Etiology and Pathophysiology of Stroke as a Complex Trait
Speranza Rubattu, Rosangela Giliberti, and Massimo Volpe

Stroke (brain attack) is currently the third leading cause of death in Western societies. Recent advances in molecular genetics have finally demonstrated what has long been suggested by the clinical observation, that is, stroke is not only the complication of major pathologic conditions such as atherosclerosis, hypertension, or cardiac diseases, but rather it represents a complex trait itself. Thus, the pathogenesis of stroke is often the result of the combined effects of genes exerting a direct contributory role and of their interactions with several environmental determinants.

A genetic dissection of stroke has been attempted in suitable animal models and in humans. With this approach, the genetic defects underlying monogenic disorders associated with stroke were identified. Moreover, important findings have recently highlighted the contribution of genes encoding cardiovascular hormones, such as the atrial natriuretic peptide, for the pathogenesis of multifactorial, polygenic forms of stroke.

A more thorough understanding of the fine mechanisms, dependent from mutations within stroke susceptibility genes and underlying the disease pathogenesis, may help to introduce new specific tools to achieve better prevention and treatment of stroke.

KEY WORDS: Cerebrovascular accident, epidemiology, genetics.
Thus, more accurate prevention, diagnosis, and treatment will necessarily rely on the combined results of epidemiologic, basic research, and clinical studies. In this regard, the major task of the basic research approach is the attempt to identify the precise determinants of the disease and to better understand the pathophysiologic mechanisms. The results of such studies, along with the epidemiologic evidence, may provide fundamental advances in the understanding of the human disease to get a more appropriate clinical approach.

The current article reviews available data related to the etiopathogenesis of stroke, with particular emphasis on the recent achievements on the genetic basis of this complex disease associated with hypertension.

**ROLE OF MODIFIABLE RISK FACTORS**

Several epidemiologic studies carried out during the past decades have pointed out that, along with non-modifiable risk factors such as age, gender, race, ethnicity, and genes, there are definite modifiable determinants of stroke. They include smoking, physical inactivity, dietary habits, obesity, high blood pressure (BP), cardiac diseases and atrial fibrillation, carotid stenosis, and diabetes.\(^8\)\(^,\)\(^9\) In fact, several reports in the literature have described a significant increase in the risk of stroke in the presence of these factors, particularly when clustering of these risk or predisposing factors occurs. On the other hand, the role of blood levels of cholesterol is still controversial.\(^10\)\(^,\)\(^11\) Of note, although stroke represents a heterogeneous pathologic condition, risk factors for cerebrovascular accidents have been usually considered for stroke in general, probably because the different subtypes were not distinguished well enough in most epidemiologic studies.\(^9\) Thus, selective predisposing conditions for lacunar or atherothrombotic ischemic strokes have not been clearly dissected out,\(^12\) although it is apparent that conditions such as atrial fibrillation play a role in the lacunar subtypes. On the other hand, intracerebral hemorrhagic strokes associated with hypertension may also derive from rupture of intracerebral microaneurysms or even represent the complication of the atherosclerotic or lacunar process.

The common denominator by which modifiable risk factors predispose to stroke can be ultimately identified with the mechanisms leading to increased damage of the vessel wall components and to increased atherosclerotic, thrombotic, or embolic processes. Thus, an appropriate control of such predisposing modifiable conditions will definitively lead to a significant decline of the risk of stroke, as it has been repeatedly documented by large epidemiologic studies. In fact, the results of recent clinical trials have demonstrated that a better pharmacologic control of high BP levels is able to reduce stroke occurrence.\(^13\)\(^,\)\(^14\) The Hypertension Optimal Treatment (HOT) study has recently shown that lowering BP levels to 140 mm Hg systolic and 90 mm Hg diastolic may drastically reduce the incidence of all cardiovascular events, including stroke, whereas no significant improvement was observed by further lowering BP levels.\(^15\) In addition, undertreatment of hypertension was indicated as a major cause of increased stroke occurrence in hypertensive subjects.\(^16\)

