Epigenetics and stroke risk – beyond the static DNA code

Abstract: Advances in high-throughput genome sequencing and genome-wide association studies indicate that only a fraction of estimated variability in stroke risk can be explained by genetic variation in protein-coding genes alone. Epigenetics is defined as chromatin-based mechanisms important in the regulation of gene expression that do not involve changes in the DNA sequence per se. Epigenetics represents an alternative explanation for how traditional risk factors confer increased stroke risk, provide a newer paradigm to explain heritability not explained by genetic variation, and provide insight into the link between how the environment of a cell can interact with the static DNA code. The nuclear-based mechanisms that contribute to epigenetic gene regulation can be separated into three distinct but highly interrelated processes: DNA methylation and hydroxymethylation; histone density and posttranslational modifications; and RNA-based mechanisms. Together, they offer a newer perspective on transcriptional control paradigms in blood vessels and provide a molecular basis for understanding how the environment impacts the genome to modify stroke susceptibility. This alternative view for transcriptional regulation allows a reassessment of the \textit{cis}/\textit{trans} model and even helps explain some of the limitations of current approaches to genetic-based screens. For instance, how does the environment exert chronic effects on gene expression in blood vessels after weeks or years? When a vascular cell divides, how is this information transmitted to daughter cells? This review provides an introduction to epigenetic concepts and a conceptual framework for understanding the shortcomings of an approach to stroke research that focuses solely on the static DNA code. Additionally, it will discuss classical and emerging mechanisms of epigenetic gene regulation that are especially relevant to large-vessel ischemic stroke.

Keywords: DNA methylation, endothelial nitric oxide synthase, histone posttranslational modifications, long noncoding RNA, ANRIL, heritability

Introduction

On average in the United States, someone has a stroke every 40 seconds and stroke-related deaths occur every 4 minutes. Stroke is ranked number four among all causes of death and is a leading cause of acquired disability in adults.\textsuperscript{1} With an estimated 7 million Americans over the age of 20 years having suffered a stroke, it represents a tremendous socioeconomic burden, with direct costs of $34.3 billion in 2008. Despite intensive and sustained research efforts, it is estimated that between 2010 and 2030 the direct costs of stroke will escalate further by 238\% as the prevalence of stroke increases by 25\% over the same time period.\textsuperscript{2} It must be underscored that the problem of stroke is not confined to the US or other high-income countries, but is rather a global epidemic with broad socioeconomic relevance.\textsuperscript{3} For example, more than 85\%
of all strokes worldwide occur in low-and middle-income countries. Stroke ranks as the second-leading cause of mortality worldwide in adults over the age of 15 years. It represents the fifth-leading cause of disability-adjusted life years lost, just after HIV/AIDS. The problem is worsening due to an aging population and changes in the distribution of modifiable cardiovascular risk factors in developing and third-world countries. Newer approaches for understanding stroke pathogenesis are urgently needed.

Stroke is defined as a clinical syndrome characterized by an acute loss of neurological function, with symptoms lasting greater than 24 hours, that results from a vascular problem. Implicit in this definition of “clinical syndrome” is the idea that stroke is not a single disease. This definition encompasses both ischemic and hemorrhagic stroke, although the former accounts for nearly 90% of the overall stroke burden.

Ischemic stroke does not represent a homogeneous group: large-vessel atherosclerosis, cardioembolism, and small-vessel occlusion are important pathomechanistic subtypes. Of the common forms of ischemic stroke, perhaps the greatest strides have been made in understanding large-vessel atherosclerosis, due to its prevalence and commonality with atherosclerotic disease of the coronary and peripheral vasculatures. In this stroke subtype, traditional risk factors that increase the risk of stroke are well known, such as hypertension, dyslipidemia, and smoking. But the molecular mechanisms that impart an increased risk to the individual are poorly understood. Undoubtedly, genetic predisposition is an important contributory factor for defining an individual’s stroke risk. Evidence from twin and other studies of familial aggregation of stroke provides estimates of heritability.

However, recent genome-wide association studies (GWASs) suggest that only a fraction of the estimated heritability can be explained by genetic variation alone. Perhaps the genetic contribution to stroke risk is overestimated, since uninterrogated genetic elements may contribute to heritable risk. We argue that the impact of gene–environment interactions has been underestimated and can be more fully addressed by an epigenetic theory of complex, non-Mendelian disease.

Epigenetics is broadly defined as chromatin-based mechanisms of gene expression that do not involve changes to the DNA sequence per se. Prominent examples include DNA methylation, the density of histones and their posttranslational modifications, and RNA-based pathways (Figure 1). These molecular mechanisms are the “nuts and bolts” on which three founding principles of epigenetic theory are grounded. First, the same DNA sequence can demonstrate variable expressivity depending on its chromatin state. This chromatin state can be inherited in mitosis. Second, epigenetic modifications are more malleable than the static genetic code, and therefore potentially more responsive to environmental stimuli, both intrinsic and extrinsic to the cell, as well as short-term therapeutic intervention. Third, the possibility of meiotic inheritance of epigenetic modifications allows for the transfer of disease susceptibility from parent to child outside the context of the classical genetic code. Taken together, epigenetic theory may help to explain how traditional risk factors confer increased stroke risk, and account for the heritability that is not explained by known genetic variation.

The aims of this review are to provide an introduction to epigenetic concepts applicable to future stroke research and to discuss classical and emerging mechanisms of epigenetic gene regulation, especially those that are relevant to large-vessel ischemic stroke.

**Ischemic stroke risk: why the static code is not enough**

The evidence for a genetic liability in ischemic stroke is very strong. Perhaps the most definitive evidence comes from animal studies. For example, in the well-established mouse model of focal cerebral ischemia involving permanent occlusion of the distal middle cerebral artery, final infarct volume is highly dependent on the inbred mouse strain. In one recent study, there was as much as a 30-fold difference in infarct volume between inbred mouse strains at the phenotypic extremes. Heritability of the trait of infarct volume was 0.88, arguing for a powerful genetic contribution. However, it is important to remember that such studies rigorously control the environment of the animal. It is reasonable to infer that gene–environment interaction will be much more relevant in the human setting.

