with lower vascular density and fewer bifurcations. In general, venous diameter variability was associated with almost all diseases, including the only associations with pulmonary embolism and stroke (Fig. 4c).

While vascular densities and the number of bifurcations showed strong phenotypic correlation, their associations with diseases sometimes differed. For instance, SBP and hypertension were associated with arterial vascular density and the number of bifurcations but not with venous vascular density. Interestingly, atherosclerosis and stroke shared one positive association, the standard deviation of vein diameter (already reported above for diabetes, angina, and heart attack). No other parameters were significant for the two conditions.

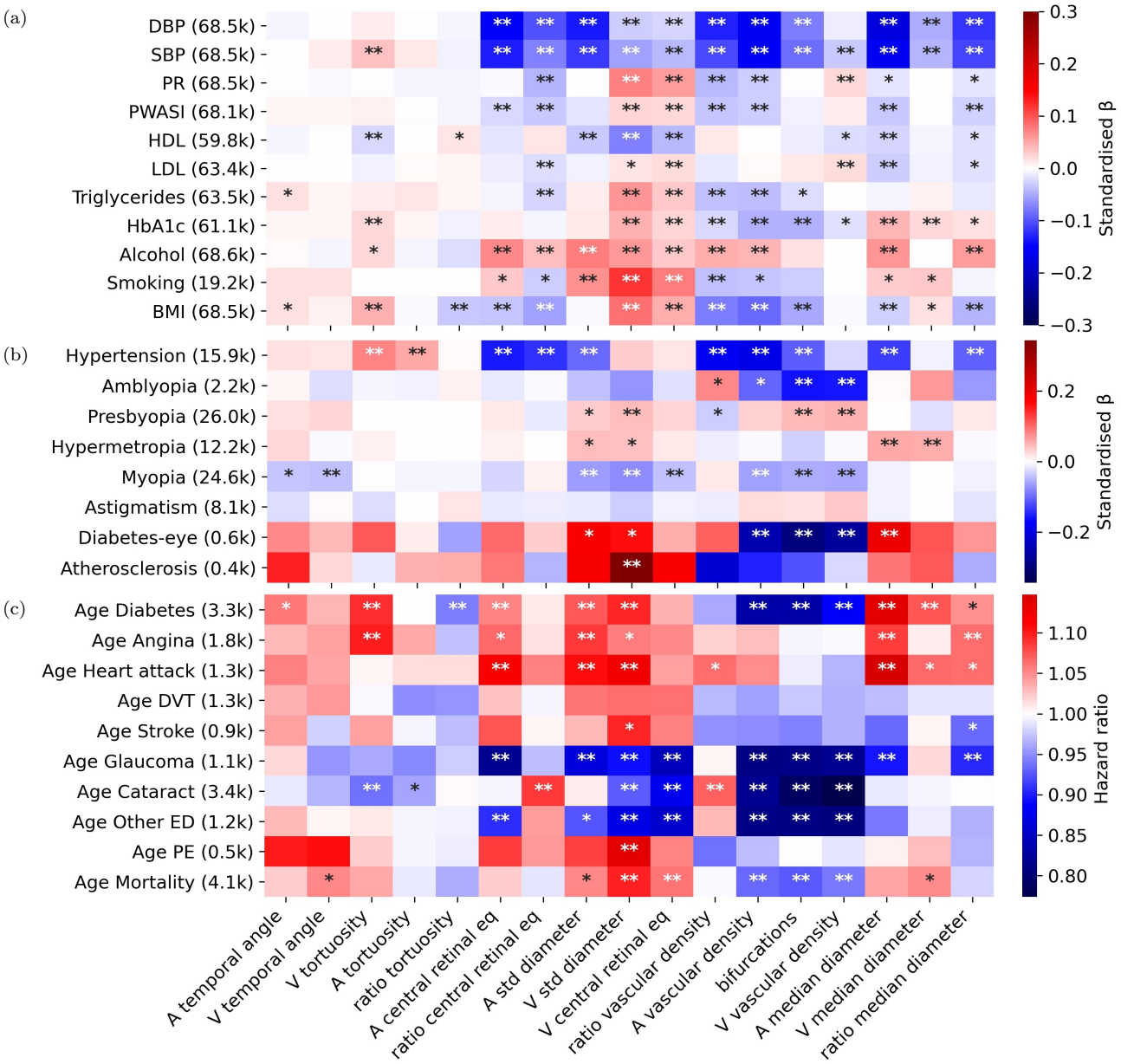


Figure 4: Phenotypic association of IDPs with risk factors and diseases. The x-axis shows IDPs and the y-axis shows risk factors and diseases. For continuous traits, the numbers in parentheses correspond to the number of subjects with trait information for which we were able to measure at least one of the 17 IDPs. For binary case/control traits, it represents the number of subjects who are cases and have data for at least one of the 17 IDPs. Linear (**a**) and logistic (**b**) regressions were used for continuous and case/control diseases, respectively. For age-of-death and other severe diseases with the age-of-onset information, Cox proportional hazards regression was performed (**c**). In all models, traits were corrected for age, sex, eye geometry, batch effects, and ethnicity. The color indicates standardized effect sizes for linear and logistic regressions, or hazard ratios for Cox models. The asterisks indicate level of statistical significance ($* : p < 0.05/N_{test}$, $** : p < 0.001/N_{test}$, where $N_{test} = N_{IDPs} \times N_{Diseases}$, and $N_{Diseases} = N_{LinearDiseases}$, $N_{LogisticDiseases}$, or $N_{CoxDiseases}$ respectively). Labels: 'PR': Pulse Rate, 'PWASI': Pulse wave Arterial stiffness index, 'HDL': High-Density Lipoprotein, 'HbA1c': Glycated haemoglobin, 'Alcohol': Alcohol intake frequency, 'Smoking': pack-years, 'Diabetes-eye': diabetes related to the eye, 'DVT': Deep Vein Thrombosis, 'Other ED': all types of severe eye diseases not included explicitly, 'PE': pulmonary embolism.

Genetic associations with risk factors and causality analysis

To investigate the extent of common genetic architectures between the vascular IDPs and risk factors, we first used LDSR to estimate their genetic cross-correlations. Notably, we observed the strongest negative correlations (≈ -0.30) between BP measures and the ratios of the central retinal equivalents, and vascular densities. Body mass index (BMI) was correlated positively with median vein diameter (≈ 0.15), and its variation (≈ 0.18), and negatively with arterial vascular density and the number of bifurcations (Fig. 5a). HbA1c was also negatively correlated with the number of bifurcations, while HDL was negatively correlated with venous diameter variability, but only with marginal significance.

To compare the phenotypic and genetic associations between vascular IDPs and risk factors, we plotted the effect sizes of the phenotypic linear regressions against genetic correlations for each IDP/risk factor pair (Fig. 5b). The slopes of the best-fitting lines through the corresponding points were positive for all risk factors, except for LDL cholesterol.

A strong genetic association between two traits does not imply the existence of a causal link. To systematically assess potential causal relationships between vascular IDPs and risk factors, we performed bidirectional two-sample Mendelian Randomisation (MR) analyses [53, 54]. Causal estimates were derived from the inverse variance weighted method using the TwoSampleMR R package [55, 56] (see Methods).

