Rupture of an intracranial aneurysm (IA) leads to aneurysmal subarachnoid haemorrhage (ASAH), a severe type of stroke. Some rare variants that cause IA in families have been identified, but still, the majority of genetic causes, as well as the biological mechanisms of IA development and rupture, remain unknown. We aimed to identify rare, damaging variants for IA in three large Dutch families with multiple affected members with IA (N = 9, 11, and 6). By combining linkage analysis and genome sequencing (GS), we identified six rare and damaging variants for which all cases within one of the families were heterozygous. These variants were p.Tyr87Cys in SYCP1, p.Phe1077Leu in FMLL2, p.Thr754Lys in TBC1D2, p.Arg321His in ZNF782, p.Arg979Trp in CCDC180, and p.Val125Met in NCBP1. None of the variants showed association with IA status in a large cohort of 937 patients from the general IA patient population and 1046 controls. Gene expression in IA and cerebral artery tissue further prioritized FMLL2 and TBC1D2 as potential important players in IA pathophysiology. Further studies are needed to characterize the functional consequences of the identified variants and their role in the biological mechanisms of IA.

INTRODUCTION

An intracranial aneurysm (IA) is a dilation of a cerebral blood vessel which are found in about 3% of the general population [1]. Rupture of an IA leads to aneurysmal subarachnoid haemorrhage (ASAH), a severe type of stroke with substantial morbidity and mortality with one in three patients dying [3]. Of the patients who survive ASAH, approximately half remain dependent on help from others in daily life [3]. Endovascular or surgical treatment of IA can prevent ASAH, thus avoiding its devastating consequences. However, currently, no clinically relevant risk models are available to identify persons at risk of developing and rupturing of IA.

The most important risk factors for IA and ASAH are age, female sex, hypertension, smoking, and a positive family history for ASAH [1, 2]. In persons with two or more first-degree family members with ASAH the life-time risk of ASAH is 50 times higher than in persons without such a family history [1, 4]. Moreover, individuals with an IA and a positive family history for ASAH have an almost 3-fold increased risk of rupture compared to individuals with IAs without such a family history (i.e., sporadic IAs) [5]. The heritability of ASAH has been estimated at 41% [6] and genome-wide association studies (GWAS) of large cohorts of IA cases have identified many common genetic risk factors with small effects, which combined explain roughly half of the total heritability [7]. Besides GWAS, sequencing studies in families with multiple affected family members with IA have identified rare IA-causing genetic variants with a presumable large effect located in the following genes: ADAMTS15 [8], THSD1 [9], ANGTL6 [10], PCNT [11], and ARHGEF17 [12]. For the genes THSD1 and ADAMTS15, linkage analyses had previously already highlighted regions containing these genes [13, 14], showing that linkage analysis with subsequent sequencing is a viable method to identify risk genes for IA.

Despite recent findings of IA risk genes and their associated biological functions, the primary causes of IA remain unknown. Here, we aim to identify rare, damaging genetic variants segregating with IA in three large Dutch families with multiple affected first-degree family members with an unruptured IA or ASAH using a combination of linkage analysis and genome sequencing.

METHODS

A flowchart of the methods is shown in Fig. 1.

Participants

Three large Dutch families with multiple affected first-degree relatives with an unruptured IA or an ASAH were selected: pedigree 1 included nine affected relatives in one generation, pedigree 2 included eleven relatives in three successive generations, while pedigree 3 included six in three successive generations (Fig. 2). Genealogical research showed no shared ancestor between these families for at least six generations back. For the variants identified in these families we further determined their role in the general IA patient population, using a cohort of 937 Dutch IA patients, of whom 160 had a positive family history for ASAH, from the University Medical Center Utrecht (Table 1). For the control cohort, GS genotypes from 1046 persons recruited within Project MinE were used [15, 16]. Presence of IA was determined with computed tomography, magnetic
Inclusion of unaffected family members. Depending on age categories and/or other variables, thereby allowing the parametric linkage analysis, was used to specify disease penetrance values. Heterozygous for a pathogenic variant, making the use of unaffected penetrance of IA, unaffected family members are may also be by-descent between affected family members. Due to the age-dependent control are described in the Supplementary Note. Two linkage methods polymorphism (SNP) genotyping was performed with the Global Screening Blood10k kit (PerkinElmer, Inc., Waltham, MA, USA). Single nucleotide resonance, or digital subtraction angiogram. Families or persons with Loeys-Dietz syndrome, Ehlers-Danlos disease, Marfan Syndrome, or autosomal dominant polycystic kidney disease were excluded.

