Human Molecular Genetics, 2006, Vol. 15, No. 16 **2468–2478**
doi:10.1093/hmg/ddl169
Advanced Access published on July 11, 2006

Phosphodiesterase 4D polymorphisms and the risk of cerebral infarction in a biracial population: the Stroke Prevention in Young Women Study

Qing Song[1,†], John W. Cole[2,3,*,†], Jeffrey R. O'Connell[2], Oscar C. Stine[2], Margaret Gallagher[4], Wayne H. Giles[4], Braxton D. Mitchell[2], Marcella A. Wozniak[2,3], Barney J. Stern[2,3], John D. Sorkin[2,3], Patrick F. McArdle[2], Adam C. Naj[2,5], Qin Xu[1], Gary H. Gibbons[1] and Steven J. Kittner[2,3]

[1]Morehouse School of Medicine, Atlanta, GA 30314, USA, [2]University of Maryland School of Medicine, Baltimore, MD 21201, USA, [3]Veterans Administration Medical Center, Baltimore, MD 21201, USA, [4]Centers for Disease Control, Atlanta, GA 30333, USA and [5]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA

Received March 14, 2006; Revised June 29, 2006; Accepted July 2, 2006

An association between polymorphisms within the phosphodiesterase 4D gene (*PDE4D*) and ischemic stroke was initially reported in older adults from Iceland and has been supported by studies in several other primarily elderly populations. In the present study, we examined the association between *PDE4D* polymorphisms and early-onset ischemic stroke in a biracial female population. A systematic search for polymorphisms in the highly evolutionary conserved regions of *PDE4D* was performed on 48 African-American and 48 Caucasian participants. Novel and known polymorphisms were then prioritized and genotyped in the entire study population of 224 cases of first ischemic stroke among women aged 15–49 and 211 age- and ethnicity-balanced control subjects. Forty novel and previously reported polymorphisms with a minor allele frequency greater than 0.05 were determined, with 23 polymorphisms selected for analysis in the full case–control sample. Single nucleotide polymorphism (SNP), linkage disequilibrium and haplotype analyses were performed. SNP rs918592, found in an intron near the 5′ end of the gene, was significantly associated with stroke (age- and race-adjusted odds ratio (OR = 1.5, *P* = 0.007), with four other SNPs showing significant, albeit less strong, associations. The magnitude of association was similar across African-Americans and Caucasians and across multiple stroke subtypes (e.g. atherosclerotic, lacunar and non-lacunar of undetermined etiology). The association of rs918592 with stroke was confined exclusively to current smokers (OR = 3.2, *P* = 0.00014), with no association observed among never-smokers (OR = 0.9, *P* = 0.75) or former smokers (OR = 1.2, *P* = 0.66), demonstrating a gene–environment interaction (*P* = 0.03). A strong dose–response relationship was also seen among current smokers. No specific risk haplotypes were identified.

INTRODUCTION

The phosphodiesterase 4D gene (*PDE4D*) has been implicated in the etiology of stroke, particularly the atherosclerotic and cardioembolic subtypes, based on linkage and association analyses carried out by deCODE Genetics in the Icelandic population (1). The human *PDE4D* gene spans a 1.6 Mb region on chromosome 5q12 and contains 24 exons (OMIM 600129). Through differential promoter and alternative splicing in gene expression, the gene can express nine different functional protein isoforms (Fig. 1). The associated single nucleotide polymorphisms (SNPs) are located among the first exons at the 5′ end of the gene. The association with stroke is hypothesized to be mediated through cAMP, a key signal transduction

*To whom correspondence should be addressed at: Maryland Stroke Center, University of Maryland School of Medicine, 655 West Baltimore Street, Room 12-006, Baltimore, MD 21201, USA. Tel: +1 4107060414; Fax: +1 4107060816; Email: jcole@som.umaryland.edu
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

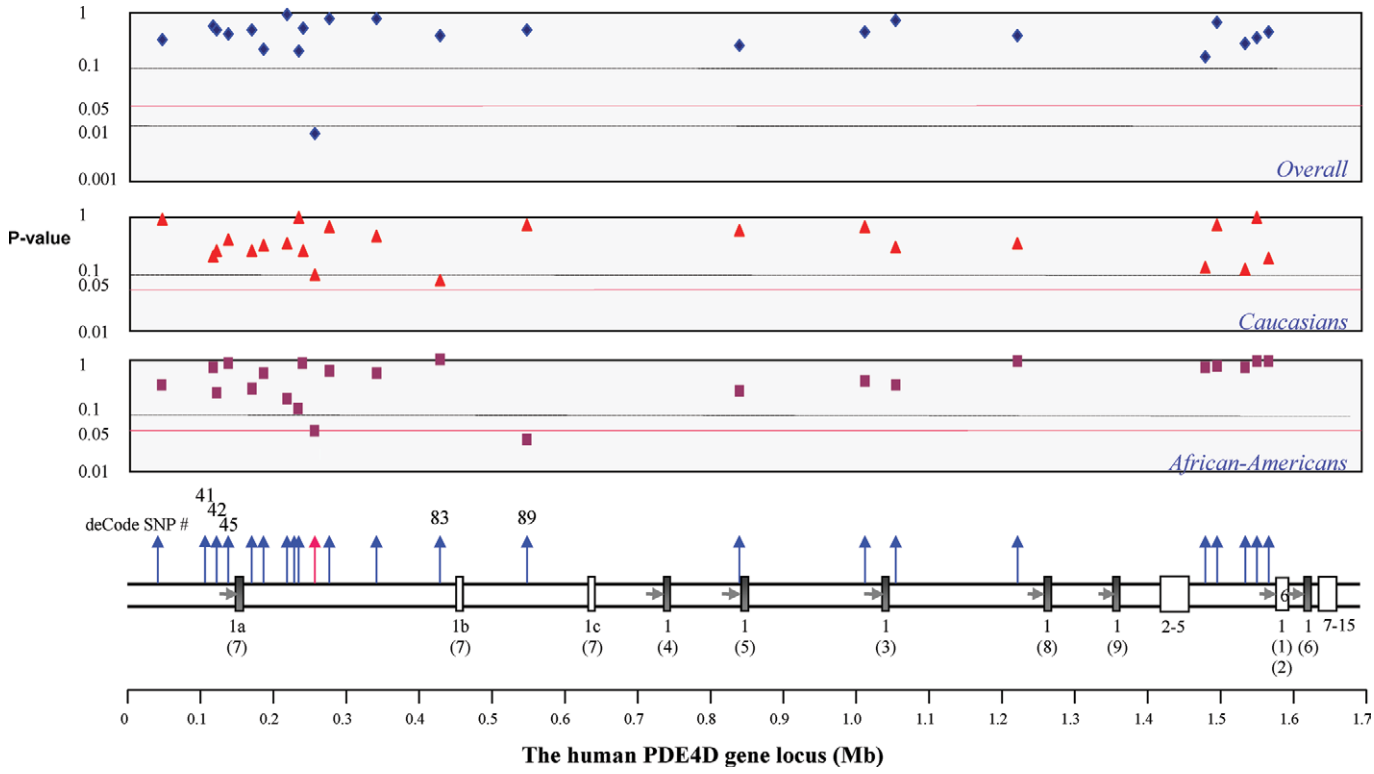

Figure 1. Association mapping of the PDE4D gene polymorphisms with ischemic stroke in the SPYW. Age-adjusted *P*-values among the overall population (race adjusted), Caucasians and African-Americans for all SNPs analyzed using an additive model. The PDE4D gene structure is shown at the bottom. Solid boxes indicate the first exons for various isoforms, other exons are indicated by open boxes. Exons 2–5 are shown together, and exons 7–15 are shown together. Sideways arrows indicate promoters. Parenthesis indicate isoforms. The upward arrows on the gene demonstrate the SNP position within the gene as listed in Table 2.

molecule in multiple tissues and cell types, including vascular cells (2). *PDE4D* selectively degrades cAMP, which exerts protean effects on the vasculature and nervous system.

