

ARTICLE

Copy number variation in patients with cervical artery dissection

Caspar Grond-Ginsbach^{*1,10}, Bowang Chen^{2,10}, Rastislav Pjontek^{1,10}, Tina Wiest¹, Yanxiang Jiang¹, Barbara Burwinkel^{3,4}, Sandrine Tchatchou³, Michael Krawczak⁵, Stefan Schreiber⁶, Tobias Brandt⁷, Manja Kloss¹, Marie-Luise Arnold¹, Kari Hemminki², Christoph Lichy¹, Philippe A Lyrer⁸, Ingrid Hausser⁹ and Stefan T Engelter⁸

Cervical artery dissection (CeAD) occurs in healthy young individuals and often entails ischemic stroke. Skin biopsies from most CeAD-patients show minor connective tissue alterations. We search for rare genetic deletions and duplication that may predispose to CeAD. Forty-nine non-traumatic CeAD-patients with electron microscopic (EM) alterations of their dermal connective tissue (EM + patients) and 21 patients with normal connective tissue in skin biopsies (EM – patients) were analyzed. Affymetrix 6.0 microarrays (Affymetrix) from all patients were screened for copy number variants (CNVs). CNVs absent from 403 control subjects and from 2402 published disease-free individuals were considered as CeAD-associated. The genetic content of unidentified CNVs was analyzed by means of the Gene Ontology (GO) Term Mapper to detect associations with biological processes. In 49 EM + patients we identified 13 CeAD-associated CNVs harboring 83 protein-coding genes. In 21 EM – patients we found five CeAD-associated CNVs containing only nine genes (comparison of CNV gene density between the groups: Mann–Whitney $P=0.039$). Patients' CNVs were enriched for genes involved in extracellular matrix organization (COL5A2, COL3A1, SNTA1, $P=0.035$), collagen fibril organization COL5A2, COL3A1, ($P=0.0001$) and possibly for genes involved in transforming growth factor beta (TGF)-beta receptor signaling pathway (COL3A1, DUPS22, $P=0.068$). We conclude that rare genetic variants may contribute to the pathogenesis of CeAD, in particular in patients with a microscopic connective tissue phenotype.

European Journal of Human Genetics (2012) 20, 1295–1299; doi:10.1038/ejhg.2012.82; published online 23 May 2012

Keywords: carotid and vertebral artery dissection; copy number variants (CNVs); connective tissue morphology; cerebral ischemic stroke

INTRODUCTION

Spontaneous cervical artery dissection (CeAD) is a major cause of ischemic stroke in younger adults.^{1,2} The etiology of CeAD is largely unknown. A genetic predisposition seems likely as CeAD is associated with inherited microscopic and submicroscopic connective tissue alterations in skin biopsy samples from about half of the patients.^{3,4} However, the search for mutations in candidate genes in patients with CeAD has yielded disappointing results.⁵

Rare copy number variants (CNVs) have a role in the etiology of tetralogy of Fallot, sporadic epilepsy syndromes or autism.^{6–8} In patients with sporadic aortic aneurysms and dissections rare CNVs were identified that disrupt genes involved in vascular smooth muscle function.⁹ The current study was the first analysis of CNVs in CeAD. The aim of the study was to identify causative genetic deletions and duplications in patients with and without connective tissue alterations.

MATERIALS AND METHODS

Patients and control subjects

The diagnosis of CeAD was based upon at least one of the following widely accepted criteria: (1) intimal flap visible on carotid ultrasound, (2) mural hematoma visible on magnetic resonance imaging or computed tomography or (3) a non-atherosclerotic, tapered, flame-shaped internal carotid artery (ICA) occlusion or a string-like ICA stenosis.² For this study we selected non-traumatic CeAD-patients from the databases of the University hospitals of Basel and Heidelberg who met the following inclusion criteria: (1) written informed consent, (2) Caucasian ethnicity (ie, both parents white Europeans), (3) information about the presence or absence of dermal connective tissue alterations based on criteria published previously^{4,10} and (4) availability of sufficient DNA. Seventy-seven patients met the inclusion criteria. For the current study, patients were classified according to their dermal connective tissue morphology as visualized by electron microscopy (EM + patients; $n=52$ versus EM – patients; $n=25$).