**Linkage analysis**

The following number of family members were included for genotyping: nine affecteds and five unaffecteds from pedigree 1; six affecteds and 12 unaffecteds from pedigree 2; and four affecteds and eight unaffected from pedigree 3 (Fig. 2). DNA was extracted from blood with the Chemagic DNA Blood10k kit (PerkinElmer, Inc., Waltham, MA, USA). Single nucleotide polymorphism (SNP) genotyping was performed with the Global Screening Array (Illumina, San Diego, CA, USA). Detailed methods on the SNP quality control are described in the Supplementary Note. Two linkage methods were performed. First, we performed non-parametric linkage (NPL) analysis as it is an effective method to detect segments of DNA that are identical-by-descent between affected family members. Due to the age-dependent penetrance of IA, unaffected family members may also be heterozygous for a pathogenic variant, making the use of unaffected family members in NPL analysis ineffective. Therefore, the second method, parametric linkage analysis, was used to specify disease penetrance values depending on age categories and/or other variables, thereby allowing the inclusion of unaffected family members.

Non-parametric linkage analysis was performed using MORGAN *lm_ibdtest* [17]. MORGAN uses a Markov Chain Monte Carlo approach to estimate inheritance vectors: a representation of allele sharing by descent. In this iterative process, the first 5,000 iterations were discarded, before saving every 30th iteration from a total of 90,000 iterations for analysis. NPL scores Spairs and Srobdom were calculated [18, 19]. Spairs can detect both recessive and dominant inheritance loci, while Srobdom has the most power to detect linkage in a dominant model [19].

Parametric linkage analysis was performed in three steps: (1) obtaining inheritance vectors using MORGAN *gl_auto*, (2) determining pedigree baseline log likelihoods with MORGAN base_trait_lods, and (3) calculating logarithm of odds (LOD) scores with MORGAN *gl_lods* [20]. The *gl_auto* option was run with the same setting to estimate inheritance vectors as non-parametric linkage analysis. Four inheritance models; two dominant inheritance models and two recessive models, with either low or high penetrance, were used to calculate LOD-scores, using the liability class-specific penetrance matrix option of MORGAN. Phenocopy rate was set at the population prevalence of 3% [1]. Prevalence of IA tends to increase with age until, but the precise interaction between age and penetrance of high impact genetic variants is unknown [1]. Therefore, we selected two age-dependent penetrance models, reaching a maximum penetrance at age 70 (95% for the high penetrance models, 80% for the low penetrance models), and decreasing by 15% each younger decade until reaching the phenocopy rate (Supplementary Note). Expected allele frequency of the trait locus was set at 0.1% for dominant models and 1% for recessive models.

Regions with statistically significant linkage were defined as having LOD ≥ 3.3, or an equivalent NPL-score ≥ 3.90, derived from [21]:

\[
\text{LOD} = NPL^2 \div 2 \times \ln 10
\]

Or equivalent:

\[
NPL = \sqrt{\text{LOD} \times 2 \times \ln 10}
\]

Regions with suggestive statistically significant linkage were defined as LOD ≥ 1 or NPL-score ≥ 2.15.

**Genome sequencing**

From each family, two or three of the affected family members were selected for genome sequencing (GS), totaling seven patients (Fig. 2). Library preparation and sequencing was performed by Macrogen Inc. (Seoul, South Korea). The TruSeq DNA PCR-Free Shotgun Library was used on 350 bp fragments. Sequencing was done on the Illumina HiSeq system. Detailed methods about mapping, variants calling and quality control of GS data are given in the Supplementary Note.

**Variant selection**

After variant calling, variants located on regions identified by linkage analysis with suggestive statistical significance (LOD ≥ 1 or NPL-score ≥ 2.15) and for which all affected members of the family in which the region was identified were heterozygous, were retained. We also included variants located on previously identified linkage loci (loci from ref. [22], and specified in Supplementary Table 1). SnpEff was used to annotate variants (more details in the Supplementary Note) [23]. Variants in loci with suggestive significance were selected if they (1) had MAF ≤ 1% in non-Finnish Europeans in ExAC [24], Dutch controls from Project MinE [16], and Dutch controls from GoNL [25], or were not reported in any of these datasets, and (2) were disruptive of non-disruptive missense, which included any of the following SnpEff annotations: exon loss, frameshift, splice acceptor, splice donor, start site loss, stop site gain, stop site loss or transcript ablation, rare amino acid variant, inframe deletion, inframe insertion, or missense.

**Sanger sequencing**

In order to validate variants found using GS, and to test segregation in additional affected and unaffected family members, we used Sanger sequencing. In pedigree 1 an additional seven affected family members (resulting in nine affected family members with genotyping information in total) and five unaffected members were included; in pedigree 2 an additional three affected (six in total) and twelve unaffected, and in pedigree 3 an additional two affected (four in total) and eight unaffected members were included. PCR primers were designed using Primer3Plus, or if not possible by manual selection. Surrounding sequences (500 bp on both ends) were obtained from Ensembl v99. In order to design primers and probes that only bind to positions without genetic variants, the genetic variants with MAF ≥ 0.01% were converted to an N before uploading in the TaqMan assay design tool. An overview of the primers
is shown in Supplementary Table 2. Primer annealing temperature was optimized for each primer pair, by selecting the temperature resulting in the strongest specific amplification, as determined by gel electrophoresis. Genotypes were called by visually inspecting the Sanger sequencing chromatograms.