Consistently with the preliminary epidemiologic observations, reduction of cigarette smoking and changes of dietary habits leading to a better control of obesity also resulted in the expected decrease of cardiovascular diseases.\(^17\)\(^-\)\(^19\) In particular, a role of potassium supplement in decreasing the risk of stroke was demonstrated, although the precise mechanisms underlying such beneficial effect remain to be assessed.\(^20\)\(^,\)\(^21\) Remarkably, a favorable effect of fish consumption on stroke mortality was also reported.\(^22\)

Beyond the results obtained with the investigations of the influence of some of the major modifiable determinants for stroke, a few studies were designed to test the clinical relevance of minor, less characterized predisposing risk factors. For example, attempts to identify the pathogenic role of alcohol consumption in cardiovascular and cerebrovascular diseases raised intriguing questions. In fact, a “protective” role of light-to-moderate alcohol consumption was suggested in several studies,\(^23\)\(^,\)\(^24\) including a recent analysis on the well-characterized prospective cohort of the Physician’s Health Study.\(^25\) The effects of ethanol on lipid metabolism, hemostasis, and nitric oxide synthase activity in the vessel wall may contribute to the surprisingly favorable influence of a moderate alcohol intake on the incidence of stroke.

For obvious reasons, the role of aspirin on stroke occurrence was investigated. At this regard, whereas a beneficial effect was described by some reports,\(^26\) a larger benefit on myocardial infarction but not on stroke incidence was reported when aspirin treatment was added to aggressive antihypertensive therapy in the HOT study.\(^14\) Thus, although aspirin treatment remains a cornerstone for primary prevention of myocardial infarction as well as for secondary prevention of both stroke and myocardial infarction, evidence for its beneficial effect on prevention of a first stroke is still inconclusive.

Epidemiologic research also contributed to the identification of the role played by several components of the coagulation cascade, such as factor V, factor VII, fibrinogen, and tissue plasminogen activator,\(^27\)\(^-\)\(^30\) and by metabolites, such as homocysteine,\(^31\) in the increased risk of cardiovascular diseases including stroke. In particular, most of the studies reported a positive association between higher plasma levels of these factors and occurrence of vascular diseases, with odd ratios ranging between 1.3 and 1.6 for stroke.
Thus, measurements of plasma fibrinogen and homocysteine levels and of factor VII activity have become relevant parameters for the identification of individuals at higher risk for stroke. The causal link between these factors and vascular diseases remains to be defined, although it may well be explained by mechanisms leading to an increased rate of vascular damage, particularly of the atherosclerotic or thrombotic type. In particular, whereas the components of the coagulation cascade are interpreted as thrombogenic markers, plasma homocysteine levels have been established as an independent risk factor for atherosclerosis.

Moreover, these results, while providing interesting pathophysiologic clues, suggest a possible direct contributory role of the genes encoding the corresponding proteins or enzymes (when mutated) into the pathogenesis of vascular diseases.

**ROLE OF NONMODIFIABLE RISK FACTORS: THE GENETIC BACKGROUND**

Although significant benefits were obtained through the improved control of major predisposing conditions such as hypertension, cardiac diseases, smoking, and obesity, the incidence of stroke, particularly ischemic stroke, remains very high, and it is increasing in Western societies. A recent epidemiologic study carried out in Minnesota, has reported that a decline of the incidence of ischemic stroke in men and women, observed during the period 1955 to 1980, was followed by a remarkable increase of disease occurrence in both sexes. This trend has been confirmed in different populations. A study performed in Taiwan, for instance, demonstrated that although the secular trend of hemorrhagic stroke declined, ischemic strokes were not modified during the period 1974 to 1988, despite the progress in cardiovascular therapy. Thus, different pathogenetic factors may be involved in the determination of different forms of stroke. On the other hand, it is a common notion that incidence of stroke varies among individuals despite the presence of identical predisposing conditions. Therefore, control of other factors is required to achieve better preventive and therapeutic results and particularly, to cut down the enormous health and social burden bound to stroke.

