

Dissecting the Genetic Architecture of Intracranial Aneurysms

Shaunak S. Adkar^{1,2,3}, Julie Lynch^{4,5}, Ryan B. Choi⁶, Tanmoy Roychowdhury⁷, Renae L. Judy^{8,9},
Kaavya Paruchuri^{10,11,12}, Dong-Chuan Go¹³, Sharika Bamezai^{2,14}, John Cabot^{1,2,3}, Sabina
Sorondo^{1,2,3}, Michael G. Levin^{15,16}, Dianna M. Milewicz¹³, Cristen J. Willer^{17,18,19}, Pradeep
Natarajan^{10,11,12,20}, Saiju Pyarajan^{21,22}, Kyong-Mi Chang^{23,24}, Scott Damrauer^{25,26,27,28}, Phil
Tsao^{3,29,2}, Stephen Skirboll^{3,30}, Nicholas J. Leeper^{1,2,29*}, Derek Klarin^{1,3*}

¹Department of Surgery, Division of Vascular Surgery, Stanford University School of Medicine, Palo Alto, CA

²Stanford Cardiovascular Institute, Stanford University, Palo Alto, CA

³Veterans Affairs (VA) Palo Alto Healthcare System, Palo Alto, CA

⁴Veterans Affairs (VA) Salt Lake City Healthcare System, Salt Lake City, UT

⁵Epidemiology, School of Medicine, University of Utah, Salt Lake City, UT

⁶Stanford University School of Medicine, Palo Alto, CA

⁷Department of Genetics, Yale School of Medicine, New Haven, CT

⁸Department of Surgery, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

⁹Research, Corporal Michael J. Cresenz VA Medical Center

¹⁰Department of Medicine, Harvard Medical School, Boston, MA

¹¹Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA

¹²Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA

¹³Department of Internal Medicine, Division of Medical Genetics, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX

¹⁴University of Michigan School of Medicine, Ann Arbor, MI

¹⁵Department of Medicine, Division of Cardiovascular Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

¹⁶Department of Medicine, Corporal Michael J. Cresenz VA Medical Center, Philadelphia, PA

¹⁷Department of Internal Medicine, Division of Cardiology, University of Michigan School of Medicine, Ann Arbor, MI

¹⁸Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI

¹⁹Department of Human Genetics, University of Michigan, Ann Arbor, MI

²⁰Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA

²¹Center for Data and Computational Sciences, VA Boston Health Care System, Boston, MA

²²Research, Harvard Medical School, Boston, MA

²³Department of Medicine, Division of Gastroenterology and Hepatology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

²⁴Research & Medicine, Corporal Michael J. Cresenz VA Medical Center, Philadelphia, PA

²⁵Department of Surgery, Corporal Michael J. Cresenz VA Medical Center, Philadelphia, PA

²⁶Department of Surgery, Division of Vascular Surgery and Endovascular Therapy, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

²⁷Department of Genetics, University of Pennsylvania Perelman School of Medicine,

1 Philadelphia, PA

2 ²⁸Penn Cardiovascular Institute, University of Pennsylvania Perelman School of Medicine,
3 Philadelphia, PA

4 ²⁹Department of Medicine, Stanford University School of Medicine, Stanford, CA

5 ³⁰Department of Neurosurgery, Stanford University School of Medicine, Palo Alto, CA

6 * Contributed equally

7

8 **Corresponding Author:**

9 Derek Klarin, MD

10 dklarin@stanford.edu

11 3801 Miranda Avenue; Building 100 B3-106B, Palo Alto, CA 94304

12 Tel: (650) 493-5000 x64996

13 Fax: (650) 852-3430

14

1 ABSTRACT

2

3 Background: The genetic risk of intracranial aneurysm (IA) development has been ascribed
4 largely to the genetic risk of smoking exposure and hypertension. However, the relationship of
5 IA to other cardiovascular traits and the contribution of IA risk loci to aberrant gene programs
6 within cerebrovascular cell types remains unclear.

7

8 Methods: We performed a genome-wide association study in the Million Veteran Program
9 testing association of roughly 25 million DNA variants with unruptured IA (3,165 cases and
10 592,927 controls) in veterans of European, African, and Hispanic ancestries. This was meta-
11 analyzed with publicly available summary statistics to yield a final cohort of 15,438 cases and
12 1,183,973 controls. Candidate causal genes were prioritized through expression quantitative
13 trait loci colocalization, fine-mapping transcriptome wide association studies, and multi-trait
14 colocalization. We constructed a cerebrovascular single nucleus RNA sequencing (snRNA-seq)
15 dataset and integrated IA summary statistics to prioritize candidate causal cell types. We then
16 constructed a polygenic risk score to identify patients at greater risk of developing IA.

17

18 Results: We identified five novel loci association with IA, increasing the number of known
19 susceptibility loci to 22. At these susceptibility loci, we prioritized 16 candidate causal genes.
20 Amongst other cardiovascular traits, we found a significant positive genetic correlation of IA
21 with coronary artery disease and abdominal aortic aneurysm, and we identified 13 IA risk loci
22 that colocalize with aneurysmal, atherosclerotic, and blood pressure traits. Integration of an IA
23 gene set with human cerebrovascular snRNA-seq data revealed significant association with
24 matrix-producing pericytes, smooth muscle cells (SMCs), and other mural subtypes. In addition,
25 gene expression analysis revealed enrichment of several candidate genes within SMCs and
26 pericytes. Finally, a high polygenic risk score (PRS) was significantly associated with IA across
27 European (OR: 1.87, CI: 1.61-2.17, $P = 8.8 \times 10^{-17}$), African (OR: 1.62, CI: 1.19-2.15, $P = 1.2 \times 10^{-3}$),
28 and Hispanic (OR: 2.28, CI: 1.47-3.38, $P = 1.0 \times 10^{-4}$) ancestries.

29

30 Conclusion: Here, we identify five novel loci associated with IA. Integration of summary statistics
31 with cerebrovascular snRNA-seq reveals association of cell-types involved in matrix production.
32 We constructed and validated a PRS that predicts IA, while controlling for demographic
33 variables including smoking status, sex, and blood pressure. Taken together, our findings suggest
34 that an intrinsic deficit in matrix production and vascular integrity may drive IA pathogenesis
35 independent of systemic hypertension and smoking exposure.

36

37

38

1
2 INTRODUCTION

3
4 Intracranial aneurysms (IAs) are dilations of the cerebral arterial vasculature. These
5 aneurysms represent a major cause of cerebrovascular morbidity, and the most feared
6 complication – aneurysm rupture - results in intracranial hemorrhage, carrying a mortality of
7 50%.¹ While a minority of patients develop syndromic forms of IAs due to connective tissue
8 disorders, such as Marfan Syndrome, Ehlers Danlos Syndrome, or adult dominant polycystic
9 kidney disease (ADPKD), 80-90% of patients with IA have no known predisposing genetic
10 syndrome or immediate family history.²⁻⁴ In these patients, IA risk, similar to that of many
11 complex traits, appears to be mediated by a combination of polygenic inheritance and
12 environmental factors. The only multi-ancestry genome wide association study (GWAS) to-date
13 estimated a narrow-sense genetic liability of 21% through common single nucleotide
14 polymorphisms (SNPs).⁵ Hypertension and smoking, two of the strongest clinical risk factors for
15 development of IA, not only share substantial genetic overlap with IA but likely have a causal
16 genetic role in IA development as demonstrated by Mendelian Randomization.⁶

17 Although the clinical risk factors such as age, hypertension, smoking, and female sex
18 have been well documented, the underlying etiology of IA and its relationship to other related
19 cardiovascular traits such as atherosclerosis or aortic aneurysm is not well understood.
20 Epidemiological studies have described an increased likelihood of IA development in patients
21 with aneurysms in other vascular beds, suggesting common genetic or environmental risk
22 factors.⁷⁻¹¹ In addition, recent work has suggested common pathways and drug targets with
23 atherosclerosis and abdominal aortic aneurysms (AAA), including *PCSK9*.^{12,13} Appreciation of
24 the shared genetic architecture of IA with these cardiovascular traits may help reveal other
25 shared therapeutic targets.

26 To this end, we sought to expand upon known IA genetic susceptibility loci by performing
27 a GWAS leveraging genetic data of patients in the Million Veteran Program (MVP), a precision

1 medicine effort by the Department of Veteran’s Affairs to study how genes affect human health
2 and disease. This cohort is enriched for patients with vascular pathologies and is ethnically
3 diverse with patients from Hispanic and African ancestries in addition to European ancestries.^{13–}
4 ¹⁵ We then used a combination of expression quantitative trait locus (eQTL) colocalization, fine-
5 mapping transcriptome-wide association study (TWAS), and multi-trait colocalization to prioritize
6 candidate causal genes and variants. We also assessed the pleiotropy of IA susceptibility loci,
7 the genetic correlation of IAs to other cardiovascular traits, and the putative causal cell types
8 involved through analysis of human cerebrovascular single nucleus RNA-sequencing (sn-RNA
9 seq) data. Finally, we developed and validated the utility of a polygenic risk score (PRS)
10 controlling for concomitant clinical risk factors to predict IA diagnosis.

11

12 METHODS

13 *Study Populations*

14 We conducted a discovery genome-wide association study (GWAS) using DNA samples and
15 phenotypic data from the MVP and FinnGen (DF8), a publicly available biobank of Finnish
16 participants’ clinical and genetic data. In MVP, individuals aged 18 to over 100 years have been
17 recruited from 63 Veterans Affairs (VA) Medical Centers across the United States. After quality
18 control, we identified 2,280 participants of European, 627 of African, and 258 of Hispanic
19 ancestry with IA and 592,927 controls without evidence of IA. Within FinnGen, we identified
20 1,519 participants with IA and 284,164 controls.

21

22 *Intracranial Aneurysm Phenotype Definitions*

23 From the participants passing quality control in MVP, individuals were defined as having IA
24 based on possessing at least one of the following ICD codes for unruptured intracranial
25 aneurysms: ICD9 473.3 and/or ICD10 I67.1 in their electronic health record (EHR). Patients with
26 a diagnosis code for Marfan, Ehlers Danlos, ADPKD, and traumatic subarachnoid hemorrhage

1 were excluded (Supplemental Table 1). After exclusion, controls (individuals without IA) were
2 defined as participants without an IA diagnosis code.

3

4 *PheWAS of IA Risk Variants*

5 Multiple databases with publicly available summary statistics were used to assess phenotype
6 associations with IA lead variants. The NHIGRI-EBI GWAS catalog is a publicly available online
7 database comprising 60,071 summary statistics from a wide array of phenotypes.¹⁶
8 PhenoScanner v2 is a similar dataset comprising 65 billion associations of 150 million
9 variants.^{17,18} Finally, the Cardiovascular Disease Knowledge Portal (CDKP)
10 (<https://cvd.hugeamp.org/>) is a database part of the Common Metabolic Diseases Knowledge
11 Portal (CMDKP) project, and contains 207 GWAS summary statistics for 157 anthropometric
12 and cardiovascular traits and conditions. IA lead risk variants, defined as those with the lowest
13 p-value in a megabase window identified in our GWAS analysis, were queried in each of these
14 three databases separately to perform the PheWAS (Phenome-wide Association Study).
15 Phenotypic associations related to lab values, family history, medication use, imaging variables,
16 or self-reported diagnoses were excluded. After exclusion, each trait was re-coded into a broad
17 phenotypic category.

18

19 *Causal IA Gene and Variant Identification*

20 We prioritized candidate causal genes at each of the IA risk loci by collating evidence from the
21 following: 1) Literature review: prior genetic, clinical, or functional studies of genes at the locus
22 were reviewed, 2) Proximity: the closest gene to the lead risk variant, 3) eQTL query: cis-eQTLs
23 from the Genotype-Tissue Expression Project (GTEx)¹⁹ aortic and tibial tissue datasets with
24 significant association $P < 5 \times 10^{-6}$, 4) TWAS: results from FUSION²⁰, a fine-mapping technique to
25 identify causal genes in a transcriptome-wide association study (TWAS) using bulk RNA-seq
26 data from post-mortem aortic and tibial tissue and IA meta-analysis summary statistics, 5) eQTL

1 colocalization analysis: colocalizing eQTLs were determined using IA meta-analysis summary
2 statistics and annotated GTEx eQTLs in tibial and aortic tissue (echolocatoR²¹), 6) genes with
3 protein-altering variants in high linkage disequilibrium ($R^2 > 0.8$) with the lead IA risk variant
4 (LDLinkR)²², 7) Multi-trait colocalization: prioritization of colocalizing variants among at least 1
5 set of cardiovascular summary statistics (Hyprcoloc).²³ Genes prioritized as causal (beyond
6 *SOX17* where prior literature has established this as the likely causal gene) were identified
7 based on having: 1) two prioritization strategies highlighting it as a likely causal gene *and* 2)
8 plausible biological evidence for a role in IA pathogenesis based on prior genetic, clinical, or
9 functional studies.