**TaqMan assay**

Segregating variants identified by linkage analysis and GS and validated using Sanger sequencing were additionally genotyped in a large cohort of 937 IA patients (Table 1) using TaqMan assay (Thermo Fisher Scientific Inc. Waltham, MA, USA). Details on assay design and optimization are given in the Supplementary Note. Quantitative polymerase chain reaction was conducted using the QuantStudio 6 Flex Real-Time PCR system (Applied biosystems). Polymerase chain reaction protocol was as follows: 10 min at 95 °C, 40 cycles of [15 s at 92 °C, 1 min at 60 °C], and 1 min at 60 °C. Fluorescence was read by the QuantStudio 6 Flex Real-Time PCR system and analyzed using the QuantStudio 6 and 7 system software. Genotypes of Dutch controls from Project MinE were obtained by GS as described before [15]. Association of the variants with IA was determined using two-tailed Firth logistic regression in PLINK v2.0. Sex and the first 10 genetic principal components projected on the HapMap v3 reference dataset were used as covariates. Only European ancestry individuals were included. These were defined by a maximum of five standard deviations distance from the mean of the genetically Dutch control samples from Project MinE in principal components 1–4. Since only persons homozygous for the ancestral allele, or heterozygous, were observed, a dominant inheritance model was selected. The analysis was also performed including only cases with positive family history versus controls. A P-value of 0.05 divided by the number of genes included with at least one observed mutated allele (N = 5) was considered statistically significant.

**Rare variant association and burden analysis**

To assess if additional rare variants in the genes containing potentially pathogenic variants segregating with IA, may be associated with IA, we performed two analyses using exome sequencing data from the UK Biobank: a rare variant burden analysis and a variant-level association analysis of potentially pathogenic variants. Variants were annotated based on predicted functional impact using SNPeff [23]. The association between IA and the number of rare, coding, minor alleles of the selected genes was calculated by two-tailed Firth logistic regression. For the variant-level

---

Fig. 2  Families recruited for linkage analysis and GS. Black: affected with an unruptured IA and/or ASAH. Grey: not included. White: no IA found at time of screening. Small white square: included in GS. Persons included in linkage analysis are depicted by an asterisk.
RESULTS

Gene expression analysis
For the genes in which segregating variants were found, we performed a lookup of the gene expression, and of differentially expressed genes between IA tissue and control cerebral artery tissue. We used RNA sequencing read counts, and differential expression data for IA tissue and control cerebral artery tissue analyzed and described before [26].

For the variant-level association analysis was defined as a P-value below 0.05 divided by the number of variants tested for that gene. More details are given in the Supplementary Note.

Gene expression analysis was assessed using two-tailed Firth logistic regression. Statistical significance was considered if the 95% confidence interval did not include 1.

Differential expression with a false discovery rate below 0.05 was considered statistically significant (LOD ≥ 3.3; statistically significant).

Eleven variants on four distinct regions with at least suggestive evidence for linkage (Table 2) were found using parametric linkage analysis. Non-parametric linkage analysis also identified five of these regions (Table 3). A total of 39 regions were found using positional linkage analysis, with a LOD ≥ 1.0. Regions are defined as least three consecutive markers passing this threshold, with gaps allowed if all markers in between have LOD > 1. Positions on GRCh37. Previous evidence are prior associations with IA, another aneurysm type, or another cerebrovascular disease. Power may not add up to the total N due to missing values. *Available for 84 patients with ASAH.

Table 2. Linkage analysis regions with LOD-score ≥ 3.3.

| Pedigree | Chromosome | Position | Locus | Statistic | Highest NPL | Previous evidence |
|----------|------------|----------|-------|-----------|-------------|-------------------|
| 1        | 2          | 150020329 – 159204593 | 2q23.3–24.1 | Spairs and Srobdom | 7.34 | Hypertension linkage analyses, IA GWAS |
| 2        | 1 217211217 – 218756117 | 2q35 | Srobdom | 7.48 |
| 9        | 9 93471627 – 101092018 | 9q22.2–22.33 | Srobdom | 7.24 |
| 2        | 9 24756027 – 25973292 | 9p21.3–21.2 | Srobdom | 6.00 |
| 10       | 10 72515775 – 72716955 | 10q22.1–22.2 | Srobdom | 6.92 |
| 14       | 14 2878618 – 34682723 | 14q12–13.1 | Srobdom | 6.73 |
| 3        | 5 81207272 – 90645137 | 5q14.2–14.3 | Srobdom | 5.15 |
| 5        | 5 171960073 – 174214185 | 5q35.1–35.2 | Srobdom | 5.21 |
| 6        | 6 84533341 – 94104527 | 6q14.2–16.1 | Srobdom | 4.40 |
| 7        | 7 36588659 – 39942647 | 7p14.2–14.1 | Srobdom | 5.26 |
| 20       | 20 53425217 – 54482619 | 20q13.2 | Srobdom | 5.28 |
| 22       | 22 48258293 – 51151350 | 22q13.3–13.33 | Srobdom | 5.37 |

Regions are defined as least three consecutive markers passing this threshold, with gaps allowed if all markers in between have LOD > 1. Positions on GRCh37. Previous evidence are prior associations with IA, another aneurysm type, or another cerebrovascular disease.