Other linkage studies to the *PDE4D* region of chromosome 5q12 have been mixed, with a Swedish study (3) replicating the Icelandic finding, but an American study (4) reporting a negative result. Several studies have examined the deCODE markers in the region of *PDE4D* where associations were found in Iceland. Some (3–8) but not all (9,10) studies have found associations in this region to all ischemic stroke or some stroke subtypes. Most (3,4,6,9) but not all (5,8,10) studies have examined predominantly older and Caucasian populations and, with one exception (3), have examined only deCODE markers. Furthermore, no prior study had tested for gene–environment interactions. Therefore, we sought to extend the genetic epidemiology of *PDE4D* and stroke risk by examining a biracial case–control population of early-onset stroke, by performing novel SNP discovery and by assessing interactions with environmental exposures.

RESULTS

Subject characteristics

Demographic and risk factor characteristics by case–control status are described in Table 1. The mean age of the cases and control subjects was 41.7 and 39.6 years, respectively.

Table 1. Characteristics by case–control status

	Cases ($n = 224$)	Controls ($n = 211$)	*P*-value
Mean age (years)	41.7	39.6	0.0026
African-American (%)	47.3	43.1	0.579
Hypertension (%)	41.1	14.2	<0.0001
Diabetes mellitus (%)	17.9	6.2	0.0002
Current smokers (%)	47.8	23.7	<0.0001
Angina/MI (%)	11.6	2.8	0.0005
OCP (%)[a]	12.2	6.2	0.032

[a]Two cases and one control could not recall their last OCP use, therefore cases $n = 222$ and controls $n = 210$.

Cases were significantly more likely to have a history of hypertension ($P < 0.0001$), diabetes ($P = 0.0002$) and angina/MI ($P = 0.0005$) and to currently smoke cigarettes ($P < 0.0001$) and to report the use of oral contraceptive pills (OCP) within the month (31 days) prior to their stroke ($P = 0.032$).

Novel SNP discovery on the non-coding conserved segments of the PDE4D gene

We identified a total of 71 polymorphisms (66 SNPs and five insertion/deletion polymorphisms). Of these, 40 were confirmed to have a minor allele frequency ≥0.05 in controls of

at least one population with 19 observed in African-Americans only, four in Caucasians only and 17 in both races. Many of these SNPs were not listed in public or private databases [National Center for Biotechnology Information (NCBI), Celera's Human SNP Reference Database, The SNP Consortium, Human Genome Variation Database and others].

Twenty-three SNPs were genotyped in the entire case–control sample based upon the SNP selection criteria detailed in Materials and Methods. In brief, SNPs were chosen based upon their location within the *PDE4D* gene, with a special emphasis on the regions around the first exons at the 5′ end of the gene, and based upon an MAF ≥ 0.05 among control subjects. Table 2 lists the 23 genotyped SNPs as ordered by their physical position within the gene, and additionally provides the allelic variants, genotype call rates and minor allele frequencies among cases and controls stratified by race. Seventeen SNPs had a minor allele frequency ≥ 0.05 among controls of both races. There were three SNPs unique to Caucasians and three SNPs unique to African-Americans that satisfied this criterion. All genotyped SNPs were in Hardy–Weinberg equilibrium in the combined case/control sample stratified by race.

SNPs associated with stroke

Minimally adjusted model. Figure 1 illustrates the *P*-values from the additive model (primary analysis) of the 23 SNPs and their position within *PDE4D*. Results are shown for the entire population with adjustment for age and race and for African-American and Caucasian women separately with adjustment for age only. In the entire population, one SNP was significantly associated with stroke, rs918592 (OR = 1.49, $P = 0.007$). Age-adjusted race-stratified analyses were not consistently significant, although the risk allele for SNP rs918592 (the A allele) showed a similar degree of association in both races (African-American OR = 1.50, $P = 0.029$ and Caucasian OR = 1.52, $P = 0.08$). The A allele occurred more frequently in African-Americans (frequencies among cases and controls are 0.66 and 0.55, respectively) than in Caucasians (frequencies among cases and controls are 0.25 and 0.18, respectively). These findings are detailed in Tables 2 and 3. Using the GG genotype as the reference group, the age- and race-adjusted genotype effect of the AG heterozygotes risk allele demonstrated no stroke risk (OR = 1.21, $P = 0.51$) but significant stroke risk among AA homozygotes (OR = 2.22, $P = 0.014$). Age-adjusted race-specific genotype risk estimates demonstrated increased risk among African-Americans (AG: OR = 1.18, $P = 0.63$; AA: OR = 2.73, $P = 0.024$) but not among Caucasians (AG: OR = 1.08, $P = 0.91$; AA: OR = 1.77, $P = 0.39$).

As detailed in Materials and Methods, the permutation method was used to determine the multiple comparison-corrected *P*-value for the age–race-adjusted association of rs918592 with stroke. When the 17 SNPs with minor allele frequency greater than 0.05 in both races were considered, the empirically determined 'experiment-wide' *P*-value was 0.08. When this subset was further reduced to include only our eight genotyped SNPs that were within the stroke associated Icelandic haplotype block (block B) (1), the empirically determined experiment-wide *P*-value was 0.03. A similar analysis that included our four genotyped SNPs in Icelandic block A (1) also yielded an experiment-wide *P*-value was 0.03.

Vascular model. Further analyses evaluating SNP rs918592 in an additive model controlling for age, smoking, OCP, HTN, DM and angina/MI (see Materials and Methods section) are shown in Table 3. The relationship was slightly attenuated (OR = 1.38, $P = 0.052$).

Table 4 shows the risks among all participants (Caucasian, African-American and all others) for SNP rs918592, as stratified by age and other stroke risk factors with associated adjusted ORs, *P*-values and 95% confidence intervals. We found that this SNP was not associated with stroke among never-smokers (OR = 0.93, $P = 0.75$) or former smokers (OR = 1.16, $P = 0.66$), but was highly associated with stroke among smokers (OR = 3.22, $P = 0.00014$). Further analyses demonstrated a strong dose–response relationship between stroke risk and smoking quantity. Among light smokers (1–10 cigarettes per day), the risk genotype had a weaker association with stroke (OR = 2.27, $P = 0.033$), whereas among heavy smokers (11+ cigarettes per day), the risk genotype has a stronger association with stroke (OR = 8.39, $P = 0.0004$). These results are demonstrated in Figure 2. The smoking by genotype interaction term was significant ($P = 0.03$) in a model including the covariates of age, race, hypertension, diabetes mellitus and oral contraception. The risk genotype was not significantly associated with any of the stroke risk factors.

Of the remaining 22 SNPs evaluated using these and other models, we found that several additional SNPs were associated with stroke, including: rs966221 (SNP 83) and rs1396476 (SNP 89) under an additive model, rs153031 (SNP 42) and rs918592 under a dominant model and rs1498606 under a recessive model. Details regarding these results are provided in Supplementary Material, Table S1. Risk factor stratified analyses of these associated SNPs are provided in Supplementary Material, Table S2.

Stroke subtype analyses

In an exploratory analysis, stroke subtypes were analyzed separately. Table 5 demonstrates the results for SNP rs918592 as stratified by stroke subtype with associated adjusted ORs, *P*-values and 95% confidence intervals. Similar analyses regarding the other associated SNPs are shown in Supplementary Material, Table S3. In brief, the atherosclerotic group included 27 cases with either probable or possible atherosclerotic mechanism, the cardiac group included 14 cases with a probable cardiac source of embolism, the probable dissection group included 13 cases confirmed by neuroimaging, the lacunar group included 45 cases of symptomatic small deep lesions on neuroimaging studies or classic lacunar syndromes regardless of other potential causes and the probable hematologic group included nine cases. These categories were not mutually exclusive. There were 125 non-lacunar stroke cases of undetermined etiology. SNP rs918592 was significantly associated with all stroke subtypes except the cardiac and hematologic subtypes. The analysis for the dissection subgroup was indeterminate as the model estimates did not converge.