10

11

12 *LD Score Regression Analysis*

13 Genetic correlation of cardiovascular traits was assayed using LDSC function in the
14 *GenomicSEM* R package. We obtained and used summary statistics for CAD²⁴, PAD¹⁵,
15 SBP/DBP (<http://www.nealelab.is/uk-biobank/>), AAA²⁵, TAA²⁶, along with descending and
16 ascending aortic diameters.²⁷ We approximated the LD structure from Europeans within the
17 1000 Genomes reference panel²⁸, and set a Bonferroni corrected $P < 0.0014$ (0.05/36 pairwise
18 comparisons) for statistical significance.

19

20 *Construction and Annotation of Human Cerebrovascular snRNA-seq Data*

21 Using publicly available data,^{29,30} we identified cerebrovascular snRNA-seq data from 21 healthy
22 control patients without any clinical or histopathological evidence of neurological or
23 cerebrovascular disease. We also included patients from the Religious Orders Study and
24 Memory and Aging Project (ROSMAP) without any neurologic pathology as used in Garcia et
25 al.³⁰ Data from each sample were processed in R Studio with *Seurat* (4.3.0). After conversion to
26 a single cell experiment (SCE), doublets and ambient RNA were removed. Cells within each

1 sample were filtered based on the following parameters: $200 < n\text{FeatureRNA} < 5000$; $200 <$
2 $n\text{Count_RNA} < 20,000$; and $\text{percent.mt} < 0.5\%$. Normalization was performed using
3 SCTransform v2. Each processed sample was then carried forward for integration. The
4 integrated data set was normalized and an integrated analysis was applied to all ~140,000 cells.
5 Level 1 annotation was performed by transferring cell type labels with a high prediction value
6 from the Tabula Sapiens³¹ vasculature single-cell reference set and using canonical markers for
7 vascular and cerebral cell types as described in Garcia et al. and Yang et al. After subsetting
8 cerebrovascular cell types including, endothelial cells, mural cells, fibroblasts, and astrocytes,
9 we performed Level 2 annotation of these cell types, again using markers described in Garcia et
10 al. and Yang et al. Differential expression of candidate causal genes was assayed using the
11 ‘SCT’ assay.

12

13 *Cell Type Prioritization of IA Using Integrated snRNA-seq Data and IA Summary Statistics*

14 A customized gene set using Stage 2 summary statistics effect sizes was created using
15 MAGMA. The integrated cerebrovascular snRNA-seq dataset with level 2 annotations was
16 converted to Anndata format for use with scanPy in the scDRS pipeline
17 (<https://github.com/martinjzhang/scDRS>).³² We performed individual cell-level analysis to
18 identify populations of cells with enrichment of putative candidate causal genes. Next, cell type-
19 level analysis was performed to identify clusters within the level 2 annotation that are
20 significantly associated with IA. Finally, gene-level analysis was performed to identify genes and
21 gene programs correlated with the scDRS.

22

23 *Generation of an IA Polygenic Risk Score*

24 A weighted polygenic risk score (PRS) was constructed using PRSCs, which generates posterior
25 genetic variant effect sizes throughout the genome using continuous shrinkage priors. Using
26 effect sizes (beta) from Bakker et al. Stage 2 summary statistics, we created scores for 756,520

1 variants, estimated LD structure using the 1000 genomes European reference panel. We
2 applied this score to MVP participants from European, African, and Hispanic ancestries. Each
3 individual was additively scored based on copy number of scored risk alleles using Plink -score
4 flag.

5

6 *Statistical Analysis*

7 In our primary discovery analysis, genotyped and imputed DNA sequence variants in
8 individuals of European, African, and Hispanic ancestry were tested for association with IA using
9 logistic mixed models as performed in the REGENIE v2.0 statistical software program.³³ We
10 included in step 1 of REGENIE (i.e. prediction of individual trait values based on the genetic
11 data) variants that were directly genotyped, had a minor allele frequency >1%, <10%
12 missingness, and with Hardy-Weinberg equilibrium test P-value >10⁻¹⁵. The association model
13 used in step 2 of REGENIE included as covariates age, sex, and 5 principal components of
14 ancestry. Next, associated statistics across of European, African, and Hispanic ancestry MVP
15 participants were meta-analyzed along with summary statistics available from FinnGen using an
16 inverse variance-weighted fixed effects method as implemented in the METAL software program
17 (Stage 1).³⁴

18 We next performed meta-analysis with IA summary statistics from Bakker et al (Stage 2).
19 Within the final meta-analysis, we defined IA susceptibility loci as those whose lead variant had
20 nominal significance ($p < 0.05$) in both Stage 1 and Bakker et al. summary statistics, were
21 directionally consistent across cohorts, and had an $P < 5 \times 10^{-8}$ in the meta-analysis. Novel loci
22 were annotated if greater than 500 kilobases from a previously described IA susceptibility locus.
23 All logistic regression values of P were two-sided.

24 In the PheWAS analysis, IA lead risk variants were queried using PhenoScanner v2,
25 GWAS catalog, and cardiovascular disease knowledge portal, which are all online compilations

1 of summary statistics from prior GWAS. We used a p-value threshold of $P < 5 \times 10^{-8}$ to determine
2 significance (i.e. only phenotypes with variants at genome-wide significance were considered).

3 Genetic correlation of IA with cardiovascular traits was performed using LD score
4 regression. Associations were considered significant if a Bonferroni corrected two-sided p-value
5 was less than 0.0014 (.05/36 pairwise comparisons). Multi-trait colocalization was performed
6 with Hyprcoloc, a deterministic Bayesian divisive clustering algorithms that can assess multiple
7 traits simultaneously. Colocalizing traits with a posterior probability > 0.7 were considered
8 significant. Similarly, significance of eQTL colocalization performed with echolocator, which
9 also uses a Bayesian approach based on priors, was declared for genes with posterior
10 probability > 0.7 .

11 In our snRNA-seq analysis, cell-type level analysis was performed with scDRS.
12 Benjamini-Hochberg corrected p-values were calculated and significant associations were
13 declared with P-values < 0.05 .

14 In our PRS analysis, logistic regression models were used to estimate odds ratios (OR)
15 and 95% confidence intervals for associations for prevalent IA adjusting for age, sex, diastolic
16 blood pressure, smoking status, and 5 principal components. Patients with diastolic blood
17 pressures outside a range of 50-109, corresponding to the bottom and top 0.5 percentiles, were
18 excluded. We calculated the OR of IA using the PRS as a continuous independent variable
19 (standard deviation unit) along with a high PRS (top 5th percentile of the continuous PRS).
20 Bonferroni corrected P-values < 0.017 (0.05/3 ancestries) were deemed significant.

21

22 RESULTS

23 *Multi-ancestry IA GWAS*

24 We performed a two-stage GWAS (Figure 1). In Stage 1, we performed a logistic mixed
25 model regression testing the association of TOPMED imputed genetic variants for each ancestry
26 in the MVP. We obtained 15.4 million (African), 10.1 million (Hispanic), 8.9 million (European)

1 common variants (MAF > 0.01) for analysis (Table 1). The resulting summary statistics were
2 then meta-analyzed, also including summary statistics from a Finnish cohort study (FinnGen)
3 comprising 13.0 million variants. The meta-analysis was comprised of 4,684 IA cases (2,280
4 European, 627 Black, 258 Hispanic, and 1,519 Finnish cases) along with 877,091 controls. We
5 identified four loci at genome-wide significance, all of which have been previously reported
6 (Supplemental Table 2).⁵ There was no evidence of genomic inflation in any group
7 (Supplemental Figure 1). In addition, we replicated the 16/17 known loci with $P < 0.05$ in our
8 Stage 1 analysis (Supplemental Table 3). In Stage 2, we meta-analyzed the results of Stage 1
9 with summary statistics of the prior IA GWAS (Bakker et al.) consisting of 10,754 cases and
10 306,882 controls from European and East Asian ancestries to yield a total of 15,438 cases and
11 1.2 million controls.

12 We identified 1,409 common variants (MAF > 0.01) at 22 loci associated with IA at
13 genome wide significance ($P < 5e-8$), five of which were novel associations (Table 2). With the
14 exception of the known lead variant near *BET1L* on chromosome 11, lead variants at the
15 remaining 21 loci had at least nominal significance ($P < 0.05$) in both Stage 1 and Bakker et al.
16 summary statistics (Supplemental Table 3).⁵

17 Consistent with prior results, the lead variant in Stage 2 was rs1537373 (52.3%
18 frequency of the G allele; OR = 1.20; 95% CI= 1.19-1.21; $P = 8.7 \times 10^{-46}$) within *CDKN2B-AS1*
19 which encodes for the long non-coding RNA (lncRNA) *ANRIL*, associated with several
20 cardiovascular and neoplastic disease states. We identified novel associations near *NAV1*,
21 *FGF5*, *EXOC2*, *TWIST1*, and the lncRNA *DLEU1* (Supplemental Figure 2).

22

23 *Phenome-Wide Association Study*

24 Pleiotropy of susceptibility variants offers insight into the shared pathophysiology of both
25 similar and seemingly unrelated traits. To investigate the breadth of phenotypes associated with
26 intracranial aneurysms, we performed a Phenome-wide Association Study (PheWAS) of IA

1 susceptibility lead variants using publicly available summary statistics within Phenoscanner v2,
2 NHGRI-EBI GWAS catalog, and the cardiovascular disease knowledge portal (CVDKP). We
3 identified 316 unique associations of IA susceptibility loci over 12 broad phenotypic categories
4 (Supplemental Table 4). A representative heatmap of traits with multiple associations and
5 publicly available summary statistics is shown in Figure 2. As previously described, many of the
6 IA susceptibility variants were also significantly associated with systolic and diastolic blood
7 pressure. We also observed associations with coronary and peripheral arterial disease
8 (CAD/PAD) as well as thoracic and abdominal aortic aneurysm. The lead variants at several loci
9 including *MTMR3*, *STARD13*, *SOX17*, and *EXOC2*, were not significantly associated with any
10 other cardiovascular traits, suggesting a unique association with IA. Surprisingly, at 7 loci
11 including *TWIST1* and *EDNRA*, the IA risk variant was either associated with *decreased*
12 SBP/DBP and or was *protective* for the development of CAD and PAD.

13

14 *Genetic Correlation of IA with Cardiovascular Traits*

15 Given the extensive pleiotropy of IA risk loci with cardiovascular traits, we sought to
16 assess the genetic correlation of IA with atherosclerotic, blood pressure, and aortic aneurysmal
17 traits. We assessed the pair-wise genetic correlation of 9 cardiovascular traits using LD score
18 regression (Figure 3a, Supplemental Table 5). Unsupervised clustering of traits based on
19 genetic correlation coefficients (R_g) revealed that IA shared heritability with aortic aneurysms,
20 most significantly with AAA ($R_g = 0.34$, $p = 2.9 \times 10^{-18}$). In addition, similar to AAA, IA had a
21 significant and positive genetic correlation with CAD ($R_g = 0.17$, $p = 5.5 \times 10^{-6}$).

22 While PheWAS can help identify shared phenotypic associations with lead variants and
23 suggest common underlying biological pathways, it fails to assess whether shared genetic
24 signals are likely driven by a common, causal variant. Therefore, we used multi-trait
25 colocalization (Hyprcoloc) to determine the contribution of IA susceptibility loci to the observed
26 positive correlation with atherosclerotic, blood pressure, and aneurysmal traits as seen by LD

1 score regression (Figure 3b).²³ We identified significant colocalization of *FGF5*, *TWIST1*, and
2 *ATP2B1* with at least one trait within every category (posterior probability > 0.7). Notably, among
3 the 9 IA susceptibility loci significantly colocalizing with at least one atherosclerotic trait, 6
4 prioritized variants had discordant effect directions with CAD. Similarly, 4 of the 6 IA
5 susceptibility loci colocalizing with SBP or DBP are associated with reduced blood pressure.
6 Thus, despite an overarching positive genetic correlation with CAD and SBP/DBP, it appears
7 that certain prominent pathways that cause IA are paradoxically protective for hypertension and
8 atherosclerosis.

9

10 *Prioritization of Candidate Genes*

11 Prior human genetic and functional evidence strongly suggests that *SOX17* is a causal
12 gene at one of our identified loci.^{35,36} For genes at the remaining loci, we examined the literature
13 for genetic or functional evidence of vascular involvement and identified non-synonymous
14 coding variants in high LD with the lead variant ($R^2 > 0.8$). In addition, we hypothesized that IA
15 risk variants may be acting by inducing expression changes in *cis* and performed fine-mapping
16 TWAS and colocalization analyses using aortic and tibial expression quantitative trait locus
17 (eQTL) data from the Genotype Tissue Expression Project (GTEx) (Supplemental Table 6,
18 Supplemental Figure 3). We supplemented this with multi-trait colocalization. When combining
19 the above strategies, we identified 17 additional putative causal genes - *FGF5* (4q21), *EDNRA*
20 (4q31), *P4HA2/SLC22A5* (5q31), *UFL1* (6q16), *TWIST1* (7p21), *NOC3L/PLCE1* (10q23),
21 *NT5C2* (10q24), *BET1L* (11p15), *ATP2B1* (12q21), *FGD6* (12q22), *STARD13* (13q13), *CHRNA5*
22 (15q25), *BCAR1* (16q23), *SLC24A3* (20p11), and *MTMR3* (22q12) (Supplemental Table 7).

