

49. Steinberg, G.R. and B.E. Kemp, AMPK in Health and Disease. *Physiol Rev*, 2009. 89(3): p. 1025-1078.;
50. Sun, M.W., J.Y. Lee, et al., Haplotype structures and large-scale association testing of the 5' AMP-activated protein kinase genes PRKAA2, PRKAB1, and PRKAB2 [corrected] with type 2 diabetes. *Diabetes*, 2006. 55(3): p. 849-855.;
51. Towler, M.C. and D.G. Hardie, AMP-activated protein kinase in metabolic control and insulin signaling. *Circ Res*, 2007. 100(3): p. 328-341.;
52. Weyrich, P., F. Machicao, et al., Role of AMP-activated protein kinase gamma 3 genetic variability in glucose and lipid metabolism in non-diabetic whites. *Diabetologia*, 2007. 50(10): p. 2097-2106.;
53. Xiao, B., M.J. Sanders, et al., Structure of mammalian AMPK and its regulation by ADP. *Nature*, 2011. 472(7342): p. 230-233.;
54. Xu, M., X. Li, et al., Glucose and lipid metabolism in relation to novel polymorphisms in the 5'-AMP-activated protein kinase gamma2 gene in Chinese. *Mol Genet Metab*, 2005. 86(3): p. 372-378.;
55. Zhang, B.B., G. Zhou, and C. Li, AMPK: an emerging drug target for diabetes and the metabolic syndrome. *Cell Metab*, 2009. 9(5): p. 407-416.;
56. Zhu, L., L. Chen, et al., Structural Insights into the Architecture and Allostery of Full-Length AMP-Activated Protein Kinase. *Structure*, 2011. 19(4): p. 515-522.;

**STRATEGIES FOR IDENTIFYING GENETIC FACTORS
IN MULTIFACTORIAL STROKE
(REVIEW)**

Elena Mocan

Institute of Genetics and Plant Physiology, Academy of Sciences of Moldova

Summary

Stroke remains a major cause of death and disability both in developed and developing world countries. Understanding the genetic contributions to ischemic stroke is important not only so as to explain, or predict, the minority of cases that occur in the absence of well-established risk factors, such as smoking, hypertension and diabetes. Since many of the candidate genes tested for an association with ischemic stroke. Furthermore, the effect of each underlying gene may depend upon interaction with other loci and with environmental and lifestyle factors, as reported for some of the mendelian forms of stroke. Here we would like to describe important strategies for identifying genetic factors in multifactorial stroke.

Rezumat

Strategii pentru identificarea factoriilor genetice a ictusului multifactorial

Ictusul cerebrovascular rămâne a fi cauza majoră de deces și invaliditate, atât în țările dezvoltate cât și în cele în curs de dezvoltare. Înțelegerea contribuțiilor genetice în accidentele vascular-cerebrale ischemice, este importantă nu doar pentru a explica sau a prezice minoritatea de cazuri care apar în absența unor factori de risc bine stabiliți cum ar fi fumatul, hipertensiunea și diabetul zaharat. Astfel, multe dintre genele candidate au fost testate în asociere cu ictusul ischemic. În plus, efectul fiecărei gene de bază, ar putea depinde de interacțiunea cu alți loci și cu factorii de mediu și stilul de viață, cum s-a raportat pentru unele forme mendeliene ale ictusului. În acest articol, ne-am dori să descriem strategiile importante pentru identificarea factorilor genetici în ictusul multifactorial.

Problem

Stroke is a major killer since about 5 million people die each year of this disease. Stroke is also the major cause of disability in adults, and the second most important cause of dementia both in developed and developing world countries. Among those who survive a stroke, the risk of a second stroke is very high. Currently known methods characterizing risk factors for stroke (age, cardiovascular diseases, arterial fibrillation, arterial hypertension, diabetes mellitus, carotid stenosis) are of low sensitivity and specificity.

Risk factors for stroke

Because stroke is pathologically heterogeneous it can be expected that the risk factor profiles leading to the different types and subtypes of stroke vary [1]. However, many large prospective studies on risk factors were performed before it was feasible to differentiate between the main types, let alone the various subtypes of ischemic stroke [14].

