# PLOS ONE

RESEARCH ARTICLE

# *ADAMTS12*, a new candidate gene for pediatric stroke

**Anika Witten**[1], **Frank Rühle**[1], **Marlous de Witt**[1], **Andrei Barysenka**[1], **Michael Stach**[2], **Ralf Junker**[3]☯, **Ulrike Nowak-Göttl**[3]☯, **Monika Stoll**[1,4]☯ *

1 Institute of Human Genetics, Genetic Epidemiology, University of Münster, Münster, Germany, 2 IT Service Centre, University Hospital of Münster, Münster, Germany, 3 Institute of Clinical Chemistry, University Hospital Schleswig-Holstein, Kiel, Germany, 4 Department of Biochemistry, Cardiovascular Research Institute Maastricht, Maastricht University, Maastricht, The Netherlands

☯ These authors contributed equally to this work.
* mstoll@uni-muenster.de

## Abstract

We recently reported a family-based genome wide association study (GWAS) for pediatric stroke pointing our attention to two significantly associated genes of the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) gene family *ADAMTS2* (rs469568, $p = 8 \times 10^{-6}$) and *ADAMTS12* (rs1364044, $p = 2.9 \times 10^{-6}$). To further investigate these candidate genes, we applied a targeted resequencing approach on 48 discordant sib-pairs for pediatric stroke followed by genotyping of the detected non-synonymous variants in the full cohort of 270 offspring trios and subsequent fine mapping analysis. We identified eight non-synonymous SNPs in *ADAMTS2* and six in *ADAMTS12* potentially influencing the respective protein function. These variants were genotyped within a cohort of 270 affected offspring trios, association analysis revealed the *ADAMTS12* variant rs77581578 to be significantly under-transmitted ($p = 6.26 \times 10^{-3}$) to pediatric stroke patients. The finding was validated in a pediatric venous thromboembolism (VTE) cohort of 189 affected trios. Subsequent haplotype analysis of *ADAMTS12* detected a significantly associated haplotype comprising the originally identified GWAS variant. Several ADAMTS genes such as *ADAMTS13* are involved in thromboembolic disease process. Here, we provide further evidence for *ADAMTS12* to likely play a role in pediatric stroke. Further functional studies are warranted to assess the functional role of ADAMTS12 in the pathogenesis of stroke.

## Introduction

Pediatric stroke happens with an incidence of 2.6–6.4 per 100 000 children per year and is still one of the top 10 causes of death in children [1, 2]. Extracellular matrix components like members of the ADAMTS (a disintegrin and metalloproteinase with thrombospondin motifs) gene family, primarily *ADAMTS13*, in combination with misbalanced coagulation signals appear to play an important role in pediatric stroke etiology after postnatal vascular injuries [3, 4].

Recently, we have reported results of a family-based genome wide association study (GWAS) for pediatric stroke pointing our attention to *ADAMTS2* and *ADAMTS12* [1].

*ADAMTS2* (rs469568, $p = 8 \times 10^{-6}$) and *ADAMTS12* (rs1364044, $p = 2.9 \times 10^{-6}$) were both highly associated with pediatric stroke. Furthermore, an association of rs1364044 was described having a protective effect in cerebral aneurysms (CA) pathogenesis in a case control study (OR 0.65) [5]. *ADAMTS2* is a pro-collagen N-propeptidase and known to be involved in pro-collagen processing and has properties of angiogenesis inhibition and interfering with tumor growth. The protein is involved in the innate immune system, playing a role in calming inflammation and tissue repair [6]. *ADAMTS12*, is classified as proteinase of the cartilage oligomeric matrix protein (COMP), it is inducing neutrophil apoptosis in mice, and silenced *ADAMTS12* genes were found in human tumor cells. As *ADAMTS2*, the protein is showing antiangiogenic characteristics [7].

To further investigate the molecular background and function in pediatric stroke of the two genes we conducted a targeted resequencing approach on 48 discordant sib-pairs and a subsequent fine mapping analysis in the full cohort of 270 affected trios.

## Material and methods

### Subjects

The study was performed in accordance with the Helsinki Declaration of 1975 and was approved by the ethical standards of the medical ethics committee, University of Münster (2008-161-f-S) and of the University Clinics Schleswig-Holstein (UKSH) (B 304/16), Germany. Written confirmed consent was obtained from all participants or their parents. Enrollment and genotyping of affected pediatric stroke family trios, probands selection and characteristics of 48 independent pediatric probands and 48 unaffected siblings has been described by Arning et al. and Stoll et al. [1, 3]. From the same catchment area, 189 offspring trios for pediatric venous thromboembolism (VTE) enrolled from neonates and children with confirmed diagnosis of VTE not older than 18 years at onset and having available nonaffected biological brothers, sisters and parents were utilized for the pseudo-replication study and are described in detail by Rühle et al. [8].