23 Interestingly, we found that increased *TWIST1* expression is associated with a
24 decreased risk of IA. *TWIST1* encodes a master transcription factor involved in the regulation of
25 smooth muscle cell plasticity during atherogenesis.³⁷ Increased *TWIST1* has been shown to
26 promote a synthetic and proliferative phenotype in smooth muscle cells. Using multi-trait

1 colocalization, we prioritized rs2107595 as the candidate causal variant (Supplemental Figure
2 4). This variant lies in a DNase hypersensitivity site, and directly modulates *TWIST1*
3 expression.³⁷ Taken together, these results suggest that increased *TWIST1* expression
4 promotes SMC proliferation and extracellular matrix production, reducing IA risk potentially
5 through stabilization of the vascular wall.

6

7 *Gene expression analysis facilitates identification of candidate causal cell types*

8 Our candidate causal gene set and PheWAS observations highlight directional
9 discordance of many IA susceptibility loci with blood pressure and atherosclerosis. Thus, rather
10 than systemic blood pressure and smoking status as sole mediators of IA pathogenesis, we
11 hypothesized that intrinsic aberrancies within the cerebrovasculature may directly contribute to
12 IA pathogenesis. To this end we analyzed expression of IA susceptibility loci using human
13 cerebrovasculature single-nuclear RNA sequencing (snRNA-seq) data.

14 Using publicly available human brain snRNA-seq datasets enriched for
15 cerebrovasculature we identified 20 patients without any neurological or vascular pathology.^{29,30}
16 We performed quality control and integrated snRNA-seq data from these patients with data from
17 ROSMAP (see methods) to obtain over 147,439 nuclei and 31 Louvain clusters. After
18 unsupervised clustering, we performed a Level 1 annotation to identify fibroblast, pericyte,
19 endothelial, and smooth muscle cells as well as cerebral cell types (Supplemental Figure 4).

20 After subsetting relevant cerebrovascular cell types (endothelial cells, smooth muscle
21 cells, pericytes, astrocytes, and meningeal/perivascular fibroblasts) we obtained 102,140 nuclei
22 for analysis. We performed Level 2 annotation, using previously described markers to annotate
23 endothelial, pericyte, fibroblast, and astrocyte subtypes (Figure 4a). Analysis of subpopulations
24 of EC revealed enrichment of capillary ECs relative to arterial and venous ECs (Supplemental
25 Figure 6). Our integrated analysis allowed identification of a relatively sparse population of ECs
26 expressing Laminin alpha-2 and Laminin beta-1, important constituents of the basement

1 membrane and blood brain barrier (Supplemental Figure 7). Concordant with the original
2 datasets, these cells are annotated as tip ECs. Mural cells demonstrated the greatest amount of
3 heterogeneity with four subclusters described in the cerebrovasculature. Amongst the pericytes,
4 we annotated roughly equal proportions of matrix producing pericytes (M-pericyte) and pericytes
5 involved in membrane transport (T-pericyte) (Supplemental Figure 6). M-pericytes display
6 enrichment of type 4 collagen, while T-pericytes are enriched in several solute carriers
7 (Supplemental Figure 7). Finally, the majority of fibroblasts displayed a perivascular phenotype
8 as evinced by expression of structural components, while meningeal fibroblasts expressed
9 several SLC solute transporters (Supplemental Figure 6c).

10 Having defined the major and minor cell populations in the integrated snRNA-seq
11 cerebrovascular dataset, we next sought to further prioritize candidate causal cell-types and
12 gene programs through linking our GWAS and snRNA-seq datasets at single cell resolution. We
13 integrated gene expression profiles with a gene set derived from IA summary statistics
14 (MAGMA) to associate single cells with IA through computation of a single cell disease
15 relevance score (scDRS).³² We began with a cell-type level analysis to identify cell type
16 associations and heterogeneity within each cell type. SMCs, aaSMC, M-pericytes, and tip ECs
17 were all significantly associated with IA (Figure 4b, Supplemental Table 8). With the exception of
18 tip ECs, each cell type displayed considerable and significant heterogeneity suggesting
19 subpopulations that contribute to the observed association. Single cell analysis of disease
20 relevance score revealed a sub-population of M-pericytes and SMCs with the strongest signal
21 (Supplemental Figure 7).

22 Among cell types significantly associated with the IA (mural cell populations and tip
23 ECs), we assayed the correlation of each gene's expression with the scDRS. We consistently
24 found that *EDNRA* and *FHL5*, whose dysregulation was recently found to drive vascular
25 calcification in atherosclerosis, were highly correlated with the scDRS. Both genes have high z-
26 scores within the IA gene set and likely drive the observed cell-type association. Interestingly,

1 within the M-pericyte, tip EC, and aaSMC populations we also found basement membrane
2 components *COL4A1/COL4A2* as one of the most strongly correlated genes despite the lack of
3 significant association signal within our GWAS summary statistics (Figure 4c-f, Supplemental
4 Figure 8).

5 Finally, we sought to identify relevant cell types for the individual putative causal IA risk
6 genes examining the cell-type enrichment of each candidate causal gene. Within ECs, we found
7 enrichment of *BCAR1*, which encodes a scaffold protein involved in mechanotransduction.^{38–40}
8 Within the mural cell types, we found enrichment of *EDNRA* and *PLCE1*, with relative
9 enrichment of these genes in the M-pericyte population. Amongst the fibroblasts, *TWIST1* and
10 *FGD6* demonstrated enrichment within both fibroblast populations, while *SLC24A3* expression
11 was clearly enriched within the meningeal fibroblast population (Figure 5). Together, this
12 suggests that although SMCs and pericytes appear to be the primary drivers of IA,
13 dysregulation of ECs and fibroblasts may also contribute to IA development.

14

15 *Generation of an IA polygenic risk score*

16 Finally, we sought to evaluate the contribution of polygenic inheritance on IA risk. Using
17 the Bakker et al. summary statistics comprising over 10,000 cases, we generated a PRS using
18 756,520 variants (PRSs⁴¹, see methods). We assessed the performance of the PRS using
19 patients within the MVP from European, African, and Hispanic ancestries separately. We
20 identified patients with a high PRS (top 5th percentile) who we hypothesized were most
21 susceptible to IA development. Amongst the European cohort, patients in the highest ventile had
22 an 87% increased risk of prevalent diagnosis (OR: 1.87, CI: 1.61-2.17, $P = 8.82 \times 10^{-17}$) while
23 controlling for age, diastolic blood pressure, female sex, and smoking status in the regression
24 model (Figure 6a). Similarly, patients with a high PRS with African and Hispanic ancestries show
25 an increased risk of IA diagnosis (African: OR 1.62, CI 1.19-2.15, $P = 1.18 \times 10^{-3}$; Hispanic: OR
26 2.28, CI 1.47-3.38, $P = 1.02 \times 10^{-4}$) (Figure 6b-c). We found a similar effect while examining a 1-

1 SD increase in PRS across ancestries. Finally, we stratified patients from each ancestry by
2 ventile and observed an increasing prevalence and risk of IA, controlling for the aforementioned
3 comorbidities and demographic parameters (Supplemental Figure 9).

4

5 DISCUSSION

6 In this study, we performed a GWAS to identify 5 novel IA susceptibility loci, raising the
7 total number of replicated IA loci to 22. As seen in our PheWAS, the vast majority of loci
8 overlapped with susceptibility loci of several complex cardiovascular traits, namely
9 atherosclerosis, blood pressure, and aneurysms. Consistently, we found a positive genetic
10 correlation of IA with many cardiovascular traits and identified the causal genes contributing to
11 the observed correlation by colocalizing signal from multiple sets of summary statistics. We
12 identified causal cerebrovascular cell types by constructing a large snRNA-seq dataset from
13 cerebrovascular tissue and integrating gene expression with IA summary statistics. Finally, we
14 developed a polygenic risk score for IA and performed validation using the MVP.

15 These findings permit several conclusions. First, we show through several lines of
16 evidence that IA has substantial shared heritability with other cardiovascular traits and
17 hypothesize that this information may be useful for therapeutic intervention. Assessment of
18 genetic correlations with LD score regression supports significant overlap with atherosclerosis,
19 hypertension, and other aneurysmal traits. Of all cardiovascular traits, IA demonstrated the
20 strongest genetic correlation with AAA, consistent with epidemiological and observational
21 studies.^{8,9} Furthermore, we applied fine-mapping with multi-trait colocalization to reveal the loci
22 largely responsible for the observed correlation. As GWAS become increasingly powered and
23 the number of significant associations with cardiovascular pathologies grows, assessing
24 commonalities between phenotypes will be important to postulate shared mechanisms and
25 therapeutic targets. Here, we find that the overlapping genetic correlation with atherosclerosis,
26 aortic aneurysm, and hypertension is largely driven by six loci, 9p21 (atherosclerosis, aortic

1 aneurysm), *NT5C2* (hypertension and atherosclerosis), *NAV1* (aortic aneurysm), *CHRNA5*
2 (aortic aneurysm, smoking exposure⁴²), *PLCE1* (hypertension), and *FGF5* (aortic aneurysm,
3 atherosclerosis, hypertension). While further mechanistic validation of these loci remains to be
4 performed, our data suggests the presence of a subset of shared, targetable pathways that may
5 treat hypertension, aneurysmal disease, and atherosclerosis. In particular, the causal gene at
6 the *FGF5* locus (prioritized here to be *FGF5*) may be targeted to treat all three categories of
7 cardiovascular pathology.

8 Second, despite an overall positive genetic correlation of IA with atherosclerosis and
9 blood pressure traits, the presence of opposing risk alleles at multiple, shared genome-wide
10 significant loci highlights the importance of vessel wall integrity in IA pathogenesis. In our
11 PheWAS, the effect directions of the IA risk allele are protective for hypertension or CAD at
12 several loci, including *EDNRA*. Increased expression of *EDNRA* is associated with
13 atherosclerosis and hypertension through promotion of systemic vasoconstriction. A paradoxical
14 *increased* risk of IA conferred by *EDNRA* variants associated with *decreased* blood pressure
15 suggests an alternative mechanism by which this variant causes IA. *EDNRA* encodes a G-
16 protein coupled receptor, which after binding its primary ligand endothelin-1, activates a plethora
17 of signaling cascades involved in matrix deposition, contraction, and proliferation.⁴³ The
18 *EDN1/EDNRA* axis has both a critical role in neural crest specification, which ultimately gives
19 rise to the craniofacial vasculature, and promotes extracellular matrix contraction which is likely
20 to be critical in vessel wall integrity.^{44,45} Similarly, variants at the *UFL1-FHL5* locus, also with
21 discordant effect direction in the contexts of IA and CAD, have recently been shown to promote
22 phenotypic switching of smooth muscle cells to a synthetic phenotype in the context of
23 atherosclerosis.^{37,46} Consistent with the concept that differences in matrix production can
24 differentially drive atherosclerosis and IA, we find through integration of snRNA-seq and GWAS
25 summary statistics that tip ECs and M-pericytes, both involved in basement membrane
26 production through expression of type IV collagen and laminins, are significantly associated with

1 IA. Moreover, amongst these cell types, *COL4A1/COL4A2* expression, which encode subunits
2 of type IV collagen, the primary component of the basement membrane, is strongly correlated
3 with the weighted expression of an IA gene set. Together these data suggest that shared causal
4 pathways seen in IA and other cardiovascular traits may result in divergent phenotypic
5 consequences. Therefore, despite considerable genetic overlap with atherosclerosis and
6 hypertension, we hypothesize that *targeted* IA therapeutics augmenting extracellular matrix
7 production and vascular wall stability will provide the most utility for IA treatment or prevention.

8 Third, our results suggest that focusing on SMC biology is likely to be particularly key for
9 understanding IA pathologic mechanisms and identifying novel therapies. Through our genome-
10 wide and gene-based transcriptomic analyses, SMCs and M-pericytes consistently emerged as
11 likely causal cell types. Often referred to as ‘phenotype switching’, SMCs are capable of
12 remarkable plasticity which typically involves downregulation of SMC markers and transitioning
13 from a contractile phenotype to a proliferative, synthetic, or migratory phenotype.⁴⁷ In the
14 context of atherosclerosis, smooth muscle cells are thought to migrate to the intima and trans-
15 differentiate to a fibrochondrocyte-like cell to secrete a collagen-rich ECM of the fibrous cap. In
16 this manner, modulated SMCs can contribute to plaque burden. By contrast, in our study we
17 annotated mediators of phenotype switching, including *TWIST1*, *EDNRA*, and *UFL1*, as likely to
18 reduce risk of IA.^{37,48–50} These genes likely promote vascular wall stability through ECM
19 generation and smooth muscle cell proliferation. Consistent with this mechanism, inherited risk
20 at *9p21* may cause SMC apoptosis and medial thinning through dysregulation of *CDKN2B*.⁵¹
21 Therefore, in the context of aneurysmal disease, harnessing the plasticity of SMCs to promote
22 matrix production may restore vascular wall integrity and decrease the risk of aneurysm rupture.
23 We also caution that emerging endeavors to generate therapeutics inhibiting SMC phenotype
24 switching for atherosclerosis also consider potential deleterious effects of matrix destabilization
25 in the context of concomitant aneurysmal degeneration.