In human ischemic stroke, several well-studied monogenic disorders increase stroke risk and account for some familial aggregation of ischemic stroke. Prominent examples include cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) due to NOTCH3 mutations, Fabry disease due to GLA mutations, and sickle-cell disease due to HBB mutations. These stroke syndromes are rare, and responsible for only a fraction of the overall stroke burden. In the general population, however, demonstrating a genetic predisposition to stroke risk has been more challenging. Nonetheless, accruing evidence from larger and better-designed studies argues for a significant heritable contribution. For example, a recent analysis of the prospectively collected Framingham Heart Study demonstrated a threefold increased risk of offspring stroke with documented parental stroke by 65 years of age. These results parallel those from earlier twin, case-control, and cohort studies pointing to a genetic contribution to ischemic
stroke risk.\textsuperscript{10,20} Studies of intermediate phenotypes, such as carotid-artery intima-media thickness, also demonstrate significant heritability ranging from 0.55 to 0.71.\textsuperscript{22–25} Finally, recent GWASs of many thousands of patients have identified genetic susceptibility loci for ischemic stroke: most prominently and consistently the chromosome 9p21 (chr9p21) locus.\textsuperscript{26–28} In a recent meta-analysis of eight ischemic stroke studies, the chr9p21 locus was shown to be a risk factor for large-vessel ischemic stroke with a modest effect size (odds ratio [OR], 1.20 [95% CI, 1.08–1.33, \( P = 0.0006 \))).\textsuperscript{26} Importantly, this increased risk was independent of traditional cardiovascular risk factors, suggesting that it may exert its influence via novel mechanistic pathways. We will discuss this locus in more detail below. More recently, a large GWAS study reported for 1162 large vessel ischemic stroke cases. When compared with 1244 control patients a new susceptibility locus on chromosome 6p21.1 was identified (OR = 1.62 [95% CI 1.36 to 1.93, \( P = 3.9 \times 10^{-8} \)]. Importantly, this was replicated in multiple independent population cohorts.\textsuperscript{29} Interestingly, this SNP falls within a region of the genome that does not contain known protein coding genes.

Epigenetic mechanisms may also play a role in the regulation of gene expression after ischemic insult in adults. There is small but accumulating evidence to suggest that DNA methylation and chromatin changes occur in cells following ischemia.\textsuperscript{30,31} For example, classical trans/cis regulators of neuronal gene expression seem to interface with epigenetic pathways, especially in disease. A major mechanism relevant to neuronal-enriched gene expression is the neuron-restricted silencing factor/repressor element-1 silencing transcription factor (REST). REST was originally described as a repressor of neuronal gene expression in nonexpressing cell types. REST is a transfactor that binds to the RE-1 cis DNA element. REST/RE-1 trans/cis interactions are key to establishing and maintaining neuronal gene-expression pathways, and hence phenotype.\textsuperscript{30,32,33} In mature neurons, REST is quiescent but can be activated by ischemia.\textsuperscript{34} Recent findings have shown that epigenetic remodeling of chromatin occurs in ischemia-induced neuronal cell death and can be abrogated with knockdown of the REST trans complex.\textsuperscript{30} Therefore, epigenetic pathways interface with REST and thus is relevant in disease. The use of these epigenetic mediators in therapeutics should be approached with caution. As much as there are theoretical applications to the modulation of epigenetic regulators, as with the use of microRNAs in therapeutics, they must be fully understood for safe and effective use in therapy.\textsuperscript{32,35–37}
Currently, epigenetic modulators as indicators of prognosis have shown some promise.35,38

Taken together, a genetic contribution to human ischemic stroke susceptibility is certain, but less so is the magnitude of that contribution. The remainder of this section will highlight three observations regarding ischemic stroke syndromes that appear to be poorly explained by traditional models. An epigenetic theory of complex, non-Mendelian disease may offer newer insights.

**Discordance of monozygotic twins for ischemic stroke**

Classical twin studies are a powerful research tool to assess the genetic contribution to human disease.39,40 By comparing the prevalence of a disease phenotype among monozygotic (MZ) versus dizygotic (DZ) twins, the relative genetic contribution can be estimated, and is referred to as heritability. This strategy is based on the principle that MZ twins are (nearly) genetically identical, while DZ twins share 50% of segregating DNA sequence variation. In theory, an increased prevalence of disease amongst MZ versus DZ twins points to a genetic contribution to disease phenotype. For example, in a population-based study of Danish twins, the heritability estimate for the liability to stroke death was 0.32, suggesting a strong genetic predisposition.41 This calculation was based in part on concordance rates for stroke death of 10% and 5% for MZ and DZ twins, respectively. Similarly, in a separate study of male twin pairs born between 1917 and 1927 who were US veterans, proband concordance rates were 17.7% and 3.6% for MZ and DZ twins, respectively, again suggesting a strong genetic component of stroke risk.42 Closer inspection of these data, however, reveals a paradox. Although stroke is more prevalent amongst MZ versus DZ twins, the concordance for stroke within MZ twin pairs is poor. What explains the discordance for stroke within MZ twin pairs? The familiar argument proposes that environmental exposures, both intrinsic and extrinsic to the cell, eg, cigarette smoking, poor diet, and a sedentary lifestyle, differentially affect the members of a twin pair, thereby altering their individual stroke risk. But how these differential exposures impact the genome to impart enhanced disease susceptibility remains poorly understood.12,13

**Phenotypic variability in CADASIL and other monogenic stroke syndromes**

CADASIL is an autosomal dominant disease and the most common cause of inherited stroke in adults.10,20 It is characterized by five main symptoms – migraine with aura, subcortical strokes, mood disorder, apathy, and cognitive dysfunction – and initially becomes manifest in young and middle-aged adults.43 CADASIL is a severe disease that progresses to dementia, mutism, and a bedridden state over the course of approximately 25 years. Characteristic magnetic resonance imaging (MRI) findings typically precede clinical symptoms by several years. Eventually, they include confluent white matter changes in periventricular areas and the centrum semiovale. CADASIL is a systemic arteriopathy of small and medium-sized vessels caused by mutations in a single gene, NOTCH3, a transmembrane receptor predominantly expressed in vascular smooth muscle cells (VSMCs). Despite tremendous advances in the pathophysiology and molecular genetics of CADASIL, and the recent description of murine transgenic models, a specific treatment is not yet available.43,44 Complicating treatment and prognostication of patients with CADASIL is the well-recognized extreme variability in clinical phenotype of patients of the same age, patients within the same family, and even patients carrying the same causative NOTCH3 mutation.45–49 For example, Mykkanen et al reported that 18 of 21 Finnish CADASIL pedigrees shared a common ancestral mutation – a C457T missense mutation in exon 3 resulting in an R133C substitution.49 The age at first-ever stroke among these family members varied from 28 to 71 years.48,49 Opherk et al studied 151 CADASIL patients from 95 families to better understand the heritability of MRI lesion volume.46 Consistent with previous reports, no differential effect of NOTCH3 genotypes on lesion volumes was observed; however, heritability estimates were high, ranging from 0.634 to 0.738, depending on adjustments for identified covariates. From a genetics perspective, these data argue for the strong contribution of modifier genes to MRI lesion volume, ie, a quantitative marker for disease severity, in CADASIL. However, such modifier genes have yet to be identified. Epigenetic theory offers a different perspective. Consider the recent report of different clinical phenotypes in MZ CADASIL twins with a causative NOTCH3 mutation.48 In this example, modifier genes are not likely to be contributory, and so alternative explanations are required. Herein, we have used CADASIL as a specific example; however, other monogenic stroke syndromes, ie, Fabry disease and sickle-cell disease, demonstrate similar phenotypic variability poorly explained by current disease models.10,20