### Targeted resequencing and data analysis

The resequencing approach and subsequent next generation sequencing (NGS) data analysis was the same described in detail in our previous study [3]. In brief, 1μg genomic DNA isolated from EDTA blood samples was applied to perform a TruSeq DNA Sample Prep (Illumina) followed by target capture using the NimbleGen SeqCap EZ Choice Library Preparation System (Roche). The resulting libraries were pooled equimolar and sequenced in a paired-end mode (100 cycles) using the TruSeq PE Cluster Kit v3 and TruSeq SBS chemistry v2 (Illumina) on a HiScanSQ sequencing system (Illumina). After quality control, on average of 382,245 read pairs per sample were retained for further analysis. High-quality reads were mapped to the reference genome (hg19) using BWA software (ver. 0.6.2) [9]. Over 99% of the reads were mapped resulting in a mean coverage of the target region of about 227. Reads were realigned around indels and quality score recalibration was performed using GATK 2.0 [10]. GATK tool UnifiedGenotyper was used to produce a raw variant set (SNPs, short indels). In order to filter false positive variants from this set, "hard filtering" of the variants was performed [11, 12]. The Protein Variation Effect Analyzer (PROVEAN) web tool was used to assign Sorting Intolerant From Tolerant (SIFT) scores for the (non-synonymous) variants [13]. Combined Annotation Dependent Depletion (CADD) scores were looked up on https://cadd.gs.washington.edu/. Sequencing data is deposited to the Sequence Read Archive (SRA) and accessible with the following link: http://www.ncbi.nlm.nih.gov/bioproject/647652.

## Association and haplotype testing

For association analysis, a family-based approach was implemented: A custom software based on the Sibship-Disequilibrium-Test (SDT) allows to test for association and linkage and does not require parental genotypes [14]. Rare-variant association analysis was performed on SNPs with an minor allele frequency (MAF) of ≤0.02 without ambiguous genotypes (NoCalls) using a C-alpha test statistic [15]; both tests are implemented in PLINK/SEQ (http://atgu.mgh.harvard.edu/plinkseq). Empirical *p*-values were estimated using 10000 permutations per gene. SNP haplotypes for the 48 discordant sib-pairs were tested by FBAT [16] using a sliding window approach ranging from 13 (*ADAMTS12*) to 15 (*ADAMTS2*) SNPs, the optimal window size was defined by the minimal *p*-value corrected for 1 million permutations. Median joining (MJ) networks were drawn by Network 4.6.1.1 (http://fluxus-engineering.com).

## Genotyping and association analyses

Genotyping in 270 affected offspring trios of 15 SNPs selected by sequencing was carried out with TaqMan SNP Genotyping Assays (Life Technologies) as described in Stoll et al. [3]. The missing-genotype rate for all SNPs was < 1%, Hardy-Weinberg-Equilibrium (HWE) for founders was tested using an exact test as implemented in PLINK software package version 1.07 [17] and was met for all. The total calling rate was 94% (non-synonymous SNPs *ADAMTS2*), 98.6% (non-synonymous SNPs *ADAMTS12*) and 95.6% (tagging SNPs). Association of pediatric stroke affection status was calculated using the Transmission Disequilibrium Test (TDT) as implement in PLINK without adjustment for covariates. The Bonferroni-method was used for multiple testing correction. Association of haplotypes was calculated using the family trio association test implemented in Haploview 4.2 [18] defining blocks by the 4 gamete rule, examine haplotypes >5% and applying a HW *p*-value cutoff of 0.05. Pseudo-replication in 189 pediatric VTE family trios for rs77581578 was performed using the same settings.

A *p*-value cutoff of 0.05 was applied for all statistical test. The overall study design and workflow is shown in S1 Fig.

## Results

### Targeted resequencing of *ADAMTS2* and *ADAMTS12* in 48 discordant sib-pairs

Based on our association findings in the previous GWA study [1] for *ADAMTS2* and *ADAMTS12* we selected two candidate regions for subsequent NGS based on linkage disequilibrium (LD) information using the postgwas package [19]. For *ADAMTS2* a 275 kbp region (Chr5, 178,525,000–178,800,000) and for *ADAMTS12* a 415 kbp region (Chr5, 33,510,000–33,925,000) was selected for subsequent target enrichment and paired-end sequencing in 48 children affected with pediatric stroke and their unaffected siblings. After data processing, we were able to identify 1201 SNPs and indels in the region of *ADAMTS2* and 1435 for *ADAMTS12* (see S1 and S2 Tables). Out of these variants, eight SNPs (*ADAMTS2*) and six SNPs (*ADAMTS12*) were annotated as non-synonymous in the coding regions in the two genes and represent interesting new disease candidates.

For association analysis of the identified variants a Sibship Disequilibrium Test (SDT) was used, no significant *p*-values were detected likely due to power restrictions (see S3 Table).

### Testing for rare variants within *ADAMTS2* and *ADAMTS12*

Our targeted resequencing approach identified 287 (*ADAMTS2*) and 406 (*ADAMTS12*) rare variants (MAF≤0.02) in the region of the two genes not considered for the SDT association

analysis and we applied a C-alpha test statistic to test for an accumulation of rare variants in the affected children. This collapsing test yielded a significant association with a combined $p$-value of $p = 3.8x10^{-2}$ for *ADAMTS2* and $p = 4.5 x10^{-2}$ for *ADAMTS12*.