Using IDPs as exposures, we observed potential causal effects on vascular risk factors for 10 of them, but only with marginal significance (see Fig. 5c). Specifically, venous diameter variability had a positive effect on DBP and SBP, while the ratio of central equivalent had a negative effect on BP. The ratio of vascular density and the venous median diameter had a negative effect on HDL levels, while arterial central equivalent had a positive effect on HbA1c levels. The venous temporal angle, the venous central central retinal equivalent and the ratio of tortuosity had an effect on BMI. Finally, the venous temporal angle showed a negative effect on BMI and BP. In contrast, using risk factors as exposures, we observed evidence for causal effects on all IDPs, except the ratio of tortuosities (see Fig. 5d). Notably, we found strong evidence for negative effects of DBP and SBP on many different vascular IDPs, including vascular density, the arterial (and ratio) central retinal equivalent and median diameter. Also, SBP had a positive effect on arterial and venous tortuosity. BMI had a positive effect on venous and arterial diameter variability, while it showed a negative effect on ratio central equivalent and ratio vascular density, among others. Also, alcohol intake had a positive effect on arterial (but not venous) diameter variability, while PR had a positive effect on venous diameter variability. Finally, HbA1c levels had a weak negative effect on arterial (and ratio) vascular density and on the number of bifurcations (see Suppl. data "Data Mendelian Randomization risk factors" for complete results).

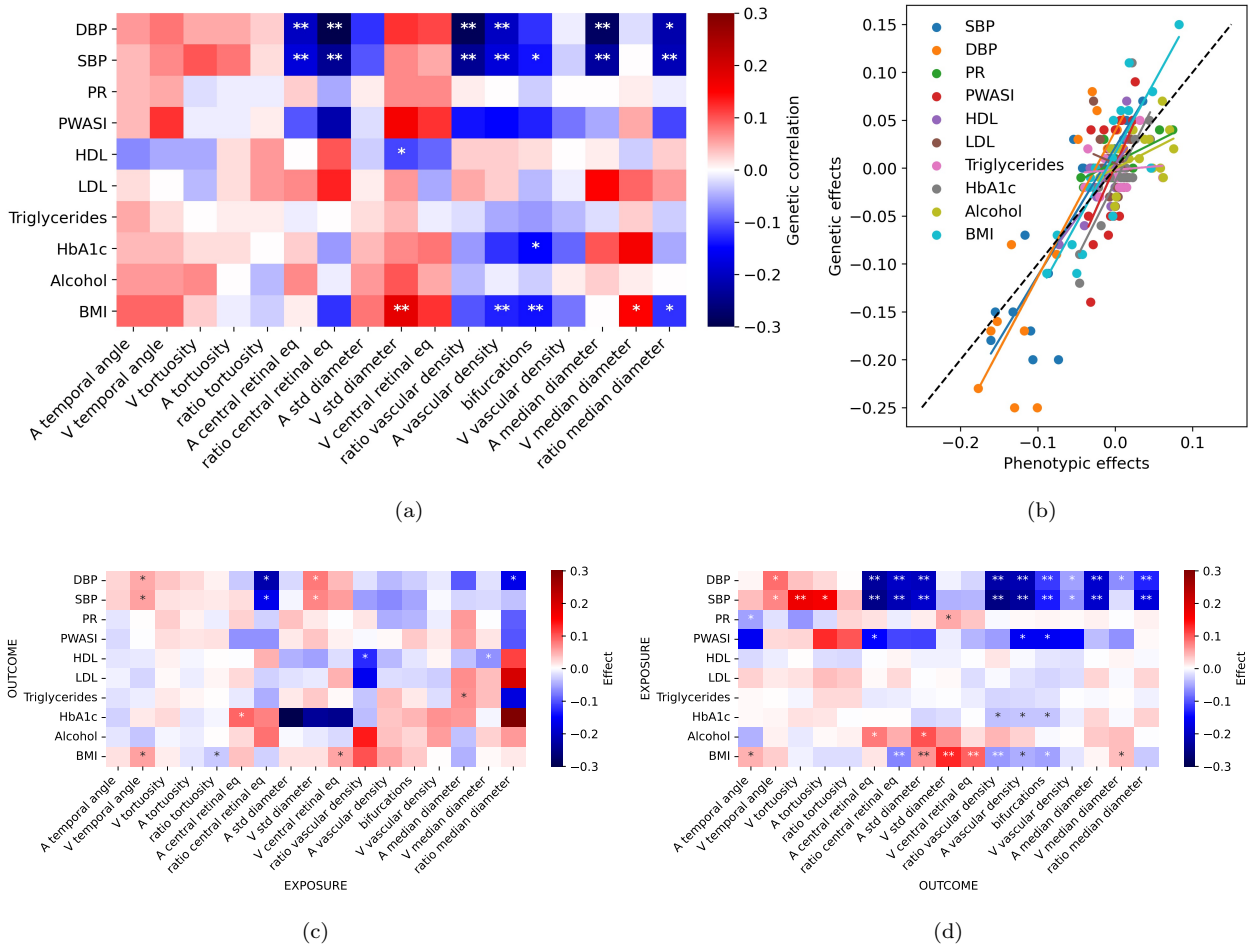


Figure 5: **a)** Genetic correlation between IDPs and continuous risk factors, computed using LDSR [50]. The color indicates the genetic correlation coefficient and the asterisks indicate the level of statistical significance ($* : p < 0.05/N_{test}$, $** : p < 0.001/N_{test}$, being $N_{test} = N_{IDPs} \times N_{LinearDiseases}$). **b)** Correlation between phenotypic and genetic correlations of IDPs with risk factors. **c)** Causal effect estimates with IDPs as exposures and risk factors as outcomes. **d)** Causal effect estimates with risk factors as exposures and IDPs as outcomes. The color indicates the causal effect estimates based on the inverse variance-weighted MR method. The stars indicate the level of statistical significance ($* : p < 0.05$, $** : p < 0.001$).

To identify the genes that were jointly associated with IDPs and risk factors, we first computed simple intersections between the corresponding gene sets of such pairs (Fig. 6a). This revealed some sizable overlaps, in particular for vein diameter variability and central equivalent with BP measures, triglycerides, HbA1c, and BMI. Applying *PascalX* cross-GWAS analysis [52], as previously for the pairs of two vascular IDPs, we were able to identify many new candidates for pleiotropic genes (Fig. 6b and 6c). We observed that the sets of coherent genes (Fig. 6b) tended to be largest for trait pairs whose sets of associated genes already had a sizable overlap, while we found large sets of anti-coherent genes (Fig. 6c) for some trait pairs whose associated genes had no or only little overlap, notably for DBP and the ratio of the central retinal equivalents, the ratio of the vascular densities and the median arterial diameter and its variability. Also, the ratio of central retinal equivalents shared multiple anti-coherent genes with BP and BMI (see Suppl. data "Data PascalX

retina-disease cross gene scores" for a comprehensive list of genes).