**Genome-wide association studies of large-vessel ischemic stroke: missing heritability and epigenetic determinants of stroke risk**

GWAS is a powerful technique to discover associations between variation in the genome and phenotypic variance.
in a patient cohort. Pragmatically, this approach involves evaluating several hundred thousand to more than a million single-nucleotide polymorphisms (SNPs) in thousands of individuals using commercial “SNP chips” that capture most but not all common variation in the genome. The comparison group, or cohort population, needs to be carefully controlled for confounding variables. Especially important is the degree of genetic admixture and background. The power of the technique is threefold: (1) it is unbiased, not assuming any biological knowledge about a clinical phenotype; (2) it takes advantage of haplotype blocks, which limits the number of SNPs that need to be tested; and (3) it is statistically powerful in its ascertainment of common genomic variation. Replication in independent patient and control cohorts is key. For these reasons, GWASs represent a major advance compared to traditional candidate gene and familial linkage studies for exploring the genetic landscape of complex, non-Mendelian diseases. Indeed, the method has been successfully applied to such myriad conditions as Crohn’s disease, systemic lupus erythematosus, early onset myocardial infarction, and many others, including ischemic stroke.50,51

However, a discrepancy has emerged in these reams of data. Conceptually, the ability of GWASs to identify genetic susceptibility loci is reliant on the “common disease, common variant” hypothesis, which posits that common diseases are attributable to common genomic variations present in at least 1%–5% of the population.51 The discrepancy in GWASs as applied to a vast majority of common human diseases, or complex traits, is that the identified common variant(s) confer incremental risk (1.1–1.5-fold) and explain only a small fraction of the estimated heritability. Perhaps the best example is human height, a classic complex trait, with an estimated heritability of 80%. GWASs have identified more than 50 susceptibility loci that together explain only 5% of phenotypic variance.51–53 Large-vessel ischemic stroke and its intermediate phenotypes are no exception.11 What accounts for this “missing heritability” in GWASs of complex diseases? Several explanations have been suggested. Some argue that as-of-yet uninterrogated common genomic variants of even smaller effect sizes will additively account for the missing heritability. Others argue that difficult-to-study gene–gene interactions will compound effect sizes. Still others suggest that rare variants with possibly larger effect sizes, poorly detected using currently available SNP chips, are the missing pieces in the heritability puzzle. However, empirical data for these hypotheses are lacking at this time in ischemic stroke.51,54 We argue, as others have, that epigenetic theory provides an alternative conceptual framework to account for, at least in part, the missing heritability of complex diseases.55 It is tantalizing, therefore, that a method designed to identify genomic susceptibility loci has instead identified epigenetic effectors as the major determinants of large-vessel ischemic stroke liability. For example, the chr9p21 locus is the most replicated marker of coronary artery disease, myocardial infarction, and large-vessel ischemic stroke.26–28 Intriguingly, this genomic region is devoid of protein-coding genes, but instead contains a long noncoding RNA (lncRNA) – antisense noncoding RNA in the INK4 locus (ANRIL) – a mediator of epigenetic gene regulation.56 Most recently, the largest GWAS of ischemic stroke to date identified a new association with the HDAC9 gene in large-vessel ischemic stroke with a relatively large estimated effect size (OR = 1.38 [95% CI, 1.22–1.57, P = 1.87 × 10^{-11}]).28 HDAC9 is a member of the histone deacetylases, classical epigenetic regulators of gene expression. Further examination of this finding is needed, as the authors suggest distinct genetic architectures for different stroke subtypes, implying a function for HDAC9 in distinct epigenetic events. Finally, a more limited study uncovered an association between carotid-artery intima-media thickness and HDAC4.37 The pattern is clear: unbiased views of genetic variation in large-vessel ischemic stroke liability point to a prominent role for epigenetics.

Taken together, these three observations – discordance of MZ twins for stroke, phenotypic variability in monogenic stroke syndromes, and the missing heritability in GWASs – provide examples of stroke susceptibility (including heritability) that are poorly explained by traditional disease models. The next section will introduce the “nuts and bolts” of epigenetic gene regulation and articulate an epigenetic theory of complex non-Mendelian disease.

Molecular nuts and bolts: epigenetic effectors of gene expression and theory of complex disease

The compaction required of genomic DNA to fit into a eukaryotic cell nucleus is staggering. The haploid human genome is made up of approximately 3.3 billion DNA base pairs.59 If we estimate that an average human adult is comprised of 10^{13} cells, the total length of DNA in a single individual would span 2 × 10^{13} meters, or roughly the distance from the earth to the sun and back 70 times.59 This phenomenal degree of compaction is achieved by the packaging of DNA into a DNA-protein complex referred to as chromatin. Chromatin structure is based on a fundamental repeating unit conserved across all eukaryotic genomes – the nucleosome. A human nucleosome comprises 146 bp of DNA
wrapped around an octamer of core histone proteins; namely, two molecules each of H2A, H2B, H3, and H4. Adjacent nucleosomes are linked by shorter, species-specific stretches of DNA associated with a fifth histone protein – histone H1. The molecular pathways that regulate the structure and accessibility of chromatin without changing the A-C-G-T genetic code constitute the “nuts and bolts,” or effectors, of epigenetic theory. They are discussed separately below, but constitute a highly integrated and evolutionarily conserved system for regulated gene expression (Figure 1).

DNA methylation: adding fifth and sixth bases to the genetic code

DNA methylation refers to the covalent modification of the 5-position of cytosine to create 5-methylcytosine (5mC), the “fifth base of DNA.” This apparently simple modification adds a significant layer of epigenetic complexity to regulated gene expression. DNA methylation is generally viewed as a repressive mark associated with the inhibition of transcriptional initiation. For example, it is strongly implicated in the silencing of repetitive (parasitic) DNA sequences, X-chromosome inactivation, genomic imprinting, mammalian embryogenesis, and cellular differentiation.

The molecular mechanisms underlying DNA methylation and their influence on gene expression are increasingly well understood. This unique pyrimidine (5mC) continues to base-pair with guanine, and in mammals is restricted (almost exclusively) to CpG dinucleotides. A curiosity in the evolution of the mammalian genome is that it is CpG-depleted. However, certain regions of the genome are relatively spared from this depletion and are referred to as CpG islands. They are associated with the 5′-regulatory regions (or promoters) of ~60%–70% of human genes and intergenic, repetitive DNA sequences such as Alu elements.

