Stroke genomics in people of African ancestry: charting new paths

RO Akinyemi, B Ovbiagele, A Akpalu, C Jenkins, K Sagoe, L Owolabi, F Sarfo, R Obiako, M Gebreziabher, E Melikam, S Warth, O Arulogun, D Lackland, A Ogunniyi, H Tiwari, RN Kalaria, D Arnett, MO Owolabi, for the SIREN investigators as members of the H3Africa Consortium

Abstract

One in six people worldwide will experience a stroke in his/her lifetime. While people in Africa carry a disproportionately higher burden of poor stroke outcomes, compared to the rest of the world, the exact contribution of genomic factors to this disparity is unknown. Despite noteworthy research into stroke genomics, studies exploring the genetic contribution to stroke among populations of African ancestry in the United States are few. Furthermore, genomics data in populations living in Africa are lacking. The wide genomic variation of African populations offers a unique opportunity to identify genomic variants with causal relationships to stroke across different ethnic groups. The Stroke Investigative Research and Educational Network (SIREN), a component of the Human Health and Heredity in Africa (H3Africa) Consortium, aims to explore genomic and environmental risk factors for stroke in populations of African ancestry in West Africa and the United States. In this article, we review the literature on the genomics of stroke with particular emphasis on populations of African origin.

Stroke is the clinical culmination of several complex processes and interacting pathways that involve various genetic and environmental factors.[1] However, the exact nature and level of the contribution of genetic factors to stroke and its different subtypes have not been clearly established. Presumably, genetic contributions to stroke may result from common variants with small effect sizes, rare variants with large effect sizes, or a combination of both.[2-4] Nevertheless, studies exploring the genetic underpinnings of the peculiarities of stroke in populations of African ancestry in the United States are few,[3,5-7] while there are hardly any data on populations living in Africa.

The diverse genomic variation of African populations[8-10] offers a unique opportunity to identify novel genes and molecular pathways of stroke that may lead to new and better prevention and treatment options for stroke in people of African ancestry and other global populations. Understanding the interplay of genetic and environmental risk factors for stroke is critical to the prediction of its occurrence, severity and outcome as well as the formulation of successful tailored treatment and prevention programmes. In addition, the biology of stroke subtypes will be better deciphered.

In this review article, we provide an overview of the changing global and in particular, African epidemiology of stroke, the known peculiarities of stroke in Africa, extant literature on the genomics of stroke and cerebrovascular risk factors, with particular attention to people of African ancestry, as well as opportunities for charting new paths through the Human, Health and Heredity in Africa (H3Africa) initiative.[11,12]

Division of Neurology, Federal Medical Centre Abeokuta, Nigeria
RO Akinyemi, MB BS, PhD

Department of Neurosciences, Medical University of South Carolina, USA
B Ovbiagele, MD, MSc
M Gebreziabher, PhD
S Warth, BSc
D Lackland, PhD

College of Health Sciences, University of Ghana, Accra, Ghana
A Akpalu, MB BS
K Sagoe, PhD

Department of Medicine, Bayero University, Kano, Nigeria
L Owolabi, MB BS

School of Medical Sciences, Kwame Nkrumah University of Science and Technology, Kumasi, Ghana
F Sarfo, MB BS, PhD

Department of Medicine, Ahmadu Bello University, Zaria, Nigeria
R Obiako, MB BS, MSc

Department of Public Health, University of Alabama at Birmingham, USA
H Tiwari, PhD
D Arnett, PhD

Institute of Neuroscience, Newcastle University, UK
RO Akinyemi, MB BS, PhD
RN Kalaria, PhD, FRCPath

College of Medicine, University of Ibadan, Ibadan, Nigeria
E Melikam, MSc
O Arulogun, PhD
A Ogunniyi, FRCP
MO Owolabi, MB BS, DM, mayowaowolabi@yahoo.com

Changing global and African epidemiology of stroke

Stroke has attracted global attention, as one in six people will develop stroke in their lifetime.[13,14] Stroke is a significant medical and public health problem, with loss of productivity and burden on families, caregivers and society.[13-15] The lifetime risk of stroke is one in five women and one in six men.[13] Stroke is the most common cause of acquired disability and the second most common cause of death worldwide. The World Health Organisation (WHO) estimates for 2001 indicated that death from stroke and disability-adjusted life years (DALYs) due to stroke was at least seven times higher in low- and middle-income countries (LMIC) than in high-income countries (HIC).[16] Recent incidence estimates[17,18] indicate that whereas stroke incidence declined 12% in HIC, it increased by 12% in LMIC over the last decade.

Africa, in epidemiological transition,[19,20] is currently faced with an exploding but neglected burden of non-communicable diseases (NCDs), including hypertension, diabetes mellitus and dyslipidaemia, which often culminate in stroke. The recent United Nations high-level meeting on the global burden of NCDs highlighted their disproportionately high burden and stressed the urgent need to tackle them, particularly in developing countries.

The burgeoning incidence of stroke in Africa is attributable to rising cardiovascular risk profile, which is in turn driven by epidemiological transition, an aging population, rapid urbanisation and accompanying lifestyle changes. Africa, with a current population of over one billion, has a stroke prevalence rate of up to 963/100 000 population,[21] an incidence rate of up to 315/100 000 population and a three-year mortality rate as high as 84%. About 3.2 million Africans develop incident stroke every year.[15] Recent data from Nigeria, Tanzania and Sudan showed that stroke was the leading cause of elderly medical admissions,[22] while up to 78% of neurological hospital admissions were due to stroke.[23] The impact of this on mental capacity, quality of life and economic productivity portends great danger for the emerging economies of Africa.[24-26]

Peculiarities of stroke in people of African ancestry

Enhanced predisposition, different pattern of subtypes, worse severity and often poorer outcome of stroke in people of African descent is quite well established. According to data from the INTERSTROKE study, ischaemic stroke accounts for 66% while haemorrhagic stroke accounts for 34% in Africa, compared to 91 and 9%, respectively for ischaemic and haemorrhagic stroke in HIC. Ischaemic stroke subtypes diagnosed in African populations were small vessel (27%), cardio-embolism (25%), large vessel (14%), others (20%) and undetermined (14%).[27]

Among sub-Saharan Africans, stroke affects a relatively younger age group and productive workforce than in developed economies.[27,28] Data from the USA suggest that African Americans have a higher predisposition, worse severity and often poorer outcomes compared to Caucasian Americans.[29,30] In a recent report from the multi-ethnic South London Stroke Registry study, black stroke survivors had worse cognitive outcome compared to other racial groups.[31] Although, this may be due to socio-economic differences, disparities in healthcare-seeking practices and differential access to healthcare services, the influence of underlying differences in genetic factors cannot be underestimated.[32,33]

Genomics and health disparities

Genetics and genomics research offer insight into disparities in the risk profile, phenotypes and outcome of diseases among different populations as a result of accumulated small differences in common alleles or rare variants, interactions among multiple genetic loci and interactions between genes and environmental factors, which may include cultural practices and health-seeking behaviour.[34,35] The potential of treatment approaches tailored to individual, unique genomic profiles represents a distinct potential impact of genomics on improving health disparities. Also, the globalisation of complex chronic diseases further suggests that all populations are susceptible, and that variation in rates may also be explained as a result of differential exposure to environmental causes, including lifestyles, cultural practices and health-seeking behaviours.[36]

African human genomic variation

African populations present the highest genomic diversity, the lowest levels and most divergent patterns of linkage disequilibrium, as well as smaller haplotype block sizes across human populations.[8,37] Although the human species is believed to have originated from Africa about 200 000 years ago, studies of genomic variation in Africa suggest that the present pattern of variation within and between populations is a product of several factors. These include demographic history, population structure, diversities of geographical location, language classification and different patterns of subsistence, dietary differences, multiple migrations with accompanying high levels of genetic admixture and survival related to exposure to infectious diseases.[38,39]

For example, Tishkoff and colleagues[8] identified 14 ancestral population clusters in Africa with four predominant clusters that broadly represent populations from major African geographical regions and the four dominant African language families. These are Niger-Kordofanian (spoken primarily by agriculturalist populations located in large contiguous regions of sub-Saharan Africa from West Africa to eastern and southern Africa), Nilo-Saharan (spoken predominantly by pastoralist populations in central and eastern Africa), Afro-Asiatic (spoken predominantly by agro-pastoralists and pastoralist populations in northern and eastern Africa), and Khoisan (a language family that contains click consonants, spoken by hunter–gatherer San populations in southern Africa as well as the Hadza and Sandawe hunter–gatherers in Tanzania). The remaining 10 are mainly restricted to specific geographic regions, languages, or in some cases, individual populations.