Table 1. General IA patient cohort characteristics.

| Characteristic | Value |
|----------------|-------|
| Number of IAs  | 937   |
| Number of cases | 643   |
| Ever smoked     | 641 (71.6%) |
| Hypertension    | 376 (40.7%) |
| Women           | 663 (70.8%) |
| Age at ASAH in years [mean (standard deviation)]* | 56.6 (13.4) |
| Women (N (%))   | 644 (69.3%) |
| Men (N (%))     | 273 (29.7%) |
| Positive family history (N (%)) | 317 (33.1%) |
| Women (N (%)    | 461 (68.8%) |
| Women (N (%))   | 169 (26.7%) |
| Men (N (%))     | 707 (14.3%) |
| Controls (N)    | 31.1 |
| Positive family history (N (%)) | 161 (44.3%) |
| Women (N (%))   | 663 (70.8%) |
| Women (N (%))   | 160 (18.1%) |
| Number of controls | 707 (14.3%) |

Participants (N) 937

Characteristics of the participants with ASAH: aneurysmal subarachnoid hemorrhage. Percentage multiplied by 100.

Power may not add up to the total N due to missing values. *Available for 84 patients with ASAH.
pedigree 1, and seven on 9q22.2–22.33 identified in pedigree 1 (Supplementary Table 5). On region 9q22.2–22.33, variant COL15A1 p.Lys1001Arg was found in all affected members of pedigree 2, while the region was identified by linkage in pedigree 1. Sanger sequencing in the remaining family members confirmed segregation of six of the 11 variants among all affected members, while for the other five, not all cases among one family were heterozygous, and these variants were excluded from further analysis (See Table 3 for both segregating and not segregating variants). The six variants that remained were: SYCP1 p.Tyr87Cys on 1p13.1–31.1 in pedigree 3, FMNL2 p.Phe1077Leu on 2q23.3–24.1 in pedigree 1, and four variants on 9q22.2–22.33 in pedigree 1: TBC1D2 p.Thr754Lys, ZNF782 p.Arg321His, CCD1C180 p. Arg979Trp, and NCBP1 p.Val125Met. All affected family members in pedigree 1 were heterozygous for both FMNL p.Phe1077Leu and TBC1D2 p.thr754Lys. Among the unaffected members of pedigree 1, one was heterozygous for both variants. Despite this genotype, this male member (II-8) had a negative screen for IA at age 65. Three unaffected family members were not heterozygous for either variant (II-4, II-7, and II-12), and one for neither (II-9).

Association of segregating variants in the general IA patient population
In the cohort of 937 IA patients (including 777 sporadic and 160 familial patients), for all variants except TBC1D2 p.Thr754Lys, at least one person was heterozygous for the risk allele among both cases and controls (Table 4), but none of the variants showed a statistically significant association with IA. When analysing only cases with a positive family history for ASAH also no association was found. In the UK Biobank dataset, no statistically significant rare variant burden was found for gene containing one of the segregating variants (Supplementary Table 6), and no additional potentially pathogenic variants with a statistically significant association with IA status were found (Supplementary Table 7).

Gene expression in intracranial aneurysm and cerebral artery tissue
The genes FMNL2 and TBC1D2 had high median expression levels (among top 20% highest expressed genes with non-zero median expression) in IA tissue as well as control cerebral artery tissue (median RPKM = 11.1 [range = 1.18–38.5] and 10.3 [1.55–61.9], respectively in the combined tissues), whereas NCBP1 and ZNF782 showed a lower expression and SYCP1 was expressed in too low amounts to be reliably detected (Supplementary Table 8). Differential expression analysis of these genes did not show difference in expression between aneurysms and control tissue for any of the genes.

| Table 3. Results of the segregation analysis of variants identified by genome-sequencing in additional family members. |
|-----------------|-----------------|-----------------|-----------------|-----------------|
| Pedigree variant | Transcript variant | Protein variant | MAF | Gene |
| 1   | 1:14574656:A   | NM_001354918.1:c.3892C > T | NP_001341847.1:p.(Arg1298Trp) | 7.24 | SYCP1 |
| 2   | 9:99054627:A   | NM_001855.4:c.3002A > G | NP_001846.3:p.(Lys1001Arg) | 7.24 | COL15A1 |
| 3   | 2:165593045:A  | NM_001345911.1:c.2382G > A | NP_001341847.1:p.(Arg1298Gln) | 7.24 | SYCP1 |
| 4   | 2:165593045:A  | NM_001345911.1:c.2382G > A | NP_001341847.1:p.(Arg1298Gln) | 7.24 | SYCP1 |