Table 2. SNPs analyzed in complete case–control study population including allelic variants, position, genotype call rate and minor allele frequencies as stratified by race and case/control status

SNP name or rs number, alleles	Position[a]	Call rate	Allele frequency[b]/*n*					
			African-American			Caucasians		
			Cases	Controls	*P*-value[c]	Cases	Controls	*P*-value[c]
PDE4D-32913 (Novel), C/**T**[d]	32913	85	0.235/78	0.272/76	0.343	0.589/82	0.568/74	0.918
rs152312 (Decode 41), C/**T**	144510	99	0.028/102	0.044/88	0.246	0.133/93	0.085/99	0.259
rs153031 (Decode 42), C/**T**[d]	145035	93	0.353/75	0.338/65	0.851	0.607/75	0.657/70	0.413
rs12188950 (Decode 45), C/**T**	149009	99	0.129/101	0.137/89	0.750	0.163/94	0.116/98	0.207
rs27224, G/**T**	154835	99	0.033/102	0.050/89	0.293	0.133/93	0.085/99	0.259
rs153067, **C**/T[d]	168927	99	0.355/83	0.358/71	0.555	0.612/76	0.669/74	0.322
rs42222, **A**/C	216712	99	0.043/101	0.079/87	0.186	0.104/90	0.066/98	0.366
PDE4D-229902 (Novel), G/**A**[d]	229902	97	0.245/100	0.170/86	0.130	0.011/95	0.005/86	0.976
rs918590, **G**/T	234518	95	0.462/99	0.463/83	0.833	0.246/91	0.200/98	0.264
rs918592, A/**G**	235023	97	0.340/102	0.455/89	0.056	0.750/94	0.825/99	0.100
PDE4D-249037 (Novel), C/**T**[d]	249037	90	0.282/75	0.271/69	0.615	0.040/75	0.052/67	0.706
PDE4D-327964 (Novel), **A**/G[d]	327964	93	0.247/94	0.222/85	0.542	0.006/85	0.016/90	0.467
rs966221(Decode 83), C/**T**	429806	95	0.495/101	0.489/88	0.990	0.608/91	0.525/98	0.078
rs1396476 (Decode 89), **G**/T	535684	99	0.152/101	0.073/88	**0.034**	0.145/92	0.157/98	0.716
PDE4D-866896 (Novel), **G**/A[d]	866896	95	0.026/76	0.044/65	0.273	0.207/75	0.225/71	0.371
PDE4D-1024373 (Novel), **A**/G[d]	1024373	79	0.119/62	0.161/53	0.410	0.083/66	0.070/64	0.676
rs294496, C/**T**[d]	1053899	78	0.213/78	0.261/68	0.337	0.233/73	0.187/74	0.303
rs525099, **C**/T[d]	1242015	78	0.187/76	0.201/66	0.924	0.336/74	0.395/80	0.349
PDE4D-1483844 (Novel), C/**T**	1483844	83	0.020/97	0.023/83	0.748	0.051/88	0.098/91	0.133
PDE4D-1483907 (Novel), **A**/G[d]	1483907	84	0.121/86	0.111/78	0.703	0.218/78	0.195/81	0.765
rs929820, **C**/T[d]	1538994	95	0.380/96	0.447/83	0.721	0.710/87	0.643/97	0.129
PDE4D-1568979 (Novel), C/**T**[d]	1568979	94	0.064/98	0.077/82	0.908	0.006/83	0/93	0.990
rs1498606, C/**T**	1594777	95	0.257/99	0.202/87	0.908	0.176/90	0.116/98	0.189

[a]Relative position (148436–148557 exon 1a) (7).
[b]As listed by minor allele frequency in African-Americans (bolded in column 1).
[c]In race-stratified age-adjusted additive model, (primary analyses) as denoted in Fig. 1.
[d]SNP genotyped using the UHT system.

Table 3. SNP rs918592—additive model

Race	Controls/cases[a]	Age-adjusted				Vascular model (age, smoking, OCP, HTN, DM, angina/MI)			
		OR	95% CI		*P*-value	OR	95% CI		*P*-value
All[b]	**207/219**	**1.49**	**1.11**	**2.00**	**0.007**	1.38	0.99	1.90	0.052
African-American	**89/102**	1.50	0.99	2.26	0.029	1.40	0.89	2.20	0.146
Caucasian	**99/94**	1.52	0.93	2.48	0.080	1.35	0.78	2.31	0.284

Age-adjusted and vascular models adjusted ORs, 95% confidence intervals and *P*-values. Significant results ($P < 0.05$) are in bold.
[a]Indicates number of controls and cases used in crude analyses.
[b]Includes Hispanics, Asian and all others. Adjusted for race (black, white and others).

Haplotype analyses

Figures 3 and 4 show the haplotype block structure among African-Americans and Caucasians, respectively. High correlations existed within both races between eight SNPs (excluding SNP PDE4D-229902), which were in the promoter 1a isoform region of the gene, between SNP rs152312 (SNP 41) and rs918592. Of note, these figures demonstrate that linkage disequilibrium exists between SNPs rs918592 [associated with stroke in the Stroke Prevention in Young Women Study 2 (SPYW2) population] and rs152312 (SNP 41, associated with stroke in the Iceland population) that is stronger among African-Americans (LD = 1.0) than Caucasians (LD = 0.66). Our stroke-associated SNPs rs918592 and rs153031 (SNP 42) (Supplementary Material, Tables) (Figs 3 and 4) also demonstrate varying linkage disequilibrium among African-Americans (LD = 0.59) and Caucasians (LD = 0.95).

Haplotype analyses including the eight SNPs identified through Haploview (as described in preceding paragraph) were performed among all SPYW2 participants, among the Caucasian and African-American subgroups and by stroke subtype. After correction for multiple comparisons testing, these analyses failed to demonstrate additional associations.

DISCUSSION

Replication studies of the Icelandic findings have yielded mixed results. Figure 5 shows the SNPs examined by all

Table 4. Among all participants, SNP rs918592—additive model, stratified by age and other stroke risk factors demonstrating adjusted ORs, 95% confidence intervals and *P*-values

rs918592—additive						
Risk factor[a]	Strata	Controls/cases	OR[a]	95% CI		*P*-value
Age	**14–39**	79/57	0.95	0.54	1.67	0.871
	40–49	128/162	**1.85**	**1.29**	**2.65**	**0.001**
OCP use	−	193/191	**1.55**	**1.14**	**2.10**	**0.005**
	+	13/26	1.20	0.33	4.35	0.787
Smoking	Never	115/82	0.93	0.60	1.46	0.752
	Former	44/36	1.16	0.61	2.21	0.658
	Current	48/101	**3.22**	**1.76**	**5.87**	**0.00014**
HTN	−	178/127	**1.79**	**1.24**	**2.57**	**0.002**
	+	29/92	0.90	0.49	1.64	0.714
DM	−	193/181	**1.57**	**1.14**	**2.15**	**0.005**
	+	14/38	0.71	0.25	2.00	0.513
Angina/MI	−	202/195	**1.53**	**1.13**	**2.08**	**0.006**
	+	5/24	NC	NC	NC	NC

Significant results ($P < 0.05$) are in bold. NC indicates insufficient sample size to perform calculation; model did not converge.
[a]Each stroke risk factor adjusted for age and race (black, white and others).

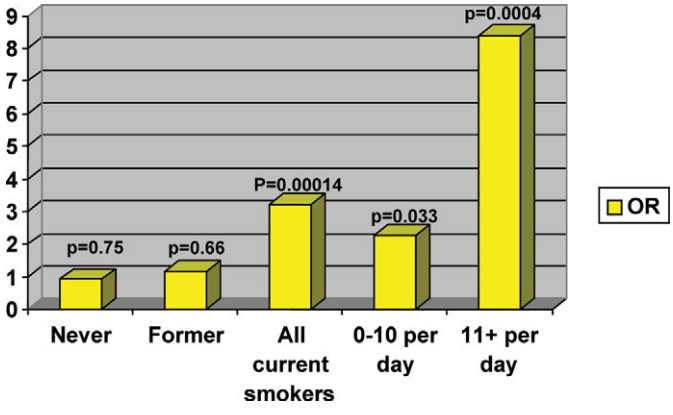

Figure 2. ORs and *P*-values for SNP rs918592 among never-, former-, all current-smokers and among light smokers (1–10 per day) and heavy smokers (11+ per day) demonstrating a strong dose–response relationship.

published studies of *PDE4D* and stroke, demonstrating that work has primarily focussed on the portion of the gene where deCODE associations were found.