1 Our findings should be interpreted within the context of the following limitations. First, our
2 IA phenotype is based on EHR diagnosis code data, and may result in misclassification of case
3 status. However, such misclassification should, on average, reduce statistical power for
4 discovery and bias results toward the null. Second, with available ICD codes we were unable to
5 ascertain patients who presented with intracranial hemorrhage from an IA, which limits our
6 ability to identify risk loci for rupture. Third, the Veterans Affairs Healthcare System population is
7 overwhelmingly male and our ability to detect sex-specific genetic associations in discovery was
8 limited.⁴⁵ Finally, although MVP participants with the highest PRS are at increased risk for IA, the
9 PRS mechanism of action represents a combination of many causal risk factors, rather than a
10 single pathway that leads to disease. Assessment of individual risk can aid in identifying those
11 at highest risk for IA and more likely to obtain benefit from screening, regardless of mechanism.

12 In conclusion, we have performed the largest IA GWAS study to date and discovered 5
13 novel loci. We identify causal risk factors, cell types, and genes that may be used to inform
14 clinical care.

15
16
17
18
19
20
21
22
23
24
25
26

1 References

2

- 3 1. Hop JW, Rinkel GJE, Algra A, Gijn J van. Case-Fatality Rates and Functional Outcome After
4 Subarachnoid Hemorrhage: A Systematic Review. *Stroke* 1997;28:660–664. doi:10.1161/01.str.28.3.660.
- 5 2. Brown RD, Broderick JP. Unruptured intracranial aneurysms: epidemiology, natural history,
6 management options, and familial screening. *Lancet Neurology* 2014;13:393–404. doi:10.1016/s1474-
7 4422(14)70015-8.
- 8 3. Hitchcock E, Gibson WT. A Review of the Genetics of Intracranial Berry Aneurysms and Implications
9 for Genetic Counseling. *J Genet Couns* 2017;26:21–31. doi:10.1007/s10897-016-0029-8.
- 10 4. Shabani M, Abdollahi A, Brar BK, MacCarrick GL, Venkatesh BA, Lima JAC, et al. Vascular
11 aneurysms in Ehlers-Danlos syndrome subtypes: A systematic review. *Clin Genet* 2023;103:261–267.
12 doi:10.1111/cge.14245.
- 13 5. Bakker MK, Spek RAA van der, Rheenen W van, Morel S, Bourcier R, Hostettler IC, et al. Genome-
14 wide association study of intracranial aneurysms identifies 17 risk loci and genetic overlap with clinical
15 risk factors. *Nat Genet* 2020;52:1303–1313. doi:10.1038/s41588-020-00725-7.
- 16 6. Karhunen V, Bakker MK, Ruigrok YM, Gill D, Larsson SC. Modifiable Risk Factors for Intracranial
17 Aneurysm and Aneurysmal Subarachnoid Hemorrhage: A Mendelian Randomization Study. *J Am Hear*
18 *Assoc* 2021;10:e022277. doi:10.1161/jaha.121.022277.
- 19 7. Song J, Lim YC, Ko I, Kim J-Y, Kim D-K. Prevalence of Intracranial Aneurysms in Patients With
20 Systemic Vessel Aneurysms. *Stroke* 2020;51:115–120. doi:10.1161/strokeaha.119.027285.
- 21 8. Lee D, Ahn SJ, Cho E-S, Kim YB, Song S-W, Jung WS, et al. High prevalence of intracranial
22 aneurysms in patients with aortic dissection or aneurysm: feasibility of extended aorta CT angiography
23 with involvement of intracranial arteries. *J NeuroInterventional Surg* 2017;9:1017.
24 doi:10.1136/neurintsurg-2016-012619.
- 25 9. Kurtelius A, Väntti N, Jahromi BR, Tähtinen O, Manninen H, Koskenvuo J, et al. Association of
26 Intracranial Aneurysms With Aortic Aneurysms in 125 Patients With Fusiform and 4253 Patients With
27 Saccular Intracranial Aneurysms and Their Family Members and Population Controls. *J Am Hear Assoc*
28 2019;8:e013277. doi:10.1161/jaha.119.013277.
- 29 10. Rouchaud A, Brandt MD, Rydberg AM, Kadirvel R, Flemming K, Kallmes DF, et al. Prevalence of
30 Intracranial Aneurysms in Patients with Aortic Aneurysms. *Am J Neuroradiol* 2016;37:1664–1668.
31 doi:10.3174/ajnr.a4827.
- 32 11. Rantasalo V, Gunn J, Kiviniemi T, Hirvonen J, Saarenpää I, Kivelev J, et al. Intracranial aneurysm is
33 predicted by abdominal aortic calcification index: A retrospective case-control study. *Atherosclerosis*
34 2021;334:30–38. doi:10.1016/j.atherosclerosis.2021.08.027.

- 1 12. Punch E, Klein J, Diaba-Nuhoho P, Morawietz H, Garelnabi M. Effects of PCSK9 Targeting:
2 Alleviating Oxidation, Inflammation, and Atherosclerosis. *J Am Hear Assoc Cardiovasc Cerebrovasc Dis*
3 *n.d.*;11:e023328. doi:10.1161/jaha.121.023328.
- 4 13. Klarin D, Verma SS, Judy R, Dikilitas O, Wolford BN, Paranjpe I, et al. Genetic Architecture of
5 Abdominal Aortic Aneurysm in the Million Veteran Program. *Circulation* 2020;142:1633–1646.
6 doi:10.1161/circulationaha.120.047544.
- 7 14. Tcheandjieu C, Zhu X, Hilliard AT, Clarke SL, Napolioni V, Ma S, et al. Large-scale genome-wide
8 association study of coronary artery disease in genetically diverse populations. *Nat Med* 2022;28:1679–
9 1692. doi:10.1038/s41591-022-01891-3.
- 10 15. Klarin D, Lynch J, Aragam K, Chaffin M, Assimes TL, Huang J, et al. Genome-wide association
11 study of peripheral artery disease in the Million Veteran Program. *Nature Medicine* 2019;25:1–22.
12 doi:10.1038/s41591-019-0492-5.
- 13 16. Sollis E, Mosaku A, Abid A, Buniello A, Cerezo M, Gil L, et al. The NHGRI-EBI GWAS Catalog:
14 knowledgebase and deposition resource. *Nucleic Acids Res* 2022;51:D977–D985.
15 doi:10.1093/nar/gkac1010.
- 16 17. Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: an
17 expanded tool for searching human genotype-phenotype associations. *Bioinformatics* 2019;35:4851–
18 4853. doi:10.1093/bioinformatics/btz469.
- 19 18. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: a database of
20 human genotype-phenotype associations. *Bioinformatics* 2016;32:3207–3209.
21 doi:10.1093/bioinformatics/btw373.
- 22 19. Aguet F, Brown AA, Castel SE, Davis JR, He Y, Jo B, et al. Genetic effects on gene expression
23 across human tissues. *Nature* 2017;550:204–213. doi:10.1038/nature24277.
- 24 20. Gusev A, Ko A, Shi H, Bhatia G, Chung W, Penninx BWJH, et al. Integrative approaches for large-
25 scale transcriptome-wide association studies. *Nat Genet* 2016;48:245–252. doi:10.1038/ng.3506.
- 26 21. Schilder BM, Humphrey J, Raj T. echolocatoR : an automated end-to-end statistical and functional
27 genomic fine-mapping pipeline. *Bioinformatics* 2021;38:536–539. doi:10.1093/bioinformatics/btab658.
- 28 22. Myers TA, Chanock SJ, Machiela MJ. LDlinkR: An R Package for Rapidly Calculating Linkage
29 Disequilibrium Statistics in Diverse Populations. *Frontiers Genet* 2020;11:157.
30 doi:10.3389/fgene.2020.00157.
- 31 23. Foley CN, Staley JR, Breen PG, Sun BB, Kirk PDW, Burgess S, et al. A fast and efficient
32 colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat Commun*
33 2021;12:764. doi:10.1038/s41467-020-20885-8.
- 34 24. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Atri DS, et al. Discovery and systematic
35 characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat*
36 *Genet* 2022;54:1803–1815. doi:10.1038/s41588-022-01233-6.

- 1 25. Roychowdhury T, Klarin D, Levin MG, Spin JM, Rhee YH, Deng A, et al. Multi-ancestry GWAS
2 deciphers genetic architecture of abdominal aortic aneurysm and highlights PCSK9 as a therapeutic
3 target. *Medrxiv* 2022:2022.05.27.22275607. doi:10.1101/2022.05.27.22275607.
- 4 26. Klarin D, Devineni P, Sendamarai A, Angueira A, Graham S, Shen Y, et al. Genome-wide
5 Association Study of Thoracic Aortic Aneurysm and Dissection in the Million Veteran Program 2022.
6 doi:10.21203/rs.3.rs-1507463/v1.
- 7 27. Pirruccello JP, Chaffin MD, Chou EL, Fleming SJ, Lin H, Nekoui M, et al. Deep learning enables
8 genetic analysis of the human thoracic aorta. *Nat Genet* 2022;54:40–51. doi:10.1038/s41588-021-00962-
9 4.
- 10 28. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global
11 reference for human genetic variation. *Nature* 2015;526:68–74. doi:10.1038/nature15393.
- 12 29. Yang AC, Vest RT, Kern F, Lee DP, Agam M, Maat CA, et al. A human brain vascular atlas reveals
13 diverse mediators of Alzheimer’s risk. *Nature* 2022;603:885–892. doi:10.1038/s41586-021-04369-3.
- 14 30. Garcia FJ, Sun N, Lee H, Godlewski B, Mathys H, Galani K, et al. Single-cell dissection of the human
15 brain vasculature. *Nature* 2022;603:893–899. doi:10.1038/s41586-022-04521-7.
- 16 31. Consortium* TS, Jones RC, Karkanas J, Krasnow MA, Pisco AO, Quake SR, et al. The Tabula
17 Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science* 2022;376:eabl4896.
18 doi:10.1126/science.abl4896.
- 19 32. Zhang MJ, Hou K, Dey KK, Sakaue S, Jagadeesh KA, Weinand K, et al. Polygenic enrichment
20 distinguishes disease associations of individual cells in single-cell RNA-seq data. *Nat Genet*
21 2022;54:1572–1580. doi:10.1038/s41588-022-01167-z.
- 22 33. Mbatchou J, Barnard L, Backman J, Marcketta A, Kosmicki JA, Ziyatdinov A, et al. Computationally
23 efficient whole-genome regression for quantitative and binary traits. *Nat Genet* 2021;53:1097–1103.
24 doi:10.1038/s41588-021-00870-7.
- 25 34. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association
26 scans. *Bioinformatics* 2010;26:2190–2191. doi:10.1093/bioinformatics/btq340.
- 27 35. Laarman MD, Geeven G, Barnett P, Bank NB, Rinkel GJE, Laat W de, et al. Chromatin
28 Conformation Links Putative Enhancers in Intracranial Aneurysm–Associated Regions to Potential
29 Candidate Genes. *J Am Hear Assoc* 2019;8:e011201. doi:10.1161/jaha.118.011201.
- 30 36. Lee S, Kim I-K, Ahn JS, Woo D-C, Kim S-T, Song S, et al. Deficiency of Endothelium-Specific
31 Transcription Factor Sox17 Induces Intracranial Aneurysm. *Circulation* 2015;131:995–1005.
32 doi:10.1161/circulationaha.114.012568.
- 33 37. Nurnberg ST, Guerraty MA, Wirka RC, Rao HS, Pjanic M, Norton S, et al. Genomic profiling of
34 human vascular cells identifies TWIST1 as a causal gene for common vascular diseases. *Plos Genet*
35 2020;16:e1008538. doi:10.1371/journal.pgen.1008538.