Microarray data sets from 70 (out of 77) CeAD-patients passed quality control (see below) and were included in the final analyses. CeAD-patients

¹Department of Neurology, University of Heidelberg, Heidelberg, Germany; ²Division of Molecular Genetic Epidemiology, German Cancer Research Center, Heidelberg, Germany; ³Helmholtz University Group Molecular Epidemiology, German Cancer Research Center, Heidelberg, Germany; ⁴Department of Gynaecology, University of Heidelberg, Heidelberg, Germany; ⁵Institute of Medical Informatics and Statistics, Christian-Albrechts University of Kiel, Kiel, Germany; ⁶Department of Internal Medicine, University Hospital Schleswig Holstein, Kiel, Germany; ⁷Kliniken-Schmieder, Heidelberg, Germany; ⁸Stroke Unit and Department of Neurology, University Hospital Basel, Basel, Switzerland; ⁹Department of Dermatology and Electron Microscopy Core Facility, University of Heidelberg, Heidelberg, Germany

*Correspondence: Dr C Grond-Ginsbach, Department of Neurology, University of Heidelberg, Im Neuenheimer Feld 400, Heidelberg D-69120, Germany. Tel: +49 6221 568213; Fax: +49 6221 565461; E-mail: Caspar.Grond-Ginsbach@med.uni-heidelberg.de

¹⁰These authors contributed equally to this work

Received 14 September 2011; revised 9 March 2012; accepted 12 April 2012; published online 23 May 2012

included in the final study sample had a mean age of 42.5 ± 9.8 years. Forty-six patients (66%) were male and 21 patients (30%) presented with dissections in other cervical arteries. Two of these 21 patients suffered from additional malformations in other arterial territories (aorta and iliac arteries). Skin biopsy samples from all patients had been investigated by electron microscopy (EM). Forty-nine patients (70%) patients showed alterations of their connective tissue morphology (EM + patients).

Affymetrix 6.0 (Affymetrix, Santa Clara, CA, USA) data sets from 1262 white Northern Germans, collected in the PopGen study,¹¹ were subjected to more stringent quality control criteria (see below), in order to avoid rejection of patients' CNV due to filtering against false positive control findings. 403 of the data sets were selected for the final analyses.

Ethics

The study protocol has been approved by the local ethical committees in Heidelberg and Basel.

Microarray analysis

DNA was extracted from 3 ml venous blood after SDS-proteinase-K digest, phenol/chloroform extraction and ethanol precipitation following standard procedures. The DNA was dissolved in distilled water and quantified by UV spectrophotometry. All DNA samples were extracted from native peripheral blood samples (lymphoblastoid cell lines or fibroblast cell cultures were not used).

Two hundred and fifty nanograms of DNA was hybridized on Affymetrix GeneChip Human Mapping SNP6.0 arrays (Affymetrix), using Human Mapping SNP6.0 assay kit and following the manufacturer's instructions. The microarrays were washed, stained by streptavidin-phycoerythrin conjugates, and scanned using Affymetrix GeneChip Command Console Software (Affymetrix). Probe cell intensity data, sample and array registration, data management, instrument control, as well as automatic and manual image gridding were performed according to the manufacturer's instructions.

