GWAS of Retinal Vessel Tortuosity Identifies 173 Novel Loci Revealing Genes and Pathways Associated with Vascular Pathomechanics and Cardiometabolic Diseases

Mattia Tomasoni, PhD[1,2,7,†]; Michael Johannes Beyeler, MSc[1,2,†]; Sofia Ortin Vela, MSc[1,2,†]; Ninon Mounier, PhD[2,3]; Eleonora Porcu, PhD[2,3,4]; Tanguy Corre[1,2,3]; Daniel Krefl, PhD[1,2]; Alexander Luke Button, PhD[1,2]; Hana Abouzeid, MD[5,6]; Konstantinidis Lazaros, MD[7]; Murielle Bochud, MD, PhD[3]; Reinier Schlingemann, MD, PhD[7]; Ciara Bergin, PhD[7]; Sven Bergmann, PhD[1,2,8]

[1]Dept. of Computational Biology, University of Lausanne, Lausanne, Switzerland; [2]Swiss Institute of Bioinformatics, Lausanne, Switzerland; [3]Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne, Switzerland; [4]Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland; [5]Division of Ophthalmology, Geneva University Hospitals, Switzerland; [6]Clinical Eye Research Center Memorial Adolphe de Rothschild, Geneva, Switzerland; [7]Jules-Gonin Eye Hospital, Lausanne, Switzerland; [8]Dept. of Integrative Biomedical Sciences, University of Cape Town, Cape Town, South Africa; [†]Authors contributed equally to this work.

Address for Correspondence:
Sven Bergmann, PhD
University of Lausanne
Genopode
1016 Lausanne
Switzerland
E-mail: sven.bergmann@unil.ch

Short Title: Tomasoni et al.; GWAS of Retinal Vessel Tortuosity
Count: 11□401 word

Abstract

Background: Fundus images allow for non-invasive assessment of the retinal vasculature whose features provide important information on health. Blood vessel tortuosity is a morphological feature associated with many diseases including hypertension.

Methods: We analyzed 116□639 fundus images of suitable quality from 63□662 participants from three cohorts, namely the UK Biobank (n = 62□751), SKIPOGH (n = 397), and *OphtalmoLaus* (n = 512). We used a fully automated image processing pipeline to annotate vessels and a deep learning algorithm to determine the vessel type, characterizing these subjects in terms of their median retinal vessel tortuosity specific to arteries and to veins. Tortuosity was measured by the *distance factor* (the length of a vessel segment over its chord length), as well as measures that integrate over vessel curvature. Using these measures as traits, we performed the largest genome-wide association study (GWAS) of vessel tortuosity to date. We assessed gene set enrichment using the novel high-precision statistical method *PascalX*.

Results: Higher tortuosity was significantly associated with higher incidence of angina, myocardial infarction, stroke, deep vein thrombosis, and hypertension. We identified 175 significantly associated genetic loci in the UK Biobank; 173 of these were novel and 4 replicated in our second, much smaller, meta-cohort. We estimated heritability at ~25% using linkage disequilibrium score regression. Vessel type specific GWAS revealed 114 loci for arteries and 63 for veins. Genes with significant association signals included COL4A2, ACTN4, LGALS4, LGALS7, LGALS7B, TNS1, MAP4K1, EIF3K, CAPN12, ECH1, and SYNPO2. These tortuosity genes were overexpressed in arteries and heart muscle and linked to pathways related to the structural properties of the vasculature. We demonstrated that tortuosity loci served pleiotropic functions as cardiometabolic disease variants and risk factors. Concordantly, Mendelian randomization revealed causal effects between tortuosity, BMI and LDL.

Conclusions: Several alleles associated with retinal vessel tortuosity point to a common genetic architecture of this trait with cardiovascular diseases and metabolic syndrome. Our results shed new light on the genetics of vascular diseases and their pathomechanisms and highlight how GWASs and heritability can be used to improve phenotype extraction from high-dimensional data, such as images.

Keywords: GWAS, retina, microvasculature, tortuosity, genomics, cardiometabolic disease, Mendelian randomization

Nonstandard Abbreviations and Acronyms

BMI: body mass index
BRB: blood-retina barrier
CAD: coronary artery disease
CVD: cardiovascular diseases
DBP: diastolic blood pressure
DF: distance factor
DVT: deep vein thrombosis
EC: endothelial cells
GO: gene ontology
GWAS: genome-wide association study
LD: linkage disequilibrium
LDL: low-density lipoprotein
SBP: systolic blood pressure
SMC: smooth muscle cell
SNP: single nucleotide polymorphism
VEGF: vascular endothelial growth factor

Clinical Perspective

What is new?

- We automatically estimated arterial and venous tortuosity in over 100k retinal fundus images using image analysis and deep learning.
- GWAS revealed 173 novel loci.
- Mendelian randomization showed that increased venous tortuosity reduces BMI whereas elevated LDL levels reduce the tortuosity of both arteries and veins.
- Measuring tortuosity in terms of the *distance factor*, which is sensitive to total vessel elongation, had higher heritability and more associated loci than other tortuosity measures that are sensitive to local vessel bending.

What are the clinical implications?

- Tortuosity genes were overexpressed in the aorta, tibial artery, coronary artery, and in two heart tissues.
- Higher tortuosity was associated with higher incidence of angina, myocardial infarction, stroke, deep vein thrombosis and hypertension.
- We demonstrated a shared genetic architecture between retinal tortuosity and certain diseases related to the vasculature, and the associations included several cardiometabolic disease variants and risk factors. Further research is needed to investigate the potential of the retinal vessel tortuosity as a clinically relevant biomarker for cardiovascular disease and metabolic syndrome.
- Enriched pathways include a well-known therapeutic target for ocular diseases (VEGFA-VEGFR2) affecting tissue remodeling. We highlight several transcription factors as interesting targets for further experimentation.

INTRODUCTION

Cardiovascular diseases (CVD) are the leading cause of death in developed countries[1–3] and a major societal health burden. Though several risk factors for CVD development, such as age, smoking, and hypertension, have been firmly established, the degree of importance of vascular properties as risk factors is unclear. Retinal fundus photos allow non-invasive in-vivo assessment of the vascular system of the superficial inner retina, i.e. the central and branch veins and arteries plus the venules and arterioles. These vessels are composed of tightly sealed endothelial cells (ECs), forming the inner blood-retina barrier (BRB), encased by smooth muscle cells (SMCs) form to the vessel wall [4,5]. Automatic segmentation of retinal vessels in fundus images is well established, and is commonly used in standard clinical care to screen and diagnose ocular and systemic diseases [6]. In diabetes, for example, hyperglycemia induces damage to the ECs and pericytes of the inner BRB contributing to retinal edema and hemorrhage [7].

Pathological changes in the retinal vessels often coincide with those in the microvasculature of other organs and may precede the progression of systemic vascular diseases. The retinal vasculature can provide insights into neuro-degenerative diseases, such as Alzheimer's, Parkinson's, and vascular dementia [8–12]. In addition, abnormalities in retinal parameters are of diagnostic value for systemic diseases, including increased risk of diabetes [13–15], obesity [16] and CVD [17,18] (such as stroke [19–22], coronary heart disease [23], peripheral artery disease [24], hypertension [21,25–33], atherosclerosis [19,21,34], myocardial infarction [35,36], and nephropathies [37,38]).

In recent years, genome-wide association studies (GWAS) have been used to link genes with phenotypes extracted from fundus images, such as vessel size [39,40], optic disc morphology [41,42], vascular density [43], fractal dimensions [43] and vessel tortuosity [44]. The diameter of the retinal microvasculature was associated with genes TEAD1, TSPAN10, GNB3 and OCA2 [39]. A recently published study [43] on vascular density and fractal dimensions reported 7 and 13 single nucleotide polymorphism (SNPs) associated with these traits respectively, including OCA2, MEF2C and GNB3. Retinal vessel tortuosity has been associated with SNPs that map to the genes ACTN4 and COL4A2 [44] Tortuosity of the vasculature was reported in the context of CAD [45] and connective tissue disease [46]. These results demonstrated that GWAS can reveal genes with a potential role in modulating vascular properties and related pathomechanisms, though their power has been often limited.

Here, we report the results of the largest GWAS on vessel tortuosity to date. Our study was motivated by the clinical relevance of this trait to diseases [9,13,28,46–48], as significant associations were already reported in smaller sample sizes, making further discoveries likely. We used images and genotypes from 62 751 subjects in the UK Biobank and from 397 and 512 subjects of the much smaller, yet independent, population-based cohorts SKIPOGH [49,50] and *OphtalmoLaus* [51]. We constructed an automated image analysis pipeline to extract retinal tortuosity from these data as a biomarker. We report the correlation with patient records, SNPs, genes, pathways, tissue expression, pathomechanisms, and causal effects associated with this biomarker.

METHODS

Data: genotypes, phenotypes and fundus images

The UK Biobank is a population-based cohort of approximately 488□000 subjects with rich, longitudinal phenotypic data and a median 10-year follow-up [52,53]. We analyzed 173□837 retinal

fundus images from 84□825 individuals. 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. We performed an additional quality control step by filtering out SNPs with MAF < 5×10^{-4} (which translates to an expected minimum number of ~60 individuals expected to have at least one minor allele). Finally, we applied a filtering procedure [54] to remove SNPs with imputation quality < 0.3. In addition to genomic information, we obtained phenotypic information from the patient records: type-2 diabetes, angina, myocardial infarction, deep vein thrombosis (DVT), stroke, hypertension and smoking status. Age, sex, and principal components of genotypes were used to correct for biases in the genetic associations.

We performed replication via a meta-analysis of two independent, population-based cohorts: the *Swiss Kidney Project on Genes in Hypertension* (SKIPOGH) [49,50] and *OphtalmoLaus* [51]. SKIPOGH is a family-based, cross-sectional study exploring the role of genes and kidney hemodynamics in blood pressure regulation and kidney function in the general population, comprising 1□054 genotyped individuals. 1□352 retinal fundus images were available from 518 participants. The genotyping was performed with the Illumina Omni 2.5 chip. *OphtalmoLaus* is a sub-study of *Cohorte Lausannoise* (*CoLaus*), a population-based cohort comprising 6□188 genotyped individuals. 7□252 fundus images were available from 1□015 subjects. *CoLaus* has as its objective to investigate the epidemiology and genetic determinants of CVD risk factors and metabolic syndrome: participants were phenotyped accordingly. The genotyping was performed using the 500K Affymetrix chip technology.

Automated analysis of color fundus images and quality control

We extended the software ARIA [55] to perform batch segmentation and positional annotation of blood vessels, using the default parameters [56]. We applied a quality filter based on the total length of the vasculature and on the number of vessels (see Supplemental Text 1). Roughly two out of three images passed this strict quality control (116□639 out of 173□837 in the UK Biobank). Based on ARIA's vessel annotations, we calculated a tortuosity measure known as the *distance factor* (DF) [57], defined as:

$$DF = \frac{s(C)}{chord(C)}$$

where the total vessel length, $s(C)$, is divided by the Euclidean distance between the vessel segment endpoints, $chord(C)$. DF is referred to in a recent review as the arc over chord ratio [58]. In addition to DF, we also calculated six other tortuosity phenotypes based on alternative measures using integrals over the curvature along the vessel (see Supplemental Text 2).

We phenotyped each individual by calculating median retinal tortuosities, then averaging the values derived from the left and right eye, when available (for the resulting distribution, refer to Supplemental Text 3).

Deep Learning classification of arteries and veins

We calculated pixel-wise artery and vein classifications using the Deep Learning algorithm *Little W-Net* [59]. For each vessel segment recognized by ARIA, we used the difference between pixels classified as arterial and venous as a score that was required to be positive or negative for the segment to be annotated as artery or vein, respectively. On a set of 44 images, manually annotated by an ophthalmologist, we obtained an area under the curve of 0.93 and an accuracy of 0.88. Thus, we performed vessel type classification for the entire set of retinal fundus images, computing artery- and vein-specific tortuosity values (see Supplemental Text 4).

Genome-wide association analyses

We ran genetic association studies on tortuosity of arteries, of veins, and combining both vessel types. We used BGENIE [60], applying linear regression to confounder-corrected, quantile-quantile normalized, retinal vessel tortuosity on the genotypes of the matching subjects imputed to a panel of approximately 15M genetic variants. In order to account for confounding effects [61], the following variables were provided as covariates: age, sex, PC of the genotypes (we considered only PCs with a significant correlation to tortuosity, namely 1, 2, 5, 6, 7, 8, 16, 17 and 18). We applied a Bonferroni threshold of 5×10^{-8}. A list of independent SNP was obtained by performing linkage disequilibrium (LD) pruning using the LDpair function of the R package LDlinkR [62]. Two SNPs were considered independent if they had LD $r^2<0.1$ or were more than 500K bases apart (see Supplemental Dataset 1).

Replication meta-cohort

As the SKIPOGH cohort includes subjects with a high degree of relatedness, we used the EMMAX function of the EPACTS software [63] and the kinship matrix in the model to account for family structure. We also included the recruitment center as a covariable. For the GWAS on the *OphtalmoLaus* cohort, we used the same parameters and tools as for the discovery cohort. Results from SKIPOGH and *OphtalmoLaus* were meta-analyzed using an inverse-variance weighting scheme for the respective effect sizes as well as a random-effects model (Supplemental Text 5).

To keep the multiple hypotheses testing burden low, we took a candidate gene approach. Specifically, we used the Benjamini–Hochberg procedure [64], which applies no multiple hypothesis testing correction for the top hits, and then corrects the association of the candidate locus of rank *n* for *n* tests.

Heritability estimates

We used LD Score Regression [65] to estimate the SNP-based heritability of our tortuosity measures.

Novel method for gene-based tests

We used *PascalX* [66], a novel high-precision pathway scoring algorithm that we developed building on our *Pascal* [67] tool, to aggregate SNP-wise summary statistics into gene scores using a sum of χ^2 statistics: *PascalX* takes into account LD by effectively transforming the sum of χ^2 from all SNPs within the gene window into a new basis of independent "Eigen-SNPs" corresponding to a weighted sum of χ^2 statistics. Using multiple-precision arithmetics, *PascalX* computes the corresponding null cumulative probability distribution to essentially arbitrary precision, while other tools usually only approximate the underlying distribution. We thus computed p-values up to a precision of 10^{-100}, allowing for accurate scoring of genes with contributions from extremely significant SNPs, which become increasingly frequent in highly powered GWASs such as this one.

We used the following configurations: We computed gene scores from SNPs lying within a window of 50kb before the transcription start site and 50kb after the transcript end. The annotation of the gene positions was based on the Genome Reference Consortium Human genome build 37 (GRCh37/hg19) downloaded from the Ensembl biomart [68]; we considered only protein-coding and lincRNA genes. The reference panel from the UK10K project [69] was used to estimate the SNP-SNP correlations (LD effects). *PascalX* uncovered 265 significant genes (after Bonferroni correction for 25 489 gene-based tests $p < 0.05 / 25489 \simeq 2.0\times10^{-6}$).

Gene set enrichment

We used *PascalX* [66] to compute gene set enrichment scores based on ranking derived from the gene-based tests. As a large number of genes have inflated p-values in highly powered GWASs, this ranking approach was more conservative. We first computed scores for 2□868 canonical pathways (BioCarta, KEGG, PID, Reactome, and WikiPathways), then extended our analysis to the 31□120 pathways in MSigDB (v7.2)□ [70]. To adjust for statistical dependence and co-expression, genes that are less than 100kb apart were "fused" (i.e. considered as single entities termed "fusion genes" [67]).

Tissue-wide gene expression analysis

We performed tissue-wide gene expression analysis using *PascalX* [66] on the whole GTEx [71] (v8) dataset, comprising 54 tissues. We defined gene sets based on the significant genes from each of our three GWAS on DF tortuosity (artery, vein and combined). *PascalX* was used to perform an enrichment analysis that indicated whether these sets were over-expressed in any particular tissue. *PascalX* corrected for the co-expression of gene sub-clusters within each gene set by merging nearby genes to fusion genes. We computed the fusion genes expression values in transcripts per kilobase million from the raw read counts. These values values were made uniform via ranking, transformed to $\square^2$-distributed random variables, summed, and tested against a $\square^2$ distribution with as many degrees of freedom as there were "fusion genes" in each set. We applied a Bonferroni threshold: $p = 0.05 / 54 = 9.2\times10^{-4}$.

Shared genetic architecture with disease

We computed the overlap between DF tortuosity SNPs (from the combined-vessel GWAS) and disease-related SNPs. To this end, we first identified which of the independent SNPs in the combined-vessel GWAS were listed in the GWAS Catalog [72]. We then extended this analysis by considering DF tortuosity SNPs in LD ($r^2 > 0.8$) with disease-related SNPs in the GWAS Catalog.

Mendelian randomisation analysis

We performed two-sample bidirectional Mendelian randomisation [73,74] to search for evidence of causal effects between DF tortuosity (from the combined-vessel GWAS) and the following traits: body mass index (BMI), coronary artery disease (CAD), systolic blood pressure (SBP), and lipid traits, namely high-density lipoprotein, low-density lipoprotein (LDL), total cholesterol, and triglycerides. For each trait, we used independent ($r^2 < 0.01$) significant ($P < 5\times10^{-8}$) SNPs as instrumental variables. All summary statistics (estimated univariate effect size and standard error) originated from the most recent meta-analyses (not including UK Biobank individuals) and were downloaded from the publicly available NIH Genome-wide Repository of Associations Between SNPs and Phenotypes□[75]. We only used SNPs on autosomal chromosomes available in the UK10K reference panel□[69], which allowed us to estimate the LD among these SNPs and prune them. We removed strand ambiguous SNPs. Causal estimates were based on the inverse variance weighted method [76] and calculated using the Mendelian randomisation R package [77].

Code Availability

Code for all computations is available on request to the corresponding author.

RESULTS

Baseline characteristics and tortuosity quantification

Following quality control measures, we analyzed 116□639 images from 62□751 subjects of the UK Biobank (mean±SD age = 56±8 years; 35□098 females at birth [54%]; 4□618 smokers [7%]). We analyzed 1□352 images from 379 subjects of the SKIPOGH cohort (mean±SD age = 48±16 years; 211 females [53%]; 107 smokers [27%]). We analyzed 7□254 images from 512 subjects of the *OphtalmoLaus* cohort (mean±SD age = 51±10 years; 270 females [53%]). Baseline characteristics and disease prevalence are presented in Supplemental Text 6. For an overview of our pipeline see Figure 1.

The distributions of DF tortuosity were similar across cohorts: long-tailed, left-skewed, with means ranging from 1.030 (UK Biobank) to 1.034 (*OphtalmoLaus*). DF was higher in the elderly population (Cohen's d = 0.49, $p = 1\times10^{-195}$), while sex had limited impact (Cohen's d = 0.049, $p = 9\times10^{-10}$). Overall, DF was higher in veins (Cohen's d = 0.13, $p = 9\times10^{-142}$). For details about the stratified analysis of the DF phenotype in the UK Biobank see Supplemental Text 3.

We extracted six additional tortuosity measures based on alternative mathematical definitions. Correlations analysis and dimensionality reduction in terms of principle components showed that the DF is most similar to the path integral of the squared curvature (τ_3) and least similar to the path integral of the curvature (τ_2). The other alternative measures ($\tau_{4\text{-}7}$) were similar to each other, very different from τ_2 and of intermediate similarity to the DF and τ_3 (see Supplemental Text 2).

Vessel tortuosity correlates with disease status

We found DF tortuosity (from the combined-vessel GWAS) to be associated with angina, myocardial infarction, stroke, DVT and hypertension. Analyzing arteries and vein separately, we found that the DF of arteries was significantly associated with hypertension (*beta* = 0.19, $p = 3\times10^{-56}$) and angina (*beta* = 0.09, $p = 6\times10^{-4}$), but not with myocardial infarction, stroke or DVT. In the case of veins, the DF was significantly associated with hypertension (*beta* = 0.25, $p = 7\times10^{-99}$), angina (*beta* = 0.18, $p = 2\times10^{-10}$), myocardial infarction (*beta* = 0.12, $p = 2\times10^{-4}$), stroke (*beta* = 0.16, $p = 5\times10^{-5}$) and DVT (*beta* = 0.11, $p = 5\times10^{-4}$). For predictive power over disease status, see Supplemental Text 7.

Vessel tortuosity GWASs identify 173 novel loci

We identified 7□072 significantly associated SNPs in the combined-vessel GWAS on DF tortuosity in the UK Biobank (Supplemental Dataset 4A). The vessel type specific GWAS resulted in 6□563 significantly associated SNPs for arteries, and 2□896 SNPs for veins (Supplemental Dataset 4B and 4C). We applied LD pruning, identifying 128 independent loci in the combined-vessel GWAS, 116 in the artery-specific GWAS, and 63 in the vein-specific GWAS. Accounting for overlap between these sets (see Supplemental Text 9), we obtained a total of 175 independent lead SNPs (see Figure 2a-c). The top 10 SNPs are listed in Table 1, ordered by significance (for complete listings, see Supplemental File 1). Among the significantly associated variants, rs1808382 and rs7991229 had been previously reported [45] (Supplemental Text 8), whereas the remaining 173 independent lead SNPs represented novel loci associated (see Supplemental File 5).