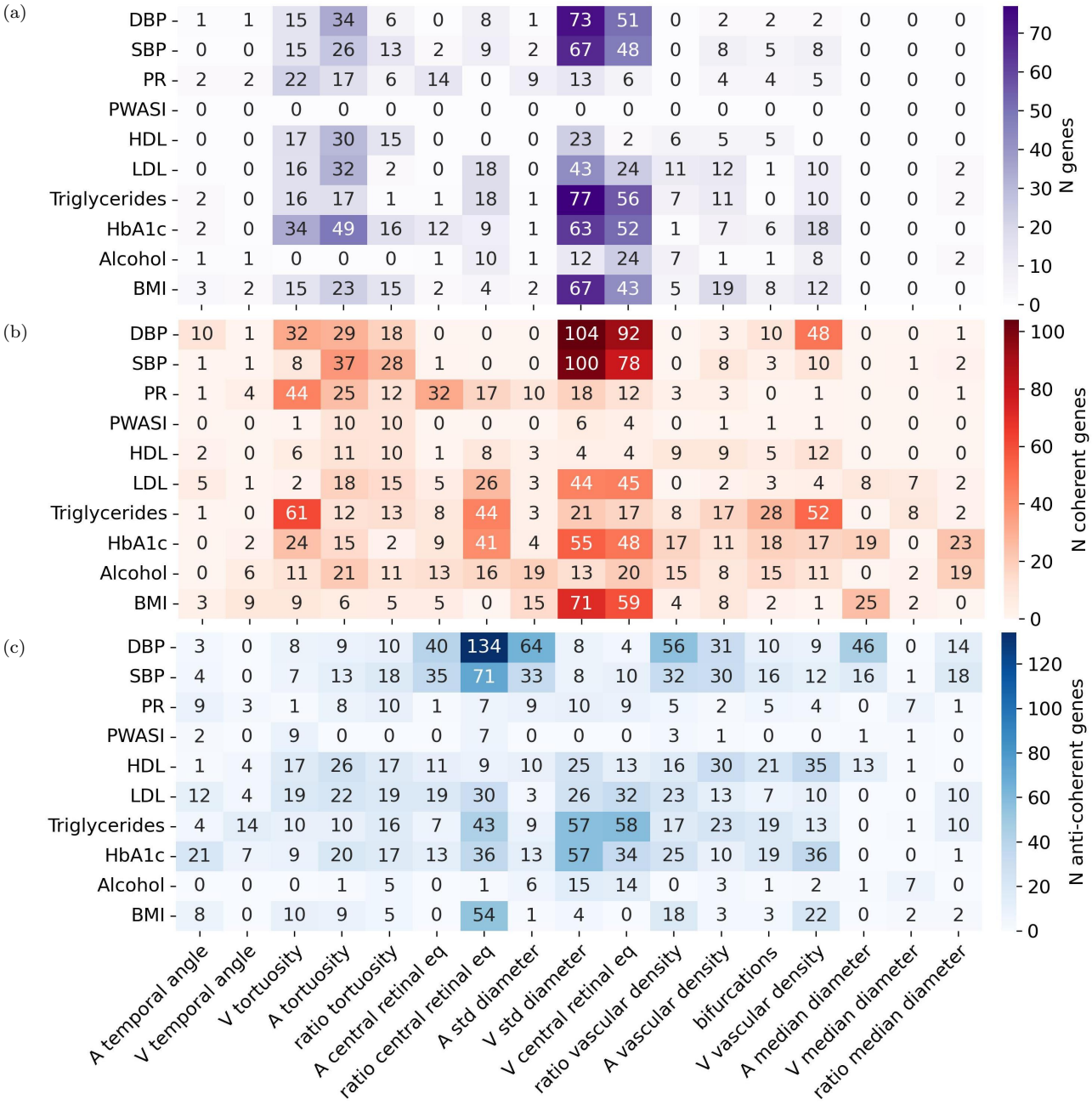


Figure 6: **a)** Gene-scoring plain intersection showing genes in common between IDPs and continuous risk factors. Each cell shows the number of intersected genes in trait pairs. **b)** and **c)** Cross-trait coherence analysis showing the number of coherent **b)** and anti-coherent **c)** genes between trait pairs. Summary statistics for risk factors were obtained from http://www.nealelab.is/uk-biobank, and significant genes were computed using *PascalX* [52].

Discussion

In this study, we established an automated analysis pipeline to extract 17 retinal vascular traits from CFIs and applied it to over 130k CFIs of close to 72k UKBB subjects. While some of these traits have been studied before individually, our work is the first to provide a common reference for these vascular parameters. This allows us to gain insights into their phenotypic and genetic relationships, and to identify shared and specific genes and pathways.

The phenotypic correlations revealed distinct clusters of vascular IDPs. One cluster includes vascular arterial and venous density, previously linked to fractal dimension [48], and the number of bifurcations. Notably, the number of bifurcations, which is challenging to identify, can be reliably estimated using vascular densities. Interestingly, diameter variabilities showed stronger correlations with central retinal equivalents (of the same vessel type) than with the median diameters, suggesting that larger vessels dominate diameter variability. Indeed, median diameters are confounded by vascular density, which is probably due to the fact that subjects with lower vascular density tend to exhibit less blood vessels of small caliber. The latter may be due to degeneration of the vasculature, but also to poorer image quality, in particular blurriness, which is impacted by corneal opacity. More work is needed to disentangle the different distributions to establish robust diameter measures for blood vessels not proximal to the OD. Finally, the temporal angles exhibited little correlation with other vascular IDPs, indicating that they may be influenced by non-vascular factors, such as eye anatomy.

The genetic correlation between IDPs largely mirrored their phenotypic correlation, supporting Cheverud's conjecture which states that phenotypic correlation can be used as a proxy for genetic correlation [57]. However, in our study genetic correlations were slightly larger on average, indicating potential differential effects of environmental factors on these traits.

We observed significant variation in heritability estimates among different vascular IDPs. Tortuosity and vascular density showed comparatively high heritability, consistent with previous findings [46, 48]. Notably, venous diameter variability also exhibited high heritability, which may be partially attributed to vascular beading, severe forms of which are known to be inherited [58]. Interestingly, the heritability of temporal angles was relatively low, likely reflecting sensitivity to environmental factors, as well as measurement noise and methodological variations, such as: (1) variability in the position of the OD across images, (2) variation due to refraction that may be insufficiently corrected by spherical and cylindrical power as our covariates, (3) variability introduced by gestational age [59]. Future work could use heritability as a guide for developing more robust measures of these angles both in terms of the methodology for extracting them (see Suppl. text "Methods for phenotype extraction" for our procedure) and implementing proper corrections for potential confounders. Finally, median vessel diameters obtained the smallest heritabilities, consistent with their known dependence on environment, such as vasoconstriction induced by stress, and diseases, as well as the aforementioned confounding by image quality and age-dependent reduction in detectable small blood

vessels.

Gene analysis of individual IDPs replicated some previously identified gene associations with explicit vascular phenotypes [46, 48], as well as DL vascular phenotypes representing the latent space of an autoencoder [39]. For the latter, such genes included *CTNNB1*, *PDE3A*, *HERC2*, *PDE6G* and *WNT7B* (see Suppl. text "Comparison with previous GWAS" for complete list). Diameter variability and central retinal equivalent, particularly for veins, received numerous gene associations, providing evidence for the genetic complexity of these traits. Finally, the ratios of vascular measures, combining arterial and venous components, showed associations with genes not found in individual measures, indicating sensitivity to vessel-type specific effects (e.g. *FGF21* and *IZUMO1* were associated with the ratio of median diameter, and the genes *CUX2*, *MCF2L*, *NR5A2*, and *FAM109A* with the central retinal equivalent ratio, while none of them appeared to be associated with the IDPs contributing to these ratios). Generally, higher heritability traits tended to have more associated genes, with some exceptions like arterial diameter variability.

No single gene was significantly associated with all the vascular IDPs. The most frequently observed gene was *LINC00461*, also known as visual cortex-expressed gene (VISC), an evolutionarily conserved long non-coding RNA that produces several alternatively spliced transcripts [60]. Other frequently observed genes are related to eye diseases (*PDE6G* has been linked to Retinitis Pigmentosa, and *SIX6* to glaucoma, myopia, and retinal degeneration) or vascular processes (*FLT1* linked to hypertension and heart disease). Also, *HERC2* and *OCA2*, two neighboring genes related to pigmentation, were linked to multiple IDPs. *OCA2* is involved in the production of melanosomes, and therefore directly contributes to the formation of pigments, while mutations in *HERC2* have been found to modulate *OCA2* expression [61]. We speculate that less eye pigmentation leads to reduced protection from damaging light, which could cause global changes on the retinal surface, including vascular morphology.