Analysis of CNVs

CNVs were identified with three software packages: Affymetrix Power Tools (APT), Birdsuite version 1.5.5 and PennCNV (2010May01 version).^{12,13}

Quality filtering of the data sets. In a first exploration we analyzed the data sets from all patients and from 501 control subjects (Supplementary material 1), and identified for each patient all markers with copy number (CN) state $\neq 2$ but with CN-state = 2 in all the analyzed controls. Moreover, we visualized the signal intensities (LRR, log R ratios) and the BAF (B-allele frequency)-values for each individual patient and control subject to appreciate the noise in the data sets. On the basis of this preliminary exploration, we identified seven patients with outlier numbers of markers with CN-state $\neq 2$, as well as with prominent genomic waves.¹⁴ In a second step we filtered the control cohort in a more stringent way and selected four hundred three data sets of high quality (genotyping call rates $\geq 98.0\%$ and the absence of strong genomic waves) from the 1262 PopGen subjects.

CNVs were identified with both PennCNV and Birdsuite. Subsequently, CNV calls containing < 20 markers in PennCNV and calls with a confidence below or equal 10 in Birdsuite were excluded from the subsequent analysis. CNVs found by both software packages were visualized individually before further analysis. Data sets were visualized with a newly designed software (<http://noise-free-cnv.sourceforge.net/index.php>) that also permits visual comparison of several data sets and easy correction of genomic waves. Microarrays from 11 patients had outlier number (< 10 or > 250) of CNV calls with PennCNV. After visual inspection of the data sets, seven patients were not analyzed further (Supplementary Material 1).

CeAD-associated CNVs. Patient CNV data were first compared with CNVs from control individuals to identify rare CNVs. In the second step, we compared the rare CNVs detected in our patient samples with published CNVs in 2402 Caucasian healthy subjects (HapMap project, $n = 90$; CHOP database, $n = 1327$; and a recent CNV population study, $n = 985$).^{15–17} Rare CNVs that were detected among CeAD-patients but were absent from own control

subjects, and from the aforementioned data sets were considered as CeAD-associated.

Molecular characterization of CNVs

CNV and their breakpoints were mapped on genome assembly NCBI36. For the validation of the CNVs by independent molecular methods, we established CNV-specific PCRs with primers located near the presumptive breakpoints, as described by Vissers *et al.*¹⁸ Primers located in the 5' and 3' flanking sequences outside of a large deletion allow proof of existence of the deletion and identification of the deletion breakpoints (Supplementary Material 2). Duplications were validated with primers at the 3' (forward) and 5' (reverse) ends within the duplication in order to amplify the joining of the 3' end and the 5' end of two tandem copies. A positive PCR therefore confirms both the existence of the duplication and its tandem arrangement. The specificity of each CNV-specific PCR was confirmed by the absence of amplification in 10 control samples, randomly selected from 204 age- and sex-matched healthy German subjects, recruited for earlier genetic studies.¹⁹

Gene ontology (GO) analysis

The GO Term Finder software (<http://go.princeton.edu/cgi-bin/GOTermMapper>) was used to search for genetic enrichment among GO process terms. We analyzed the following GO terms: apoptosis (GO:0006915), blood vessel development (GO:0001568), collagen fibril organization (GO:0030199), immune response (GO:0006955), inflammatory response (GO:0006954), proteoglycan biosynthetic process (GO:0030166), regulation of blood pressure (GO:0008217), smooth muscle cell proliferation (GO:0048659) and transforming growth factor beta (TGF-beta) receptor signaling pathway (GO:0007179). We selected these GO process terms as medial degeneration, the putative cause of arterial dissections, is characterized by smooth muscle proliferation, apoptosis, connective tissue remodeling and proteoglycan accumulation.^{20,21} Disruption of TGF-beta receptor signaling, inflammation and connective tissue disorders are known risk factors for arterial dissections.^{3,22,23}

RESULTS

Data sets from 70 (out of 77) CeAD-patients and 403 (out of 1262) control subjects were included in the final analysis.