## Median joining networks

To assess the genetic structure of our two candidate genes *ADAMTS2* and *ADAMTS12* we performed a haplotype analysis based on our targeted resequencing data of the 48 discordant sib-pairs and visualized the results applying median joining (MJ) networks (Fig 1). The MJ networks enable to infer the evolutionary relationship of the haplotypes given as the length of the edges. The ancestral *ADAMTS2* haplotype H_7 showed the strongest association signal ($p = 9.2x10^{-5}$). The $p$-value for the most strongly associated haplotype for *ADAMTS12* (H_1) was $p = 5.2x10^{-4}$. Both haplotypes are located in the propeptide domain of the respective ADAMTS gene and point on potential functional changes fixed in our study population.

## *ADAMTS2* and *ADAMTS12* non-synonymous variant validation in 270 family trios

The non-synonymous called variants identified in our applied NGS approach on 48 discordant sib-pairs for *ADAMTS2* and *ADAMTS12* and not implemented on the previously used array for genome-wide genotyping [1] (rs1054480, rs398829 *ADAMTS2*; rs25754, rs3813474
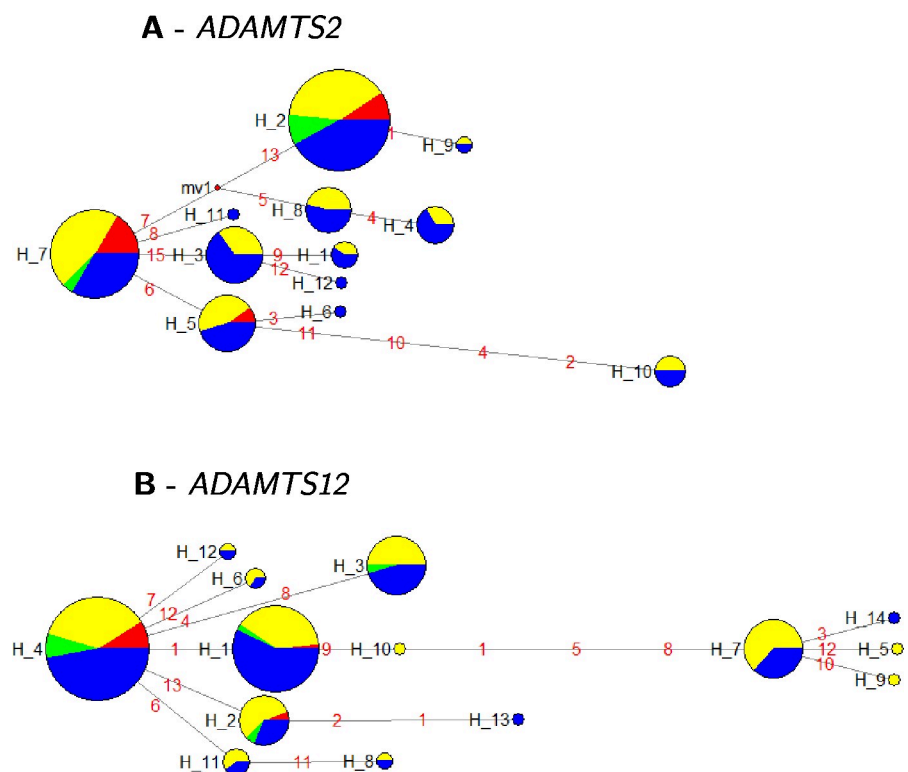


**Fig 1. Median joining networks for *ADAMTS2* and *ADAMTS12*.** Median joining networks of our detected haplotypes from *ADAMTS2* (A) and *ADAMTS12* (B) genes with indicated domain positions. Node size indicates the quantity of probands showing the haplotype either hetero- or homozygous. Color code represents the distribution of haplotypes in cases and controls (case/heterozygous = yellow, case/homozygous = red, control/heterozygous = blue, control/ homozygous = green). Evolutionary distance between haplotypes is weighted for allele frequencies and illustrated by edge length.

*ADAMTS12*) were forwarded to validation in the full cohort of 270 affected offspring trios for pediatric stroke already applied in our former GWAS study [1] and subsequent family-association testing using TDT (see Table 1). While no significant association was found for variants residing in *ADAMTS2*, we observed the *ADAMTS12* variant rs77581578 to be significantly under-transmitted ($p = 6.3x10^{-3}$) to pediatric stroke patients. This protective variant resides within a thrombospondin 1 domain of *ADAMTS12* and leads to replacement of proline by threonine at amino acid position 1329 (transcript ID NP_001311441). This change is predicted as damaging to the protein according to the respective SIFT score of 0.029 and a CADD score of 23.3 and potentially influences the protein's function and adhesion properties.

A separate cohort of 189 affected child trios suffering from venous thromboembolism (VTE) was used for pseudo-replication of the variant rs77581578 and subsequent TDT analysis revealed a significant validation signal in the replication cohort of $p = 0.025$ (Bonferroni corrected; T = 0, U = 5).

