Published in final edited form as:

Stroke. 2018 October ; 49(10): 2549–2554. doi:10.1161/STROKEAHA.118.020371.

Copy Number Variation and Risk of Stroke

Caspar Grond-Ginsbach, PhD[1], **Philip Erhart, MD**[2], **Bowang Chen, PhD**[3], **Manja Kloss, MD**[1], **Stefan T. Engelter, MD, PhD**[4,5], and **John W. Cole, MD**[6]

[1.] Department of Neurology, University Hospital Heidelberg, Heidelberg, Germany [2.] Department of Vascular and Endovascular Surgery, University Hospital Heidelberg, Heidelberg, Germany [3.] Department of Biology, Southern University of Science and Technology, Shenzhen, China [4.] Neurorehabilitation Unit, University of Basel and University Center for Medicine of Aging, Felix Platter Hospital, Basel, Switzerland [5.] Department of Neurology and Stroke Center, University Hospital Basel, Basel, Switzerland [6.] Department of Neurology, Veterans Affairs Medical Center and University of Maryland School of Medicine, Baltimore, USA

Keywords

genomics; stroke; cervical artery dissection; SNP-microarray; incidental findings

Copy number variation (CNV) describes both genomic deletions, defined as “loss” of genetic material, and genomic duplications, defined as a “gain” of an additional copy of an existing DNA sequence. As illustrated in Figure 1, CNV can range in size from 50 base pairs (bp) up to several megabases (Mb) or even entire chromosomes, in contrast to a single nucleotide polymorphism (SNP) altering only a single nucleotide base. CNV occurs ubiquitously throughout the genome and constitutes an important part of the genetic diversity in the human population with increasingly recognized clinical impact.[1] Presently, the identification of CNV is limited by several factors including DNA-quality, the data

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

Conflict of Interest Disclosures

None.

Social media handles for *Stroke* posts:

University of Heidelberg Medical School:
Facebook: www.klinikum.uni-heidelberg.de/facebook
Twitter: https://twitter.com/uniklinik_hd
University Hospital Basel:
Facebook: https://www.facebook.com/unispitalbasel
Twitter: https://twitter.com/unispitalbasel
Felix-Platter Hospital Basel:
Facebook: https://www.facebook.com/felixplatterspital/
Twitter: https://twitter.com/hashtag/felixplatterspital
University of Maryland Medical Center:
Facebook: @UMDMedCenter https://www.facebook.com/UMDMedCenter
Twitter: @UMMC https://twitter.com/UMMC
University of Maryland School of Medicine:
Facebook: @Maryland.Medicine https://www.facebook.com/Maryland.Medicine?sk=wall
Twitter: @UMmedschool https://twitter.com/ummedschool

generation platform, and computational analysis. As a consequence, the study of CNV in the clinical setting lags far behind the analysis of SNPs.

CNV-detection and interpretation

Different data generation platforms (SNP-microarrays, chromosome microarrays (also known as comparative genomic hybridization or CGH) and next generation sequencing) allow genome-wide detection of CNV.[1] The current review will focus on SNP-microarray-based CNV analysis, since large collections of stroke patients have previously been genotyped using SNP-microarrays.[2] From an efficiency standpoint, these samples can be directly re-utilized for CNV-analysis without extra material costs. However, the use of SNP arrays for CNV detection has some limitations:

1. The resolution of SNP-microarray analyses (i.e. the minimal size of the variants that can be reliably detected) is inversely related to SNP-density of the platform. Hence, low SNP-density arrays are less useful. Moreover, as shown in Figure 2, the distribution of annotated SNPs across the genome (Figure 2A) and on a typical microarray (Figure 2B) is highly non-random. The SNP density is particularly low in genomic regions with segmental duplications, as well as around the centromere (Figure 2D). These regions are prone to CNV, as indicated in Figure 2C. As a consequence, the resolution of SNP-microarray studies varies across the genome and is particularly low in regions with high frequency of CNV.
2. Noise in microarray datasets is an important source of false-positive CNV-findings. Some noise components can be adjusted for by pair-wise comparison of samples or by more sophisticated identification of independent noise components. [3,4]

The clinical interpretation of a specific CNV-finding may be difficult or uncertain.[5] In general, very large (> 500 kilobase) and rare (<1%) CNVs are more likely to be disease-associated than small and common ones, but size alone is not crucial. Large CNVs can be benign while small ones can be clinically important.[5,6] Rather than the physical length, the total number of genes within the CNV, as well as the function of the affected genes (protein-coding or non-coding, coding for dosage-sensitive or dosage-insensitive proteins) are likely to determine the clinical importance.[5,7] Currently, the clinical interpretation of many CNV-findings remains unclear.[7,8] Moreover, in large CNVs covering many different genes, it may be difficult to pinpoint the specific disease-causing gene.