Interestingly, CpG dinucleotides in repetitive DNA sequences are densely methylated (and transcriptionally silent), whereas those associated with gene promoters are typically unmethylated. In mammals, the addition of a methyl group to the 5-position of cytosine is catalyzed by three distinct DNA methyltransferases (DNMTs) encoded by distinct genes, each on different chromosomes: DNMT1, DNMT3a, and DNMT3b (Figure 2). The latter two enzymes are responsible for de novo methylation and the establishment of DNA methylation patterns during embryogenesis and early development. In contradistinction, DNMT1 attends to maintenance methylation and the propagation of DNA methylation patterns during mitotic cell division from parent to daughter cells (Figure 2). To accomplish this task, it preferentially binds to hemimethylated DNA and is localized to the replication fork. In this way, methylation on the nascent DNA strand is informed by the methylation pattern of the complementary strand.

DNA methylation patterns are transmitted to daughter cells in a semiconservative fashion. The error rate for DNA methylation conservation is significantly greater than the error rate for transmission of the A-C-G-T static DNA code. Notwithstanding this increased error rate, DNA methylation is a remarkably stable epigenetic modification. However, DNA demethylation does occur and is best characterized in specific developmental windows, eg, in preimplantation embryos and primordial germ cells. This has generally been thought to occur via a passive (replication-dependent) process mediated by a regulated absence or reduction of DNMT activity. Examples of rapid and active (replication-independent) DNA demethylation events in response to cell signaling and cellular differentiation are accruing, but a plausible mechanism has been elusive. Most recently, great

---

**Figure 2** The addition of a methyl group at the 5-position of cytosine is mediated by the DNA methyltransferase (DNMT) family, where S-adenosylmethionine serves as the methyl donor.

**Notes:** The ten-eleven translocation (TET) family of enzymes is able to oxidize 5-methylcytosine to 5-hydroxymethylcytosine in an oxygen-dependent reaction requiring adenosine triphosphate and 2-oxoglutarate. Cytosine can then be generated from the action of putative demethylases on 5-hydroxymethylcytosine, but has not yet been fully described.
excitement in the field has been generated over the rediscove
of 5hmC as an intermediate in an active demethylation
pathway.172 The formation of 5hmC is an oxidative reaction
from 5mC that is catalyzed by the ten-eleven translocation
(TET) family in mammals. The TET1, TET2, and TET3
proteins appear to play a role in diverse biological processes
development from pluripotent stem-cell differentiation to explo
from leukemia.73 The TET enzymes are capable of catalyzing
5mC to 5hmC, 5-formylcytosine, and 5-carboxylcytosine
through successive oxidation reactions.8 The enzymatic
players responsible for active DNA demethylation are still
not well understood, and may also include the AID/APOBEC
(activation-induced cytidine deaminase/apoprotein B
mRNA–editing catalytic polypeptides) family, a family of
base excision repair glycosylases, in addition to the TET
family discussed above.75 At a pragmatic level, the two classical
techniques for the assessment of DNA methylation are
bisulfite sequencing and methylation-sensitive restriction
enzyme digests, neither of which is capable of distinguishing
between 5mC and 5hmC.73 These fundamental discoveries
poise the field for a rethinking of DNA methylation dynamics
in regulated gene expression, and the establishment and
maintenance of cellular identity.

How does DNA methylation regulate gene expression?
Two putative mechanisms are best characterized. First,
a methyl group added to the 5-position of cytosine alters the
structure of DNA by projecting into the major groove of the
DNA double helix, thereby potentially impacting sequence-
specific recruitment of DNA-binding proteins, including
transcriptional regulators.13,61 Prominent examples include
Myc, activator protein–2, hypoxia-inducible factor-1α, and
the insulator protein CTCF.13,61 The relative contribution of
DNA methylation in defining trans factor binding is poorly
understood. The second mechanism attests to the highly
integrated nature of epigenetic pathways and is reliant on
a family of methyl-CpG binding domain (MBD) proteins.
These include MBD1, MBD2, MBD4, MeCP2, and Kaiso.
These proteins can block the binding of activating trans
factors and recruit to the chromatin histone–modifying
enzymes, chromatin remodelers, and even RNA molecules
that cooperatively act to silence gene expression.13,61,76

Intriguingly, mammalian MBD3 does not bind to methyl-
ated DNA, but is an integral member of the nucleosome
remodeling and deacetylase complex and is associated with
repressive chromatin structure.77,78 The recent demonstration
that it specifically recognizes and binds to 5hmC but
not 5mC supports the emergence of 5hmC as the putative
sixth base of DNA.77–79

Histone proteins and posttranslational modifications
The second layer of epigenetic gene regulation resides in
the histone proteins around which the DNA double helix is
wound. These evolutionarily conserved proteins are comprised
of a globular domain and a histone amino-terminal tail.
The latter structure, in particular, provides a robust platform
for a myriad of posttranslational modifications that can inform
gene expression.61,80 To date, more than 100 distinct modification
sites have been characterized (Table 1).81 A major chal-
lenge of the postgenomic era is to understand their individual
and combinatorial effect on gene expression.65,82 Examples of
well-studied histone posttranslational modifications include
lysine acetylation, lysine and arginine methylation, serine
and threonine phosphorylation, and lysine ubiquitylation.82,80
These marks are mutually exclusive for any given histone
amino acid residue. For example, histone 3 lysine 9 (H3K9)
can either be acetylated or methylated, but not both.

How do histone posttranslational modifications regulate
gene expression? Two general mechanisms are postulated.
First, histone posttranslational modifications can alter
the physical structure of chromatin and its accessibility to DNA-
binding proteins, such as transcriptional regulators.82 The
best example is histone lysine acetylation, whereby acetyl
marks neutralize the basic charge of lysine residues and

Table 1 Known epigenetic modifications and their effect on gene
transcription

<table>
<thead>
<tr>
<th>Mechanism</th>
<th>Transcriptional effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DNA modifications</strong></td>
<td></td>
</tr>
<tr>
<td>Methylation (CpG dinucleotides)</td>
<td>↓</td>
</tr>
<tr>
<td>Hydroxymethylation (CpG dinucleotides)</td>
<td>↑</td>
</tr>
<tr>
<td><strong>Histone posttranslational modifications</strong></td>
<td></td>
</tr>
<tr>
<td>Histone H3</td>
<td></td>
</tr>
<tr>
<td>Acetylation (K9, K14)</td>
<td>↑</td>
</tr>
<tr>
<td>Methylation</td>
<td></td>
</tr>
<tr>
<td>K4</td>
<td>↑</td>
</tr>
<tr>
<td>K9</td>
<td>↓</td>
</tr>
<tr>
<td>K27</td>
<td>↓</td>
</tr>
<tr>
<td>K36</td>
<td>↑</td>
</tr>
<tr>
<td>Phosphorylation (S10)</td>
<td>↑↓</td>
</tr>
<tr>
<td>Histone H4</td>
<td></td>
</tr>
<tr>
<td>Acetylation (K5, K8, K12, K16)</td>
<td>↑</td>
</tr>
<tr>
<td>Methylation (K20)</td>
<td>↓</td>
</tr>
<tr>
<td><strong>Long noncoding mRNA (lncRNA) mechanisms</strong></td>
<td></td>
</tr>
<tr>
<td>Repressive lncRNAs</td>
<td></td>
</tr>
<tr>
<td>eg, Xist, ANRIL, HOTAIR</td>
<td>↓</td>
</tr>
<tr>
<td>Activating lncRNAs</td>
<td></td>
</tr>
<tr>
<td>eg, HOTTIP</td>
<td>↑</td>
</tr>
</tbody>
</table>

