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

Running Title: Tomasoni et al.; GWAS of Retinal Vessel Tortuosity

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

[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]Dept. of Integrative Biomedical Sciences, University of Cape Town, Cape Town, South Africa; [8]Jules-Gonin Eye Hospital, Lausanne, Switzerland.

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

Abstract

Background: Fundus images of the eye allow for a non-invasive assessment of the retinal vasculature, whose features provide important information on cardiovascular health. In particular, blood vessel tortuosity (or “bendiness”) is a morphological feature associated with hypertension and other common diseases.

Methods: We analysed a total of 175 821 fundus images across 63 662 participants from three cohorts of genotyped subjects, namely the UK Biobank (n=62 751), SKIPOGH (n=397), and OphtalmoLaus (n=514). We set up a fully automated image processing pipeline, using the *ARIA* software for positional annotation of the vasculature and the DL tool *Little W-Net* to determine vessel type, enabling the massively parallel characterisation of such images in terms of seven distinct vessel tortuosity measures for arteries, veins, and combined vessels. Using these measures as traits, we performed the largest Genome-Wide Association Study (GWAS) of vessel tortuosity conducted to date. We calculated gene set enrichment using a novel high-precision statistical method *PascalX*.

Results: Higher tortuosity was significantly associated with the incidence of angina, myocardial infarction, stroke, deep vein thrombosis, and hypertension. Heritability was highest (25%) when using the tortuosity measure known as *Distance Factor* (curve over chord length). Using this measure, we identified 175 significantly associated loci, 173 of which were novel, 4 of which replicated in an independent meta-cohort. Vessel-type specific GWAS revealed 114 and 63 loci for arteries and veins, respectively. Aggregating SNP-wise effects revealed key genes, such as COL4A2, ACTN4, LGALS4, LGALS7, LGALS7B, TNS1, MAP4K1, EIF3K, CAPN12, ECH1, and SYNPO2. Tortuosity genes were overexpressed in arteries and heart muscle, forming pathways modulating the structural properties of the vasculature, and involved in tissue remodelling and angiogenesis. We demonstrated that tortuosity loci served pleiotropic functions as cardiometabolic disease variants and risk factors. In line with this, Mendelian Randomisation revealed causal effects of LDL on tortuosity.

Conclusions: We showed that 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 GWAS 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.

Introduction

Cardiovascular diseases (CVD) are the leading cause of death globally [1–3] and a major societal health burden. While many aspects of CVD development have been acknowledged and many risk factors like age, smoking, and hypertension have been firmly established, the degree of importance of vascular properties as predilection factors is still unclear. The retina is the only place in the body where the microvasculature can be directly inspected. Non-invasive colour retinal photography through the pupil allows the capture of Color Fundus Images (CFI), which provide two-dimensional representations of retinal vasculature. CFIs are used in standard clinical care to screen and diagnose ocular and systemic circulatory diseases (see [4] for review). The retinal fundus is of interest beyond ophthalmology since pathological changes of the retinal vessels often reflect those in the microvasculature of other organs and may precede the progression of systemic microvascular diseases [5]. The retinal vasculature can provide insights into neuro-degenerative diseases, such as Alzheimer's, Parkinson's, and vascular dementia [6–10]. It is also well known that abnormalities in retinal parameters are of diagnostic value for systemic diseases, including increased risk of diabetes [11–14], obesity [15] and cardiovascular disease (CVD) [16–20], specifically stroke [18,21–23], coronary heart disease [24], coronary artery disease [25], peripheral artery disease [26], hypertension [5,13,22,27–37], atherosclerosis [21,22,38], myocardial infarction [39,40], and nephropathies [41,42]. Thus, CFI analysis has the potential of becoming a powerful screening tool for systemic diseases.

In recent years, measuring retinal features from images in large genotyped cohorts has paved the way for studying the genetic underpinning [28] of these phenotypes. Published GWAS using relatively small sample sizes have already identified several loci [13,14] associated with renal vessel size [43–45], optic disc morphology [46,47] and vessel tortuosity [25]. According to the GWAS Catalogue [48], two SNPs, rs1808382 and rs7991229, mapping to the ACTN4 and COL4A2 genes, respectively, have been associated with retinal vessel tortuosity [25] with a p-value below $5 \cdot 10^{-8}$. The association of a third locus with retinal vessel tortuosity [25], rs73157566, is controversial as it was only marginally significant in the discovery cohort and failed to be replicated in the independent cohort of the study that proposed it [25]. There is also a report of SNPs associated with the size of the retinal microvasculature [44], specifically rs7926971 in TEAD1 ($p=3.1 \cdot 10^{-11}$), rs201259422 in TSPAN10 ($p=4.4 \cdot 10^{-9}$), rs5442 in GNB3 ($p=7.0 \cdot 10^{-10}$) and rs1800407 in OCA2 ($p=3.4 \cdot 10^{-8}$). The latter SNP was also associated with central retinal arteriole equivalent ($p=6.5 \cdot 10^{-12}$). These studies demonstrated that GWAS has great potential to reveal genes with a key role for modulating vascular properties and potential pathomechanisms, yet, due to their limited power, revealed at best the tip of the iceberg. Indeed, a recently published study [49] performed a GWAS on vascular density and fractal dimension extracted from 97 895 images across 54 813 participants from the UK Biobank [54], finding 7 and 13 loci associated with these traits, respectively. The most significant association with the vascular fractal dimension was found for a locus containing the OCA2 gene ($p \sim 10^{-80}$), while for the vascular density, the strongest associations were for MEF2C and GNB3.

Here we focused on vessel tortuosity as a retinal trait for GWAS because preliminary analyses indicated smaller genetic contributions of other traits, such as vessel diameter, fractal dimensions and number of branching points (we will report the complete analyses elsewhere). Furthermore, tortuosity is known to be linked to disease [7,11,18,28,50–52]. Our analysis is based on the largest collection of retinal images performed to date: we used images [53] and genotypes from 62 751

subjects of the UK Biobank and 436 and 514 subjects of the much smaller, yet independent, population-based cohorts SKIPOGH [55,56] and OphtalmoLaus [57], respectively. Our fully automated image processing pipeline extended and combined different existing tools to annotate the retinal vessels in each image and assign the vessel type as either artery or vein. Tortuosity was quantified using seven different measures computed for all vessel segments in a given image. We then used the median value of the most heritable tortuosity measure (i.e. the Distance Factor) across arterial, venous and combined segments as our three principle traits to be studied by GWAS, which confirmed two known trait variants and identified 173 new ones. Many of the variants we identified had previously been associated with other traits, specifically CVD, metabolic syndrome and some of their risk factors. Furthermore, our downstream analysis aggregating the SNP-wise association signals onto genes and gene-sets provides new insights into the molecular pathways modulating the structural properties of the retinal vasculature. Importantly, many of the underlying mechanisms and their response to disease risk factors are likely relevant to vascular pathomechanisms in other organs.

Results

Automated processing of retinal images defines vessel tortuosity phenotypes

We processed 175 821 images from 62 751 individuals available in the UK Biobank. We first applied the *ARIA* [58] software for automated vessel annotation of these images. To this end, we modified this software to operate in batch mode. Using classical image processing procedures, *ARIA* delivers a set of identified blood vessel segments for each image, each represented by a list of dense points along the vessel midline. Using these annotated vessel points, we measured the tortuosity for each vessel segment in terms of the so-called *Distance Factor* (DF), i.e. the ratio between the path length $s(C)$ along the segment and the distance $chord(C)$ between its start and end point, as was first suggested in Smedby, O. et al. [59] (see Supplementary text 1). We then used the median retinal vessel tortuosity over all annotated vessels (averaged over multiple images, if available) as our trait (see Methods). Median tortuosity provides a single robust representative value of the tortuosity distribution across vessel segments which is insensitive to outliers (such as segments representing artefacts or erroneous vessel fusions).

Independently, we also applied the recently published retina analysis tool *Little W-Net*, which uses a deep neural network approach [60] exclusively to assign vessel type predictions to each image pixel. Then, using a simple majority vote for all points attributed to a vessel segment by *ARIA*, we classified vessel segments as being part of an artery or a vein. This procedure allowed us to define tortuosity specific to arteries and veins (see Materials and Methods, Classification of vessels into arteries and veins) (see Figure 1 for an overview of our retinal phenotype extraction and imaging GWAS pipeline).

Heritability varies for the different tortuosity definition and vessel types

We used linkage disequilibrium (LD) Score Regression [61] to estimate the SNP-based heritability of tortuosity defined using the *Distance Factor* (DF), as well as that of the six alternative measures. Heritability differed substantially across measures with DF receiving the highest estimate ($h^2_{SNP} = 0.25$, $SE = 0.025$), around double ($0.11 \leq h^2_{SNP} \leq 0.13$, $0.011 \leq SE \leq 0.012$) the heritability of the six alternatives measures (see Supplementary Table 1). For this reason, we adopted DF as

our tortuosity measure for all subsequent analyses. Indeed, the higher heritability of DF is likely to imply that this measure better captures the molecular pathways contributing to vessel tortuosity modulated by genetic variation. We did not observe any significant genomic inflation: the LD Score Regression [61] *Intercept* was 1.01 (SE=0.01), indicating that most of the signal was due to polygenicity, with little influence of confounders (see Table 4 for details). Heritability also varied depending on vessel type: for DF, we obtained $h^2_{SNP} = 0.23$, SE = 0.020 for arteries, and $h^2_{SNP} = 0.15$, SE = 0.021 for veins.