Age, gender, race, ethnicity, and heredity have been identified as markers of risk for stroke [3, 7, 9]. Although these factors cannot be modified, their presence helps identifying those at greatest risk, in whom treatment of modifiable risk factors can be initiated. High blood pressure, hypertension, is the most prevalent and modifiable risk factor for stroke [10]. A number of other modifiable risk factors have been identified and include cigarette smoking, diabetes mellitus, certain cardiac conditions, obesity, hypercholesterolemia and physical inactivity [8].

In recent years there has been considerable interest in identifying novel risk factors for stroke [14]. Examples of these are infection, hemostatic factors, inflammatory markers, plasma homocysteine and various genetic polymorphisms. However, because of small sample sizes, differing inclusion criteria between studies, and other methodological issues, data on the impact of these novel risk factors on stroke are still limited [3]. It is important to identify new markers to more appropriately predict the possible development of stroke and to identify those subjects with potential benefit from preventive therapy.

Ischemic stroke

The majority of strokes (approximately 85% of strokes) is ischemic and occurs when a blood vessel becomes occluded and the blood supply to part of the brain is totally or partially blocked. In the majority of ischemic strokes, intravascular thrombus formation plays an important role for vessel occlusion. The thrombus commonly forms around atherosclerotic plaques where it gradually narrows the lumen of the affected artery (stenosis) [9, 12]. A thrombotic stroke represents 52% of all ischemic strokes. It is a consequence of atherosclerotic disease and is caused by destabilization of atheromatous plaque with ensuing thrombosis and vessel occlusion locally or distally due to embolism [9].

Genetics in stroke

Stroke is both a heterogeneous and a multifactorial disease, in which heritable and environmental factors equally contribute [5, 9]. The heritable component has been investigated in family [6], twin [2, 3] and animal studies [6]. Twin studies provide the most reliable evidence of a genetic component in complex diseases, as they are least confounded by environmental factors. In twins, concordance rates were reported to be about 1.6 times greater in monozygotic twins than in dizygotic twins. However, most of these studies have been relatively small and have not differentiated between stroke types [6].

Cohort and case-control studies on family history of stroke support a hereditary component in both ischemic and hemorrhagic stroke. However, study designs and possible publication and recall bias have made it difficult to reliably estimate the strength of the association [6].

Most studies combined ischemic and hemorrhagic stroke and failed to differentiate between the various ischemic strokes subtypes. Recent data have suggested that a family history of stroke is a risk factor for various ischemic stroke subtypes but not for stroke of undetermined etiology [4].

In all subtypes, the family history effect was stronger in patients with a young age of onset [Schulz 2004]. Concern has been raised that part of the increased risk may be explained by heritability of common intermediate phenotypes, such as hypertension [6, 11].

Of note is also that some recent studies suggest that there are sex-specific differences in stroke heritability, with women being about 50% more likely to have a maternal than a paternal history of stroke [13]. No similar effect was detected in men. This mother-to-daughter mechanism is hard to explain by classical genetic mechanisms, but could perhaps be explained by non-genetic factors or by transmitted epigenetic factors [13].

Multifactorial stroke

In contrast to the mendelian forms of stroke, such as CADASIL, MELAS, Fabry's disease and etc., for the more common polygenic trait it can be expected that the contribution from each single susceptibility locus is relatively small [4,12]. The identification of rare, Mendelian forms of stroke has greatly benefited our understanding of the common, multifactorial forms of stroke, because they allow us to identify key pathways that lead to a stroke, which are likely to be involved also in the development of other forms of stroke.

Furthermore, the effect of each underlying gene may depend upon interaction with other loci and with environmental and lifestyle factors, as reported for some of the mendelian forms of stroke. To further complicate the issue, many conventional risk factors such as diabetes mellitus, hypertension and cardiovascular diseases, are themselves complex genetic diseases that may interact with environmental exposures [4].

Still, on the individual level genetic variants may interact in an additive or multiplicative manner with other genetic variants and with environmental exposures, thus making certain individuals more vulnerable to certain exposures.

