Review

Public Health Genomics

Public Health Genomics 2015;18:40–51
DOI: 10.1159/000367962

Received: March 28, 2014
Accepted after revision: August 29, 2014
Published online: November 26, 2014

Genome-Wide Association Studies in Africans and African Americans: Expanding the Framework of the Genomics of Human Traits and Disease

Emmanuel Peprah[a] Huichun Xu[c] Fasil Tekola-Ayele[b] Charmaine D. Royal[d]

[a]Center for Translation Research and Implementation Science, National Heart, Lung, and Blood Institute, and [b]Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, Md., and [c]Division of Endocrinology, Diabetes and Nutrition, Department of Medicine, University of Maryland School of Medicine, Baltimore, Md., and [d]Department of African and African American Studies, Duke University, Durham, N.C., USA

Key Words
African American · African ancestry · African populations · Genetic conditions · Genome-wide association studies · Genomics · Global health · Health

Abstract
Genomic research is one of the tools for elucidating the pathogenesis of diseases of global health relevance and paving the research dimension to clinical and public health translation. Recent advances in genomic research and technologies have increased our understanding of human diseases, genes associated with these disorders, and the relevant mechanisms. Genome-wide association studies (GWAS) have proliferated since the first studies were published several years ago and have become an important tool in helping researchers comprehend human variation and the role genetic variants play in disease. However, the need to expand the diversity of populations in GWAS has become increasingly apparent as new knowledge is gained about genetic variation. Inclusion of diverse populations in genomic studies is critical to a more complete understanding of human variation and elucidation of the underpinnings of complex diseases. In this review, we summarize the available data on GWAS in recent African ancestry populations within the western hemisphere (i.e. African Americans and peoples of the Caribbean) and continental African populations. Furthermore, we highlight ways in which genomic studies in populations of recent African ancestry have led to advances in the areas of malaria, HIV, prostate cancer, and other diseases. Finally, we discuss the advantages of conducting GWAS in recent African ancestry populations in the context of addressing existing and emerging global health conditions.

Introduction

Since the first genome-wide association studies (GWAS) publication a few years ago, there has been a significant increase in studies that use GWAS as a tool to identify genetic variants associated with communicable and non-communicable diseases. GWAS have also been used to identify variants that influence human traits including eye color, hair color and anthropometric characteristics [1–3]. Genomic variations have been associated with drug resistance and treatment outcome [4]. The number of GWAS and pharmacogenomics studies that have identified single nucleotide changes associated with

1662–4246/14/0181–0040$39.50/0

E-Mail karger@karger.com
www.karger.com/phg

Emmanuel Peprah, PhD
Center for Translation Research and Implementation Science
National Heart, Lung, and Blood Institute, National Institutes of Health
6705 Rockledge Drive, Suite 6070, Bethesda, MD 20817 (USA)
E-Mail peprahek@mail.nih.gov

treatment outcomes is rising [5–7]. In other areas of research, genomic studies have led to a greater comprehension of disease progression and pathogenesis in Crohn's disease and several neurological disorders [8, 9]. Increased identification of variants associated with monogenic and polygenic disorders could contribute to more accurate diagnosis, effective clinical management, and specialized treatment improving the health of individuals and populations affected by these disorders [10, 11].

An Overview of Diversity in GWAS Data

Despite the expansion of GWAS tools to identify disease- or trait-related variants in human populations, adequate studies inclusive of diverse populations are lacking. Previous assessments of the representation of non-European populations in GWAS utilizing publically available databases have revealed significant underrepresentation of these populations in published GWAS [12–14]. We queried the Research Portfolio Online Reporting Tools (http://projectreporter.nih.gov/reporter.cfm) to determine the numbers of NIH-funded GWAS focused on or utilizing non-European populations. Our search yielded 2,267 current NIH-funded studies matching the search term 'genome-wide association studies' as of mid 2014. For a more accurate comparison, we used search terms such as African American, Hispanic, or Jewish ancestry also utilized by aforementioned publications. For ongoing studies (e.g. 2011) in the National Institutes of Health Research Portfolio Online Reporting Tools, GWAS accounted for only ~14, ~3, and <1% for African American, Hispanic, or Jewish ancestry populations, respectively. Subsequent analysis of published GWAS from PubMed using the same search terms produced 4,942 publications, of which African Americans accounted for ~3% with Hispanics and Jewish accounting for <1%, which are essentially in agreement with the previous reports. From the 2011 data, African Americans represented the largest non-European population included in GWAS.

Recently, several GWAS have included African ancestry populations in their discovery (initial stage) or replication datasets. From January 2009 to May 2010, 21 such studies have been included in the National Human Genome Research Institute (NHGRI) GWAS catalogue [13]. We chose to use the NHGRI GWAS catalogue for our review on African Americans and African ancestry populations because the catalogue includes literature that meets stringent inclusion criteria (http://www.genome.gov/gwastudies/). Moreover, analysis using the NHGRI GWAS catalogue produced results that reflect paucity of data for the African American population. In this review, we examine and briefly summarize several GWAS in diasporic African populations and continental African populations from publications in the NHGRI GWAS catalogue until mid 2014.

Advantages of Conducting GWAS in African Americans and African Ancestry Populations

Arguments have been made for broader inclusion of populations of African ancestry in GWAS [12–16]. With the underrepresentation of African ancestry populations in GWAS, one might ask: what is the utility in genetically characterizing populations of African ancestry? The occurrence of some genetic variants shows considerable frequency variation across populations which also includes the frequency of the risk allele, frequency of causal and correlated variants, and prevalence of diseases. Genetic determinants of disease and their effect sizes have also been shown to vary significantly between European and non-European populations such as populations of African ancestry [12]. It has been demonstrated that variants associated with diseases found in European ancestry populations do not always replicate in non-European populations [9, 17, 18]. GWAS conducted in European populations have frequently failed replication in non-European populations for several reasons, including differences in allelic architecture, linkage disequilibrium, and confounding of environmental factors across populations [14]. As the continent where the human species originated, Africa harbors populations with longer histories compared to European populations [19, 20]. GWAS provide opportunities for characterization of longer population histories and greater genetic heterogeneity found in African populations, allowing more accurate construction of ancestral haplotypes which cannot occur in non-African populations [21]. If one cannot determine the original haplotype, African populations offer opportunities to determine the evolutionary histories of these variants [22]. For example, the genetic origins of fragile X syndrome via haplotype analysis of the X chromosome have been well characterized in European populations [23]. However, the haplotypic origins of this mutation in African populations are yet unknown, even though fragile X syndrome has been found in an African population [24]. This syndrome and other Mendelian disorders offer a model to understand the evolutionary history of monogenetic dis-

orders as humans migrated out of Africa. The information could also contribute to greater understanding of recombination and evolution of genes in the X chromosome. Thus, inclusion of populations of African ancestry and other diverse populations in GWAS is critical to understanding human variation in both monogenetic and more complex disorders.