Four European studies exclusively examined the strongest genotype and haplotype associations with all ischemic stroke or the combined atherosclerotic/cardioembolic group from the Iceland study and failed to replicate these findings (3,6,7,9). A German study examined SNP 41(rs152312), SNP 45(rs12188950) and one microsatellite marker (AC008818-1), which had shown the most significant association with stroke in Iceland (9). They further examined a set of haplotype tagging SNPs that distinguished 95% of all chromosomes within linkage disequilibrium blocks 'B' and 'C', which contained the haplotypes with the strongest association with stroke in Iceland. No positive associations with ischemic stroke were found.

An English study (6) examined the positively associated Icelandic SNPs and haplotypes in relationship to carotid

Table 5. Among all participants, SNP rs918582—additive model, stratified by stroke subtype demonstrating adjusted ORs, 95% confidence intervals and *P*-values

rs918592—additive					
Stroke subtype	Controls/cases	OR[a]	95% CI		P
Atherosclerotic	**207/27**	**2.35**	**1.22**	**4.53**	**0.011**
Cardiac	207/14	0.94	0.42	2.08	0.870
Dissection	207/13	NC	NC	NC	NC
Lacunar	**207/45**	**1.85**	**1.12**	**3.05**	**0.017**
Hematologic	207/9	0.21	0.04	1.10	0.065
Non-lacunar, undetermined	**207/125**	**1.50**	**1.07**	**2.12**	**0.020**

Significant results ($P < 0.05$) are in bold. NC indicates insufficient sample size to perform calculation; model did not converge.
[a]Adjusted for age and race (black, white and other).

intima-media thickness and carotid plaque as well as risk of ischemic stroke and ischemic stroke subtypes. No SNPs or haplotyes were associated with all ischemic stroke or with the combined atherosclerotic/cardioembolic groups. Two SNPs, SNP 19(rs4133470) and SNP 87(rs2910829) were found to be associated with large vessel stroke after adjustment for other risk factors and one of these two, SNP 87(rs2910829), was associated with carotid intima-media thickness at the carotid bulb. Six deCODE SNPs [SNP2(rs152341), SNP13(rs26949), SNP 14(rs26950), SNP 15(rs35382), SNP 20(rs16878206), SNP 26(rs40512)] showed significant association with cardioembolic stroke after multivariate adjustment, and haplotypes derived from these SNPs were less strongly associated than the individual SNPs.

A Swedish study (3) of predominantly ischemic strokes found no association with the deCODE-associated variants, the A allele of AC008818-1 marker and deCODE SNPs 45(rs12188950) and 41(rs152312) near PDE4D7 region. However, this study found associations with the B allele of AC008818-1 and an SNP near PDE4D3 region.

A study in a geographically isolated population in the Netherlands (7) examined deCODE SNPs 39(rs3887175), 45(rs12188950), 83(rs966221) and their derived haplotypes. For the overall ischemic stroke group, no associations were found either in the total population or within inbred individuals. The subgroup of inbred individuals with small vessel disease did show an association with SNPs 39(rs3887175) and 45(rs12188950).

In contrast, recent American (4,5) and Pakistani (8) studies found significant associations between deCODE SNPs and all ischemic stroke. One American (4) study, predominately consisting of Caucasians, found associations with two of four deCODE-associated variants, SNP 83(rs966221) and SNP 56(rs702553), with the most significant SNP 83(rs966221) having a stronger association with large artery atherosclerotic stroke. The other American (5) study found SNP 87(rs2910829) to be associated with cardioembolic stroke among both whites and blacks and SNP 41(rs152312) to be associated with cardioembolic stroke among whites, all after multiple comparison corrections. The Pakistani study (8) examined SNP 32(rs456009), 83(rs966221) and 87(rs2910829) and found an association with SNP 83(rs966221).

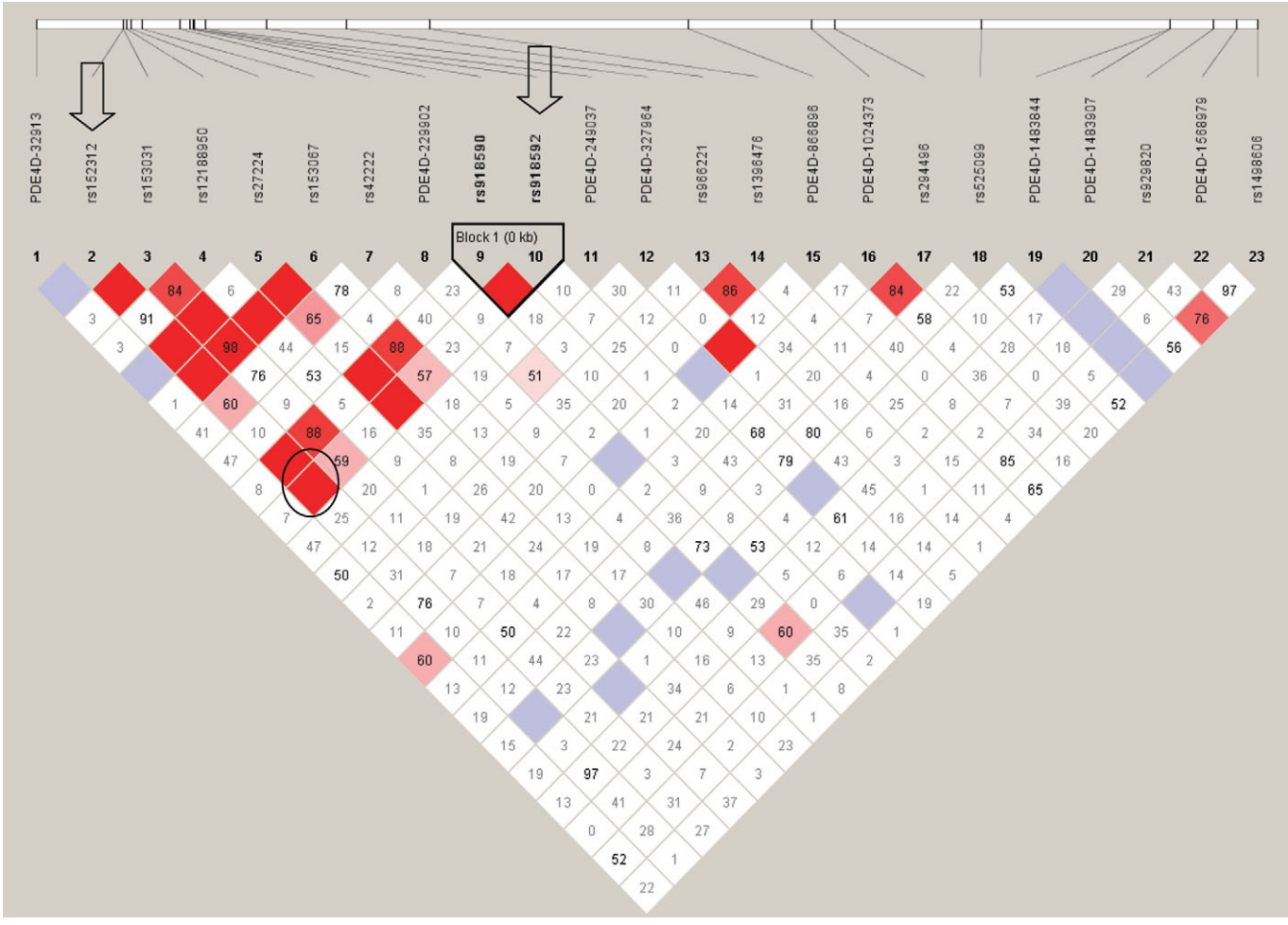

Figure 3. Haplotype block structure among African-Americans. Arrows indicate SNPs rs918592 (associated with stroke in the SPYW2 population) and rs152312 (deCODE SNP 41, associated with stroke in the Iceland population) are in strong linkage disequilibrium as indicated by the circle.