DISCUSSION
Using genome-wide linkage analysis combined with GS in large families with multiple affected members with IA, we identified six rare, damaging variants located in the genes SYCP1, FMNL2, TBC1D2, ZNF782, CCD1C180, and NCBP1 which segregated among all cases of one of these families. All genes except SYCP1 were expressed in intracranial aneurysm and control cerebral artery tissue, while FMNL2 and TBC1D2 showed high gene expression in these tissues, further prioritizing these genes. Genotyping the identified variants in a large cohort of IA patients did not show additional evidence that these variants play a role in IA patients beyond the families in which they were discovered.

Of the rare, damaging variants we identified, five were located on either of two regions with strong evidence (2q23.3–24.1 and 9q22.2–22.33), and one on a region with suggestive evidence (1p13.1–31.1). By Sanger sequencing, we confirmed that all affected family members of the respective families were heterozygous for the identified variants. For SYCP1 p.Tyr87Cys on 1p13.1–31.1, all four affected family members of pedigree 3 and
### Table 4: Variant associations in the general IA patient cohort.

| Gene variant ID | Carriers among controls (N = 1047) | P | Carriers among cases (N = 937) | OR (95%CI) | P | OR (95%CI) |
|-----------------|-------------------------------------|---|--------------------------------|------------|---|------------|
| 9q22.23 G>T    | 103 (0.43–2.54)                     | 0.035 | 1.11 (0.26–4.84)                      | 0.09      | - | -          |
| 9q22.23 C>T    | 0.59 (0.16–2.16)                     | 0.43 | 0.71 (0.10–4.97)                      | 0.73      | - | -          |
| 9q22.23 G>A    | 1.08 (0.27–4.43)                     | 0.21 | 1.08 (0.17–6.77)                      | 0.19      | - | -          |
| 9p21.3 T>0     | 0.68 (0.28–1.68)                     | 0.41 | 0.47 (0.07–2.83)                      | 0.28      | - | -          |

**Positions on GRCh38.**

**IA** intracranial aneurysm.

**OR** odds ratio from Firth logistic regression. 95%CI: 95% confidence interval.

**Table 4**

The table shows the gene variants and their associations with IA in the general and familial cases. The data includes the number of carriers among controls and cases, along with the odds ratio and 95% confidence interval. The table also indicates that further research is needed to determine whether one of the loci is a false positive and if a genetic variant on both loci is required to develop IA.

**Gene 9q22.23:** This locus contains the TGFBR2 gene, which predisposes mainly to thoracic aortic aneurysms but also to IA. The role of this variant is largely unclear, and its contribution to IA pathophysiology is not well understood.

**Gene 9p21.3:** This locus contains the EDIL3 gene. An exome sequencing study on a family of three members with ASAH found that EDIL3 p.Cys128Tyr was shared by all cases. In a cohort of 35 unrelated IA cases, the researchers found another variant in EDIL3 (p.Lys387Gln) in 2 cases. Here, we did not find any rare, damaging variants in EDIL3, nor in other genes in this locus. Locus 9p21.3 is a well-known GWAS risk locus for IA and several other diseases.

**Gene 9q22.23:** This locus contains the TGFBR2 gene, which predisposes mainly to thoracic aortic aneurysms but also to IA. The role of this variant is largely unclear, and its contribution to IA pathophysiology is not well understood.

**Gene 9p21.3:** This locus contains the EDIL3 gene. An exome sequencing study on a family of three members with ASAH found that EDIL3 p.Cys128Tyr was shared by all cases. In a cohort of 35 unrelated IA cases, the researchers found another variant in EDIL3 (p.Lys387Gln) in 2 cases. Here, we did not find any rare, damaging variants in EDIL3, nor in other genes in this locus. Locus 9p21.3 is a well-known GWAS risk locus for IA and several other diseases.
been identified as risk variant in GWAS of IA [7]. Both regions identified in this study are located ±3 megabases away from the lead GWAS variants, which could mean that these are located too far away to be driven by the same causal variants as the GWAS loci. No rare, damaging variants were found in these regions. This study has considerable strengths. First, the combination of linkage analysis with genome sequencing in large families provides an opportunity to detect rare damaging variants on segregating genomic regions. Second, using GS instead of exome sequencing provides a more uniform coverage of the genome and can detect more coding variants than exome sequencing [48], which was used in previous IA sequencing studies [8–12]. Another strength is the use of MORGAN for the linkage analysis. This allows joint analysis of the full pedigrees and leverages dense SNP information to provide high-resolution allele sharing information.