Notable Discoveries from GWAS in African American and African Ancestry Populations

Non-Communicable Diseases

GWAS have had considerable influence on the identification of genetic variants associated with non-communicable diseases. We have identified several GWAS on non-communicable diseases and also traits that have been conducted with African ancestry populations. Publications from these studies are listed in online supplementary table 1 (for all online suppl. material, see www.karger.com/doi/10.1159/000367962). Asthma is a complex non-communicable disease that significantly affects both the young and old. Approximately 15 million Americans are affected, with ~5,500 deaths attributed to complications [25]. Hospitalizations for asthma are 2–3 times higher for African Americans compared to other ethnic groups [25, 26]. GWAS exploring the genetic basis of this multifactorial disease in African ancestry populations have identified several associated genes (online suppl. table 1). Susceptibility loci for asthma were first identified in children of European ancestry (from Europe and the US) and then replicated in African American children [27]. Variants in the genes *CRB1* and *PDE4D* were associated with asthma in children of European ancestry [27, 28]. For African American populations, association with asthma was found for variation in the gene *DENND1B*, but not in *CRB1* or *PDE4D* [27, 28]. SNP variants found in intergenic regions near *ADRA1B, PRNP, GNA13*, and *DPP10* have been implicated in asthma and allergic diseases in African American and African Caribbean populations, but have not been replicated in European populations [29]. PDE4D, a regulator of airway smooth muscle contractility has been replicated in several European and Hispanic populations, but not in populations of recent African origin [28]. Over 43 genes associated with asthma have been identified [30]. However, few of them have been replicated via subsequent GWAS in other populations [30].

Asthma is a serious chronic disorder characterized by airway inflammation [31]. The inflammatory process in asthma is characterized by a significant elevation of white blood cell count (e.g. eosinophils, neutrophils and other white blood cells) [32]. GWAS have provided evidence that variation in white blood cell count in African Americans has a genetic component. GWAS have replicated genes on 17q21.1 in 2 cohorts of African Americans [33, 34]. Interestingly, 17q21.1 is also associated with inflammation and with childhood asthma [35]. This provides evidence suggesting that in some disorders the genetics of traits and diseases could be similar, as observed in white blood cell count, asthma and 17q21.1 [33].

Unlike asthma, the primary genetic variant underlying sickle cell disease (SCD) has been identified, but the tools of GWAS are being used to determine the nature of the genetic modifiers associated with phenotypic diversity in SCD. Genes implicated in SCD severity include *BCL11A*, which has been shown to modulate fetal hemoglobin levels and has been confirmed in GWAS in African Americans [36]. Sex-stratified analysis indicated *GLP2R* is an important modifier of fetal hemoglobin in males [36]. There has not been an established connection between SCD and asthma. Evidence supports asthma and SCD as distinct comorbid conditions [37]. However, some argue that the high prevalence of asthma in individuals with SCD suggests that the underlying mechanisms of the 2 diseases might be similar [38, 39]. GWAS of sickle cell anemia in African Americans have produced associations with *OTUB3* and *KCNK6* [40]. *KCNK6,* a member of the K^+ channel protein family, is expressed in cardiomyocytes and in airway epithelial cells [41, 42]. *KCNK6* also showed differential expression in pulmonary artery endothelial cells when exposed to sickle cell plasma [40]. This suggests an uncharacterized role of *KCNK6* in SCD. For both asthma and SCD, the complexity of the phenotype would suggest that genetic heterogeneity is a major contributor to both diseases, so careful classifications of individuals based on clinical subphenotype or other classification schemes and the use of appropriate genetically matched controls might be worthwhile.

GWAS have identified several loci that are associated with prostate cancer (online suppl. table 2). Eeles et al. [43] found 9 SNPs in 7 genomic regions dispersed among 5 chromosomes in patients with prostate cancer. Some of these variants were replicated in GWAS of African American males. However, the majority of associations were not replicated, which may have been due to the small sample sizes of the 2 studies [44]. Several of these genes including *TET2, NKX3.1* and *ITGA6* have been found to have oncogenic properties. *ITGA6* is necessary for tumor formation in breast cancer cells, but its expression is down-

regulated in cervical cancer [45, 46]. Some groups have reported TET2 associated with hematological cancers including myeloid leukemia and thrombocytosis [47–49]. *NKX3.1* encodes an androgen-regulated transcription factor expressed in luminal epithelial cells in the prostate [50]. The function of NKX3.1 is to prevent other genetic insults from initiating tumor growth because overexpression of NKX3.1 in vitro inhibits cell proliferation and formation of anchorage-independent cell groups, which has been associated with early stages of prostate tumorigenesis [51]. Additional evidence suggests that the loss of NKX3.1 leads to aberrant regulation of gene clusters, which is a major initiating factor for prostate tumorigenesis [50]. NKX3.1 has an indispensable role in prostate cancer and is regarded as the 'gatekeeper' to prevent tumorigenesis [50, 52]. Subsequent GWAS using Tobagonian men as the replication sample have not found an association between prostate tumor development and TGFBR3, a tumor suppressor with a significant role in prostate tumor development [53, 54]. These data suggest that the mechanism which produces tumorigenesis in prostate cells or other cell types could have different associated genetic variants in some non-European populations.

Recently, a GWAS and a family-based association study were conducted on podoconiosis (geochemical non-filarial elephantiasis of the lower legs). So far, this study is the only published GWAS on non-communicable diseases among continental African population groups. The findings have shown that variants in *HLA-DRB1*, *HLA-DQB1* and *HLA-DQA1* genes are associated with risk of susceptibility to podoconiosis [55]. Further evidence for the HLA class II association was found with HLA typing that showed *HLA-DRB1*0701*, *HLA-DQA1*0201* and *HLA-DQB1*0202* are risk variants. These alleles may have a functional role in presenting antigens (currently unknown) to T cells and imply that podoconiosis may be a T-cell-mediated inflammatory disease [55].