Heritability of DF is larger than for other tortuosity measures

The SNP-based heritability differed substantially across tortuosity measures, with DF receiving the highest estimate ($h^2_{SNP} = 0.25$, SE = 0.025). This was approximately twice the heritability estimate of the six alternative curvature-based measures ($0.11 \le h^2_{SNP} \le 0.13$, $0.011 \le SE \le 0.012$,

Supplemental Text 2). We did not observe any significant genomic inflation (see Table 2). Heritability also varied depending on vessel type (h^2_{SNP} = 0.23 [SE = 0.020] for arteries, and h^2_{SNP} = 0.15 [SE = 0.021] for veins). The distribution of the DF phenotype for each vessel type is shown in Supplemental Text 3.

Replication of lead SNPs and genes in a small meta-cohort

We replicated four of the lead SNPs from the UK Biobank GWAS (see Figure 2e), and discovered three associated genes (TNS1, AC011243.1, LHFPL2; see Figure 3e) [64]. Given the limited sample size of the replication meta-cohort (n = 911), we performed additional analyses on trend of effect sizes. We observed a Pearson correlation of $r = 0.53$ ($p = 1.2\times10^{-11}$) between the effect size estimates in the two studies (see Figure 2d, Figure 3d and Supplemental Text 5).

Tortuosity genes and pathways affect vascular tissue remodeling and angiogenesis

Mapping the SNP-wise association signals onto genes (Methods) we identified 265 significant genes in the discovery GWAS combining vessel types, 203 in the artery-specific GWAS, and 123 in the vein-specific GWAS. Accounting for overlap between these sets (see Supplemental Text 9), we obtained a total of 312 genes (see Figure 3a-c). Top genes are reported in Table 3 (for a complete listing, see Supplemental Dataset 6A / 6B / 6C). A large fraction of these genes carried annotations related to vessel integrity, vascular tissue remodeling and angiogenesis. Specifically, we identified a cluster of highly significant genes on chromosome 19, including *ACTN4* (related to actin filament bundling), *TNS1* (cross-linking of actin filaments), and *CAPN12* (involved in structural integrity to blood vessel walls). This locus also included three genes involved in adhesion to the connective tissue [78]: *LGALS7*, *LGALS7B* and *LGALS4*. We also replicated the highly significant association of tortuosity with two type IV collagen genes, *COL4A2 and COL4A1* [44]. *SYNPO2* (related to actin polymerisation, vascular injury [79] and ocular growth [80], also received a highly significant association. Finally, among the artery-specific genes, we found *FLT1* coding for VEGFR1, which plays a role in vessel formation and vascular biology [81]. (See Discussion for further details and interpretation of these results.)

Gene set enrichment (Methods) yielded 78 significant sets in total (see Figure 4), with the strongest signals arising from the combined and artery-specific analysis (see Supplemental Text 9 and Supplemental Dataset 7A / 7B / 7C). Similarly to genes, many of the pathways pointed to specific biological processes, cellular components, and molecular functions related to vessel integrity and remodeling. These included “human retinal fibroblasts”, “vascular smooth muscle cells” (both in the kidney and the neuroepithelium), and “epithelium development”. We also observed a pathway related to “vascular endothelial growth factors”, VEGFA-VEGFR2, which is a well-known therapeutic target for ocular diseases. We highlight several transcription factors and binding motifs for further experimentation (see Figure 4b). The role of integrity and development of blood vessels for tortuosity was supported by the enrichment of several GO terms such as “circulatory system development”, “anatomical structure morphogenesis” and “tube development”. The enriched terms “cell-substrate junction”, “anchoring junction”, “actin” and “actomyosin” revealed some of the molecular players involved. (See Discussion for more details).

Compared to the DF analysis, the alternative tortuosity measures had lower heritability and fewer enriched genes and pathways. However, some were unique and disease-relevant, such as a pathway related to “abnormal cardiac ventricle morphology” (see Supplemental Text 2).

Tortuosity genes are overexpressed in arteries and heart tissues

Performing enrichment analyses across expression data from 54 tissues, we found that tortuosity genes were overexpressed in three types of arteries (i.e., aorta, tibial artery and coronary artery), two heart tissues (i.e. ventricle and atrial appendage), and, less significantly, fibroblasts and

muscular tissues. The profile of enrichment significance values across tissues for tortuosity genes detected by combined-vessel type GWAS analysis is more similar to that of the artery-specific GWAS than that of vein-specific one (see Figure 5), which did not result in any significant tissue associations (for a strict Bonferroni threshold of $p = 0.05/54 = 9.2\times10^{-4}$).

Tortuosity loci are known disease variants

Nine of the discovered tortuosity loci had been previously reported as disease variants that mapped to specific genes (Table 4): four loci were linked to vascular diseases (coronary heart disease, myocardial infarction, arterial hypertension, and diverticular disease), two loci were linked to ocular diseases (glaucoma and myopia), and three loci were linked to other systemic diseases (chronic lymphocytic leukemia, type 2 diabetes, and Alzheimer's disease). Similarly, we identified 12 loci influencing both tortuosity and disease risk factors. We also uncovered 26 additional disease variants that have not been confidently mapped to a specific gene (see Supplemental Text 10).

Genetic overlap with cardiometabolic risk factors

We expanded our analysis of disease variants to SNPs belonging to the same LD block (Figure 6). We observe a sizable number of tortuosity-associated variants that overlap with CVD (54 SNPs). Several traits related to metabolic syndrome also stand out: blood pressure (55 SNPs for SBP, 49 for DBP, 15 for pulse pressure), blood cholesterol levels (54 SNPs), BMI (54 SNPs), blood pressure linked to alcohol intake and smoking (44 SNPs for SBP□+□alcohol, 27 for DBP□+□alcohol) and type2 diabetes (5 SNPs). In addition, other CVD risk factors share a high number of variants associated with tortuosity, such as protein levels (27 SNPs) and type1 diabetes (9 SNPs). Finally, we detected an overlap with various eye morphology traits, including optic disc morphometry (40 SNPs).

Causal effects between tortuosity, BMI and LDL

Using inverse-variance weighting MR, we observed that exposure to elevated (standardized) levels of LDL reduced the tortuosity of veins by 3% ($p = 0.02$) and arteries by 5% ($p = 0.001$). Conversely, increased venous (but not arterial nor combined) tortuosity reduced BMI by 4.4% ($p = 0.01$) (see Supplemental Text 11).

DISCUSSION

Blood vessel tortuosity is a complex trait whose variation is induced in part during developmental angiogenesis and vascular differentiation and in part through vessel remodeling due to pathological processes in adult life. Both sources of variation are modulated by the environment, but also genetically through gene and regulatory variants that subtly modulate these processes. In order to better understand the involved genetic architecture we conducted the largest GWAS on retinal vessel tortuosity to date, identifying 173 novel loci and pinpointing genes and gene-sets enriched with these primary association signals. Leveraging the unprecedented number of hits, we performed MR that revealed the causal relationships between retinal tortuosity, BMI and blood lipids. This provides context for the considerable overlap we observed between variants associated with vessel tortuosity and cardiometabolic diseases as well as their risk factors. Our results were consistent with the overexpression of tortuosity genes in the aorta, tibial artery, coronary artery, and heart tissues. We found tortuosity genes to be involved in the development of blood vessels, the maintenance of vessel integrity and the remodeling as a consequence of disease processes.

Vessel integrity

Several enriched GO categories that are integral to vessel development were enriched, namely "morphogenesis of anatomical structures", "development of circulatory system", and "tube development". Similarly GO categories pertinent to the structural integrity of vessels and the stability of specific tissues were highlighted: "cell-substrate junction" and "anchoring junction" which are responsible for the mechanical attachment of a cell and its cytoskeleton to the extracellular matrix. Molecularly, "actin cytoskeleton", "actin binding", "actin filament bundle organization", and "positive regulation of actin filament bundle assembly" highlighted the important role of actin.

Among the top hits, we found genes directly related to vessel integrity. The product of *ACTN4*, contributes to cell adhesion and to assembly of the tight junction by mediating actin filament bundling. The paralogues *COL4A1* and *COL4A2* provide structural support and elasticity to connective tissues by forming the hetero-trimer α1α1α2, which is the most abundant collagen in the basement membrane [82]. We found both *COL4A2* and *ACTN4* to be over-expressed in vascular tissues (see Supplemental Text 12). Two more genes with actin-related activity were also among our top hits: *TNS1*, which promotes cell migration and regulates angiogenesis [83], and *SYNPO2*, which is activated by actin polymerization, highly expressed in SMCs [79] and known to provide structural integrity to blood vessel walls [84]. Finally, we identified three genes coding for galectins, which are involved in adhesion to the connective tissue via modulation of cell-cell and cell-matrix interactions [78]: *LGALS7*, its paralog *LGALS7B* and *LGALS4*.

Vessel remodeling

Pathological stresses such as inflammation, infection, or injury, can cause remodeling of vessels, manifesting as occlusions, kinks, tubulations, or other collateral formation of vessels. Pathway analysis identified gene sets of ECs (four sets), SMCs (2 sets), fibroblasts (1 set) and pericytes (1 set) which are the basic cell types composing vessel walls. Dysregulated response of vascular SMC can induce hypertension, and excessive proliferation of these cells contributes to CVD progression [85]. ECs dysfunction can lead to hyperpermeability, neurovascular decoupling and proinflammatory responses [7]. We identified a gene set for "human retinal fibroblasts" consistent with the fact that this cell type is the most common in connective tissue and involved in maintaining the extracellular matrix. Under stress, fibroblasts proliferate resulting in the accumulation of extracellular materials that ultimately limits elasticity [86]. In addition, we found enrichment in a gene set related to "mesangial cells", which are kidney-specific pericyte cells. Retinal capillaries are composed of endothelial cells and pericytes. These contractile cells control blood flow in capillaries [87] and their function is inhibited under stress, such as in high glucose conditions typical in diabetes [88]. Therefore dysregulation of these gene sets has the potential to induce vessel remodeling under stress.

We identified genes directly involved in vessel remodeling. In particular, *FLT1* plays a role in the process of collateral vessel formation, which is a form of vascular remodeling in response to stress, such as hypoxia or hypertension [89]. *FLT1* is transcribed in several tissues including arteries and heart [71] and translated into VEGFR1. VEGFR1 is upregulated in response to micro-inflammation in the early stages of several vascular diseases [89]. In the retina, VEGFR1 is observed in ECs, SMCs, pericytes and RPE cells (which modulate fibroblast proliferation), and excess VEGFR1 contributes to vessel leakage and angiogenesis [89].

Associations with diseases

We detected pleiotropic effects of tortuosity loci, which we showed to be independently associated with CAD, myocardial infarction, hypertension, diabetes, chronic lymphocytic leukemia, Alzheimer's disease, myopia and glaucoma. We also found tortuosity genes to be involved in disease pathomechanisms. *ACTN4*, our top hit, was recently associated with vasorelaxation [90], a mechanism that can lead to hypertension when malfunctioning. The lead SNP in *ACTN4* tortuosity (rs1808382) is also independently associated with CAD [45]. *COL4A1* mutation has been reported

as the cause of retinal arteriolar tortuosity [91] and cerebral small vessel disease [92] vessel leakage and hyperpermeability [93]. Fittingly, *COL4A2* also figured among our variants with pleiotropic effects on disease risk (see Table 4). Variants in the fetal genome near *FLT1* have been associated with preeclampsia [94], a condition of pregnant women presenting with hypertension and damage to the liver and kidneys, whose underlying mechanism involves abnormal formation of blood vessels in the placenta [95]. Retinal vessel modifications have been observed to precede clinical onset of preeclampsia and persist upto 12 months postpartum [96–98].

We elucidated causal links between tortuosity and disease risk factors by applying MR. Specifically, we established that elevated LDL exposure causally reduces arterial tortuosity. High LDL is known to cause the buildup of atherosclerotic plaque [99], which has been clinically linked to arterial tortuosity [100,101]. In fact, arteriosclerosis may make retinal arterial walls less flexible and thereby reduce their DF. We observed a *negative* causal effect of venous tortuosity on BMI, despite the known *positive* correlation between BMI and retinal tortuosity [102], suggesting that environmental factors may play a role in the relationship between BMI and vascular tortuosity.

Limitations

This study was subject to the following limitations. First, we focused on the DF as a tortuosity measure, since the corresponding GWAS revealed many more significant loci, genes and pathways, as well as a higher heritability estimate in comparison to the alternative curvature-based tortuosity measures. These measures are more sensitive to local physiological vessel features, such as aneurysms or sharp bending ("kinks"), while DF only captures the total vessel elongation. Interestingly, the GWAS for these measures revealed several specific genes and pathways that were not significant in the DF analysis, which may be associated with pathologies manifesting as local disruptions in the microvascular network. Further work is needed to elucidate to what extent the stronger association signals for the DF are due to its robustness as a tortuosity measure or its quality to capture *total* vessel elongation as the most physiologically relevant trait**.** Second, due to the small size of our replication meta-cohort we only had enough power to replicate 4 of our 173 hits. Nevertheless, the effect sizes of SNPs and genes in the replication studies correlated strongly with those in the discovery cohort, providing independent evidence that they were not driven by any artifacts specific to the UK Biobank [52]. Moreover, our gene and pathway analyses indicate that these signals make biological sense. Finally, we did not adjust for spherical equivalent refractive error, which may have confounded our measurements to some degree. Investigating the impact of this and other corrections on the heritability estimates could shed more light on their usefulness.

Conclusion

This study exploits advanced automated image processing and deep learning to characterize different *vessel type specific* retinal tortuosity measures from >100k fundus images to conduct a high-powered GWAS on this trait. The resulting significant association signals for 175 genetic loci allowed us to estimate the heritability of tortuosity, identify genes, annotated gene-sets and tissues relevant for this trait, and reveal pleiotropic links with and causal effects to or from disease-related traits. This range of analyses was largely possible due to the large data set made available by the UK Biobank, in terms of both genetic analysis and the high-dimensional data available from fundus images. Combining these data provided novel insights into the genetic architecture of retinal tortuosity, the associated morphological features, the potential disease processes, and the identification of candidate targets for ocular and systemic treatments.

ACKNOWLEDGEMENTS

This work was conducted using the UK Biobank (application ID 43805) and SKIPOGH. Thanks to Micha Hersch for inspiring this project, to the UK Biobank team for their support and

responsiveness, and to all UK Biobank participants for sharing their personal data. We also thank aSciStance Ltd for their help in revising the manuscript.

SOURCE OF FUNDING

This work was supported by the Swiss National Science Foundation (#FN 310030_152724/1 to SB) and by the Swiss Personalized Health Network (2018DRI13 to Thomas J. Wolfensberger). The SKIPOGH study was also supported by the Swiss National Science Foundation (#FN 33CM30-124087 to MB). The *OphtalmoLaus* study was supported by the Claire et Selma Kattenburg Foundation.

DISCLOSURES

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

MT and SB designed the study. MT and MJB performed QC on the raw images. MT and SOV extracted tortuosity measurements from the image data (UK Biobank, *OphalmosLaus* and SKIPOGH). MJB performed classification of arteries and veins. MT carried out the median DF tortuosity GWAS, with the guidance of SB, NM and EP. SOV carried out the tortuosity GWASs based on alternative measures with input from MT. MJB and SOV carried out gene and pathway scoring with the guidance of DK. ALB performed LD Score Regression analysis. SOV evaluated the correlation between different tortuosity measures and their impact on genetic associations. MT and TC performed the replication analysis in SKIPOGH with input from MB. HA, LK, RS and CB provided ophthalmological expertise and manually annotated the raw image data. MT, CB and SB lead the writing of the manuscript with contributions from all other authors.

SUPPLEMENTAL

Supplemental Methods (Supplemental Text 1–5)
Supplemental Results (Supplemental Text 6–13)
Supplemental Datasets 1–3, 4A–C, 5, 6A–C, 7A–C

TABLES

Chr	SNP	EA	RA	freq	beta	$-\log_{10}$ *p*	GWAS type
13	rs9559797	G	C	0.580	-0.162	182.4	artery and vein
19	rs16972767	G	A	0.473	-0.150	164.6	artery and vein
4	rs17008193	T	C	0.403	-0.089	55.0	artery and vein
7	rs187691758	A	G	0.005	0.627	53.2	vein
4	rs12506823	G	A	0.406	0.083	47.9	artery and vein
2	rs2571461	T	G	0.601	-0.082	47.1	artery and vein
15	rs12913832	A	G	0.744	0.080	43.6	artery and vein
12	rs11045245	A	G	0.375	0.073	37.3	artery
5	rs784420	A	G	0.281	0.078	37.0	artery and vein
4	rs11727963	G	A	0.166	0.092	35.1	artery

Table 1 | Top retinal tortuosity SNPs. The 10 most significant DF tortuosity SNPs, ordered by p-value. For full results, refer to the list of 175 independent lead SNPs in Supplemental Dataset 1. **Chr**: chromosome; **SNP**: rsIDs of the single nucleotide polymorphism; **EA**: effect allele; **RA**: reference allele; **freq**: allele frequency of effect allele; **beta**: effect size estimate; **$-\log_{10}$ *p***: normalized p-value in the discovery cohort; **GWAS type**: vessel type to which the signal applied.

GWAS type	h2SNP	lambda GC	mean Chi2	intercept	ratio
combined-vessel	0.25 (0.025)	1.14	1.31	1.01 (0.01)	0.03 (0.03)
artery	0.23 (0.020)	1.12	1.27	1.00 (0.01)	< 0
vein	0.15 (0.021)	1.10	1.18	1.00 (0.01)	< 0

Table 2 | SNP-based heritability. h^2_{SNP}: portion of phenotypic variance cumulatively explained by the SNPs; **lambda GC**: inflation, measure of the effect of confounding and polygenicity acting on the trait; **intercept**: LD score regression intercept (values close to 1 indicates little influence of confounders, mostly of population stratification); **ratio**: ratio of the proportion of the inflation in the **mean Chi^2** that is not due to polygenicity (a ratio close to, or smaller than, 0 is desirable as it indicates low inflation from population stratification). SE are given in parentheses.

Gene	Chr	base pair	$-\log_{10}$ *p* combined	$-\log_{10}$ *p* artery	$-\log_{10}$ *p* vein
ACTN4	19	39,138,289	> 100	> 100	> 100
CAPN12	19	39,220,827	> 100	> 100	> 100
EIF3K	19	39,109,735	> 100	> 100	> 100
LGALS7	19	39,261,611	> 100	> 100	> 100
LGALS7B	19	39,279,851	> 100	> 100	> 100
COL4A2	13	110,958,159	> 100	> 100	9.5
LGALS4	19	39,292,311	> 100	34.5	> 100
MAP4K1	19	39,078,281	> 100	34.3	> 100
TNS1	2	218,664,512	> 100	32.7	16.9

ECH1	19	39,306,062	> 100	15.3	> 100
AC104534.3	19	39,310,806	> 100	14.0	> 100

Table 3 | Top retinal tortuosity genes. The 15 most significant DF tortuosity genes, for each GWAS (combining all vessels, considering only arteries, and only veins). P-values were computed by *PascalX [66]* (precision cutoff: 1×10^{-100}). For full results, refer to Supplemental Dataset 6A / 6B / 6C.