Some limitations apply to this study. First, the genetic association between variants discovered in this study and IA does not prove a causal relationship. Functional follow-up studies are needed to further prioritize these variants. Second, most regions of linkage were identified through an NPL statistic optimized for dominant inheritance (Srobdom), while only one and five of those were also identified using the Spairs statistic or parametric linkage analysis, respectively. Srobdom uses an exponential function that increases quickly with the number of affected persons who share an allele identical-by-descent [19]. This can lead to very high Srobdom in large pedigrees with many closely related persons such as pedigree 1 in this study. To deal with this, we used parametric and non-parametric linkage analysis, allowing for both dominant and recessive inheritance (by defining penetrance categories for parametric linkage analysis, and Spairs and Srobdom for NPL) for the purpose of detecting regions of linkage, and subsequently validated segregation among all affected members of a family using Sanger sequencing. Thereby we identified regions truly shared identical-by-descent among all affected family members, and thus eliminating this potential limitation. Another limitation is that we identified multiple variants of interest within a single locus of linkage, in particular 9q22.2–22.33. It can be expected that only a single causal variant is present on this locus, indicating the majority of variants segregates with IA simply by being in close proximity to this variant and are not causal. Including additional family members to increase resolution, identifying structural genetic variants, and performing functional experiments are ways to further prioritize variants in this locus.

In summary, we conducted a genome-wide linkage analysis paired with GS in three large families with IA and ASAH patients. This led to the identification of six potential IA-causing variants in families. We did not find evidence that these variants also contribute to IA in the general IA population. Regions of linkage showed overlap with previous genetic studies of IA and cerebrovascular and connective tissue disorders, providing insight in the pathogenetic mechanisms of IA. As a future direction of research, we suggest functional follow-up studies to discover the roles of the IA-associated genes, in particular FNNL2 and TBC1D2 which are highly expressed in cerebral artery and intracranial aneurysm tissue, and to discover how the damaging variant in those genes contribute to IA pathophysiology.