Anticipating the disclosure of incidental (“unsolicited”) pathogenic findings is an ongoing challenge for all genome-wide diagnostic methods, including CNV detection. For example, well-established pathogenic variants may be found in genes not related to the phenotype of interest. Genetic counseling of patients prior to CNV analysis, as usually offered to patients in a clinical context, could anticipate and assist with incidental findings as well as unclear findings (variants of unknown significance).[9] However, patient-data in large epidemiologic studies are usually anonymized –with neither the scientists nor participants well prepared to deal with incidental findings.

CNV in stroke patients

As listed below and as demonstrated in Table 1, ischemic stroke has been associated with several different types of CNV-findings:

1. *Common risk-variants*: The lipoprotein (a) gene (*LPA*) contains a repeated domain of 114 amino-acids that occurs in a highly variable number of copies (1 to >40). Individuals with lower copy numbers (<22 repeats) have an approximately two-fold higher risk of ischemic stroke than those with larger isoforms.[25] Multiplex ligation-dependent probe amplification (MLPA) is the gold-standard to analyze this CNV.

2. *Rare disease-causing variants*: A large (749,000 bp) duplication encompassing six protein-coding genes (*COL4A1, COL4A2, RAB20, NAXD, CARS2, ING1*) genes as well as several non-protein-coding genes was found in a young patient with recurrent lacunar infarcts due to small vessel disease and in eleven affected family members.[16] The CNV was identified during next-generation sequencing analysis and was confirmed by array comparative genome hybridization.

3. *Global genomic imbalance*: An excess burden of large, gene-disrupting CNVs was found in stroke patients with unfavorable functional outcome after three months, compared to patients with favorable outcome.[26] SNP-microarrays from previous GWAS were re-utilized to study CNV.

4. *Variants of unknown significance (VUS)*: a large (> 3.1 Mb) duplication encompassing eight protein-coding genes (*SCOC, CLGN, ELMOD2, TBC1D9, RNF150, ZNF330, IL15, INPP4B*) was detected in a 19-year-old boy with ischemic stroke due to spontaneous carotid artery dissection.[11] This finding should be conservatively considered as variant of unknown significance (VUS), since the variant was novel and sporadic (i.e. non-familial) and since none of the duplicated genes are known candidate genes for cervical artery dissection.

Stroke patients with chromosome aberrations

Chromosome aberrations that are large enough to be analyzed by microscopy are rare in patients with non-syndromic cardiovascular diseases or stroke.[27,28] An abnormal chromosome 13 was detected by light microscopic analysis of lymphocyte metaphase chromosomes in a young patient with recurrent stroke. Parallel SNP-microarray analysis revealed a complex rearrangement with multiple duplications.[29] Some studies associated stroke with numerical aberrations including trisomy 21,[30] Klinefelter syndrome[31] and Turner syndrome.[32] Somatic loss of the Y-chromosome is a common acquired aberration in male blood cells associated with age and smoking.[33] An association between loss of the Y-chromosome and cardiovascular outcomes is suspected,[34] but rigorous analysis of mosaicism in cardiovascular patients is lacking.

CNV and ischemic stroke

Common CNV in several candidate genes was tested for association with ischemic stroke. A lower copy number of the *DEFB4* gene was associated with ischemic stroke in a single

study.[18] A meta-analysis of four CNV studies did not replicate prior CNV associations of *GSTM1* and *GSTT1* with stroke.[35] Positive associations with ischemic stroke were reported with common CNVs in *LPA* and *LDLR*.[36]

As early as 2008, the impact of CNV was explored in 263 ischemic stroke patients and 275 control subjects by analyzing microarray data from the first GWAS in ischemic stroke.[37] A total of 408,000 SNPs were genotyped in each study subject, resulting in a resolution of CNV-detection of about 50,000 bp. In the stroke cohort, 231 CNVs were identified, widely distributed throughout the genome, with sizes up to 2.1 Mb. All reported variants were low-frequency findings. None of the observed variants were unequivocally linked to ischemic stroke. In the stroke cohort, ischemic strokes were classified according to TOAST criteria,[38] but subtype analysis was not performed as the numbers were small and power was deemed insufficient.