CNV analysis identified 65 CNVs in the patient group that were not found in the 403 control subjects (Supplementary Material 3). Thirty-four of these CNVs were absent in 2402 Caucasian controls^{16,17} and were considered as CeAD-associated. Eighteen of the CeAD-associated CNVs contained coding sequences (Table 1). Thirteen CeAD-associated CNVs in the EM + patients harbored 83 protein-coding genes, whereas five CeAD-associated CNVs in the EM – group contained only 9 protein-coding genes, indicating a higher density of coding genes in the CNVs of the EM + patients (Mann–Whitney *U*-test: $P = 0.039$). The sample of 403 population controls contained 87 CNVs that were absent from 2402 published controls. Fifty-five of these rare CNVs from the population controls contained 147 protein-coding genes (Supplementary Material 3). The number of genes in the CeAD-associated CNVs of the EM + patients was higher than in the control population (Mann–Whitney *U*-test; $P = 0.012$). However, there was no difference in gene density between the (few) CeAD-associated CNVs of the EM – patients and the CNVs of healthy controls (Mann–Whitney *U*-test; $P = 0.28$).

Ten CeAD-associated CNVs containing coding sequences (5 deletions and 5 duplications) were validated by detailed molecular analysis (Supplementary Material 2). Several oligonucleotide primers close to the indicated CNV borders were tested by trial and error to amplify the joining segment of each of these ten CNVs. For 9 out of 10 (ie, 5 deletions and 4 duplications) a CNV-specific PCR was successfully established. Subsequent DNA sequence analysis of the amplified joining segment permitted the precise identification of the

Table 1 CeAD-specific CNVs containing coding sequences detected in this study

Position (chr: start-end)	Size (kb)	CN-state	DGV	No. of genes	Gene names
<i>CeAD-associated CNVs detected in 49 patients with aberrant dermal connective tissue morphology</i>					
18:59640388–59694035	54	Loss	Yes	1	SERPINB2
4:144570037–144634141	64	Gain	Yes	1	GAB1
12:91609585–91681005	71	Loss	No	2	C12orf74, PLEKHG7
9:107430028–107517525	87	Loss	Yes	3	FKTN, TAL2, TMEM38B
6:161612276–161734297	122	Gain	Yes	2	AGPAT4, PARK2
3:113368966–113538579	170	Loss	Yes	2	SLC9A10, CD200
19:52912535–53094035	182	Gain	Yes	7	GLTSCR1, EHD2, GLTSCR2, SEPW1, TPRX1, CRX, SULT2A1
6:183515–671736	488	Gain	Yes	5	EXOC2, IRF4, DUSP22, HUS1B, AL031770
1:202619544–203219581	600	Gain	Yes	5	PPP1R15B, PIK3C2B, MDM4, LRRN2, NFASC
2:189109859–189763802	654	Loss	No	4	GULP1, DIRC1, COL3A1, COL5A2
19:1971710–2632482	661	Gain	Yes	24	MKNK2, C19orf36, MOBKL2A, AP3D1, DOT1L, IZUMO4, AC004410, PLEKHJ1, C19orf35, SF3A2, AMH, JSRP1, OAZ1, LINGO3, AC104537.1, LSM7, TIMM1, TMPRSS9, LMNB2, GADD45B, GNG7, DIRAS1, SLC39A3, SGTA
20:31395708–32782473	1387	Loss	Yes	19	CDK5RAP1, SNTA1, CBFA2T2, NECAB3, C20orf144, PXMP4, C20orf134, E2F1, ZNF341, RALY, EIF2S2, ASIP, AHCY, ITCH, DYNLRB1, MAP1LC3A, PIGU, TP53INP2, NCOA6
4:141190354–144311522	3121	Gain	No	8	SCOC, CLGN, ELMOD2, TBC1D9, RNF150, ZNF330, IL15, INPP4B
<i>CeAD-associated CNVs detected in 21 patients with normal dermal connective tissue morphology</i>					
8:14006431–14131717	125	Loss	Yes	1	SGCZ
7:132844963–132988175	143	Loss	Yes	1	EXOC4
10:68972491–69137046	165	Gain	Yes	1	CTNNA3
16:11935326–12115916	181	Gain	Yes	5	RP11-166B2.1, TNFRSF17, RUNDC2A, SNX29, AC00760.1
2:133324676–133563950	239	Loss	Yes	1	NCKAP5