Shared SNP	Disease GWAS	Gene	$-\log_{10} p$	ref
rs875107	diverticular disease	ANO1	22.5	[105]
rs7588567	glaucoma	NCKAP5	21.2	[106]
rs7119	type 2 diabetes	HMG20A	17.5	[107]
rs936226	hypertension	CYP1A2	14.7	[108]
rs757978	chronic lymphocytic leukemia (CLL)	FARP2	10.4	[109]
rs2753462	myopia	FARP2	10.4	[110]
rs6725887 (*)	coronary heart disease	WDR12	9.6	[111]
rs6725887 (*)	myocardial infarction	WDR12	9.6	[112]
rs9381040	Alzheimer's disease	TREML2	6.0	[113]
rs11083475	heart rate (rhythm disorders)	ACTN4	> 100	[114]
rs9555695	waist-hip ratio (obesity)	COL4A2	> 100	[115]
rs2571445	lung function (pulmonary disease)	TNS1	> 100	[116]
rs3791979	intraocular pressure (open angle glaucoma)	TNS1	> 100	[117]
rs17263971	eGFR (Chronic Kidney Disease) and retinal dysfunction	SYNPO2	28.7	[79,118]
rs35252676	pulse pressure (CVD)	LHFPL2	26.6	[119]
rs1378942 (*)	diastolic blood pressure (CVD)	CSK	23.4	[120]
rs1378942 (*)	mean arterial pressure (CVD)	CSK	23.4	[121]
rs17355629	pulse pressure (CVD)	LRCH1	19.6	[122]
rs7655064	waist-hip ratio (obesity)	MYOZ2	14.5	[115]
rs6495122	diastolic blood pressure (CVD)	CPLX3	14.5	[123]
rs12913832	intraocular pressure (open angle glaucoma)	HERC2	12.3	[124]
rs9303401	cognitive ability (mental disorders)	PPM1E	10.01	[125]

Table 4 | Pleiotropic disease-variants. List of variants identified in the tortuosity GWAS (combined-vessel analysis) which were found to be associated with a disease outcome or risk factor in an independent study. We report only exact variants (same rsID in both tortuosity and disease GWAS), which we could confidently map to a gene. Gene p-values were computed by *PascalX* [66]. Variants associated with more than one disease are marked by a star (*)

FIGURES

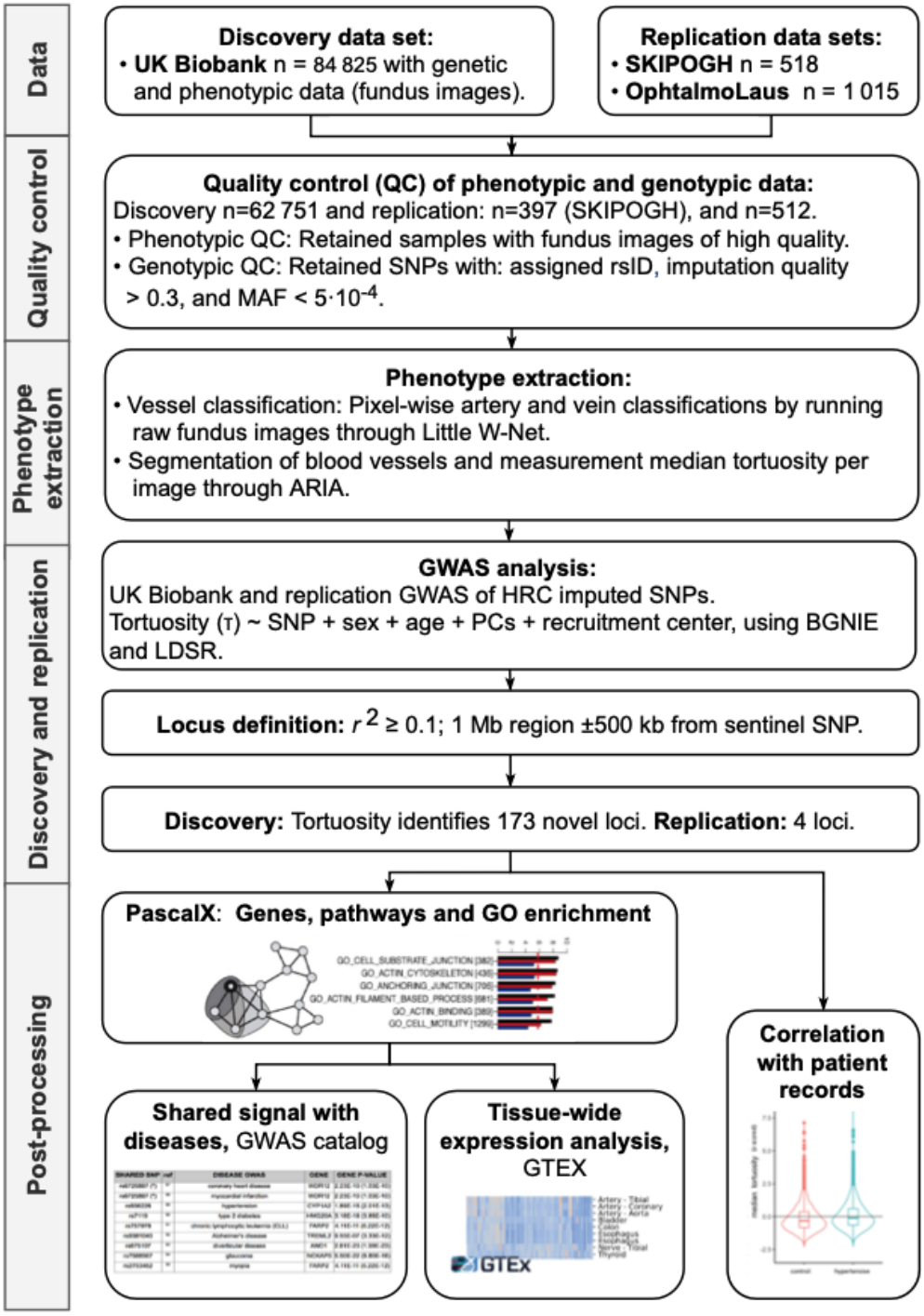


Figure 1 | Pipeline and results. Relevant phenotypes, genotypes and fundus images were collected from the UK Biobank, *OphtalmoLaus* and SKIPOGH. After quality control, the images were processed by deep learning, classifying arteries and veins. A range of tortuosity measures were then calculated, which provided the phenotypes for the GWASs. The primary results were 173 novel genetic trait loci. The association signals pointed to a shared genetic architecture of retinal tortuosity and disease (metabolic syndrome and CVD). Their aggregation on annotated gene-sets identified relevant pathways and GO terms. Tissue-wide expression analysis revealed expression in the arteries and heart. Correlation analysis revealed associations between retinal tortuosity and cardiometabolic diseases.

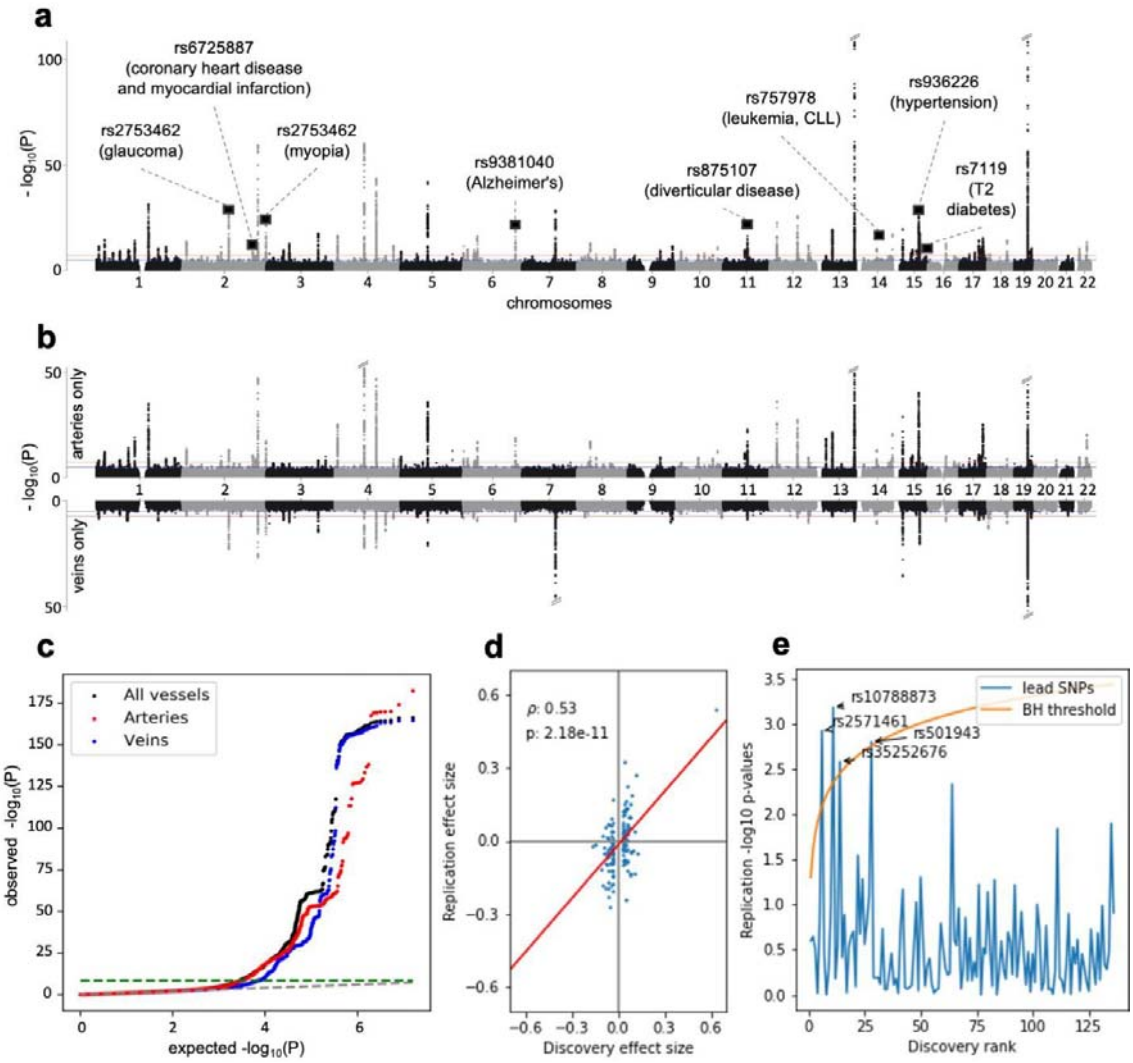


Figure 2 | SNP p-values and effects. a, Manhattan plot of Genome-Wide Association Study (GWAS) of retinal vessel tortuosity, combining all vessel types (both arteries and veins). The red line indicates the genome-wide significance level after Bonferroni correction ($p = 5\times10^{-8}$). Oblique dashes on top of peaks mark extremely significant p-values that have been cropped. Squares mark the position of disease SNPs (see Table 4). The trait was corrected for phenotypic variables which showed a statistically significant association, i.e.: age, sex, and a subset of principal components of genotypes. **b**, Manhattan plots of the vessels-specific GWAS (artery-specific on top, vein-specific at the bottom). Confounder correction, significance level and cropping of extremely significant p-values as in the (a). **c**, GWAS q-q plot: arteries in red, veins in blue, combined-vessels signal in black; the genome-wide significance level is represented as a green dashed line. **d**, Statistically significant correlation between the measured effect sizes in the discovery cohort (UK Biobank, n = 62 751) and replication meta-cohort (SKIPOGH plus *OphtalmoLaus*, n = 911). We considered all lead (independent) SNPs in the UK Biobank. We could find 136 with matching rsIDs in the replication meta-cohort, 90 of which had the same sign of their effect size estimate in the UK Biobank. The resulting Pearson correlation is $r = 0.53$; $t=7.28$, $p = 1.18\times10^{-11}$. **e**, Benjamini-Hochberg procedure on discovery lead SNPs from the UK Biobank yields 4 hits in the replication cohort.

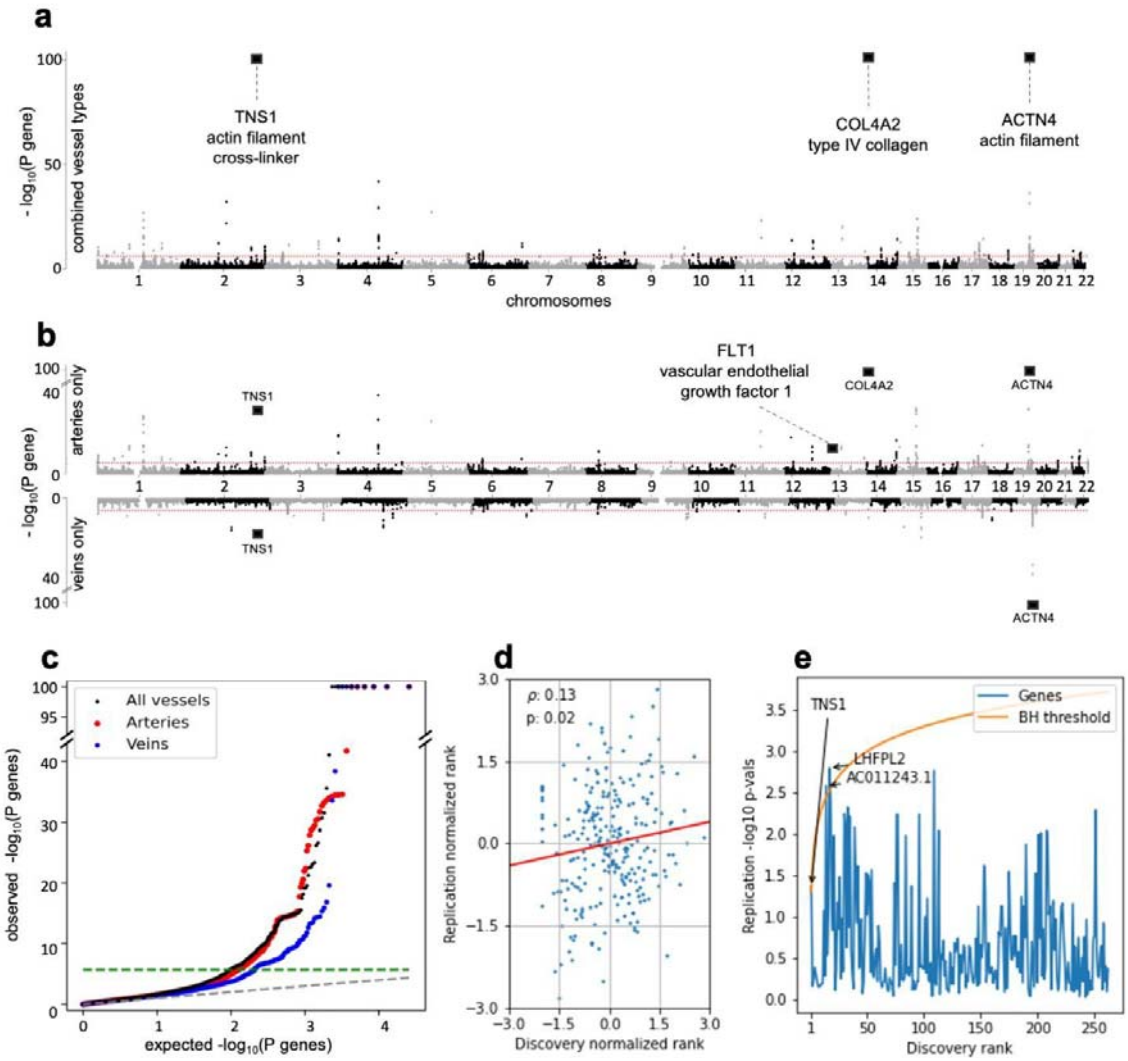


Figure 3 | Gene p-values and replication scores. a, Gene-based Manhattan plot of retinal vessel tortuosity, combining all vessel types (both arteries and veins). Gene-based tests were computed by *PascalX* [66]. The red line indicates the genome-wide significance level after Bonferroni correction ($p = 5\times10^{-8}$). Squares mark the position of particularly relevant genes (see corresponding Results section). **b,** Gene-based Manhattan plots of the vessels-specific GWAS (artery-specific on top, vein-specific at the bottom). **c,** q-q plot of gene p-values: arteries in red, veins in blue, combined-vessel signal in black; the genome-wide significance level is represented as a green dashed line. **d,** Statistically significant correlation between q-q normalized genes' p-values in the discovery (UK Biobank) and in the replication meta-cohort (SKIPOGH + *OphtalmoLaus*). Only genes that were significant in the discovery cohort were considered. The resulting Pearson correlation is $r = 0.13$; $p = 0.02$. **e,** Benjamini-Hochberg generates 3 hits in the replication meta-cohort. Only genes that were significant in the discovery cohort were considered.

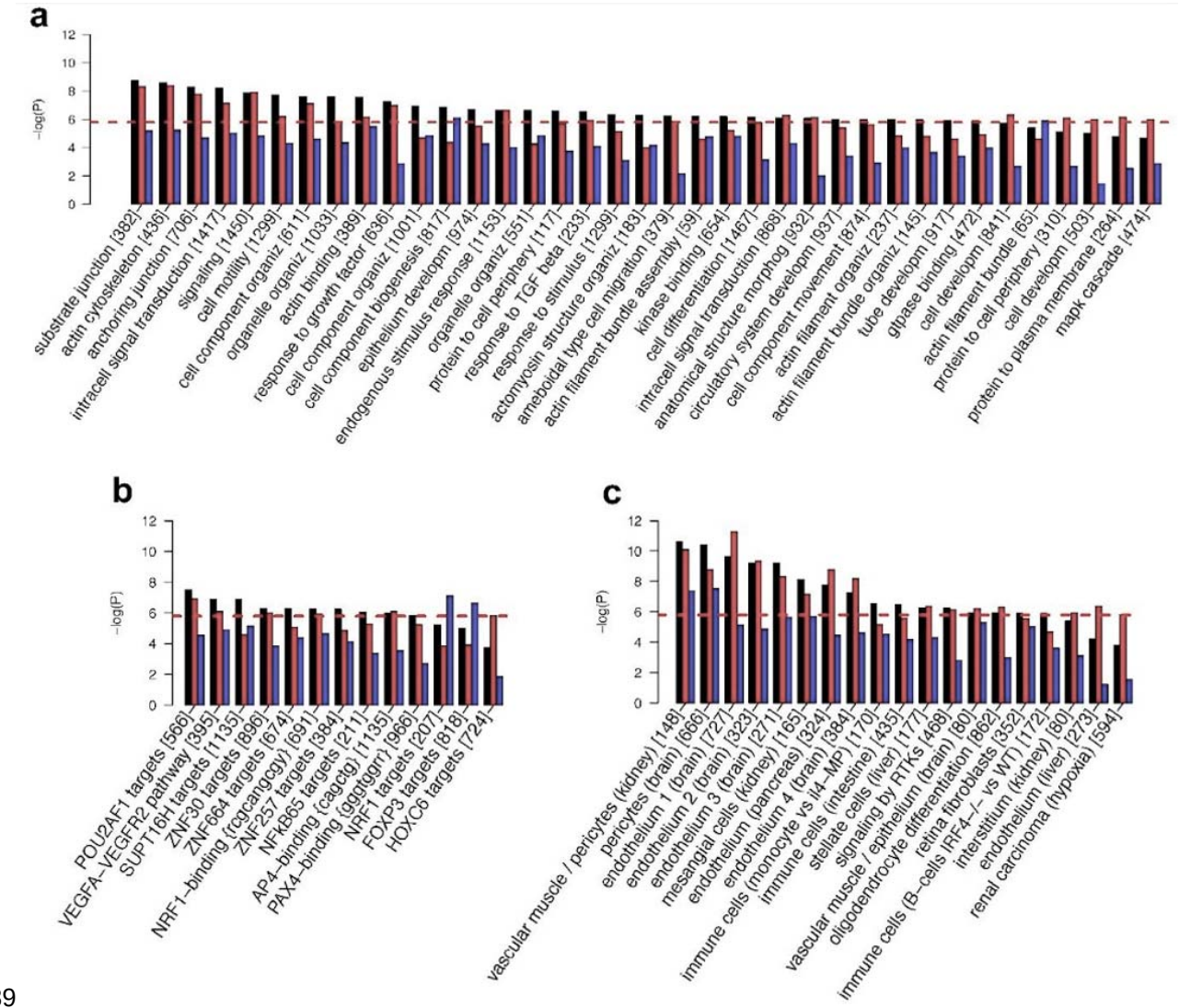


Figure 4 | Enriched pathways and gene-sets. Arteries in red, veins in blue, combined-vessel signal in black: scores for 31□120 gene-sets in MSigDB (v7.2) [103] were calculated by *PascalX* [66]. Only gene-sets for which significance was reached by at least one GWAS are shown. The red dashed line indicates Bonferroni-threshold ($-\log_{10} p = 5.7$). The number of genes in each set is indicated in squared brackets. Gene-set names have been shortened and some redundant GO categories are not shown. For details, refer to the extended plot in Supplemental Text 13. **a,** Enrichment in GO categories. **b,** Enrichment in pathways referring to a particular molecule (typically a transcription factor) and/or binding motif. **c,** Enrichment in gene-set obtained from transcriptomic analysis of tissues of treated cell types.