DATA AVAILABILITY STATEMENT
The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES
1. Vlak MH, Algra A, Brandenburg R, Rinkel GJ. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. Lancet Neurol. 2011;10:262–36.
2. Etminan N, Chang HS, Hackenberg K, de Rooij NK, Vergouwen MDI, Rinkel GJE, et al. Worldwide incidence of aneurysmal subarachnoid hemorrhage according to region, time period, blood pressure, and smoking prevalence in the population: a systematic review and meta-analysis. JAMA Neurol. 2019;76:588–97. https://doi.org/10.1001/jamaneurol.2019.0006.
3. Nieuwkamp DJ, Setz LE, Algra A, Linn FH, de Rooij NK, Rinkel GJ. Changes in case fatality of aneurysmal subarachnoid haemorrhage over time, according to age, sex, and region: a meta-analysis. Lancet Neurol. 2009;8:635–42.
4. Bor AS, Rinkel GJ, Adams J, Koopberg H, Emborn A, Buskens E, et al. Risk of subarachnoid haemorrhage according to number of affected relatives: a population based case-control study. Brain. 2008;131:2662–5. https://doi.org/10.1093/brain/awn187.
5. Mansfield LA, Greving JP, Verhoef TA, Rinkel GJE, Ruigrok YM. Comparison of rupture risk of intracranial aneurysms between familial and sporadic patients. Stroke. 2019;50:1380–3. https://doi.org/10.1161/STROKEAHA.118.023783.
6. Korja M, Silventoinen K, McCarron P, Zdravkovic S, Skytte A, Haaparanta A, et al. Genetic epidemiology of spontaneous subarachnoid hemorrhage: Nordic Twin Study. Stroke. 2010;41:2458–62.
7. Bakker MK, van der Spek RAA, van Rheenen W, Morel S, Bourcier R, Hostettler KC, et al. Genome-wide association study of intracranial aneurysms identifies 17 risk loci and genetic overlap with clinical risk factors. Nat Genet. 2020;52:1303–13. https://doi.org/10.1038/s41588-020-00725-7.
8. Yan J, Hitomi T, Takenaka K, Kato M, Kobayashi H, Okuda H, et al. Genetic study of intracranial aneurysms. Stroke. 2015;46:620–6.
9. Santiago-Sim T, Fang X, Hennessy ML, Naballach SV, DePalma SR, Lee MS, et al. THSD1 (Thrombospondin Type 1 domain containing Protein 1) mutation in the pathogenesis of intracranial aneurysm and subarachnoid hemorrhage. Stroke. 2016;47:3005–13.
10. Bourcier R, Le Scouarnec S, Bonnau S, Karakachoff M, Bourcereau E, Heutebise-Chétrêit S, et al. Rare coding variants in ANGPTL6 are associated with familial forms of intracranial aneurysms. Ann J Hum Genet. 2018;102:133–47.
11. Lorenzo-Betancor D, Blackburn PR, Edwards E, Vazquez-do-Campo R, Klee EW, Labbe C, et al. PCNT point mutations and familial intracranial aneurysms. Neurology. 2018;91:e2170–e2181. https://doi.org/10.1212/WNL.000000000006614.
12. Yang X, Li J, Fang Y, Zhang Z, Jin D, Chen X, et al. Rho guanine nucleotide exchange factor argef17 is a risk gene for intracranial aneurysms. Circ Genom Prec Med. 2018;11:e002099. https://doi.org/10.1161/CIRCGEN.117.002099.
13. Ozturk AK, Nahed BV, Bydon M, Bilgukar G, Goksu E, Bademci G, et al. Molecular genetic analysis of two large kindreds with intracranial aneurysms demonstrates linkage to 11q24-25 and 14q23-21. Stroke. 2006;37:1021–7. https://doi.org/10.1161/01.STR.0000206153.92675.b9.
14. Santiago-Sim T, Depalma SR, Ju KL, McDonough B, Seidman CE, Seidman JG, et al. Genomewide linkage in a large Caucasian family maps a new locus for intracranial aneurysms to chromosome 13q. Stroke. 2009;40:557–60. https://doi.org/10.1161/STROKEAHA.108.534396.
15. van der Spek RAA, van Rheenen W, Pult SL, Kenna KP, van den Berg LH, Veldink JH, et al. The project MiNE database: bringing large-scale whole-genome sequencing in ALS to researchers and the public. Amyotroph Lateral Scler Frontotemporal Degener. 2019;20:432–40. https://doi.org/10.1080/21678421.2019.1606244.
16. Project MiNE. ALS Sequencing Consortium. Project MiNE: study design and pilot analyses of a large-scale whole-genome sequencing study in amyotrophic lateral sclerosis. Eur J Hum Genet. 2018;26:1537–46. https://doi.org/10.1038/s41431-018-0177-4.
17. Basu S, Di Y, Thompson EA. Exact trait-model-free tests for linkage detection in pedigrees. Ann Hum Genet. 2008;72:576–82. https://doi.org/10.1111/j.1469-1809.2008.00451.x.
18. Whittemore AS, Halpern J. A class of tests for linkage using affected pedigree members. Biometrics. 1994;50:118–27.
19. McPeek MS. Optimal allele-sharing statistics for genetic mapping using affected relatives. Genet Epidemiol. 1999;16:225–49. 10.1002/(SICI)1098-2272(1999) 16:3–225:AID-GEPI1-3.0.CO;2-4
20. Thompson E. The structure of genetic linkage data: from LIPED to 1M SNPs. Hum Hered. 2011;71:86–96. https://doi.org/10.1159/000313555.
21. Abreu PC, Greenberg DA, Hodge SE. Direct power comparisons between simple LOD scores and NPL scores for linkage analysis in complex diseases. Am J Hum Genet. 1999;65:847–57. https://doi.org/10.1086/302536.
22. Zhou S, Dion PA, Rouleau GA. Genetics of intracranial aneurysms. Stroke. 2018;49:780–7. https://doi.org/10.1161/STROKEAHA.117.018152.
23. Cipriani P, Wu Q, Platts A, Wang W, Dion PA. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly. 2012;6:80–92. https://doi.org/10.1212/WNL.0b013e3181f6965.
24. Karczewski KJ, Francioli LC, Tsao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 genomes. Nature. 2020;581:434–43. https://doi.org/10.1038/s41586-020-2308-7.
25. Genome of the Netherlands Consortium. Whole-genome sequence variation, population structure and demographic history of the Dutch population. Nat Genet. 2014;46:818–25. https://doi.org/10.1038/ng.3021.

26. Kleinloog R, Verweij BH, van der Vlies P, Deelen P, Swertz MA, de Muynck L, et al. RNA Sequencing Analysis of Intracranial Aneurysm Walls Reveals Involvement of Lysosomes and Immunoglobulins in Rupture. Stroke. 2016;47:1286–93. https://doi.org/10.1161/STROKEAHA.116.012541.

27. Kim CJ, Park SS, Lee HS, Chung HJ, Choi W, Chung JH, et al. Identification of an autosomal dominant locus for intracranial aneurysms through a model-based family collection in a geographically limited area. J Hum Genet. 2011;56:464–6. https://doi.org/10.1038/jhg.2011.27.

28. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019;47:D506–D515. https://doi.org/10.1093/nar/gkz1049.

29. Block J, Breitbrecher D, Kuhn S, Winterhoff M, Kage F, Geffers R, et al. FMN2 drives actin-based protrusion and migration downstream of Cdc42. Curr Biol. 2012;22:1005–12. https://doi.org/10.1016/j.cub.2012.03.064.

30. Ostberg, NP, Zafar, MA, Ziganshin, BA & Elefteriades, JA. The genetics of thoracic aortic aneurysms and dissection: a clinical perspective. Biomedecules. 2020;10:182.