In this regard, lessons learned from epidemiologic surveys, major clinical trials, and common clinical observations have highlighted the potential role of the genetic background, which exerts a key role by playing both indirect and direct effects. In fact, some of the predisposing conditions for stroke also recognize a genetic background. Thus, genes involved in hypertension, cardiac diseases, diabetes, vascular disease, and atherosclerosis can be indirectly involved in the pathogenesis of cerebrovascular accidents. Most important, recent studies have clearly documented the existence of several genes exerting a direct contributory role in the pathogenesis of stroke associated with hypertension.

**Genetic Approach in Animal Models** Availability of animal models for human diseases presents the advantage of reducing the obvious complexity inherent in the study of human populations. In addition, due to their genetic homogeneity, animal models favor a direct approach of genetic dissection of the most common diseases. Finally, the short lifespan of rodents permits genetic analysis of different stroke phenotypes.

In this regard, important results were achieved during the past 5 years by using the well-characterized model of the stroke-prone spontaneously hypertensive rat (SHRsp). This rat strain, originally obtained by selective inbreeding from the spontaneously hypertensive rat (SHR), is characterized by a very high frequency of cerebrovascular events as compared to the stroke-resistant SHR under a high salt, low potassium, low protein diet (stroke proneness). In addition, this strain shows a greater sensitivity to the middle cerebral artery occlusion (MCAO)-induced brain infarct as compared to the normotensive strain (stroke sensitivity).

Remarkably, this animal model shares several similarities with the human disease. In fact, hypertension and dietary factors behave as predisposing risk factors for both rats and humans. In particular, a high salt dietary content exerts a causative role, whereas high potassium intake is a protective factor toward stroke.

Thus, the stroke-prone rat represents a suitable model for the investigation of the etiopathogenetic mechanisms underlying the human disease. On the basis of these considerations, it has been largely used for characterization of hormonal and vascular abnormalities or other intermediate phenotypes associated with cerebrovascular disease and, only recently, to attempt a genetic dissection of stroke.

In this latter regard, studies based on linkage analysis design were performed on F2 hybrid populations derived either from the SHRsp/SHR or from the SHRsp/WKY cross. The purpose of these studies was to identify whether any part of the stroke-prone genome cosegregated with the occurrence of the stroke phenotype. The necessary requisite for this type of investigation is the accurate definition of a clear-cut pathologic phenotype. On the other hand, the subsequent statistical analysis requires quantitative parameters. Thus, the two existing linkage studies chose as phenotypes the number of days necessary to develop strokes under the stroke permissive dietary regimen (stroke proneness) and the extent of ischemic damage after MCAO (stroke sensitivity).
were able to demonstrate the existence of different genes involved in the two distinct phenotypes and exerting a direct pathogenetic role on stroke associated with hypertension, independently from the BP levels. In fact, in our studies assessment of BP levels by tail–cuff sphygmomanometry and of stroke occurrence distributions in the parental strains and among the F2 offsprings consistently documented the lack of a relationship between the two variables (Fig. 1). Thus, three quantitative trait loci (QTL), accounting altogether for 28% of the phenotype variance in our study,46 and one QTL, accounting for 67% of the phenotype variance in the study by Jeffs et al, 47 were detected. In addition, evidence for the existence of both causing and protective genes was obtained.46 In fact, we identified one major locus on chromosome 1 containing a stroke-causing gene and two other loci on chromosomes 4 and 5 containing genes with a delaying effect toward the disease. A significant epistatic interaction was observed between the chromosome 5 and the chromosome 1 QTL. Candidate genes were identified for both areas, that is, the gene encoding adrenomedullin on chromosome 1 (mapping close to the peak of significance) and the genes encoding atrial and brain natriuretic peptides on chromosome 5 (mapping exactly at the peak of significance).