A recent Japanese (10) study examined three microsatellites, one VNTR and 31 deCODE SNPs, finding no associations with any of the polymorphisms examined after correcting for multiple comparisons. However, haplotype analyses within two separate blocks revealed significant association with cerebral infarction. The first block, consistent in location with the deCODE findings, revealed a significant haplotype ($P = 0.002$) comprising SNP83(rs966221), rs153031 and VNTR(AC008818). The second block in a different *PDE4D* gene region revealed a significant haplotype ($P < 0.001$) comprising of rs878567, rs3756739, rs6875372 and rs1423351.

Using existing information from databases, our SNP discovery efforts, as well as the Iceland results, we examined SNPs throughout the entire *PDE4D* gene with greater emphasis on the regions around the first exons at the 5′ end of the gene. Among the deCODE stroke-associated variants evaluated in our study, several SNPs were found to be associated with ischemic stroke including SNP 83(rs966221) among Caucasians in an additive model, SNP 89(rs1396476) among African-Americans using an additive model and SNP 42(rs153031) in African-Americans using a dominant model. Additionally, rs918592, which was most associated with stroke in our study, was in linkage disequilibrium with rs152312(SNP 41), one of the polymorphisms more strongly associated with stroke in the Iceland population. Furthermore, the risk haplotype in the Iceland study contained rs152312 and had lower expression of the *PDE4D*7 isoform (1).

Although many of the studies of the *PDE4D* gene focussed on replicating the Icelandic findings, the association between specific SNPs and stroke have been inconsistent. Only the Japanese study (10) and our study have examined the *PDE4D* gene more extensively, but neither study has comprehensively examined the genetic variation within the *PDE4D* gene. The Japanese study identified two stroke-associated haplotype blocks. The more strongly associated haplotype was towards the 3′ end of the gene. The other associated haplotype was in the 5′ region of *PDE4D*, near the first alternative exon for *PDE4D*7, found to be associated with stroke in Iceland. Our strongest associated SNP was also in the 5′ region of *PDE4D*.Thus, our study adds independent support to the findings of the Iceland investigators and extends these findings to early-onset stroke among African-Americans as well as Caucasians.

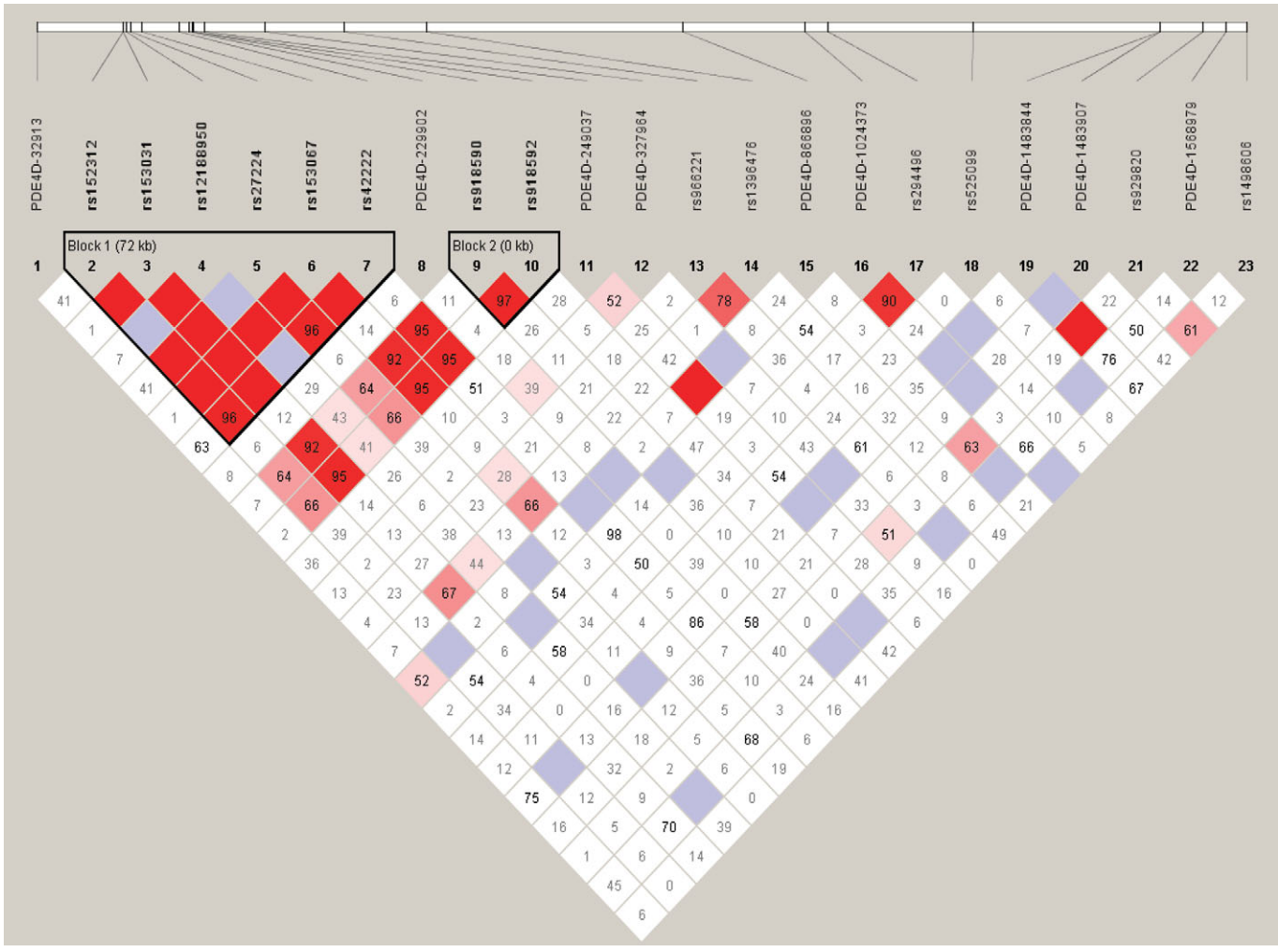

Figure 4. Haplotype block structure among Caucasians.

Our findings do not support the hypothesis that atherogenesis mediates the association. We have demonstrated an effect of the *PDE4D* locus on stroke risk among young adults, a population with a very low prevalence of atherosclerotic disease. In contrast to the findings from Iceland (1) and the recent American (4) study, we found that our most strongly associated SNP, rs918592, was associated with multiple ischemic stroke subtypes. Significant findings were found not only for atherosclerotic stroke, but also for lacunar stroke and non-lacunar stroke of undetermined etiology. This later group included large artery strokes without evidence for a significant degree of proximal atherosclerosis. Although our findings differed from the Iceland results in that there was no association with cardiac embolism, the underlying cardiac conditions associated with stroke in a young population are quite different from those among the elderly.

Ours is the first study to examine the possibility of a *PDE4D* gene–environment interaction with smoking. The risk genotype was not associated with stroke among never-smokers or former smokers, but highly associated with stroke among current smokers. Additionally, a strong dose–response relationship was seen among current smokers. Replication of these findings, including studies of males and older stroke patients, will be essential because of the small sample size, the number of risk factors tested for interaction and the lack of a similar interaction among the other stroke-associated *PDE4D* SNPs. However, an interaction is biologically plausible. Cigarette smoking causes endothelial dysfunction (11) and is known to modify the expression of many genes in endothelial cells, including cAMP response element binding protein (CREB) (12). CREB is a transcription factor that induces the expression of a panel of genes with established roles in cell survival, metabolism and plasticity in the nervous system, including a potential role in ischemic preconditioning (13). Thus, both *PDE4D* and smoking may mediate effects via CREB providing a potential mechanism for the smoking interaction. An alternative mechanism could be that smoking may potentiate the effect of increased *PDE4D* expression. In a mouse model, prenatal exposure to cigarette smoking increased *PDE4D* messenger RNA expression and decreased cAMP in the lung (14).