Abbreviations: K, lysine; S, serine.
thereby prevent the higher-order compaction of chromatin. This conformational change makes chromatin more open and accessible to DNA-binding proteins.\textsuperscript{83,84} More important than the direct physical effect on the structure of chromatin, perhaps, is the regulatory information contained within the specific combinations of histone marks. This mechanistic concept is captured in a hypothesis termed the “histone code,” eloquently put forward by Strahl and Allis more than 10 years ago.\textsuperscript{85} At its core, the hypothesis states that a given combination of modifications is “read” by a combination-specific protein or protein complex to effect a specific gene-expression outcome. Given the hundreds of histone-modification sites, there are innumerable combinatorial possibilities.\textsuperscript{86} Cracking the histone code will take years of detailed and systematic studies. It remains a top priority for the epigenetics research community.

Tremendous strides have been made in understanding the impact of specific histone posttranslational modifications in the control of mammalian gene expression. To date, the best-understood histone tail modifications are lysine acetylation and lysine methylation.\textsuperscript{13,80} Each modification is catalyzed by an increasingly well-characterized group of “writers” that can then be interpreted by a distinct group of “readers.” For example, in mammalian cells, lysine acetylation is catalyzed by three families of histone acetyltransferases (HATs): GNAT, MYST, and CBP/p300. In general, they have poor specificity for individual histone tail lysine residues, and are targeted to relevant promoters in the context of large, multiprotein complexes. Indeed, the demonstration that classical transcriptional coactivators possess intrinsic HAT activity helped to establish histone lysine acetylation as a permissive, epigenetic mark strongly correlated with transcriptional activation.\textsuperscript{13,80} The removal of histone lysine acetylation is accomplished by four families of histone deacetylases (HDACs): class I (HDAC1–3, HDAC8), class II (HDAC4–7, HDAC9–10), class III sirtuins (SIRT1–7), and class IV (HDAC11). Experimentally, these families of HDACs are further categorized by their sensitivity to inhibition by the pharmacological agent, trichostatin A (TSA). Class I and II HDACs are TSA-sensitive, while class III and IV HDACs are TSA-insensitive. Like HATs, HDACs demonstrate poor specificity for individual lysine residues and are recruited to target promoters in large, multiprotein complexes. This targeting is accomplished by a group of acetyl lysine “readers” that contain a specific protein domain—the bromodomain.\textsuperscript{13,80} Importantly, HDAC inhibitors have found early clinical utility in the treatment of hematological malignancies, in particular, T-cell lymphomas.\textsuperscript{87} Their efficacy in the treatment of solid cancers,\textsuperscript{88} neurodegenerative disease,\textsuperscript{89} and immune-mediated pathologies,\textsuperscript{90} among a myriad of other human ailments, is eagerly being investigated.

Recent studies on histone lysine methylation underscore the tremendous complexity of epigenetic gene regulation. Unlike histone lysine acetylation, the downstream effects of lysine methylation on gene expression are residue-specific and context-dependent.\textsuperscript{13,80,91} For example, histone H3 lysine 4 (H3K4) is strongly associated with transcriptional activation, while H3K9 is a classical repressive mark associated with heterochromatin formation and transcriptional silencing. These epigenetic modifications are controlled by distinct groups of readers and writers. Lysine methylation biology is even more finessed. An individual histone lysine residue can be mono-, di-, or trimethylated with profound effects on mammalian gene expression. The functional relevance of these nuanced modifications is supported by their differential localization in the genome; for example, H3K4 monomethylation preferentially localizes to enhancers and H3K4 trimethylation to active promoters.\textsuperscript{13,92}

In addition to the dizzying complexity of histone posttranslational modifications, more recent studies have established important roles for nucleosome density,\textsuperscript{93,94} adenosine triphosphate–dependent chromatin remodeling,\textsuperscript{95} and the regulated, replication-independent incorporation of histone variants in the control of mammalian gene expression.\textsuperscript{96} For example, across all eukaryotic species, promoters and enhancers are more nucleosome-depleted than transcribed regions. Moreover, this state of relative histone depletion is dynamic and correlated with gene expression.\textsuperscript{97} These additional layers of epigenetic control provide an intricate and responsive system superimposed on the static genetic code.

RNA-based mechanisms

The central dogma of genetics states that DNA is transcribed to mRNA, which is in turn translated into proteins. However, the discovery of an extensive catalog of long RNA transcripts that do not code for proteins—lncRNAs—provides a new perspective on the importance of RNA in gene regulation (Figure 3).\textsuperscript{37,98} Of the three general modes of epigenetic regulation, RNA-based mechanisms are the most recently described and presently the least well understood. It is a matter of argument whether microRNAs fall into the definition of epigenetic regulation. They are considered by some to be epigenetic modifiers. Though not the focus of this review, readers are directed to work that has shown the involvement of microRNAs in gene regulation.\textsuperscript{99,100} MicroRNAs can exert an effect on gene transcription. For example,
DNA methylation states of the E-cadherin gene can be modified by RNA interference.\(^9\) This miRNA field will also be exciting to follow given their emerging role in the pathophysiology of stroke.\(^1\)\(^0\) As there are interesting overlaps between predominantly posttranscriptional regulators (eg, microRNAs) and transcription factors, similar overlaps in pathways are expected for epigenetic regulators, and readers are directed to a review in their role in various aspects of stroke.\(^1\)\(^0\)\(^1\)\(^0\) Amongst others, RNA-based mechanisms of epigenetic gene regulation can involve the coordinated activities of lncRNAs with other epigenetic activities, such as DNA methylation and histone posttranslational modifications.\(^1\)\(^0\)\(^1\)\(^0\) These lncRNAs are functionally distinct from small noncoding RNAs, such as microRNAs, that primarily mediate posttranscriptional repression in the cytoplasm. In mammalian systems, the most studied lncRNA is the 17-kb Xist nuclear RNA, which is expressed exclusively from the inactivated X chromosome (Xi) in women, and is essential for silencing one of the X chromosomes in somatic female mammalian cells.\(^1\)\(^0\)\(^4\)