Abbreviations: Chr, chromosome; CN-state, copy number state; Chr, chromosome; Kb, kilobasepairs. DGV = database of genomic variants (<http://projects.tcag.ca/variation/>). DGV = yes, if overlapping CNV in DGV contains coding exons. DGV = no, if DGV does not contain any CNVs overlapping with coding sequences of any gene in a rare patient CNV. The position of the CNV indicates chromosome and first and last nucleotide position mapped on genome assembly NCBI36.

breakpoints in these 9 validated CNVs (Supplementary Material 3). The validation of one CNV by CNV-specific PCR was unsuccessful, but the presence of a duplication in the region was confirmed by quantitative PCR. Because of substantial differences in an estimated size of this CNV according to different CNV calling algorithms (Supplementary Table 4) we did not include it in subsequent analyses.

GO Term Mapper analysis was performed to search for enrichment of genes involved in specific biological processes. We considered 10 different biological processes, which were assumed to have a role in the etiology of CeAD. The genetic content of the rare CNVs found in the control population was not enriched for any of these candidate processes, but the CeAD-specific CNVs were significantly enriched for genes involved in extracellular matrix organization ($P=0.035$), collagen fibril organization ($P=0.0001$) and showed a tendency ($P=0.068$) toward enrichment for genes involved in TGF-beta receptor signaling pathway (Table 2). The deletion of the COL3A1/COL5A2 locus in patient 442 was associated with a characteristic electron microscopic (EM) phenotype (Supplementary Material 4), resembling findings in patients with vascular Ehlers–Danlos syndrome. Several other CeAD-associated CNVs contained interesting candidate genes involved in additional pathways, like interleukin 15 (IL15), zeta-sarcoglycan (SGCZ) or plasminogen activator inhibitor-2 (SERPINB2). Owing to the small size of the current study, however, these findings had to remain anecdotic.

DISCUSSION

In this genome-wide search for small chromosome aberrations we analyzed 70 CeAD-patients with a well-characterized connective tissue

phenotype. Rare CNVs of EM+ patients harbored more coding genes than those of EM– patients or of population controls. The genetic content of the CeAD-associated CNVs was enriched for genes involved in collagen fibril organization and perhaps for genes involved in TGF-beta receptor signaling. These findings suggest that rare mutations affecting different biological processes may increase the risk for CeAD.

Previous genetic findings from patients with cervical and aortic dissections indicated that the genetic background of arterial dissection is heterogeneous. Mutations were identified in a variety of genes (COL3A1, COL1A1, COL5A2, TGFBR2, TGFBR1, ACTA2) in a minority of the patients, whereas the search for mutation remained without positive results in the majority of analyzed CeAD-patients.²⁴ Mutations in FBN1, COL3A1, MYH11, ACTA2, TGFBR1 and TGFBR2 were found in patients with aortic dissection, particularly in familial cases, but no mutations were detected in the majority of sporadic patients.²⁴ A recent study identified various rare CNVs in patients with aortic dissections, disrupting different genetic processes, in particular affecting vascular smooth muscle adhesion and contractility.⁹ A large deletion encompassing the COL3A1/COL5A2 locus was identified in another patient with aortic dissection.²⁵ Our current findings indicate that the genetic disposition to CeAD might be heterogeneous: none of the rare CNVs was found in more than one patient. Moreover, the identified deletions and duplications disrupt different genes, belonging to different GO terms. Instead of being associated with a single genetic process the genetic predisposition to CeAD is probably highly complex.