CNV and cervical artery dissection (CeAD)

Dissection of the carotid or vertebral artery is a major cause of ischemic stroke in patients younger than 50 years. Dissection can occur spontaneously in young adults without known vascular risk factors, suggesting an underlying structural defect of the arterial wall, which was subsequently confirmed by electron-microscopic investigation of arterial biopsies.[39] Genetic analysis revealed rare point-mutations in different candidate genes associated with inherited connective tissue diseases, but this was only in a minority of the patients.[40,41]

In an early CNV study, 70 CeAD patients were phenotyped by an electron-microscopic analysis of a skin biopsy in order to detect connective tissue alterations.[11] One patient with carotid artery dissection and a history of aortic disease had a large deletion covering the entirety of the *COL3A1* and *COL5A2* genes.[11] Another patient carried a large recurrent duplication of chromosome 16p13 including the *MYH11* and *ABCC6* genes, a rare finding in the normal population that predisposes to aortic aneurysm and dissection.[42,43] Four further patients with CNV of the *MYH11/ABCC6* locus were identified in a subsequent exploration of 833 CeAD patients.[12] This latter CNV-study of CeAD did not detect association with variation in a particular locus, but found association with variation in a predefined set of genes involved in cardiovascular system development.

CNV and hemorrhagic stroke

CNV was studied in 23 Korean patients with ruptured intracranial aneurysms by comparative genomic hybridization (CGH) with 4,030 BAC (bacterial artificial chromosome)-clones covering the entire human genome at a resolution of 1 Mbp. Each patient was analyzed separately and compared with a pooled reference DNA sample from 10 gender-matched healthy subjects, but no definitive risk CNVs were detected.[44] The CGH system used in this study was designed for the detection of very large (>1Mb) variants, which are usually rare. Further, for the study of rare variants, the sample size was notably small.

In another study, high-density (300K Illumina) SNP-microarrays from a GWAS of 191 Japanese patients with aneurysmal subarachnoid hemorrhage and 282 controls were used for CNV exploration.[45] CNV-findings were carefully validated by visual examination of the

genoplot images and overlapping analysis with the Database of Genomic Variants (DGV - http://dgv.tcag.ca/dgv/app/home). Moreover, selected findings were validated by quantitative PCR. Most CNV findings were distributed evenly across the chromosomes, but common variants in two regions (chr4:153210505–153212191 and chr10:6265006–6267388) were significantly associated with the risk of SAH. These findings are pending replication in independent studies. No subsequent genome-wide CNV-studies of intracranial aneurysm have been published. Systematic explorations of the impact of CNV on intracerebral hemorrhage have yet to be published.

CNV and stroke pharmacokinetics

CNV occurs in many genes associated with drug absorption, distribution, metabolism and excretion, but until recently the influence of CNV on drug response was not well recognized.[46,47] Currently, the impact of CNV on drug response is increasingly perceived as a potential driver of drug efficacy. This may lead in the near future to more precise pharmacological targeting, including stroke-specific medications.

Outlook

Worldwide, stroke researchers have collected large numbers well-characterized stroke patients and genotyped them on high-density SNP-microarrays for genome-wide association studies of common SNPs.[2] These microarrays could be re-analyzed to detect CNV without the need for extra material costs. Unfortunately, several obstacles may prevent scientists from initiating such large-scale CNV studies using these data, including: 1) the high proportion of false positive CNV-calls when using current CNV-detection algorithms; 2) the huge work-load of pair-wise comparing each sample with a referent sample; 3) the uncertain clinical interpretation of many CNV-findings, and; 4) the lack of a universally accepted reference set of CNV-findings (size, frequency) across ethnically-diverse human populations. Moreover, a significant fraction of the large CNV-findings seen in SNP-microarray based CNV studies is expected to be rare (i.e. population frequency <0.01), which may have power implications regarding associations with specific phenotypes.[48,49]

In closing, the concept of CNV describes a large and highly heterogeneous set of genomic variants, including rare vs. common, benign vs. pathogenic, and inherited vs. de-novo. While current genetic epidemiological methods for the analysis of common variants with small relative risks are well-suited for large patient populations, optimal analytical methodologies for rare disease-causing mutations such as CNV remain uncertain. Rather than mere risk factors, pathogenic CNVs may also be considered as “inborn errors”. Thus, the study of rare pathogenic CNVs may lead stroke genetics back to a more individual or patient-centered clinical focus.

Acknowledgments

Sources of funding

Dr. Cole’s efforts on this project were supported in part by NIH grants U01 NS069208, R01 NS100178, and R01 NS105150; the U.S. Department of Veterans Affairs, and the American Heart Association Cardiovascular Genome-Phenome Study (Grant# 15GPSPG23770000), and an American Heart Association Discovery Grant supported by Bayer Group (Grant# 17IBDG33700328).