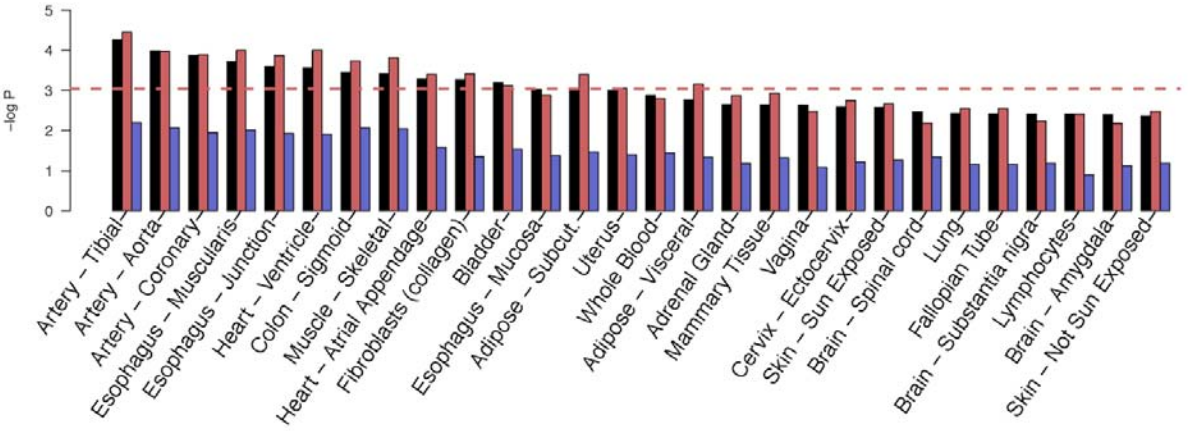


Figure 5 | Tissue expression results. Arteries in red, veins in blue, combined-vessel signal in black: tissue-specific gene expression analysis of GTEx (v8) [71] performed using *PascalX* [66]. We defined sets based on the significant genes from each of the three GWAS we carried out and asked whether they were over-expressed in a particular tissue.

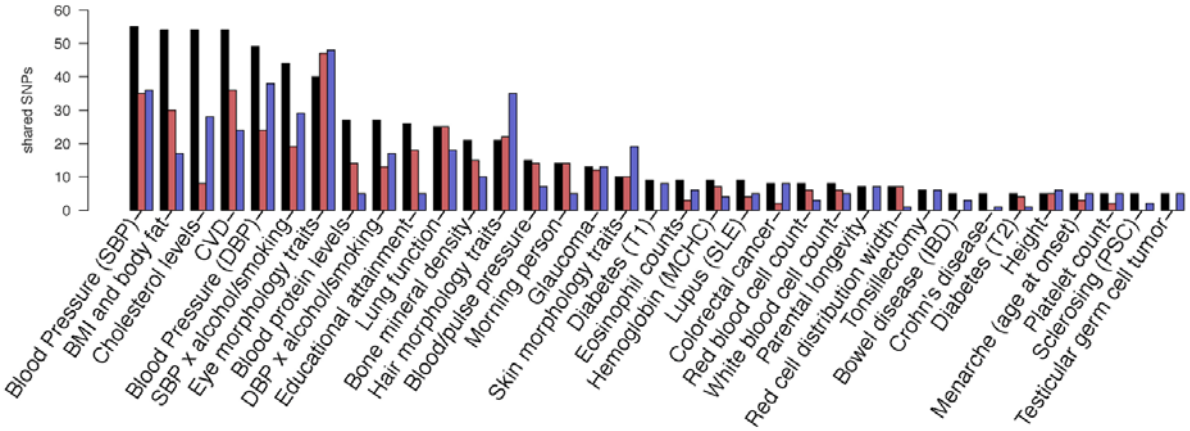


Figure 6 | Overlap in genetic signals with diseases and other complex traits. Arteries in red, veins in blue, combined-vessel signal in black: number of variants shared with other traits reported in the GWAS Catalog [72] (also considering SNPs in high LD with the lead SNP, $r^2 > 0.8$). Only traits with at least 5 shared associations are included (for a full list, including rsIDs, refer to the Supplemental Dataset 3). The traits with the highest number of shared SNPs belong to metabolic syndrome (blood pressure, BMI, blood cholesterol levels) and CVD. This analysis was generated using FUMA [104].

REFERENCES

1. Wilkins E, Wilson L, Wickramasinghe K, Bhatnagar P, Leal J, Luengo-Fernandez R, et al. European cardiovascular disease statistics 2017. 2017 [cited 25 May 2021]. Available: https://researchportal.bath.ac.uk/en/publications/european-cardiovascular-disease-statistics-2017

2. Federal Statistical Office. Cause of death statistics. Bundesamt für Statistik (BFS); 2021.

3. Rana JS, Khan SS, Lloyd-Jones DM, Sidney S. Changes in Mortality in Top 10 Causes of Death from 2011 to 2018. J Gen Intern Med. 2021;36: 2517–2518.

4. Díaz-Coránguez M, Ramos C, Antonetti DA. The inner blood-retinal barrier: Cellular basis and development. Vision Res. 2017;139: 123–137.

5. Klaassen I, Van Noorden CJF, Schlingemann RO. Molecular basis of the inner blood-retinal barrier and its breakdown in diabetic macular edema and other pathological conditions. Prog Retin Eye Res. 2013;34: 19–48.

6. Liew G, Wang JJ, Mitchell P, Wong TY. Retinal vascular imaging: a new tool in microvascular disease research. Circ Cardiovasc Imaging. 2008;1: 156–161.

7. Duh EJ, Sun JK, Stitt AW. Diabetic retinopathy: current understanding, mechanisms, and treatment strategies. JCI Insight. 2017;2. doi:10.1172/jci.insight.93751

8. MacCormick IJC, Czanner G, Faragher B. Developing retinal biomarkers of neurological disease: an analytical perspective. Biomark Med. 2015;9: 691–701.

9. Patton N, Aslam T, Macgillivray T, Pattie A, Deary IJ, Dhillon B. Retinal vascular image analysis as a potential screening tool for cerebrovascular disease: a rationale based on homology between cerebral and retinal microvasculatures. J Anat. 2005;206: 319–348.

10. Liao H, Zhu Z, Peng Y. Potential Utility of Retinal Imaging for Alzheimer's Disease: A Review. Front Aging Neurosci. 2018;10: 188.

11. Dumitrascu OM, Qureshi TA. Retinal Vascular Imaging in Vascular Cognitive Impairment: Current and Future Perspectives. J Exp Neurosci. 2018;12: 1179069518801291.

12. Baker ML, Hand PJ, Wang JJ, Wong TY. Retinal signs and stroke: revisiting the link between the eye and brain. Stroke. 2008;39: 1371–1379.

13. Weiler DL, Engelke CB, Moore ALO, Harrison WW. Arteriole tortuosity associated with diabetic retinopathy and cholesterol. Optom Vis Sci. 2015;92: 384–391.

14. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 2016;316: 2402–2410.

15. Mookiah MRK, Acharya UR, Fujita H, Tan JH, Chua CK, Bhandary SV, et al. Application of different imaging modalities for diagnosis of Diabetic Macular Edema: A review. Comput Biol Med. 2015;66: 295–315.

16. Wang JJ, Taylor B, Wong TY, Chua B, Rochtchina E, Klein R, et al. Retinal vessel diameters and obesity: a population-based study in older persons. Obesity . 2006;14: 206–214.

17. Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nat Biomed Eng. 2018;2: 158–164.

18. Flammer J, Konieczka K, Bruno RM, Virdis A, Flammer AJ, Taddei S. The eye and the heart. Eur Heart J. 2013;34: 1270–1278.

19. Seidelmann SB, Claggett B, Bravo PE, Gupta A, Farhad H, Klein BE, et al. Retinal Vessel Calibers in Predicting Long-Term Cardiovascular Outcomes: The Atherosclerosis Risk in Communities Study. Circulation. 2016;134: 1328–1338.

20. Ikram MK, Ong YT, Cheung CY, Wong TY. Retinal vascular caliber measurements: clinical significance, current knowledge and future perspectives. Ophthalmologica. 2013;229: 125–136.

21. Kawasaki R, Cheung N, Wang JJ, Klein R, Klein BE, Cotch MF, et al. Retinal vessel diameters and risk of hypertension: the Multiethnic Study of Atherosclerosis. J Hypertens. 2009;27: 2386–2393.

22. Ikram MK, de Jong FJ, Bos MJ, Vingerling JR, Hofman A, Koudstaal PJ, et al. Retinal vessel diameters and risk of stroke: the Rotterdam Study. Neurology. 2006;66: 1339–1343.

23. Liew G, Mitchell P, Rochtchina E, Wong TY, Hsu W, Lee ML, et al. Fractal analysis of retinal microvasculature and coronary heart disease mortality. Eur Heart J. 2011;32: 422–429.

24. Wintergerst MWM, Falahat P, Holz FG, Schaefer C, Schahab N, Finger R. Retinal Vasculature assessed by OCTA in Peripheral Arterial Disease. Invest Ophthalmol Vis Sci. 2020;61: 3203–3203.

25. Konstantinidis L, Guex-Crosier Y. Hypertension and the eye. Curr Opin Ophthalmol. 2016;27: 514–521.

26. Smith W, Wang JJ, Wong TY, Rochtchina E, Klein R, Leeder SR, et al. Retinal arteriolar narrowing is associated with 5-year incident severe hypertension: the Blue Mountains Eye Study. Hypertension. 2004;44: 442–447.

27. Wong T, Mitchell P. The eye in hypertension. Lancet. 2007;369: 425–435.

28. Cheung CY-L, Zheng Y, Hsu W, Lee ML, Lau QP, Mitchell P, et al. Retinal vascular tortuosity, blood pressure, and cardiovascular risk factors. Ophthalmology. 2011;118: 812–818.

29. Wong TY, Shankar A, Klein R, Klein BEK, Hubbard LD. Prospective cohort study of retinal vessel diameters and risk of hypertension. BMJ. 2004;329: 79.

30. Dimmitt SB, West JN, Eames SM, Gibson JM, Gosling P, Littler WA. Usefulness of ophthalmoscopy in mild to moderate hypertension. Lancet. 1989;1: 1103–1106.

31. Leung H, Wang JJ, Rochtchina E, Wong TY, Klein R, Mitchell P. Impact of current and past blood pressure on retinal arteriolar diameter in an older population. J Hypertens. 2004;22: 1543–1549.

32. Wong TY, Klein R, Sharrett AR, Duncan BB, Couper DJ, Klein BEK, et al. Retinal arteriolar diameter and risk for hypertension. Ann Intern Med. 2004;140: 248–255.

33. Ikram MK, Witteman JCM, Vingerling JR, Breteler MMB, Hofman A, de Jong PTVM. Retinal vessel diameters and risk of hypertension: the Rotterdam Study. Hypertension. 2006;47: 189–194.

34. Sharrett AR, Hubbard LD, Cooper LS, Sorlie PD, Brothers RJ, Nieto FJ, et al. Retinal arteriolar diameters and elevated blood pressure: the Atherosclerosis Risk in Communities Study. Am J Epidemiol. 1999;150: 263–270.

35. Woo SCY, Lip GYH, Lip PL. Associations of retinal artery occlusion and retinal vein occlusion to mortality, stroke, and myocardial infarction: a systematic review. Eye . 2016;30: 1031–1038.

36. Rim TH, Han JS, Oh J, Kim DW, Kang S-M, Chung EJ. Retinal vein occlusion and the risk of acute myocardial infarction development: a 12-year nationwide cohort study. Sci Rep. 2016;6: 22351.

37. Sabanayagam C, Xu D, Ting DSW, Nusinovici S, Banu R, Hamzah H, et al. A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. The Lancet Digital Health. 2020;2: e295–e302.

38. Park HC, Lee Y-K, Cho A, Han CH, Noh J-W, Shin YJ, et al. Diabetic retinopathy is a prognostic factor for progression of chronic kidney disease in the patients with type 2 diabetes mellitus. PLoS One. 2019;14: e0220506.

39. Jensen RA, Sim X, Smith AV, Li X, Jakobsdóttir J, Cheng C-Y, et al. Novel Genetic Loci Associated With Retinal Microvascular Diameter. Circ Cardiovasc Genet. 2016;9: 45–54.

40. Ikram MK, Sim X, Jensen RA, Cotch MF, Hewitt AW, Ikram MA, et al. Four novel Loci (19q13, 6q24, 12q24, and 5q14) influence the microcirculation in vivo. PLoS Genet. 2010;6: e1001184.

41. Springelkamp H, Mishra A, Hysi PG, Gharahkhani P, Höhn R, Khor C-C, et al. Meta-analysis of Genome-Wide Association Studies Identifies Novel Loci Associated With Optic Disc Morphology. Genet Epidemiol. 2015;39: 207–216.

42. Han X, Qassim A, An J, Marshall H, Zhou T, Ong J-S, et al. Genome-wide association analysis of 95 549 individuals identifies novel loci and genes influencing optic disc morphology. Hum Mol Genet. 2019;28: 3680–3690.

43. Zekavat SM, Raghu VK, Trinder M, Ye Y, Koyama S, Honigberg MC, et al. Deep Learning of the Retina Enables Phenome- and Genome-wide Analyses of the Microvasculature. Circulation. 2021. doi:10.1161/CIRCULATIONAHA.121.057709

44. Veluchamy A, Ballerini L, Vitart V, Schraut KE, Kirin M, Campbell H, et al. Novel Genetic Locus Influencing Retinal Venular Tortuosity Is Also Associated With Risk of Coronary Artery Disease. Arterioscler Thromb Vasc Biol. 2019;39: 2542–2552.

45. Veluchamy A, Ballerini L, Vitart V, Schraut KE, Kirin M, Campbell H, et al. Novel Genetic Locus Influencing Retinal Venular Tortuosity Is Also Associated With Risk of Coronary Artery Disease. Arterioscler Thromb Vasc Biol. 2019;39: 2542–2552.

46. Welby JP, Kim ST, Carr CM, Lehman VT, Rydberg CH, Wald JT, et al. Carotid Artery Tortuosity Is Associated with Connective Tissue Diseases. AJNR Am J Neuroradiol. 2019;40: 1738–1743.

47. Tapp RJ, Owen CG, Barman SA, Welikala RA, Foster PJ, Whincup PH, et al. Associations of Retinal Microvascular Diameters and Tortuosity With Blood Pressure and Arterial Stiffness. Hypertension. 2019. pp. 1383–1390. doi:10.1161/hypertensionaha.119.13752

48. Heneghan C, Flynn J, O'Keefe M, Cahill M. Characterization of changes in blood vessel width and tortuosity in retinopathy of prematurity using image analysis. Med Image Anal. 2002;6: 407–429.

49. Pruijm M, Ponte B, Ackermann D, Vuistiner P, Paccaud F, Guessous I, et al. Heritability, determinants and reference values of renal length: a family-based population study. Eur Radiol. 2013;23: 2899–2905.

50. Ponte B, Pruijm M, Ackermann D, Vuistiner P, Eisenberger U, Guessous I, et al. Reference values and factors associated with renal resistive index in a family-based population study. Hypertension. 2014;63: 136–142.

51. Firmann M, Mayor V, Vidal PM, Bochud M, Pécoud A, Hayoz D, et al. The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. BMC Cardiovasc Disord. 2008;8: 6.

52. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, 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. 2015. p. e1001779. doi:10.1371/journal.pmed.1001779

53. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018;562: 203–209.

54. Pistis G, Porcu E, Vrieze SI, Sidore C, Steri M, Danjou F, et al. Rare variant genotype imputation with thousands of study-specific whole-genome sequences: implications for cost-effective study designs. Eur J Hum Genet. 2015;23: 975–983.

55. Bankhead P, Scholfield CN, McGeown JG, Curtis TM. Fast retinal vessel detection and measurement using wavelets and edge location refinement. PLoS One. 2012;7: e32435.

56. Al-Diri B, Hunter A, Steel D, Habib M, Hudaib T, Berry S. REVIEW - a reference data set for retinal vessel profiles. Conf Proc IEEE Eng Med Biol Soc. 2008;2008: 2262–2265.

57. Smedby O, Högman N, Nilsson S, Erikson U, Olsson AG, Walldius G. Two-dimensional tortuosity of the superficial femoral artery in early atherosclerosis. J Vasc Res. 1993;30: 181–191.

58. Abdalla M, Hunter A, Al-Diri B. Quantifying retinal blood vessels' tortuosity — Review. 2015 Science and Information Conference (SAI). 2015. doi:10.1109/sai.2015.7237216

59. Adrian Galdran, André Anjos, José Dolz, Hadi Chakor, Hervé Lombaert, Ismail Ben Ayed. The Little W-Net That Could: State-of-the-Art Retinal Vessel Segmentation with Minimalistic Models. arXiv. 2020. doi:The Little W-Net That Could: State-of-the-Art Retinal Vessel Segmentation with Minimalistic Models

60. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. bioRxiv. 2017. p. 166298. doi:10.1101/166298

61. Pain O, Dudbridge F, Ronald A. Are your covariates under control? How normalization can re-introduce covariate effects. Eur J Hum Genet. 2018;26: 1194–1201.

62. Myers TA, Chanock SJ, Machiela MJ. LDlinkR: An R Package for Rapidly Calculating Linkage Disequilibrium Statistics in Diverse Populations. Frontiers in Genetics. 2020. doi:10.3389/fgene.2020.00157

63. Kang HM. EPACTS: efficient and parallelizable association container toolbox. 2016. Available: https://genome.sph.umich.edu/wiki/EPACTS

64. Benjamini Y, Hochberg Y. On the Adaptive Control of the False Discovery Rate in Multiple Testing with Independent Statistics. Journal of Educational and Behavioral Statistics. 2000. p. 60. doi:10.2307/1165312

65. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. Bioinformatics. 2017;33: 272–279.

66. Krefl D, Bergmann S. PascalX v0.0.1. 2021. doi:10.5281/zenodo.4429922

67. Lamparter D, Marbach D, Rueedi R, Kutalik Z, Bergmann S. Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. PLoS Comput Biol. 2016;12: e1004714.

68. Kinsella RJ, Kähäri A, Haider S, Zamora J, Proctor G, Spudich G, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database . 2011;2011: bar030.

69. Huang J, Howie B, McCarthy S, Memari Y, Walter K, Min JL, et al. Improved imputation of low-frequency and rare variants using the UK10K haplotype reference panel. Nat Commun. 2015;6: 8111.

70. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27: 1739–1740.

71. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45: 580–585.

72. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. 2019;47: D1005–D1012.

73. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res. 2017;26: 2333–2355.

74. Smith GD, Ebrahim S. “Mendelian randomization”: can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32: 1–22.

75. Genome-wide Repository of Associations Between SNPs and Phenotypes. In: National Institutes of Health (NIH) [Internet]. [cited Feb 2021]. Available: https://grasp.nhlbi.nih.gov/

76. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37: 658–665.

77. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. Int J Epidemiol. 2017;46: 1734–1739.

78. Kuwabara I, Kuwabara Y, Yang R-Y, Schuler M, Green DR, Zuraw BL, et al. Galectin-7 (PIG1) exhibits pro-apoptotic function through JNK activation and mitochondrial cytochrome c release. J Biol Chem. 2002;277: 3487–3497.

79. Turczyńska KM, Swärd K, Hien TT, Wohlfahrt J, Mattisson IY, Ekman M, et al. Regulation of smooth muscle dystrophin and synaptopodin 2 expression by actin polymerization and vascular injury. Arterioscler Thromb Vasc Biol. 2015;35: 1489–1497.

80. Karouta C, Kucharski R, Hardy K, Thomson K, Maleszka R, Morgan I, et al. Transcriptome-based insights into gene networks controlling myopia prevention. FASEB J. 2021;35: e21846.

81. Shibuya M. Vascular endothelial growth factor receptor-1 (VEGFR-1/Flt-1): a dual regulator for angiogenesis. Angiogenesis. 2006;9: 225–30; discussion 231.

82. J.M.B. Sand, F. Genovese, N.S. Gudmann, M.A. Karsdal. Type IV collagen. In: Karsdal MA, editor. Biochemistry of Collagens, Laminins and Elastin 2nd edition. 2019.

83. Shih Y-P, Sun P, Wang A, Lo SH. Tensin1 positively regulates RhoA activity through its interaction with DLC1. Biochim Biophys Acta. 2015;1853: 3258–3265.

84. Fonović M, Turk B. Cysteine cathepsins and extracellular matrix degradation. Biochim Biophys Acta. 2014;1840: 2560–2570.

85. Yang D, Sun C, Zhang J, Lin S, Zhao L, Wang L, et al. Proliferation of vascular smooth muscle cells under inflammation is regulated by NF-κB p65/microRNA-17/RB pathway activation. Int J Mol Med. 2018;41: 43–50.

86. Dick MK, Miao JH, Limaiem F. Histology, fibroblast. StatPearls [Internet]. 2021. Available: https://www.ncbi.nlm.nih.gov/books/NBK541065/

87. Kur J, Newman EA, Chan-Ling T. Cellular and physiological mechanisms underlying blood flow regulation in the retina and choroid in health and disease. Prog Retin Eye Res. 2012;31: 377–406.

88. Wakisaka M, Nagao T. Sodium glucose cotransporter 2 in mesangial cells and retinal pericytes and its implications for diabetic nephropathy and retinopathy. Glycobiology. 2017;27: 691–695.

89. Uemura A, Fruttiger M, D’Amore PA, De Falco S, Joussen AM, Sennlaub F, et al. VEGFR1 signaling in retinal angiogenesis and microinflammation. Prog Retin Eye Res. 2021;84: 100954.