## Fine mapping of *ADAMTS12*

Due to the results of the single marker analysis above we decided to focus on *ADAMTS12* and to perform a fine mapping analysis. A dominant LD block was noticed (subset 1, see Fig 2) for *ADAMTS12*. Fine mapping was performed in this subset because it accumulates several marginal GWAS association results and is also carrying the strongest GWAS association signal (rs1364044). Beside rs1364044, four intronic haplotype tagging SNPs were selected within this subset for genotyping in the full cohort of 270 offspring trios. As quality control for the genotyping data an additional SNP carrying the same information as rs1364044 (rs7443937) was selected for TaqMan genotyping (see Table 2). A further promising variant, rs1530507 detected by our targeted resequencing approach, was introduced to the genotyping step due to its stop gained function. None of these variants except for rs7443937 ($p = 3.1x10^{-3}$) showed a significant association with pediatric stroke using a single marker TDT analysis.

The genotype data was utilized for LD and haplotype analysis and it was possible to define five haplotypes in the full cohort that occur with a frequency greater than 5% (see Table 3). Two haplotypes (CCCTC and CCCTT) having a frequency of 42% and 21% respectively were

**Table 1. Family-association testing results of coding *ADAMTS2* and *ADAMTS12* variants.**

| SNP | position | A1 | A2 | MAF | T | U | OR (95% C.I.) | Chi$^2$ | *p* | Bonf |
|---|---|---|---|---|---|---|---|---|---|---|
| *ADAMTS2* | | | | | | | | | | |
| rs35445112 | 178,555,097 | T | C | 0.026 | 11 | 17 | 0.65 (0.3–1.38) | 1.29 | 0.256 | 1 |
| rs35372714 | 178,563,002 | T | C | 0.003 | 1 | 1 | 1 (0.06–15.99) | 0 | 1 | 1 |
| - | 178,634,681 | T | C | 0.003 | 2 | 1 | 2 (0.18–22.06) | 0.33 | 0.563 | 1 |
| rs11750821 | 178,634,683 | T | C | 0.107 | 47 | 57 | 0.82 (0.56–1.21) | 0.96 | 0.327 | 1 |
| rs59567206 | 178,634,704 | C | T | 0.286 | 131 | 128 | 1.02 (0.8–1.31) | 0.035 | 0.852 | 1 |
| *ADAMTS12* | | | | | | | | | | |
| rs61753559 | 33,546,207 | C | T | 0.005 | 3 | 2 | 1.5 (0.25–8.98) | 0.20 | 0.655 | 1 |
| rs112196098 | 33,549,350 | G | C | 0.045 | 21 | 25 | 0.84 (0.47–1.5) | 0.35 | 0.555 | 1 |
| **rs77581578** | 33,549,374 | T | G | 0.009 | 0 | 10 | 0 | 10 | 2x10$^{-3}$ | 6.3x10$^{-3}$ |
| rs140486982 | 33,891,894 | A | G | 0.006 | 2 | 3 | 0.67 (0.11–3.99) | 0.20 | 0.655 | 1 |

Family-association testing of nine non-synonymous variants of *ADAMTS2* and *ADAMTS12*. Based on our targeted resequencing result for *ADAMTS2* and *ADAMTS12* variants were selected and forwarded to subsequent genotyping in the full cohort of 270 affected offspring trios for pediatric stroke.

(position = genomic location on chromosome 5, A1 = minor allele, A2 = major allele, MAF = minor allele frequency, T = transmitted minor allele count,