In the past decades, considerable progress towards understanding the neurogenetics of social behavior has used twin studies to establish the heritability of several behavioral phenotypes [56]. These studies have produced substantial amounts of data including the association of oxytocin and vasopressin with several aspects of social behavior and also suggest that structural variants might predispose families or individuals to such behaviors and neuropsychiatric disorders [56–58]. Many of these studies need replication in non-European populations.

Numerous reports have explored the genetic underpinnings of schizophrenia (online suppl. table 1). This psychiatric disorder is a complex disease with heredity often playing a substantial role in predisposition to the disorder [59]. Several genes have been implicated in schizophrenia; however, previous findings have not been replicated, causing some to argue that new approaches and fundamental rethinking of schizophrenia are required [60, 61]. For example, *HIST1H2AG, HIST1H2AB* and other genes are implicated in schizophrenia in African Americans and populations of European ancestry without achieving genome-wide significance [62]. Bipolar disorder is a neuropsychiatric disorder characterized by episodic mania and depression. Similar to schizophrenia, bipolar disorder has a significant genetic component with heritability of ~80% [63]. GWAS of this disorder in African Americans and European ancestry populations yielded an association with *DPY19L3* which was found to be not significant after correcting for genome-wide multiple testing [64].

Several groups have used GWAS to understand the genetic components of dependent behaviors ranging from alcohol dependence to smoking behavior [65]. In European ancestry populations, a variant located in a potential transcription factor-binding site upstream of *IL15* is correlated to smoking frequency [65]. Cigarette smoke has been found to have decreased *IL15* expression and diminished activation of downstream signaling molecules [66]. It would be interesting to look at the mRNA/protein levels of these individuals to determine the expression of IL15. This would suggest a mechanism of smoking behavior that could also be associated with interleukins characterized in immunity [67]. If IL15 is truly associated with smoking behavior, therapeutic interventions could be developed based on the characterized function of IL15 in immunity [68]. For other behaviors, including alcohol dependency, GWAS in which African American patients were included have not yielded SNPs with genome-wide significance [69, 70]. These studies have found areas in the genome that are correlated with dependence behaviors. For example, a region on chromosome 11 is described as a candidate locus for alcohol dependence [69]. For dependency behaviors, replication of previous associations found in African Americans has been difficult.

GWAS exploring cardiovascular diseases have revealed several association signals (online suppl. table 1). GWAS of African American hypertensive patients found associations with genes involved in systolic blood pressure including an $Na/K^+/Ca$ exchanger *(SLC24A4)*, and a class of calcium channel blockers *(CACNA1H)*, *PMS1, YWHAZ,* and *IPO7* [71]. With the largest sample size using multiple cohorts of American Americans, Fox et al.

[72] were not able to replicate initial genetic associations for systolic or diastolic blood pressure. They suggest that within this large multiple cohort study, several issues (e.g. the heterogeneity in blood pressure measurement across centers, the number of individuals on blood pressure-lowering medication, antihypertensive medication, and the median age of cohorts) may have biased the results toward the null hypothesis [72].

GWAS on cardiac left ventricular hypertrophy and stroke have showed associations with variants in *RAI14*, and *CD36* for left ventricular hypertrophy and *NINJ2* and *CD36* for stroke, showing that CD36 is implicated in both conditions [73, 74]. CD36 is a receptor with diverse functions including negative regulation of angiogenesis, internalizing pathogens in immune response, and facilitating lipid transport and muscle lipid utilization [75]. The lipid utilization function of CD36 is also implicated in signaling for systolic blood pressure for hypertension and adipose energy storage in obesity and diabetes [71, 75]. CD36-mediated signaling could provide an example of common signaling paradigms in which other effectors could interact between pathways of hypertension, diabetes, stroke, and cardiovascular disease. Further characterization in non-European populations of CD36 function in these different conditions is warranted.

In addition to identifying genes associated with non-communicable disorders, GWAS can also be used to identify genotypes associated with various treatment outcomes. In hypertension, variants in *LYZ, YEATS4,* and fibroblast growth factor receptor substrate 2 *(FRS2)* are associated with differential responses to antihypertensive drugs (e.g. thiazide diuretics) commonly used to regulate blood pressure [17]. In African Americans, the variant located in YEATS4 was significantly associated with good response to thiazide diuretics [17]. Thiazide diuretics can be used to distinguish good responders (patients that show improvement on diuretics) from non-responders or bad responders (patients with a negative response to diuretic medications for the treatment of hypertension).

Communicable Diseases

One of the major infectious diseases affecting sub-Saharan African populations is malaria. Studies have been conducted to characterize the regions of the *Plasmodium falciparum* genome under positive selection in order to monitor the development of drug resistance in the parasite [18]. Additionally, GWAS have utilized cases in endemic areas to determine loci in the human host that influence disease outcomes [76]. The rationale is that host genetics accounts for ~25% of the risk for life-threatening malaria, or more specifically, the anticipated effects of hemoglobinopathies (e.g. α-thalassemia) would only explain ~2.5% of the total variation. Thus, it is hypothesized that protective genes of hemoglobinopathies result in small population effects suggesting that malaria resistance is under multigenic control with each individual gene having a relatively small impact [77]. Jallow et al. [76] identified genetic variants close to *SOC1, HBB* and *DDC* as protective against severe malaria in children (online suppl. table 3). SOC1 is a scaffolding protein with multiple functional motifs mediating interactions with other signaling molecules [78]. Similar to other Gab proteins, SOC1 is involved in signaling pathways mediated by receptor and non-receptor protein tyrosine kinases and is important in immunity [79]. Dihydroxyphenylalanine decarboxylase converts specific chemical precursors to dopamine and serotonin via decarboxylation which also functions in cellular immunity and resistance to parasitic infections via melanization [80]. The hemoglobin β locus *(HBB)* has been replicated in several studies as a major locus for malaria resistance and is reviewed in other papers [81, 82]. The locus has been under positive selection in humans because of its role in malaria resistance [83]. These data suggest that resistance to malaria infection also includes immune response mechanisms, which need further characterization.