More recently, Shriner and colleagues[9] analysed ancestry data from 12 global and regional diversity projects with genome-wide genotype data for 3 528 unrelated individuals from 163 samples from around the world. They identified 19 ancestral components with 94.4% of individuals showing mixed ancestry. Furthermore, they validated the earlier findings of Tishkoff and colleagues and identified an additional ancestral component in Africa, the Omotic-speaking peoples of Ethiopia.

Our knowledge of African human genomic variation is growing. This was previously limited by the small number of African populations involved in landmark projects such as the International HapMap project[40] and the more recent 1 000 Genomes project.[41] In these projects, participation was limited

to largely Niger-Kordofanian-speaking Yoruba and Esan from Nigeria, Mende from Sierra Leone, Bantu-speaking Luhya from Kenya, and Nilo-Saharan-speaking Maasai from Kenya. A large proportion of African human genomic variation therefore remained unexplored.

However, recent data from the African Genome Variation project (AGVP) have provided evidence and more detailed characterisation of African genomic diversity.[10] The AGVP utilised dense genotypes from 1 481 individuals and whole-genome sequences from 320 individuals across sub-Saharan Africa. Novel evidence of complex, regionally distinct widespread hunter–gatherer and Eurasian admixture across sub-Saharan Africa was apparent and substantial hunter–gatherer and Eurasian ancestry admixture of up to 23 and 50%, respectively, were found in many African populations with detailed chronology of the timing of the admixture. For instance, whereas the Eurasian admixture among the Yoruba occurred 7 500–10 500 years ago, it was more recent among the Fula tribe of Gambia, occurring only about 320–780 years ago.

These admixtures provide evidence for back-to-Africa migration, the existence of hunter–gatherer populations in West Africa and a pattern of gene flow consistent with the Bantu expansion. The AGVP also found new loci related to susceptibility, pathogenesis, severity and outcome of several diseases, including malaria, Lassa fever, trypanosomiasis, trachoma and hypertension. For instance, they identified highly differentiated variants within genes involved in osmoregulation (*ATP1A1* and *AQP2*), deregulation of *AQP2* expression, and loss-of-function mutations in *ATP1A1* have been associated with essential and secondary hypertension, respectively.[42,43] The study also established an efficient genotype array design capturing common genetic variation in Africa, which would be useful for future African genomic studies.[11,12]

African American genomic variation

African Americans have mixed ancestry originating from Africa and other continents, especially Europe. Studies have shown the average amount of African ancestry in African Americans to be about 80% (predominantly of western and central African origin),[8,44,45] although there is substantial variation in the level of African ancestry in individual African Americans, as the proportion of African ancestry in a given individual can range from one to 99%.[39]

Genomic variation and related phenotype data on variable traits contribute novel information useful for identifying population-specific variants that play a role in gene function, phenotypic adaptation and susceptibility to complex diseases, such as stroke in Africans and populations of African descent. The APOE ε4 allele, a well-studied example that contributes to a small extent to individual and population risks of traits such as stroke, heart disease and dementia, is found in virtually all populations, albeit at varying rates. The frequency of homo- or heterogeneous APO ε4 alleles varies across populations but confers different attributable risks of Alzheimer's disease; the risk being higher among the Japanese but much lower among people of African ancestry with higher allele frequencies. This suggests possible intervening roles for epigenetic interactions from certain modifier genes or some other environmental factors.[34]

Genomics of stroke and cerebrovascular risk factors

Stroke is a complex polygenic, heterogeneous and multifactorial disorder involving many complex mechanisms, intermediate phenotypes and the interplay of genetic and non-genetic factors. Evidence from twin studies, family history studies, animal models and heritability studies of vascular risk factors and intermediate phenotypes suggests a likely significant contribution of genetic factors to the neurobiology and phenomenology of stroke.[1,4,46]

Family history and heritability

Among individuals with a positive family history of stroke, there is an increased risk of stroke, which may be due to expression of genetic susceptibility, a shared environment or both.[47] In the Family Heart study, personal and familial histories of stroke were assessed in 3 168 individuals (probands) who were at least 45 years old and 29 325 of their first-degree relatives. The odds of stroke were 2.00 (1.13–3.54) for a positive paternal and 1.41 (0.80–2.50) for a positive maternal history of stroke after adjusting for age, gender, ethnicity and presence of vascular risk factors, and the pattern was similar between African Americans and European Americans.[48]

In a systematic review of the heritability of stroke in 53 independent studies (three twin studies, 33 case–control studies and 17 cohort studies), it was found that monozygotic twins were more likely to be concordant than dizygotic twins (OR, 1.65; 95% CI, 1.2–2.3; $p = 0.003$) while a positive family history was a risk factor for stroke in both case–control (OR, 1.76; 95% CI, 1.7–1.9; $p < 0.00001$) and cohort (OR, 1.30; 95% CI, 1.2–1.5; $p < 0.00001$) studies. Besides, positive family history was more associated with small-vessel and large-vessel strokes.[49]

Cerebrovascular risk factors

Genomic factors may contribute to the neurobiology of stroke through their influence on established risk factors, such as hypertension, diabetes, dyslipidaemia, obesity and cigarette smoking or through their influence on intermediate phenotypes, such as white matter hyperintensities (WMH) and carotid intima–media thickness (CIMT). For instance, research evidence has shown racial and ethnic disparities in cardiovascular and cerebrovascular diseases, with Americans of African ancestry showing a higher prevalence of hypertension and earlier onset, and faster and more severe end-organ damage, including stroke.[50] Apart from non-inherited factors such as lifestyles, health-related practices, socio-economic profile and differential access to healthcare, genetic factors contributed significantly to this disparity.[32]

A recent genome-wide association study (GWAS) of hypertension and blood pressure in African Americans using the pathway-focused approach established the genome-wide significant association of the genetic variants *PMS1, SLC24A4, YWHA7, IPO7* and *CACANA1H* with systolic blood pressure levels, with significant replication of some single-nucleotide polymorphisms (SNPs) in a sample of West Africans.[51] Using a similar approach, a more recent study has found association between multiple variants in several genes in the adrenergic alpha-1 receptor (ADRA1) pathway and hypertension in Yoruba Nigerians.[52] A meta-analysis of genome-wide linkage scans for blood pressure variation in Nigerians and African Americans

reported association in two loci: 2p14 –p13.1 and 7p21.3 –p15.3, the second locus being attributed to the Nigerian sample and suggesting a unique locus for blood pressure variation in people of African ancestry.[53]

In the GenHAT study evaluating the pharmacogenetic effects of candidate gene complexes on stroke, significant genetic difference was found between hypertension drug treatment groups in patients who had experienced stroke, especially among African Americans and non-Hispanic whites.[54] Given the fact that hypertension is the most dominant risk factor for stroke among people of African ancestry in Africa,[15,24,27,28] and the diaspora,[30,50] it would be worthwhile exploring the possible contribution of these hypertension-related genotypes in people of African ancestry.

A Nigerian study assessed glucose and insulin responses to an oral glucose load among offspring of parents with type 2 diabetes mellitus (T2DM) and found higher levels of fasting plasma glucose, fasting plasma insulin, and two-hour post-glucose load plasma insulin, indicating a higher risk for developing diabetes.[55] A Cameroonian case–control pedigree study showed increased prevalence of diabetes and impaired glucose tolerance in the offspring of parents with T2DM.[56]

The Africa America Diabetes Mellitus (AADM) study has utilised genome-wide linkage and association studies to provide insight into the genomics of T2DM in Niger-Kordofanian African populations of Nigeria and Ghana. Multiple linkage analysis provided evidence of regions of chromosome 12, 19 and 20 (the strongest being 20q13.3).[57] The loci found to influence C-peptide plasma levels (10q23, 4p15) were found to harbour multiple T2DM candidate genes [phosphatase and tensin homolog (PTEN), protein phosphatase 1, regulatory subunit 3C (PPP1R3C), insulin degrading enzyme (IDE), and peroxisome proliferator activated receptor gamma, coactivator 1 alpha (PPARGC1)].[58] Collaborative GWAS and other studies have identified further susceptibility (CDKAL1, CAPN10, TCF7L2 variants and PPARG variants) and protective loci (TCF2, AGRP -38C/T).[57]