However, studies aimed at the identification of chromosomal linkage areas present major limitations, among which the considerable length of such areas, extending from about 15 up to 50 cM and, thus, containing millions of genes. This may represent a major problem in the further precise localization of the disease gene (positional cloning). For these reasons, the discovery of QTL should be viewed as the first, albeit significant, step of a laborious and time-consuming approach aimed at the progressive narrowing of the initial area down to the size required for the positional cloning, approximately 1 cM. The latter procedure requires a fine, dense map of the area of interest and it involves the generation of congenic strains, first, and then of congenic substrains carrying smaller and smaller pieces of the original QTL and still showing the disease phenotype.48 Therefore, physical reconstruction of the smallest area can be carried out as described in details elsewhere.49 The generation of congenic lines for stroke QTL is currently in progress. Recently, we reported preliminary observations derived from congenic lines carrying the chromosome 1 QTL in the SHRsp configuration within a stroke-resistant genetic background.50 These lines showed a 50% incidence of stroke, versus 0% of the SHR and 100% of the SHRsp parental strains, during the 3 months of observation. Moreover, a significantly lower body weight, similar to what is observed in the stroke-prone parental strain, was associated with the stroke phenotype. Thus, these data confirm the pathogenetic role of the gene contained within the chromosome 1 QTL and they certainly justify the need for the identification of the specific genetic component hidden inside this area.

Once a QTL has been identified, another type of experimental approach may be undertaken, in parallel with the development of congenic strains. This is the
identification and strain characterization of “hot” candidate genes contained within the area. “Candidates” for the disease of interest are all those genes encoding proteins, hormones, or enzymes with known biological activities potentially related to the phenotype under investigation. Although this procedure has been criticized, it represents a logical follow-up of QTL analysis and it may truly help in identifying substantial abnormalities within putative genes mapping inside the linkage area or, in the presence of only negative evidence, it may help rule out potential candidates. Obviously, this approach does not have the power to definitively prove that the candidate gene and the disease gene within the QTL are the same, unless a substantial series of abnormalities is consistently found only in the gene of the diseased individuals.

A successful result of candidate gene analysis, applied to the investigation of hypertension in the rat, was obtained with the work on the Dahl rat model of salt-sensitive hypertension. In fact, the characterization of the 11β-hydroxylase and of the aldosterone synthase genes, mapping within a linkage area detected on chromosome 7 in an F2 intercross from Dahl-S/Dahl-R rats, revealed structural and functional abnormalities in the salt-resistant versus the salt-sensitive rat.51–54 Furthermore, these data paralleled quite closely a similar situation found in some forms of human hypertension.55,56

When this type of approach was applied to our investigation of the genetic basis of stroke, a very intriguing result was achieved.