GWAS have also been performed to identify genetic variants associated with pulmonary tuberculosis in West Africans [84]. The nearest genes to the variants identified were *GATA6, CTAGE1, RBBP8*, and *CABLES1* along with several unannotated open reading frames [84]. GATA6 is a transcription factor identified as regulating the local epithelial innate immune response, with an essential function maintaining the lung epithelium [85, 86]. In addition, other reports indicated that *GATA6* regulates the temporal appearance and number of bronchioalveolar stem cells, making *GATA6* essential for proper lung epithelial regeneration [87]. Human *GATA6* expression has a protective function in lung epithelial cells exposed to *Pseudomonas aeruginosa* [85]. Other genes involved in immunity have also been identified in diseases caused by viruses. For example, *HLA-B* located at chromosome 6p21 has been found to influence the HIV viral load in both African Americans and Europeans [88–90]. Pelak et al. [89] found that *HLA-B*5703* influences the viral load variation in African Americans. *HLA-B*5703* has not been found in European populations studied and has an allele frequency of ~5.8% in the Yoruba, indicating that its presence in African Americans is likely due to African ancestry [91]. The closely related allele *HLA-B*5701* is

also important in HIV-1 control; it has an allele frequency ~6.1% in Europeans, but has not been found in the Yoruba [91, 89]. Other determinants of HIV infection and viral load could lie within the HLA locus which has also been linked to HIV viral load [88]. In general, these data suggest that the mechanistic control of HIV in different populations (i.e. African Americans and Europeans) is similar [90].

One major influence on treatment outcomes for hepatitis C viral infections is a variant upstream of *IFNL3*, encoding interferon-λ-3 [6]. This variant is associated with sustained absence of detectable virus or a sustained virological response (SVR). Using GWAS, Ge et al. [6] observed that almost half of the difference in SVR between populations could be accounted for by the difference in the frequency of the C allele between African Americans and Europeans. In viral infections, the mechanism of clearance seems to suggest that the variants are similar across populations, but it is the frequency of the variants that lead to differences in responses in diverse populations [6]. Additional research is needed to determine common variants that influence viral response.

Lessons from Previous GWAS

Some lessons have emerged from GWAS that can inform other genomics-related studies. The unique lessons acquired from studies of non-European populations, such as African Americans, Caribbeans, Hispanic Americans, and Africans are largely due to their increased genetic diversity. First, GWAS in one African population or diasporic African population may not always be replicated in another African or diasporic African population because of population diversity and allele frequency differences. Thorough understanding of population structure between and within these populations is needed when performing GWAS analysis. Currently, reference sequences for the majority of African or diverse populations are not available (with the exception of the limited number of African populations in the HapMap and 1000 Genome Project databases). To circumvent this major problem, DNA sequencing is required to characterize population genetic diversity which can be used to generate very robust results [76]. This method overcomes issues of incorrect imputation due to the lack of inappropriate reference populations. Second, the use of genetically matched controls from reference populations is also critical in GWAS. The introduction of non-genetically matched controls can decrease the statistical power to detect associations in GWAS or may result in spurious association because of population structure. Third, careful ascertainment of the phenotype is crucial to prevent residual confounding effects of unmeasured phenotypes [9]. Fourth, generally weaker linkage disequilibrium pattern in African populations is both an advantage and a disadvantage. Weaker linkage disequilibrium allows for improved localization of associations. On the other hand, coverage is lower, and tag SNPs are less efficient in African or African ancestry populations, making the initial discovery of association signals using the same commercial platforms used in European population GWAS less efficient in these populations. Moreover, improved localization and refinement of association signals have been demonstrated in African Americans and African populations leading to successful discoveries in infectious disease susceptibility [76, 92]. Some initial studies using European populations have failed to replicate findings in African ancestry populations suggesting some phenotypes may show genetic heterogeneity between populations [14]. Finally, many GWAS that involved African ancestry populations do not have adequate sample sizes. It is recommended that investigators should use appropriate sample sizes when performing GWAS in non-European populations; essentially, the samples sizes should be larger when compared to European populations to account for additional genetic diversity within the African ancestry populations.

Limitations of current GWAS are related to both the tools and methods. Currently, many GWAS in African populations are performed using arrays designed for European populations, directly or indirectly capturing ~60% of the genomic variation in any African population, leaving 40% of these variants unassayed [93, 94]. These limitations also result in the lack of discovery of variants that could significantly contribute to the phenotype. Several non-profit entities have developed partnerships with for-profit corporations to enhance variant coverage and power of genotyping arrays for African ancestry populations. Contemporary efforts to develop an African Diaspora Power SNP Chip that aims to contribute to a deeper understanding of genetic variation particularly in admixed African ancestry populations in the Americas is underway. As part of this effort, whole genome sequencing of ~700 African ancestry samples via the Consortium on Asthma among African-ancestry Populations in the Americas has identified about 22 million novel variants, and the design of the array that includes common tag SNPs is anticipated to be valuable, particularly for admixed African ancestry populations in the Americas

(http://www.cidr.jhmi.edu/supported/20140519_AfricanDiasporaPowerChip2.pdf). However, given the vast genetic diversity of continental African populations that are not adequately represented by admixed Africans in the diaspora [19], sequencing of more divergent and diverse African population groups will uncover novel genetic variation and will be valuable for developing widely representative sequence reference panels and genotyping arrays.

Finally, methods that capture gene-environment interactions are not well developed, despite the acknowledged complexity of disease. Statistical genetics methods only capture genetic variation, but gene-environment interaction in the African and African ancestry context may be more important due to the different environments in which African populations have evolved and live. As GWAS move forward, methods will be needed to detect gene and environment interactions in disease associations that could be significantly different within and among different African and African ancestry populations.

The Future of Genomic Studies in African and African Ancestry Populations

Recently, initiatives such as the Human Heredity and Health in Africa (H3Africa) Initiative jointly funded by the US National Institutes of Health and the Wellcome Trust in the UK have awarded genomic research funding to Africa-based researchers with an aim to build genomic research capacity locally [95]. H3Africa along with other ongoing projects, such as the African Genome Variation Project (http://www.sanger.ac.uk/research/initiatives/globalhealth/research/africangenome.html), will help to elucidate the genetic landscape of population groups across sub-Saharan Africa and advance future genomic research endeavors in Africa.

Inclusion of African ancestry populations offers substantial advantages to investigators, but many challenges also exist in study design, implementation and data interpretation [95]. While problems in study design can often be localized and efficiently addressed, data interpretation in these populations could be difficult because of several challenges which include: (1) ascertainment of markers, (2) decreased levels of tag-SNP portability and (3) increased genotype imputation error [12]. Available software can be utilized to address many issues that arise from analyzing GWAS data in diverse populations [96]. Several methods have been designed to address unique issues in assessing population stratification of populations with admixed ancestries [97–99]. These new methods improve the ability to correctly account for several confounding factors which, if not properly addressed, could lead to serious errors in modeling data.