Our study of phenotypic associations between vascular IDPs and disease-relevant traits confirmed some previously known associations, such as the link between vessel diameter and hypertension, likely due to alterations in vascular resistance and blood flow [3, 4, 62]. Interestingly our MR analysis provides support for a low ratio of the central retinal equivalent being causal for high BP. This is consistent with a positive (but not significant) phenotypic and genetic correlation between central retinal venular equivalent and hypertension/BP. We also observed a significant associations between IDPs capturing arterial vessel caliber and heart attack, consistent with previous findings [18, 42–44, 63, 64], and the previously observed link between tortuosity and cardiovascular risk factors [10, 65]. For a more complete discussion refer to Suppl. text "Replication of previously identified associations of retinal vascular phenotypes with diseases".

Additionally, we found diameter variability in both vessel types to be positively associated with heart attack, possibly due to the formation of plaques that enlarge the vessel while reducing lumen diameter and negatively impacting blood flow. While atherosclerosis has a similar association profile across the vascular traits, interestingly the association with venous diameter variability was the strongest and the only one passing statistical significance. This signal may be driven by artery-vein crossings and calls for further investigation.

Type-2 diabetes is known to affect the microvascular circulation of the retina resulting in a range of structural changes unique to this tissue, such as neovascularization [66], which can affect vascular density [48] and arteriolar tortuosity [67], among others. We note that the association profile of our IDPs with HbA1c, which is used to diagnose type-2 diabetes [68], is similar, but typically a bit weaker in terms of effect size and significance, suggesting that our Cox model taking into account the age of onset of the disease endpoint has more power. Given that diabetes is a known risk factor for cardiovascular disease, including angina, this could explain why many associations between type-2 diabetes and IDPs related to vessel caliber are also observed for angina and heart attack.

Age at death is the only trait, besides myopia, that is associated with the venous temporal angle. We speculate that this might be an indirect effect on this angle by gestational age [59], as it is well known that longer gestation is associated with reduced morbidity and mortality [69]. Moreover, age at death also shares some of the associations of specific diseases with our IDPs, such as a negative association with the number of bifurcations (in line with the results for hypertension and type-2 diabetes reported above), and positive associations with several vein-related parameters besides the temporal angle, including the diameter variability, the median diameter (in line with the results for stroke, type-2 diabetes, heart attack and atherosclerosis reported above). It seems plausible that the reduced lifespan for these common diseases explains these associations.

Also our analysis of genetic associations between IDPs and risk factors revealed that the strengths of genetic associations tend to mirror those of phenotypic associations. However, in this case, the trend was much weaker than for IDP pairs, and several phenotypic associations were not significant at the genetic level, including the PR and cholesterol associations. Among the genetic associations that best reflect the phenotypic ones are those with BP and BMI.

MR analysis allowed assessment of potential causal directions for the links observed in the correlation analysis. Overall the effects of CVD risk factors on our vascular IDPs tend to be stronger and more significant than the reverse, underlining the usefulness of IDPs for early diagnosis of CVD. Consistent with previous findings, we found that individuals with genetically elevated BP tend to have lower retinal vascular density [48]. In addition, we found several other IDPs being affected, suggesting that vascular remodeling can be caused by elevated BP. Also, elevated BMI may cause higher variability in venous and arterial diameter, consistent with the finding that obesity may decrease venous return of blood from the lower extremities thereby increasing the risk of chronic venous insufficiency [70]. Amongst the strongest potential reverse causal effects is the aforementioned decrease of BP due to a higher ratio of the central retinal equivalent. Interestingly, we also found that vein diameter variability may cause higher BP.

In the gene analysis, we observed that the pairs of diseases and IDPs with the highest positive correlations tended to have a greater number of shared genes, both in the intersection and coherence analyses (see

Suppl. text "LDSR genetic correlation against PascalX"). It is worth noting that this was not limited to pairs with high significance values. Besides, we again observed that the patterns differed in the analysis of anti-coherent genes. For example in the case of vein diameter variability and BMI, they predominantly had genes associated coherently, implying that the genes modulating these traits acted in the same direction. Conversely, between BMI and central equivalent ratio, it appears that the majority of genetic effects acted in the opposite direction. This is consistent with the previous finding that vein diameter variability and ratio central equivalent shared many anti-coherent genes.

While our study pushes the boundaries of analysing retinal vascular traits, it has several important limitations: First, limiting our study to data from the UKBB makes our findings specific to a population of mostly European ancestry. Second, there are some other potentially relevant vascular IDPs that we did not analyze, including branching angles, artery-vein crossings, and neovascularization. Third, our GWAS did not include the analysis of sex chromosomes or rare variants. The exclusion of these factors may have limited the scope of the study and prevented the identification of potential associations between genetic variations and the development of certain diseases or traits. Finally, the analysis of binary disease traits for genetic associations was not included in the study for methodological reasons (see Methods section Genetic association with diseases).

In summary, this study establishes a common framework for studying multiple vascular parameters of the retina. Our analysis of genes and pathways revealed remarkably low overlap suggesting largely decoupled mechanisms modulating these traits. Our disease trait associations confirmed some previously known findings and revealed many novel connections. Notably, diameter variability, especially for the veins, seems to have many new and interesting associations with multiple disease traits, including age of death, pulmonary embolism, and heart attack. While further work is needed to replicate and extend our findings, our analyses provide evidence that we start to have sufficient power for obtaining functional, as well as some initial causal insights into the genetic and disease-related processes modulating the retinal vasculature.

Methods

UK Biobank data

The UKBB is a population-based cohort of approximately 488k subjects with rich, longitudinal phenotypic data including a complete medical history, and a median 10-year follow-up [71, 72]. Standard retinal 45° color fundus images (CFIs) were captured using a Topcon Triton 3D OCT 1000. 173 814 images from 84 813 individuals, were analyzed. Genotyping was performed on Axiom arrays for a total of 805 426 markers, from which approximately 96 million genotypes were imputed. We used the subset of 15 599 830 SNPs that had been assigned an rsID. Baseline and disease information about the subjects whose images were analyzed can be found in Suppl. text "Baseline characteristics".

Image segmentation and quality control

Raw CFIs were segmented into pixel-wise segmentations of arteries and veins using the DL model LWNET [36], which had been trained on the publicly available DRIVE dataset. To detect the OD, a random set of 100 CFIs of varying quality from the UKBB were first annotated, and the resulting ground truths were then used to retrain a standard U-Net previously trained to detect the OD in various public datasets [37]. Last, branch points of vessels and vessel-segment-wise and centerlines were extracted using skeletonization, and diameters were extracted using the distance transform, both provided in ARIA [49].

A published QC method [48] to assess image quality in the UKBB was used, and the 75% highest-quality CFIs according to this method were retained for further analysis. In this method, the image quality of 1 000 CFIs was quantified by professional graders, and a CNN was then trained to imitate the graders' quality assessment. A significant negative correlation between image quality and age was observed, r=-0.21, but was not corrected. For more information refer to Suppl. text "Image segmentation". Additionally, in order to see how the threshold on the QC can affect the results refer to Suppl. text "Quality control threshold effect on results".

Phenotyping

Retinal vessel morphology was broadly phenotyped, drawing on a set of known relevant ophthalmological traits, including a few that were previously undescribed, such as diameter variability. Due to the lack of consensus definitions and methods for their measurements, their implementation can vary. Therefore, we implemented different definitions and methods for most traits, see Suppl. text "Methods for phenotype extraction" and Suppl. text "Correlation structure and heritability of extended list of retinal vascular IDPs". In this study, we focus on 17 representative traits with significant heritability and relevant disease associations. These traits were selected for independence, removing highly redundant traits with similar definitions, but some highly correlated traits when resulting from significantly different definitions. The mean between left and right eye measurements was taken whenever measurements from both eyes passed quality control, otherwise, the single eye measurement was used. Measurements from only the first measured time point were used whenever measurements from multiple time points were present. Cohen's d, $\frac{\overline{r^{(g)}}-\overline{r^{(p)}}}{\sqrt{\frac{SD^{(g)2}+SD^{(p)2}}{2}}}$, was used to quantify the mean difference between genetic and phenotypic correlations.