Using chromatin-state maps, actively transcribed genes show unique signatures of H3K4 trimethylation at promoters and H3K36 trimethylation within transcribed regions, now termed K4-K36 domains.\(^1\)\(^0\)\(^5\) Using these K4-K36 domains to demarcate transcriptional units has led to the identification of thousands of lncRNAs in mammalian cells with broad cellular functions.\(^1\)\(^0\)\(^6\) The discovery and action of lncRNAs are an evolving story, as their cataloging and functions continue to expand. Initially, approximately 1600 lncRNAs were identified and current studies have expanded that to more than 8000 lncRNAs. Since the initial reports of the first lncRNAs in 2008 and 2009, emphasis has been placed on defining their functional interactions with chromatin-modifying complexes rather than their location with respect to other genes.\(^1\)\(^0\)\(^7\)

lncRNAs have been shown to direct both repressive and activating complexes and allow chromatin modifications to occur at specific loci (Figure 3).\(^5\)\(^6\)\(^6\)\(^0\)\(^1\)\(^0\)\(^8\) One such lncRNA, HOTAIR, was found to regulate the expression of developmental HOX genes, and has been implicated in promoting the invasiveness and metastasis of breast cancer.\(^3\)\(^8\)\(^1\)\(^0\)\(^9\)

**Figure 3** Long noncoding RNAs (lncRNAs) can act in cis or trans to activate or repress gene transcription through modification of chromatin.

**Notes:** In cis, the lncRNA is transcribed and recruits chromatin-modifying complexes to the site of transcription. For example, antisense noncoding RNA in the INK4 locus (ANRIL) recruits PRC2, which serves to repress the transcription of adjacent genes that have histones with the mark H3K27me3. In trans, lncRNAs act at a distal location from their transcription site, where they interact with target genes and chromatin-modifying complexes. The lncRNA HOTTIP acts at distal HoxA genes and interacts with the histone methyltransferase MLL to mediate the addition of trimethylation on lysine 4 of histone H3 (H3K4me3) to activate transcription.
**HOTAIR** functions in **HOXD** silencing by recruiting polycomb repressive complex 2 (PRC2) and its H3K27 trimethylation activity. The mechanism of how this directing occurs remains unclear. Although these RNAs were initially identified as long intergenic noncoding RNAs, this concept has become broadened to encompass IncRNAs that may exist near or antisense to protein-coding genes. **ANRIL** is an example of an IncRNA, as it is antisense to protein-coding genes but interacts with the chromatin-modifying complexes PRC1 and PRC2 to repress the target loci (see below). Additionally, IncRNAs can mediate transcriptional activation via recruitment of the H3K4 methyltransferase MLL1. One intriguing model is **HOTTIP** that acts at distal HOXA genes to activate transcription by mediating H3K4 trimethylation. Additional examples of IncRNA in mammals include **AIR** and **KCNQ1OT1**, which are involved in genomic imprinting, a process that mediates the expression of only one allele of a gene in a parent-of-origin-dependent manner.

It is clear that the clinical implications and the basic science discovery of IncRNAs are just emerging. A hint to their biological relevance, especially in the pathophysiology of stroke, may be found in unbiased GWAS results. GWAS hits also fall into large “gene deserts,” genomic regions with no protein-coding genes that are not artifacts and may have functional importance in complex disease. This is interesting given that approximately 80% of SNPs exist within intergenic and noncoding intronic regions, which make up a significant proportion of GWAS hits. The possibility that a gene can be implicated in the genetic cause of disease because it produces a functional RNA that is an epigenetic modifier is an emerging concept.

**An epigenetic theory of complex, non-Mendelian disease**

These epigenetic effectors are more than novel regulators of gene expression, but provide for a fundamentally different perspective on common human diseases, like stroke, with apparently strong genetic and environmental components. Epigenetics sits at the interface between the static genetic code and ever-changing environment with potentially long-standing influences on gene expression and inherited disease susceptibility. It provides a molecular framework for studying gene–environment interactions and thereby opens new avenues for greater understanding of human diseases and their treatment. Fundamental to this epigenetic perspective is the duality of the epigenetic code as both a dynamic and responsive landscape to environmental change and a potential carrier of heritable information.

Is there evidence to suggest that the discordance of MZ twins for stroke liability or the phenotypic variability of family members with CADASIL are determined by epigenetic mechanisms? The short answer is “maybe.” In 2005, Fraga et al performed an elegant experiment assaying for three epigenetic marks (DNA methylation and histone H3 and H4 acetylation) in 40 twin pairs of varying ages. Those MZ twins that demonstrated the greatest intra-twin epigenetic differences were older, spent more of their lifetime apart, and reported the largest differences in medical health. These data suggest that epigenetic differences can accrue over time and account for phenotypic differences, such as disease expression, later in life. An analogous observation is made in genetically identical mammals cloned by somatic cell nuclear transfer. While some offspring are phenotypically normal, a majority are runted, a phenomenon at least partially explained by the inefficient epigenetic reprogramming of transplanted nuclei. These data are compelling.

And what of the missing heritability of GWASs? If an epigenetic explanation could account for this missing heritability, there would have to be evidence for transgenerational inheritance of an epimutation. Although this concept remains contentious and hotly debated, these data are accruing in animal models, with a few notable examples in man.

Epigenetic theory encompasses molecular pathways that regulate the structure and accessibility of chromatin without changing the genetic code. The three main effectors of epigenetics are DNA methylation, histone density and posttranslational modifications, and RNA-based mechanisms, such as IncRNAs. In the next section, we illustrate and discuss the roles of these three effector pathways using case studies of two genes—endothelial nitric oxide synthase (eNOS) and **ANRIL**. eNOS is a protein-coding gene that is regulated by epigenetics, while **ANRIL** is an IncRNA that mediates epigenetic regulation.

**Case study 1**

**eNOS:** from molecular cloning to a model of vascular epigenetic gene regulation

Given that endothelial cells (ECs) figure prominently in the pathophysiology of large-vessel atherosclerosis, it is important to understand how gene expression in these cells becomes perturbed in disease (Figure 4). We do not discount the contribution of other pathophysiological factors beyond blood-vessel pathology, especially the coagulation pathway, complement, and platelets, among others. Indeed, the involvement of genetic and epigenetic influences in these complementary events has been well argued. The **eNOS/NOS3** gene is very restricted in expression to ECs, where it provides the dominant source
of nitric oxide in the vasculature. Nitric oxide is a vasodilator that plays a vital role in the maintenance of vascular homeostasis, by the virtue of potent antithrombotic and anti-atherogenic activities. The generation of nitric oxide by nitric oxide synthases maintains the vasculature in the vessels of the brain to inhibit ischemic or hemorrhagic stroke. It does so via the inhibition of platelet aggregation and leukocyte-endothelium adhesion, inducing vasodilation and decreasing VSMC proliferation. A large facet of stroke pathology is the attenuation of large vessels caused by atherosclerosis; therefore, it is important to understand changes in eNOS regulation during endothelial dysfunction. For example, eNOS mRNA and protein levels are decreased in endothelial cells overlying atherosclerotic plaques. Seminal work in dyslipidemia, a risk factor for stroke, has shown strong evidence that low-density lipoprotein levels play a key role in the induction of atherosclerotic lesions and decreased nitric oxide availability. Changes in eNOS expression occur at characteristic locations in the vasculature due to changes in local hemodynamic forces, such as the curvature of the aorta and branch points. Examination of the mouse aorta at regions predisposed to atherosclerosis shows decreased eNOS activity and an increase in priming of the nuclear factor kappa B signal-transduction cascade. The regulation of vascular homeostasis is subject to nitric oxide availability through contributions of eNOS, the expression of which is abnormal in disease.