The art of conducting genomic studies in populations identified by ancestry is as important and daunting as the science itself. The unique and shared historical, social, cultural, and economic contexts within and between populations must be taken into account throughout the conceptualization and implementation of these studies. A growing body of research is exposing and exploring the range of ethical and policy issues pertaining to informed consent, data and sample sharing, management of results and secondary findings, and other aspects of genomic research in African and African ancestry populations [96, 100–107]. In addition, ongoing community and public engagement in these and other populations involved in or affected by genomic research is increasingly highlighted as a means of building relations, facilitating understanding through exchange of information, and enhancing collaboration between researchers and communities [108–112]. As cutting edge genomic research gradually proliferates throughout Africa and the African diaspora, opportunities and challenges related to the identification and resolution of emerging issues will multiply. It is essential that our current training efforts adequately foster the development of researchers in these locations who can rigorously examine these issues.

Conclusions

GWAS have uncovered genetic variants associated with susceptibility for several different traits and disease conditions including asthma, cancer, obesity, hypertension, blood pressure, and malaria in recent African ancestry populations. GWAS involving non-Europeans such as African American and African ancestry populations have greatly lagged behind the increasing number of GWAS that have been published. Some existing GWAS that use non-European populations at replication omit these populations in the initial discovery phase, resulting in subsequent failure to replicate the associated variant(s). Failure to replicate genetic variants associated with the phenotype of interest in these populations could be erroneously attributed to differences in population structure, allele frequencies and environmental factors suggesting the need to include larger cohorts of non-European populations in both discovery and replication.

Whereas non-communicable diseases impose huge burdens on US populations, infectious diseases have traditionally been more relevant and problematic on the African continent. However, within Africa, the increase in a more sedentary lifestyle, in conjunction with high-caloric diet is contributing to increases in the prevalence of several categories of non-communicable diseases. Both the relevance of infectious diseases and increasing prevalence of disease due to lifestyle offer unparalleled advantages of African populations in understanding the genetic and environmental underpinnings of diseases in human populations. Another unique advantage of African populations in discerning the evolutionary history of disease variants and haplotypes has not been well exploited and may provide directions in understanding pathways and mechanisms of human diseases [12–14]. Finally, the potential role of genomics research with more diverse non-European populations in understanding health disparities should be emphasized. Several projects are underway to characterize human genetic and phenotypic variation (http://1000genomes.org and http://news.sciencemag.org/scienceinsider/2010/06/uk-to-sequence-10000-genomes.html). However, without the inclusion of multiple populations of African ancestry, the data generated by these multinational collaborative efforts might have a limited benefit to humanity as a whole. Advances in genomics should help ameliorate rather than contribute to health disparities among diverse populations. With the collaborative efforts of global scientists and local communities, and the development of new data analysis capacity, state-of-the-art genomics research is poised to take advantage of humanity's most diverse populations and return the greatest benefit in the improvement of human health.

Acknowledgement

The authors kindly acknowledge and also thank Dr. Daniel Shriner at the Center for Research on Genomics and Global Health (CRGGH), and the NIH Fellows Editorial Board for editorial advice and careful reading of the manuscript. This research was supported in part by an NIA training grant to the University of Maryland (T32-AG000219).

References

1 Eriksson N, Macpherson JM, Tung JY, Hon LS, Naughton B, Saxonov S, Avey L, Wojcicki A, Pe'er I, Mountain J: Web-based, participant-driven studies yield novel genetic associations for common traits. PLoS Genet 2010; 6:e1000993.

2 Sulem P, Gudbjartsson DF, Stacey SN, Helgason A, Rafnar T, Jakobsdottir M, Steinberg S, Gudjonsson SA, Palsson A, Thorleifsson G, Palsson S, Sigurgeirsson B, Thorisdottir K, Ragnarsson R, Benediktsdottir KR, Aben KK, Vermeulen SH, Goldstein AM, Tucker MA, Kiemeney LA, Olafsson JH, Gulcher J, Kong A, Thorsteinsdottir U, Stefansson K: Two newly identified genetic determinants of pigmentation in Europeans. Nat Genet 2008;40: 835–837.

3 Kang SJ, Chiang CW, Palmer CD, Tayo BO, Lettre G, Butler JL, Hackett R, Adeyemo AA, Guiducci C, Berzins I, Nguyen TT, Feng T, Luke A, Shriner D, Ardlie K, Rotimi C, Wilks R, Forrester T, McKenzie CA, Lyon HN, Cooper RS, Zhu X, Hirschhorn JN: Genome-wide association of anthropometric traits in African- and African-derived populations. Hum Mol Genet 2010;19:2725–2738.

4 Steidl C, Telenius A, Shah SP, Farinha P, Barclay L, Boyle M, Connors JM, Horsman DE, Gascoyne RD: Genome-wide copy number analysis of Hodgkin Reed-Sternberg cells identifies recurrent imbalances with correlations to treatment outcome. Blood 2010;116: 418–427.

5 Shi JY, Ren ZH, Jiao B, Xiao R, Yun HY, Chen B, Zhao WL, Zhu Q, Chen Z, Chen SJ: Genetic variations of DNA repair genes and their prognostic significance in patients with acute myeloid leukemia. Int J Cancer 2011;128: 233–238.

6 Ge D, Fellay J, Thompson AJ, Simon JS, Shianna KV, Urban TJ, Heinzen EL, Qiu P, Bertelsen AH, Muir AJ, Sulkowski M, McHutchison JG, Goldstein DB: Genetic variation in *IL28B* predicts hepatitis c treatment-induced viral clearance. Nature 2009;461:399–401.

7 Fellay J, Marzolini C, Meaden ER, Back DJ, Buclin T, Chave JP, Decosterd LA, Furrer H, Opravil M, Pantaleo G, Retelska D, Ruiz L, Schinkel AH, Vernazza P, Eap CB, Telenti A: Response to antiretroviral treatment in HIV-1-infected individuals with allelic variants of the multidrug resistance transporter 1: a pharmacogenetics study. Lancet 2002;359: 30–36.

8 Mathew CG: New links to the pathogenesis of Crohn disease provided by genome-wide association scans. Nat Rev Genet 2008;9:9–14.