We used two validation procedures: (1) visual inspection of random images in the UKBB to identify and refine potential confounders affecting certain phenotypes, and (2) the use of the DRIVE dataset as a proxy to assess performance in other datasets with similar characteristics, using ground truth on it for the number of bifurcations and temporal angles (Suppl. text "Validation of IDPs").

Correction for covariates

IDPs were corrected for sex, age, age-squared, sex-by-age, sex-by-age-squared, spherical power, spherical power-squared, cylindrical power, cylindrical power-squared, instance, assessment center, genotype measurement batch, and genomic PCs 1-20. Their effects on each trait are visualized in Suppl. text "Covariate effects". For GWAS, raw phenotypes were transformed with the rank-based inverse normal transformation (rb-INT) before correction.

Disease association

The list of diseases analyzed includes vascular-related diseases, eye-related, risk factors, mortality, and other conditions previously found to be associated with the retina vascular system (see Suppl. text "Previous associations with diseases"). The disease data were collected from the UKBB, and the official data field identifier corresponding with each disease used can be found in the Suppl. text "Disease list and corresponding UKBB datafield IDPs", and in Suppl. text "Data Disease information".

Different regression models were employed based on the nature of the diseases. For continuous risk factors, linear regression was used to estimate standardized effects using ordinary least squares, OLS, a regression method from the 'statsmodels.formula.api' library in Python 3.8.13. For binary and categorical disease traits, logistic regression was applied, the logit function from the 'statsmodels.formula.api' library.

Prior to conducting the regression analyses, a preprocessing step was performed to address potential confounding effects. Covariates were regressed out of retinal IDPs and the obtained residuals were then used as regressors in the linear/logistic regression analyses. For Cox models, the covariates were added again to the models to adjust for potential non-linear effects.

The regression models were fitted using the adjusted independent variables, and the estimates of regression coefficients, betas, and their corresponding standard errors, std, or odds ratios were obtained. Significance testing was conducted using p-values derived from the regression models, with the significance level set at a predetermined alpha threshold. In order to determine the significance of the regression results, p-values were computed and compared to predefined alpha thresholds (0.05 and 0.001), divided by the total number of tests conducted (i.e., the number of independent variables multiplied by the number of diseases analyzed).

In cases where time-to-event data was available and accurate diagnosis was feasible, multivariate Cox proportional hazards regression was utilized. This approach allowed us to estimate hazard ratios for diseases with time-to-event outcomes and reliable diagnoses. 'coxph' from the R packages 'survival' and 'survminer' were used with default parameters.

Genome-wide analyses

GWAS was performed using BGENIE [72]. SNP-wise heritabilities and genetic correlations between IDPs were derived using LDSR [50]. Gene and pathway scores were computed using *PascalX* [52]. Both protein-

coding genes and lincRNAs were scored using the novel, approximate "saddle" method, taking into account all SNPs within a 50kb window around each gene. All pathways available in MSigDB v7.2 were scored using *PascalX*' ranking mode, fusing and rescoring any co-occurring genes less than 100kb apart. *PascalX* requires LD structure to accurately compute gene scores, which in our analyses was provided with the UK10K (hg19) reference panel. Gene-level cross-GWAS coherence test between IDP pairs and between IDPs and risk factors was computed using the *PascalX* cross-scoring *zsum* method, testing for both coherence and anti-coherence of GWAS signals. Variants with a minor allele frequency of at least 0.001 were considered. Correction for bias due to sample overlap was done using the intercept from pairwise LDSR genetic correlation. Significance threshold was set at 0.05 divided by the number of tested genes. Top hits GWAS summary statistics can be found in Suppl. data "Data GWAS top hits".

Genetic association with diseases

For the genetic correlation between IDPs and diseases, the summary statistics of the diseases were obtained from the Neale lab. LDSR format summary statistics were used nealelab-ldsc-sumstat-files). LDSR was computed for the diseases and the IDPs. We limited ourselves to diseases that were categorized as 'High confidence' by the Neale lab (column 'Neale_ldsr' in Suppl. text "Data Disease information").

Regarding the genes shared between IDPs and diseases, we used the GWAS format summary statistics for diseases nealelab-sumstat-files. However, in this case, we applied some additional filters deleting values in rsid that do not start with rs, delete the p-values with nans, and low confidence variants (Suppl. data "Data Significant genes per continuous diseases").

No binary disease was included in the genetic associations with diseases since their results require additional caution. Firstly, the GWAS summary statistics from the Neale lab used linear regression for all the disease traits, which is not ideal for discrete ones. For this reason, for many binary diseases that we used for the phenotypic correlation, their corresponding GWAS results were categorized by the Neale lab as having a low confidence level. Furthermore, both LDSR and *PascalX* are not tuned to work as well for unbalanced binary trait summary statistics, which makes their results for the binary (especially unbalanced) data, less reliable. It should be noted that the variants used for the GWAS analysis of our IDPs and those used for the risk factors are not exactly the same.

Mendelian Randomisation

To perform the bidirectional two-sample MR analyses we used the TwoSamplesMR package in R [55]. We only considered independent ($r_2 < 0.001$) SNPs that are all significantly associated with the exposure. We used UK10K reference panel to estimate the LD among these SNPs and prune them [73]. As selection threshold, we adopted the commonly used $p < 5{\times}10^{-8}$. In addition, in cases where the number of instruments was very low($n_{SNPS} < 6$) the MR analysis was repeated with a more liberal ($p < 10^{-6}$) selection threshold. Due to the different sample sizes, the number of instruments (genome-wide significant SNPs) was on average

higher when risk factors were used as exposure.

To assess instrument strength, we computed the F-statistic from the regression of the exposure on the instruments [74]. The parameters required to test the F-statistics are R^2 (explained variance of genetic instruments on exposure), N (sample size of GWAS for exposure), and k (number of instruments). F is then defined as, $F = \frac{R^2 \times (N-k-1)}{(1-R^2) \times k}$. A rule of thumb is that the F statistic should be greater than 10 for strong instruments. For all the forward and reverse MR pairs, the F-statistic values were > 50 suggesting that the selected SNPs were suitable instruments.

Causal estimates were based on the inverse variance-weighted method [56]. The TwoSamplesMR package also applies a sensitivity analysis to evaluate the presence of palindromic SNPs, pleiotropy, and heterogeneity which might influence the results of the study.

Since the number of UKBB subjects for which we extracted IDPs was much smaller than the number of UKBB subjects used to study disease traits, our analysis can still be considered a two-sample MR setup and potential bias due to sample overlap is expected to be small and in direction of the null [75, 76]. Importantly, considering the fact that an asymmetry in power due to different sample size of IDP and nIDP GWAS exists, MR analyses with IDPs as outcome generally have higher power than MR analyses that involve IDPs as exposure.

Data and code availability

The genetic and phenotypic UKB data are available upon application to the UKBB, https://www.ukbiobank.ac.uk. GWAS summary statistics will be available on the GWAS catalog following publication. Code will be made available in this public Github repository following publication: https://github.com/BergmannLab/retina-phenotypes.

Author contributions

SOV, MJB, and SB designed this study. MJB performed QC and image preprocessing. SOV extracted the number of bifurcations, temporal angles, central retinal equivalents, and diameter variability. MJB extracted OD positions, vascular density, and fractal dimension. MT extracted median diameter and tortuosity. MJB performed postprocessing (covariate correction and normalization). GWAS was performed by MJB and SOV. MJB estimated heritability and cross-trait correlations. SOV and MJB performed gene and pathway analysis. Gene-level cross-trait analysis was done by OT. Linear and logistic regressions were done by SOV and OT, and Cox model analyses by MJB. Cross-trait LDSR and gene analysis between retina traits and diseases were done by SOV and OT. FH helped SOV with the validation of the angles and bifurcation traits. II performed MR analysis. SOV, MJB, OT, II, MT, and SB wrote the manuscript. All other authors contributed to the writing.