Studies conducted in vivo have shown that deregulation of eNOS contributes to vessel attenuation and also contributes to postinfarction pathology. Studies have shown that eNOS-deficient mice show decreased neovascularization, larger infarcts, and smaller penumbral regions, although the mechanisms of how this occurs are still unclear. This contrasts with studies done in neuronal NOS (nNOS)-deficient mice. nNOS−/− mice show smaller cerebral infarctions, although blood flow is equivalent in nNOS knockout and wild-type mice. This suggests a potential role for nNOS, but not eNOS, in tissue damage after cerebral ischemia. Clearly, the nNOS gene is relevant to global hypoxia and anemia-induced cerebral hypoxia, as recently shown by Tsui and colleagues. Deficiencies in the inducible form of NOS – iNOS – have revealed no changes in infarction size after cerebral ischemia. Interestingly, iNOS shows changes in mRNA splicing after infarction as differential expression of exons 2 and 3 are seen. It remains unclear as to the functional effects of these changes.

In humans, there are a number of EC-enriched genes in addition to eNOS, such as von Willebrand factor (VWF), CD31, vascular endothelial growth factor 2 (VEGFR2),
and intercellular adhesion molecule 2 (ICAM2), among others. However, no master transcriptional control factor has been found in endothelial cells. Although some trans factors are relatively EC-enriched, such as ETS and GATA factors, these factors alone cannot explain the EC-enriched expression of these target genes. This contrasts with other terminally differentiated adult cell types that have master transcription-factor regulators, such as PPAR and MYOD, in adipocytes and skeletal muscle cells, respectively. Thus, the typical cis/trans paradigms may not be applicable to the restricted gene expression in EC. Epigenetic regulation has provided an interesting explanation of cell-specific expression in endothelial cells, particularly through the study of eNOS. Molecular characterization of the eNOS gene started with its cloning, which revealed a protein-coding unit with 26 exons spanning 21 kb of genomic DNA that produces a 4052-nucleotide mRNA. The eNOS promoter does not contain a TATA box or a CpG island, but it does contain two regulatory domains termed positive regulatory domains I and II at −104/−95 and −144/−115, respectively, relative to the transcription start site. The eNOS promoter contains cis binding sites for the Ets family members, GATA, SP1, AP-1, and YY1, which are common between endothelial enriched genes. Remarkably, these transcription factors are not endothelial-restricted, in contrast to eNOS expression. The finding that only nonendothelial restricted transcription factor–binding sites were detected at the promoter of eNOS is an important clue that the simplistic view that cis/trans interactions may not suffice for endothelial gene enrichment.

The initial discovery of the role of epigenetics in eNOS gene expression came from the use of episomal eNOS promoter–reporter constructs. These constructs showed in vitro activity regardless of cell type while native eNOS mRNA was EC-enriched. In contrast, stably integrated eNOS promoter–reporter transgenic mice recapitulated eNOS expression patterns in humans. Since the expression of the eNOS promoter reporter in transgenic mice is dependent on the chromatin structure of its integrated site, this suggests that epigenetics plays a role in the endothelial-enriched expression of eNOS. Indeed, the first direct evidence of eNOS epigenetic regulation is DNA methylation. DNA methylation plays a role in eNOS transcription, as suggested by the finding that the eNOS promoter showed DNA hypomethylation in ECs, but dense methylation in nonexpressing cells, such as VSMCs. DNA methylation abrogated synergistic binding of Sp1, Sp3, and Ets1 to the eNOS promoter. Moreover, recovery of eNOS expression was observed in HeLa cells and VSMCs after treatment with the DNA methylation inhibitor 5-azacytidine. Additional studies of epigenetic mechanisms indicated that the histone code plays an active role in the proper function of the eNOS gene in ECs. Interrogation of various histone marks at the eNOS promoter showed the presence of acetyl H3 and H4 marks and trimethylated lysine 4 of H3, specifically H3K9 and H4K12, which are hallmark of actively transcribed chromatin in ECs (Table 1). In contrast, non-ECs lacked these histone posttranslational modifications. Indeed, repression of eNOS in non-ECs is abrogated by the addition of trichostatin A, a histone deacetylase inhibitor. Taken together, eNOS expression is regulated, in part, by epigenetics. Thus, it follows that epigenetics may play a role in disease pathology where eNOS expression is perturbed.

Does the environment affect gene expression in ECs via epigenetic pathways? Newer evidence suggests this may be the case. eNOS mRNA and protein decreases in hypoxic cells; both in vitro and in vivo. In acute hypoxia, the eNOS promoter shows a decrease in histone-activating marks such as acetylation and lysine-4 methylation, which is mediated by the loss of histone proteins due to histone eviction. Under chronic hypoxia, the histone octamers return to normal levels, but are not modified by the activating marks, that are observed under normoxic conditions. It is worth highlighting that hypoxia upregulated the expression of an antisense eNOS gene, sONE (eNOS antisense, NOS3AS) that regulates eNOS expression post-transcriptionally. Little is known about this noncoding gene. The sONE gene overlaps the 3′ end of the eNOS gene and appears to regulate eNOS mRNA through a posttranscriptional mechanism, different from that of the antisense gene ANRIL (see below).

Overall, eNOS is a remarkable case study of the fundamentals of epigenetic regulation. In a quiescent endothelium, eNOS regulation is maintained by several aspects of epigenetics, while it can change during disease states. Moreover, this gene provides valuable insight into an epigenetic framework that can help to explain the heritability of complex diseases such as ischemic stroke. The non-Mendelian inheritance of atherosclerosis and stroke together with the contribution of eNOS gene regulation argues that epigenetics can be a participant in stoke burden.