Retinal vessel tortuosity GWAS identify 173 novel loci

Applying linear regression of confounder-corrected, quantile-quantile (q-q)-normalised, median retinal vessel tortuosity on the genotypes of the matching subjects imputed to a panel of about 15M genetic variants, we identified 7 072 significantly associated SNPs in a combined-vessel-type GWAS (Supplementary File 1A). Vessel-type-specific GWAS resulted in 6 563 and 2 896 significantly associated SNPs for arteries and veins, respectively (Supplementary File 1B and C). We applied LD pruning with a threshold of $r^2<0.1$ within a window of 500k bases to define independence. This procedure resulted in the identification of 128 independent loci in the combined-vessel-type GWAS, 116 in the artery-specific GWAS, and 63 in the vein-specific GWAS. In total, we obtained a list of 175 independent lead SNPs (Supplementary text 2.1). The top 10 are listed in Table 1, ordered by statistical significance (for complete listings, see Supplementary File 2).

Among the significantly associated variants, rs1808382 and rs7991229 had previously been reported [62], while the remaining 173 independent lead SNPs represented novel loci associated with this trait (see Figure 2 for a Manhattan plot of our genome-wide signals). According to the GWAS Catalogue [48], a third locus is known for retinal vessel tortuosity [62], namely rs73157566. However, its association is controversial as it was only marginally significant in the discovery cohort and failed to replicate in the independent cohort of the study that proposed it [62]. Similarly, we did not replicate this finding (for details about these three variants and the respective genomic regions, see Supplementary text 3 and Supplementary File 3).

Replication analysis of lead SNPs in a small meta-cohort

We combined data from two small Swiss cohorts, namely *OphtalmoLaus* [57] (n=514) and *SKIPOGH* [55,56] (n=397) to set up a replication meta-study with a total sample size of n=911. While no SNPs reached genome-wide significance in this meta-study, four of the lead SNPs from the UK Biobank (see Figure 2e), as well as three associated genes discovered via *PascalX* (TNS1, AC011243.1, LHFPL2; see Figure 3e), replicated when using the Benjamini Hochberg procedure [63]. Given the limited sample size of the replication meta-cohort, we performed additional analyses on the trend of effect sizes. First, 90 of the 136 lead SNPs from the UK Biobank had the same sign of their effect size in the replication cohort (binomial test $p=4.95\cdot10^{-5}$). Second, we observed a Pearson correlation of $r=0.53$ ($p=1.18\cdot10^{-11}$) between the effect size estimates in the two studies (see Figure 2d). Both results remain significant when removing an outlier (see Supplementary text 4).

Tortuosity genes and pathways affect tissue remodeling and angiogenesis

Using *PascalX* [64], a new high-precision tool to aggregate SNP-wise effects, we obtained significant gene-wise association scores in 312 genes: 203 from the artery-specific GWAS, 123 from the vein-specific GWAS, and 51 genes shared between them. Furthermore, 37 of the 265 genes from the GWAS combining vessel types were specific to this trait (Supplementary text 2.2). We list the top genes in Table 3 (for a complete listing, see Supplementary File 4A/4B/4C).

Specifically, we identified a cluster of genes with very low p-values on chromosome 19, including ACTN4, a gene related to actin filament bundling (previously implicated in vessel tortuosity and disease [25]), TNS1, a functionally related gene involved in cross-linking of actin filaments, and CAPN12, a calcium-activated neutral protease involved in extracellular matrix degradation [65]. Moreover, this locus includes three genes coding for galectins, involved in adhesion to the connective tissue via modulation of cell-cell and cell-matrix interactions [66]: LGALS7, its paralog LGALS7B and LGALS4. We also replicated the highly significant association of tortuosity with the COL4A2 gene (coding for type IV collagen alpha 2 chain) in both arteries and veins. The COL4A1 gene coding for a collagen subunit also received a very low p-value in our artery-specific analysis (replicating previous findings [25]). SYNPO2, an actin-binding gene whose expression has been linked to actin polymerisation and vascular injury [67] as well as ocular growth [68], also received a highly significant association. Finally, among the artery-specific genes, we found FLT1, which plays a role in vessel formation [69].

Pathway scoring using *PascalX* [64] yielded 78 significant pathways (Supplementary text 2c). These pathways included gene-sets annotated as “human retinal fibroblasts”, “vascular smooth muscle cells” (both in the kidney and the neuroepithelium), and “epithelium development”. *PascalX* also identified a pathway containing Vascular Endothelial Growth Factors (VEGF), specifically VEGFA-VEGFR2, a major signalling pathway to activate endothelial angiogenesis and a drug target against age-related macular degeneration. Enrichment signals of three gene-sets annotated as “response to growth factor” further supported the involvement of tortuosity genes in VEGF-mediated angiogenesis. The role of tortuosity genes in the development of blood vessels was supported by significant gene-sets annotated as “circulatory system development”, “anatomical structure morphogenesis”, and “tube development”. Another significantly enriched gene-set was “kidney mesangial cells” (which contract to regulate capillary flow and initiate the formation of convoluted capillaries). Finally, the annotation terms “cell-substrate junction” and “anchoring junction” points to the importance of genes involved in the structural stability of vessels for tortuosity, and several terms related to “actin” and “actomyosin” revealed some of the molecular players involved.

Tortuosity genes are overexpressed in arteries and appendages of the heart

In order to assess whether genes associated with tortuosity are predominantly expressed in specific tissues, we performed a tissue-specific gene expression analysis on 54 tissues from GTEx [70] using *PascalX* [64]. The significance values across tissues are similar between the artery and combined-vessel type analyses revealing that genes associated with tortuosity are overexpressed in three types of arteries (i.e., aorta, tibial artery and coronary artery) and two tissues from the heart (i.e., ventricle and atrial appendage). We also found overexpression of these genes in fibroblasts and muscular tissues (including heart muscles). Interestingly, the enrichment analysis of loci specific to venous tortuosity did not result in any significant tissue associations (for

a strict Bonferroni threshold of p=0.05/54=9.2·10^{-4}), which may be due to the fact that blood pressure is lower in veins, such that structural modulations due to genetic variants may manifest themselves less than in arteries (see Supplementary text 5).

Some tortuosity loci are known cardiac and other disease variants

We queried the GWAS Catalogue [48] for our hits, revealing 9 loci carrying genes previously reported as disease variants: 4 loci were linked to vascular diseases (coronary heart disease, myocardial infarction, arterial hypertension, and diverticular disease), 2 loci were linked to ocular diseases (glaucoma and myopia), and 3 loci were linked to other systemic diseases (chronic lymphocytic leukaemia, type 2 diabetes, and Alzheimer's disease) (see Table 2). We also uncovered 12 loci linked to genes influencing both tortuosity and disease risk factors. Finally, we report 26 additional disease variants which could not be confidently mapped to a specific gene (see Supplementary text 6).

Genetic overlap with cardiometabolic risk factors

We expanded our analysis of overlap with known genetic signals to SNPs belonging to the same LD block (see Figure 4). We observe a sizable number of tortuosity-associated variants overlapping with CVD (54 SNPs). Several traits related to metabolic syndrome [71] 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, 27 for BDP) and type2 diabetes (5 SNPs). In addition, other CVD risk factors share a high number of associated variants with tortuosity, such as protein levels [72] (27 SNPs) and type1 diabetes (9 SNPs). Finally, we detected an overlap with various eye morphology traits, like optic disc morphometry (40 SNPs).

Mendelian Randomisation reveals causal effects of LDL on tortuosity

We performed bi-directional Mendelian Randomisation (MR) [73,74] to search for evidence of causal effects between tortuosity and selected traits within the metabolic syndrome, namely: BMI [75], Coronary artery disease (CAD) [76], SBP [77], and blood lipids [78] (see Methods for details). We used arterial, venous, and combined-vessel tortuosity both as outcome and exposure. Using inverse variance weighting MR, we observed that exposure to elevated levels of low-density lipoprotein (LDL) induced significantly reduced arterial tortuosity (p=0.001). Conversely, increased venous (but not arterial or combined) tortuosity causes higher BMI (p=0.01) (see Supplementary text 7).