Funding

This work was supported by the Swiss National Science Foundation grant no. CRSII5_209510.
Conflict of Interest statement: None declared.

References

[1] M Kamran Ikram, Yi Ting Ong, Carol Y Cheung, and Tien Y Wong. Retinal vascular caliber measurements: clinical significance, current knowledge and future perspectives. *Ophthalmologica*, 229(3): 125–136, 2013.

[2] Sara B Seidelmann, Brian Claggett, Paco E Bravo, Ankur Gupta, Hoshang Farhad, Barbara E Klein, Ronald Klein, Marcelo Di Carli, and Scott D Solomon. Retinal vessel calibers in predicting long-term cardiovascular outcomes: the atherosclerosis risk in communities study. *Circulation*, 134(18):1328–1338, 2016.

[3] Ryo Kawasaki, Ning Cheung, Jie Jin Wang, Ronald Klein, Barbara EK Klein, Mary Frances Cotch, A Richey Sharrett, Steven Shea, FM Amirul Islam, and Tien Y Wong. Retinal vessel diameters and risk of hypertension: the multiethnic study of atherosclerosis. *Journal of hypertension*, 27(12):2386, 2009.

[4] M Kamran Ikram, Frank Jan De Jong, Ewoud J Van Dijk, Niels D Prins, Albert Hofman, Monique MB Breteler, and Paulus TVM De Jong. Retinal vessel diameters and cerebral small vessel disease: the rotterdam scan study. *Brain*, 129(1):182–188, 2006.

[5] Raviv Allon, Michael Aronov, Michael Belkin, Elad Maor, Michael Shechter, and Ido Didi Fabian. Retinal microvascular signs as screening and prognostic factors for cardiac disease: a systematic review of current evidence. *The American Journal of Medicine*, 134(1):36–47, 2021.

[6] Gerald Liew, Paul Mitchell, Elena Rochtchina, Tien Yin Wong, Wynne Hsu, Mong Li Lee, Alan Wainwright, and Jie Jin Wang. Fractal analysis of retinal microvasculature and coronary heart disease mortality. *European heart journal*, 32(4):422–429, 2011.

[7] Maximilian WM Wintergerst, Peyman Falahat, Frank G Holz, Christian Schaefer, Nadjib Schahab, and Robert Finger. Retinal vasculature assessed by octa in peripheral arterial disease. *Investigative Ophthalmology & Visual Science*, 61(7):3203–3203, 2020.

[8] Wayne Smith, Jie Jin Wang, Tien Yin Wong, Elena Rochtchina, Ronald Klein, Stephen R Leeder, and Paul Mitchell. Retinal arteriolar narrowing is associated with 5-year incident severe hypertension: the blue mountains eye study. *Hypertension*, 44(4):442–447, 2004.

[9] Tien Wong and Paul Mitchell. The eye in hypertension. *The Lancet*, 369(9559):425–435, 2007.

[10] Carol Yim-lui Cheung, Yingfeng Zheng, Wynne Hsu, Mong Li Lee, Qiangfeng Peter Lau, Paul Mitchell, Jie Jin Wang, Ronald Klein, and Tien Yin Wong. Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors. *Ophthalmology*, 118(5):812–818, 2011.

[11] Andrea Grosso, Franco Veglio, Massimo Porta, FM Grignolo, and TY Wong. Hypertensive retinopathy revisited: some answers, more questions. *British Journal of Ophthalmology*, 89(12):1646–1654, 2005.

[12] Tien Yin Wong, Anoop Shankar, Ronald Klein, Barbara EK Klein, and Larry D Hubbard. Prospective cohort study of retinal vessel diameters and risk of hypertension. *bmj*, 329(7457):79, 2004.

[13] Tien Yin Wong, Ronald Klein, Barbara EK Klein, Stacy M Meuer, and Larry D Hubbard. Retinal vessel diameters and their associations with age and blood pressure. *Investigative ophthalmology & visual science*, 44(11):4644–4650, 2003.

[14] SB Dimmitt, SM Eames, P Gosling, JNW West, JM Gibson, and WA Littler. Usefulness of ophthalmoscopy in mild to moderate hypertension. *The Lancet*, 333(8647):1103–1106, 1989.

[15] Harry Leung, Jie Jin Wang, Elena Rochtchina, Tien Y Wong, Ronald Klein, and Paul Mitchell. Impact of current and past blood pressure on retinal arteriolar diameter in an older population. *Journal of hypertension*, 22(8):1543–1549, 2004.

[16] Jie Jin Wang, Paul Mitchell, Harry Leung, Elena Rochtchina, Tien Yin Wong, and Ronald Klein. Hypertensive retinal vessel wall signs in a general older population: the blue mountains eye study. *Hypertension*, 42(4):534–541, 2003.

[17] Tien Yin Wong, Ronald Klein, A Richey Sharrett, Bruce B Duncan, David J Couper, Barbara EK Klein, Larry D Hubbard, and F Javier Nieto. Retinal arteriolar diameter and risk for hypertension. *Annals of internal medicine*, 140(4):248–255, 2004.

[18] Tien Yin Wong, Ronald Klein, A Richey Sharrett, Bruce B Duncan, David J Couper, James M Tielsch, Barbara EK Klein, and Larry D Hubbard. Retinal arteriolar narrowing and risk of coronary heart disease in men and women: the atherosclerosis risk in communities study. *Jama*, 287(9):1153–1159, 2002.

[19] A Richey Sharrett, Larry D Hubbard, Lawton S Cooper, Paul D Sorlie, Rosemary J Brothers, F Javier Nieto, Joan L Pinsky, and Ronald Klein. Retinal arteriolar diameters and elevated blood pressure: the atherosclerosis risk in communities study. *American journal of epidemiology*, 150(3):263–270, 1999.

[20] SCY Woo, GYH Lip, and PL Lip. Associations of retinal artery occlusion and retinal vein occlusion to mortality, stroke, and myocardial infarction: a systematic review. *Eye*, 30(8):1031–1038, 2016.

[21] Tyler Hyungtaek Rim, John Seungsoo Han, Jaewon Oh, Dong Wook Kim, Seok-Min Kang, and Eun Jee Chung. Retinal vein occlusion and the risk of acute myocardial infarction development: a 12-year nationwide cohort study. *Scientific reports*, 6(1):1–9, 2016.

[22] Danielle L Weiler, Carla B Engelke, Anna LO Moore, and Wendy W Harrison. Arteriole tortuosity associated with diabetic retinopathy and cholesterol. *Optometry and Vision Science*, 92(3):384–391, 2015.

[23] Varun Gulshan, Lily Peng, Marc Coram, Martin C Stumpe, Derek Wu, Arunachalam Narayanaswamy, Subhashini Venugopalan, Kasumi Widner, Tom Madams, Jorge Cuadros, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 316(22):2402–2410, 2016.

[24] Muthu Rama Krishnan Mookiah, U Rajendra Acharya, Hamido Fujita, Jen Hong Tan, Chua Kuang Chua, Sulatha V Bhandary, Augustinus Laude, and Louis Tong. Application of different imaging modalities for diagnosis of diabetic macular edema: a review. *Computers in biology and medicine*, 66: 295–315, 2015.