**Case study 2**

**ANRIL: a long noncoding RNA and epigenetic risk stratification in large-vessel ischemic stroke**

ANRIL was first identified by a germ-line deletion in a melanoma-neural system tumor family in an unrelated
The discovery of the **INK/ARF** locus. This discovery happened only months prior to the identification of a coronary artery disease (CAD) risk locus in the genomic region that traverses the **ANRIL** lncRNA. Four independent groups reported this genetic association. This CAD risk locus sits in a region that is devoid of protein-coding genes, but is located within the 3’ end of the **ANRIL** gene. The identification of a CAD risk locus spurred investigations into a potential association with stroke due to the cardiovascular-related phenotype. In 2008, Bilguvar and colleagues were the first to identify an association of intracranial aneurysm and an SNP located on chr9p21 within the **ANRIL** locus. Although further examination of associations between stroke and this risk locus showed conflicting results, Anderson et al. established the likely basis for this confusion by demonstrating that chr9p21 variants are risk factors for ischemic stroke related to large-artery atherosclerosis but not other pathomechanistic stroke subtypes. Furthermore, a meta-analysis of one SNP (rs10757274) within the locus showed an association with ischemic stroke with a modest effect size (OR, 1.11 [95% CI, 1.05–1.17, \( P = 0.0001 \))). Moreover, it is expressed in tissues and cell types relevant to stroke, such as adult human brain tissue and blood vessels (endothelial cells, vascular smooth muscle cells), making it a prime candidate for involvement in cardiovascular disease, but full functional analysis remains to be performed. Knockdown of the **ANRIL** lncRNA in vascular smooth-muscle cells has been recently demonstrated to have a profound effect on gene expression in this cell type. Moreover, epigenetic mediators are being tested as potential targets for therapy in stroke. Readers are referred to recent reviews discussing these findings.Taken together, the evidence suggests that the chr9p21 haplotype influences stroke risk.

During large-scale GWASs at chr9p21 for stroke associations, molecular analysis was conducted on the **ANRIL** lncRNA transcript. The study of the RNA was elusive due to complexities in its splicing. **ANRIL** RNA has been found to form multiple rare variants from alternative splicing. Additionally, these variants showed linear and circular forms. Subsequent studies suggest that **ANRIL** functions through the *cis* recruitment of chromatin-modifying complexes – specifically, PRC1 and PRC2 – to neighboring genes of the **INK4b/ARF/INK4a** locus. This locus spans approximately 42 kb on human chromosome 9p21 and is an important regulator of cellular senescence. These genes act through the regulation of retinoblastoma protein, cyclin-dependent kinases, and p53 signaling. Regulation of this locus is mediated by H3K27 trimethylation catalyzed by PRC2 and recognized by PRC1 that results in silencing of transcription. PRC1 and PRC2 are recruited to the locus by the lncRNA **ANRIL**, expressed antisense to **ARF/INK4b**. It is believed that **ANRIL** transcription and secondary RNA structure formation are able to interact and direct PRC1 and PRC2 to the adjacent **INK4b/ARF** locus to mediate the establishment of repressive marks H2AK119 and H3K27me3, respectively (Figure 3). The presence of these epigenetic marks is associated with chromatin compaction and transcriptional repression of these protein-coding genes. More importantly, the previously discussed SNP appears to play a molecular role in **ANRIL** function. **ANRIL** alleles harboring the disease-associated SNP are associated with a reduction in **INK4b, INK4a, and ARF** mRNAs. One theory is that the presence of the disease-associated SNP allele changes the abundance or function of **ANRIL** splice variants, resulting in their reduced ability to exert full repression on the **INK4/ARF** locus. This work suggests that SNPs within the risk locus may play a role in **ANRIL** function.

**ANRIL** has become a model for the role of epigenetics in complex, non-Mendelian diseases such as stroke, and CAD in general. Although we are far from a complete understanding of **ANRIL** function in stroke, it is of great interest that this epigenetic modifier represents an exciting newer perspective on stroke pathophysiology.

**Concluding remarks: the emerging field of stroke epigenetics**

This review has primarily considered the potential of epigenetic pathways to better understand stroke liability, both inherited and as the result of environmental exposures. A related field of study has emerged that focuses on the role of epigenetic pathways in response to cerebral ischemia and their modulation as novel targets for therapeutic intervention. For example, Endres et al. have demonstrated increased levels of DNA methylation after ischemia/reperfusion in a mouse model of mild focal brain ischemia. Genetic reduction of DNMT levels and treatment with a pharmacological DNMT inhibitor were neuroprotective. However, this neuroprotective effect was not observed in severe stroke/excitotoxic cell death, underscoring the complexity of these pathways. Similar studies have investigated histone acetylation pathways in rodent models of focal cerebral ischemia. These studies highlighted the potential of pharmacological HDAC inhibitors to serve as powerful neuroprotective agents, and were recently reviewed by Langley et al. Less well studied but equally compelling is the role of noncoding RNAs in ischemic stroke pathogenesis. While the use of
epigenetic modifiers in ischemic stroke is enticing, caution is warranted, as our understanding of these mechanisms is still in its infancy. The interested reader is referred to an excellent three-part review by Qureshi and Mehler on the emerging role of epigenetics in stroke pathogenesis. Finally, it must be noted that ischemic stroke pathogenesis is complex and multifactorial, implicating a diversity of cell types and molecular processes. While we have focused on vascular-wall pathology in this review, epigenetic concepts also apply to other facets of ischemic stroke pathogenesis, eg, platelet aggregation, thrombus formation, and cholesterol homeostasis. These topics are beyond the scope of this manuscript and expertly reviewed elsewhere.

Epigenetic pathways offer a new perspective in the control of gene regulation with relevance to human cardiovascular disease and stroke research. As epigenetic processes are dynamic and respond to environmental cues, they provide the molecular substrate for understanding and experimentally verifying gene–environment interactions. The observation that a large number of GWAS hits fall within regions of the genome without classical coding function highlights the importance of epigenetics in human disease, including stroke. The ANRIL case is the best example of this phenomenon in human cardiovascular disease. Valuable insight has also been gained by studies exploring the epigenetic regulation of the eNOS gene in response to physiological and pathophysiological stimuli, including hypoxia. As we expand our understanding of epigenetics, the hope is that newer insight will be gained in prevention, diagnosis, and treatment of stroke. This is important because of its implications for clinical diagnosis and therapeutic intervention in vascular diseases, especially in large-vessel ischemic stroke.

Acknowledgments
We gratefully thank members of the Marsden lab for critical review of the manuscript. Philip Marsden holds the Keenan Chair in Medical Research and is supported by operating grants from CIHR MOP 79475 and NIH/HHLBI P01 HL076540-06A1. Charles Matouk gratefully acknowledges support from Yale University, jointly from the School of Medicine and Department of Neurosurgery, and the Brain Tumor Gift Fund.

Disclosure
The authors report no conflicts of interest in this work.

References
11. Lakhteeb MB, Dichgans M, Hegele RA. Epigenetic modifiers in ischemic stroke is enticing, caution is warranted, as our understanding of these mechanisms is still in its infancy.