9 Simón-Sánchez J, Singleton A: Genome-wide association studies in neurological disorders. Lancet Neurol 2008;7:1067–1072.

10 Innocenti F, Cox NI, Dolan ME: The use of genomic information to optimize cancer chemotherapy. Semin Oncol 2011;38:186–195.

11 Feero WG, Guttmacher AE, Collins FS: Genomic medicine – an updated primer. N Engl J Med 2010;362:2001–2011.

12 Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, Boehnke M: Genome-wide association studies in diverse populations. Nat Rev Genet 2010;11:356–366.

13 Need AC, Goldstein DB: Next generation disparities in human genomics: concerns and remedies. Trends Genet 2009;25:489–494.

14 Haga SB: Impact of limited population diversity of genome-wide association studies. Genet Med 2010;12:81–84.

15 Ramos E, Rotimi C: The A's, G's, C's, and T's of health disparities. BMC Med Genomics 2009;2:29.

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

17 Turner ST, Bailey KR, Fridley BL, Chapman AB, Schwartz GL, Chai HS, Sicotte H, Kocher JP, Rodin AS, Boerwinkle E: Genomic association analysis suggests chromosome 12 locus influencing antihypertensive response to thiazide diuretic. Hypertension 2008;52:359–365.

18 Mu J, Myers RA, Jiang H, Liu S, Ricklefs S, Waisberg M, Chotivanich K, Wilairatana P, Krudsood S, White NJ, Udomsangpetch R, Cui L, Ho M, Ou F, Li H, Song J, Li G, Wang X, Seila S, Sokunthea S, Socheat D, Sturdevant DE, Porcella SF, Fairhurst RM, Wellems TE, Awadalla P, Su XZ: *Plasmodium falciparum* genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. Nat Genet 2010;42:268–271.