In conclusion, we identified several *PDE4D* SNPs associated with stroke in both African-American and Caucasian

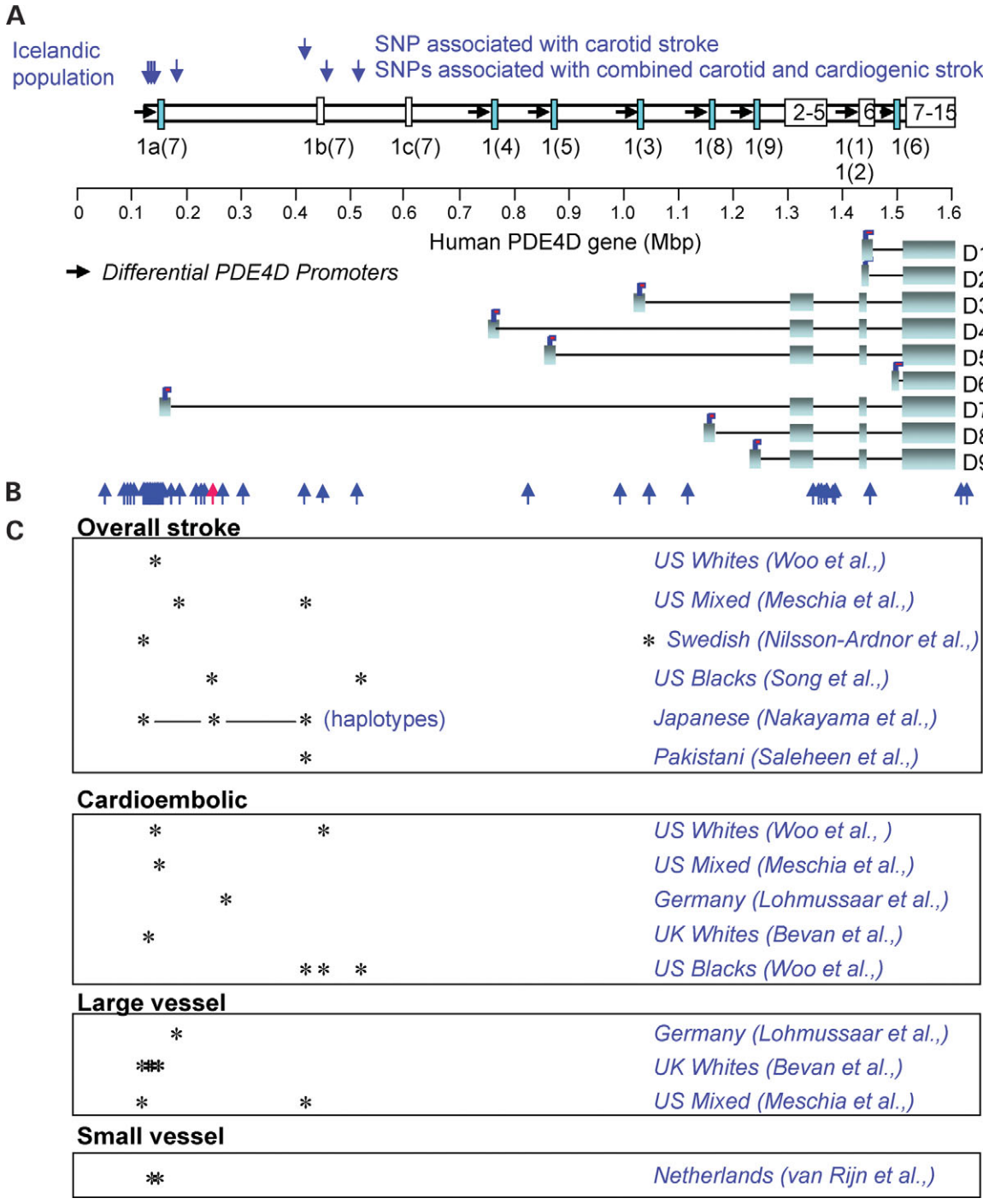

Figure 5. A summary of *PDE4D* replication association studies in various populations. (**A**) The *PDE4D* gene structure. Exons are indicated by boxes. Differential promoters for *PDE4D* isoforms are indicated by horizontal arrows on the gene. The alternative splicing patterns of *PDE4D* isoforms are shown below the gene structure. The SNPs that were associated with stroke in the loelandic population are shown above the gene structure by the vertical arrows. (**B**) The SNP position throughout the *PDE4D* gene used in the replication studies. (**C**) A summary of these replication studies with asterisks (*) representing SNPs that reached statistical significance in association with stroke.

females, providing support for association of this gene with stroke in non-Icelandic populations and among young adults. Our most significantly associated SNP was found to be in strong linkage disequilibrium with another previously associated with stroke in Iceland. The effect of the highly prevalent risk allele in this SNP was not subtype-specific but was specific to smokers. More comprehensive studies of the association between *PDE4D* genetic variation and both functional correlates and stroke risk in larger, ethnically diverse populations are needed.

MATERIALS AND METHODS

Study subjects

The SPYW2 is a population-based case–control study that was initiated to examine genetic risk factors for ischemic stroke in young women. The term 'population-based' indicates that the cases and their comparison group were identified from the same defined population, which included all of Maryland (except the far Western panhandle), Washington DC and the southern portions of both Pennsylvania and Delaware. Cases were 239 female patients 15–49 years of age with a first cerebral infarction as identified by discharge surveillance at 51 regional hospitals and through direct referral by regional neurologists. The methods for discharge surveillance, chart abstraction and case adjudication have been described previously (15–17). The adjudication of stroke cases was performed blinded to genetic information. Stroke cases were classified as having a probable, possible or undetermined etiology as per prior description (15,16). Using predetermined exclusions modified from the Siblings With Ischemic Stroke Study (SWISS) protocol (18), we excluded 15 (no. 15) cases with the following characteristics: sickle cell disease (no. 1), CNS vasculitis by angiogram and clinical criteria (no. 3), post-radiation arteriopathy (no. 1), endocarditis (no. 3), neurosyphillis (no. 1), mechanical prosthetic heart valves (no. 2), left atrial myxoma (no. 1) and cocaine use in the 48 h prior to their stroke (no. 3). Control subjects were 212 women without a history of stroke, identified by random digit dialing and were frequency matched to the cases by age and geographic region of residence. One control was excluded from analyses based on a history of sickle cell disease. Thus, the sample for genetic analyses consisted of 224 cases and 211 controls.

Cases and controls were grouped into the following race/ethnic categories: Caucasian (non-Hispanic) ($n = 95$ cases and 99 controls), African-American ($n = 105$ cases and 91 controls) and others (including Hispanic, Asian, American-Indian, etc.) ($n = 24$ cases and 21 controls). Because of the small size and heterogeneity of the latter group, it was not analyzed separately, but was included within the combined total study group ($n = 224$ cases and 211 controls). Haplotype analyses were conducted only on the Caucasian (non-Hispanic) and the African-American groups.

Previously identified SNPs

We evaluated stroke risk within our population for several previously reported positively associated Icelandic deCODE Genetics SNPs (1). The following SNPs were evaluated: rs12188950 (deCODE SNP 45), rs152312 (SNP 41), rs153031 (SNP 42), rs966221 (SNP 83) and rs1396476 (SNP 89). All of these SNPs are in the region near the alternative forms of exon 1. SNPs were also chosen from the NCBI database based upon their location in the *PDE4D* gene and their allele frequencies (minor allele frequency >0.05). Thus, rs27224, rs153067, rs42222, rs918590, rs918592, rs294496, rs525099, rs929820 and rs1498606 were included in this study.