Abbreviation of gene names

ABCC6	ATP Binding Cassette Subfamily C Member 6
AMY1	Amylase, Alpha 1A (Salivary)
CARS2	Cysteinyl-TRNA Synthetase 2, Mitochondrial
CLGN	Calmegin
COL3A1	Collagen Type III Alpha 1 Chain
COL4A1	Collagen Type IV Alpha 1 Chain
COL4A2	Collagen Type IV Alpha 2 Chain
COL5A2	Collagen Type V Alpha 2 Chain
DEFB4	Defensin Beta 4A
ELMOD2	ELMO Domain Containing 2
GSTM1	Glutathione S-Transferase Mu 1
GSTT1	Glutathione S-Transferase Theta 1
IL15	Interleukin 15
ING1	Inhibitor Of Growth Family Member 1
INPP4B	Inositol Polyphosphate-4-Phosphatase Type II B
KCNIP1	Potassium Voltage-Gated Channel Interacting Protein 1
LDLR	Low Density Lipoprotein Receptor
LPA	Lipoprotein(A)
MYH11	Myosin Heavy Chain 11
NAXD	NAD(P)HX Dehydratase
NOTCH3	Notch (Drosophila) Homolog 3
RAB20	RAB20, Member RAS Oncogene Family
RNF150	Ring Finger Protein 150
SCOC	Short Coiled-Coil Protein
SERPINC1	Serpin Family C Member 1
TBC1D9	TBC1 Domain Family Member 9
ZNF330	Zinc Finger Protein 330

References

1. Valsesia A, Macé A, Jacquemont S, Beckmann JS, Kutalik Z. The Growing Importance of CNVs: New Insights for Detection and Clinical Interpretation. Front Genet. 2013;4:92. [PubMed: 23750167]
2. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nat Genet. 2018;50:524–537. [PubMed: 29531354]
3. Diskin SJ, Li M, Hou C, Yang S, Glessner J, Hakonarson H, et al. Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. Nucleic Acids Res. 2008;36,e126. [PubMed: 18784189]
4. Ginsbach P, Chen B, Jiang Y, Engelter ST, Grond-Ginsbach C. Copy Number Studies in Noisy Samples. Microarrays (Basel). 2013;2:284–303. [PubMed: 27605193]
5. Korbel JO, Kim PM, Chen X, Urban AE, Weissman S, Snyder M, et al. The current excitement about copy-number variation: how it relates to gene duplications and protein families. Curr Opin Struct Biol. 2008;18:366–374. [PubMed: 18511261]
6. Nowakowska B: Clinical interpretation of copy number variants in the human genome: J appl Genetics. 2017;58:449–457. [PubMed: 28963714]
7. Rice AM, McLysaght A. Dosage sensitivity is a major determinant of human copy number variant pathogenicity. Nat Commun. 2017;8:14366. [PubMed: 28176757]
8. Itsara A, Cooper GM, Baker C, Girirajan S, Li J, Absher D, et al. Population analysis of large copy number variants and hotspots of human genetic disease. Am J Hum Genet. 2009;84:148–161. [PubMed: 19166990]
9. Burke W, Antommaria AH, Bennett R, Botkin J, Clayton EW, Henderson GE, et al. Recommendations for returning genomic incidental findings? We need to talk! Genet Med. 2013;15:854–859. [PubMed: 23907645]
10. Weiming F, Yuliang W, Youjie L, Xinsheng L, Shuyang X, Zhaoxia L. A novel Notch3 deletion mutation in a Chinese patient with cerebral autosomal dominant arteriopathy with subcortical infarcts and leucoencephalopathy (CADASIL). J Clin Neurosci. 2013;20:322–323. [PubMed: 23151434]
11. Grond-Ginsbach C, Chen B, Pjontek R, Wiest T, Jiang Y, Burwinkel B, et al. Copy number variation in patients with cervical artery dissection. Eur J Hum Genet. 2012;20:1295–1299. [PubMed: 22617347]
12. Grond-Ginsbach C, Chen B, Krawczak M, Pjontek R, Ginsbach P, Jiang Y,et al. Genetic Imbalance in Patients with Cervical Artery Dissection. Curr Genomics. 2017;18:206–213. [PubMed: 28367076]
13. Rosenberg RE, Egan M, Rodgers S, Harter D, Burnside RD, Milla S, et al. Complex chromosome rearrangement of 6p25.3->p23 and 12q24.32->qter in a child with moyamoya. Pediatrics. 2013;131:e1996–2001.
14. Toldo I, Po' C, Morao V, Talenti G, Causin F, D'Avella D, et al. Moyamoya syndrome and 6p chromosome rearrangements: Expanding evidences of a new association. Eur J Paediatr Neurol. 2016;20:766–771. [PubMed: 27236536]
15. Felbor U, Gaetzner S, Verlaan DJ, Vijzelaar R, Rouleau GA, Siegel AM. Large germline deletions and duplication in isolated cerebral cavernous malformation patients. Neurogenetics. 2007;8:149–153. [PubMed: 17211633]
16. Saskin A, Sillon G, Palfreeman N, Buhas D. Col4A1/2 CNVs and cerebral small vessel disease. Neurology. 2018;90:1026–1028. [PubMed: 29703772]
17. Renard D, Miné M, Pipiras E, Labauge P, Delahaye A, Benzacken B, et al. Cerebral small-vessel disease associated with COL4A1 and COL4A2 gene duplications. Neurology. 2014;83:1029–1031. [PubMed: 25098541]
18. Tiszlavicz Z, Somogyvári F, Szolnoki Z, Sztriha LK, Németh B, Vécsei L, et al. Genetic polymorphisms of human β-defensins in patients with ischemic stroke. Acta Neurol Scand. 2012;126:109–115. [PubMed: 22050386]