19 Campbell MC, Tishkoff SA: African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. Annu Rev Genomics Hum Genet 2008;9:403–433.
20 Tishkoff SA, Williams SM: Genetic analysis of African populations: human evolution and complex disease. Nat Rev Genet 2002;3:611–621.
21 International Hap Map Consortium: The International HapMap Project. Nature 2003; 426:789–796.
22 Peprah EK, Allen EG, Williams SM, Woodard LM, Sherman SL: Genetic diversity of the fragile X syndrome gene *(FMR1)* in a large sub-Saharan West African population. Ann Hum Genet 2010;74:316–325.
23 Crawford DC, Acuña JM, Sherman SL: *FMR1* and the fragile X syndrome: Human genome epidemiology review. Genet Med 2001;3:359–371.
24 Goldman A, Krause A, Jenkins T: Fragile X syndrome occurs in the South African black population. S Afr Med J 1997;87:418–420.
25 Redd SC: Asthma in the United States: burden and current theories. Environ Health Perspect 2002;110(suppl 4):557–560.
26 Lenoir M, McGill CL, Graham LM Jr, Foggs M, Robinson S, Crim C, Stempel DA: A new focus on assessing and treating asthma control in the African-American community: a call to action. J Natl Med Assoc 2008;100(9 suppl):1–23.
27 Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, Wang K, Rafaels NM, Michel S, Bonnelykke K, Zhang H, Kim CE, Frackelton EC, Glessner JT, Hou C, Otieno FG, Santa E, Thomas K, Smith RM, Glaberson WR, Garris M, Chiavacci RM, Beaty TH, Ruczinski I, Orange JM, Allen J, Spergel JM, Grundmeier R, Mathias RA, Christie JD, von Mutius E, Cookson WO, Kabesch M, Moffatt MF, Grunstein MM, Barnes KC, Devoto M, Magnusson M, Li H, Grant SF, Bisgaard H, Hakonarson H: Variants of *DENND1B* associated with asthma in children. N Engl J Med 2010;362:36–44.
28 Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, Wilk JB, Willis-Owen SAG, Klanderman B, Lasky-Su J, Lazarus R, Murphy AJ, Soto-Quiros ME, Avila L, Beaty T, Mathias RA, Ruczinski I, Barnes KC, Celedon JC, Cookson WOC, Gauderman WJ, Gilliland FD, Hakonarson H, Lange C, Moffatt MF, O'Connor GT, Raby BA, Silverman EK, Weiss ST: Genome-wide association analysis identifies *PDE4D* as an asthma-susceptibility gene. Am J Hum Genet 2009;84:581–593.
29 Mathias RA, Grant AV, Rafaels N, Hand T, Gao L, Vergara C, Tsai YJ, Yang M, Campbell M, Foster C, Gao P, Togias A, Hansel NN, Diette G, Adkinson NF, Liu MC, Faruque M, Dunston GM, Watson HR, Bracken MB, Hoh J, Maul P, Maul T, Jedlicka AE, Murray T, Hetmanski JB, Ashworth R, Ongaco CM, Hetrick KN, Doheny KF, Pugh EW, Rotimi CN, Ford J, Eng C, Burchard EG, Sleiman PM, Hakonarson H, Forno E, Raby BA, Weiss ST, Scott AF, Kabesch M, Liang L, Abecasis G, Moffatt MF, Cookson WO, Ruczinski I, Beaty TH, Barnes KC: A genome-wide association study on African-ancestry populations for asthma. J Allergy Clin Immunol 2010;125: 336–346.
30 Weiss ST, Raby BA, Rogers A: Asthma genetics and genomics 2009. Curr Opin Genet Dev 2009;19:279–282.
31 Nakagome K, Nagata M: Pathogenesis of airway inflammation in bronchial asthma. Auris Nasus Larynx 2011;38:555–563.
32 Bernstein DI: ABCs of asthma. Clin Cornerstone 2008;8:9–25.
33 Crosslin DR, McDavid A, Weston N, Nelson SC, Zheng X, Hart E, de Andrade M, Kullo IJ, McCarty CA, Doheny KF, Pugh E, Kho A, Hayes MG, Pretel S, Saip A, Ritchie MD, Crawford DC, Crane PK, Newton K, Li R, Mirel DB, Crenshaw A, Larson EB, Carlson CS, Jarvik GP; Electronic Medical Records and Genomics (eMERGE) Network: Genetic variants associated with the white blood cell count in 13,923 subjects in the eMERGE network. Hum Genet 2012;131:639–652.
34 Reiner AP, Lettre G, Nalls MA, Ganesh SK, Mathias R, Austin MA, Dean E, Arepalli S, Britton A, Chen Z, Couper D, Curb JD, Eaton CB, Fornage M, Grant SF, Harris TB, Hernandez D, Kamatini N, Keating BJ, Kubo M, LaCroix A, Lange LA, Liu S, Lohman K, Meng Y, Mohler ER, 3rd, Musani S, Nakamura Y, O'Donnell CJ, Okada Y, Palmer CD, Papanicolaou GJ, Patel KV, Singleton AB, Takahashi A, Tang H, Taylor HA Jr, Taylor K, Thomson C, Yanek LR, Yang L, Ziv E, Zonderman AB, Folsom AR, Evans MK, Liu Y, Becker DM, Snively BM, Wilson JG: Genome-wide association study of white blood cell count in 16,388 African Americans: the continental origins and genetic epidemiology network (COGENT). PLoS Genet 2011; 7:e1002108.
35 Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, von Mutius E, Farrall M, Lathrop M, Cookson WO; GABRIEL Consortium: A large-scale, consortium-based genomewide association study of asthma. N Engl J Med 2010;363:1211–1221.
36 Bhatnagar P, Purvis S, Barron-Casella E, DeBaun MR, Casella JF, Arking DE, Keefer JR: Genome-wide association study identifies genetic variants influencing F-cell levels in sickle-cell patients. J Hum Genet 2011;56:316–323.
37 Field JJ, DeBaun MR: Asthma and sickle cell disease: two distinct diseases or part of the same process? Hematology Am Soc Hematol Educ Program 2009:45–53.
38 Morris CR: Asthma management: reinventing the wheel in sickle cell disease. Am J Hematol 2009;84:234–241.
39 Caboot JB, Allen JL: Pulmonary complications of sickle cell disease in children. Curr Opin Pediatr 2008;20:279–287.
40 Sebastiani P, Solovieff N, Hartley SW, Milton JN, Riva A, Dworkis DA, Melista E, Klings ES, Garrett ME, Telen MJ, Ashley-Koch A, Baldwin CT, Steinberg MH: Genetic modifiers of the severity of sickle cell anemia identified through a genome-wide association study. Am J Hematol 2010;85:29–35.
41 Liu W, Saint DA: Heterogeneous expression of tandem-pore K^+ channel genes in adult and embryonic rat heart quantified by real-time polymerase chain reaction. Clin Exp Pharmacol Physiol 2004;31:174–178.
42 Davis KA, Cowley EA: Two-pore-domain potassium channels support anion secretion from human airway Calu-3 epithelial cells. Pflugers Arch 2006;451:631–641.
43 Eeles RA, Kote-Jarai Z, Al Olama AA, Giles GG, Guy M, Severi G, Muir K, Hopper JL, Henderson BE, Haiman CA, Schleutker J, Hamdy FC, Neal DE, Donovan JL, Stanford JL, Ostrander EA, Ingles SA, John EM, Thibodeau SN, Schaid D, Park JY, Spurdle A, Clements J, Dickinson JL, Maier C, Vogel W, Dork T, Rebbeck TR, Cooney KA, Cannon-Albright L, Chappuis PO, Hutter P, Zeegers M, Kaneva R, Zhang H-W, Lu Y-J, Foulkes WD, English DR, Leongamornlert DA, Tymrakiewicz M, Morrison J, Ardern-Jones AT, Hall AL, O'Brien LT, Wilkinson RA, Saunders EJ, Page EC, Sawyer EJ, Edwards SM, Dearnaley DP, Horwich A, Huddart RA, Khoo VS, Parker CC, Van As N, Woodhouse CJ, Thompson A, Christmas T, Ogden C, Cooper CS, Southey MC, Lophatananon A, Liu J-F, Kolonel LN, Le Marchand L, Wahlfors T, Tammela TL, Auvinen A, Lewis SJ, Cox A, FitzGerald LM, Koopmeiners JS, Karyadi DM, Kwon EM, Stern MC, Corral R, Joshi AD, Shahabi A, McDonnell SK, Sellers TA, Pow-Sang J, Chambers S, Aitken J, Gardiner RA, Batra J, Kedda MA, Lose F, Polanowski A, Patterson B, Serth J, Meyer A, Luedeke M, Stefflova K, Ray AM, Lange EM, Farnham J, Khan H, Slavov C, Mitkova A, Cao G, Easton DF: Identification of seven new prostate cancer susceptibility loci through a genome-wide association study. Nat Genet 2009;41:1116–1121.
44 Hooker S, Hernandez W, Chen H, Robbins C, Torres JB, Ahaghotu C, Carpten J, Kittles RA: Replication of prostate cancer risk loci on 8q24, 11q13, 17q12, 19q33, and xp11 in African Americans. Prostate 2010;70:270–275.