Novel SNP discovery

Screening for novel polymorphisms was performed on a subset of stroke cases and controls including both African-Americans and Caucasians (24 unrelated individuals for each group, in total 96 individuals). Because the *PDE4D* gene locus is very large and coding SNPs tend to occur very infrequently in exons (1), we performed a computational VISTA analysis prior to SNP discovery. Aiming to identify potentially functional non-coding variants, we selected highly evolutionarily conserved regions throughout the entire human *PDE4D* gene that were revealed in the VISTA map as the targets of our novel SNP discovery effort. Special focus was paid to the *PDE4D*7 isoforms and *PDE4D*1/2 isoforms, the 5′ flanking regions (5 kb) for all *PDE4D* isoform-specific promoters and the exon–intron junctions.

VISTA alignments. Postulating that highly evolutionary-conserved genomic regions among species are likely to be functionally important, we used the VISTA web server (19) (http://www-gsd.lbl.gov/VISTA/) to align genomic sequences in our *in silico* analysis of global sequence conservation on the *PDE4D* gene locus. We aligned human (UCSC human genome assembly 2004 May chr5: 59968082.58282468), mouse (UCSC mouse genome assembly 2004 May chr13: 104923788.106765297) and rat (UCSC rat genome assembly 2003 June chr2: 39230274.41294264) genomic sequences containing the entire *PDE4D* gene and its upstream (100 kb) and downstream (20 kb) intergenic sequences. Transposable elements in the mammalian sequences were selected to be masked. Human sequence and its annotation were always used as the base sequence. Pairwise sequence comparisons were calculated with a threshold of 75% identity in a 100 bp window. All segments with a conservation $\geq 80\%$ over a 100 bp window or a conservation $\geq 75\%$ over 100 bp but within proximal 5 kb of the 5′-flanking regions for each *PDE4D* isoform-specific promoter were selected as the targets for the denaturing high-performance liquid chromatography (DHPLC) analysis.

DHPLC analysis. PCR primers were designed to amplify the targeted regions. Genomic DNA (10 ng) was amplified by PCR using AmpliTaq Gold (Applied Biosystems, Foster City, CA, USA), denatured at 94°C for 5 min and slowly re-annealed at room temperature for heteroduplex formation. Samples were then analyzed in the WAVE DNA Fragment Analysis System (Transgenomic Inc., Omaha, NE, USA) based on DHPLC, with the buffer gradient and oven temperature controlled by the WAVEMaker software. Heterozygous amplicons were identified by elution profiles. For each type of elution profile, three representative heterozygous samples were further analyzed on an ABI 3100 automatic sequencer (Applied Biosystems) in the forward and reverse directions. Sequence trace files were analyzed using the Mutation Surveyor software (SoftGenetics, State College, PA, USA) to locate nucleotide variants and were verified by manual inspections. In this study, only those polymorphic PCR amplicons (with minimal of three individuals for a DHPLC heteroduplex elution profile) were subjected to further sequence analysis.

Genotyping methods for the case/control population

After removing markers with minor allele frequency <0.05 by race and among controls, we prioritized the remaining markers as follows: (i) greater minor allele frequencies, (ii) distribution throughout the gene, (iii) VISTA peaks and (iv) deCODE positive SNPs. Twenty-three SNPs were genotyped in our stroke cohort (Table 2).

Genotyping was conducted using DNA isolated from whole blood using the QIAamp DNA Blood Maxi Kit (Qiagen, Valencia, CA, USA). SNP typing was performed by one of two methods. The first method types 12 SNPs simultaneously and was developed for an SNPstream Ultra-High Throughput machine (Beckman Coulter, Inc., Fullerton, CA, USA) Sequences surrounding the SNPs were obtained from GenBank (http://www.ncbi.nlm.nih.gov/Genbank/index.html) and submitted to Autoprimer.com (Beckman Coulter, Inc.). For each SNP, three primers were designed, two for PCR amplification and an internal primer with a 5′ DNA sequence tag. Twelve pairs of primers were used to initiate PCR amplification. The free primers were removed by enzymatic digestion using Exonuclease I and Shrimp Alkaline Phosphotase (Beckman Coulter, Inc.). The internal primers were used to initiate a sequencing reaction that adds one labeled base for the alternative nucleotides of each SNP to have distinct labels. The labeled products are separated on a SNP-IT plate consisting of 384 mini-arrays with 16 spots each (Beckman Coulter, Inc.). For each individual DNA sample, 16 spots hybridized to the two homozygotes, the heterozygote, a negative control and the 12 labeled primers associated with the 12 SNPs. Thus, every PCR and labeling reaction has internal controls to confirm the success of the reactions and the appropriateness of the fluorescent outputs for each DNA sample.

The second method was the Taqman method (Applied Biosystems). This method is based on four primers, two flanking the SNP that are used to amplify the DNA surrounding the SNP and two, one for each alternative allele, that were labeled with different fluorescent dyes. The original form of the labeled primer has a quencher in close proximity to the dye. However, when the exonuclease activity of DNA polymerase disrupts the primer hybridized to the single-strand DNA during the PCR, the quencher and the dye are released and the fluorescence can be measured. The reaction itself follows manufacturer's instructions included with each individual primer set.

Analyses

All statistical analyses were performed using SAS®, Version 9.1 (SAS Institute, Cary, NC, USA). We compared means by *t*-tests and proportions by χ^2 tests using two-sided *P*-values. For the primary analysis, we considered an age-adjusted additive model to test the effect of genotype of stroke. Analyses were conducted in the total population and in African-American and Caucasian women separately. As secondary analyses, dominant and recessive genetic models were also considered. Genotype-specific ORs were also computed.

Adjusted ORs derived by logistic regression were used to determine whether the presence of the risk allele was associated with an increased risk for stroke after controlling for potential confounders. As previously mentioned, the primary analysis or minimally adjusted model included age and race. A secondary vascular risk factor model was also implemented that included age, race, OCP use, hypertension, diabetes mellitus, current cigarette smoking and history of angina or myocardial infarction (angina/MI). Age, race, cigarette smoking status and OCP use were determined by subject reports (or proxy report, if a participant was unable to answer). Hypertension and diabetes mellitus were determined by asking study participants (or a proxy) whether a physician had ever told them that they had the condition.

Significant SNPs underwent additional race-specific stratified analyses to evaluate groups stratified by standard risk factors (age, OCP use, current cigarette smoking, hypertension, diabetes mellitus and history of angina or myocardial infarction). Additionally, significant SNPs underwent a stroke subtype (atherosclerotic, cardiac, dissection, lacunar, hematologic and all other stroke) analyses adjusting for age and race. Tests for interaction were performed using logistic regression.

Haplotype block structure among African-Americans and Caucasians was determined by implementing Haploview (http://www.broad.mit.edu/mpg/haploview/). Haplotypes were estimated and haplotypic association analyses were performed using the HaploStats package (20). Haplotypes based on eight SNPs demonstrating high pairwise correlation, as demonstrated through Haploview, were studied for association with stroke adjusting for multiple comparisons. The eight SNPs included were deCODE SNPs 41, 42 and 45, as well as rs27224, rs153067, rs42222, rs918590 and rs918592. Analyses were performed for all SPYW2 participants and stratified by race and stroke subtype.

Permutation testing was used to empirically determine the multiple comparison-corrected *P*-value for the age–race-adjusted association of rs918592 with stroke.

Because we considered multiple SNPs in our analyses, we are sensitive to the fact that a 'significant' association may have been observed with any one of them by chance alone. However, because many of the SNPs are correlated with each other, a Bonferroni-type adjustment would be too conservative. To account for the non-independence of SNPs, we therefore carried out a permutation test in which we permuted the genotype vector among subjects. In this analysis, we included all of our genotyped SNPs with a minor allele frequency greater than 0.05 in controls of both races; 17 SNPs (of the 23 analyzed) satisfied these criteria. Restricting the permutation to the genotype vector preserves the original allelic correlations between the SNPs, but disassociates any phenotype–genotype correlation that may have existed. To construct the distribution of expected *P*-values under the null hypothesis, we performed 5000 permutations, taking the smallest *P*-value from the association analysis of the 17 SNPs each time. We then defined our multiple comparison-corrected *P*-value as the proportion of the 5000 *P*-values obtained from the permutation that were lower than the *P*-value observed from our analysis of the real data. Following these analyses, we then evaluated our results in the context of the deCODE findings using a similar permutation testing methodology. In this case, the empiric experiment-wide

P-value was based upon a reduced subset of SNPs which included only the eight SNPs (listed earlier) genotyped by our group that were within the stroke-associated Icelandic haplotype block (block B) (1). A similar analysis was performed on our four genotyped SNPs in Icelandic Haplotype block A (1), which included SNPs PDE4D-249037, PDE4D-327964 and deCODE SNPs 83 and 89.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at HMG Online.