Chronic kidney disease (CKD) is an identified risk factor for cerebral vascular disease.[59] Multiple common SNPs in the gene that encodes non-muscle myosin heavy-chain type II isoform A (MYH9) have been associated with an increase in the risk of focal segmental glomerulosclerosis and end-stage renal disease,[60] while more recently the apolipoprotein L1 (APOL1) gene has been identified as a risk locus for CKD in African Americans, and replications confirmed in Nigerian Yoruba CKD patients.[61,62]

In Africa, APOL1 confers resistance to infection from Trypanosoma brucei brucei, one of the trypanosomes that cause African sleeping sickness and it is believed that its evolutionary history lies in its positive selection due to its protection against sleeping sickness.[63] Interestingly, an inverse relationship between high-density lipoprotein cholesterol (HDL-C) and kidney function in African ancestry populations has also been described in individuals with the nephropathy risk APOL1 gene.[64] Higher HDL-C was associated with worse kidney function in those with the risk genotype, while no association was observed among those without the genotype. Therefore, the increasing incidence of cardiovascular disorders (CVDs) in Africa along with the evidence of genetic variants that increase susceptibility to CVDs signals the need for large-scale genomic epidemiology studies in Africa in search of other putative protective and susceptibility loci.[12]

Population-attributable risks of genetic variants differ depending on whether they are monogenic, common variants or rare variants of multiple genes of polygenic disorders.[2,3] More important, however, is the functional significance of the variants in the biological pathways where their gene products contribute to the biology of the disease and may possibly be of therapeutic or preventative importance.[65] This is the major thrust of our proposed study of genetic variants relevant to stroke in people of African ancestry. African representation in the 1 000 Genomes study is limited,[41] while the H3Africa projects[11] offer robust opportunities for detailed exploration of genomic data relevant to African and global populations.

Intermediate phenotypes: WMH and CIMT

Both twin and family studies have shown that magnetic resonance imaging of white matter hyperintensities has shown a heritability (proportion of variation explained by genetic factors) of up to 70%.[66] CIMT measured by ultrasound and believed to represent the early stages of atherosclerosis and related to large-artery stroke has been estimated to have a heritability of between 30 and 70%.[67]

Monogenic stroke disorders

Monogenic disorders may cause stroke as part of multi-systemic manifestations or solely as a clinical phenotype limited to the central nervous system. They are important for individual patients but may not account for much population-attributable risk.[46]

Sickle cell disease (SCD) is of particular importance in people of African ancestry. It is caused by a point mutation at codon 6 of the beta-globin gene, leading to a glutamic acid to valine (Glu→Val) substitution in the beta-globin chain of human adult haemoglobin, and producing sickle haemoglobin (HbS). Inherited autosomal recessively, either two copies of HbS or one copy of HbS plus another beta-globin variant (such as HbC) are required for disease expression. HbS carriers are protected from malaria infection, and this selective pressure is believed to have led to the high frequency of HbS (up to 40%) in individuals of African ancestry, especially in areas of high malaria endemicity.[68,69]

The spread of SCD to the Americas is inextricably linked to slavery and the large-scale forced translocation of populations from West Africa.[70] Large- or small-vessel cerebral 'vasculopathy' characterised by proximal intracranial arterial stenoses, often leading to a moyamoya pattern, commonly complicates SCD and may manifest as abnormal transcranial Doppler velocity (> 200 cm/s) or frank stroke, particularly in younger patients with sickle cell anaemia, while complicated (hemiplegic) migraine was previously reported in Nigerian adults with sickle cell trait (HbAS).[71] By middle age, up to 25% of SCD patients develop overt stroke.[72]

Certain genetic polymorphisms may be associated with stroke in SCD as modifier genes. For instance whereas α-thalassaemia genes may be protective, mutations in the glucose-6-phosphate dehydrogenase (G6PD) genes and certain SNPs, including ANXA2, rs11853426, TEK rs489347, and TGFBR3 rs284875 variants, have been associated with increased stroke risk.[73] A recent whole-exome sequencing (WES) study identified two

modifier mutations *GOLGB1* (*Y1212C*) and *ENPP1* (*K173Q*) associated with protection from stroke in a cohort of children with sickle cell anaemia.[74]

However, the interactions between SCD, its associated modifier genes, and environmental factors to produce an intermediate phenotype (TCD velocity > 200 cm/s) and stroke have not been examined in people of indigenous sub-Saharan Africa. Knowledge of these interactions and the metabolic pathways involved may unmask targets for preventative and therapeutic interventions in the sub-population of people living with SCD.

Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) is associated with mutations in the *NOTCH3* gene and presents with migraine headache, followed by depression and ischaemic stroke in the deep gray structures and subcortical white matter, cognitive decline, and dementia.[75,76] A model of small-vessel disease, the first case of CADASIL in populations of African ancestry, was recently reported in a 73-year-old African American with a 15-base-pair heterozygous duplication of the exon 7 of the *NOTCH 3* gene.[77] Other related monogenic small-vessel cerebrovascular disorders include cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy (CARASIL), retinal vasculopathy with cerebral leukodystrophy (RVCL), and Fabry disease.[78]

Genetic linkage studies

Genetic linkage studies have contributed to our understanding of the heritability of stroke and especially chromosomal regions and sub-regions involved, even though most studies have focused more on the ischaemic phenotype. Table 1 shows findings from a few linkage studies in stroke,[79-82] including genome-wide linkage studies. Much of these findings are further confirmed by more specific candidate gene analysis and the more rigorous approaches of association studies. A relative strength of linkage studies is the feasibility of working with a few hundred subjects using the case–control approach.

Stroke candidate genes

Identification of the phosphodiesterase 4D (*PDE4D*) and 5-lipoxygenase activating protein (*ALOX5AP*) genes through linkage analysis by the Icelandic Decode group was a significant landmark in the history of stroke genomics.[1,2,4] The rs918592 SNP variant of *PDE4D* was found to be significantly associated with stroke in current smokers in an African American cohort,[7] while mutations in the *NOS3* have also been significantly associated with large-artery stroke in African Americans.[6] Other variants significantly associated with ischaemic stroke in African Americans include the *IL6R* polymorphisms and the kappa-casein gene *CSN3* found on chromosome 4, the latter through exome sequencing.[3,5]

The genetics of intracerebral haemorrhage (ICH) has also been explored through a range of candidate gene and GWAS approaches. Genes involved in the renin–angiotensin–aldosterone system, coagulation pathway, lipid metabolism, homocysteine metabolism and inflammation are among the most explored.[83] The *APOE* ε2 and *APOE* ε4 genes have been associated with lobal ICH in Caucasian, Asian and African American populations with a high prevalence of cerebral amyloid angiopathy (CAA).[84,85]

The association of *APOE* ε4 genes with deep ICH is rather inconsistent.[85] It is widely accepted that CAA is a frequent cause of ICH and the presence of an APOE ε4 allele substantially increases the risk of CAA.[86] Whether the risk of CAA or ICH is different in Africans is uncertain but it has long been known that the frequency of the APOE ε4 alleles, irrespective of tribal origin, is highly represented in the general African population.[87]

Deep ICH is also more aetiologically related to hypertensive chronic small-vessel disease and is likely to be more relevant in African populations where hypertension is the dominant stroke risk factor. Other genes with significant polymorphisms related to ICH include the methylenetetrahydrofolatereductase (*MTHFR*), interleukin-6, tumour necrosis factor-α, angiotensin converting enzyme (ACE), factor VII, factor XIII, platelet activating factor and β-tubulin, although most are described in populations of non-African ancestry.[83]

Stroke GWAS and WES studies

The candidate genes approach has proved disappointing in identifying genes contributing to the risk of multifactorial or polygenic stroke. This is a situation shared with other complex diseases.[88] Recently, the GWAS approach has revolutionised the field of stroke genetics. GWAS enables markers spanning the whole genome to be genotyped in a single experiment. Using a case–control methodology and rigorous statistical methods to account for the multiple comparisons made, associations between completely unexpected chromosomal loci and disease can be identified.[88,89]

Table 1. Genetic linkage studies in stroke				
First author (year)[74]	*Study type*	*Phenotype*	*Sample*	*Salient findings*
Craig *et al.* (1998)[74]	Linkage analysis	Cerebral cavernous malformation	20 non-Hispanic Caucasian families	CCM – 1(7q) (found in Hispanic Americans), CCM2 (7p13-15) and CCM3 at 3q25.2-27 all found in non-Hispanic Caucasian families.
Nilsson-Ardnor *et al.* (2007)[76]	Genome-wide linkage analysis	All strokes; ischaemic stroke	56 Swedish families with familial stroke	LOD scores > 1.2 at 1ocations: 1p34, 5q13, 7q35, 9q22, 9q34, 13q32, 14q32, 18p11, and moderate linkage on chromosomes 5q, 9q, 13q, and 18p.
			Additional 53 families with familial strokes	Analysis of 53 additional families, further confirmed linkage on chromosomes 5q, 13q, and 18p.
Janunger *et al.* (2009)[75]	Genome-wide linkage analysis	All strokes	7 nuclear Swedish families with a common ancestor and connected over 8 generations	A maximum allele-sharing LOD score of 4.81 on chromosome 9q31-q33 was detected. Haplotype analysis identified a region for intracerebral haemorrhage.
Wang *et al.* (2014)[77]	Linkage and association analysis	Ischaemic stroke	227 Chinese families with ischaemic stroke	SNP rs1800798 in the IL-8 gene is signficantly linked to ischaemic stroke (*p* = 0.002) and small arterial occlusion (small-vessel disease) (*p* = 0.022).