Correlations with patient records link tortuosity and disease

Associations between retinal vessel tortuosity and disease have been reported in the literature for arterial stiffness [79] and type 2 diabetes [80]. Measuring the direct association between subjects' retinal tortuosity and their disease status reported in patient medical records (UK Biobank [54], medical phenotypes), we confirmed these associations and obtained five novel associations, namely with angina, myocardial infarction, stroke, deep vein thrombosis (DVT) and hypertension. P-values were obtained via Welch's t-test (comparing the distribution of tortuosity measurements between cases and controls), while effect sizes were estimated through a linear model evaluating disease status as a function of tortuosity. Z-scored median tortuosity of retinal arteries was

significantly associated with hypertension (*beta*=0.19, $p=3\cdot10^{-56}$) and angina (*beta*=0.09, $p=6\cdot10^{-4}$), but not with myocardial infarction, stroke or DVT. Median tortuosity of veins significantly associated with hypertension (*beta*=0.25, $p=7\cdot10^{-99}$), angina (*beta*=0.18, $p=2\cdot10^{-10}$), myocardial infarction (*beta*=0.12, $p=2\cdot10^{-4}$), stroke (*beta*=0.16, $p=5\cdot10^{-5}$) and DVT (*beta*=0.11, $p=5\cdot10^{-4}$). Median venous (but not arterial) tortuosity was significantly associated with myocardial infarction (*beta*=0.12, $p=2\cdot10^{-4}$). Similarly, for stroke and deep vein thrombosis (DVT), a significant association was only observed for veins (stroke *beta*=0.16, $p=5\cdot10^{-5}$, DVT *beta*=0.11, $p=5\cdot10^{-4}$). Overall, the clearest association was measured between hypertension and median tortuosity of both arteries (*beta*=0.19, $p=3\cdot10^{-56}$) and veins (*beta*=0.25, $p=7\cdot10^{-99}$), confirming a known association between tortuosity and systolic blood pressure [79] (see Supplementary text 8).

Discussion

Adapting the retina image processing tool *ARIA* [58] to run on a cluster facilitated the fully automated annotation of retinal vasculatures from colour fundus images from ~63 000 subjects of the UK Biobank. Here we focused on vessel tortuosity as a clinically relevant endpoint and performed Genome-Wide Associations Studies for median tortuosity with substantially increased power compared to previous studies. This gain in power resulted in the identification of 173 novel loci and the replication of all associations known from previous studies, providing a substantially improved picture of the genetic architecture of this trait. Furthermore, we replicated 4 SNPs in a small meta-cohort (n=911).

Our gene analysis revealed multiple signals pointing towards vascular and connective tissue stability in the context of disease, in line with observational studies [50]. One of the strongest genome-wide associations belonged to an actin-related gene: ACTN4, coding for the alpha-actinin-4 protein, mediates actin filament bundling, thereby contributing to cell adhesion and motility and is involved in the tight junction assembly in epithelial cells. Additionally, ACTN4 has recently been associated with vasorelaxation [81] (the widening of vessels) which lowers blood pressure, a mechanism that, when malfunctioning, can lead to hypertension. The ACTN4 tortuosity lead SNP (rs1808382) is also independently associated with Coronary Artery Disease [62], underlining the usefulness of retinal vascular traits as biomarkers for cardiovascular diseases [61]. Another top hit, TNS1, is functionally connected to ACTN4: its protein cross-links actin filaments and localises to focal adhesions (regions of the plasma membrane where the cell attaches to the extracellular matrix).

Among the most significant genes, we also found COL4A1 and COL4A2. These paralogues form the hetero-trimer α1α1α2, a tertiary helical structure that serves as the most abundant fundamental unit of the collagen 4 lattice-like network. This network is a part of the basement membrane and forms an extracellular barrier between tissue compartments [82] that potentially influences blood vessel stability. A missense mutation in COL4A1 has been reported in a case study [83] as the cause of familial retinal tortuosity (a rare disorder with autosomal dominant inheritance [84]). COL4A2 figured among our variants with pleiotropic effects on disease risk (see Table 2). High expression levels of ACTN4 and COL4A2 in relevant tissues from the Genotype-Tissue Expression database [70] further corroborated evidence for their involvement in tortuosity (see Supplementary text 9). Another exciting hit was FLT1—coding for VEGFR1—which is transcribed in most tissues, including arteries and heart (albeit at lower levels than in some tissues like the thyroid [70]). It plays a suppressive role during embryonic vascular development and promotes atherosclerosis during

adulthood [69]. Our strong association with retinal tortuosity suggests that it continues to play a significant role in vessel formation beyond this stage (at least for the retina). The fact that VEGFR-1 is also expressed in macrophages, suggests that it may be one of the drivers for the observed genetic overlap of the tortuosity-associated variants with those of colorectal cancer and inflammatory diseases.

Our GO enrichment results contribute further details to the overall picture of the genetic architecture. Some of the enriched GO categories are “morphogenesis of anatomical structures”, “development of circulatory system”, and “tube development”, as well as GO terms related to the structural stability of cells and tissues. Specifically, we found enrichment of “cell-substrate junction” and “anchoring junction”, responsible for the mechanical attachment of a cell and its cytoskeleton to the extracellular matrix. A related GO category, “actin cytoskeleton”, points to the actin filament molecule in particular. Several of our enriched GO categories point to actin biology (“actin binding”, “actin filament bundle organization”, “positive regulation of actin filament bundle assembly”), in line with the identification of ACTN4, TNS1 and SYNPO2 (cf. above). Part of the involvement of actin in the tortuosity phenotype might involve its interaction with myosin, as suggested by the enriched GO term “actomyosin structure organization”. Indeed, perturbations in processes generating or maintaining the vascular system [85] and their structural integrity are likely to modulate vessel tortuosity by irreversible elongation of vessels, in particular due to elevated blood pressure.

Our pathway analysis also retrieved gene-sets that had been generated by experiments conducted on the retina itself or on organs for which microcirculation is known to develop similarly to that of the retina (namely the kidneys, brain, and pancreas). For instance, we found enrichment in a gene set from transcriptomic analysis of human retinal fibroblasts [86]. Fibroblasts are cells that synthesize structural collagen and the extracellular matrix, which is known to remodel itself in response to signals from the embedded vasculature. Thus, this enrichment suggests that genetic variants specific to fibroblasts may modulate the response of vessel shape to external mechanical factors, such as those induced by elevated blood pressure.

Our analysis further illustrates how mechanical stimuli act on specific tissues inside and around blood vessels, leading to a change in their morphometric properties, which we measure as tortuosity. Specifically, we recovered two gene sets belonging to “vascular smooth muscle cells”, which compose the walls of blood vessels. A dysregulated response in the natural tendency of vascular smooth muscle cells to contract leads to high blood pressure, and their excessive proliferation [87] contributes to the CVD progression: this matches our analysis of shared genetic signals between tortuosity and disease GWAS. The first of these two gene sets referred to the epithelium, the protective layer of cells which lines the outer surfaces of blood vessels (separate support was offered by enrichment for a related GO category, “epithelium development”). The second referred to pericytes, which are mural cells that wrap around the endothelium, lining capillaries and venules. The endothelium itself was the source of four other gene sets. As endothelial cells maintain blood vessel tone, a lack of them or their dysfunction might result in the tissue bulging and deforming, increasing vessel tortuosity. Since endothelial dysfunction is a hallmark of vascular diseases (i.e., coronary artery disease, diabetes mellitus and arterial hypertension), this fits our analysis of direct correlation of tortuosity with patient records, and our analysis of shared genetic signals with disease GWAS. In the retina, the endothelium plays a direct role in angiogenesis, while two major eye diseases may be associated with abnormal neovascularization, namely Age-related Macular Degeneration (AMD) and proliferative Diabetic Retinopathy (DR). Fittingly, we specifically recovered VEGFA-VEGFR2, the major signalling

pathway to activate angiogenesis (with further support from two GO categories related to the regulation of response to growth factors). Intravitreal anti-VEGF injections is a widely used therapy for both neovascular DR and AMD. These findings thus highlight the potential of the present tortuosity GWAS to identify potential drug targets. Further directions of exploration are opened: we recovered a GO category related to the nuclear transforming growth factor-kappa B (NF-kappaB) which interacts with VEGFs and promotes proliferation of vascular smooth muscle cells in several vascular disorders [87]. Moreover, our gene set enrichment suggests a specific mechanism linking TF-mediated angiogenesis and tortuosity in the endothelial cells, namely, tissue stretching in response to elevated blood pressure. Specifically, we found enrichment in a gene set originating from mesangial cells (part of the so-called vascular pole of the kidneys). During development, local endothelial precursors produce a growth factor (PDGF-B) in response to capillary stretching, which induces mesangial cells to initiate the development of convoluted capillaries[88]. Mesangial cells also play a role in circulation once they have reached maturity, contracting to regulate capillary blood flow.

The proposed pathways and GO enrichments confirm the results of our analysis of individual gene functions, which highlighted processes related to cell structure maintenance, adhesion, and smooth muscle function. Furthermore, in line with those results, our tissue expression results indicate that arterial tortuosity genes are over-expressed in arteries, heart, fibroblasts and muscular tissue. Expression of tortuosity genes was generally elevated across most tissue types, indicating they might be required to maintain basic cellular functions (see Supplementary text 5).

We detected pleiotropic effects of 8 tortuosity loci, which we showed to be independently associated with several diseases, specifically: CVD (Coronary Artery disease, Coronary Heart disease, Myocardial infarction, Hypertension), systemic disorders (diabetes, chronic lymphocytic leukaemia, Alzheimer's disease), and eye conditions (myopia, glaucoma). All these traits received further support in our extended analysis of SNPs in high LD. An additional locus was shared with diverticular disease, a condition characterized by a hyperproliferation of the epithelium [89], a tissue to which our gene and pathway analysis clearly pointed. Our results link these diseases through common associated genetic variants to retinal vessel tortuosity. Moreover, the effects obtained from our tortuosity GWAS were sufficiently strong to apply Mendelian Randomisation [73,74], allowing us to establish that elevated LDL exposure causally reduced arterial tortuosity. Indeed, it is well known that high LDL causes atherosclerosis [90], which has been clinically linked to arterial tortuosity [91,92]: atherosclerotic plaque buildup could make arterial walls less flexible and thereby reduce their tortuosity. The potential causal effect of venous tortuosity on BMI we observed is more enigmatic, yet its significance is only marginal and does not survive correction for multiple hypothesis testing.