[25] Jie J Wang, Bronwen Taylor, Tien Y Wong, Brian Chua, Elena Rochtchina, Ronald Klein, and Paul Mitchell. Retinal vessel diameters and obesity: a population-based study in older persons. *Obesity*, 14 (2):206–214, 2006.

[26] Dora Elisa Alvarado-Carrillo, Iván Cruz-Aceves, Martha Alicia Hernández-González, and Luis Miguel López-Montero. Robust detection and modeling of the major temporal arcade in retinal fundus images. *Mathematics*, 10(8):1334, 2022.

[27] Faraz Oloumi, Rangaraj M Rangayyan, and Anna L Ells. A graphical user interface for measurement of temporal arcade angles in fundus images of the retina. In *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4. IEEE, 2012.

[28] Álvaro S Hervella, José Rouco, Jorge Novo, Manuel G Penedo, and Marcos Ortega. Deep multi-instance heatmap regression for the detection of retinal vessel crossings and bifurcations in eye fundus images. *Computer Methods and Programs in Biomedicine*, 186:105201, 2020.

[29] Harry Pratt, Bryan M Williams, Jae Yee Ku, Charles Vas, Emma McCann, Baidaa Al-Bander, Yitian Zhao, Frans Coenen, and Yalin Zheng. Automatic detection and distinction of retinal vessel bifurcations and crossings in colour fundus photography. *Journal of Imaging*, 4(1):4, 2018.

[30] Xiayu Xu, Joseph M Reinhardt, Qiao Hu, Benjamin Bakall, Paul S Tlucek, Geir Bertelsen, and Michael D Abramoff. Retinal vessel width measurement at branchings using an improved electric field theory-based graph approach. *PloS one*, 7(11):e49668, 2012.

[31] Alauddin Bhuiyan, Baikunth Nath, and Kotagiri Ramamohanarao. Detection and classification of bifurcation and branch points on retinal vascular network. In *2012 International Conference on Digital Image Computing Techniques and Applications (DICTA)*, pages 1–8. IEEE, 2012.

[32] M Elena Martinez-Perez, Alun D Hughes, Alice V Stanton, Simon A Thom, Neil Chapman, Anil A Bharath, and Kim H Parker. Geometrical and morphological analysis of vascular branches from fundus retinal images. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2000: Third International Conference, Pittsburgh, PA, USA, October 11-14, 2000. Proceedings 3*, pages 756–765. Springer, 2000.

[33] Adria Perez-Rovira, T MacGillivray, Emanuele Trucco, KS Chin, K Zutis, C Lupascu, Domenico Tegolo, Andrea Giachetti, Peter J Wilson, A Doney, et al. Vampire: vessel assessment and measurement platform for images of the retina. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 3391–3394. IEEE, 2011.

[34] Behdad Dashtbozorg, Ana Maria Mendonca, Susana Penas, and Aurelio Campilho. Retinacad, a system for the assessment of retinal vascular changes. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6328–6331. IEEE, 2014.

[35] Jonathan Fhima, Jan Van Eijgen, Ingeborg Stalmans, Yevgeniy Men, Moti Freiman, and Joachim A Behar. Pvbm: A python vasculature biomarker toolbox based on retinal blood vessel segmentation. In *Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part III*, pages 296–312. Springer, 2023.

[36] Adrian Galdran, André Anjos, José Dolz, Hadi Chakor, Hervé Lombaert, and Ismail Ben Ayed. State-of-the-art retinal vessel segmentation with minimalistic models. *Scientific Reports*, 12(1):1–13, 2022.

[37] Artem Sevastopolsky. Optic disc and cup segmentation methods for glaucoma detection with modification of u-net convolutional neural network. *Pattern Recognition and Image Analysis*, 27(3):618–624, 2017.

[38] Harry Pratt, Bryan M Williams, Jae Yee Ku, Charles Vas, Emma McCann, Baidaa Al-Bander, Yitian Zhao, Frans Coenen, and Yalin Zheng. Automatic detection and distinction of retinal vessel bifurcations and crossings in colour fundus photography. *Journal of Imaging*, 4(1):4, 2017.

[39] Ziqian Xie, Tao Zhang, Sangbae Kim, Jiaxiong Lu, Wanheng Zhang, Cheng-Hui Lin, Man-Ru Wu, Alexander Davis, Roomasa Channa, Luca Giancarlo, et al. igwas: image-based genome-wide association of self-supervised deep phenotyping of human medical images. *medRxiv*, pages 2022–05, 2022.

[40] Ryan Poplin, Avinash V Varadarajan, Katy Blumer, Yun Liu, Michael V McConnell, Greg S Corrado, Lily Peng, and Dale R Webster. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering*, 2(3):158–164, 2018.

[41] Tao Li, Wang Bo, Chunyu Hu, Hong Kang, Hanruo Liu, Kai Wang, and Huazhu Fu. Applications of deep learning in fundus images: A review. *Medical Image Analysis*, 69:101971, 2021.

[42] Xueling Sim, Richard A Jensen, M Kamran Ikram, Mary Frances Cotch, Xiaohui Li, Stuart MacGregor, Jing Xie, Albert Vernon Smith, Eric Boerwinkle, Paul Mitchell, et al. Genetic loci for retinal arteriolar microcirculation. *PloS one*, 8(6):e65804, 2013.

[43] Richard A Jensen, Xueling Sim, Albert Vernon Smith, Xiaohui Li, Jóhanna Jakobsdóttir, Ching-Yu Cheng, Jennifer A Brody, Mary Frances Cotch, Barbara Mcknight, Ronald Klein, et al. Novel genetic loci associated with retinal microvascular diameter. *Circulation: Cardiovascular Genetics*, 9(1):45–54, 2016.

[44] M Kamran Ikram, Sim Xueling, Richard A Jensen, Mary Frances Cotch, Alex W Hewitt, M Arfan Ikram, Jie Jin Wang, Ronald Klein, Barbara EK Klein, Monique MB Breteler, et al. Four novel loci (19q13, 6q24, 12q24, and 5q14) influence the microcirculation in vivo. *PLoS genetics*, 6(10):e1001184, 2010.

[45] Abirami Veluchamy, Lucia Ballerini, Veronique Vitart, Katharina E Schraut, Mirna Kirin, Harry Campbell, Peter K Joshi, Devanjali Relan, Sarah Harris, Ellie Brown, et al. Novel genetic locus influencing retinal venular tortuosity is also associated with risk of coronary artery disease. *Arteriosclerosis, thrombosis, and vascular biology*, 39(12):2542–2552, 2019.

[46] Mattia Tomasoni, Michael Johannes Beyeler, Sofia Ortin Vela, Ninon Mounier, Eleonora Porcu, Tanguy Corre, Daniel Krefl, Alexander Luke Button, Hana Abouzeid, Konstantinidis Lazaros, et al. Genome-wide association studies of retinal vessel tortuosity identify numerous novel loci revealing genes and pathways associated with ocular and cardiometabolic diseases. *Ophthalmology Science*, page 100288, 2023.

[47] Xiaofan Jiang, Pirro G Hysi, Anthony P Khawaja, Omar A Mahroo, Zihe Xu, Christopher J Hammond, Paul J Foster, Roshan A Welikala, Sarah A Barman, Peter H Whincup, et al. Gwas on retinal vasculometry phenotypes. *Plos Genetics*, 19(2):e1010583, 2023.

[48] Seyedeh Maryam Zekavat, Vineet K Raghu, Mark Trinder, Yixuan Ye, Satoshi Koyama, Michael C Honigberg, Zhi Yu, Akhil Pampana, Sarah Urbut, Sara Haidermota, Declan P O'Regan, Hongyu Zhao, Patrick T Ellinor, Ayellet V Segrè, Tobias Elze, Janey L Wiggs, James Martone, Ron A Adelman, Nazlee Zebardast, Lucian Del Priore, Jay C Wang, and Pradeep Natarajan. Deep learning of the retina enables phenome- and genome-wide analyses of the microvasculature. *Circulation*, 11 2021. ISSN 1524-4539. doi: 10.1161/CIRCULATIONAHA.121.057709. URL http://www.ncbi.nlm.nih.gov/pubmed/34743558.