GWAS has been employed to identify genetic loci for many other cardiovascular diseases such as coronary heart disease, diabetes and hypertension, and is just being applied to stroke. The pitfalls of previous studies of genomic contributions to stroke include poor phenotyping, underpowered studies, confounders, winner's curse, and non-validation in independent populations.[88,90] For example, the Siblings With Ischemic Stroke study (SWISS) did not demonstrate any significant genome-wide association.[91] However, certain novel genetic variants have been identified as risk factors in stroke populations, with some being replicated in other populations.

The International Stroke Genetics Consortium and the Wellcome Trust Case–Control Consortium published the largest GWAS for ischaemic stroke carried out to date. This study successfully demonstrated the importance of very large multicentre study samples, identified a new associated genetic variant and replicated findings of previous stroke GWAS. The findings also demonstrated the value of clear phenotyping and the fact that different stroke phenotypes may differ in their genetic architectures. Table 2 summarises the findings of salient recent GWAS studies in stroke, including a single study by Cole *et al*. in 2012, which utilised exome sequencing.[3,92-103]

Genetic studies of stroke in Africa

To date, only a few stroke genetic studies (Table 3) have been reported from North Africa and remarkably, none from sub-Saharan Africa where the burden of stroke is disproportionately heavy and the phenomics of stroke appears relatively different. Saidi and colleagues working consistently with a growing Tunisian stroke cohort have reported significant association between ischaemic stroke and polymorphisms in several genes, including plasminogen activator inhibitor, *APOE* ε4, human plasminogen activator, human platelet antigen, angiotensin converting enzyme *Del/Del* genotype, angiotensinogen, endothelial nitric oxide synthase and aldosterone synthase.[104-111]

A single study from Egypt noted that the presence of the *ACE* D allele significantly predisposed to stroke in children with sickle cell anaemia.[112] It is, however, significant to note that the people of North Africa have a different ancestral origin (predominantly Arabian and Berber) from sub-Saharan African populations.[9] Therefore, significant differences may be anticipated in the genomic profile of stroke and subtypes in sub-Saharan Africans.

Problems and perspectives

Apart from the lack of community-based ideal stroke epidemiological data sets and the challenge of accurate phenotypic characterisation of cases in sub-Saharan Africa, there are other inherent problems of genomic research ranging from the negative impact of cultural and religious beliefs, issues of autonomy of decision making and voluntary participation, as well as poor understanding of the health impact of genomics.[113-115] In a qualitative study assessing knowledge and attitude towards personal genomics testing for complex diseases among Nigerians, even though respondents felt the outcome of genomic testing might aid healthful lifestyle modifications, attitude was influenced by religion and culture, especially aspects that might directly contradict beliefs and practices or lead to actions contradicting religious beliefs.[115]

All these aspects introduce critical ethical issues into the framework of genomics research in Africa, which need to be addressed in order to achieve success and popularise the prospects of personalised genomic medicine. In addition, there are also the challenges of adequate infrastructure for genomic studies and analysis of genomic data, a paucity of appropriately trained scientists and physicians who have the capacity to design, implement and interpret such studies and lead translational applications, and insufficient bio-informaticians with analysis expertise and research managers. Unstable power supply and political instability are other bottlenecks.

Opportunities through H3 Africa: SIREN charting new paths

Although African populations harbour the greatest human genomic diversity, the potential of this for understanding human evolutionary biology and disparities in health and disease are

Table 2. Recent GWAS and WES studies in stroke

First author	Study type	Phenotype	Sample size	Sample ancestry	Associated regions
Hata *et al.* (2011)[85]	GWAS	Ischaemic stroke	1 112 cases, 1 112 controls	Japanese	14q22 (PRKCH), 11q12 (AGTRL1)
Matarin *et al.* (2009)[89]	GWAS	Ischaemic stroke	249 cases, 268 controls	White	None
Gretasrdottri and Gudjartsson *et al.* (2008, 2009)[87,88]	GWAS	Ischaemic stroke	1 661 cases, 10 815 controls	Icelandic	4q25 (PITX2), 16q22.3 (ZFHX3)
Bilguvar *et al.* (2008)[89]	GWAS	Intracranial aneurysms	2 100 cases, 8 000 controls	Finish, Dutch, Japanese	2q33 (PLCL1), 8q12 (SOX17), 9p21.3 (CDKN2A, CDK N2B, ANRIL)
lkram *et al.* (2009)[90]	GWAS	Ischaemic stroke	Cohort of 19 602, 1 164 events	Caucasian	12p13.33 (NINJ2)
Yamada *et al.* (2009)[91]	GWAS	Ischaemic stroke	992 cases, 5 349 controls	Japanese	22q13 (CELSR1)
Zhang *et al.* (2012)[92]	GWAS	Ischaemic and haemorrhagic stroke	1 657 cases, 1 664 controls	Chinese	9p21.3 (ANRIL)
Matsushita *et al.* (2010)[93]	GWAS	Atherothrombotic stroke	2 775 cases, 2 839 controls	Japanese	ARHGEF 10
ISGC and WTCCC (2012)[94]	GWAS	Large-vessel stroke	3 548 cases, 5 972 controls	European	7p21.1 (HDAC9); replicated previous finding for cardio-embolic stroke near PITX2 and ZFHX3
Holliday *et al.* (2012)[95]	GWAS	Large-vessel stroke	1 162 cases, 1 244 controls	Australian	6p21.1
Cole *et al.* (2012)[3]	WES	Lacunar stroke	889 cases, 927 controls (10 for exome sequencing)	African American European American	4q21.1 (CSN3) **identified by exome sequencing following previous GWAS
Zhou *et al.* (2014)[96]	GWAS	Lacunar strokes systemic vasculopathy	9 subjects (exome sequencing)	European American, European	*ADA*2 gene