19. Kibe T, Mori Y, Okanishi T, Shimojima K, Yokochi K, Yamamoto T. Two concurrent chromosomal aberrations involving interstitial deletion in 1q24.2q25.2 and inverted duplication and deletion in 10q26 in a patient with stroke associated with antithrombin deficiency and a patent foramen ovale. Am J Med Genet A. 2011;155A:215–220. [PubMed: 21204235]
20. Tsai CT, Hsieh CS, Chang SN, Chuang EY, Ueng KC, Tsai CF, et al. Genome-wide screening identifies a KCNIP1 copy number variant as a genetic predictor for atrial fibrillation. Nat Commun. 2016;7:10190. [PubMed: 26831368]
21. Macé A, Tuke MA, Deelen P, Kristiansson K, Mattsson H, Nõukas M, et al. CNV-association meta-analysis in 191,161 European adults reveals new loci associated with anthropometric traits. Nat Commun. 2017;8:744. [PubMed: 28963451]
22. Voll SL, Boot E, Butcher NJ, Cooper S, Heung T, Chow EW, et al. Obesity in adults with 22q11.2 deletion syndrome. Genet Med. 2017;19:204–208. [PubMed: 27537705]
23. Falchi M, El-Sayed Moustafa JS, Takousis P, Pesce F, Bonnefond A, Andersson-Assarsson JC, et al. Low copy number of the salivary amylase gene predisposes to obesity. Nat Genet. 2014;46:492–497. [PubMed: 24686848]
24. Iacocca MA, Hegele RA. Role of DNA copy number variation in dyslipidemias. Curr Opin Lipidol. 2018;29:125–132. [PubMed: 29303791]
25. Erqou S, Thompson A, Di Angelantonio E, Saleheen D, Kaptoge S, Marcovina S, et al. Apolipoprotein(a) isoforms and the risk of vascular disease: systematic review of 40 studies involving 58,000 participants. J Am Coll Cardiol. 2010;55:2160–2167. [PubMed: 20447543]
26. Pfeiffer D, Chen B, Bevan S, Jern C, Jimenez-Conde J, Lee JM, et al. Genetic imbalance is associated with poorer outcome after ischemic stroke. Neurol Genet. 2017; 3(1 Suppl 1): S12–S18. [PubMed: 28428978]
27. Nik-Zainal S, Cotter PE, Willatt LR, Abbott K, O'Brien EW. Ring chromosome 12 with inverted microduplication of 12p13.3 involving the Von Willebrand Factor gene associated with cryptogenic stroke in a young adult male. Eur J Med Genet. 2011;54:97–101. [PubMed: 20933620]
28. Luukkonen TM, Mehrjouy MM, Pöyhönen M, Anttonen AK, Lahermo P, Ellonen P, et al. Breakpoint mapping and haplotype analysis of translocation t(1;12)(q43;q21.1) in two apparently independent families with vascular phenotypes. Mol Genet Genomic Med. 2018;6:56–68. [PubMed: 29168350]
29. Burnside RD, Harris A, Speyer D, Burgin WS, Rose DZ, Sanchez-Valle A. Constitutional Chromoanagenesis of Distal 13q in a Young Adult with Recurrent Strokes. Cytogenet Genome Res. 2016;150:46–51. [PubMed: 27825145]
30. Sobey CG, Judkins CP, Sundararajan V, Phan TG, Drummond GR, Srikanth VK. Risk of Major Cardiovascular Events in People with Down Syndrome. PLoS One. 2015;10:e0137093. [PubMed: 26421620]
31. Calogero AE, Giagulli VA, Mongioì LM, et al. Klinefelter syndrome: cardiovascular abnormalities and metabolic disorders. J Endocrinol Invest. 2017;40:705–712. [PubMed: 28258556]
32. Irioka T1, Mizusawa H. Ischemic stroke in a young adult with Turner syndrome. Neurol Sci. 2011;32:317–319. [PubMed: 21153600]
33. Forsberg LA. Loss of chromosome Y (LOY) in blood cells is associated with increased risk for disease and mortality in aging men. Hum Genet. 2017;136:657–663. [PubMed: 28424864]
34. Fuster JJ, Walsh K. Somatic Mutations and Clonal Hematopoiesis: Unexpected Potential New Drivers of Age-Related Cardiovascular Disease. Circ Res. 2018;122:523–532. [PubMed: 29420212]
35. Nørskov MS, Frikke-Schmidt R, Loft S, Sillesen H, Grande P, Nordestgaard BG,et al. Copy number variation in glutathione S-transferases M1 and T1 and ischemic vascular disease: four studies and meta-analyses. Circ Cardiovasc Genet. 2011;4:418–428. [PubMed: 21562205]
36. Iacocca MA, Hegele RA. Role of DNA copy number variation in dyslipidemias. Curr Opin Lipidol. 2018;29:125–132. [PubMed: 29303791]
37. Matarin M, Simon-Sanchez J, Fung HC, Scholz S, Gibbs JR, Hernandez DG, et al. Structural genomic variation in ischemic stroke. Neurogenetics. 2008;9:101–108. [PubMed: 18288507]