Table 2 Association of CeAD-specific CNVs with predefined candidate GO terms

Gene ontology term	Patient cluster frequency	Control cluster frequency	genome Frequency	P-Value
	<i>n</i> = 69 genes	<i>n</i> = 91 genes	<i>n</i> = 27 358 genes	
Apoptosis (GO:0006915)	9 (13.0%)	5 (5.5%)	1965 (7.2%)	0.10
Blood vessel development (GO:0001568)	1 (1.5%)	0 (0%)	234 (0.9%)	0.59
Collagen fibril organization (GO:0030199)	2 (2.9%) (COL3A1, COL5A2)	0 (0%)	34 (0.1%)	0.0001
Extracellular structure organization (GO:0043062)	3 (4.4%) (SNTA1, COL3A1, COL5A2)	0 (0%)	283 (1.3%)	0.035
Immune response (GO:0006955)	5 (7.3%)	4 (4.4%)	1727 (6.3%)	0.94
Inflammatory response (GO:0006954)	2 (2.9%)	0 (0%)	522 (1.9%)	0.87
Proteoglycan biosynthetic process (GO:0030166)	0 (0%)	0 (0%)	32 (0.1%)	0.78
Regulation of blood pressure (GO:0008217)	0 (0%)	0 (0%)	164 (0.6%)	0.52
Smooth muscle cell proliferation (GO:0048659)	1 (1.5%)	0 (0%)	100 (0.4%)	0.62
TGF-beta receptor signaling pathway (GO:0007179)	2 (2.9%) (COL3A1, DUSP22)	0 (0%)	149 (0.5%)	0.068

Abbreviation: TGF, transforming growth factor.

Among the 83 genes found in the rare coding CNVs of the patients, 69 genes were annotated in GO terms and analyzed. Analysis of the healthy controls suggested specificity of the observed associations. *P*-values were calculated for Chi-square comparison of patient cluster frequency with genome frequency. Patient cluster frequency = frequency of genes belonging to the GO term in the patient-specific CNVs. Control cluster frequency = frequency of genes belonging to the GO term in the rare CNVs from the control group.

For the present study we selected CeAD-patients with known dermal connective tissue phenotype. As the analysis of the dermal connective tissue requires taking a deep knife skin biopsy, the number of analyzed patients was small. Moreover, the sample was slightly biased toward more severe cases and contained more patients with multiple dissection compared with other study samples. The small size of the study samples and its possible bias toward complex cases are the main potential flaws of our study. The sample of EM – patients (*n* = 21) was particularly small, and the absence of any association with GO process terms might be related to low power in this subset of patients. In this group we detected a rare duplication of the MYH11/ABCC6 locus (Supplementary Table 2), which was considered to be disease-related in a study of aortic dissection,⁹ as well as a CeAD deletion of the gene encoding SGCZ leading to a frame shift in the coding transcript. It is not unlikely that these CNVs affect the contractility of the smooth muscle cells and predispose to CeAD. Because of the small size of our study sample we did not perform an open GO Term Finder search, but restricted the GO search to predefined biological processes. These GO terms were chosen as candidates that may be involved in the development of medial degeneration,^{20,21} that were assumed to be associated with the risk for CeAD, or were involved in the development of the arterial system.

Strength of our study was the precise phenotype of the patients. (1) The diagnosis of CeAD was based upon widely accepted criteria.^{1,2} (2) A traumatic cause of the CeAD as confounding variable was excluded by careful anamnesis in all patients. (3) A skin biopsy from each patient was studied by light and electron microscopy in order to detect connective tissue alterations according to published criteria.^{4,10} Microarray data sets were analyzed with different CNV calling algorithms. The Birdsuite algorithm yielded significantly more CNVs calls than PennCNV, with a higher proportion of Birdsuite calls being false positives, as was also found by others.^{26,27} Further strength of our study was the interactive evaluation of all calls by visual inspection of each data set, the stringent selection of control data sets (to avoid filtering of patients' CNVs against false positive control findings) and the careful validation of the CNV calls by breakpoint identification. The genetic analysis of the CNVs using GO Term Finder was helpful for the identification of significant results.