U = untransmitted allele count, OR = odds ratio, C.I. = confidence interval, *p* = TDT *p*-value, Bonf = Bonferroni corrected *p*-value).
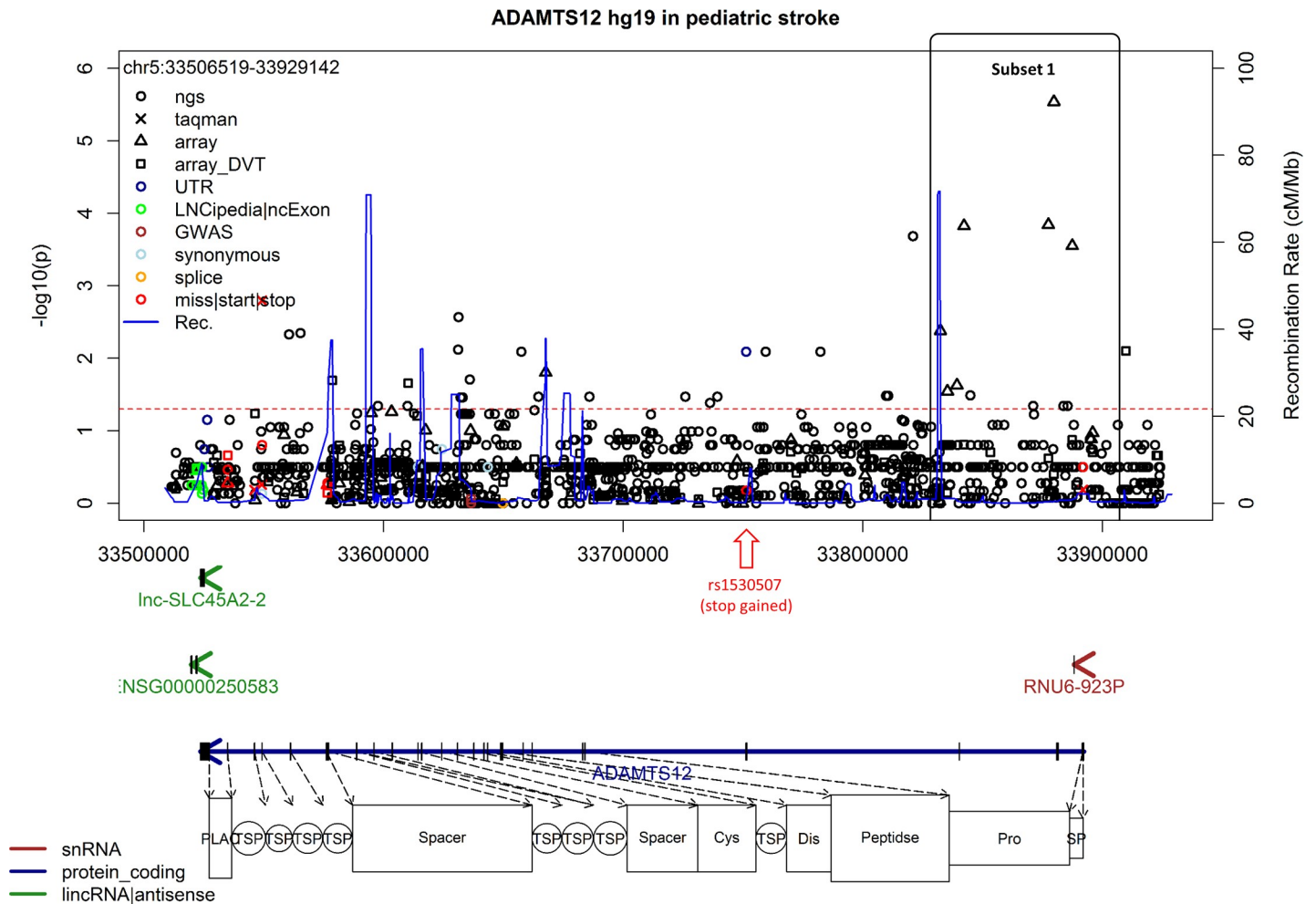
**Fig 2. *ADAMTS12* regional plot.** Regional plot of *ADAMTS12* showing the dominant haplotype block and association of the detected SNPs marker with pediatric stroke.

https://doi.org/10.1371/journal.pone.0237928.g002

significantly associated with pediatric stroke. Haplotype CCCTC ($p = 5 \times 10^{-4}$) is under-transmitted from parents to children and the corresponding haplotype CCCTT ($p = 6.6 \times 10^{-5}$) over-transmitted.

**Table 2. Single marker association analysis of *ADAMTS12*.**

| subset | SNP ID | A1 | A2 | MAF | T | U | OR (95% C.I.) | p | Bonf |
|---|---|---|---|---|---|---|---|---|---|
| 1 | rs114803348 | A | C | 0.22 | 81 | 108 | 0.75 (0.56–1.00) | 0.049 | 0.73 |
| 1 | rs17501324 | T | C | 0.08 | 47 | 36 | 1.458 (0.85–2.01) | 0.227 | 1 |
| 1 | rs139318738 | T | C | 0.05 | 28 | 24 | 1.17 (0.68–2.01) | 0.579 | 1 |
| 1 | rs75989721 | G | A | 0.29 | 101 | 117 | 0.86 (0.66–1.13) | 0.279 | 1 |
| 1 | **rs7443937** | C | T | 0.38 | 153 | 98 | 1.56 (1.21–2.01) | $5.2 \times 10^{-4}$ | $3.1 \times 10^{-3}$ |
| | rs1530507 | A | T | 0.27 | 79 | 70 | 1.13 (0.82–1.56) | 0.461 | 1 |

Single marker association analysis for the tagging SNPs in the two subsets of *ADAMTS12* for subsequent genotyping in the full cohort of 270 affected offspring trios for pediatric stroke.

(A1 = minor allele, A2 = major allele, MAF = minor allele frequency, T = transmitted minor allele count, U = untransmitted allele count, OR = odds ratio, C.I. = confidence interval, *p* = TDT *p*-value, Bonf = Bonferroni corrected *p*-value).

https://doi.org/10.1371/journal.pone.0237928.t002

**Table 3. Haplotype association results for *ADAMTS12*.**

| Haplotype | Frequency | T | U | Chi² | *p* |
|---|---|---|---|---|---|
| CCCTC | 0.42 | 104.2 | 161.2 | 12.26 | $5.0 \times 10^{-4}$ |
| CCCTT | 0.21 | 121.8 | 67.0 | 15.94 | $6.55 \times 10^{-5}$ |
| CCACC | 0.17 | 67.6 | 94.9 | 4.59 | 0.032 |
| CTCTT | 0.09 | 48.7 | 34.8 | 2.33 | 0.127 |
| CCCCC | 0.07 | 40.9 | 29.4 | 1.89 | 0.69 |

Haplotype association results for *ADAMTS12*, Subset1. Association analysis was performed using a TDT test (given *p*-value).