This study was subject to several limitations. First, the small size of our replication meta-cohort meant that we could provide full replication for 4 of our 173 hits. Nevertheless, the effect sizes in the replication studies correlated strongly with those in the discovery cohort, providing independent evidence that they were not driven by any artefacts specific to the UK Biobank [54]. Further support was provided by a similar trend in the relationships between gene scores in discovery and replication cohorts. Moreover, our gene and pathways analyses give further support for these signals making biological sense. Another limitation is that we did not adjust for spherical equivalent refractive error, which might 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. Indeed, highly powered GWAS may turn out to be a formidable tool to inform method

development in defining the most biologically relevant phenotypes from high-dimensional data, such as images.

Materials and Methods

UK Biobank phenotypes

Our data were collected as part of the UK Biobank [54] effort, a large-scale study that includes over half a million volunteers from the UK (502 505 participants, collection years 2006-2010). It includes a repeated assessment phase (20 000 participants, collection years 2012-2013). The age of participants ranges between 40 and 69 (median age 59), roughly balanced between sexes (229 122 males and 273 383 females).

175 821 fundus eye images (87 562 images of left eyes and 88 259 images of right eyes) were available at the time of data extraction. We processed all images, including those from the reassessment time point. Other phenotypes were used to correct for biases in the genetic associations, namely: age, sex, and 9 Principal Components (PCs) of genotypes (PCs: 1, 2, 5, 6, 7, 8, 16, 17, 18). Other phenotypes were used to study the correlation with disease and lifestyle. The following health statistics are of interest to interpret the medical implications of our analysis: 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); 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).

Data extraction

We used the software ARIA [58] to perform the segmentation and positional annotation of blood vessels. The ARIA parameters were kept at their default values (originally fixed via images from the database REVIEW [95] as part of the tool's system tests). The software is designed to perform vessel diameter measurements at regular intervals along the centerline of each vessel so that the number of measured diameters could be used as a proxy for the total length of the vascular system depicted in one image. We only retained images in which the total number of equally spaced diameters fell between two thresholds: a minimum of 11 000 and a maximum of 20 000 diameters per image. Similarly, we set thresholds on the number of vessels contained in each image: a minimum of 100 and a maximum of 250 vessels per image were required (see Supplementary text 10). By visual inspection, we confirmed the effectiveness of these thresholds in discriminating low-quality images that were too dark, too light, out of focus (lower-bound thresholds), or that contained spurious vessels, i.e., artefacts of the picture that were being erroneously segmented as vessels (higher-bound thresholds). Roughly two out of three images passed this strict quality control, with a total of 116 639 images being sent forward in the pipeline.

Tortuosity measures (see Supplementary text 1) were calculated using the control points through the midline of each annotated vessel. The Distance Factor (DF) [59] is defined as:

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

where the total vessel length, $s(C)$, is divided by the Euclidean distance between the vessel segment endpoints, $chord(C)$. Estimates for $s(C)$ and $chord(C)$ were provided by the *ARIA* software [58]. τ_{DF} is referred to in a recent review as the AOC measure (Arc Over Chord ratio) [93]. Some authors subtract 1 from AOC to obtain a tortuosity measure that vanishes for straight vessels ($s(C) = chord(C)$), but this additive constant is immaterial for the GWAS regression analysis.

Post-processing the data consisted in averaging the values derived from each participant's left and right eye (for the resulting distribution, refer to Supplementary text 11). The data extraction pipeline was written in *Python* and bash and was run on a cluster using the SLURM job scheduler.

Classification of vessels into arteries and veins

We calculated pixel-wise artery and vein classifications by running raw fundus images through a publicly available pre-trained convolutional neural network, *Little W-Net* [60]. We defined a score for each vessel as the difference between *ARIA*'s midline vessel points classified by *Little W-Net* to as arterial (based on the closest pixel) minus those classified as being venous. We used this score to call the type of vessel. Testing our approach on a set of 44 images manually annotated by an ophthalmologist, we obtained an AUC of 0.93 and an accuracy of 0.88. We then performed vessel type classification for the entire set of retina images, allowing us to compute artery- and vein-specific tortuosity values for all vessel segments in each image, from which we calculated the median artery tortuosity and mean vein tortuosity phenotypes (averaged over multiple images, when available). (See Supplementary text 12).

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 (Supplementary Figure 15).

UK Biobank genotype data

Around 488 000 participants were genotyped on Axiom arrays for a total of 805 426 markers. About 96 million genotypes were imputed using a combined reference panel from the 1000 Genomes and UK10K projects [96]. The annotation used to report variant positions is the Genome Reference Consortium Human genome build 37 (GRChb37 or hg19). We subset the genotypes using the software BGENIX [97], shrinking the list of investigated variants to those assigned a rsID (15 599 830 SNPs). An additional Quality Control was performed via a postprocessing step on the GWAS output. We filtered out SNPs with MAF < $5 \cdot 10^{-4}$ (which, given the sample size of 62 751 subjects, translates to an expected minimum number of ~60 individuals having at least one minor allele). We applied a filtering procedure [98] removing SNPs with imputation quality < 0.3.

Genome-wide association analysis

We normalised the raw tortuosity measures extracted from the image data using a rank-based inverse normal transformation, and we ran genetic association studies using the software BGENIE [99]. In order to account for confounding effects [100], the following variables were provided as regression covariates: age, sex, Principal Components (PCs) of the genotypes (only 9 PCs, which showed a statistically significant correlation to tortuosity, were considered).

Using a Bonferroni threshold of $5 \cdot 10^{-8}$, the (unpruned) sets of significant SNPs consist of 6 725 SNPs for the combined-vessel-type GWAS, 6 332 SNPs for the artery-specific GWAS, and 2 895 SNPs for the vein-specific GWAS (see Supplementary File 1A, 1B and 1C).

We extracted all significant genetic associations, and calculated the list of independent SNPs by performing LD pruning using the LDpair function of the R package LDlinkR [101], selecting as a reference panel "GBR" (Great Britain) from the 1k genome project. Two SNPs were considered independent if they had LD $r^2<0.1$ or were more than 500K bases apart (see Supplementary File 2), resulting in 175 independent lead SNPs.

Summary plots were generated using the R packages GWASTools [102].

Replication meta-cohort

We performed replication by performing a meta-analysis of two independent, population-based cohorts: SKIPOGH [55,56] and OphtalmoLaus [57]. SKIPOGH is a family- and population-based cohort including 1 042 participants (493 males and 549 females), aged between 18 and 96 years old, extensively phenotyped at baseline and in a 3-year follow-up. Participants were recruited from three different locations in Switzerland, namely, Bern, Geneva and Lausanne. For 397 participants (186 males and 211 females), retina images from both eyes were available. The genotyping was performed with the Illumina Omni 2.5 chip, followed by an imputation based on the HRC v1.1 panel using Minimach3. The annotation used to report variant positions is the Genome Reference Consortium Human genome build 37 (GRCh37/hg19).

OphtalmoLaus is a substudy of CoLaus, a population-based cohort comprising 6 188 genotyped individuals (3 251 women and 2 937 men) from the Lausanne area, aged 35 and 75 years at the time of recruitment. CoLaus has as its objective to investigate the epidemiology and genetic determinants of CVD risk factors and metabolic syndrome: participants were phenotyped accordingly. Fundus images were available for 514 *OphtalmoLau*s subjects. The genotyping was performed using the 500K Affymetrix chip technology. Annotations followed the same Genome Reference Consortium as SKIPOG and UK Biobank.

Since SKIPOGH includes subjects with a high degree of relatedness, we used the Emmax function of the Epacts software [103] to account for family structure by using the kinship matrix in the model. Additionally, we included the recruitment centre as a covariable. We performed the association analysis for OphtalmoLaus [57] using the same parameters and tools as the discovery cohort. Results from SKIPOGH and *OphtalmoLau*s were meta-analysed using an inverse variance weighting scheme for the respective effect sizes as well as a random-effects model (Supplementary Figure 8).

To keep the multiple hypotheses testing burden low, we took a candidate gene approach. Specifically, we used the Benjamini–Hochberg (BH) procedure [63], 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.

Gene-based tests

We used *PascalX* [64], a novel high precision implementation of the *Pascal* [104] pathway scoring algorithm, 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 arithmetic, *PascalX* computes the corresponding null cumulative probability distribution to essentially arbitrary precision (while other tools only approximate the underlying distribution). With this novel software implementation, we computed p-values up to a precision of $1 \cdot 10^{-100}$, allowing for accurate scoring of genes with contributions from extremely significant SNPs, which become increasingly frequent in highly powered GWAS such as this one.