38. Adams HP, Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, et al. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. Stroke. 1993;24:35–41. [PubMed: 7678184]
39. Brandt T, Morcher M, Hausser I. Association of cervical artery dissection with connective tissue abnormalities in skin and arteries. Front Neurol Neurosci. 2005;20:16–29. [PubMed: 17290108]
40. Grond-Ginsbach C, Brandt T, Kloss M, Aksay SS, Lyrer P, Traenka C, et al. Next generation sequencing analysis of patients with familial cervical artery dissection. Eur Stroke J. 2017;2:137–143.
41. Pezzini A, Drera B, Del Zotto E, Ritelli M, Carletti M, Tomelleri G, et al. Mutations in TGFBR2 gene cause spontaneous cervical artery dissection. J Neurol Neurosurg Psychiatry. 2011;82:1372–1374. [PubMed: 21270064]
42. Kuang SQ, Guo DC, Prakash SK, McDonald MLN, Johnson RJ, Wang M, et al. Recurrent chromosome 16p13.1 duplications are a risk factor for aortic dissections. PLoS Genetics. 2011:7, e1002118. [PubMed: 21698135]
43. Erhart P, Brandt T, Straub BK, Hausser I, Hentze S, Böckler D, et al. Familial aortic disease and a large duplication in chromosome 16p13.1. Mol Genet Genomic Med. 2018;6:441–445. [PubMed: 29441698]
44. Choi JS, Kim SR, Jeon YW, Lee KH, Rha HK. Identification of DNA copy number aberrations by array comparative genomic hybridization in patients with ruptured intracranial aneurysms. J Clin Neurosci. 2009;16:295–301. [PubMed: 19056275]
45. Bae JS1, Cheong HS, Park BL, Kim LH, Park TJ, Kim JY, et al. Genome-wide association analysis of copy number variations in subarachnoid aneurysmal hemorrhage. J Hum Genet. 2010;55:726–730. [PubMed: 20703242]
46. Santos M, Niemi M, Hiratsuka M, Kumondai M, Ingelman-Sundberg M, Lauschke VM, et al. Novel copy-number variations in pharmacogenes contribute to interindividual differences in drug pharmacokinetics. Genet Med. 2018;20:622–629. [PubMed: 29261188]
47. Willyard C Copy number variations' effect on drug response still overlooked. Nat Med. 2015;21:206. [PubMed: 25742449]
48. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45 (D1): D353–D361. [PubMed: 27899662]
49. Wang M, Lin S. Detecting associations of rare variants with common diseases: collapsing or haplotyping? Brief Bioinform. 2015;16:759–768.. [PubMed: 25596401]