Strategies for identifying genetic factors in multifactorial stroke

The most popular approach for identifying genes in human polygenic ischemic stroke has been the candidate gene approach using case-control methodologies. More recently this has been extended to family-based association studies.

Linkage-based approaches have been used less frequently. In the future genome wide association studies are likely to become more widely used.

We will illustrate gene finding for behavioural traits with examples of phenotypes taken from various research fields such as addiction and personality and psychophysiological traits considered to be risk markers or risk factors for disease.

Linkage studies

A general strategy to find genes for Mendelian traits is called classical linkage and is based on Fisher's theory of likelihood inference [3]. It is referred to as being parametric or model-based because an explicit genetic model for the disease or trait locus has to be provided. Classical linkage analysis models the distance between a DNA marker locus and a putative disease locus in small numbers of large multigenerational families (pedigrees) consisting of both affected and unaffected family members. It is the method of choice for the genetic mapping of single-gene diseases, especially when these diseases are rare. Classical linkage requires that a model for the disease or trait locus is specified a priori, in terms of allelic frequencies, penetrance and mode of action (recessive or dominant). Complete penetrance implies that all individuals with a high-risk genotype (genotype dd in the case of a recessive disorder and genotypes Dd and DD in the case of a dominant disorder) will develop the disorder. If there are individuals with a high-risk genotype who do not develop the disease, then the penetrance of the genotype is said to be incomplete. Individuals without a high-risk genotype who develop a disorder that is phenotypically indistinguishable from the genetic form, are called phenocopies [8, 12].

In linkage analysis a number of DNA markers of known location, evenly dispersed throughout the entire genome, are measured in individuals from multiple generations. DNA markers can be mutations in a single base pair (Single Nucleotide Polymorphisms (SNPs)) or a variable number of repeats of two or more base pairs (microsatellites). Genetic markers (normally microsatellites) covering the whole genome are used to genotype patients and affected family members [14,15].

An advantage of linkage analysis is that it is performed in a hypothesis-free manner and does not require any prior knowledge of the underlying disease mechanism. The method also is insensitive to spurious results due to problems with population stratification. For several reasons the linkage approach is difficult to apply in the search of genes contributing to stroke. Because of the late-onset of the disease the collection of information from other family members becomes difficult, and the affection status of siblings and offspring uncertain. The polygenic nature of stroke, and shared environmental exposures, also contributes to the difficulties with the linkage approach, because linkage is unable to detect genes with minimal or modest effect on stroke risk [8, 9].

Candidate gene allelic association studies

Linkage is usually genome-wide, while association studies are limited to candidate genes or candidate regions. Furthermore, linkage analyses must be carried out in pedigrees (families and sibling pairs), while association can be performed at the population level. With allelic association studies an association between a disease and a specific allele can be detected in groups of unrelated cases (e.g. patients) and controls (e.g. healthy subjects) [7, 15]. Association can be found either with functional genetic variants that have biological consequences related to disease, or with other variants that are in linkage disequilibrium with these variants. Linkage disequilibrium occurs when a marker allele (i.e. a SNP) and the QTL are so close on the chromosome that they co-segregate in the population over many generations of meiotic recombination. Association studies are similar in design to classic case-control studies in epidemiology. DNA is collected from all participants and the trait is compared across the various allelic variants of the DNA marker. The ideal candidate gene has been shown to be functional: it influences the concentration of the (iso)form of a protein, its functionality or efficiency, or perhaps most importantly, its responsiveness to environmental factors triggering the expression of the gene [14, 15]. The problem with a candidate gene approach for most complex traits is the potentially huge proportion of genes, which can serve as candidates. Several strategies are possible to select an optimal set of candidate genes.

Genes are normally chosen based on their known function. First, genes that are part of physiological systems known to influence the trait can be tested as candidates. Secondly, genes or chromosomal regions that are known to influence the trait in animals can be tested as candidate genes (or regions) in humans [15].