- 1 38. Dreier B, Gasiorowski JZ, Morgan JT, Nealey PF, Russell P, Murphy CJ. Early responses of vascular
2 endothelial cells to topographic cues. *Am J Physiol-Cell Physiol* 2013;305:C290–C298.
3 doi:10.1152/ajpcell.00264.2012.
- 4 39. Barrett A, Pellet-Many C, Zachary IC, Evans IM, Frankel P. p130Cas: A key signalling node in health
5 and disease. *Cell Signal* 2013;25:766–777. doi:10.1016/j.cellsig.2012.12.019.
- 6 40. Leal M del PC, Sciortino M, Tornillo G, Colombo S, Defilippi P, Cabodi S. p130Cas/BCAR1 scaffold
7 protein in tissue homeostasis and pathogenesis. *Gene* 2015;562:1–7. doi:10.1016/j.gene.2015.02.027.
- 8 41. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and
9 continuous shrinkage priors. *Nat Commun* 2019;10:1776. doi:10.1038/s41467-019-09718-5.
- 10 42. Buchwald J, Chenoweth MJ, Palviainen T, Zhu G, Benner C, Gordon S, et al. Genome-wide
11 association meta-analysis of nicotine metabolism and cigarette consumption measures in smokers of
12 European descent. *Mol Psychiatry* 2021;26:2212–2223. doi:10.1038/s41380-020-0702-z.
- 13 43. Dagamajalu S, Rex DAB, Gopalakrishnan L, Karthikkeyan G, Gurtoo S, Modi PK, et al. A network
14 map of endothelin mediated signaling pathway. *J Cell Commun Signal* 2021;15:277–282.
15 doi:10.1007/s12079-020-00581-4.
- 16 44. Square T, Jandzik D, Cattell M, Hansen A, Medeiros DM. Embryonic expression of endothelins and
17 their receptors in lamprey and frog reveals stem vertebrate origins of complex Endothelin signaling. *Sci*
18 *Reports* 2016;6:34282. doi:10.1038/srep34282.
- 19 45. Shi-Wen X, Chen Y, Denton CP, Eastwood M, Renzoni EA, Bou-Gharios G, et al. Endothelin-1
20 Promotes Myofibroblast Induction through the ETA Receptor via a rac/Phosphoinositide 3-Kinase/Akt-
21 dependent Pathway and Is Essential for the Enhanced Contractile Phenotype of Fibrotic Fibroblasts. *Mol*
22 *Biology Cell* 2004;15:2707–2719. doi:10.1091/mbc.e03-12-0902.
- 23 46. Wong D, Auguste G, Cardenas CLL, Turner AW, Chen Y, Song Y, et al. FHL5 Controls Vascular
24 Disease-Associated Gene Programs in Smooth Muscle Cells. *Circ Res* 2023.
25 doi:10.1161/circresaha.122.321692.
- 26 47. Gomez D, Owens GK. Smooth muscle cell phenotypic switching in atherosclerosis. *Cardiovasc Res*
27 2012;95:156–164. doi:10.1093/cvr/cvs115.
- 28 48. Ma L, Bryce NS, Turner AW, Narzo AFD, Rahman K, Xu Y, et al. The HDAC9-associated risk locus
29 promotes coronary artery disease by governing TWIST1. *Plos Genet* 2022;18:e1010261.
30 doi:10.1371/journal.pgen.1010261.
- 31 49. Su M, Yue Z, Wang H, Jia M, Bai C, Qiu W, et al. Ufmylation Is Activated in Vascular Remodeling
32 and Lipopolysaccharide-Induced Endothelial Cell Injury. *DNA Cell Biology* 2018;37:426–431.
33 doi:10.1089/dna.2017.4073.
- 34 50. Hu W, Wu X, Jin Z, Wang Z, Guo Q, Chen Z, et al. Andrographolide Promotes Interaction Between
35 Endothelin-Dependent EDNRA/EDNRB and Myocardin-SRF to Regulate Pathological Vascular
36 Remodeling. *Frontiers Cardiovasc Medicine* 2022;8:783872. doi:10.3389/fcvm.2021.783872.

1 51. Leeper NJ, Raiesdana A, Kojima Y, Kundu RK, Cheng H, Maegdefessel L, et al. Loss of CDKN2B
2 Promotes p53-Dependent Smooth Muscle Cell Apoptosis and Aneurysm Formation. *Arteriosclerosis*
3 *Thrombosis Vasc Biology* 2018;33:e1–e10. doi:10.1161/atvbaha.112.300399.

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

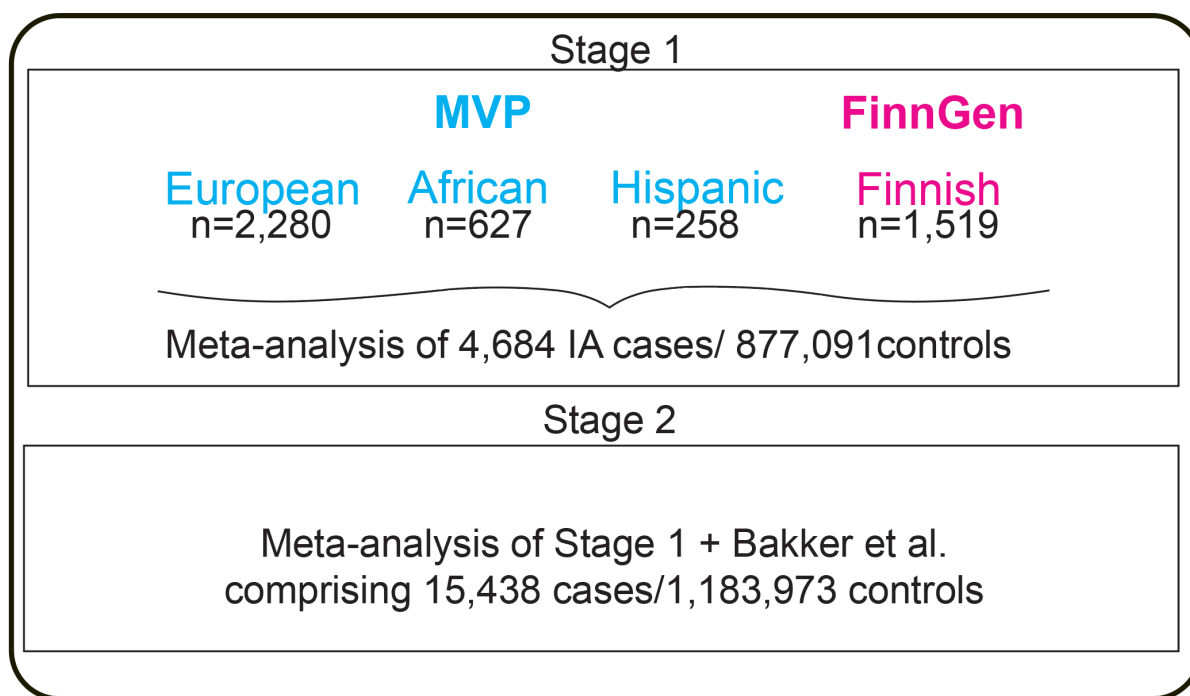
23

24

25

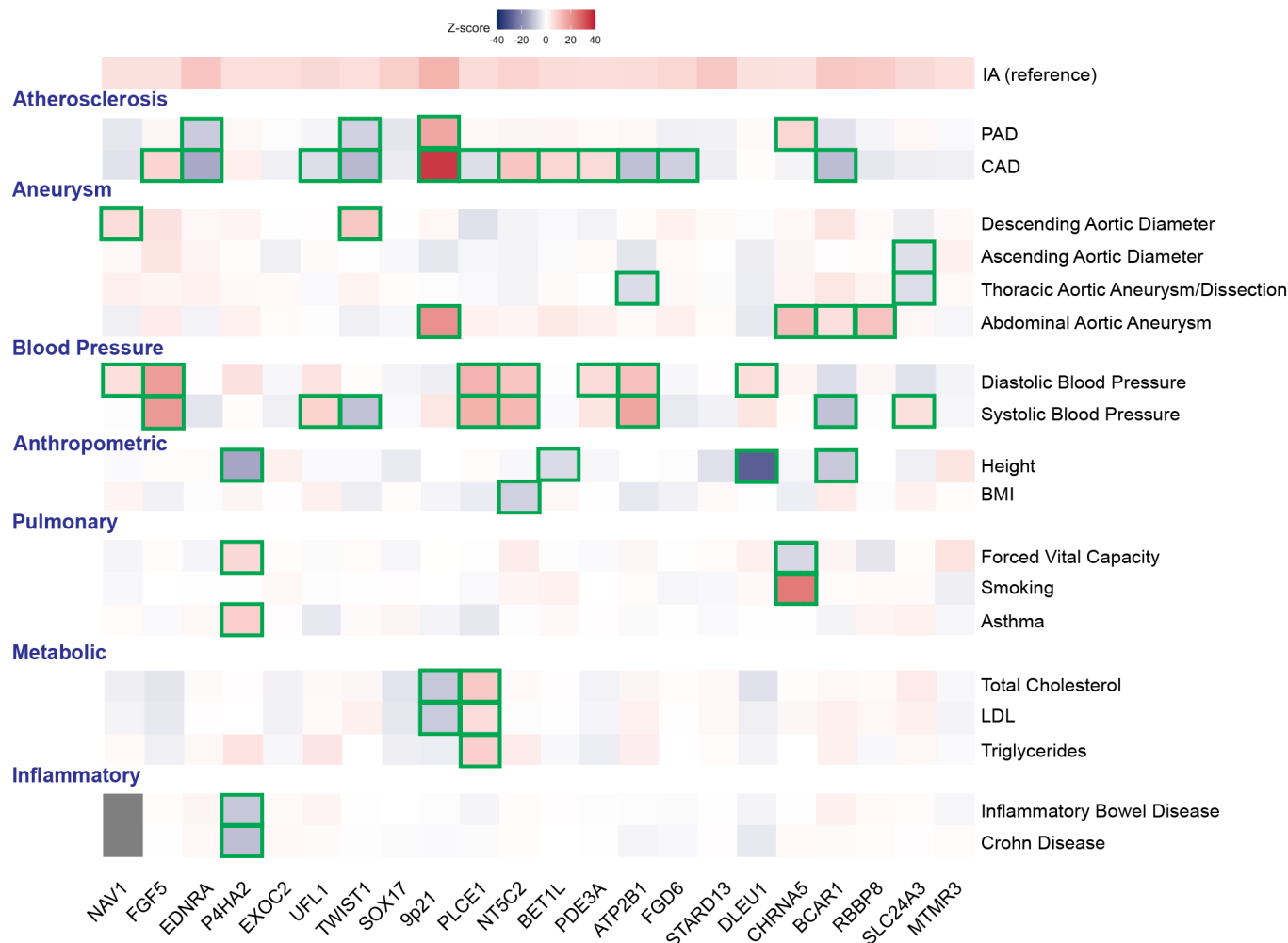
26

27



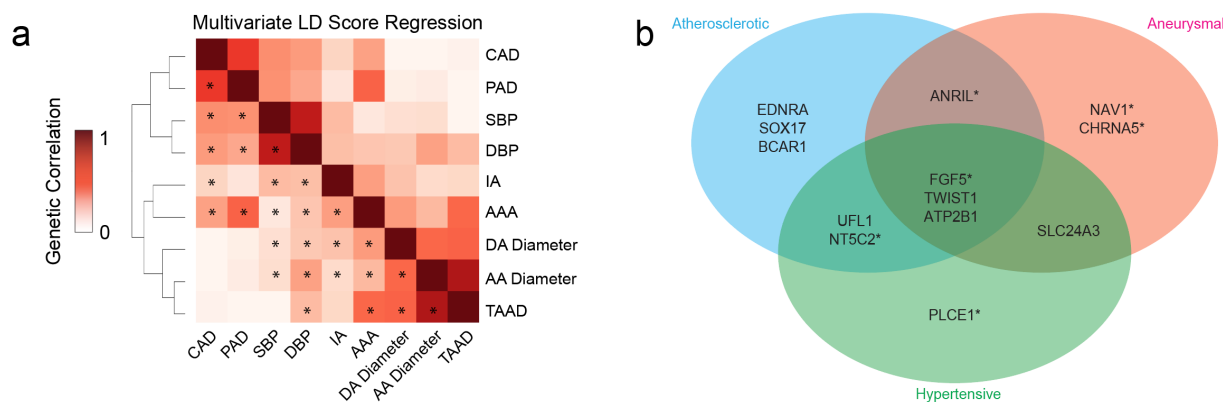
1
2 **Figure 1: Intracranial Aneurysm GWAS Study Design.** Stage 1: Meta-analysis of IA genome-
3 wide associations across European, African, and Hispanic ancestries from MVP along with a
4 Finnish cohort from FinnGen. Stage 2: Meta-analysis of Stage 1 with summary statistics from a
5 prior intracranial aneurysm GWAS, to comprise 15,438 IA cases.

6
7
8
9
10
11
12
13
14
15



1
2 **Figure 2: Representative Heatmap of Phenome-wide Association Results and Biologic**
3 **Pathways Underlying IA Susceptibility Loci.** Logistic regression Z-scores (aligned to the IA
4 risk allele) from summary statistics of traits identified in PheWAS results at each of the 22 IA
5 susceptibility loci. A positive Z-score (red) indicates a positive association between the PAD risk
6 allele and the disease, whereas a negative Z-score (blue) indicates an inverse association.
7 Boxes are outlined in green if the variant is associated with the indicated disease at genome-
8 wide significance (logistic regression two-sided $P < 5.0 \times 10^{-8}$). IA, intracranial aneurysm; PAD,
9 peripheral arterial disease; CAD, coronary artery disease; BMI, body mass index; LDL, low-
10 density lipoprotein.

1



2

3 **Figure 3: Genetic Correlation of Intracranial Aneurysm and other Cardiovascular Traits. a)**

4 Heatmap of pair-wise LD score regression across 9 cardiovascular traits. Dendrogram

5 represents results of unsupervised clustering of traits based on correlation coefficient (R_g). *

6 denotes $p < 0.001$. b) Venn-diagram illustrating colocalization of IA susceptibility loci with at least

7 one trait categorized as atherosclerotic (CAD/PAD), hypertensive (SBP/DBP), or aneurysmal

8 (AAA, DA size, AA size, or TAAD). * indicates significant colocalization (PP > 0.7) and

9 concordant z-score direction of the lead variant.