90. Won K-J, Lee KP, Kim D-K, Jung SH, Lee C-K, Lee DH, et al. Monoclonal antibody against α-actinin 4 from human umbilical vein endothelial cells inhibits endothelium-dependent vasorelaxation. J Vasc Res. 2013;50: 210–220.

91. Zenteno JC, Crespí J, Buentello-Volante B, Buil JA, Bassaganyas F, Vela-Segarra JI, et al. Next generation sequencing uncovers a missense mutation in COL4A1 as the cause of familial retinal arteriolar tortuosity. Graefes Arch Clin Exp Ophthalmol. 2014;252: 1789–1794.

92. Vahedi K, Alamowitch S. Clinical spectrum of type IV collagen (COL4A1) mutations: a novel genetic multisystem disease. Curr Opin Neurol. 2011;24: 63–68.

93. Trouillet A, Lorach H, Dubus E, El Mathari B, Ivkovic I, Dégardin J, et al. Col4a1 mutation generates vascular abnormalities correlated with neuronal damage in a mouse model of HANAC syndrome. Neurobiol Dis. 2017;100: 52–61.

94. McGinnis R, Steinthorsdottir V, Williams NO, Thorleifsson G, Shooter S, Hjartardottir S, et al. Variants in the fetal genome near FLT1 are associated with risk of preeclampsia. Nat Genet. 2017;49: 1255–1260.

95. Al-Jameil N, Aziz Khan F, Fareed Khan M, Tabassum H. A brief overview of preeclampsia. J Clin Med Res. 2014;6: 1–7.

96. Nagy ZZ. Review of the ophthalmic symptoms of preeclampsia. Developments in Health Sciences. 2020. pp. 21–23. doi:10.1556/2066.2020.00005

97. Lupton SJ, Chiu CL, Hodgson LAB, Tooher J, Ogle R, Wong TY, et al. Changes in retinal microvascular caliber precede the clinical onset of preeclampsia. Hypertension. 2013;62: 899–904.

98. Soma-Pillay P, Pillay R, Wong TY, Makin JD, Pattinson RC. The effect of pre-eclampsia on retinal microvascular caliber at delivery and post-partum. Obstet Med. 2018;11: 116–120.

99. Ference BA, Ginsberg HN, Graham I, Ray KK, Packard CJ, Bruckert E, et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. Eur Heart J. 2017;38: 2459–2472.

100. Han H-C. Twisted blood vessels: symptoms, etiology and biomechanical mechanisms. J Vasc Res. 2012;49: 185–197.

101. Kwa VIH, van der Sande JJ, Stam J, Tijmes N, Vrooland JL, Amsterdam Vascular Medicine Group. Retinal arterial changes correlate with cerebral small-vessel disease. Neurology. 2002;59: 1536–1540.

102. Tapp RJ, Owen CG, Barman SA, Welikala RA, Foster PJ, Whincup PH, et al. Retinal Vascular Tortuosity and Diameter Associations with Adiposity and Components of Body Composition. Obesity . 2020;28: 1750–1760.

103. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102: 15545–15550.

104. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8: 1826.

105. Maguire LH, Handelman SK, Du X, Chen Y, Pers TH, Speliotes EK. Genome-wide association analyses identify 39 new susceptibility loci for diverticular disease. Nat Genet. 2018;50: 1359–1365.

106. Osman W, Low S-K, Takahashi A, Kubo M, Nakamura Y. A genome-wide association study in the Japanese population confirms 9p21 and 14q23 as susceptibility loci for primary open angle glaucoma. Hum Mol Genet. 2012;21: 2836–2842.

107. Sim X, Ong RT-H, Suo C, Tay W-T, Liu J, Ng DP-K, et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. PLoS Genet. 2011;7: e1001363.

108. German CA, Sinsheimer JS, Klimentidis YC, Zhou H, Zhou JJ. Ordered multinomial regression for genetic association analysis of ordinal phenotypes at Biobank scale. Genet Epidemiol. 2020;44: 248–260.

109. Slager SL, Skibola CF, Di Bernardo MC, Conde L, Broderick P, McDonnell SK, et al. Common variation at 6p21.31 (BAK1) influences the risk of chronic lymphocytic leukemia. Blood. 2012;120: 843–846.

110. Tedja MS, Wojciechowski R, Hysi PG, Eriksson N, Furlotte NA, Verhoeven VJM, et al. Genome-wide association meta-analysis highlights light-induced signaling as a driver for refractive error. Nat Genet. 2018;50: 834–848.

111. Mehta NN. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Circulation. Cardiovascular genetics. 2011. pp. 327–329.

112. Kathiresan S, Myocardial Infarction Genetics Consortium, Voight BF, Purcell S, Musunuru K, Ardissino D, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet. 2009;41: 334–341.

113. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45: 1452–1458.

114. den Hoed M, Eijgelsheim M, Esko T, Brundel BJJM, Peal DS, Evans DM, et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. Nat Genet. 2013;45: 621–631.

115. Kichaev G, Bhatia G, Loh P-R, Gazal S, Burch K, Freund MK, et al. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. Am J Hum Genet. 2019;104: 65–75.

116. Wain LV, Shrine N, Miller S, Jackson VE, Ntalla I, Soler Artigas M, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. Lancet Respir Med. 2015;3: 769–781.

117. Khawaja AP, UK Biobank Eye and Vision Consortium, Cooke Bailey JN, Wareham NJ, Scott RA, Simcoe M, et al. Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. Nature Genetics. 2018. pp. 778–782. doi:10.1038/s41588-018-0126-8

118. Stapleton CP, Heinzel A, Guan W, van der Most PJ, van Setten J, Lord GM, et al. The impact of donor and recipient common clinical and genetic variation on estimated glomerular filtration rate in a European renal transplant

population. Am J Transplant. 2019;19: 2262–2273.

119. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nat Genet. 2017;49: 403–415.

120. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet. 2009;41: 666–676.

121. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet. 2011;43: 1005–1011.

122. Giri A, Hellwege JN, Keaton JM, Park J, Qiu C, Warren HR, et al. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. Nat Genet. 2019;51: 51–62.

123. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, et al. Genome-wide association study of blood pressure and hypertension. Nat Genet. 2009;41: 677–687.

124. Craig JE, Han X, Qassim A, Hassall M, Cooke Bailey JN, Kinzy TG, et al. Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. Nat Genet. 2020;52: 160–166.

125. Seshadri S, DeStefano AL, Au R, Massaro JM, Beiser AS, Kelly-Hayes M, et al. Genetic correlates of brain aging on MRI and cognitive test measures: a genome-wide association and linkage analysis in the Framingham Study. BMC Med Genet. 2007;8 Suppl 1: S15.

Supplemental References

1. Bankhead P, Scholfield CN, McGeown JG, Curtis TM. Fast retinal vessel detection and measurement using wavelets and edge location refinement. PLoS One. 2012;7: e32435.

2. Galdran A, Anjos A, Dolz J, Chakor H, Lombaert H, Ayed IB. The Little W-Net That Could: State-of-the-Art Retinal Vessel Segmentation with Minimalistic Models. 2020. Available: http://arxiv.org/abs/2009.01907

3. Lee CH, Eskin E, Han B. Increasing the power of meta-analysis of genome-wide association studies to detect heterogeneous effects. Bioinformatics. 2017;33: i379–i388.

4. Veluchamy A, Ballerini L, Vitart V, Schraut KE, Kirin M, Campbell H, et al. Novel Genetic Locus Influencing Retinal Venular Tortuosity Is Also Associated With Risk of Coronary Artery Disease. Arterioscler Thromb Vasc Biol. 2019;39: 2542–2552.

5. Schunkert H, König IR, Kathiresan S, Reilly MP, Assimes TL, Holm H, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet. 2011;43: 333–338.

6. C4D Consortium. A genome-wide association study in Europeans and South Asians identifies five new loci for coronary artery disease. Nat Genet. 2011;43: 339–344.

7. Osman W, Low S-K, Takahashi A, Kubo M, Nakamura Y. A genome-wide association study in the Japanese population confirms 9p21 and 14q23 as susceptibility loci for primary open angle glaucoma. Hum Mol Genet. 2012;21: 2836–2842.

8. Craig JE, Han X, Qassim A, Hassall M, Cooke Bailey JN, Kinzy TG, et al. Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. Nat Genet. 2020;52: 160–166.

9. German CA, Sinsheimer JS, Klimentidis YC, Zhou H, Zhou JJ. Ordered multinomial regression for genetic association analysis of ordinal phenotypes at Biobank scale. Genet Epidemiol. 2020;44: 248–260.

10. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet. 2009;41: 703–707.

11. Plagnol V, Howson JMM, Smyth DJ, Walker N, Hafler JP, Wallace C, et al. Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases. PLoS Genet. 2011;7: e1002216.

12. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nat Commun. 2020;11: 163.

13. Dubois PCA, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet. 2010;42: 295–302.

14. Jin Y, Birlea SA, Fain PR, Ferrara TM, Ben S, Riccardi SL, et al. Genome-wide association analyses identify 13 new susceptibility loci for generalized vitiligo. Nat Genet. 2012;44: 676–680.

15. Kichaev G, Bhatia G, Loh P-R, Gazal S, Burch K, Freund MK, et al. Leveraging Polygenic Functional Enrichment to Improve GWAS Power. Am J Hum Genet. 2019;104: 65–75.

16. Comuzzie AG, Cole SA, Laston SL, Voruganti VS, Haack K, Gibbs RA, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS One. 2012;7: e51954.

17. Levy D, Ehret GB, Rice K, Verwoert GC, Launer LJ, Dehghan A, et al. Genome-wide association study of blood pressure and hypertension. Nat Genet. 2009;41: 677–687.

18. Giri A, Hellwege JN, Keaton JM, Park J, Qiu C, Warren HR, et al. Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. Nat Genet. 2019;51: 51–62.

19. Newton-Cheh C, Johnson T, Gateva V, Tobin MD, Bochud M, Coin L, et al. Genome-wide association study identifies eight loci associated with blood pressure. Nat Genet. 2009;41: 666–676.

20. Wain LV, Verwoert GC, O'Reilly PF, Shi G, Johnson T, Johnson AD, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet. 2011;43: 1005–1011.

21. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun. 2018;9: 2098.

22. Krefl D, Bergmann S. PascalX v0.0.1. 2021. doi:10.5281/zenodo.4429922

23. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet. 2013;45: 580–585.

24. Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, et al. GeneCards Version 3: the human gene integrator. Database . 2010;2010: baq020.

25. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2018;47: D607–D613.

SUPPLEMENTAL MATERIAL

(intended for publication as an online data supplement)

GWAS of Retinal Vessel Tortuosity Identifies 173 Novel Loci Revealing Genes and Pathways Associated with Vascular Pathomechanics and Cardiometabolic Diseases

TABLE OF CONTENTS

SUPPLEMENTAL DATASETS 0

SUPPLEMENTAL METHODS 1

Text 1: Quality Control in tortuosity measurements 1

Text 2: DF and other tortuosity measures 1

Definition of DF 1

Definitions of six, alternative, curvature-based measures 1

PCA of DF and curvature-based measurements 2

Heritability of DF and curvature-based measurements **Error! Bookmark not defined.**

Enrichment analysis of curvature-based measurement associations 3

Text 3: Distribution of DF 4

DF tortuosity across cohorts 5

Stratified DF analysis: sex, age and vessel type 5

Text 4: Deep Learning classification of arteries and veins 6

Accuracy of vessel type classification 6

Censoring unclearly classified vessels 7

GWAS with random vessel type calling 7

Text 5: Replication Analysis 8

Correlation of effect sizes in the meta-cohort 8

Replication of hits in the meta-cohort 8

SUPPLEMENTAL RESULTS 10

Text 6: Baseline Characteristics 10

Text 7: Correlation with disease status 10

Text 8: Replication of known hits 10

Text 9: Vessel-type comparisons for SNPs, Genes and Pathways 13

Text 10: Genetic associations with disease and risk 14

Tortuosity variants associated with disease outcome 14

Tortuosity variants associated with disease risk factors 14

Text 11: Mendelian Randomization 15

Text 12: ACTN4 and COL4A2 over-expression 16

Text 13: Full gene set enrichment results 16

SUPPLEMENTAL DATASET

Supplemental Dataset 1 - 175 lead SNPs
Supplemental Dataset 2 - overlap with GWAS Catalog (SNPs in LD)
Supplemental Dataset 3 - overlap with GWAS Catalog (SNPs with matching rsIDs)
Supplemental Dataset 4A - significant SNPs (combined)
Supplemental Dataset 4B - significant SNPs (artery)
Supplemental Dataset 4C - significant SNPs (vein)
Supplemental Dataset 5 - replication of lead SNPs
Supplemental Dataset 6A - gene scores (combined)
Supplemental Dataset 6B - gene scores (artery)
Supplemental Dataset 6C - gene scores (vein)
Supplemental Dataset 7A - gene set enrichment (combined)
Supplemental Dataset 7B - gene set enrichment (artery)
Supplemental Dataset 7C - gene set enrichment (vein)

SUPPLEMENTAL METHODS

Text 1: Quality Control in tortuosity measurements

The quality control (QC) procedure is performed during data extraction. The objective is to remove (i) lower quality images and (ii) images containing artifacts. QC relies on thresholding being applied to two distributions: 1) Distribution of the total number of equally spaced diameters that were fitted to the vessels of each eye. Images between with values 11□000 and 20□000 passed QC. 2) Distribution of the number of vessels contained in each eye. Images with values between 100 and 250 passed QC. By visual inspection, we fine tuned these thresholds to discriminate (i) low-quality images that were too dark, too light, out of focus (lower-bound thresholds), or (ii) images that contained spurious vessels, i.e., artifacts of the picture that were being erroneously segmented as vessels (higher-bound thresholds). Images that pass QC (roughly two out of three) are further processed to extract several tortuosity measures.

Text 2: DF and other tortuosity measures

Definition of DF

We consider a vessel as a curve in a two dimensional space on the interval [t_0, t_1]. For a given curve there are different tortuosity measures: The simplest one, the "Distance Factor" (DF), is calculated as the total arc length over total chord length. This measure computes the tortuosity of the segment by examining how long the curve is relative to its chord length.

$$DF = \frac{s(C)}{chord(C)}$$

Let $s(C)$ be the arc length of the curve C, and $chord(C)$ its chord length:

$$s(C) = \int_{t_0}^{t_1} \sqrt{x'(t)^2 + y'(t)^2}\, dt \quad , \quad chord(C) = \sqrt{(x(t_1) - x(t_0))^2 + (y(t_1) - y(t_0))^2}$$

where the prime denotes derivation with respect to *t*.

Definitions of six, alternative, curvature-based measures

A first curvature-based tortuosity measure is obtained as the integral over the absolute value of curvature along the entire curvature:

$$\tau_2 = \int_{t_0}^{t_1} |\kappa(t)| dt$$, where the curvature is defined as: $|\kappa(\square)| = \frac{|x\prime(t)y\prime\prime(t) - x\prime\prime(t)y\prime(t)|}{[y\prime(t)^2 + x\prime(t)^2]^{3/2}}$

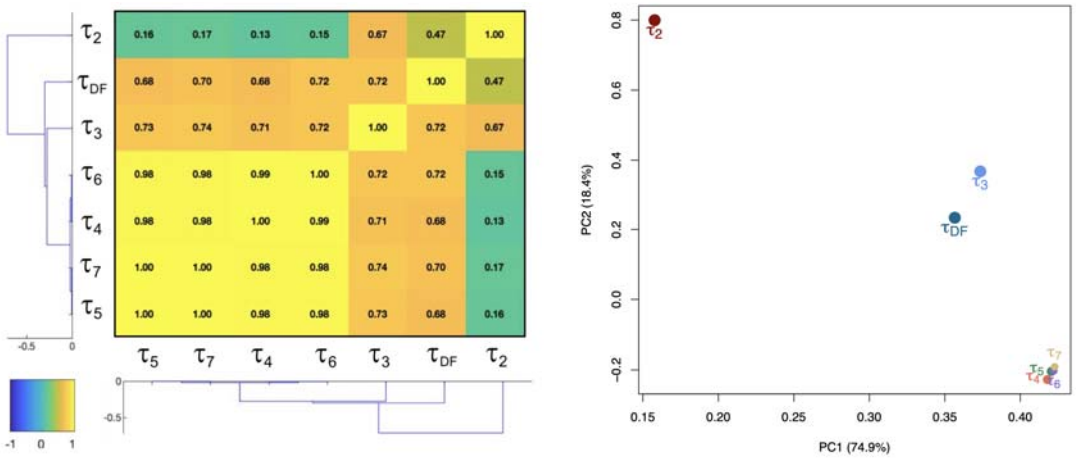
This is equivalent to the inverse curvature radius *R(t)*, i.e. the radius of a circle that is tangent to the curve at t: $|\kappa| = 1/R$.

A second tortuosity measure is the integral over the curvature squared along the entire curve:

Finally, four tortuosity measures arise from normalizing and τ_3 by division through either the curve or the arc length:

We note that, by definition, τ_2 and τ_3 depend on the length of the curve (with non-zero curvature). In the case of vessels with constant curvature radius: $\tau_2 = s(C)/R$ and $\tau_3 = s(C)/R^2$. Our analysis shows that τ_2 differs from other measures the most, followed by τ_3, while $\tau_{4\text{-}7}$ are indeed quite similar to each other (see Supplemental Figure 1)

PCA of DF and curvature-based measurements



Supplemental Figure 1 | Dimensionality reduction on tortuosity measurements across 62□751 individuals. Left: Hierarchical clustering of the pairwise correlations-matrix. Correlations are measured using the Pearson correlation coefficient. Right: We performed dimensionality reduction using a Principal Component Analysis (PCA) approach and plotted PC2 against PC1.

Heritability of DF and curvature-based measurements

Measure	h^2_{SNP}	*Lambda GC*	*Mean Chi²*	*Intercept*	*Ratio*
DF (Distance Factor)	0.25 (0.025)	1.14	1.31	1.01 (0.01)	0.03 (0.03)
τ_2	0.11 (0.011)	1.08	1.11	0.98 (0.01)	< 0
τ_3	0.11 (0.012)	1.09	1.13	0.99 (0.01)	< 0
τ_4	0.12 (0.011)	1.11	1.15	1.00 (0.01)	0.02 (0.05)
τ_5	0.12 (0.011)	1.10	1.14	1.00 (0.01)	< 0
τ_6	0.13 (0.012)	1.11	1.16	1.00 (0.01)	0.01 (0.04)
τ_7	0.12 (0.011)	1.10	1.15	1.00 (0.01)	< 0

Supplemental Table 1 | SNP-based heritability of alternative tortuosity measures. h^2_{SNP} is the

alternative tortuosity measures were less heritable than the one defined by the Distance Factor. *Lambda GC* is the measure of inflation (it measures the effect of confounding and polygenicity acting on the trait). *Intercept* is the LD Score regression intercept (values close to 1 indicates little influence of confounders, mostly of population stratification). *Ratio* is the ratio of the proportion of the inflation in the *Mean Chi*^{2} that is not due to polygenicity (a ratio close to, or smaller than, 0 is desirable as it indicates low inflation from population stratification). SE are given in parentheses.

Enrichment analysis of curvature-based measurement associations

The alternative tortuosity measures have less significant SNPs, genes and pathways overall compared to the distance factor, but harbor few unique pathways, among which is a pathway, in $_{6,}$ called “abnormal cardiac ventricle morphology” (-$\log_{10}$ p=5.8) from the human phenotype oncology (HP) group. Being the gene scores of them highly significant in our analysis.

On the phenotypic level, we found that tau 2 (total curvature) is least similar to all the others, that tau 3 (total squared curvature) is comparably similar to the DF, and that tau 4-7 (average curvature) form a distinct cluster of similar measurements.

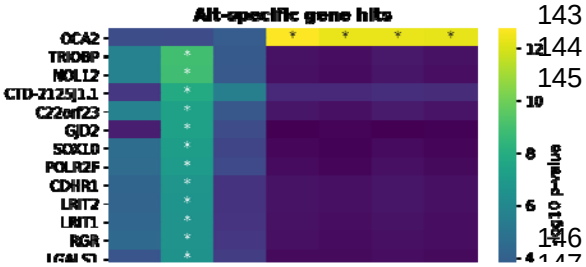
Here, despite all alternative measures having significantly lower SNP-wise heritability than the DF, we found eighteen genes (Figure 2a) and four pathways (Figure 2b) specific to them, i.e. not present in the DF. Among them are genes with potentially relevant annotations, as described by *GeneCards* [24]: OCA2 (Oculocutaneous Albinism 2), the overall top hit, is a determinant of eye color, and associated with albinism. TRIOBP, the most significant specific hit in tau 2, is a structural protein binding to F-actin, and acts as a stabilizer of the cytoskeleton. LGALS1 is implicated in modulation of cell-cell and cell-matrix interactions, and thus a structural protein that might affect, among other things, vessel bendiness. GLIS3 is a zinc-finger protein that plays a role in eye development.