Table 3. Genetic studies of stroke in Africa				
First author (year)	*Study type*	*Stroke phenotype*	*Sample*	*Salient findings*
Saidi *et al.* (2007)[97]	Genotyping	Ischaemic stroke	135 cases, 118 controls (Tunisian)	Altered plasminogen activator inhibitor 1 (PAI-1) and tissue-type plasminogen activator (tPA) levels: Significant ↑ in PAI-1 and marked ↓ in tPA levels correlated with 4G/5G, but not with -844G/A, PAI-1 variants 4G/4G carriers had reduced risk of stroke compared with other genotypes
Saidi *et al.* (2007)[98]	Genotyping	Ischaemic stroke	216 cases, 282 controls (Tunisian)	ApoE ε3 lower (0.546 vs 0.736; $p < 0.001$) in stroke vs control ApoE ε4 higher (0.370 vs 0.181; $p < 0.001$) in stroke vs control Prevalence of Apo ε4-containing phenotypes higher in: • ischaemic versus haemorrhagic ($p < 0.001$) • small-vessel versus large-vessel stroke cases ($p < 0.001$) • increased need for statin drugs ($p = 0.040$).
Mourad *et al.* (2008)[105]	Genotyping	Sickle cell anaemia	20 SCA cases, 10 controls (Egyptian)	Presence or ACE D allele significantly predisposed to stroke in children with sickle cell anaemia (SCA).
Saidi *et al.* (2008)[99]	Genotyping	Ischaemic stroke	216 stroke patients, 318 controls (Tunisian)	Human platelet alloantigen (HPA) – 1 a/b ($p < 0.001$) and HPA-5 a/b ($p < 0.001$) alleles were associated with stroke-susceptible genotypes: 1a/b-2a/a-3a/b-4a/a-5a/b protective genotypes: 1a/a-2a/a-3a/a-4a/a-5a/a; 1a/a-2a/a-3a/b-4a/a-5a/a; 1a/b -2a/a-3a/a-4a/a-5a/a; 1a/b-2a/a-3a/b-4a/a-5a/a)
Saidi *et al.* (2008)[100]	Genotyping	Ischaemic stroke	329 cases, 444 controls	Lower human platelet alloantigen, HPA-1a ($p < 0.001$) and higher HPA-1b ($p < 0.001$) allele frequencies were seen in cases than control subjects. Homozygosity for HPA-1b ($p < 0.001$) alleles was more prevalent in stroke cases than in controls.
Saidi *et al.* (2009)[101]	Genotyping	Ischaemic stroke	228 cases, 323 controls	Frequency of APOE ε3 allele and Apo E3/E3 genotype lower ($p < 0.001$) in stroke vs controls Frequency of Apo ε4 allele and genotypes (E3/E4 and E4/E4) elevated ($p < 0.001$) in stroke vs controls Higher proportion of Apo ε4-carrying + ACE Del/Del positive cases seen in young (< 50 years) patients ($p = 0.012$) and associated with large-vessel stroke ($p = 0.035$).
Saidi *et al.* (2009)[102]	Genotyping	Ischaemic stroke	329 cases, 444 controls	Angiotensinogen AGT 174T/235M/-6A, AGT 174T/235T/-6G. AGT 174T/235T/-6A and AGT 174M/235T/-6A haplotypes were significantly associated with an increased risk of stroke.
Saidi *et al.* (2010)[103]			329 IS patients, 444 controls	Endothelial nitric oxide synthase (eNOS) gene polymorphisms (298Asp allele and 298Asp/4b/-786T and 298Asp/4b/-786C haplotypes, and in addition identified 298Asp/4a/-786T haplotypes) were significantly associated with ischaemic stroke.

not yet fully explored. The H3Africa Consortium, with funding support from the National Institutes of Health (NIH) and the Wellcome Trust, is currently executing 24 different disease-based projects involving 50 000 to 75 000 participants across the African continent.[11] This initiative will deeply enhance our understanding of human genomic variation while unravelling the genomic bases of several communicable and non-communicable diseases on the continent, while facilitating genomic infrastructural development and capacity building.

The H3Africa Consortium is revolutionising genomic research in Africa and closing the huge genomics gap between Africa and the developed world. The initiative will reduce health disparities and enhance understanding of health issues for the benefit of Africans and the human race through the discovery of new genes and disease pathways with therapeutic and preventative potentials.

The Stroke Investigative Research and Education Network (SIREN) project is one of the H3Africa-funded projects. The SIREN investigators propose to explore genomic factors in stroke in 6 000 native West Africans (3 000 case–control pairs) in comparison with 1 000 African Americans (80% of whom are of West African ancestral origin) and 12 000 Americans of European ancestry in the REGARDS study (comparison among three tracks).[116,117] The wide genomic variation of African populations offers a unique opportunity to identify novel genomic variants with causal relationships to stroke across different ethnic groups.

The SIREN project has three main streams: phenomics (including community engagement), genomics, and

bio-informatics (Fig. 1). An ethnically diverse sample increases the scope and generalisability of findings, because pan-ethnic
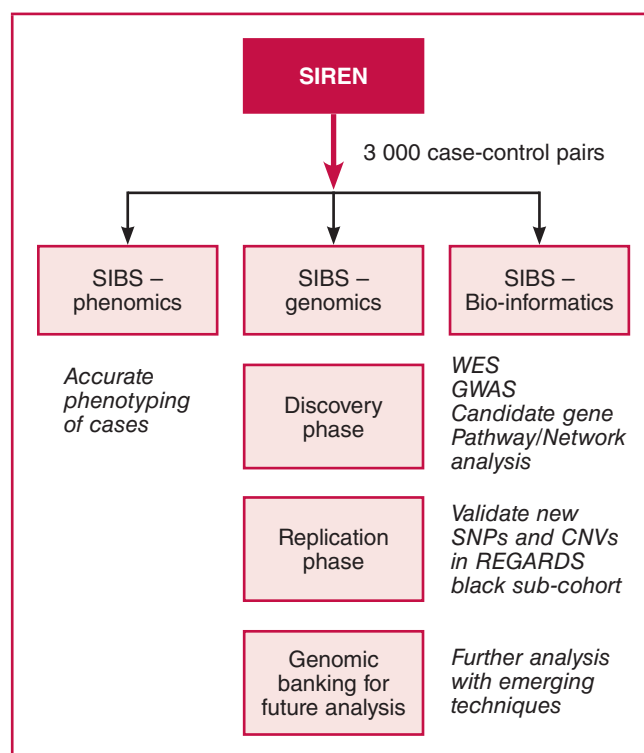

Fig 1. SIREN component projects.

Table 4. Unique features of SIREN meeting the standard criteria for stroke genomics studies

Criteria for Stroke Genomics Studies[81]	How met in SIREN – SIBS Genomics
Venice ' A' rating for sample size	large sample size > 3000 case – control pairs
Phenomic characterization	Rigorous phenotypic assessment in patients and in controls Detailed investigations (min of CT) and accurate classification using OCSP, TOAST, ASCO, CCS. Data verification and Quality control System
Control for confounders	Measurement and documentation of conventional vascular risk factors to be controlled for in the analysis
External validation	External Validation in REGARDS cohort (12,500)
Others	Low genotyping error rate (Hardy-Weinberg Equilibrium will be stated in cases and controls) Genomic controls, and other methods to account for population stratification, low P value (corrected for multiple testing)

replicability of association between a candidate SNP and trait outcome provides support for a causal relationship. However, the high levels of genomic diversity among Africans pose a potential challenge of false-positive associations due to population stratification, while heterogeneity of haplotype structure may reduce statistical power to detect true-positive signals by GWAS.

To combat these challenges, the SIREN project has been designed in compliance with the recommendations of Dichgans *et al*.[88] with due attention to adequate sample size, rigorous and accurate phenomic characterisation of cases, control for confounders, and planned validation of findings in an independent African American stroke population participating in the REGARDS study[29,116,117] (Table 4). In addition, the SIREN project will utilise GWAS approaches using customised chips including unique African variants,[10] whole-genome and whole-exome sequencing (WGS/WES) and other emergent high-throughput approaches for future analyses. Furthermore, 'pathway-based analysis' of genomic data[52] will chart new paths in our understanding of the molecular trajectories of stroke and unravel new options of stroke diagnostics and therapeutics in the emerging milieu of personalised medicine

Conclusion

Understanding the interaction between genetic and environmental conditions that predispose to stroke and impede favourable post-stroke outcomes is crucial for the formulation of targeted treatment strategies aimed at the successful prevention of and recovery from stroke. Unravelling the genomic underpinnings of stroke in populations of African ancestry will greatly improve our broad understanding of the molecular pathways of stroke and likely add substantially to ongoing efforts to mitigate the devastating global consequences of stroke.

The negative impact of cultural and religious beliefs, issues of autonomy of decision making and voluntary participation, as well as poor understanding of the health impact of genomics are potential challenges to translating genomic advances into real-world clinical applications in Africa. These suggest that caution should be exercised with regard to the expectations from stroke genomics research in Africa, while rigorous detection, evaluation, treatment and control of high blood pressure cannot be over-emphasised as a pragmatic strategy to curtail stroke in Africa.

References

1. Lanktree MB, Dichgans M, Hegele RA, Advances in genomic analysis of stroke: what have we learned and where are we headed? *Stroke* 2010; **41**(4): 825–832.

2. Bevan S, Markus HS. Genetics of common polygenic ischaemic stroke: current understanding and future challenges. *Stroke Res Treat* 2011; **2011**: 179061.

3. Cole JW, Stine OC, Liu X, Pratap A, Cheng Y, Tallon LJ, *et al*. Rare variants in ischemic stroke: an exome pilot study. *PLoS One* 2012; **7**(4): e35591.

4. Markus HS. Stroke genetics. *Hum Mol Genet* 2011; **20**(R2): R124–131.

5. Cole JW, Brown DW, Giles WH, Stine OC, O'Connell JR, Mitchell BD, *et al*. Ischemic stroke risk, smoking, and the genetics of inflammation in a biracial population: the stroke prevention in young women study. *Thromb J* 2008; **6**: 11.