The goal of these studies is to identify specific microsatellites, single nucleotide polymorphisms (SNPs) or haplotypes (combinations of microsatellites and/or SNPs) that influence the susceptibility of developing the disease. SNPs have been increasingly accepted as powerful genetic markers for the detection of susceptibility genes through association analyses as they are more frequent and stable than microsatellites. Moreover, it has been estimated that the human genome contains more than 10 million SNPs [17].

The major approach used to find stroke genes has been the candidate gene approach using case-control methodologies. This is a hypothesis-driven approach in which genes that may be involved in the pathogenesis of stroke are tested for association. One or more markers covering the gene are genotyped in a set of cases and controls, to look for allelic variants that are over- or underrepresented in cases compared with controls. To be reliably detected, small relative risks require large sample sizes, in the magnitude of 1,000 patients or more [5]. However, few studies have achieved such numbers. It can be speculated that differences in sample characteristics and limitations in study designs may explain why most reports of significant associations have not been replicated.

This has raised concerns on the validity of association studies and complex genetics in general and lead to publications suggesting standard criteria for genetic association studies in stroke [5, 14].

MOLSTROKE

Here I would like to present the beautiful example how could be unite interdisciplinary efforts of molecular geneticists, cellular biologists, physiologists, and clinicians are needed to

further our understanding of genetic susceptibility to stroke, as well as to the genetic influence on acute stroke pathology as well as recovery. Such knowledge should significantly contribute to further development of stroke prevention strategies [16]. Project MOLSTROKE with budget of 2 300 000 euros was started in 2005 and prolonged 36 months that assembled 7 partners from different European countries with multidisciplinary backgrounds.

The main scientific and technological goals of MOLSTROKE were to identify mechanisms and molecular protagonists that participate in the vascular pathological events leading to thrombotic stroke. MOLSTROKE has addressed objectives using two concomitant strategies.

During these years a tissue bank has been initiated, arterial and venous tissue arrays from atherosclerotic patients and vessel transplant donors have been constructed. Initial array investigations have been performed. Genes potentially involved in plaque rupture have been analyzed and immune-response in mouse models of atherosclerosis is under investigation. Lipid effects on angiogenesis have been studied using microarray analyses, *in vivo* and *in vitro* angiogenic assays, and cellular signaling. These initial studies have already confirmed that plaque lipids have multiple biological effects on different cell types. In particular, the initial results obtained in MOLSTROKE have shown the unique capacity of lipids to stimulate acquired immunity, innate immunity and to modulate neo-angiogenesis [16].

The proposed investigations will implement novel technologies of differential display of unknown genes and vascular tissue arrays. Thus, weighted identification of stroke denominators can be accomplished and thereby lead to improved diagnostic and treatment modalities. The research armory spans genomics, tissue arrays, molecular biology, cell biology, immunology, biochemistry, gene transfer, animal models and integrative bioinformatics software tools.

Future studies will therefore most likely have to be performed by multiple centers acting in concert to achieve sufficient power for studies on well characterized groups and subgroups of stroke patients and controls. When such a study has reported significant associations, confirmations in replication studies that are similarly designed and equally, or better, powered to detect associations will be required to confirm susceptibility loci.

Acknowledgements

This work was supported in part by the state contract and Russian Foundation of Basic Research - individual Project № 08-04-90902, and U.S. Civilian Research & Development Foundation-Academy of Sciences of Moldova Moldovan Young Scientist Support Program MYSSP №1403