We also identified a cluster of four genes (LRIT1, LRIT2, CDHR1, RGR) on locus 10q23 (Supplemental Figure 2c) through STRING [25], all of which have been associated with eye disorders. RGR has been associated with Retinitis pigmentosa. LRIT1, LRIT2 and CDHR1 have been associated with nanophthalmos, a developmental eye disorder characterized by small eyes. Molecularly, CDHR1 is a calcium-dependent cell adhesion molecule expressed in blood vessels, and thus likely to affect their morphology.

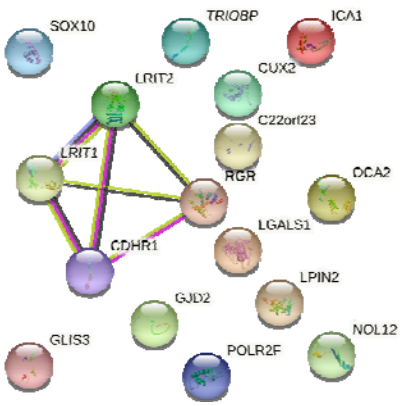
Two out of the four pathways both specific to the average curvature measures, are plausibly related to vascular changes. The first, GO_CELL_JUNCTION_ORGANISATION, is a GO set of genes influencing the tightness of connection between neighboring cells. The second, HP_ABNORMAL_CARDIAC_VENTRICLE_MORPHOLOGY, Human Phenotype Ontology set of genes associated with disease-related abnormalities in cardiovascular tissue. This pathway has three genes strongly driving the signal: 1) RYR1, a sarcoplasmic reticulum calcium release channel, 2) MYOZ2, a sarcomeric protein involved in calcium-dependent signal transduction, and 3) FADD, an apoptotic adapter molecule. Interestingly, no pathways were specific to the most dissimilar measure, tau 2.

In summary, first, this confirms two observations we made on the phenotype level: 1) tau 2 with twelve unique gene hits is the most dissimilar from the others, and 2) tau 4-7 are almost identical and can probably be treated as one measure. Second, we identified relevant genes and pathways specific to these measures, indicating that they may capture disease-relevant vascular changes the DF is not sensitive to.

a.



b.

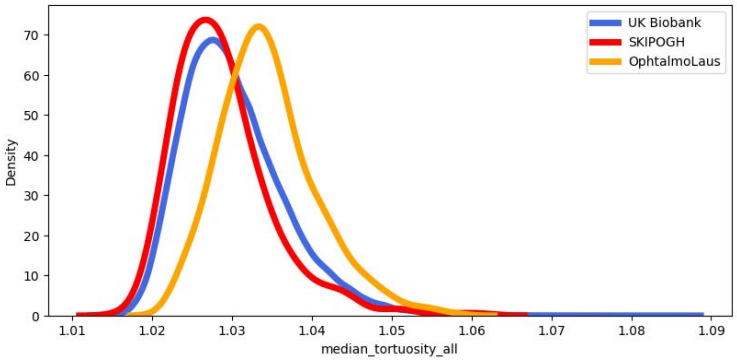


Supplemental Figure 2 | Genes and pathways specific to alternative tortuosity measures. We use a Bonferroni-corrected -log10 p-value significance threshold of 5.71 for genes, and 5.79 for pathways. Significant hits are marked with an asterisk (*). The tables are ordered by maximal significance.

a. All the genes and pathways that are significant in one of the alternative tortuosity measures, but not in the distance factor (DF), are displayed. There are twelve genes specific to tau 2, one to tau 3, and five to the very similar measures tau 4-7.
b. Four pathways are specific to the alternative tortuosity measurements: one in tau 3, three in tau 4-7, but none were found in tau 2. They contain 577, 10, 170, and 394 scored genes respectively.
c. STRING [25] cluster of four genes (LRIT1, LRIT2, CDHR1, RGR), all of which reside on locus 10q23, and have been associated with eye pathologies.

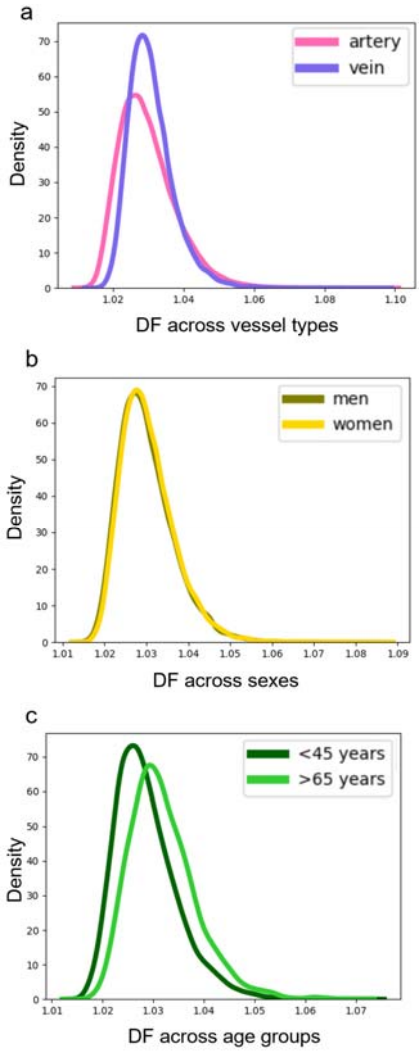
Text 3: Distribution of DF

DF tortuosity across cohorts



Supplemental Figure 3 | Distribution of DF tortuosity across cohorts. UK Biobank mean±SD = 1.030±6.5×10^{-3}; SKIPOGH mean±SD = 1.029±6.2×10^{-3}; *OphtalmoLaus* mean±SD=1.034±6.0×10^{-3}.

Stratified DF analysis: sex, age and vessel type



Supplemental Figure 4 | Median DF between age groups, sexes and vessel types in the Uk Biobank.
a. distribution in arteries (DF = 1.029) and veins (DF = 1.030); Cohen's d = 0.13, $p = 9\times10^{-142}$.
b, distribution in men (DF = 1.0300) and women (DF = 1.0304); Cohen's d = 0.049, $p = 9\times10^{-10}$.
c, median DF increases as a function of age. In the youngest decile (<45 years) DF = 1.028, while in the oldest decile (>65 years) DF = 1.031. Cohen's d = 0.49, $p = 1\times10^{-195}$.

Text 4: Deep Learning classification of arteries and veins

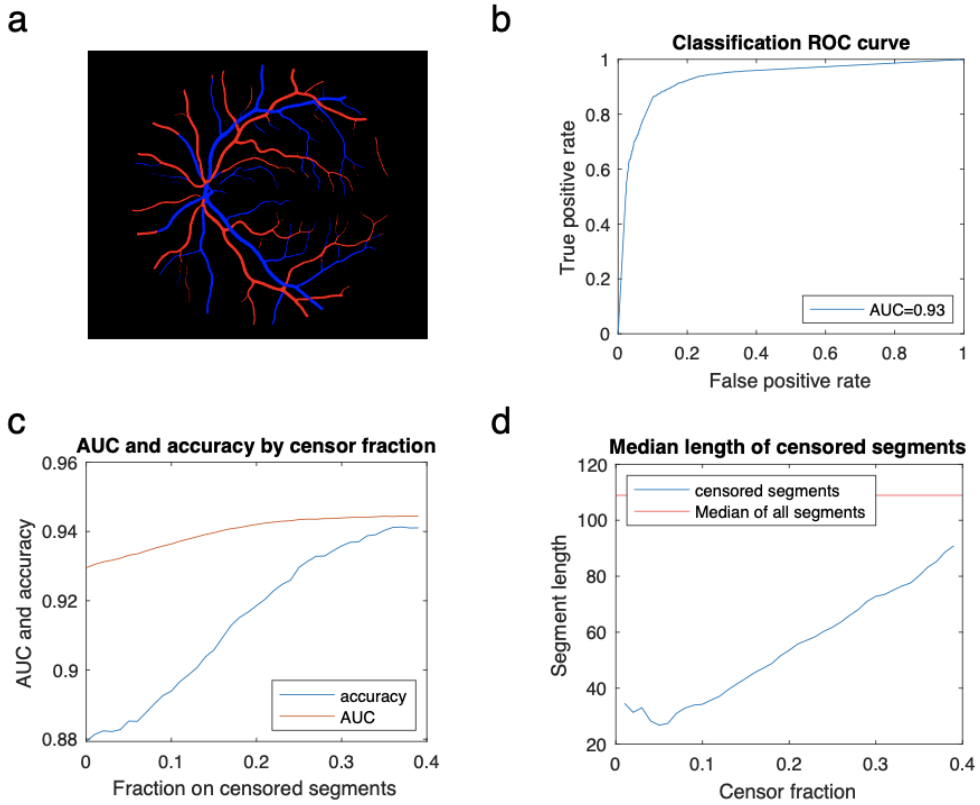
Accuracy of vessel type classification

LWNET [2] converts raw RGB fundus images into a categorical image of three categories, which are defined by the following RGB values: 1) artery: red (255,0,0), 2) vein: blue (0,0,255) and 3) background: black (0,0,0) (Supplemental Figure 5a).

We used LWNET to perform automatic artery-vein classification on our 44 ground truth images (fundus images for which we have manual annotation by ophthalmologist HA), and subsequently extracted the resulting categories for vessel segment centerlines, which we previously computed using ARIA software. For all the centerlines of the 44 images we computed a segment score based on individual centerline pixel classifications as follows:

A ROC curve and its derived AUC measure was computed on the resulting vector of segment scores using logistic regression. This resulted in AUC=0.93. Accuracy was computed simply as:

where T is the number of correctly categorized segments, and F the number of wrongly categorized segments. We chose not to censor any segments, and called classification based on the simple rule that vessels with $S>0$ were called arteries and vessels with $S<0$ were called veins. This resulted in $acc = 0.88$.



Supplemental Figure 5 | AUC and accuracy of artery-vein classification. a) LWNET output image consists of 3 categories: artery, vein and background. b) ROC curve when not censoring any segments, resulting in AUC=0.93 and accuracy=0.88. c) AUC and accuracy could be further improved by censoring the most unclearly classified vessel segments. d) Median length of censored segments is significantly shorter than the overall median length, which could be the reason why we found decreased heritability in our traits in the censoring approach.

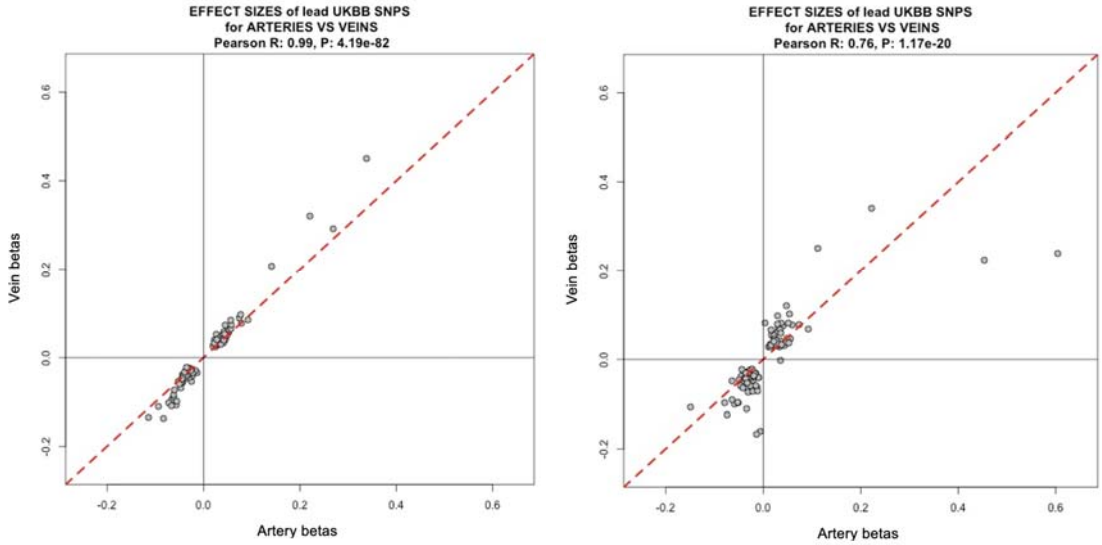
Censoring unclearly classified vessels

Performance could be further increased by removing segments with more uncertain scores (i.e. close to zero). We increasingly removed the vessel segments with the lowest absolute score, and measured how this influenced AUC and accuracy based on the remaining segments. We found that, while AUC shows only moderate increase, accuracy increased from 0.88, without censoring, to almost 0.94 when removing a third of the segments (Supplemental Figure 5c). However, we also found that by doing this we predominantly removed shorter segments (see Supplemental Figure 5d).

We then ran a GWAS based on only the approximately ⅔ most confidently scored segments in the UK Biobank, and found a doubling of the number of Genome-Wide-significant hits for arteries (+117%) (the amount of signal for veins was, on the contrary, not affected). Further analysis, though, showed an inflation in the Q-Q plot of the artery-specific GWAS based on high-confidence vessels only. This was confirmed by analysis of the parameters of the LD Score Regression, which indicated a loss in the ability of the results to explain (SNP)-based heritability of the trait (h_{SNP}^2 dropped from 0.25 to 0.11), coupled with genomic inflation (intercept had increased from 1.01 to 1.93). For these reasons, all vessels identified as arteries or veins were used in the respective vessel-type-specific analysis, as selecting vessels with the highest identification score had brought marginal improvement to the already high AUC at the price of introducing a bias.

GWAS with random vessel type calling

We estimated the independence of the signals arising from the above-described classification of arteries and veins as follows: we modified the pipeline to perform random calling of arteries and veins (by shuffling the vector of artery and veins scores computed for each eye). We then compared the similarity in the signal between two random vessel-type GWAS: we clearly show that the effect sizes (which are significant in at least one of the two GWAS) are nearly identical (r=0.99, p=$4 \cdot 10^{-82}$) when the vessel type calling is random. By comparison, the effect sizes are much less coupled (r=0.76, p=$1 \cdot 10^{-20}$) in the artery- and vein-specific GWAS based on the vessel type calling procedure that has been described

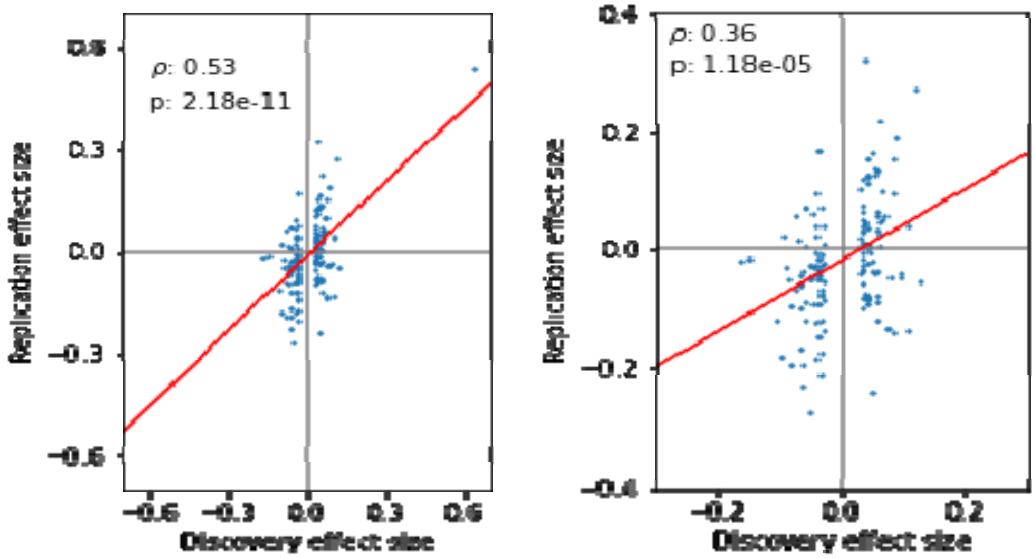


Supplemental Figure 6 | Correlation in the effect sizes from arteries and veins. To support the effectiveness of our vessel type calling procedure, we performed random calling of arteries and veins (by shuffling the vector of artery and veins scores computed for each eye). We then performed two random vessel-type GWAS (left) and compared the results with the artery and vein-specific GWAS (right): we clearly show that the effect sizes (which are significant in at least one of the two GWAS) are less coupled when the vessel type is not random.

Text 5: Replication Analysis

Correlation of effect sizes in the meta-cohort

Of the 136 SNP shared between the discovery and replication studies, the sign of the effect sizes was concordant in 90 (binomial test p = 5.0×10-5). We observed a Pearson correlation of r = 0.53 (p = 1.2×10^{-11}) between the effect size estimates in the two studies. When some outliers are removed, the correlation drops from r = 0.53 to r = 0.36, remaining highly significant.



Supplemental Figure 7 | Extended plot for correlation of effect sizes between discovery and replication cohort. Left: correlation of effect sizes in the discovery and replication cohort as shown in the main text of the manuscript. Right: same plot without considering the outlier. We removed one outlier in the top-right quadrant, corresponding to rs187691758. This shows the robustness of the result: even though the Pearson correlation dropped from r=0.53 to r=0.36, it remained highly significant (p=$1.18\cdot10^{-5}$). Out of the 135 remaining SNPs, 89 shared the sign in effect size (binomial test p=$2.94\cdot10^{-6}$).

Replication of hits in the meta-cohort

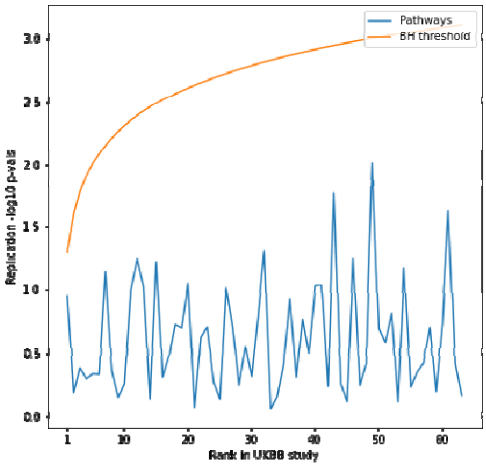
We performed a meta-analysis of the two cohorts OphtalmoLaus (N=514) and SKIPOGH (N=397), using both a fixed-effects model (see below) and a random-effects model [3]. For each SNP, the fixed-effects model computes meta-values of the standard error (SE) and effect size summary statistics, with the meta effect size being a weighted average of individual effect sizes, weighted by their corresponding inverse SE:

——— ——— where ——, and M the number of cohorts

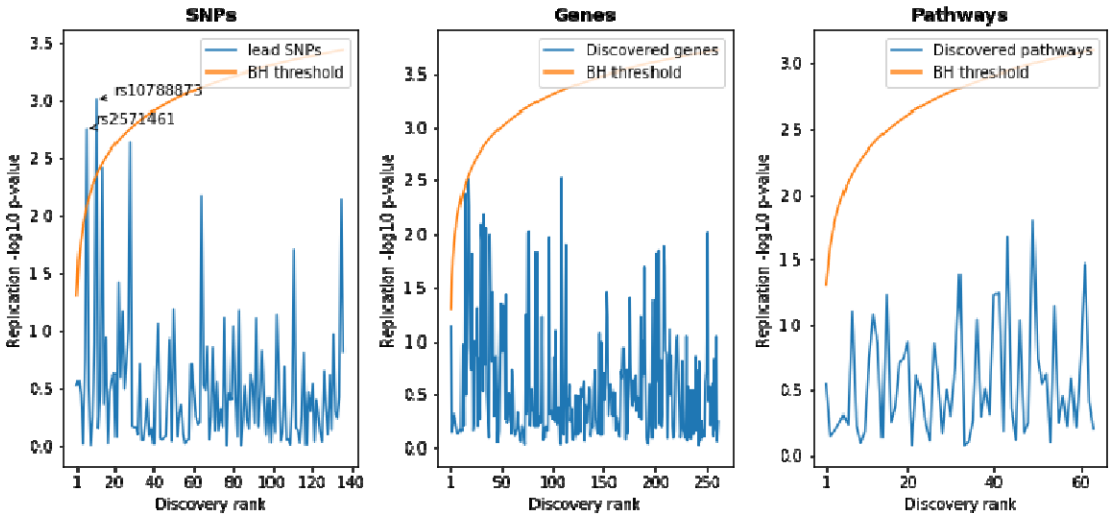
The resulting SNP P-value is then given by a two-tailed t-statistic

——— with degrees of freedom (911 in our case)

The following two figures present the results from both meta-analyses, and the subsequent gene and pathway scoring based on their results. All replication is performed on the combined-vessel median distance factor phenotype.