We restricted our analysis to those CNVs that were absent from control individuals and that contained protein-coding genes. This filtering strategy might have been too stringent. Non-coding transcripts may have a role in vascular diseases, as was recently

demonstrated for the ANRIL transcript.²⁸ A rare large duplication of the MYH11/ABCC6 locus that was also found in patients with aortic dissections was not analyzed in our study as this CNV has a frequency of about 1/1000 in the normal population.¹¹ However, such a low frequency does not rule out that this CNV is associated with an increased risk for CeAD. We filtered our findings against CNVs detected by high-density platforms (> 500 000 SNPs) in Caucasian individuals.^{15–17} Many patients' CNVs that were absent from all analyzed controls showed some overlap with rare CNVs listed in the database of genomic variants (DGV). However, the value of DGV for CNV interpretation is limited, as DGV contains non-validated CNVs and many CNVs with overestimated lengths.^{29–31} Although some of the identified rare coding CNVs in our patients might not be truly specific, we consider all identified coding CNVs as rare.

This study revealed that rare structural mutations may contribute to an increased risk for CeAD in some patients. The risk for CeAD is probably not related to a single gene or a single genetic pathway, but might be associated with a variety of different genetic variants. Rare CNVs might be one class of genetic variation associated with CeAD, in particular in patients with a connective tissue phenotype. A follow-up study of rare CNVs in a large series of CeAD-patients and in a disease control group of non-CeAD ischemic stroke patients is in preparation to confirm and extend the current findings.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

The authors are indebted to Inge Werner for excellent technical assistance and Philip Ginsbach for analyzing CNV signal strength and B-allele frequency data and developing supporting software (<http://noise-free-cnv.sourceforge.net/index.php>) and to Werner Hacke for continuous support and excellent working facilities. This work was supported by grants from the Swiss Heart Foundation (to STE and PAL), the Schlieben-Lange fellowship (to TW), and the German Academic Exchange Service DAAD (to RP). The PopGen project received funding from the German Ministry of Education and Research (BMBF) through the National Genome Research Network (NGFN).

1 Schievink WI: Spontaneous dissection of the carotid and vertebral arteries. *N Engl J Med* 2001; **344**: 898–906.

2 DeBette S, Leys D: Cervical-artery dissections: predisposing factors, diagnosis, and outcome. *Lancet Neurol* 2009; **8**: 668–678.