6. Grewal RP, Dutra AV, Liao YC, Juo SH, Papamitsakis NI. The intron 4c allele of the NOS3 gene is associated with ischemic stroke in African Americans. *BMC Med Genet* 2007; **8**: 76.

7. Song Q, Cole JW, O'Connell JR, Stine OC, Gallagher M, Giles WH, *et al*. Phosphodiesterase 4D polymorphisms and the risk of cerebral infarction in a biracial population: the Stroke Prevention in Young Women Study. *Hum Mol Genet* 2006; **15**(16): 2468–2478.

8. Tishkoff SA, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, *et al*. The genetic structure and history of Africans and African Americans. *Science* 2009; **324**(5930): 1035–1044.

9. Shriner D, Tekola-Ayele F, Adeyemo A, Rotimi CN. Genome-wide genotype and sequence-based reconstruction of the 140,000 year history of modern human ancestry. *Sci Rep* 2014; **4**: 6055.

10. Gurdasani D, Arstensen T, Tekola-Ayele F, Pagani L, Tachmazidou I, Hatzikotoulas K, *et al*. The African Genome Variation Project shapes medical genetics in Africa. *Nature* 2015; **517**(7534): 327–332.

11. H3Africa Consortium. Research capacity. Enabling the genomic revolution in Africa. *Science* 2014; **344**(6190): 1346–138.

12. Owolabi MO, Mensah GA, Kimmel PL, Adu D, Ramsay M, Waddy SP, *et al*. Understanding the rise in cardiovascular diseases in Africa: harmonising H3Africa genomic epidemiological teams and tools. *Cardiovasc J Afr* 2014; **25**(3): 134–136.

13. Seshadri S, Wolf PA. Lifetime risk of stroke and dementia: current concepts, and estimates from the Framingham Study. *Lancet Neurol* 2007; **6**(12): 1106–1114.

14. Feigin VL, Lawes CM, Bennett DA, Barker-Collo SL, Parag V. Worldwide stroke incidence and early case fatality reported in 56 population-based studies: a systematic review. *Lancet Neurol* 2009; **8**(4): 355–69.

15. Owolabi MO. Taming the burgeoning stroke epidemic in Africa: stroke quadrangle to the rescue. *West Indian Med J* 2011; **60**(4): 412–421.

16. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006; **367**(9524): 1747–1757.

17. Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah G A, Connor M, Bennett DA, *et al*. Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. *Lancet* 2014; **383**(9913): 245–254.

18. Moran A, Forouzanfar M, Sampson U, Chugh S, Feigin V, Mensah G. The epidemiology of cardiovascular diseases in sub-Saharan Africa: the

Global Burden of Diseases, Injuries and Risk Factors 2010 Study. *Prog Cardiovasc Dis* 2013; **56**(3): 234–239.

19. Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: part I: general considerations, the epidemiologic transition, risk factors, and impact of urbanization. *Circulation* 2001; **104**(22): 2746–2753.

20. Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases: Part II: variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. *Circulation* 2001; **104**(23): 2855–2864.

21. Khedr EM, *et al*. Epidemiological study and risk factors of stroke in Assiut Governorate, Egypt: community-based study. *Neuroepidemiology* 2013; **40**(4): 288–294.

22. Akinyemi RO, Izzeldin IM, Dotchin C, Gray WK, Adeniji O, Seidi OA, *et al*. Contribution of noncommunicable diseases to medical admissions of elderly adults in Africa: a prospective, cross-sectional study in Nigeria, Sudan, and Tanzania. *J Am Geriatr Soc* 2014; **62**(8): 1460–1466.

23. Owolabi LF, Shehu MY, Shehu MN, Fadare J. Pattern of neurological admissions in the tropics: Experience at Kano, Northwestern Nigeria. *Ann Indian Acad Neurol* 2010; **13**(3): 167–170.

24. Akinyemi RO, Allan L, Owolabi MO, Akinyemi JO, Ogbole G, Ajani A et al. Profile and determinants of vascular cognitive impairment in African stroke survivors: The CogFAST Nigeria Study. *J Neurol Sci* 2014; **346**(1–2): 241–249.

25. Akinyemi RO, Mukaetova-Ladinska EB, Attems J, Ihara M, Kalaria RN. Vascular risk factors and neurodegeneration in ageing related dementias: Alzheimer's disease and vascular dementia. *Curr Alzheimer Res* 2013; **10**(6): 642–53.

26. Owolabi MO. Impact of stroke on health-related quality of life in diverse cultures: the Berlin–Ibadan multicenter international study. *Health Qual Life Outcomes* 2011; **9**: 81.

27. O'Donnell MJ. Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, *et al*. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. *Lancet* 2010; **376**(9735): 112–123.

28. Adeloye D, Basquill C. Estimating the prevalence and awareness rates of hypertension in Africa: a systematic analysis. *PLoS One* 2014; **9**(8): e104300.

29. Howard VJ, Kleindorfer DO, Judd SE, McClure LA, Safford MM, Rhodes JD, *et al*. Disparities in stroke incidence contributing to disparities in stroke mortality. *Ann Neurol* 2011; **69**(4): 619–627.

30. Gutierrez J, Williams OA. A decade of racial and ethnic stroke disparities in the United States. *Neurology* 2014; **82**(12): 1080–1082.

31. Douiri A, Rudd AG, Wolfe CD. Prevalence of poststroke cognitive impairment: South London Stroke Register 1995–2010. *Stroke* 2013; **44**(1): 138–145.

32. Rotimi CN, Jorde LB. Ancestry and disease in the age of genomic medicine. *N Engl J Med* 2010; **363**(16): 1551–1558.

33. Saab KR, Kendrick J, Yracheta JM, Lanaspa MA, Pollard M, Johnson RJ. New insights on the risk for cardiovascular disease in African Americans: The role of added sugars. *J Am Soc Nephrol* 2015; **26**(2): 247–257.

34. Burchard EG, Ziv E, Coyle N, Gomez SL,Tang H, Karter AJ, *et al*. The importance of race and ethnic background in biomedical research and clinical practice. *N Engl J Med* 2003; **348**(12): 1170–1175.

35. Collins FS. What we do and don't know about 'race', 'ethnicity', genetics and health at the dawn of the genome era. *Nat Genet* 2004; **36**(Suppl 11): S13–15.

36. Cooper RS, Kaufman JS, Ward R, Race and genomics. *N Engl J Med* 2003; **348**(12): 1166–1170.

37. Adeyemo AA, Chen G, Chen Y, Rotimi C. Genetic structure in four West African population groups. *BMC Genet* 2005. **6**: 38.

38. Ramsay M, Tiemessen CT, Choudhury A, Soodyall H. Africa: the next frontier for human disease gene discovery? *Hum Mol Genet* 2011; **20**(R2): R214–220.

39. Gomez F, Hirbo J, Tishkoff SA. Genetic variation and adaptation in Africa: implications for human evolution and disease. *Cold Spring Harb Perspect Biol* 2014; **6**(7): a008524.

40. International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**(7164): 851–861.

41. Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* 2012; **491**(7422): 56–65.

42. Graffe CC, Bech JN, Lauridsen TG, Vase H, Pedersen EB. Abnormal increase in urinary aquaporin-2 excretion in response to hypertonic saline in essential hypertension. *BMC Nephrol* 2012; **13**: 15.

43. Beuschlein F, Boulkroun S, Osswald A, Wieland T, Nielsen HN, Lichtenauer UD, *et al*. Somatic mutations in ATP1A1 and ATP2B3 lead to aldosterone-producing adenomas and secondary hypertension. *Nat Genet* 2013; **45**(4): 440–4, 444e1–2.

44. Bryc K, Auton A, Nelson MR, Oksenberg JR, Hauser SL, Williams S, *et al*. Genome-wide patterns of population structure and admixture in West Africans and African Americans. *Proc Natl Acad Sci USA* 2010; **107**(2): 786–791.

45. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. The Genetic Ancestry of African Americans, Latinos, and European Americans across the United States. *Am J Hum Genet* 2015; **96**(1): 37–53.

46. Dichgans M. Genetics of ischaemic stroke. *Lancet Neurol* 2007; **6**(2): 149–161.

47. Morrison AC, Fornage M, Liao D, Boerwinkle E. Parental history of stroke predicts subclinical but not clinical stroke: the Atherosclerosis Risk in Communities Study. *Stroke* 2000; **31**(9): 2098–2102.