Supplemental Figure 8 | Replication meta-analysis with fixed-effects. Meta-GWAS summary statistics were obtained using the **fixed-effects** model described above. SNP-level and gene-level summary statistics were reported in the main text. SNP-level summary statistics were aggregated into pathway-level p-values using *PascalX*. Benjamini-Hochberg (BH) procedure on significant pathways from the discovery cohort doesn't replicate any hits in the replication meta-cohort (SKIPOGH + OphtalmoLaus)



Supplemental Figure 9 | Replication meta-analysis with random-effects. Meta-GWAS summary statistics were obtained using the **random-effects** meta-analysis tool RE2C. Left: Benjamini-Hochberg (BH) procedure on lead SNPs discovered in the UK Biobank yields 4 hits in the replication cohort. Center and Right: No discovered genes and pathways replicate in the meta-cohort. SNP-wise p-values were aggregated into gene and pathway scores using PascalX.

SUPPLEMENTAL RESULTS

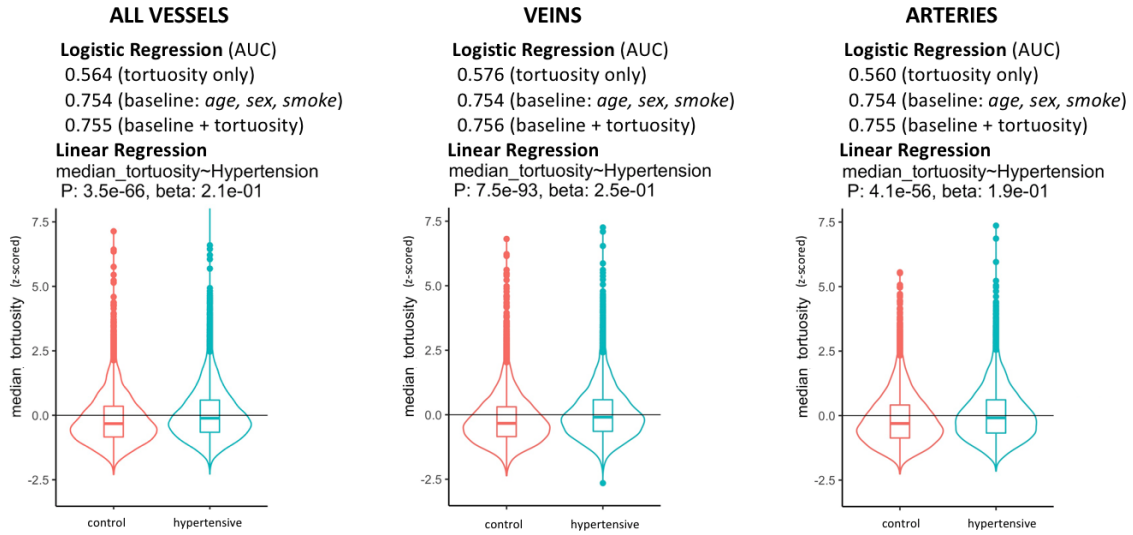
Text 6: Baseline Characteristics

Across the analyzed individuals, mean±SD age = 56±8 years; 35□098 females at birth (54%); 4□618 smokers (7%), mean±SD BMI = 27±5 kg/m2, mean±SD SBP = 140±20 mmHg, mean±SD DBP = 82±11 mmHg. 54□343 (94%) self-reported ethnicity as White, 1□243 (2.1%) as Asian, 962 (1.7%) as Back, 373 (0.6%) as Mixed, 175 (0.3%) as Chinese, and 521 (0.9%) as Other.

Among the participants for which at least one retinal fundus image was available, 2□644 had been diagnosed with type 2 diabetes, 1□448 with angina, 1□077 with myocardial infarction, 1□072 with deep-vein thrombosis (i.e., blood clot in the leg), 750 with stroke, 8□797 lived with stage 2 hypertension (i.e., automated reading of blood pressure >90 mmHg diastolic or >140 mmHg systolic); among the participants for which at least one fundus image was available, 2□644 had been diagnosed with type 2 diabetes, 1□448 with angina, 1□077 with myocardial infarction, 1□072 with deep-vein thrombosis (i.e., blood clot in the leg), 750 with stroke, 8□797 lived with stage 2 hypertension (i.e., automated reading of blood pressure >90 mmHg diastolic or >140 mmHg systolic).

Text 7: Correlation with disease status

We built a logistic regression classifier and found retinal vein tortuosity to have predictive power over disease outcome: angina AUC 55.2%, heart attack AUC 53.4%, stroke AUC 54.6%, Deep Vein Thrombosis (DVT) AUC 53.3% and hypertension AUC 56.6%. To determine whether retinal tortuosity might be used as an independent biomarker for CVD, we trained logistic regression models with known risk factors: age, sex, SBP and smoking (pack years): angina AUC 76.2%, heart attack AUC 80.6%, stroke AUC 69.3%, DVT 61.8%, The same procedure was applied to the prediction of hypertension, but without using SBP as a risk factor, resulting in AUC 75.4%. These risk factors models did not significantly increase in performance by adding any of the median vessel tortuosity measures: we conclude that, despite associations with CVD outcome, retinal tortuosity does not represent an increased health risk after correcting for known risk factors. The analysis was performed on all vessels, then repeated only on veins and only on arteries. Results varied slightly. To illustrate this, we show the distributions of median vessel tortuosity in hypertensive patients vs. controls.



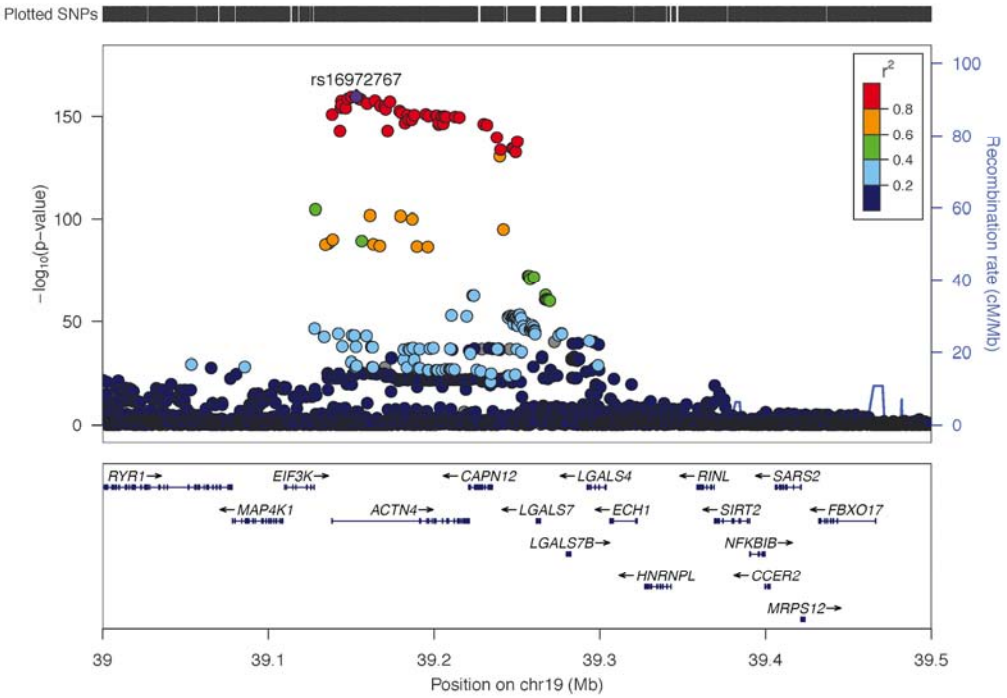
Supplemental Figure 10 | Predictive power of median tortuosity over hypertension. Effects and p-values are calculated using linear regression. AUC refers to logistic regression.

Text 8: Replication of known hits

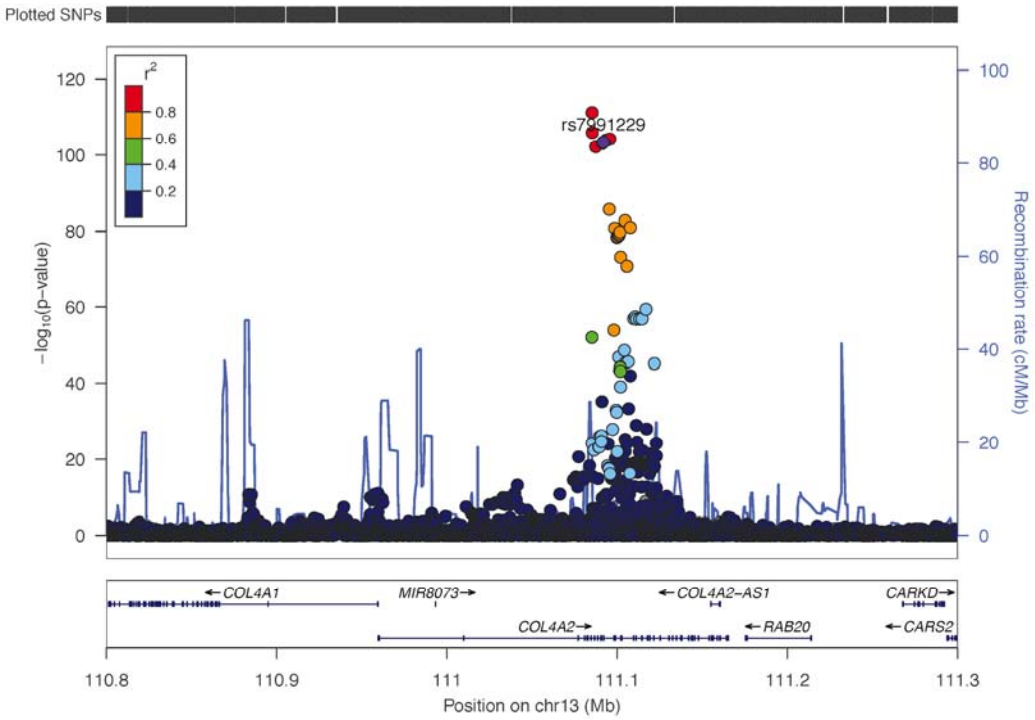
We replicated two known associations. We failed to replicate a third, for which association was controversial (it was only marginally significant in its discovery cohort and had failed replication in the independent cohort of the study that originally proposed it).

RSID	REPLICATED in our GWAS	REPLICATED in its own study	POSITION	GENE(S)	trait
rs1808382	yes	yes	19: 38.7 Kb	ACTN4 / CAPN12	retinal venular tortuosity
rs7991229	yes	yes	13: 111.1 Kb	COL4A2	retinal arteriolar tortuosity
rs73157566	no	no	12: 129.5 Kb	NLRP9P1	retinal venular tortuosity

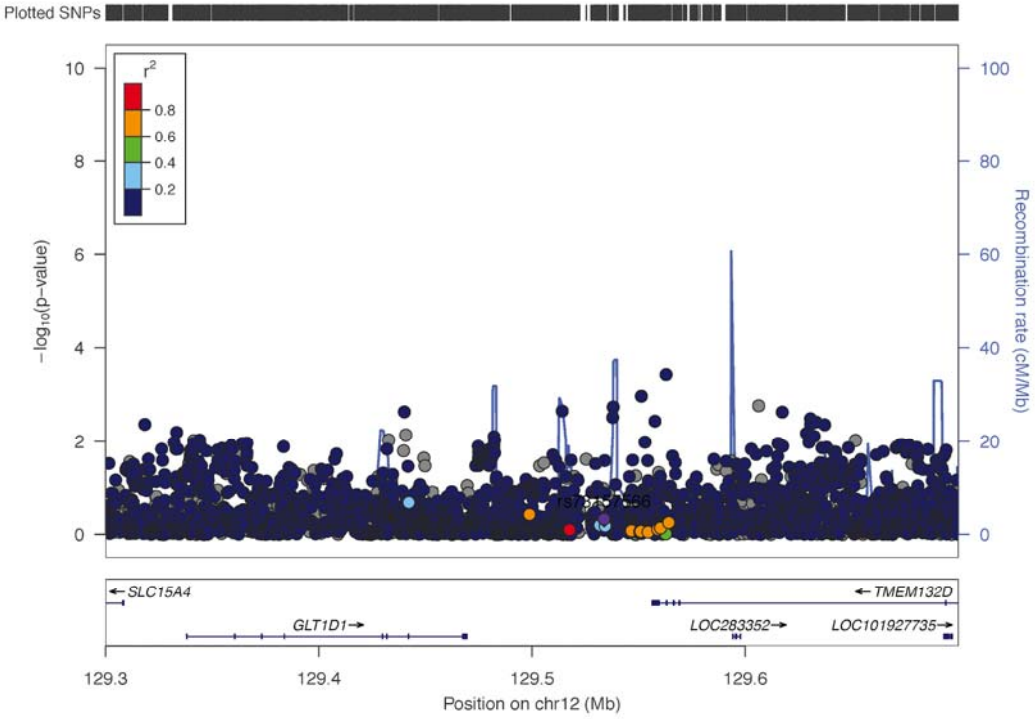
Supplemental Table 2 | Known associations with retinal vessel tortuosity. Three SNPs known to associate with a phenotype related to retinal vessel tortuosity in the literature[4]. Details can be found in the list of all statistically significant SNPs (see Supplemental Dataset 1A/1B/1C).



Supplemental Figure 11 | LocusZoom of known associations with rs1808382. Although we did not recover the exact rsid variant, we report a number of exonic variants (more likely to be causative) in extremely strong LD, which represent our strongest signal: in particular, we recovered a variant in perfect LD, rs16972767, which is an intron variant for ACTN4, the gene on which rs1808382 was reported as having potential direct regulatory effects[4]. rs16972767 is reported in the literature as a venular tortuosity hit: in accordance with this, its significance was higher in our venular tortuosity GWAS ($-\log_{10}$ p=165) rather than in our arteriolar tortuosity GWAS ($-\log_{10}$ p=65).



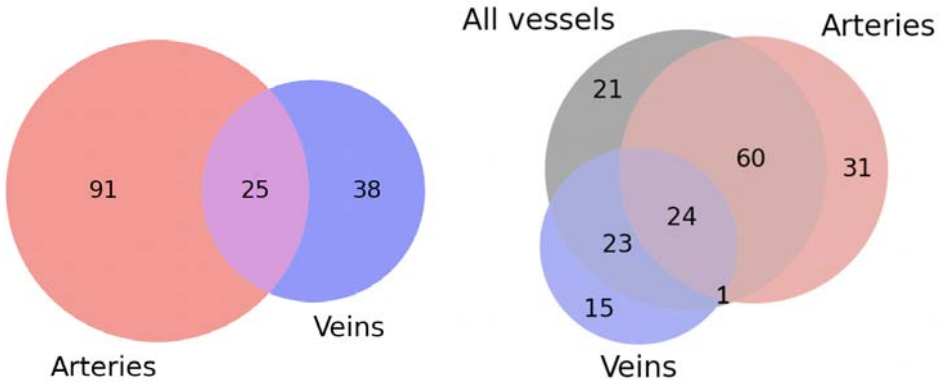
Supplemental Figure 12 | LocusZoom of known associations with rs7991229. This SNP was reported in the literature as an association to arteriolar tortuosity: indeed, its significance was substantially higher in our artery tortuosity GWAS ($-\log_{10}$ p=166) than in our vein tortuosity GWAS ($-\log_{10}$ p=8).



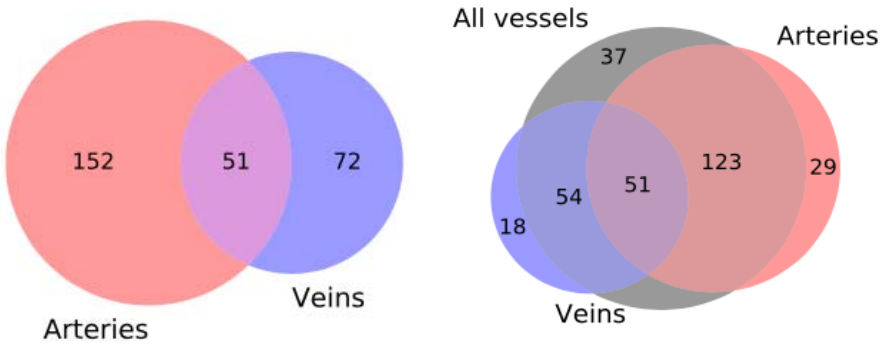
Supplemental Figure 13 | LocusZoom of (controversial) known associations with rs73157566. We did not reproduce it. Given the fact that this association had not reached genome-wide significance in its own replication cohort, we propose this locus should not be considered as associated retinal vessel tortuosity.

Text 9: Vessel-type comparisons for SNPs, Genes and Pathways

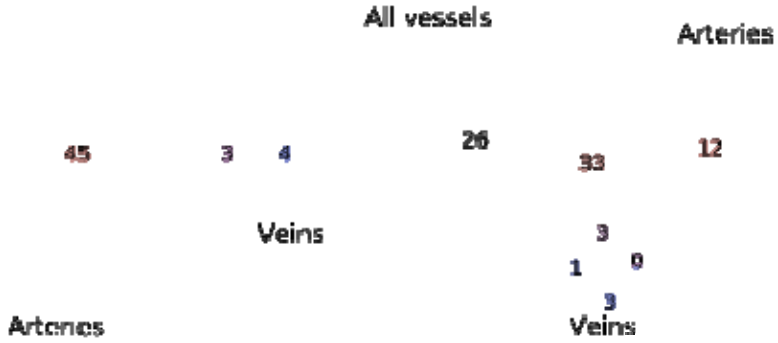
1) SNPs



2) Genes



3) Pathways



Supplemental Figure 14 | Venn diagrams. Overlap between Bonferroni-thresholded 1) SNPs, 2) genes and 3) pathways across phenotypes. Left: **Hits exclusive to artery GWAS *vs.* vein GWAS:** Venn diagram showing overlap between significant 1) SNPs, 2) genes, 3) pathways in arteries and veins. We find overall more significant SNPS, genes and pathways in arteries, and also more than twice as many artery-exclusive SNPs, genes and pathways than vein-exclusive. Right: **Hits exclusive to artery GWAS *vs.* vein GWAS *vs.* all-vessels (combined-vessel-type GWAS):** Combining arteries and veins into a single “all vessels” phenotype catches many of the SNPs, genes and pathways detected by the artery and vein phenotypes and also provide some news.

Text 10: Genetic associations with disease and risk

Tortuosity variants associated with disease outcome

SHARED SNP	ref	DISEASE GWAS
rs3184504	[5]	Coronary heart disease
rs2472299	[6]	
rs10483727	[7]	Glaucoma
rs3184504	[8]	
rs35155027	[8]	
rs10483727	[7]	
rs34935520	[7]	
rs10774625	[9]	Hypertension
rs6495127	[9]	
rs3184504	[10]	Type 1 diabetes
rs3184504	[11]	
rs653178	[12]	Chronic kidney disease
rs653178	[13]	Celiac disease
rs4766578	[12]	Heart failure
rs4766578	[14]	Vitiligo
rs1129038	[14]	
rs653178	[12]	Chronic kidney disease

Supplemental Table 3 | Variants associated with disease. List of variants influencing both tortuosity and disease outcome without statistically significant association to a gene.

SHARED SNP	ref	DISEASE GWAS
rs78058190	[15]	Cardiovascular diseases
rs10942863	[15]	
rs62434109	[15]	
rs3184504	[15]	
rs11072508	[15]	
rs189349094	[15]	
rs72938315	[15]	Respiratory diseases
rs7687906	[16]	Obesity-related traits
rs3184504	[10]	Autoimmune traits
rs7310615	[15]	

Supplemental Table 4 | Variants associated with disease (general outcome). List of variants influencing both tortuosity and general non-specific disease outcomes (e.g.,"Obesity-related traits" does not specify the exact disease outcome).