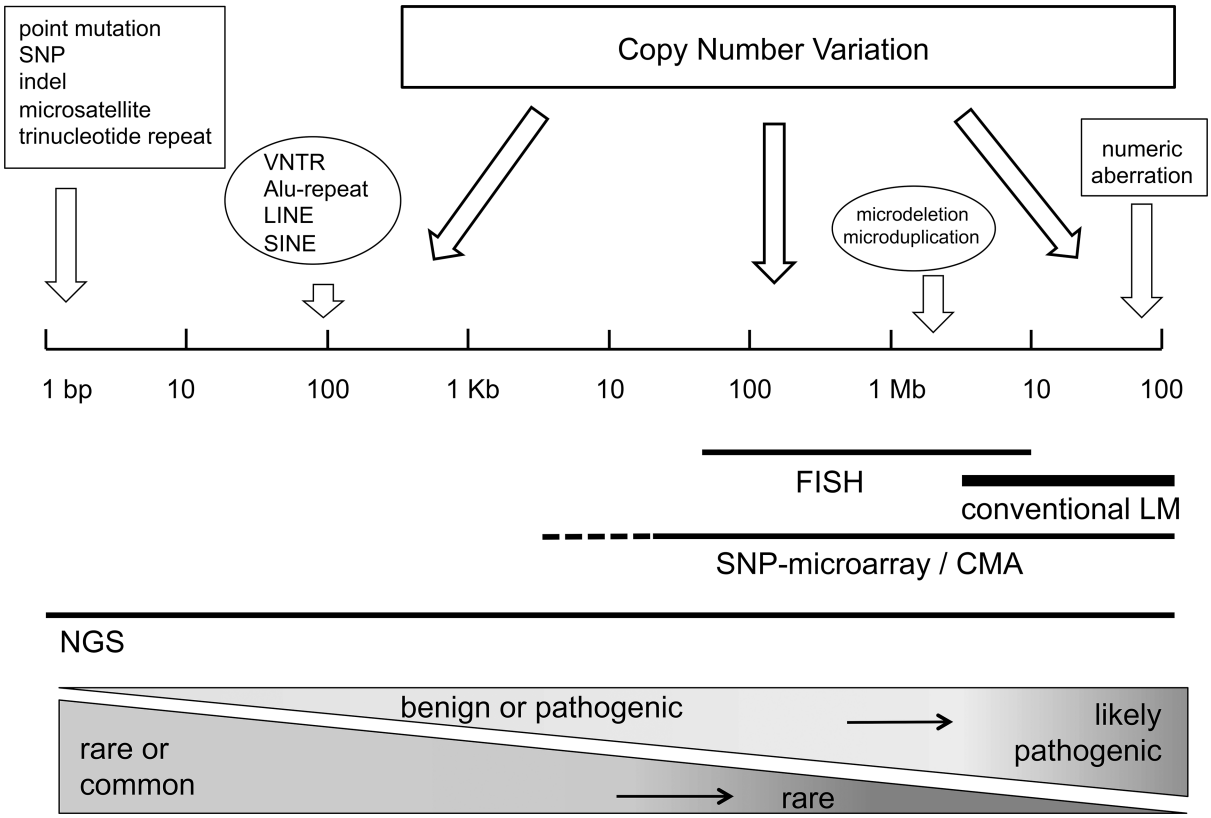

Figure 1. Types of genetic variation, ranked according to size (length of DNA in base pairs), detection methods, clinical impact and population frequency

SNP = single nucleotide polymorphism; indel = short insertion / deletion; VNTR = variable number of tandem repeats; LINE/SINE = long / short interspersed repetitive elements; FISH = fluorescent-labeled in-situ hybridization; LM = light-microscopy. NGS = next generation sequencing.