48. Liao D, Myers R, Hunt S, Shahar E, Paton C, Burke G, *et al*. Familial history of stroke and stroke risk. The Family Heart Study. *Stroke* 1997; **28**(10): 1908–1912.

49. Flossmann E, Schulz UG, Rothwell PM. Systematic review of methods and results of studies of the genetic epidemiology of ischemic stroke. *Stroke* 2004; **35**(1): 212–227.

50. Minor DS, Wofford MR, Jones DW. Racial and ethnic differences in hypertension. *Curr Atheroscler Rep* 2008; **10**(2): 121–127.

51. Adeyemo A, Gerry N, Chen G, Herbert A, Doumatey A, Huang H, *et al.* A genome-wide association study of hypertension and blood pressure in African Americans. *PLoS Genet* 2009; **5**(7): e1000564.

52. Reder NP, Tayo BO, Salako B, Ogunniyi A, Adeyemo A, Rotimi C, *et al.* Adrenergic alpha-1 pathway is associated with hypertension among Nigerians in a pathway-focused analysis. *PLoS One* 2012; **7**(5): e37145.

53. Rice T, Cooper RS, Wu X, Bouchard C, Rankinen T, Rao DC, *et al.* Meta-analysis of genome-wide scans for blood pressure in African American and Nigerian samples. The National Heart, Lung, and Blood Institute GeneLink Project. *Am J Hypertens* 2006; **19**(3): 270–274.

54. Sorensen IF, Vazquez AI, Irvin MR, Sorensen P, Davis BR, Ford CE, *et al*. Pharmacogenetic effects of 'candidate gene complexes' on stroke in the GenHAT study. *Pharmacogenet Genomics* 2014; **24**(11): 556–563.

55. Adeleye JO, Abbiyesuku FM. Glucose and insulin responses in offspring of Nigerian Type 2 diabetics. *Afr J Med Med Sci* 2002; **31**(3): 253–257.

56. Mbanya JC, Pani LN, Mbanya DN, Sobngwi E, Ngogang J, *et al*. Reduced insulin secretion in offspring of African type 2 diabetic parents. *Diabetes Care* 2000; **23**(12): 1761–1765.

57. Tekola-Ayele F, Adeyemo AA, Rotimi CN. Genetic epidemiology of type 2 diabetes and cardiovascular diseases in Africa. *Prog Cardiovasc*

Dis 2013; **56**(3): 251–260.

58. Chen G, Adeyemo A, Zhou J, Chen Y, Huang H, Doumatey A, *et al.* Genome-wide search for susceptibility genes to type 2 diabetes in West Africans: potential role of C-peptide. *Diabetes Res Clin Pract* 2007; **78**(3): e1–6.

59. Ovbiagele B. Impairment in glomerular filtration rate or glomerular filtration barrier and occurrence of stroke. *Arch Neurol* 2008; **65**(7): 934–938.

60. Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW, *et al. MYH9* is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet* 2008; **40**(10): 1175–1184.

61. Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, Freedman BI, *et al.* Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 2010; **329**(5993): 841–845.

62. Tayo BO, Kramer H, Salako BL, Gottesman O, McKenzie CA, Ogunniyi A, *et al.* Genetic variation in APOL1 and MYH9 genes is associated with chronic kidney disease among Nigerians. *Int Urol Nephrol* 2013; **45**(2): 485–494.

63. Tzur S, Rosset S, Shemer R, Yudkovsky G, Selig S, Tarekegn A, *et al.* Missense mutations in the *APOL1* gene are highly associated with end-stage kidney disease risk previously attributed to the *MYH9* gene. *Hum Genet* 2010; **128**(3): 345–350.

64. Bentley AR, Doumatey AP, Chen G, Huang H, Zhou J, Shriner D, *et al.* Variation in APOL1 Contributes to Ancestry-Level Differences in HDLc-Kidney Function Association. *Int J Nephrol* 2012; **2012**: 748984.

65. Van der Sijde MR, Ng A, Fu J. Systems genetics: From GWAS to disease pathways. *Biochim Biophys Acta* 2014; **1842**(10): 1903–1909.

66. Atwood LD, Wolf PA, Heard-Costa NL, Massaro JM, Beiser A, D'Agostino RB, *et al.* Genetic variation in white matter hyperintensity volume in the Framingham Study. *Stroke* 2004; **35**(7): 1609–1613.

67. Moskau S, Golla A, Grothe C, Boes M, Pohl C, Klockgether T. Heritability of carotid artery atherosclerotic lesions: an ultrasound study in 154 families. *Stroke* 2005; **36**(1): 5–8. Epub 2004 Nov 29.

68. Tsaras G, Owusu-Ansah A, Boateng FO, Amoateng-Adjepong Y. Complications associated with sickle cell trait: a brief narrative review. *Am J Med* 2009; **122**(6): 507–512.

69. Piel FB, atil AP, Howes RE, Nyangiri OA, Gething PW, Williams TN, *et al.* Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis. *Nat Commun* 2010; **1**: 104.

70. Lemos Cardoso G, Farias Guerreiro J. African gene flow to north Brazil as revealed by HBB*S gene haplotype analysis. *Am J Hum Biol* 006; **18**(1): 93–98.

71. Osuntokun BO, Osuntokun O. Complicated migraine and haemoglobin AS in Nigerians. *Br Med J* 1972; **2**(5814): 621–622.

72. Ohene-Frempong K, Weiner SJ, Sleeper LA, Miller ST, Embury S, Moohr JW, *et al.* Cerebrovascular accidents in sickle cell disease: rates and risk factors. *Blood* 1998; **91**(1): 288–294.

73. Flanagan JM, Frohlich DM, Howard TA, Schultz WH, Driscoll C, Nagasubramanian R, *et al.* Genetic predictors for stroke in children with sickle cell anemia. *Blood* 2011; **117**(24): 6681–6684.

74. Flanagan JM, heehan V, Linder H, Howard TA, Wang YD, Hoppe CC, *et al.* Genetic mapping and exome sequencing identify 2 mutations associated with stroke protection in pediatric patients with sickle cell anemia. *Blood* 2013; **121**(16): 3237–3245.

75. Ruchoux MM, Maurage CA. CADASIL: Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. *J Neuropathol Exp Neurol* 1997; **56**(9): 947–964.

76. Yamamoto Y, Craggs L, Baumann M, Kalimo H, Kalaria RN. Review: molecular genetics and pathology of hereditary small vessel diseases of the brain. *Neuropathol Appl Neurobiol* 2011; **37**(1): 94–113.

77. Lee SJ, Meng H, Elmadhoun O, Blaivas M, Wang MM. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy affecting an African American man: identification of a novel 15-base pair NOTCH3 duplication. *Arch Neurol* 2011; **68**(12): 1584–1586.

78. Meschia JF, Worrall BB, Rich SS. Genetic susceptibility to ischemic stroke. *Nat Rev Neurol* 2011; **7**(7): 369–378.

79. Craig HD, Gunel M, Cepeda O, Johnson EW, Ptacek L, Steinberg GK, *et al.* Multilocus linkage identifies two new loci for a Mendelian form of stroke, cerebral cavernous malformation, at *7p15-13* and *3q25.2-27*. *Hum Mol Genet* 1998;. **7**(12): 1851–1858.

80. Janunger T, Nilsson-Ardnor S, Wiklund PG, Lindgren P, Escher SA, Lackovic K, *et al.* A novel stroke locus identified in a northern Sweden pedigree: linkage to chromosome 9q31-33. *Neurology* 2009; **73**(21): 1767–1773.

81. Nilsson-Ardnor S, Janunger T, Wiklund PG, Lackovic K, Nilsson AK, Lindgren P, *et al.* Genome-wide linkage scan of common stroke in families from northern Sweden. *Stroke* 2007; **38**(1): 34–40.

82. Wang XY, Wang JW, Tang X, Li N, Chen DF, Wu YQ, *et al.* Linkage and association between interleukin-6 gene polymorphisms and ischemic stroke: a family-based study in the northern Chinese Han population. *Genet Test Mol Biomarkers* 2014; **18**(11): 761–766.

83. Liu B, Zhang L, Yang Q. Genetics of intracerebral hemorrhage: Insights from candidate gene approaches. *Neurol India* 2012; **60**(1): 3–8.

84. Biffi A, Anderson CD, Jagiella JM, Schmidt H, Kissela B, Hansen BM, *et al.* APOE genotype and extent of bleeding and outcome in lobar intracerebral haemorrhage: a genetic association study. *Lancet Neurol* 2011; **10**(8): 702–709.