Tortuosity variants associated with disease risk factors

SHARED SNP	ref	RISK FACTOR
rs35155027	[8]	Intraocular pressure (open-angle glaucoma)
rs8053277	[8]	Vertical cup-disc ratio (glaucoma)
rs3184504	[17]	Systolic blood pressure (CVD)
rs1378942	[18]	
rs3184504	[17]	Diastolic blood pressure (CVD)
rs3184504	[18]	
rs653178	[19]	
rs653178	[20]	Mean arterial pressure (CVD)
rs597808	[15]	Waist-hip ratio (obesity)
rs78058190	[15]	
rs4744056	[21]	General cognitive ability (mental disorders)

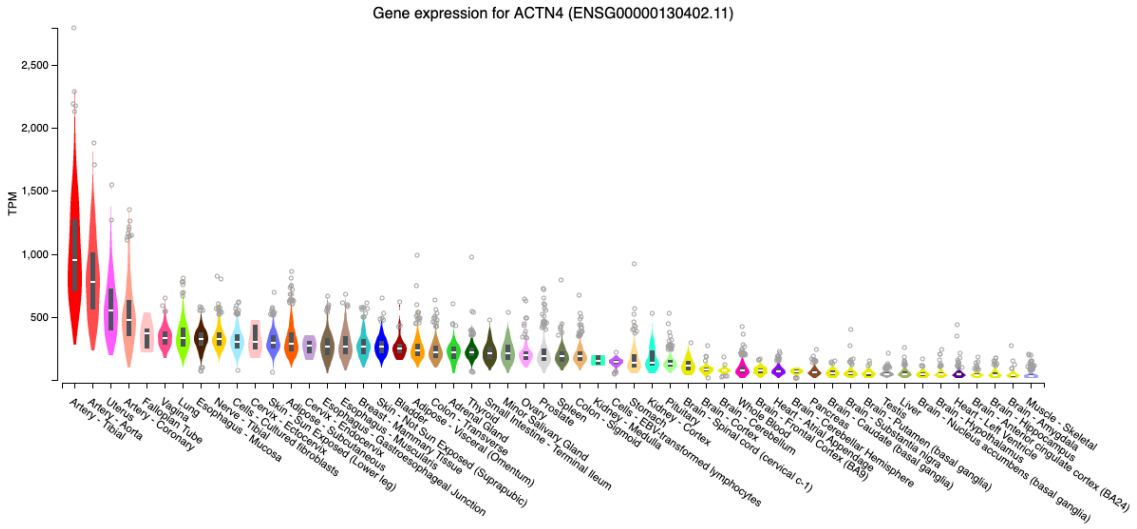
Supplemental Table 5 | Variants associated with risk of disease. List of variants influencing both tortuosity and disease outcome without statistically significant association to a gene.

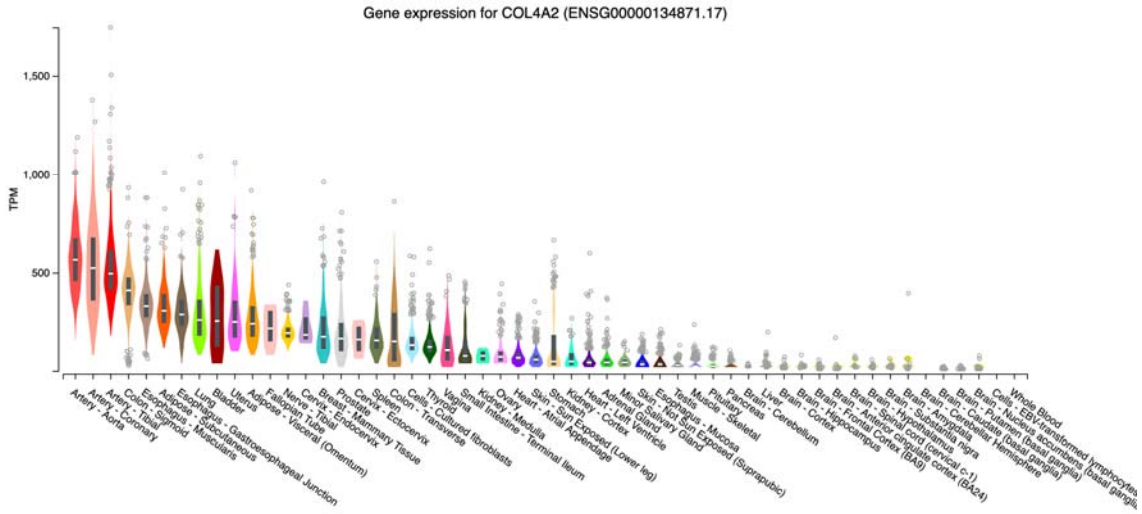
Text 11: Mendelian Randomization

Tortuosity -> Outcome					Exposure -> Tortuosity				
Exposure	**Outcome**	**Effect**	**SE**	**Pvalue**	**Exposure**	**Outcome**	**Effect**	**SE**	**Pvalue**
artery		0.019	0.015	0.200		artery	0.028	0.030	0.356
mixed	BMI	0.008	0.015	0.581	BMI	mixed	-0.011	0.029	0.715
vein		-0.044	0.018	0.013		vein	-0.047	0.027	0.078
artery		-0.053	0.041	0.200		artery	-0.018	0.028	0.524
mixed	CAD	-0.061	0.038	0.109	CAD	mixed	-0.008	0.024	0.748
vein		-0.076	0.046	0.097		vein	-0.002	0.016	0.924
artery		-0.200	0.362	0.581		artery	-0.017	0.017	0.324
mixed	SBP	0.033	0.344	0.924	SBP	mixed	-0.015	0.015	0.309
vein		-0.035	0.538	0.948		vein	-0.010	0.009	0.252
artery		0.017	0.015	0.236		artery	0.010	0.008	0.235
mixed	HDL	0.014	0.015	0.351	HDL	mixed	0.011	0.009	0.193
vein		0.009	0.024	0.722		vein	0.008	0.009	0.383
artery		-0.001	0.016	0.940		artery	-0.046	0.013	0.001
mixed	LDL	0.006	0.017	0.709	LDL	mixed	-0.043	0.013	0.001
vein		0.002	0.029	0.945		vein	-0.031	0.013	0.018
artery		0.000	0.017	0.991		artery	0.000	0.002	0.947
mixed	TC	0.006	0.019	0.741	TC	mixed	0.000	0.002	0.960
vein		-0.006	0.032	0.844		vein	0.000	0.002	0.963
artery		-0.006	0.013	0.656		artery	-0.038	0.018	0.038
mixed	TG	0.003	0.013	0.833	TG	mixed	-0.034	0.016	0.038
vein		-0.008	0.022	0.720		vein	-0.014	0.014	0.301

Supplemental Table 6 | Bi-directional Mendelian Randomization results. Causal estimates are based on the inverse variance weighted (IVW) method. Abbreviations: Body Mass Index (BMI), Coronary artery disease (CAD), Systolic Blood Pressure (SBP), High-density lipoprotein (HDL), Low-density lipoprotein (LDL), total cholesterol (TC) and triglycerides (TG). No multiple hypothesis correction was applied.

Text 12: ACTN4 and COL4A2 over-expression

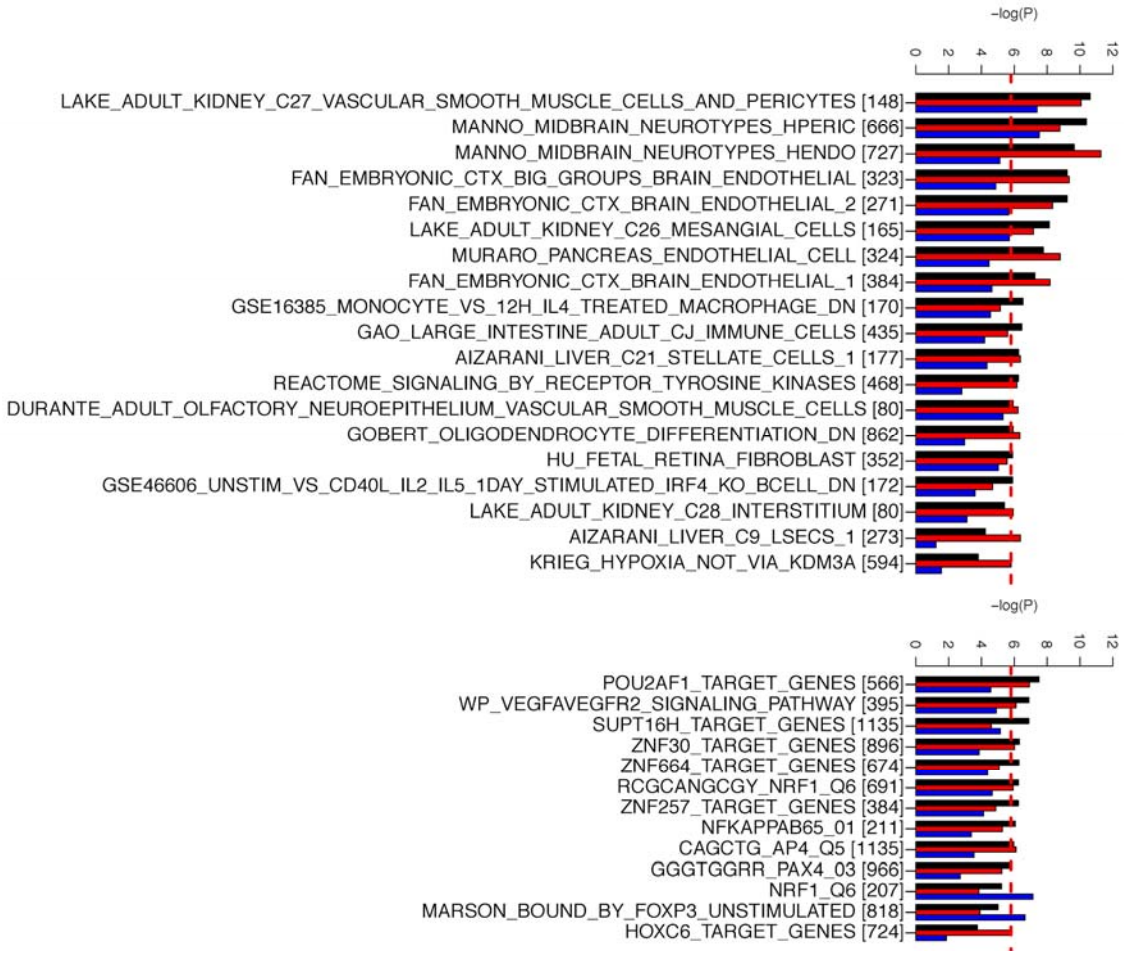




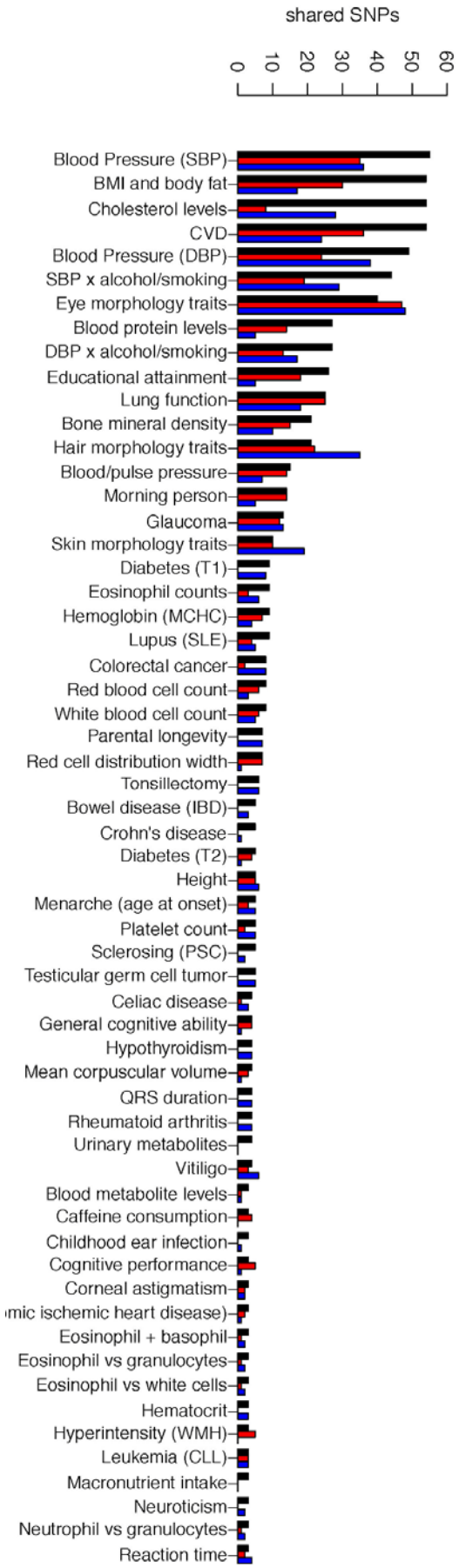
Supplemental Figure 15 | ACTN4 and COL4A2 abundance according to GTEx. The two highly significant genes' ACTN4 and COL4A2 mRNAs were both found to be highly abundant in blood vessels.

Text 13: Full gen set enrichment results

Supplemental Figure 16 | Extended version of Figure 5 from the main text (GO terms). Extended version where all enriched GO terms are shown (irrespective of them being very general or redundant) . Additionally, label names are printed out in full. The number of genes in each set is indicated in squared brackets.



Supplemental Figure 17 | Extended version of Figure 5 from the main text (pathways). Extended version where all enriched pathways are shown (irrespective of them being very general or redundant) . Additionally, label names are printed out in full. The number of genes in each set is indicated in squared brackets.



Supplemental Figure 18 | Extended version of Figure 4 from the main text. Extended version where all traits with at least 3 shared associations are included. For full data, refer to Supplemental Dataset 3.

List of label replacements

Supplemental Figure 18 was generated based on data in Supplemental Dataset 3. The following replacements were applied to the column "Trait" to homogenize or shorten some labels.

"Blood pressure"<-"Blood/pulse pressure"
"Pulse pressure"<-"Blood/pulse pressure"
"Systolic blood pressure"<-"Blood Pressure (SBP)"
"Mean arterial pressure"<-"Blood Pressure (SBP)"
"Diastolic blood pressure"<-"Blood Pressure (DBP)"
"Blood pressure traits (multi-trait analysis)"<-"Blood Pressure (DBP)"
"Body mass index"<-"BMI and body fat"
"Body fat percentage"<-"BMI and body fat"
"Fat-free mass"<-"BMI and body fat"
"Body mass index (joint analysis main effects and physical activity interaction)"<-"BMI and body fat"
"BMI and body fat (joint analysis main effects and physical activity interaction)"<-"BMI and body fat"
"Waist-to-hip ratio adjusted for BMI (additive genetic model)"<-"BMI and body fat"
"Waist-hip ratio"<-"BMI and body fat"
"Hip circumference adjusted for BMI"<-"BMI and body fat"
"Obesity-related traits"<-"BMI and body fat"
"Triglycerides"<-"Cholesterol levels"
"Cholesterol, total"<-"Cholesterol levels"
"LDL cholesterol"<-"Cholesterol levels"
"High density lipoprotein cholesterol levels"<-"Cholesterol levels"
"Low density lipoprotein cholesterol levels"<-"Cholesterol levels"
"Total cholesterol levels"<-"Cholesterol levels"
"HDL cholesterol"<-"Cholesterol levels"
"C-reactive protein levels or LDL-cholesterol levels (pleiotropy)"<-"Cholesterol levels"
"Coronary artery disease (myocardial infarction, percutaneous transluminal coronary angioplasty, coronary artery bypass grafting, angina or chronic ischemic heart disease)"<-"CVD"
"Myocardial infarction (early onset)"<-"CVD"
"Coronary artery disease or ischemic stroke"<-"CVD"
"Coronary artery disease or large artery stroke"<-"CVD"
"Cardiovascular disease"<-"CVD"
"Ischemic stroke"<-"CVD"
"Stroke"<-"CVD"
"Stroke (large artery atherosclerosis)"<-"CVD"
"Stroke (small-vessel)"<-"CVD"
"Myocardial infarction"<-"CVD"
"Ischemic stroke (small-vessel)"<-"CVD"
"Coronary artery disease"<-"CVD"
"Coronary heart disease"<-"CVD"
"CVD (small-vessel)"<-"CVD"
"Type 1 diabetes"<-"Diabetes (T1)"
"Latent autoimmune diabetes vs. type 2 diabetes"<-"Diabetes (T1)"
"Latent autoimmune diabetes vs. type 2 diabetes"<-"Diabetes (T2)"
"Type 2 diabetes"<-"Diabetes (T2)"
"Intraocular pressure"<-"Eye morphology traits"
"Macular thickness"<-"Eye morphology traits"
"Vertical cup-disc ratio"<-"Eye morphology traits"
"Optic cup area"<-"Eye morphology traits"
"Optic disc area"<-"Eye morphology traits"
"Eye color traits"<-"Eye morphology traits"
"Eye color"<-"Eye morphology traits"
"Eye color (brightness)"<-"Eye morphology traits"
"Eye color (hue)"<-"Eye morphology traits"
"Retinal vascular caliber"<-"Eye morphology traits"
"Eye color (saturation)"<-"Eye morphology traits"
"Blue vs. green eyes"<-"Eye morphology traits"
"Blue vs. brown eyes"<-"Eye morphology traits"
"Optic disc parameters"<-"Eye morphology traits"
"Optic disc parameters"<-"Hair morphology traits"
"Hair color"<-"Hair morphology traits"
"Hair morphology traits"<-"Hair morphology traits"
"Blond vs. brown/black hair color"<-"Hair morphology traits"
"Black vs. blond hair color"<-"Hair morphology traits"
"Black vs. red hair color"<-"Hair morphology traits"
"Blond vs. brown hair color"<-"Hair morphology traits"
"Brown vs. black hair color"<-"Hair morphology traits"
"Red vs. brown/black hair color"<-"Hair morphology traits"
"Skin pigmentation traits"<-"Skin morphology traits"
"Skin pigmentation"<-"Skin morphology traits"
"Perceived skin darkness"<-"Skin morphology traits"

"Glaucoma (primary open-angle)"<-"Glaucoma"
"Glaucoma (high intraocular pressure)"<-"Glaucoma"
"Lung function (FEV1/FVC)"<-"Lung function"
"Lung function (FEV1)"<-"Lung function"
"Post bronchodilator FEV1/FVC ratio"<-"Lung function"
"FEV1"<-"Lung function"
"Peak expiratory flow"<-"Lung function"
"Lung function (FVC)"<-"Lung function"
"Lung function (Lung function)"<-"Lung function"
"Mean arterial pressure x alcohol consumption (light vs heavy) interaction (2df test)"<-"SBP x alcohol/smoking"
"Mean arterial pressure x alcohol consumption interaction (2df test)"<-"SBP x alcohol/smoking"
"Blood Pressure (SBP) x alcohol consumption interaction (2df test)"<-"SBP x alcohol/smoking"
"Systolic blood pressure (alcohol consumption interaction)"<-"SBP x alcohol/smoking"
"Blood Pressure (DBP) x alcohol consumption interaction (2df test)"<-"DBP x alcohol/smoking"
"Diastolic blood pressure x smoking status (current vs non-current) interaction (2df test)"<-"DBP x alcohol/smoking"
"Diastolic blood pressure x smoking status (ever vs never) interaction (2df test)"<-"DBP x alcohol/smoking"
"Blood Pressure (SBP) x alcohol consumption (light vs heavy) interaction (2df test)"<-"SBP x alcohol/smoking"
"Blood Pressure (DBP) x alcohol consumption (light vs heavy) interaction (2df test)"<-"DBP x alcohol/smoking"
"Blood Pressure (SBP) (cigarette smoking interaction)"<-"SBP x alcohol/smoking"
"Blood Pressure (DBP) (cigarette smoking interaction)"<-"DBP x alcohol/smoking"
"Systolic blood pressure x alcohol consumption interaction (2df test)"<-"SBP x alcohol/smoking"
"Diastolic blood pressure x alcohol consumption interaction (2df test)"<-"DBP x alcohol/smoking"
"Systolic blood pressure x alcohol consumption (light vs heavy) interaction (2df test)"<-"SBP x alcohol/smoking"
"Diastolic blood pressure x alcohol consumption (light vs heavy) interaction (2df test)"<-"DBP x alcohol/smoking"
"Systolic blood pressure (cigarette smoking interaction)"<-"SBP x alcohol/smoking"
"Diastolic blood pressure (cigarette smoking interaction)"<-"DBP x alcohol/smoking"
"Educational attainment (MTAG)"<-"Educational attainment"
"Highest math class taken (MTAG)"<-"Educational attainment"
"Cognitive performance (MTAG)"<-"Cognitive performance"
"Parental longevity (combined parental attained age, Martingale residuals)"<-"Parental longevity"
"Educational attainment (years of education)"<-"Educational attainment"
"Alcohol consumption (drinks per week)"<-"Alcohol (drinks/week)"
"Mean corpuscular hemoglobin"<-"Hemoglobin (MCHC)"
"Mean corpuscular hemoglobin concentration"<-"Hemoglobin (MCHC)"
"Systemic lupus erythematosus"<-"Lupus (SLE)"
"Morning vs. evening chronotype"<-"Morning person"
"Chronotype"<-"Morning person"
"Primary sclerosing cholangitis"<-"Sclerosing (PSC)"
"Urinary metabolites (H-NMR features)"<-"Urinary metabolites"
"Heel bone mineral density"<-"Bone mineral density"
"Chronic lymphocytic leukemia"<-"Leukemia (CLL)"
"Dietary macronutrient intake"<-"Macronutrient intake"
"Eosinophil percentage of granulocytes"<-"Eosinophil vs granulocytes"
"Eosinophil percentage of white cells"<-"Eosinophil vs white cells"
"Neutrophil percentage of granulocytes"<-"Neutrophil vs granulocytes"
"Sum eosinophil basophil counts"<-"Eosinophil + basophil"
"White matter hyperintensity burden"<-"Hyperintensity (WMH)"
"Inflammatory bowel disease"<-"Bowel disease (IBD)