10

11

12

13

14

15

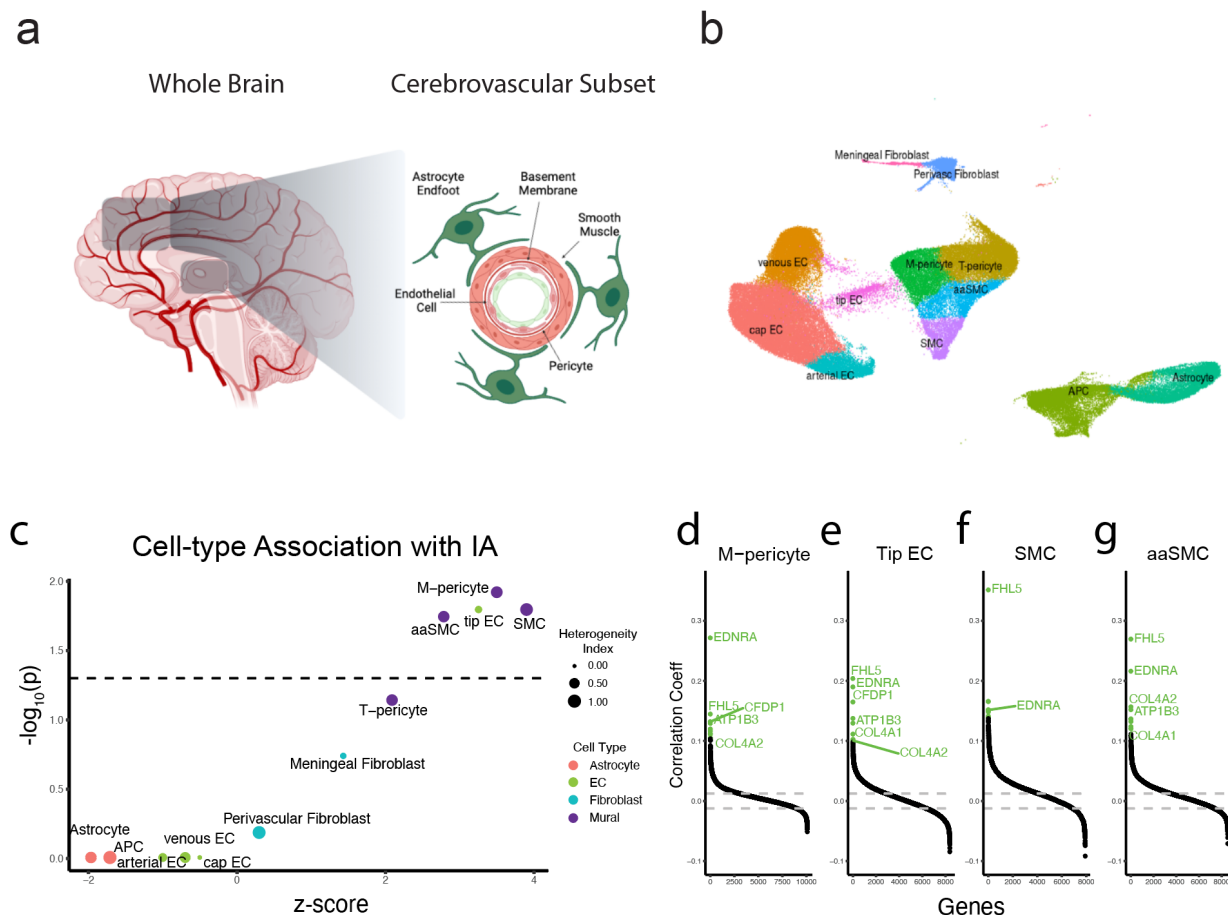
16

17

18

19

20



1
 2 **Figure 4: Generation of Human Cerebrovascular snRNA-seq Dataset and Integration with**
 3 **Intracranial Aneurysm Gene Set.** a) Schematic depicting integration of healthy human
 4 cerebrovascular samples, Level 1 annotation of all cell-types, sub-setting of cerebrovascular cell
 5 types, followed by Level 2 annotation of cerebrovascular cell types. b) Cell-type level analysis
 6 depicting IA association with each Level 2 annotated cell-type. Cell-types are color-coded by
 7 their Level 1 annotation. Heterogeneity index reflects the degree of heterogeneity of the single
 8 cell disease relevance score in each cell-type. High heterogeneity suggests the presence of a
 9 subpopulation of cells with a high disease relevance score. c-f) Gene-level analysis illustrating
 10 correlation of each gene with disease relevance score across the mural cell types. EC,
 11 endothelial cell; NPC, neural progenitor cell; OPC, oligodendrocyte progenitor cell; APC,

1 astrocyte progenitor cell; SMC, smooth muscle cell; aaSMC, arteriolar smooth muscle cell.

2 Created with BioRender.com

3

4

5

6

7

8

9

10

11

12

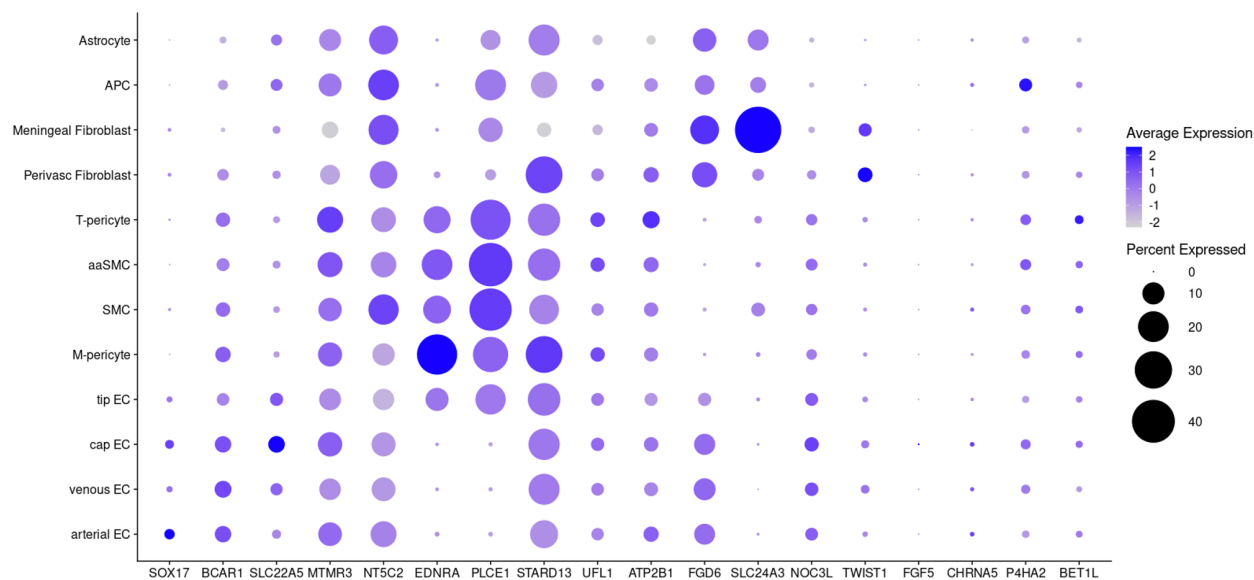
13

14

15

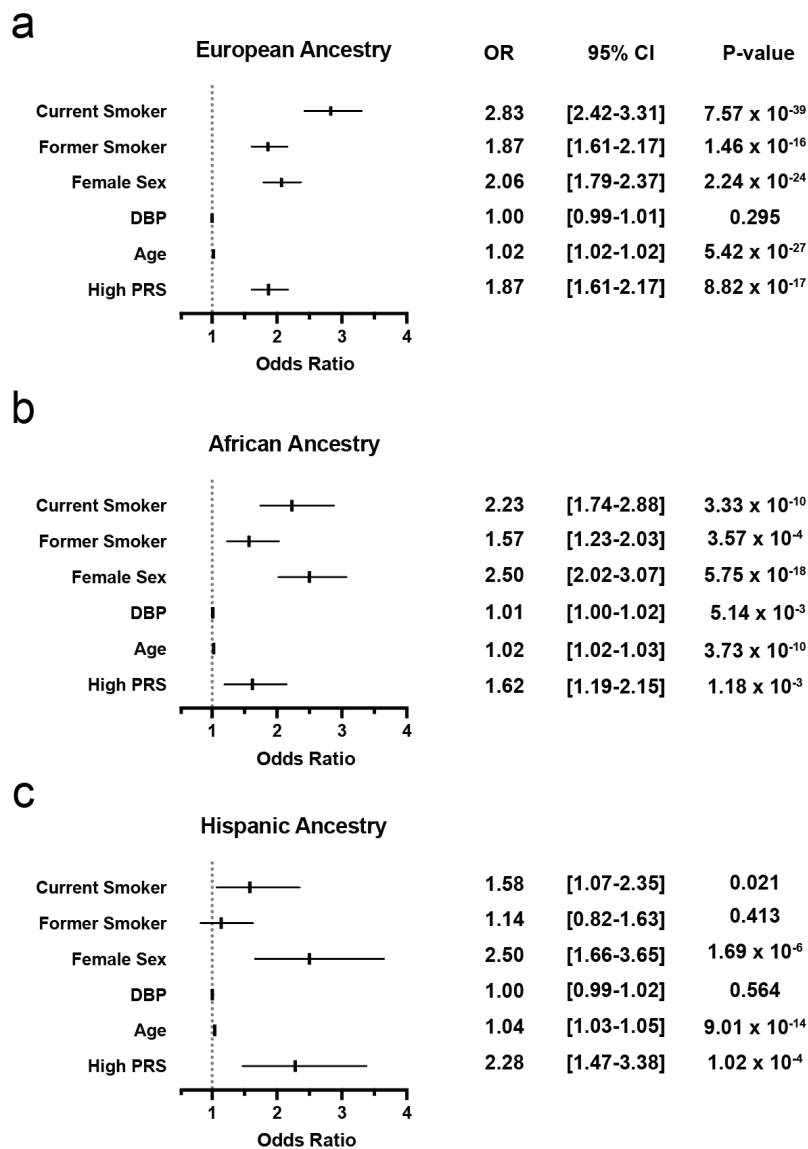
16

17



1
2 **Figure 5. Gene Expression of Candidate Causal Genes within Cerebrovascular Cell**
3 **Types.** Dot-plot representing cell-type specific expression of our candidate causal gene set. Dot
4 size represents proportion of cells with expression of the specified gene. Color gradient
5 represents the average of normalized expression values with dark blue indicating high
6 expression and light blue indicating low expression.

7
8
9
10
11
12



1
2 **Figure 6. Validation of a Polygenic Risk Score to Predict Intracranial Aneurysm.** Logistic
3 regression results in patients within the MVP stratified by a) European ancestry b) African
4 ancestry, and c) Hispanic ancestry. Forest plots (left) depict the odds ratio (OR) and 95%
5 confidence intervals (95% CI) of ICA diagnosis in patients within the top 5th percentile of a
6 polygenic risk score, controlling for current/former smoking, sex, age, diastolic blood pressure
7 (DBP) and the first five principal components.
8

Table 1: Demographic and Clinical Characteristics for Veterans in the MVP Discovery IA GWAS Analysis

	European		African		Hispanic	
	IA Cases	IA controls	IA Cases	IA controls	IA Cases	IA controls
Veterans, N	2,280	435,241	627	110,778	258	46,908
Age ± SD, years	65.8 ± 9.9	63.9 ± 13.6	59.8 ± 10.8	57.6 ± 12.3	61.6 ± 11.3	54.8 ± 15.6
Male, n (%)	2033 (89.2)	403,403 (92.7)	493 (78.6%)	95,486 (86.2)	224 (86.8)	42,560 (90.7)
Statin Therapy Prescription, n (%)	1,687 (74.0)	247,241 (56.8)	412 (65.7)	57,829 (52.2)	177 (68.6)	22,081 (47.1)
Diabetes, n (%)	886 (38.9)	146,914 (33.7)	276 (44.0)	46,594 (42.1)	126 (48.8)	17,843 (38.1)
Former Smoker, n (%)	1,329 (58.2)	245,934 (56.5)	289 (46.1)	51,015 (46.1)	150 (58.1)	26,134 (55.7)
Current Smoker, n (%)	738 (32.4)	102,728 (23.6)	248 (39.6)	34,196 (30.9)	62 (24.0)	9,526 (20.3)
Variants Passing Quality Control	8,913,234		15,458,904		10,138,272	

1

2 Table 2: Five Novel IA Susceptibility Loci Identified in Stage 2.

rsID	Chr	Locus	Consequence	EA	NEA	EAF	Effect (beta)	StdErr	Stage 1 P-value	Bakker et al. P-value	Stage 2 P-value
rs527725	1	<i>NAV1</i>	intron variant	a	c	0.631	0.0771	0.0138	1.84E-05	1.86E-04	2.25E-08
rs11099097	4	<i>FGF5</i>	intergenic variant	t	c	0.313	0.0766	0.0138	1.10E-02	4.73E-07	2.98E-08
rs59152651	6	<i>EXOC2</i>	downstream gene variant	a	g	0.74	-0.0965	0.0155	2.44E-05	4.23E-06	4.58E-10
rs57301765	7	<i>TWIST1</i>	intergenic variant	a	g	0.242	-0.0967	0.0161	5.80E-04	9.21E-07	1.95E-09
rs3118912	13	<i>DLEU1</i>	downstream gene variant	t	c	0.215	0.094	0.0169	5.31E-05	1.20E-04	2.83E-08