We used the following configurations: We computed gene scores from SNPs lying within the extended transcript window of +50kb at both the start and end sites. The annotation on the gene positions was based on the Genome Reference Consortium Human genome build 37 (GRCh37/hg19), downloaded from the Ensembl biomart [105]; we considered only protein-coding and lincRNA genes. The reference panel from the UK10K project [106] was used to estimate the SNP-SNP correlations (LD effects). *PascalX* uncovered 265 genome-wide significant genes (after Bonferroni correction for 25 489 gene-based tests $p<0.05/25{,}489=2.0 \cdot 10^{-6}$).

Pathway and GO enrichment

We also used *PascalX* to compute pathway scores based on a gene ranking derived from the gene scores (since a large number of genes have inflated p-values, this ranking approach is more conservative). We first computed scores for 2 868 canonical pathways (BioCarta, KEGG, PID, Reactome, and WikiPathways [107,108]and then extended our analysis to all of the 31 120 pathways in MSigDB (v7.2) [107]. 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"). Scores for these fusion genes were computed using the same parameters as for the ordinary gene scoring.

Tissue-wide gene expression analysis

We performed tissue-wide gene expression analysis using *PascalX* on the whole GTEx [70] (v8) dataset (54 tissues). We defined gene-sets based on the significant genes from each of our three GWAS on median vessel tortuosity (artery, vein and combined-vessel-type). *PascalX* performed an enrichment analysis to calculate whether these sets were over-expressed in any particular tissue. *PascalX* corrected for co-expression of gene sub-clusters within each gene set by merging nearby genes to meta-genes. We computed gene expression values in Transcripts Per Kilobase Million (TPM) for these meta-genes from the raw read counts. The TPM values were uniformised via ranking, transformed to χ^2-distributed random variables, summed, and tested against a χ^2

distribution with as many degrees of freedom as there were "fusion genes" in each set (see Supplementary text 5). We applied a Bonferroni threshold: p=0.05/54=9.2·10^{-4}.

Shared genetic architecture with disease

The SNP variants' overlap with disease phenotypes (matching rsid) was analysed using the GWAS Catalogue [48]. We report independent SNPs in the combined-vessel-type tortuosity GWAS that are part of the GWAS Catalogue [48], and therefore have been previously associated with disease (or with a disease-related phenotype) (see Figure 4 and Table 2). We also list independent SNPs in the combined-vessel-type tortuosity GWAS in LD (r^2>0.8) with SNPs in the GWAS Catalogue [48] (see Figure 4 and Supplementary File 5 and File 6).

Mendelian Randomisation analysis

We performed two-sample bidirectional Mendelian Randomisation (MR) to search for evidence of causal effect between tortuosity and the following traits: BMI, CAD, SBP and lipid traits, namely high-density lipoprotein (HDL), low-density lipoprotein (LDL), total cholesterol (TC) and triglycerides (TG), see Supplementary text 7. For each trait, we used independent (r^2<0.01) significant (P<5·10^{-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 [109]. We only used SNPs on autosomal chromosomes available in the UK10K reference panel [106], which allowed us to estimate the linkage disequilibrium among these SNPs and prune them. We removed strand ambiguous SNPs. Causal estimates were based on the inverse variance weighted (IVW) method [106,110] and were calculated using the Mendelian Randomisation R package [111].

Figures

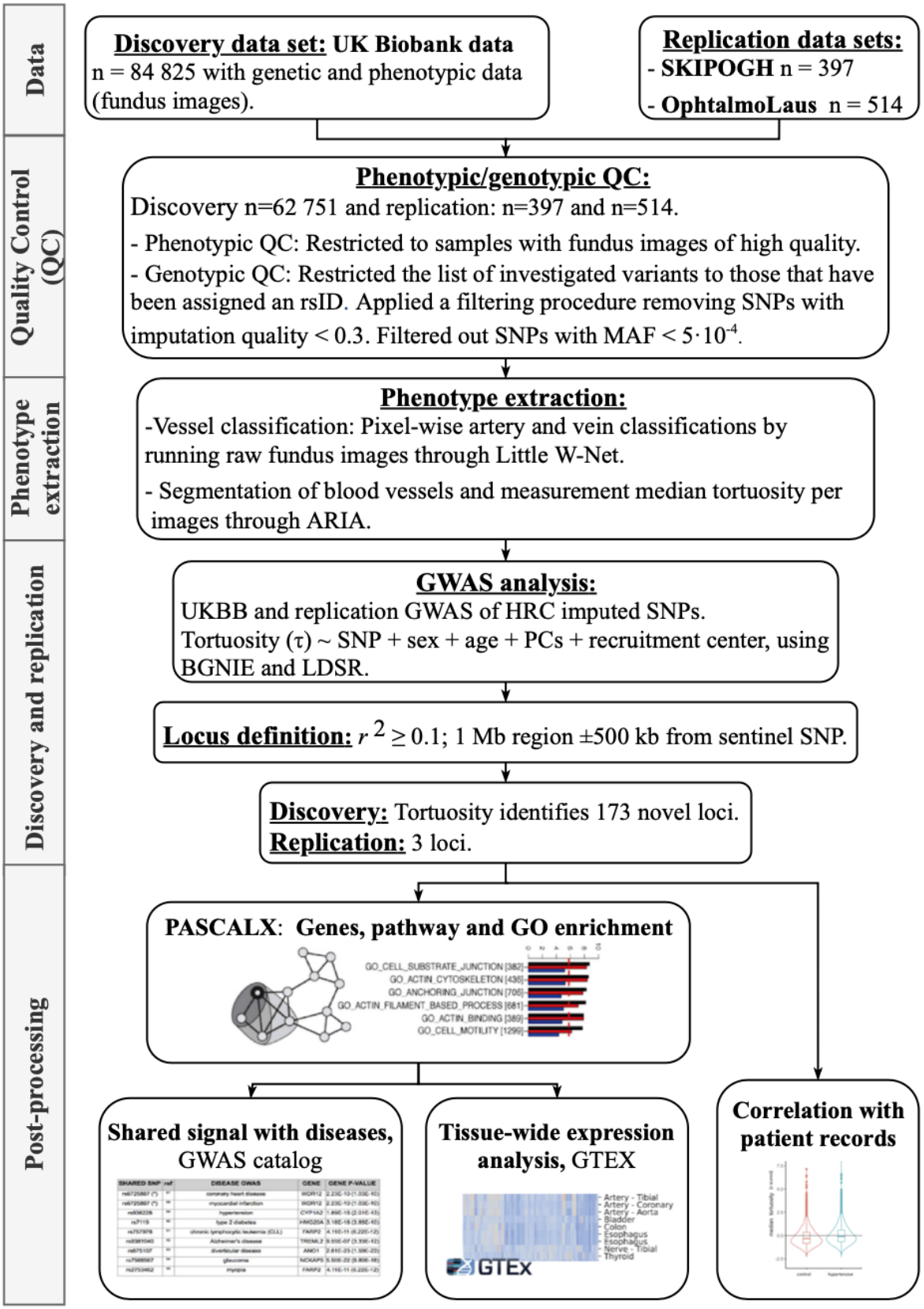
Data
Quality Control (QC)
Phenotype extraction
Discovery and replication
Post-processing
Discovery data set: UK Biobank data
n = 84 825 with genetic and phenotypic data (fundus images).
Replication data sets:
- SKIPOGH n = 397
- OphtalmoLaus n = 514
Phenotypic/genotypic QC:
Discovery n=62 751 and replication: n=397 and n=514.
- Phenotypic QC: Restricted to samples with fundus images of high quality.
- Genotypic QC: Restricted the list of investigated variants to those that have been assigned an rsID. Applied a filtering procedure removing SNPs with imputation quality < 0.3. Filtered out SNPs with MAF < 5·10-4.
Phenotype extraction:
-Vessel classification: Pixel-wise artery and vein classifications by running raw fundus images through Little W-Net.
- Segmentation of blood vessels and measurement median tortuosity per images through ARIA.
GWAS analysis:
UKBB and replication GWAS of HRC imputed SNPs.
Tortuosity (τ) ~ SNP + sex + age + PCs + recruitment center, using BGNIE and LDSR.
Locus definition: r 2 ≥ 0.1; 1 Mb region ±500 kb from sentinel SNP.
Discovery: Tortuosity identifies 173 novel loci.
Replication: 3 loci.
PASCALX: Genes, pathway and GO enrichment
GO_CELL_SUBSTRATE_JUNCTION [382]
GO_ACTIN_CYTOSKELETON [435]
GO_ANCHORING_JUNCTION [705]
GO_ACTIN_FILAMENT_BASED_PROCESS [681]
GO_ACTIN_BINDING [389]
GO_CELL_MOTILITY [1299]
Shared signal with diseases, GWAS catalog
Tissue-wide expression analysis, GTEX
Artery - Tibial
Artery - Coronary
Artery - Aorta
Bladder
Colon
Esophagus
Esophagus
Nerve - Tibial
Thyroid
GTEx
Correlation with patient records

Figure 1 **| Pipeline and results.** Analysis pipeline of this study: First, data were extracted (from the UK Biobank, OphtalmoLaus and SKIPOGH cohorts). Relevant phenotypes (i.e., confounders) and genotypes were collected, as well as fundus eye images. After passing a Quality Control (QC), the images were processed to extract positions of blood vessel segments and classify them as arterial or venous. Then, a range of tortuosity measures were calculated, which provided the phenotypes for combined and vessel-type specific GWAS. 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 the most relevant tissues for these pathways. Correlation analysis revealed the predictive value of retinal tortuosity with patient records on relevant diseases.