85. Brouwers HB, Biffi A, Ayres AM, Schwab K, Cortellini L, Romero JM, *et al.* Apolipoprotein E genotype predicts hematoma expansion in lobar intracerebral hemorrhage. *Stroke* 2012; **43**(6): 1490–1495.

86. Premkumar DR, Cohen DL, Hedera P, Friedland RP, Kalaria RN. Apolipoprotein E-epsilon4 alleles in cerebral amyloid angiopathy and cerebrovascular pathology associated with Alzheimer's disease. *Am J Pathol* 1996; **148**(6): 2083–2095.

87. Kalaria RN, Ogeng'o JA, Patel NB, Sayi JG, Kitinya JN, Chande HM, *et al.* Evaluation of risk factors for Alzheimer's disease in elderly east Africans. *Brain Res Bull* 1997; **44**(5): 573–577.

88. Dichgans M, Markus HS. Genetic association studies in stroke: methodological issues and proposed standard criteria. *Stroke* 2005; **36**(9): 2027–2031. Epub 2005 Jul 28.

89. Meschia JF. Stroke genome-wide association studies: the large numbers imperative. *Stroke* 2010; **41**(4): 579–580.

90. Markus HS, Alberts MJ. Update on genetics of stroke and cerebrovascular disease 2005. *Stroke* 2006; **37**(2): 288–290. Epub 2006 Jan 12.

91. Meschia JF, Nalls M, Matarin M, Brott TG, Brown RD Jr, Hardy J, *et al.* Siblings with ischemic stroke study: results of a genome-wide scan for stroke loci. *Stroke* 2011; **42**(10): 2726–2732.

92. Hata J, Kubo M, Kiyohara Y. [Genome-wide association study for ischemic stroke based on the Hisayama study]. *Nihon Eiseigaku Zasshi* 2011; **66**(1): 47–52.

93. Matarin M, Brown WM, Dena H, Britton A, De Vrieze FW, Brott TG, *et al.* Candidate gene polymorphisms for ischemic stroke. *Stroke* 2009; **40**(11): 3436–3442.

94. Gretarsdottir S, Thorleifsson G, Manolescu A, Styrkarsdottir U, Helgadottir A, Gschwendtner A, *et al.* Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol* 2008; **64**(4): 402–409.

95. Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, Thorgeirsson G, *et al.* A sequence variant in ZFHX3 on 16q22

associates with atrial fibrillation and ischemic stroke. *Nat Genet* 2009; **41**(8): 876–878.

96. Bilguvar K, Yasuno K, Niemelä M, Ruigrok YM, von Und Zu Fraunberg M, van Duijn CM, *et al*. Susceptibility loci for intracranial aneurysm in European and Japanese populations. *Nat Genet* 2008; **40**(12): 1472–1477.

97. Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, *et al*. Genome-wide association studies of stroke. *N Engl J Med* 2009; **360**(17): 1718–1728.

98. Yamada Y, Fuku N, Tanaka M, Aoyagi Y, Sawabe M, Metoki N, *et al*. Identification of CELSR1 as a susceptibility gene for ischemic stroke in Japanese individuals by a genome-wide association study. *Atherosclerosis* 2009; **207**(1): 144–149.

99. Zhang W, Chen Y, Liu P, Chen J, Song L, Tang Y, *et al*. Variants on chromosome 9p21.3 correlated with ANRIL expression contribute to stroke risk and recurrence in a large prospective stroke population. *Stroke* 2012; **43**(1): 14–21.

100. Matsushita T, Ashikawa K, Yonemoto K, Hirakawa Y, Hata J, Amitani H, *et al*. Functional SNP of ARHGEF10 confers risk of atherothrombotic stroke. *Hum Mol Genet* 2010; **19**(6): 1137–1146.

101. International Stroke Genetics Consortium (ISGC); Wellcome Trust Case–Control Consortium 2 (WTCCC2), Bellenguez C, Bevan S, Gschwendtner A, Spencer CC, *et al*. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet* 2012; **44**(3): 328–333.

102. Holliday EG, Maguire JM, Evans TJ, Koblar SA, Jannes J, Sturm JW, *et al*. Common variants at 6p21.1 are associated with large artery atherosclerotic stroke. *Nat Genet* 2012; **44**(10): 1147–1151.

103. Zhang Y, Tong Y, Zhang Y, Ding H, Zhang H, Geng Y, *et al*. Two novel susceptibility SNPs for ischemic stroke using exome sequencing in Chinese Han population. *Mol Neurobiol* 2014; **49**(2): 852–862.

104. Saidi S, Slamia LB, Mahjoub T, Ammou SB, Almawi WY. Association of PAI-1 4G/5G and -844G/A gene polymorphism and changes in PAI-1/tPA levels in stroke: a case-control study. *J Stroke Cerebrovasc Dis* 2007; **16**(4): 153–9.

105. Saidi S, Slamia LB, Ammou SB, Mahjoub T, Almawi WY. Association of apolipoprotein E gene polymorphism with ischemic stroke involving large-vessel disease and its relation to serum lipid levels. *J Stroke Cerebrovasc Dis* 2007; **16**(4): 160–166.

106. Saidi S, Mahjoub T, Slamia LB, Ammou SB, Al-Subaie AM, Almawi WY. Association of human platelet alloantigen 1 through 5 polymor-phisms with ischemic stroke. *Cerebrovasc Dis* 2008; **25**(1-2): 81–86.

107. Saidi S, Mahjoub T, Slamia LB, Ammou SB, Al-Subaie AM, Almawi WY. Polymorphisms of the human platelet alloantigens HPA-1, HPA-2, HPA-3, and HPA-4 in ischemic stroke. *Am J Hematol* 2008; **83**(7): 570–573.

108. Saidi S, Zammiti W, Slamia LB, Ammou SB, Almawi WY, Mahjoub T. Interaction of angiotensin-converting enzyme and apolipoprotein E gene polymorphisms in ischemic stroke involving large-vessel disease. *J Thromb Thrombol* 27(1): 68–74. Epub 2007 Nov 21.

109. Saidi S, Mallat SG, Almawi WY, Mahjoub T. Association between renin–angiotensin–aldosterone system genotypes and haplotypes and risk of ischemic stroke of atherosclerotic etiology. *Acta Neurol Scand* 2009. **119**(6): 356–363.

110. Saidi S, Mallat SG, Almawi WY, Mahjoub T. Endothelial nitric oxide synthase Glu298Asp, 4b/a, and *-786T>C* gene polymorphisms and the risk of ischemic stroke. *Acta Neurol Scand* 2010; **121**(2): 114–119.

111. Saidi S, Mahjoub T, Almawi WY. Aldosterone synthase gene (*CYP11B2*) promoter polymorphism as a risk factor for ischaemic stroke in Tunisian Arabs. *J Renin Angiotensin Aldosterone Syst* **11**(3): 180–186.

112. Mourad H, Fadel W, El Batch M, Rowisha M. Heamostatic and genetic predisposing factors for stroke in children with sickle cell anemia. *Egypt J Immunol* 2008; **15**(1): 25–37.

113. Marshall PA, Adebamowo CA, Adeyemo AA, Ogundiran TO, Strenski T, Zhou J, *et al*. Voluntary participation and comprehension of informed consent in a genetic epidemiological study of breast cancer in Nigeria. *BMC Med Ethics* 2014; **15**: 38.

114. Olaitan PB, Odesina V, Odesina V, Ademola S, Fadiora SO, Oluwatosin OM, Reichenberger EJ. Recruitment of Yoruba families from Nigeria for genetic research: experience from a multisite keloid study. *BMC Med Ethics* 2014; **15**: 65.

115. Fagbemiro L, Adebamowo C. Knowledge and attitudes to personal genomics testing for complex diseases among Nigerians. *BMC Med Ethics* 2014; **15**: 34.

116. Howard G, Wagenknecht LE, Kernan WN, Cushman M, Thacker EL, Judd SE, *et al*. Racial differences in the association of insulin resistance with stroke risk: the REasons for Geographic And Racial Differences in Stroke (REGARDS) study. *Stroke* 2014; **45**(8): 2257–2262.

117. Howard VJ, Cushman M, Pulley L, Gomez CR, Go RC, Prineas RJ, *et al*. The reasons for geographic and racial differences in stroke study: objectives and design. *Neuroepidemiology* 2005; **25**(3): 135–143.