References

1. Arboix A, Morcillo C, Garcia-Eroles L, Oliveres M, Massons J, Targa C. Different vascular risk factor profiles in ischemic stroke subtypes: a study from the "Sagrat Cor Hospital of Barcelona Stroke Registry". *Acta Neurol Scand* 2000;102:264-70.
2. Bak S, Gaist D, Sindrup SH, Skytthe A, Christensen K. Genetic liability in stroke: a long-term follow-up study of Danish twins. *Stroke* 2002;33:769-74.
3. Brass LM, Alberts MJ. The genetics of cerebrovascular disease. *Baillieres Clin Neurol* 1995;4:221-45.
4. Casas JP, Hingorani AD, Bautista LE, Sharma P. Meta-analysis of genetic studies in ischemic stroke: thirty-two genes involving approximately 18,000 cases and 58,000 controls. *Arch Neurol* 2004;61:1652-61.
5. Chanock SJ, Manolio T, Boehnke M, et al. Replicating genotype-phenotype associations. *Nature* 2007;447:655-60.
6. Floßmann E, Schultz UG, Rothwell PM. Systematic review of methods and results of studies of the genetic epidemiology of ischemic stroke. *Stroke* 2004;35:212-27.
7. Goldstein LB, Adams R, Alberts etc. Primary prevention of ischemic stroke: A guideline from the American Heart Association/American Stroke Association Stroke Council. *Circulation*. 2006;113:e873-923.

8. Gulcher JR, Gretarsdottir S, Helgadóttir A, Stefansson K. Genes contributing to risk for common forms of stroke. *Trends Mol Med* 2005;11:217-24.
9. Hassan A, Markus HS. Genetics and ischaemic stroke. *Brain* 2000;123:1784-812.
10. Kuller LH. Epidemiology and prevention of stroke, now and in the future. *Epidemiol Rev* 2000;22:14-7.
11. Lindgren A, Lovkvist H, Hallstrom B, et al. Prevalence of stroke and vascular risk factors among first-degree relatives of stroke patients and control subjects. A prospective consecutive study. *Cerebrovasc Dis* 2005;20:381-7.
12. Markus HS, Martin RJ, Simpson MA, Dong YB, Ali N, Crosby AH, Powell JF. Diagnostic strategies in CADASIL. *Neurology* 2002;59:1134-8.
13. Touzé E, Rothwell PM. Sex differences in heritability of ischemic stroke. A systematic review and meta-analysis. *Stroke* 2007. p. 45-55
14. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447:661-78.
15. Vink J.M., Boomsma D.I. Gene finding strategies *Biological Psychology* 61. 2002. p. 53-57.
16. www.molstroke.com.
17. Kruglyak, L. and Nickerson, D. A. (2001) Variation is the spice of life. *Nat Genet.*, 27, 234-236.

**IDENTIFICATION OF GENETIC RISK TO ISCHEMIC STROKE –
THE GENOME WIDE ASSOCIATION STUDY AND META-ANALYSIS
(REVIEW)**

Elena Mocan

Institute of Genetics and Plant Physiology, Academy of Sciences of Moldova

Summary

A GWAS is an approach that involves rapidly scanning markers of many samples across the complete genome, to find genetic variations associated with a particular disease. Such studies are particularly useful in finding genetic variations that contribute to common complex diseases such as ictus.

We have shown analysis of recent articles dedicated to GWA studies of stroke with scopes to demonstrate positive associations with ischemic stroke. Here we proposed next candidate genes and their polymorphisms such as factor V Leiden Gln506, ACE I/D, MTHFR C677T, prothrombin G20210A, PAI-1 5G allele, ACE I/D and glycoprotein IIIa Leu33Pro to use in research of patients with ischemic stroke from Moldavian population.

Rezumat

**Identificarea riscului genetic la ictus – studii de asociere largă
a genomului și meta-analiză**

GWAS (Genome wide association study sau Studii de asociere largă a genomului) este o metodologie care implică scanarea rapidă a markerelor de mai multe probe în genomul complet, pentru a găsi variații genetice asociate cu o anumită boală. Așa studii sunt utilizate particular în găsirea variațiilor genetice care pot să descrie predispoziția la bolile comune complexe, cum ar fi ictus cerebral.

Am arătat analiza articolelor recente dedicate studiului despre GWAS cu scopul de a demonstra asocierea pozitivă la ictusul ischemic. Aici ne-am propus următoarele gene candidate și polimorfismul lor cum ar fi factor V Leiden Gln506, ACE I/D, MTHFR C677T, prothrombina G20210A, PAI-1 5G allele, ACE I/D și glycoproteina IIIa Leu33Pro pentru a le utiliza în cercetarea pacienților cu ictus cerebral din populația Republicii Moldova.