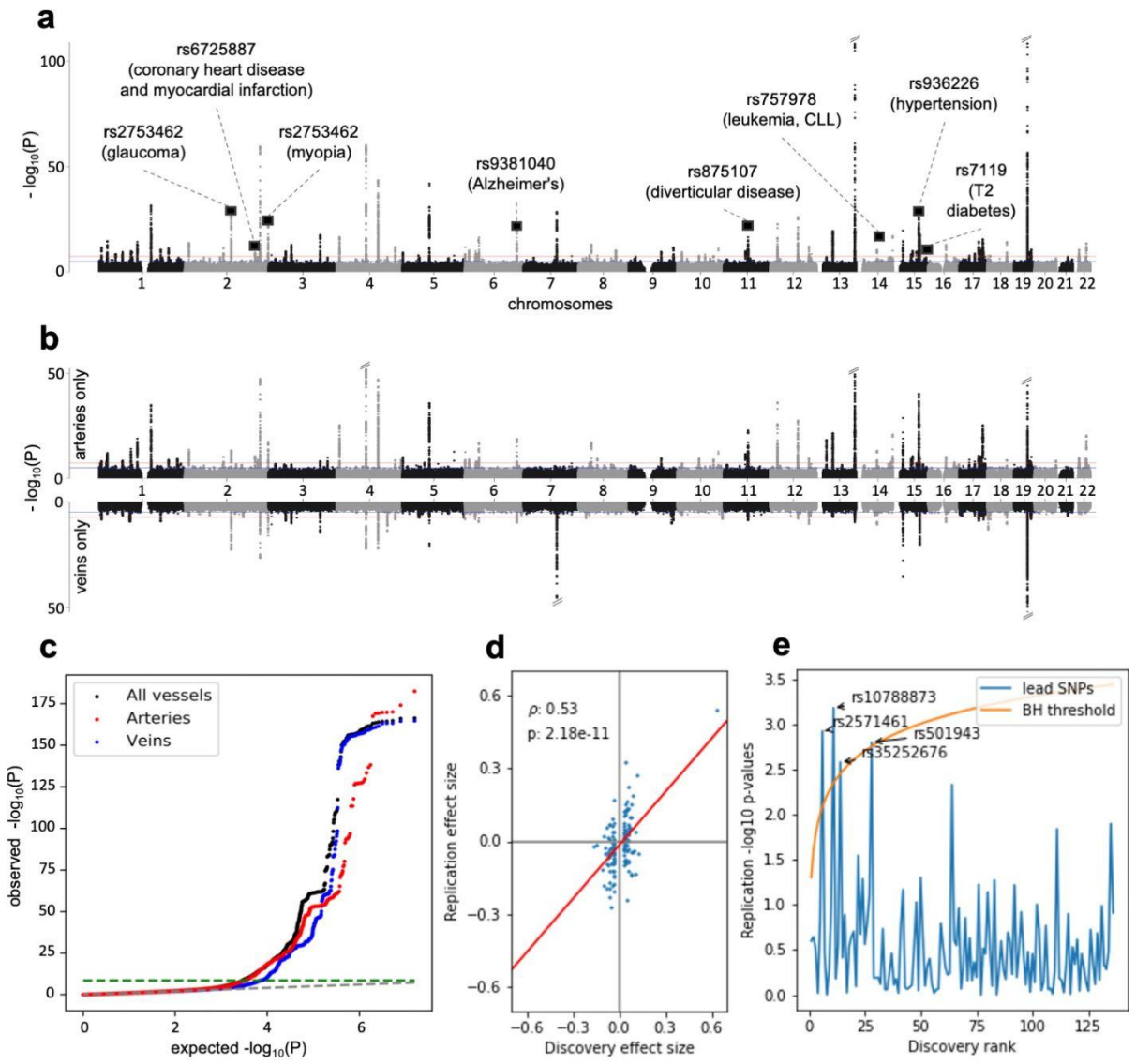

Figure 2 **| SNP p-values and effects. a**, Manhattan plot of Genome-Wide Association Study (GWAS) of median retinal vessel tortuosity, combining all vessel types (both arteries and veins). The red line indicates the genome-wide significance level after Bonferroni correction (p=5·10^{-8}). Oblique dashes on top of peaks mark extremely significant p-values that have been cropped. Squares mark the position of disease SNPs (see Table 2). The trait was corrected for phenotypic variables which showed a statistically significant association, i.e.: age, sex, and a subset of principal components (PCs) 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 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 = 1.18·10^{-11}. **e**, Benjamini-Hochberg (BH) procedure on discovery lead SNPs from the UK Biobank yields 4 hits in the replication cohort.

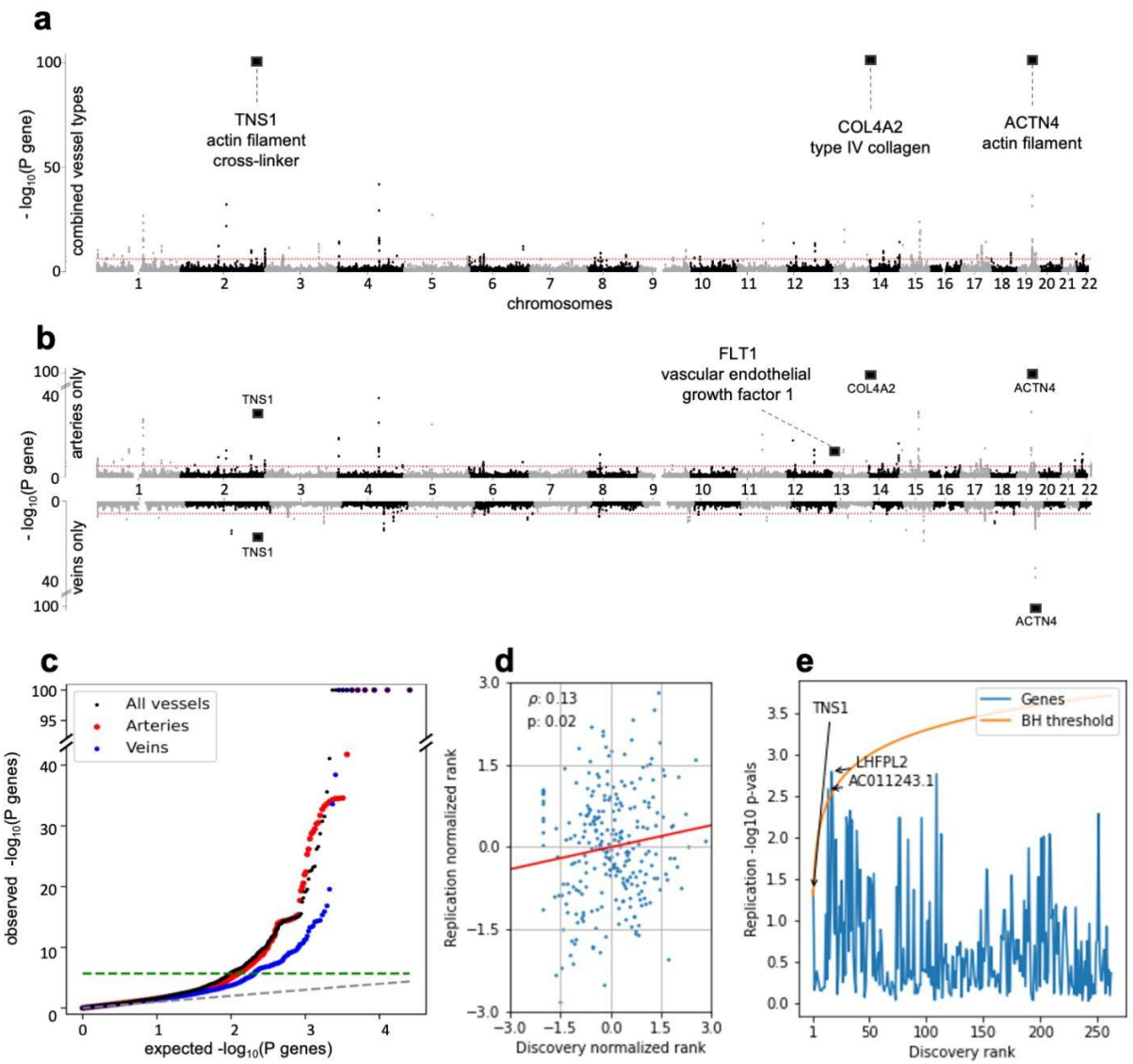

Figure 3 | **Gene p-values and replication scores. a,** Gene-based Manhattan plot of median retinal vessel tortuosity, combining all vessel types (both arteries and veins). Gene-based tests were computed by *PascalX* [64]. The red line indicates the genome-wide significance level after Bonferroni correction (p=5·10^{-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 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.