The gene encoding atrial natriuretic peptide (ANP), contained within the peak of linkage of one of the QTL discovered in the SHRsp/SHR intercross, on chromosome 5, revealed structural alterations and consequent functional abnormalities only in the disease strain.57 The ANP is a well-known physiologically important cardiovascular peptide that exerts natriuretic, diuretic, and vasorelaxant properties, and it is expressed in cardiac and cerebral tissues.58,59 Thus, our findings in the rat supported the ANP gene as a suitable candidate for the stroke proneness. Furthermore, a subsequent structural analysis of the same gene in affected individuals revealed the existence of a coding mutation responsible for a twofold increase in the risk of stroke in the subjects carrying the mutation.60 Thus, our results added further support to the existence of an intriguing parallelism between the animal model and the human disease. Moreover, they further documented the scientific relevance of candidate gene analysis as a rationale approach following the discovery of QTL, particularly when targeted to genes mapping at the peak of significance within the area. On the basis of this suggestive concordance between the disease of rats and humans, and on additional evidence provided by several experimental studies on the role of ANP in cardiac and vascular remodeling,61,62 further analyses of the ANP gene in new stroke-control populations is currently under way, along with in vitro studies aimed at the identification of the abnormal functional properties of the ANP peptide when mutated. In this regard, we have already obtained evidence that the mutant stroke-prone ANP cDNA encodes a peptide that is differentially processed by both cellular and serum proteases.57 In fact, we constantly observed the production of an additional new peptide (of about 6.5 kD) only after the processing of the stroke-prone proANP. Moreover, measurement of the cGMP levels resulting from the interaction of ANP peptides with the guanylyl cyclase-coupled receptor NPR-A receptor revealed significantly higher levels in the presence of the stroke-prone derived ANP peptides (P < .05).57 Furthermore, characterization of the regulatory mutation located within the polypeptide enhancer element-binding site of the ANP gene in the stroke-prone strain has so far demonstrated that it significantly reduces the promoter activity by 50% as compared to the stroke-resistant strain in both endothelial and cardiomyocyte cells (P < .05).57 Of note, the ANP gene is not involved in the increased stroke sensitivity to MCAO of the SHRsp strain.63

In the search for stroke-related genes, another experimental model was investigated, the SHR with MCAO-induced ischemic stroke.64 In particular, a differential display procedure was applied to the injured cerebral area and a likely candidate for stroke, the gene encoding adrenomedullin, a vasoactive peptide, was identified. In fact, this gene was overexpressed in brains of SHR after MCAO-induced ischemic stroke but not in the control SHR. Moreover, the well-known cardiovascular functions of adrenomedullin,65 very similar to those of ANP, further reinforced its possible pathogenetic contributory role. Interestingly, this gene turned out to map right inside a stroke QTL identified on chromosome 1 in the SHRsp/SHR intercross used in our studies,66 and, therefore, adrenomedullin may be also proposed as a candidate gene for the stroke proneness of our model.

Finally, mouse models were used to test the role of endothelial and neuronal nitric oxide synthase genes as putative candidates for stroke susceptibility.67 Whereas their direct contribution to stroke phenotypes had been previously excluded by linkage analysis in the rat model,66,67 the endothelial nitric oxide synthase (eNOS) gene knock-out mouse has demonstrated that eNOS is an important modulator of the degree of cerebral ischemia after stroke, by maintaining the local blood flow. Whether this finding relates to a primary or secondary involvement of the eNOS gene in stroke in this model remains to be assessed.

In summary, these experimental approaches are
fruitful, informative, and highly promising to reach the final goal, that is, a better comprehension of the human disease.

Genetic Studies in Humans Genetic studies of a complex trait such as stroke, which is usually characterized by a late onset, are hampered in humans by several difficulties. In fact, the expected lack of complete genealogic trees or of large cohorts of affected relative pairs to be used for linkage studies, the known genetic heterogeneity of human populations, along with the common coexistence of other risk factors, and the pathologic heterogeneity of stroke types do not favor at all a direct approach of genetic dissection in

**FIG. 2.** Top) Identification of a valine (Val) to methionine (Met) transposition within the human proatrial natriuretic peptide (ANP) peptide, in particular within the portion known as cardiokinin (25 to 92 amino acids). Bottom) Left, single-stranded conformational polymorphism analysis of the mutation and its direct sequencing. Gel A shows a double mutant homozygous. Gel B shows heterozygous subjects. WW = wild type; WM = heterozygous; MM = double mutant. Right, frequency of the mutation among cases and controls from all cohort and from a low-risk group. Odds ratio = 2.0 (dominant mode), 95% confidence interval = 1.17 to 3.39; P = .01. Black bars indicate cases. White bars indicate controls.
humans. Thus, important results could be achieved with the analysis of monogenic disorders associated with stroke. In contrast, with regard to the multifactorial and polygenic forms of stroke, literature mostly contains reports of either positive or negative associations between potential candidate genes and disease occurrence. These latter studies were obviously favored by the availability of polymorphic informative markers located within the genes of interest, able to distinguish wild-type from mutant individuals (ie, different allelic configurations), and they tested the hypothesis that mutations in genes encoding proteins with vascular functional properties may contribute to determine vascular diseases, including stroke. Moreover, the rational support for these investigations was provided by the epidemiologic findings demonstrating positive associations between plasma levels of some of the corresponding encoded proteins and stroke occurrence.