[49] Peter Bankhead, C Norman Scholfield, J Graham McGeown, and Tim M Curtis. Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PloS one*, 7(3):e32435, 2012.

[50] Brendan K Bulik-Sullivan, Po-Ru Loh, Hilary K Finucane, Stephan Ripke, Jian Yang, Nick Patterson, Mark J Daly, Alkes L Price, and Benjamin M Neale. Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nature genetics*, 47(3):291–295, 2015.

[51] David Lamparter, Daniel Marbach, Rico Rueedi, Zoltán Kutalik, and Sven Bergmann. Fast and rigorous computation of gene and pathway scores from snp-based summary statistics. 2016. doi: 10.1371/journal.pcbi.1004714. URL http://www2.unil.ch/cbg/index.php?title=PascalTestDatahttp://www2.unil.ch/cbg/index.php?title=Pascal.

[52] Daniel Krefl and Sven Bergmann. Cross-gwas coherence test at the gene and pathway level. *PLoS Computational Biology*, 18(9):e1010517, 2022.

[53] Stephen Burgess, Dylan S Small, and Simon G Thompson. A review of instrumental variable estimators for mendelian randomization. *Statistical methods in medical research*, 26(5):2333–2355, 2017.

[54] George Davey Smith and Shah Ebrahim. 'mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *International journal of epidemiology*, 32(1):1–22, 2003.

[55] Danielle Rasooly and Chirag J Patel. Conducting a reproducible mendelian randomization analysis using the r analytic statistical environment. *Current protocols in human genetics*, 101(1):e82, 2019.

[56] Stephen Burgess, Adam Butterworth, and Simon G Thompson. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genetic epidemiology*, 37(7):658–665, 2013.

[57] Sebastian M Sodini, Kathryn E Kemper, Naomi R Wray, and Maciej Trzaskowski. Comparison of genotypic and phenotypic correlations: Cheverud's conjecture in humans. *Genetics*, 209(3):941–948, 2018.

[58] Travis A Meredith. Inherited retinal venous beading. *Archives of Ophthalmology*, 105(7):949–953, 1987.

[59] C Wilson, M Theodorou, KD Cocker, and AR Fielder. The temporal retinal vessel angle and infants born preterm. *British journal of ophthalmology*, 90(6):702–704, 2006.

[60] Peter L Oliver, Rebecca A Chodroff, Amrit Gosal, Benjamin Edwards, Amanda FP Cheung, Julio Gomez-Rodriguez, Gene Elliot, Lisa J Garrett, Tom Lickiss, Francis Szele, et al. Disruption of visc-2, a brain-expressed conserved long noncoding rna, does not elicit an overt anatomical or behavioral phenotype. *Cerebral Cortex*, 25(10):3572–3585, 2015.

[61] Manfred Kayser, Fan Liu, A. Cecile J.W. Janssens, Fernando Rivadeneira, Oscar Lao, Kate van Duijn, Mark Vermeulen, Pascal Arp, Mila M. Jhamai, Wilfred F.J. van IJcken, Johan T. den Dunnen, Simon Heath, Diana Zelenika, Dominiek D.G. Despriet, Caroline C.W. Klaver, Johannes R. Vingerling, Paulus T.V.M. de Jong, Albert Hofman, Yurii S. Aulchenko, Andre G. Uitterlinden, Ben A. Oostra, and Cornelia M. van Duijn. Three genome-wide association studies and a linkage analysis identify herc2 as a human iris color gene. *American Journal of Human Genetics*, 82:411, 2 2008. ISSN 00029297. doi: 10.1016/J.AJHG.2007.10.003. URL /pmc/articles/PMC2427174//pmc/articles/PMC2427174/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC2427174/.

[62] M Kamran Ikram, Frank Jan de Jong, Johannes R Vingerling, Jacqueline CM Witteman, Albert Hofman, Monique MB Breteler, and Paulus TVM de Jong. Are retinal arteriolar or venular diameters associated with markers for cardiovascular disorders? the rotterdam study. *Investigative ophthalmology & visual science*, 45(7):2129–2134, 2004.

[63] Thomas J MacGillivray, James R Cameron, Qiuli Zhang, Ahmed El-Medany, Carl Mulholland, Ziyan Sheng, Bal Dhillon, Fergus N Doubal, Paul J Foster, Emmanuel Trucco, et al. Suitability of uk biobank retinal images for automatic analysis of morphometric properties of the vasculature. *PLoS One*, 10(5): e0127914, 2015.

[64] Josef Flammer, Katarzyna Konieczka, Rosa M Bruno, Agostino Virdis, Andreas J Flammer, and Stefano Taddei. The eye and the heart. *European heart journal*, 34(17):1270–1278, 2013.

[65] WT Longstreth Jr, Emily K Marino Larsen, Ronald Klein, Tien Yin Wong, A Richey Sharrett, David Lefkowitz, and Teri A Manolio. Associations between findings on cranial magnetic resonance imaging and retinal photography in the elderly: the cardiovascular health study. *American journal of epidemiology*, 165(1):78–84, 2007.

[66] Eui Seok Shin, Christine M Sorenson, and Nader Sheibani. Diabetes and retinal vascular dysfunction. *Journal of ophthalmic & vision research*, 9(3):362, 2014.

[67] MB Sasongko, TY Wong, TT Nguyen, CY Cheung, JE Shaw, and JJ Wang. Retinal vascular tortuosity in persons with diabetes and diabetic retinopathy. *Diabetologia*, 54:2409–2416, 2011.

[68] Shariq I Sherwani, Haseeb A Khan, Aishah Ekhzaimy, Afshan Masood, and Meena K Sakharkar. Significance of hba1c test in diagnosis and prognosis of diabetic patients. *Biomarker insights*, 11:BMI–S38440, 2016.

[69] Emma V Espel, Laura M Glynn, Curt A Sandman, and Elysia Poggi Davis. Longer gestation among children born full term influences cognitive and motor development. *PLoS One*, 9(11):e113758, 2014.

[70] Torsten Willenberg, Anette Schumacher, Beatrice Amann-Vesti, Vincenzo Jacomella, Christoph Thalhammer, Nicolas Diehm, Iris Baumgartner, and Marc Husmann. Impact of obesity on venous hemodynamics of the lower limbs. *Journal of vascular surgery*, 52(3):664–668, 2010.

[71] Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, et al. Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):e1001779, 2015.

[72] Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O'Connell, et al. The uk biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726):203–209, 2018.

[73] Jie Huang, Bryan Howie, Shane McCarthy, Yasin Memari, Klaudia Walter, Josine L Min, Petr Danecek, Giovanni Malerba, Elisabetta Trabetti, Hou-Feng Zheng, et al. Improved imputation of low-frequency and rare variants using the uk10k haplotype reference panel. *Nature communications*, 6(1):8111, 2015.

[74] Stephen Burgess, Simon G Thompson, and Crp Chd Genetics Collaboration. Avoiding bias from weak instruments in mendelian randomization studies. *International journal of epidemiology*, 40(3):755–764, 2011.

[75] Stephen Burgess, Neil M Davies, and Simon G Thompson. Bias due to participant overlap in two-sample mendelian randomization. *Genetic epidemiology*, 40(7):597–608, 2016.

[76] Bernd Taschler, Stephen M Smith, and Thomas E Nichols. Causal inference on neuroimaging data with mendelian randomisation. *NeuroImage*, 258:119385, 2022.