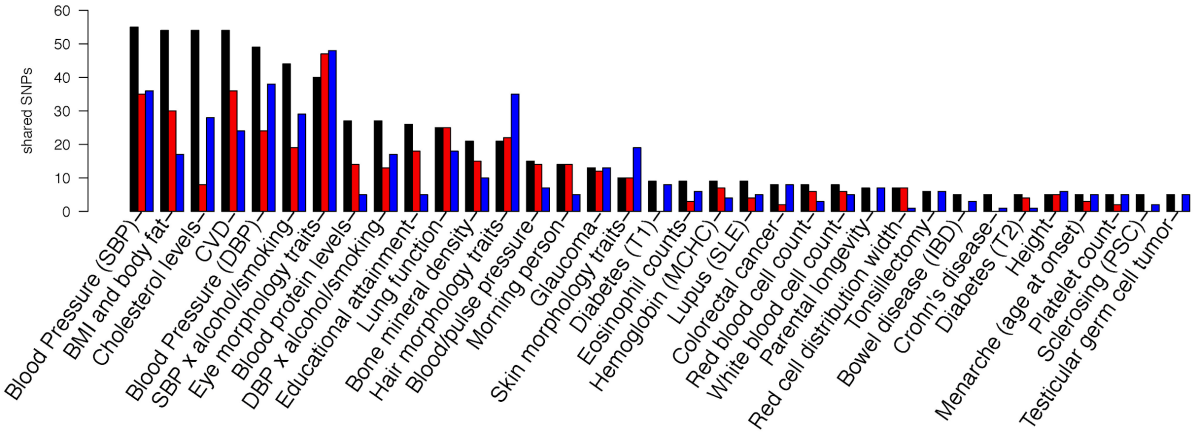

Figure 4 | **Overlap in genetic signals with diseases and other complex traits.** Number of variants shared with other traits reported in the GWAS Catalogue [48] (also considering SNPs in high LD with the lead SNP, r^2>0.8). Only traits with at least 5 shared associations are included (for a full list, including rsIDs, refer to the Supplementary File 6). 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 [112].

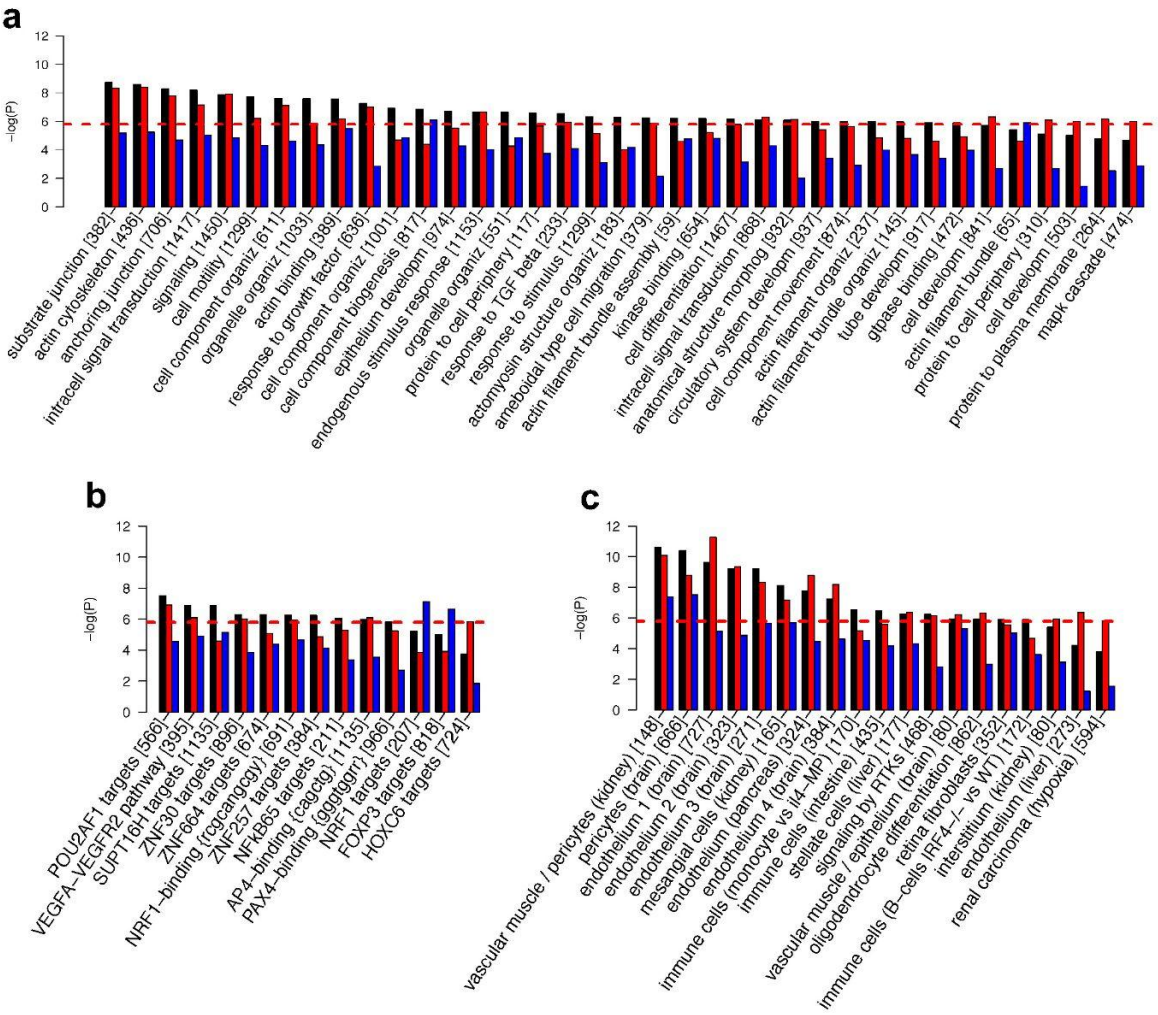

Figure 5 **| Enriched pathways and gene-sets.** Scores for 31,120 gene-sets in MSigDB (v7.2) [107] were calculated by *PascalX* [64]. Red bars correspond to enrichment results from the artery GWAS, blue correspond to veins, and black correspond to the combined-vessel-type GWAS. Only gene-sets for which significance was reached by at least one GWAS are shown. The red dashed line indicates Bonferroni-threshold (-log10 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 Supplementary Material. **a,** Enrichment in Gene Ontology (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.

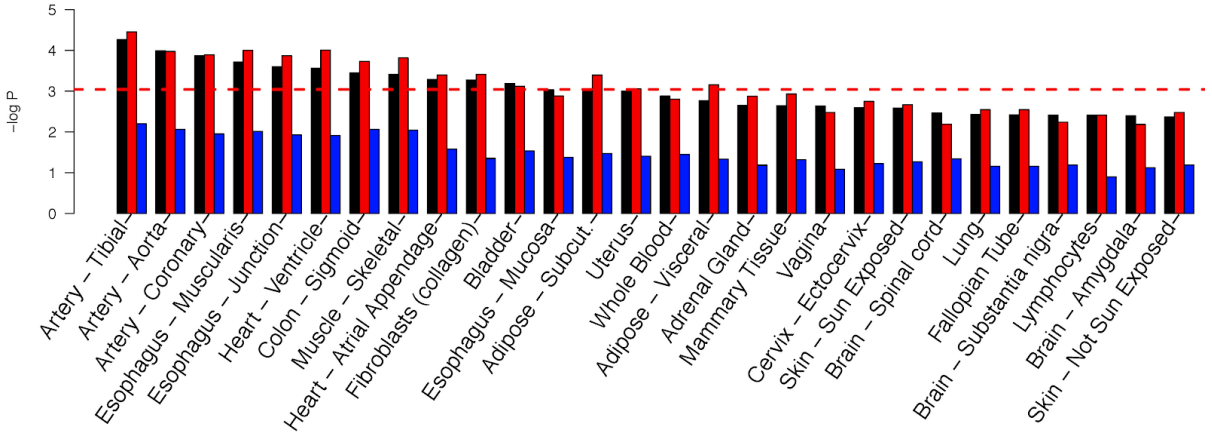

Figure 6 **| Tissue expression results.** Tissue-specific gene expression analysis of GTEx (v8) [70] performed using *PascalX* [64]. We defined sets based on the significant genes from each of the three GWAS we carried out and asked whether they were over-expressed in a particular tissue: red bars represent the estimated significance for artery genes, while significance of vein genes is in blue, and in black os the significance of genes from the combined-vessel-type GWAS.