ACKNOWLEDGEMENTS

The authors are highly indebted to Laurie Reinhart, MS for her contribution to data analyses reported in this manuscript and the SPYW2 study coordinator, Mary J. Sparks, RN. This project was partially supported through a grant from Morehouse Cardiovascular Research Institute (Enhancement of Cardiovascular and Related Research Areas, NIH/NHLBI 5 UH1 HL003676), Morehouse Clinical Research Center (Center of Clinical Research Excellence, NIH/NCRR 5U54RR014758-05), Morehouse School of Medicine (Enhancement of the Capacity of Biomedical Research-RCMI, 5G12RR003034-18), Morehouse School of Medicine (Stroke Prevention/Intervention Research Program, NIH/NINDS 5U54NS046798-02) and a Beginning-Grant-in-Aid of American Heart Association. Additionally, this project was partially supported by the Department of Veterans Affairs, Baltimore, Office of Research and Development, Medical Research Service, by the Geriatrics Research, Education and Clinical Center and by the Stroke Research Enhancement Award Program. Additional support was provided by the University of Maryland, Other Tobacco-Related Diseases Research Grant through the Maryland Cigarette Restitution Fund Program. Additional support was provided by a Cooperative Agreement with the Division of Adult and Community Health, Centers for Disease Control and Prevention; the National Institute of Neurological Disorders and Stroke and the NIH Office of Research on Women's Health; the National Institute on Aging Pepper Center Grant P60 12583 and the University of Maryland General Clinical Research Center (Grant M01 RR 165001), General Clinical Research Centers Program, National Center for Research Resources, NIH.

Conflict of Interest statement. None declared.

REFERENCES

1. Gretarsdottir, S., Sveinbjornsdottir, S., Jonsson, H.H., Jakobsson, F., Einarsdottir, E., Agnarsson, U., Shkolny, D., Einarsson, G., Gudjonsdottir, H.M., Valdimarsson, E.M. *et al*. (2003) The gene encoding phosphodiesterase 4D confers risk of ischemic stroke. *Nat. Genet.*, **35**, 131–138.
2. Dominiczak, A.F. and McBride, M.W. (2003) Genetics of common polygenic stroke. *Nat. Genet.*, **35**, 116–117.
3. Nilsson-Ardnor, S., Wiklund, P.G., Lindgren, P., Nilsson, A.K., Janunger, T., Escher, S.A., Hallbeck, B., Stegmayr, B., Asplund, K. and Holmberg, D. (2005) Linkage of ischemic stroke to the PDE4D region on 5q in a Swedish population. *Stroke*, **36**, 1666–1671.
4. Meschia, J.F., Brott, T.G., Brown, R.D., Jr, Crook, R., Worrall, B.B., Kissela, B., Brown, W.M., Rich, S.S., Case, L.D., Evans, E.W. *et al*. (2005) Phosphodiesterase 4D and 5-lipoxygenase activating protein in ischemic stroke. *Ann. Neurol.*, **58**, 351–361.
5. Woo, D., Kaushal, R., Kissela, B., Sekar, P., Wolujewicz, M., Pal, P., Alwell, K., Haverbusch, M., Ewing, I., Miller, R. *et al*. (2005) Association of Phosphodiesterase 4D with ischemic stroke: a population-based case–control study. *Stroke*, **37**, 371–376.
6. Bevan, S., Porteous, L., Sitzer, M. and Markus, H.S. (2005) Phosphodiesterase 4D gene, ischemic stroke, and asymptomatic carotid atherosclerosis. *Stroke*, **36**, 949–953.
7. van Rijn, M.J., Slooter, A.J., Schut, A.F., Isaacs, A., Aulchenko, Y.S., Snijders, P.J., Kappelle, L.J., van Swieten, J.C., Oostra, B.A. and van Duijn, C.M. (2005) Familial aggregation, the PDE4D gene, and ischemic stroke in a genetically isolated population. *Neurology*, **65**, 1203–1209.
8. Saleheen, D., Bukhari, S., Haider, S.R., Nazir, A., Khanum, S., Shafqat, S., Anis, M.K. and Frossard, P. (2005) Association of phosphodiesterase 4D gene with ischemic stroke in a Pakistani population. *Stroke*, **36**, 2275–2277.
9. Lohmussaar, E., Gschwendtner, A., Mueller, J.C., Org, T., Wichmann, E., Hamann, G., Meitinger, T. and Dichgans, M. (2005) ALOX5AP gene and the PDE4D gene in a central European population of stroke patients. *Stroke*, **36**, 731–736.
10. Nakayama, T., Asai, S., Sato, N. and Soma, M. (2006) Genotype and haplotype association study of the *STRK1* region on 5q12 among Japanese: a case–control study. *Stroke*, **37**, 69–76.
11. Celermajer, D.S., Sorensen, K.E., Georgakopoulos, D., Bull, C., Thomas, O., Robinson, J. and Deanfield, J.E. (1993) Cigarette smoking is associated with dose-related and potentially reversible impairment of endothelium-dependent dilatation in healthy young adults. *Circulation*, **88**, 2149–2155.
12. Zhang, S., Day, I.N. and Ye, S. (2001) Microarray analysis of nicotine-induced changes in gene expression in endothelial cells. *Physiol. Genomics*, **5**, 187–192.
13. Ratan, R.R. (2004) cAMP response element binding protein family transcription factors: the Holy Grail of neurological therapeutics? *Ann. Neurol.*, **56**, 607–608.
14. Singh, S.P., Barrett, E.G., Kalra, R., Razani-Boroujerdi, S., Langley, R.J., Kurup, V., Tesfaigzi, Y. and Sopori, M.L. (2003) Prenatal cigarette smoke decreases lung cAMP and increases airway hyperresponsiveness. *Am. J. Respir. Crit. Care Med.*, **168**, 342–347.
15. Johnson, C.J., Kittner, S.J., McCarter, R.J., Sloan, M.A., Stern, B.J., Buchholz, D. and Price, T.R. (1995) Interrater reliability of an etiologic classification of ischemic stroke. *Stroke*, **26**, 46–51.
16. Kittner, S.J., Stern, B.J., Feeser, B.R., Hebel, R., Nagey, D.A., Buchholz, D.W., Earley, C.J., Johnson, C.J., Macko, R.F., Sloan, M.A. *et al*. (1996) Pregnancy and the risk of stroke. *N. Engl. J. Med.*, **335**, 768–774.
17. Kittner, S.J., Stern, B.J., Wozniak, M., Buchholz, D.W., Earley, C.J., Feeser, B.R., Johnson, C.J., Macko, R.F., McCarter, R.J., Price, T.R. *et al*. (1998) Cerebral infarction in young adults: the Baltimore-Washington Cooperative Young Stroke Study. *Neurology*, **50**, 890–894.
18. Meschia, J.F., Brown, R.D., Jr., Brott, T.G., Chukwudelunzu, F.E., Hardy, J. and Rich, S.S. (2002) The Siblings With Ischemic Stroke Study (SWISS) protocol. *BMC Med. Genet.*, **3**, Epub.
19. Mayor, C., Brudno, M., Schwartz, J.R., Poliakov, A., Rubin, E.M., Frazer, K.A., Pachter, L.S. and Dubchak, I. (2000) VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics*, **16**, 1046–1047.
20. Schaid, D.J., Rowland, C.M., Tines, D.E., Jacobson, R.M. and Poland, G.A. (2002) Score tests for association between traits and haplotypes when linkage phase is ambiguous. *Am. J. Hum. Genet.*, **70**, 425–434.