- 3 Grond-Ginsbach C, Debette S: The association of connective tissue disorders with cervical artery dissection. *Curr Mol Med* 2009; **9**: 210–214.
- 4 Brandt T, Orberk E, Weber R *et al*: Pathogenesis of cervical artery dissections: Association with connective tissue abnormalities. *Neurology* 2001; **57**: 24–30.
- 5 Debette S, Markus H: The genetics of cervical artery dissection: a systematic review. *Stroke* 2009; **40**: e459–e466.
- 6 Greenway SC, Pereira AC, Lin JC *et al*: *De novo* copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet* 2009; **41**: 931–935.
- 7 Heinzen EL, Radtke RA, Urban TJ *et al*: Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. *Am J Hum Genet* 2010; **86**: 707–718.
- 8 Bucan M, Abrahams BS, Wang K *et al*: Genome-wide analyses of exonic copy number variants in a family-based study point to novel autism susceptibility genes. *PLoS Genet* 2009; **5**: e1000536.
- 9 Prakash SK, LeMaire SA, Guo DC *et al*: Rare copy number variants disrupt genes regulating vascular smooth muscle cell adhesion and contractility in sporadic thoracic aortic aneurysms and dissections. *Am J Hum Genet* 2010; **87**: 743–756.
- 10 Brandt T, Hausser I, Orberk E *et al*: Ultrastructural connective tissue abnormalities in patients with spontaneous cervicocerebral artery dissections. *Ann Neurol* 1998; **44**: 281–285.
- 11 Krawczak M, Nikolaus S, von Eberstein H, Croucher PJ, El Mokhtari NE, Schreiber S: PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet* 2006; **9**: 55–61.
- 12 Korn JM, Kuruvilla FG, McCarroll SA *et al*: Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet* 2008; **40**: 1253–1260.
- 13 Wang K, Li M, Hadley D, Liu R *et al*: PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res* 2007; **17**: 1665–1674.
- 14 Diskin SJ, Li M, Hou C *et al*: Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Res* 2008; **36**: e126.
- 15 McCarroll SA, Kuruvilla FG, Korn JM *et al*: Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* 2008; **40**: 1166–1174.
- 16 Shaikh TH, Gai X, Perin JC *et al*: High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res* 2009; **19**: 1682–1690.
- 17 Li J, Yang T, Wang L, Yan H *et al*: Whole genome distribution and ethnic differentiation of copy number variation in Caucasian and Asian Populations. *PLoS One* 2009; **4**: e7958.
- 18 Vissers LE, Bhatt SS, Janssen IM *et al*: Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture. *Hum Mol Genet* 2009; **18**: 3579–3593.
- 19 Longoni M, Grond-Ginsbach C, Grau AJ *et al*: The ICAM-1 E469K gene polymorphism is a risk factor for spontaneous cervical artery dissection. *Neurology* 2006; **66**: 1273–1275.
- 20 Ihling C, Szombathy T, Nampoothiri K *et al*: Cystic medial degeneration of the aorta is associated with p53 accumulation, Bax upregulation, apoptotic cell death, and cell proliferation. *Heart* 1999; **82**: 286–293.
- 21 He R, Guo DC, Estrera AL *et al*: Characterization of the inflammatory and apoptotic cells in the aortas of patients with ascending thoracic aortic aneurysms and dissections. *J Thorac Cardiovasc Surg* 2006; **131**: 671–678.
- 22 Tran-Fadulu V, Pannu H, Kim DH *et al*: Analysis of multigenerational families with thoracic aortic aneurysms and dissections due to TGFBR1 or TGFBR2 mutations. *J Med Genet* 2009; **46**: 607–613.
- 23 Pfefferkorn T, Saam T, Rominger A *et al*: Vessel wall inflammation in spontaneous cervical artery dissection: a prospective, observational positron emission tomography, computed tomography, and magnetic resonance imaging study. *Stroke* 2011; **42**: 1563–1568.
- 24 Grond-Ginsbach C, Pjontek R, Aksay SS, Hyhlik-Dürr A, Böckler D, Gross M-L: Spontaneous arterial dissection – phenotype and molecular pathogenesis. *Cell Mol Life Sci* 2010; **67**: 1799–1815.
- 25 Meienberg J, Rohrbach M, Neuenschwander S *et al*: Hemizygous deletion of COL3A1, COL5A2, and MSTN causes a complex phenotype with aortic dissection: a lesson for and from true haploinsufficiency. *Eur J Hum Genet* 2010; **18**: 1315–1321.
- 26 Winchester L, Yau C, Ragoussis J: Comparing CNV detection methods for SNP arrays. *Brief Funct Genomic Proteomic* 2009; **8**: 353–366.
- 27 Kidd JM, Cooper GM, Donahue WF *et al*: Mapping and sequencing of structural variation from eight human genomes. *Nature* 2008; **453**: 56–64.
- 28 Popov N, Gil J: Epigenetic regulation of the INK4b-ARF-INK4a locus: in sickness and in health. *Epigenetics* 2010; **5**: 685–690.
- 29 Zhang F, Gu W, Hurles ME, Lupski JR: Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet* 2009; **10**: 451–481.
- 30 Duclos A, Charbonnier F, Chambon P *et al*: Pitfalls in the use of DGV for CNV interpretation. *Am J Med Genet Part A* 2011; **155**: 2593–2596.
- 31 Perry GH, Ben-Dor A, Tsalenko A, Sampas N, Rodriguez-Revilla L: The fine-scale and complex architecture of human copy-number variation. *Am J Hum Genet* 2008; **82**: 685–695.

Supplementary Information accompanies the paper on European Journal of Human Genetics website (<http://www.nature.com/ejhg>)