Tables

Chr	SNP	EA	RA	freq	beta	-Log10 P	signal 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 retinal tortuosity SNPs, ordered by p-value. For full results, refer to the list of 175 independent lead SNPs in Supplementary File 2. Abbreviations: chr: chromosome; EA: effect allele; RA: reference allele; freq: allele frequency of effect allele; beta: effect size estimate; -LOG10 P: normalized p-value in the discovery cohort; -Log10 P (rep): normalized p-value in the replication meta-cohort;

Shared SNP	Disease GWAS	Gene	-Log10 P	ref
rs875107	diverticular disease	ANO1	22.5	[113]
rs7588567	glaucoma	NCKAP5	21.2	[114]
rs7119	type 2 diabetes	HMG20A	17.5	[115]
rs936226	hypertension	CYP1A2	14.7	[116]
rs757978	chronic lymphocytic leukemia (CLL)	FARP2	10.4	[117]
rs2753462	myopia	FARP2	10.4	[118]
rs6725887 (*)	coronary heart disease	WDR12	9.6	[119]
rs6725887 (*)	myocardial infarction	WDR12	9.6	[120]
rs9381040	Alzheimer's disease	TREML2	6.0	[121]
rs11083475	heart rate (rhythm disorders)	ACTN4	> 100	[122]
rs9555695	waist-hip ratio (obesity)	COL4A2	> 100	[123]
rs2571445	lung function (pulmonary disease)	TNS1	> 100	[124]
rs3791979	intraocular pressure (open angle glaucoma)	TNS1	> 100	[125]
rs17263971	eGFR (Chronic Kidney Disease) and retinal dysfunction	SYNPO2	28.7	[67,126]
rs35252676	pulse pressure (CVD)	LHFPL2	26.6	[127]
rs1378942 (*)	diastolic blood pressure (CVD)	CSK	23.4	[128]
rs1378942 (*)	mean arterial pressure (CVD)	CSK	23.4	[129]
rs17355629	pulse pressure (CVD)	LRCH1	19.6	[130]
rs7655064	waist-hip ratio (obesity)	MYOZ2	14.5	[123]
rs6495122	diastolic blood pressure (CVD)	CPLX3	14.5	[131]
rs12913832	intraocular pressure (open angle glaucoma)	HERC2	12.3	[132]
rs9303401	cognitive ability (mental disorders)	PPM1E	10.01	[133]

Table 2 **| Pleiotropic disease-variants.** List of variants identified by the UK Biobank tortuosity GWAS which were independently 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) and variants which we could confidently map to a gene (in the combined-vessel-type analysis, which included both arteries and veins). Gene p-values were computed by *PascalX* [64]. Variants associated with more than one disease are marked by a star (*).

Gene	Chr	BP	-log10 P combined	-log10 P arteries	-Log10 P veins
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 retinal tortuosity genes. P-values were computed by *PascalX* [64] (precision cutoff: $1 \cdot 10^{-100}$). For full results, refer to Supplementary File 4A/B/C.

GWAS type	h^2_{SNP}	*lambda GC*	*mean Chi²*	*intercept*	*ratio*
combined-vessel-types	0.25 (0.025)	1.14	1.31	1.01 (0.01)	0.03 (0.03)
artery-only	0.23 (0.020)	1.12	1.27	1.00 (0.01)	< 0
vein-only	0.15 (0.021)	1.10	1.18	1.00 (0.01)	< 0

Table 4 **| SNP-based heritability.** h^2_{SNP} is the portion of phenotypic variance cumulatively explained by the SNPs. *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 Chi²* 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.

Author contributions

MT and SB designed the study. MT and MJB performed QC on the raw images. MT extracted tortuosity measurements from the image data (UK Biobank and OphalmosLaus and SKIPOGH cohorts) and post-processed them. MJB performed classification of arteries and veins. MT carried out the median tortuosity GWAS (based on the Distance Factor measure), with the guidance of SB, DK, NM and EP. MJB carried out gene and pathway scoring. ALB performed LD Score Regression analysis. MT and MJB performed subsequent bioinformatics analyses. SOV evaluated the correlation between different tortuosity measures and their impact on genetic associations. TC and MT 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 and SB lead the writing of the manuscript with contributions from all other authors.

Competing interests

The authors declare no competing interests.

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.

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.

Additional information

Please refer to the Supplementary Material and files.

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. 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.

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

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

7. 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.

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

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

10. 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.

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

12. 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.

13. 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.

14. 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.

15. 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.

16. 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.

17. Grunwald JE, Pistilli M, Ying G-S, Maguire M, Daniel E, Whittock-Martin R, et al. Retinopathy and the risk of cardiovascular disease in patients with chronic kidney disease (from the Chronic Renal Insufficiency Cohort study). Am J Cardiol. 2015;116: 1527–1533.

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

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

20. Cuspidi C, Sala C, Grassi G. Updated classification of hypertensive retinopathy: which role for cardiovascular risk stratification? Journal of hypertension. 2015. pp. 2204–2206.

21. 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.

22. 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.

23. 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.

24. 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.

25. 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.

26. 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.

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. Grosso A, Veglio F, Porta M, Grignolo FM, Wong TY. Hypertensive retinopathy revisited: some answers, more questions. Br J Ophthalmol. 2005;89: 1646–1654.

30. 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.

31. Wong TY, Klein R, Klein BEK, Meuer SM, Hubbard LD. Retinal vessel diameters and their associations with age and blood pressure. Invest Ophthalmol Vis Sci. 2003;44: 4644–4650.

32. 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.

33. 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.

34. Wang JJ, Mitchell P, Leung H, Rochtchina E, Wong TY, Klein R. Hypertensive retinal vessel wall signs in a general older population: the Blue Mountains Eye Study. Hypertension. 2003;42: 534–541.

35. 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.

36. Wong TY, Hubbard LD, Klein R, Marino EK, Kronmal R, Sharrett AR, et al. Retinal microvascular abnormalities and blood pressure in older people: the Cardiovascular Health Study. Br J Ophthalmol. 2002;86: 1007–1013.

37. 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.

38. 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.

39. 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.

40. 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.

41. 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.

42. 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.

43. Nannini DR, Torres M, Chen Y-DI, Taylor KD, Rotter JI, Varma R, et al. A Genome-Wide Association Study of Vertical Cup-Disc Ratio in a Latino Population. Invest Ophthalmol Vis Sci. 2017;58: 87–95.

44. 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.

45. 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.

46. 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.

47. 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.

48. 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.

49. 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

50. 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.

51. 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

52. 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.

53. MacGillivray TJ, Cameron JR, Zhang Q, El-Medany A, Mulholland C, Sheng Z, et al. Suitability of UK Biobank Retinal Images for Automatic Analysis of Morphometric Properties of the Vasculature. PLoS One. 2015;10: e0127914.

54. 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

55. 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.

56. 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.

57. 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.

58. 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.

59. 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.

60. 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

61. 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.

62. 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.

63. 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

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

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

66. 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.

67. 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.

68. 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.

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

70. 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.

71. Lindsay RS, Howard BV. Cardiovascular risk associated with the metabolic syndrome. Curr Diab Rep. 2004;4: 63–68.

72. Halbesma N, Bakker SJL, Jansen DF, Stolk RP, De Zeeuw D, De Jong PE, et al. High protein intake associates with cardiovascular events but not with loss of renal function. J Am Soc Nephrol. 2009;20: 1797–1804.

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. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518: 197–206.

76. Nikpay M, Goel A, Won H-H, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015;47: 1121–1130.

77. International Consortium for Blood Pressure Genome-Wide Association Studies, Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature. 2011;478: 103–109.

78. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature. 2010;466: 707–713.

79. 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: United Kingdom Biobank. Hypertension. 2019;74: 1383–1390.

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

81. 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.

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. 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.

84. Sutter FKP, Helbig H. Familial retinal arteriolar tortuosity: a review. Surv Ophthalmol. 2003;48: 245–255.

85. Lubarsky B, Krasnow MA. Tube morphogenesis: making and shaping biological tubes. Cell. 2003;112: 19–28.

86. Hu Y, Wang X, Hu B, Mao Y, Chen Y, Yan L, et al. Dissecting the transcriptome landscape of the human fetal neural retina and retinal pigment epithelium by single-cell RNA-seq analysis. PLoS Biol. 2019;17: e3000365.

87. 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.

88. Schell C, Wanner N, Huber TB. Glomerular development--shaping the multi-cellular filtration unit. Semin Cell Dev Biol. 2014;36: 39–49.

89. Tursi A, Brandimarte G, Elisei W, Inchingolo CD, Aiello F. Epithelial cell proliferation of the colonic mucosa in different degrees of colonic diverticular disease. J Clin Gastroenterol. 2006;40: 306–311.

90. 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.

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

92. 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.

93. 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

94. Lisowska A, Annunziata R, Loh GK, Karl D, Trucco E. An experimental assessment of five indices of retinal vessel tortuosity with the RET-TORT public dataset. Conf Proc IEEE Eng Med Biol Soc. 2014;2014: 5414–5417.

95. 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.

96. Chou W-C, Zheng H-F, Cheng C-H, Yan H, Wang L, Han F, et al. A combined reference panel from the 1000 Genomes and UK10K projects improved rare variant imputation in European and Chinese samples. Sci Rep. 2016;6: 39313.

97. Band G, Marchini J. BGEN: a binary file format for imputed genotype and haplotype data. doi:10.1101/308296

98. 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.

99. 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

100. 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.

101. 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

102. Gogarten SM, Bhangale T, Conomos MP, Laurie CA, McHugh CP, Painter I, et al. GWASTools: an R/Bioconductor package for quality control and analysis of genome-wide association studies. Bioinformatics. 2012;28: 3329–3331.

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

104. 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.

105. 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.

106. 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.

107. 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.

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

109. 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/

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

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

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

113. 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.

114. 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.

115. 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.

116. 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.

117. 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.

118. 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.

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

120. 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.

121. 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.

122. 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.

123. 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.

124. 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.

125. 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

126. 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.

127. 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.

128. 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.

129. 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.

130. 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.

131. 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.

132. 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.

133. 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.