31. Frasa MA, Maximiano FC, Smolarchzyk K, Francis RE, Betson ME, Smolarczyk K, et al. Genome of the Netherlands Consortium. Whole-genome sequence variation, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The online version contains supplementary material available at https://doi.org/10.1038/s41431-022-01059-0.

32. Smith JR, Hayman GT, Wang SJ, Laulederkind SJF, Hoffman MJ, Kaldunski ML, et al. The year of the rat: the rat genome database at 20: a multi-species knowledgebase and analysis platform. Nucleic Acids Res. 2020;48:D731–D742. https://doi.org/10.1093/nar/gkz1041.

33. Izzurralde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW. A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell. 1994;78:657–68. https://doi.org/10.1016/0092-8674(94)90530-4.

34. Hong, EP, Kim, BJ, Cho, SS, Yang, JS, Choi, HJ, Kang, SH et al. Genomic variations in susceptibility to intracranial aneurysm in the Korean population. J Clin Med 8, https://doi.org/10.3390/jcm8020275 (2019).

35. Guo Y, Wang F, Li L, Akracki S, Wang IZ, et al. Genome-wide linkage analysis of large multiple multigenational families identifies novel genetic loci for coronary artery disease. Sci Rep. 2017;7:5472 https://doi.org/10.1038/s41598-017-0381-2.

36. Gao H, Li L, Rao S, Shen G, Xi Q, Chen S, et al. Genome-wide linkage scan identifies two novel genetic loci for coronary artery disease: in GeneQuest families. PLoS One. 2014;9:e113935 https://doi.org/10.1371/journal.pone.0113935.

37. Kim ST, Brinjikji W, Kallmes DF. Prevalence of intracranial aneurysms in patients with connective tissue diseases: a retrospective study. Am J Neuroradiol. 2016;37:122–26. https://doi.org/10.3174/ajnr.a4718.

38. Loeys BL, Schwarze U, Holm T, Callewaert BL, Thomas GH, Pannu H, et al. Aneurysm syndromes caused by mutations in the TGF-beta receptor. N. Engl J Med. 2006;355:788–98. https://doi.org/10.1056/NEJMoa055695.

39. Pomianowski P, Elefteriades JA. The genetics and genomics of thoracic aortic disease. Ann Cardiothorac Surg. 2013;2:271–9. https://doi.org/10.3978/jissn.2225-319X.2013.05.12.

40. Kurki MI, Gaal EI, Kettunen J, Lappalainen T, Menelaou A, Anttila V, et al. Genome-wide scan for premature hypertension supports linkage to chromosome 2 in a large Kyrgyz family. Hypertension. 2006;48:908–13. https://doi.org/10.1161/01.HYP.0000244107.13957.2b.

41. Williams PJ, Broughton Pipkin F. The genetics of pre-eclampsia and other hypertensive disorders of pregnancy. Best Pr Res Clin Obstet Gynaecol. 2011;25:401–17. https://doi.org/10.1016/j.bogyn.2011.02.007.

42. Mein CA, Caufield MJ, Dobson RJ, Munroe PB. Genetics of essential hypertension. Hum Mol Genet. 2004;13 Spec No 1:S169–75. https://doi.org/10.1093/hmg/ddh078.

43. Guo DC, Pannu H, Tran-Fadulu V, Papke CI, Yu RK, Avidan N, et al. Mutations in smooth muscle alpha-actin (ACTA2) lead to thoracic aortic aneurysms and dissections. Nat Genet. 2007;39:1488–93. https://doi.org/10.1038/ng.2007.3.

44. Sauvigny T, Alawi M, Krause L, Benner S, Spohn M, Busch A, et al. Exome sequencing in 38 patients with intracranial aneurysms and subarachnoid hemorrhage. J Neurol. 2020;267:2533–45. https://doi.org/10.1007/s00415-020-09865-6.

45. Nikpay M, Goel A, Won HH, 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;47:121–30. https://doi.org/10.1038/ng.3396.

46. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestery genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet. 2018;50:524–37. https://doi.org/10.1038/s41588-018-0058-3.

47. Bélkadi A, Boltz A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc Natl Acad Sci USA. 2015;112:5473–8. https://doi.org/10.1073/pnas.1418631112.

ACKNOWLEDGEMENTS
We like to thank Dr. V. Guryev for providing genetic variant information from the GoNL data resource. This research was conducted using the UK Biobank Resource under application 2532. This study was supported by the Netherlands Cardiovascular Research Initiative: An initiative with support of the Dutch Heart Foundation, CVON2015-08 ERASE. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (PRYSM, grant agreement No. 852173).

COMPETING INTERESTS
The authors declare no competing interests.

ETHICAL APPROVAL
All participants of the study signed informed consent.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41431-022-01059-0.

Correspondence and requests for materials should be addressed to Mark K. Bakker.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.