(T = transmitted minor allele count, U = untransmitted allele count, *p* = TDT *p*-value).

https://doi.org/10.1371/journal.pone.0237928.t003

## Discussion

Here we report the identification and validation of several non-synonymous *ADAMTS2* and *ADAMTS12* variants to further prove the role of the two members of the ADAMTS gene family in the disease etiology of pediatric stroke. *ADAMTS2* (rs469568, $p = 8 \times 10^{-6}$) and *ADAMTS12* (rs1364044, $p = 2.9 \times 10^{-6}$) were first associated with the occurrence of pediatric stroke in a family-based GWAS published by our group [1]. To our knowledge, this comprehensive cohort of 270 families affected with pediatric stroke is unique and offers the outstanding possibility to characterize the genetic background of the disease and to identify the respective underlying molecular mechanisms without any environmental risk factors for stroke in adults such as smoking or obesity. Interestingly, the *ADAMTS12* variant rs1364044 has recently shown to be associated with the pathogenesis of cerebral aneurysms emphasizing the role of the gene in stroke related phenotypes [5].

Therefore, we conducted a NGS based targeted resequencing approach to trace our previous GWAS findings, where we first sequenced the surrounding genomic regions based on LD information of the *ADAMTS2* and *ADAMTS12* GWAS hits in 48 discordant sib-pairs. We were able to detect several synonymous and non-synonymous variants within the coding regions of the two genes, which strengthen their role as interesting novel candidate genes for pediatric stroke. Although we applied a discordant sib-pair approach and SDT testing, which should provide an advantage to achieve sufficient statistical power while testing a reasonable sample size [20] due to sequencing costs, we were not able to achieve statistical significance for any of these variants.

We are aware that, despite the discordant sib-pair study design, resequencing of only 96 individual samples is a limitation to test for rare variants. Nevertheless, we applied a gene based collapsing test (C-alpha) on the identified rare variants. We were able to support our previous association findings for the two genes by detecting a significant C-alpha *p*-value of $p = 3.8 \times 10^{-2}$ for *ADAMTS2* and $p = 4.5 \times 10^{-2}$ for *ADAMTS12*. Based on our resequencing data, we applied haplotype MJ network analysis revealing first evidence on the contribution of the propeptide domain of both genes in disease etiology which is discussed later and an evolutionary fixation on our study population.

In contrast to our SDT association findings we were able to validate the non-synonymous *ADAMTS12* variant rs77581578 (TDT $p = 6.3 \times 10^{-3}$) in the full cohort of 270 affected offspring trios for pediatric stroke. The thrombospondin (TSP)-1 domain variant is predicted as influencing protein function and protein´s adhesion properties, this finding is in accordance with previous reports on ADAMTS gene family members such as *ADAMTS13* linking disease associated variants of the thrombospondin domain and protein function in pediatric stroke [3]. An independent replication of our association finding in another study population would strengthen the impact of our findings. We lack an independent replication cohort for pediatric

stroke as to our knowledge; no comparable collection of father-mother-child trios is available in the scientific community. Indeed, we were in the fortunate position to have access to a study cohort for pediatric VTE recruited from the same catchment area. Pediatric stroke and VTE share a similar genetic background as well as a comparable prevalence of prothrombotic risk factors, and—in our opinion—is the most suitable cohort for validation of our findings given that no other pediatric stroke cohort is available to conduct a true replication. Notably, the pediatric stroke and VTE share many associated clinical conditions and diseases such as neo-natal infections, leukemia or sickle cell disease [21, 22]. Most notably many prothrombotic abnormalities are prominent risk factors for pediatric stroke und both diseases share known susceptibility genes as *prothrombin* or *factor V-Leiden* [22, 23]. Furthermore, risk factors differ between children and adults due to the lack of classical environmental risk factors like smoking stressing the need to replicate findings within pediatric cohorts [6]. Interestingly, we were able to pseudo-replicate the variant rs77581578 (TDT $p$ = 0.025) in this cohort but we were not able to validate the stop gained variant rs1530507. Our findings point on *ADAMTS12* as a promis-ing new candidate gene for pediatric stroke hence we focused on this gene for a subsequent fine mapping analysis to refine the identified association signal. The resulting significantly associated common haplotype is, in accordance with the initial GWAS, signal located in the in the intronic regions in-between the propeptide domain of *ADAMTS12*.