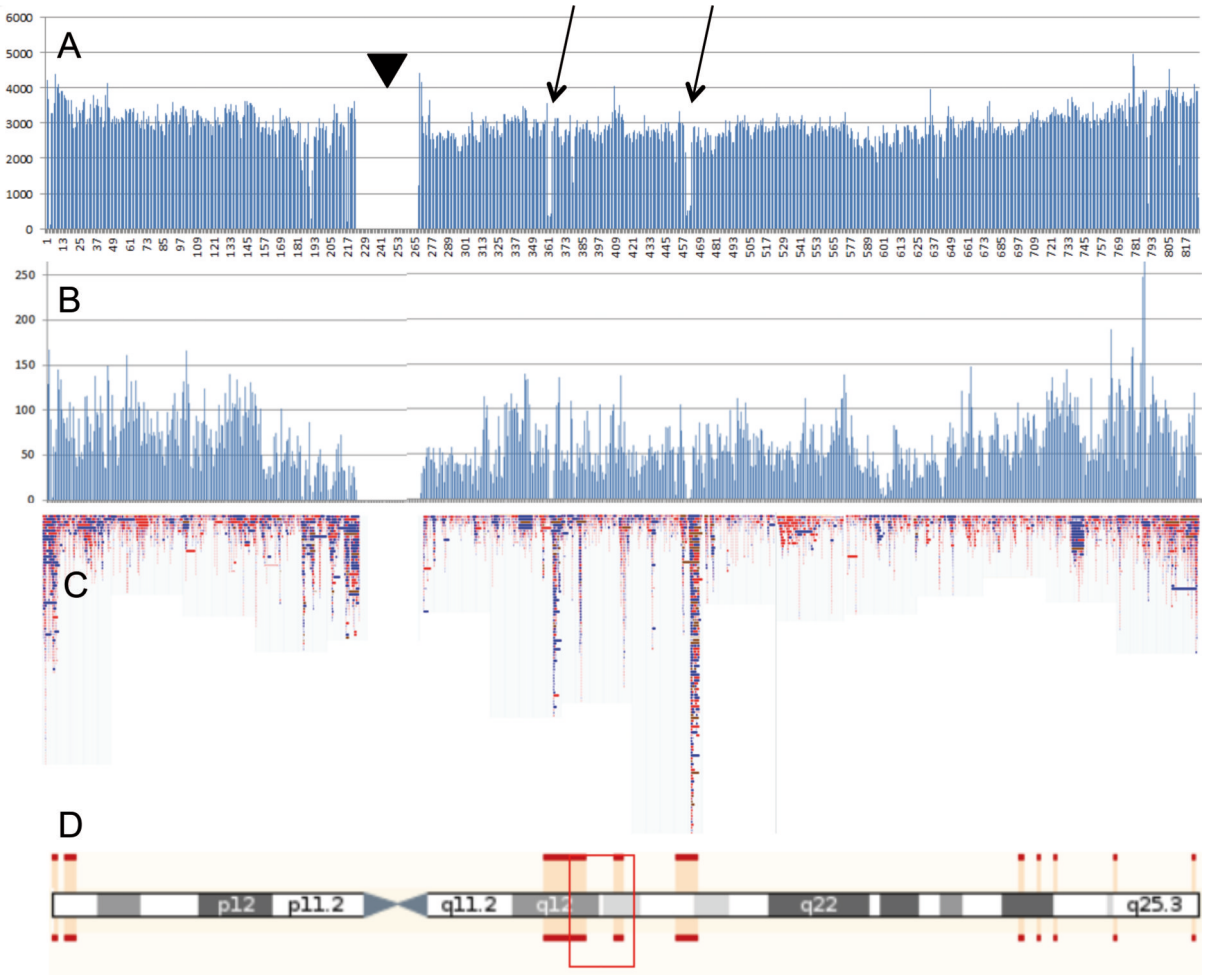

Figure 2. CNV detection using SNP-microarrays

Fig. 2A. All annotated short variants from dbSNP (https://www.ncbi.nlm.nih.gov/SNP/), distributed over 100 Kb bins along chromosome 17; Fig. 2B: all SNPs from chromosome 17 of the Illumina Omni 5 exome platform; Fig 2C. All CNVs of human chromosome 17 from the Database of Genomic Variants (DGV: http://dgv.tcag.ca/dgv/app/home); Fig. 2D. Idiogram of chromosome 17. Red bars delineate regions of uncertain mapping due to segmental duplications. Arrows point to regions with low SNP density in dbSNP. The low SNP-density is not outbalanced by SNP selection for the Illumina platform. Regions with low SNP density on the Illumina array appear to be very rich in CNV. Arrowhead indicates peri-centromeric region.

Table 1.

CNV-findings associated with ischemic stroke.

CADASIL: Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy; EDS: Ehlers Danlos syndrome; SAO: small arterial occlusive disease; LVD: large vessel disease; CeAD: cervical artery dissection; CCM: cerebral cavernous malformations; VNTR: variable number of tandem repeats.

Phenotype	CNV	affected/disrupted genes	Ref
CNV-findings in stroke due to a Mendelian disorder			
CADASIL	100 bp deletion	NOTCH3	[10]
vascular EDS	2q32 deletion	COL3A1, COL5A2	[11]
CNV associated with subtypes of ischemic stroke			
CeAD		enrichment of various CNVs affecting arterial development	[12]
CeAD	16p13.1 duplication	MYH11 /ABCC6 locus	[12]
Moya-moya	6pter duplications		[13,14]
CCM	exonic CNVs	CCM1; CCM2; CCM3	[15]
SAO	13q34 duplication	COL4A1/COL4A2 locus	[16,17]
SAO	low (<4) copy number	DEFB4	[18]
LVD	low (<4) copy number	DEFB4	[18]
CNV associated with complex developmental retardation phenotypes and pediatric stroke			
	1q24 /10q26 deletions	SERPINC1	[19]
CNV associated with stroke risk factors			
Atrial fibrillation	intronic duplication	KCNIP1	[20]
Obesity	CNV burden		[21]
Obesity	16p11.2/22q11.2 deletion		[22]
Obesity	low copy number	AMY1	[23]
Hyperlipidemia	VNTR	LDLR, LPA	[24]