Although the background was quite promising, only variable and inconsistent results, with regard to cerebrovascular disease, were obtained. In particular, different and weak associations with molecular variants of angiotensin converting enzyme, methyltetrahydrofolate, fibrinogen, and the platelet glycoprotein IIIa genes were described. The different ethnic background of the populations used and the small size of the studies conducted so far, with only one exception, may have represented a main limitation. Therefore, the inconsistency of the data does not currently support a definite scientific relevance and more studies, particularly large prospective studies, are needed to better clarify the role of these factors.

Recently, we reported a positive association between a mutation of the human ANP gene and the risk of stroke. In particular, a detailed structural analysis of this gene by single stranded conformational polymorphism revealed a coding mutation on exon 1, responsible for a valine to methionine transposition within the proANP peptide, significantly and independently associated to stroke occurrence (Fig. 2). Interestingly, the systolic BP levels were significantly lower in individuals carrying the ANP mutation and showing, at the same time, a higher incidence of cerebrovascular events. Whether this represented the overall effect of the ANP mutation or it resulted from the combination of lower BP levels in carriers of the ANP mutation remains an interesting issue to explore. Of note, our observation was obtained from the analysis of the largest matched case/control cohort prospectively studied so far in cardiovascular medicine, derived from the Physician’s Health Study. Although with the limitations of association type of studies and with the caution necessary when considering the heterogeneity of the disease in animals and in humans, our finding may be viewed as an intriguing evidence of a parallelism between the animal model and the human disease. Finally, it is important to mention that no information is yet available about the hypothetical role that genes responsible of monogenic disorders may exert with regard to the etiology of the multifactorial and polygenic forms of stroke. The recent development of a set of single nucleotide polymorphisms located within known genes certainly represents a great promise to expand the genetic investigation of stroke in humans by the association type of approach.

**SUMMARY AND OUTLOOK**

The recent introduction of molecular biology and genetic techniques is progressively changing the cultural approach of scientists and physicians toward the understanding of common diseases.

The basic mechanisms underlying hypertension, diabetes, atherosclerosis, and cardiac diseases are now finely dissected out and their interrelationships partially understood. Thus, new tools for the development of targeted preventive and therapeutic strategies may be provided in the near future to the medical community.

In particular, a large body of evidence has accumulated to show that stroke is not the mere consequence of high BP levels or cardiac diseases, although these may play an important predisposing role. Other factors, however, particularly genetic factors, need to be considered to reach a more thorough understanding of the complex pathogenesis of stroke. In this regard, the identification of individuals genetically at risk for stroke represents an important clinical end point.

As a matter of fact, this goal can now be pursued on the basis of the latest achievements in this area of research. Identification of individuals carrying either variant or wild-type allelic configurations for specific “risk” genes is already possible and it is of great help for early identification of all monogenic disorders resulting from known genetic defects. In regard to the multifactorial forms of diseases such as stroke, the usefulness of the genetic individual characterization will rely on the definite assessment of the true pathogenetic role of the risk genes as well as on our understanding of the pathophysiologic mechanisms related to gene abnormalities. This achievement will also permit a new way to design specific therapeutic interventions (pharmacogenomics). Only then, the final goal of all our research efforts will be accomplished: to cure common diseases such as stroke with the appropriate strategies and to reduce their risks while improving the life expectancy of “predisposed” individuals.

**REFERENCES**