3

4 Abbreviations: Chr, chromosome; EA, Effect Allele; NEA, Non Effect Allele; EAF, Effect Allele
5 Frequency

6

7

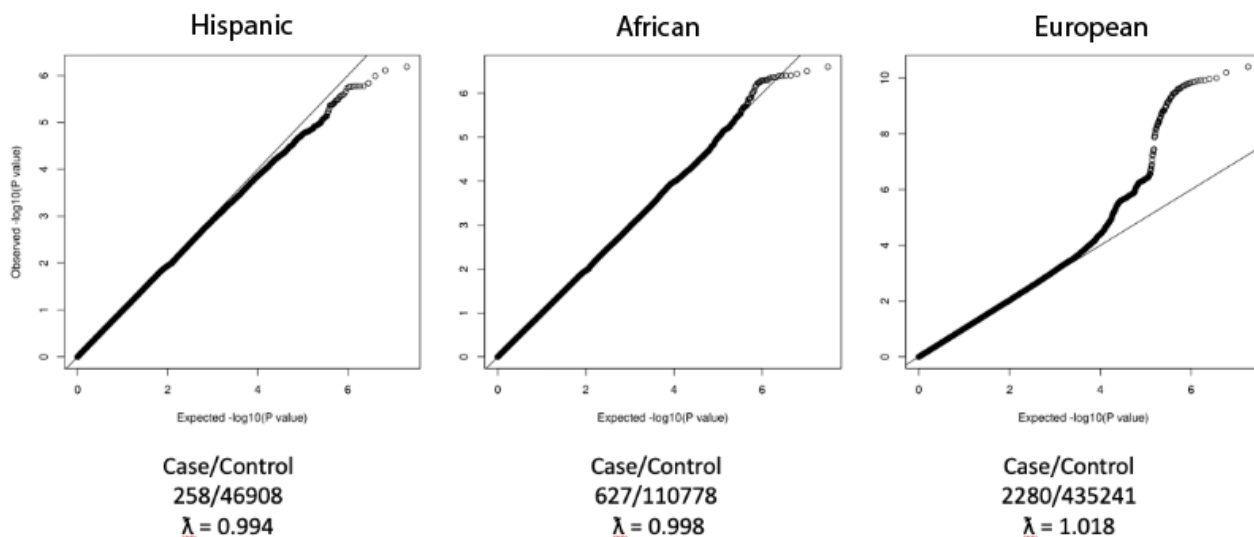
8

9

10

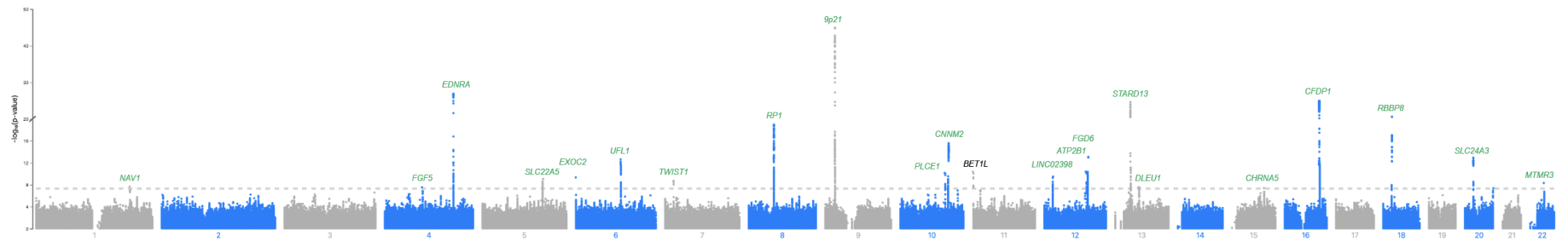
11

12



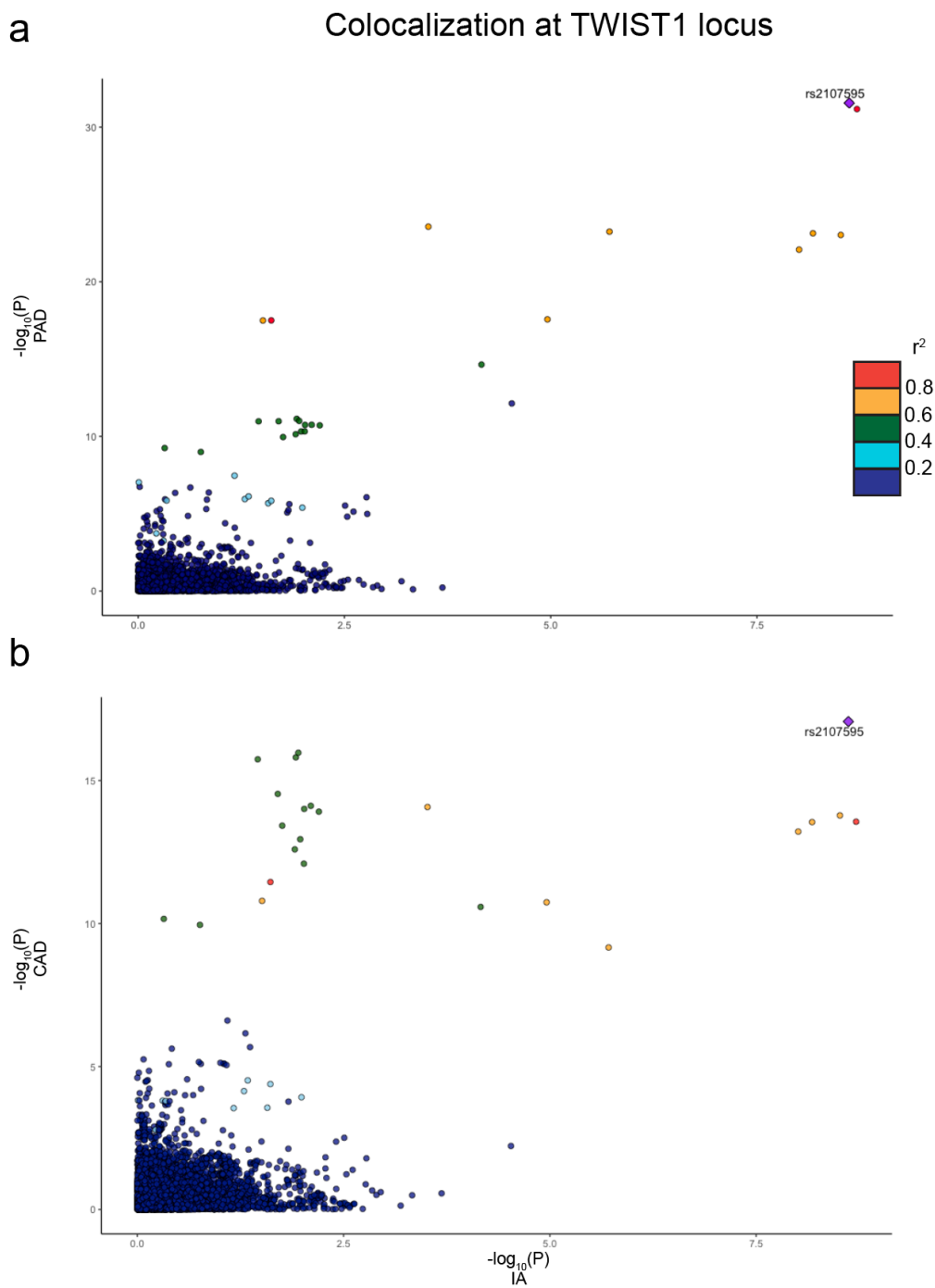
1
2 Supplemental Figure 1: Quantile-Quantile (QQ) plots depicting the expected logistic regression
3 association P-values versus the observed distribution of P-values for GWAS of each MVP
4 ancestry. No systemic inflation was observed. All P-values were two-sided. GWAS, Genome-
5 wide Association Study; MVP, Million Veterans Program

6
7
8
9
10
11
12
13
14
15
16

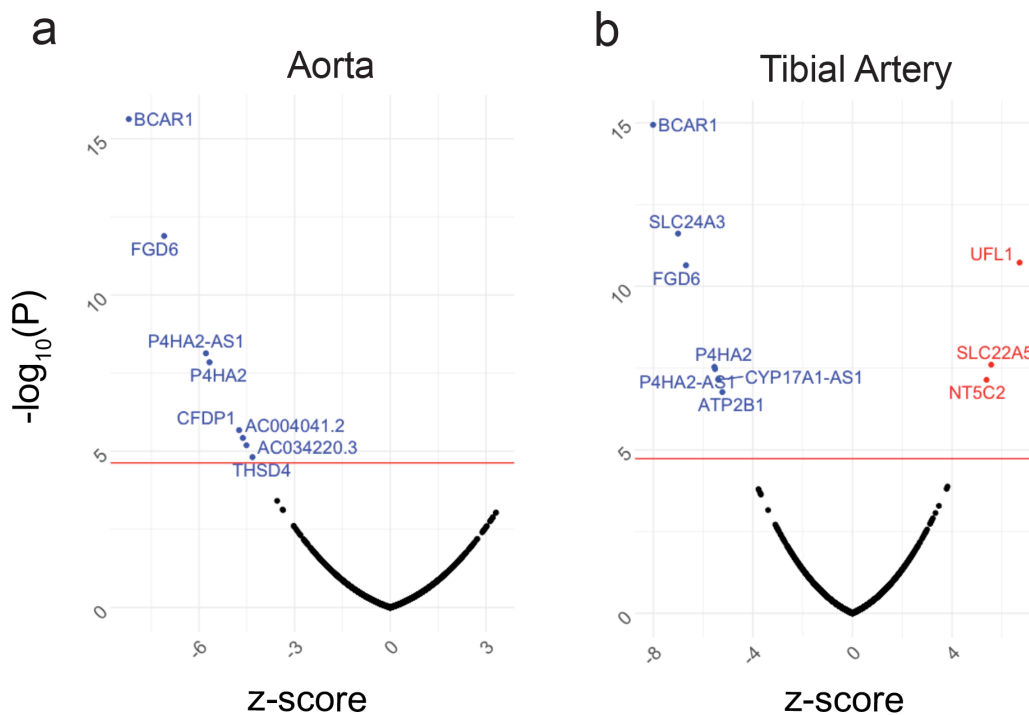


1

2 Supplemental Figure 2: Manhattan plot of Stage 2 summary statistics depicting association of genotyped and imputed variants with
 3 intracranial aneurysm. Dashed line indicates genome-wide significance threshold of $p < 5 \times 10^{-8}$. Annotated genes represent closest
 4 gene to lead variant. Green text denotes loci replicated in Bakker et al. summary statistics with at least nominal significance ($p <$
 5 0.05)

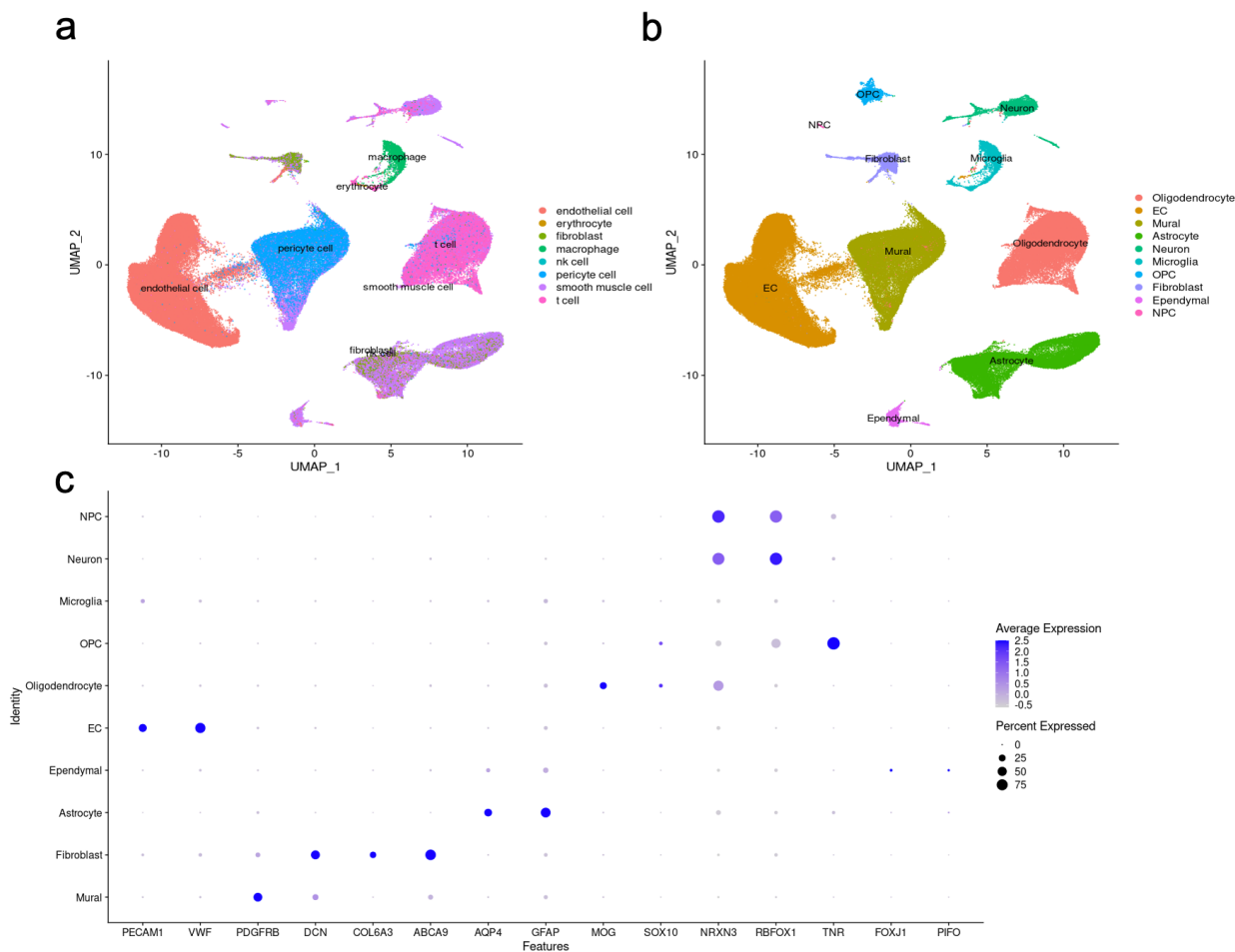


- 1
- 2 Supplemental Figure 3: Comparison of association P-values at TWIST1 locus for IA (x-axis) and
- 3 (a) CAD and b (PAD). Individual genetic variants are represented as points, color coded by
- 4 degree of genetic correlation (r^2) with the prioritized colocalizing variant (diamond).
- 5