## Conclusion

Taken together, our study confirms the role of *ADAMTS12* as a promising candidate gene for pediatric stroke. Our approach has shown that the region of the associated haplotype block in *ADAMTS12* and the variant rs77581578, is of particular importance in identifying the underly-ing functional mechanisms in pediatric stroke. *ADAMTS12* is known to degrade cartilage olig-omeric matrix protein (COMP), a non-collagenous protein in cartilage. Abnormalities in this protein usually result in skeletal dysplasia. Produced by vascular smooth muscle cells and platelets, it has also a key role in maintaining cardiovascular haemostasis and downregulating coagulation by inhibition of thrombin [24]. An increase in *ADAMTS12* activity would hypo-thetical lead to the degradation of COMP resulting in less inhibition of thrombin and an ele-vated risk for ischemic stroke. However, the *ADAMTS12*-COMP-thrombin axis has not yet been reported in the context of (pediatric) stroke. The determination of *ADAMTS12* levels in our study cohort in the future would allow to establish a causal relationship between *ADAMTS12* levels and pediatric stroke susceptibility. Nevertheless, interaction of the TSP-1 like repeats of *ADAMTS12* with the EGF-like repeat domain in COMP had been reported in absence of other domains, including the propeptide domain [25]. Therefore, the propeptide domain may only have a marginal impact on the catalytic function of *ADAMTS12*. Further functional studies are warranted to assess the role of *ADAMTS12* in disease etiology.

## Supporting information

**S1 Fig. Study design.** Workflow and design of the study including sample sizes and methods. (TIF)

**S1 Table. Sibship Disequilibrium Test (SDT) for *ADAMTS2* and *ADAMTS12* variants.** Variants detected using NGS based target enrichment in the defined genomic region of *ADAMTS2*. (DOCX)

**S2 Table. Sibship Disequilibrium Test (SDT) for *ADAMTS2* and *ADAMTS12* variants.** Variants detected using NGS based target enrichment in the defined genomic region of

*ADAMTS12*.
(DOCX)

**S3 Table. Sibship Disequilibrium Test (SDT) for *ADAMTS2* and *ADAMTS12* variants.**
Results of the SDT association analysis in 48 discordant sib-pairs for the non-synonymous variants in *ADAMTS2* and *ADAMTS12*.
(DOCX)

# Acknowledgments

# Author Contributions

**Conceptualization:** Anika Witten, Frank Rühle, Ulrike Nowak-Göttl, Monika Stoll.

**Data curation:** Michael Stach.

**Formal analysis:** Anika Witten, Frank Rühle, Marlous de Witt, Andrei Barysenka.

**Funding acquisition:** Ulrike Nowak-Göttl, Monika Stoll.

**Investigation:** Anika Witten, Marlous de Witt.

**Resources:** Ulrike Nowak-Göttl, Monika Stoll.

**Supervision:** Ralf Junker, Ulrike Nowak-Göttl, Monika Stoll.

**Visualization:** Frank Rühle.

**Writing – original draft:** Anika Witten.

**Writing – review & editing:** Ulrike Nowak-Göttl, Monika Stoll.

# References

1. Arning A, Hiersche M, Witten A, Kurlemann G, Kurnik K, Manner D, et al. A genome-wide association study identifies a gene network of ADAMTS genes in the predisposition to pediatric stroke. Blood. 2012; 120(26):5231–6. Epub 2012/09/20. https://doi.org/10.1182/blood-2012-07-442038 PMID: 22990015.

2. Association AH. Knowing No Bounds: Stroke in Infants, Children, and Youth www.strokeassociation.org: American Heart Association; [February 14, 2019].

3. Stoll M, Ruhle F, Witten A, Barysenka A, Arning A, Strauss C, et al. Rare Variants in the ADAMTS13 Von Willebrand Factor-Binding Domain Contribute to Pediatric Stroke. Circ Cardiovasc Genet. 2016; 9 (4):357–67. Epub 2016/07/15. https://doi.org/10.1161/CIRCGENETICS.115.001184 PMID: 27412500.

4. Reilly MP, Li M, He J, Ferguson JF, Stylianou IM, Mehta NN, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. Lancet. 2011; 377(9763):383–92. Epub 2011/01/18. https://doi.org/10.1016/S0140-6736(10)61996-4 PMID: 21239051; PubMed Central PMCID: PMC3297116.

5. Arning A, Jeibmann A, Kohnemann S, Brokinkel B, Ewelt C, Berger K, et al. ADAMTS genes and the risk of cerebral aneurysm. J Neurosurg. 2016; 125(2):269–74. Epub 2016/01/09. https://doi.org/10.3171/2015.7.JNS154 PMID: 26745484.

6. Stoll M, Ruhle F, Nowak-Gottl U. Advances in understanding stroke risk in children—a geneticist's view. Br J Haematol. 2014; 164(5):636–45. Epub 2013/12/21. https://doi.org/10.1111/bjh.12706 PMID: 24354735.

7. El Hour M, Moncada-Pazos A, Blacher S, Masset A, Cal S, Berndt S, et al. Higher sensitivity of Adamts12-deficient mice to tumor growth and angiogenesis. Oncogene. 2010; 29(20):3025–32. Epub 2010/03/09. https://doi.org/10.1038/onc.2010.49 PMID: 20208563.

8. Ruhle F, Witten A, Barysenka A, Huge A, Arning A, Heller C, et al. Rare genetic variants in SMAP1, B3GAT2, and RIMS1 contribute to pediatric venous thromboembolism. Blood. 2017; 129(6):783–90. Epub 2016/12/25. https://doi.org/10.1182/blood-2016-07-728840 PMID: 28011674.

9. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25(14):1754–60. Epub 2009/05/20. https://doi.org/10.1093/bioinformatics/btp324 PMID: 19451168; PubMed Central PMCID: PMC2705234.

10. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010; 20(9):1297–303. Epub 2010/07/21. https://doi.org/10.1101/gr.107524.110 PMID: 20644199; PubMed Central PMCID: PMC2928508.

11. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013; 43:11 0 1–33. Epub 2014/11/29. https://doi.org/10.1002/0471250953.bi1110s43 PMID: 25431634; PubMed Central PMCID: PMC4243306.

12. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011; 43(5):491–8. Epub 2011/04/12. https://doi.org/10.1038/ng.806 PMID: 21478889; PubMed Central PMCID: PMC3083463.

13. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One. 2012; 7(10):e46688. Epub 2012/10/12. https://doi.org/10.1371/journal.pone.0046688 PMID: 23056405; PubMed Central PMCID: PMC3466303.

14. Horvath S, Laird NM. A discordant-sibship test for disequilibrium and linkage: no need for parental data. Am J Hum Genet. 1998; 63(6):1886–97. Epub 1998/12/05. https://doi.org/10.1086/302137 PMID: 9837840; PubMed Central PMCID: PMC1377659.

15. Neale BM, Rivas MA, Voight BF, Altshuler D, Devlin B, Orho-Melander M, et al. Testing for an unusual distribution of rare variants. PLoS Genet. 2011; 7(3):e1001322. Epub 2011/03/17. https://doi.org/10.1371/journal.pgen.1001322 PMID: 21408211; PubMed Central PMCID: PMC3048375.

16. Laird NM, Horvath S, Xu X. Implementing a unified approach to family-based tests of association. Genet Epidemiol. 2000; 19 Suppl 1:S36–42. Epub 2000/10/31. https://doi.org/10.1002/1098-2272(2000)19:1+<::AID-GEPI6>3.0.CO;2-M PMID: 11055368.

17. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007; 81(3):559–75. Epub 2007/08/19. https://doi.org/10.1086/519795 PMID: 17701901; PubMed Central PMCID: PMC1950838.

18. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21(2):263–5. Epub 2004/08/07. https://doi.org/10.1093/bioinformatics/bth457 PMID: 15297300.

19. Hiersche M, Ruhle F, Stoll M. Postgwas: advanced GWAS interpretation in R. PLoS One. 2013; 8(8):e71775. Epub 2013/08/27. https://doi.org/10.1371/journal.pone.0071775 PMID: 23977141; PubMed Central PMCID: PMC3747239.

20. Boehnke M, Langefeld CD. Genetic association mapping based on discordant sib pairs: the discordant-alleles test. Am J Hum Genet. 1998; 62(4):950–61. Epub 1998/06/13. https://doi.org/10.1086/301787 PMID: 9529345; PubMed Central PMCID: PMC1377027.

21. Tsze DS, Valente JH. Pediatric stroke: a review. Emerg Med Int. 2011; 2011:734506. Epub 2012/01/19. https://doi.org/10.1155/2011/734506 PMID: 22254140; PubMed Central PMCID: PMC3255104.

22. Nowak-Gottl U, Kosch A, Schlegel N, Salem M, Manco-Johnson M. Thromboembolism in children. Curr Opin Hematol. 2002; 9(5):448–53. Epub 2002/08/13. https://doi.org/10.1097/00062752-200209000-00010 PMID: 12172465.

23. Haywood S, Liesner R, Pindora S, Ganesan V. Thrombophilia and first arterial ischaemic stroke: a systematic review. Arch Dis Child. 2005; 90(4):402–5. Epub 2005/03/23. https://doi.org/10.1136/adc.2004.049163 PMID: 15781933; PubMed Central PMCID: PMC1720343.

24. Fu Y, Kong W. Cartilage Oligomeric Matrix Protein: Matricellular and Matricrine Signaling in Cardiovascular Homeostasis and Disease. Curr Vasc Pharmacol. 2017; 15(3):186–96. Epub 2017/02/06. https://doi.org/10.2174/1570161115666170201121232 PMID: 28155614.

25. Liu CJ, Kong W, Xu K, Luan Y, Ilalov K, Sehgal B, et al. ADAMTS-12 associates with and degrades cartilage oligomeric matrix protein. J Biol Chem. 2006; 281(23):15800–8. Epub 2006/04/14. https://doi.org/10.1074/jbc.M513433200 PMID: 16611630; PubMed Central PMCID: PMC1483932.