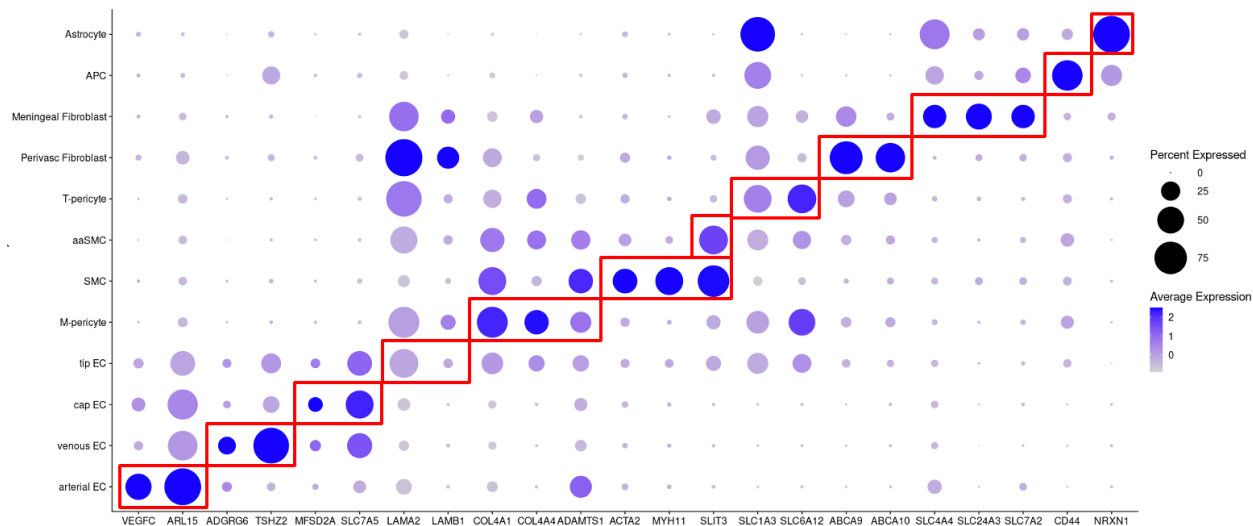


1
2 Supplemental Figure 4: Fine mapping transcriptome-wide association study using bulk RNA-
3 sequencing data from (a) aortic and (b) tibial tissue.

4
5
6
7
8
9
10
11
12
13
14

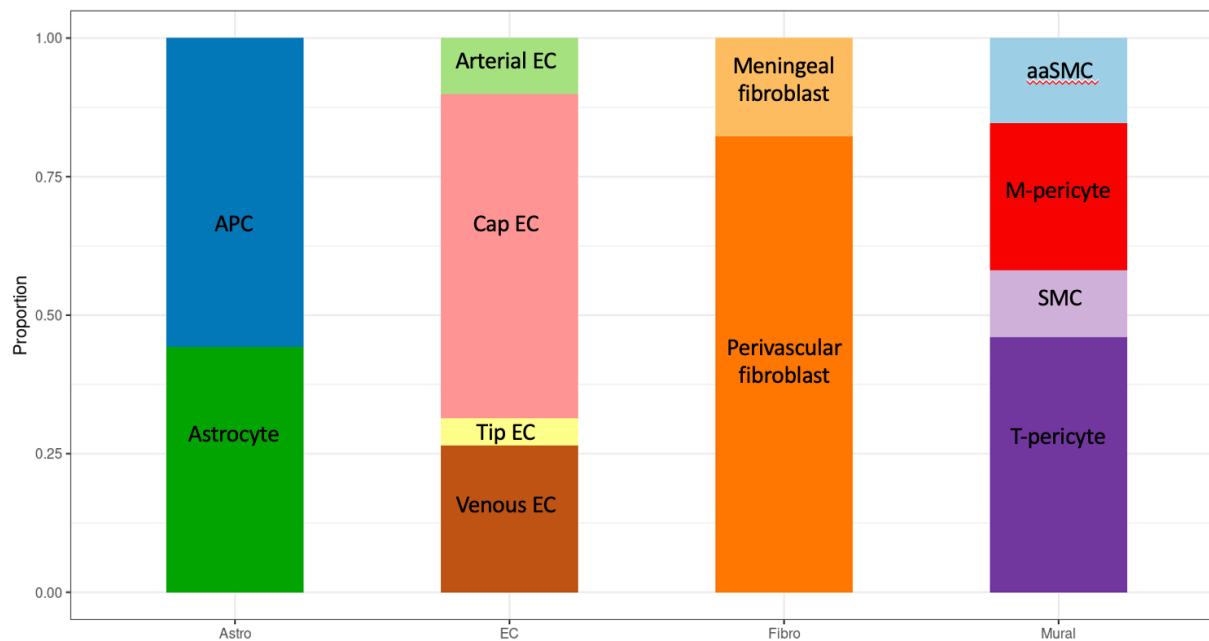


Supplemental Figure 5: Level 1 annotation of brain cerebrovasculature. a) UMAP representation of predicted vascular cell types using Tabula Sapiens reference annotation. b) Final level 1 cell type annotation using Tabula Sapiens and c) dot plot representation of canonical marker expression for each cell type.



Supplemental Figure 6: Level 2 Annotation of cerebrovasculature using previously described markers of vascular subtypes. Dot plot representing gene expression on vascular subtypes. Subtypes were annotated based on expression of genes highlighted in red boxes

1



2

3 Supplemental Figure 7: Proportions of cell sub-types amongst vascular stromal cell populations.

4 aaSMC, arteriolar SMC; Cap EC, capillary EC; M-pericyte, matrix-pericyte; T-pericyte, transport-
5 pericyte

6

7

8

9

10

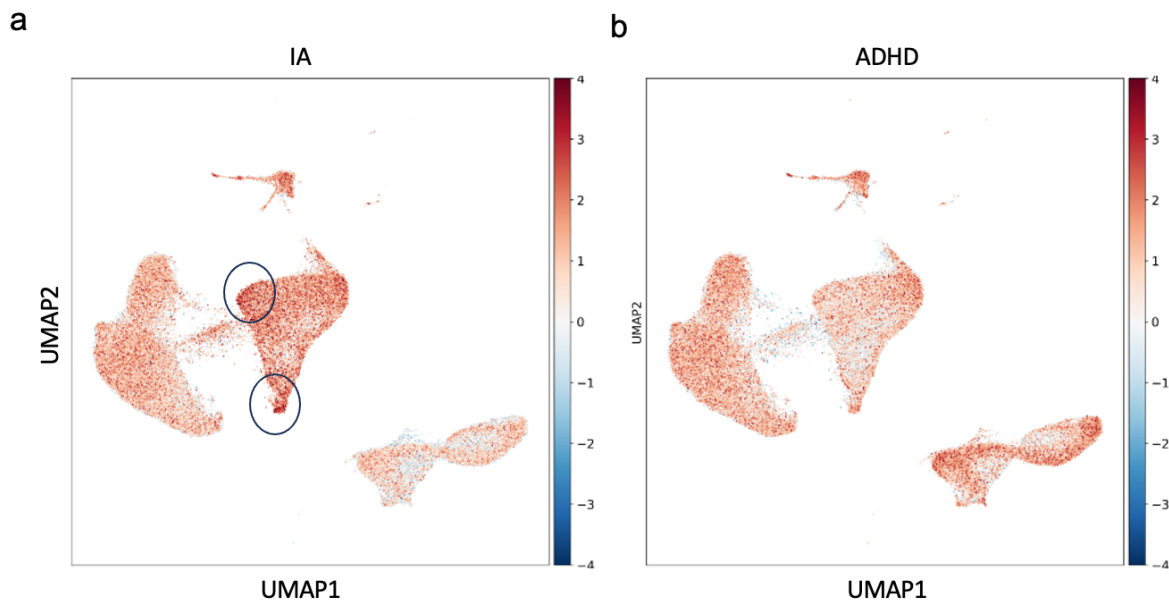
11

12

13

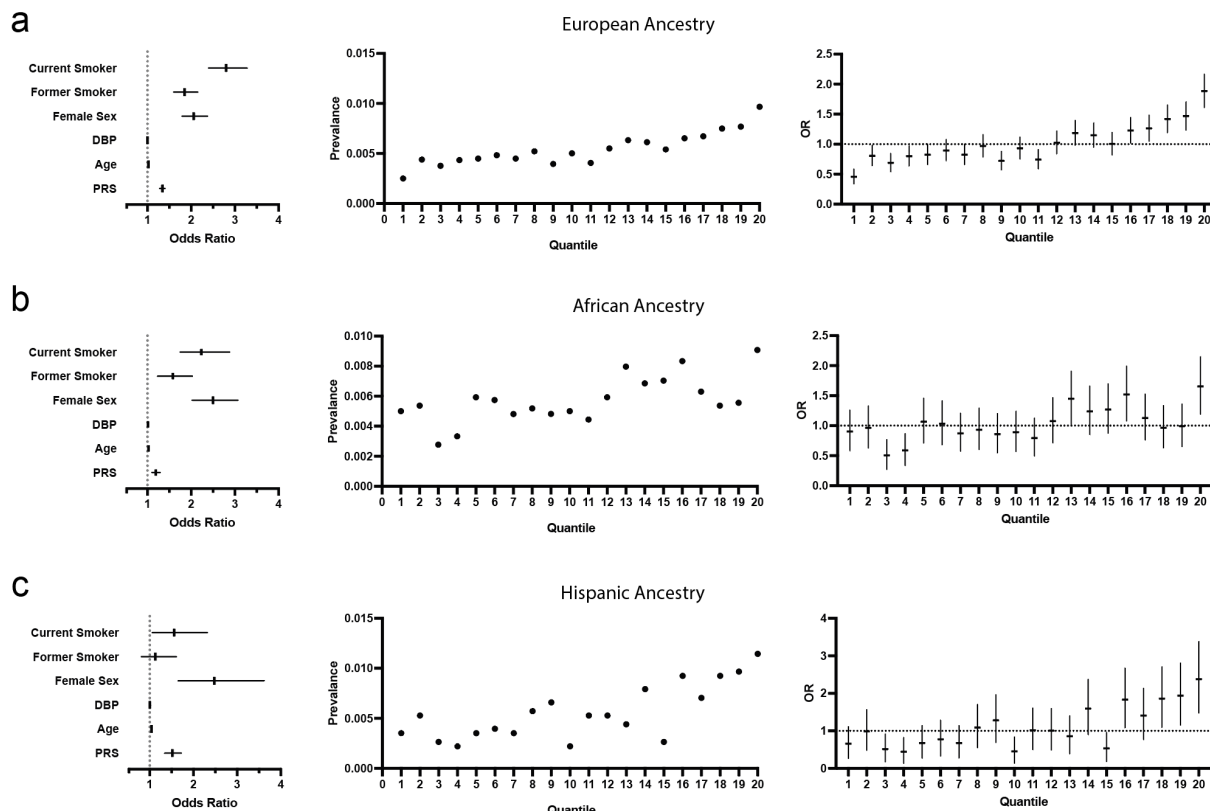
14

15



1
2 Supplemental Figure 8: Single cell resolution of single cell disease relevance scores using a)
3 Intracranial Aneurysm (IA) gene set and b) Attention deficit hyperactivity disorder (ADHD) gene
4 set as a control. Cells with high scDRS scores highlighted in (a).

5
6
7
8
9
10
11
12
13
14
15
16
17



1
2 Supplemental Figure 9: Generation of a Polygenic Risk Score to Predict Intracranial Aneurysm
3 Diagnosis. Evaluation of a PRS in a) European b) African and c) Hispanic ancestries. Risk of IA
4 diagnosis increases with higher polygenic risk scores (standard deviation unit) controlling for
5 demographic parameters (left). Prevalence (middle) and odds ratio (right) of IA increases with
6 PRS ventile.