45 Cariati M, Naderi A, Brown JP, Smalley MJ, Pinder SE, Caldas C, Purushotham AD: Alpha-6 integrin is necessary for the tumourigenicity of a stem cell-like subpopulation within the MCF7 breast cancer cell line. Int J Cancer 2008;122:298–304.
46 Manavi M, Hudelist G, Fink-Retter A, Gschwandtler-Kaulich D, Pischinger K, Czerwenka K: Gene profiling in Pap-cell smears of high-risk human papillomavirus-positive squamous cervical carcinoma. Gynecol Oncol 2007;105:418–426.
47 Abdel-Wahab O, Mullally A, Hedvat C, Garcia-Manero G, Patel J, Wadleigh M, Malinge S, Yao J, Kilpivaara O, Bhat R, Huberman K, Thomas S, Dolgalev I, Heguy A, Paietta E, Le Beau MM, Beran M, Tallman MS, Ebert BL, Kantarjian HM, Stone RM, Gilliland DG, Crispino JD, Levine RL: Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. Blood 2009;114:144–147.
48 Delhommeau F, Dupont S, Valle VD, James C, Trannoy S, Massé A, Kosmider O, Le Couedic JP, Robert F, Alberdi A, Lécluse Y, Plo I, Dreyfus FJ, Marzac C, Casadevall N, Lacombe C, Romana SP, Dessen P, Soulier J, Viguié F, Fontenay M, Vainchenker W, Bernard OA: Mutation in *TET2* in myeloid cancers. N Engl J Med 2009;360:2289–2301.
49 Flach J, Dicker F, Schnittger S, Kohlmann A, Haferlach T, Haferlach C: Mutations of *JAK2* and *TET2*, but not *CBL* are detectable in a high portion of patients with refractory anemia with ring sideroblasts and thrombocytosis. Haematologica 2010;95:518–519.
50 Song H, Zhang B, Watson MA, Humphrey PA, Lim H, Milbrandt J: Loss of *Nkx3.1* leads to the activation of discrete downstream target genes during prostate tumorigenesis. Oncogene 2009;28:3307–3319.
51 Kim MJ, Bhatia-Gaur R, Banach-Petrosky WA, Desai N, Wang Y, Hayward SW, Cunha GR, Cardiff RD, Shen MM, Abate-Shen C: *Nkx3.1* mutant mice recapitulate early stages of prostate carcinogenesis. Cancer Res 2002;62:2999–3004.
52 Abate-Shen C, Shen MM, Gelmann E: Integrating differentiation and cancer: the *Nkx3.1* homeobox gene in prostate organogenesis and carcinogenesis. Differentiation 2008;76:717–727.
53 Koller DL, Ichikawa S, Lai D, Padgett LR, Doheny KF, Pugh E, Paschall J, Hui SL, Edenberg HJ, Xuei X, Peacock M, Econs MJ, Foroud T: Genome-wide association study of bone mineral density in premenopausal European-American women and replication in African-American women. J Clin Endocrinol Metab 2010;95:1802–1809.
54 Ajiboye S, Sissung TM, Sharifi N, Figg WD: More than an accessory: implications of type III transforming growth factor-β receptor loss in prostate cancer. BJU Int 2010;105:913–916.
55 Tekola Ayele F, Adeyemo A, Finan C, Hailu E, Sinnott P, Burlinson ND, Aseffa A, Rotimi CN, Newport MJ, Davey G: HLA class II locus and susceptibility to podoconiosis. N Engl J Med 2012;366:1200–1208.
56 Ebstein RP, Israel S, Chew SH, Zhong S, Knafo A: Genetics of human social behavior. Neuron 2010;65:831–844.
57 Holden C: Parsing the genetics of behavior. Science 2008;322:892–895.
58 Tabarés-Seisdedos R, Rubenstein JL: Chromosome 8p as a potential hub for developmental neuropsychiatric disorders: implications for schizophrenia, autism and cancer. Mol Psychiatry 2009;14:563–589.
59 Gejman PV, Sanders AR, Duan J: The role of genetics in the etiology of schizophrenia. Psychiatr Clin North Am 2010;33:35–66.
60 Gershon ES, Alliey-Rodriguez N, Liu C: After GWAS: searching for genetic risk for schizophrenia and bipolar disorder. Am J Psychiatry 2011;168:253–256.
61 Lin PI, Shuldiner AR: Rethinking the genetic basis for comorbidity of schizophrenia and type 2 diabetes. Schizophr Res 2010;123:234–243.
62 Shi J, Levinson DF, Duan J, Sanders AR, Zheng Y, Pe'er I, Dudbridge F, Holmans PA, Whittemore AS, Mowry BJ, Olincy A, Amin F, Cloninger CR, Silverman JM, Buccola NG, Byerley WF, Black DW, Crowe RR, Oksenberg JR, Mirel DB, Kendler KS, Freedman R, Gejman PV: Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature 2009;460:753–757.
63 Byerley W, Badner JA: Strategies to identify genes for complex disorders: a focus on bipolar disorder and chromosome 16p. Psychiatr Genet 2011;21:173–182.
64 Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, Byerley W, Coryell W, Craig D, Edenberg HJ, Eskin E, Foroud T, Gershon E, Greenwood TA, Hipolito M, Koller DL, Lawson WB, Liu C, Lohoff F, McInnis MG, McMahon FJ, Mirel DB, Murray SS, Nievergelt C, Nurnberger J, Nwulia EA, Paschall J, Potash JB, Rice J, Schulze TG, Scheftner W, Panganiban C, Zaitlen N, Zandi PP, Zollner S, Schork NJ, Kelsoe JR: Genome-wide association study of bipolar disorder in European American and African American individuals. Mol Psychiatry 2009;14:755–763.
65 Liu YZ, Pei YF, Guo YF, Wang L, Liu XG, Yan H, Xiong DH, Zhang YP, Levy S, Li J, Haddock CK, Papasian CJ, Xu Q, Ma JZ, Payne TJ, Recker RR, Li MD, Deng HW: Genome-wide association analyses suggested a novel mechanism for smoking behavior regulated by IL15. Mol Psychiatry 2009;14:668–680.
66 Mian MF, Pek EA, Mossman KL, Stämpfli MR, Ashkar AA: Exposure to cigarette smoke suppresses IL-15 generation and its regulatory NK cell functions in poly I:C-augmented human PBMCs. Mol Immunol 2009;46:3108–3116.
67 van Leeuwen EM, Sprent J, Surh CD: Generation and maintenance of memory $CD4^+$ T cells. Curr Opin Immunol 2009;21:167–172.
68 Rodrigues L, Bonorino C: Role of IL-15 and IL-21 in viral immunity: applications for vaccines and therapies. Expert Rev Vaccines 2009;8:167–177.
69 Edenberg HJ, Koller DL, Xuei X, Wetherill L, McClintick JN, Almasy L, Bierut LJ, Bucholz KK, Goate A, Aliev F, Dick D, Hesselbrock V, Hinrichs A, Kramer J, Kuperman S, Nurnberger JI Jr, Rice JP, Schuckit MA, Taylor R, Todd Webb B, Tischfield JA, Porjesz B, Foroud T: Genome-wide association study of alcohol dependence implicates a region on chromosome 11. Alcohol Clin Exp Res 2010;34:840–852.
70 Bierut LJ, Agrawal A, Bucholz KK, Doheny KF, Laurie C, Pugh E, Fisher S, Fox L, Howells W, Bertelsen S, Hinrichs AL, Almasy L, Breslau N, Culverhouse RC, Dick DM, Edenberg HJ, Foroud T, Grucza RA, Hatsukami D, Hesselbrock V, Johnson EO, Kramer J, Krueger RF, Kuperman S, Lynskey M, Mann K, Neuman RJ, Nothen MM, Nurnberger JI Jr, Porjesz B, Ridinger M, Saccone NL, Saccone SF, Schuckit MA, Tischfield JA, Wang JC, Rietschel M, Goate AM, Rice JP; Gene, Environment Association Studies Consortium: A genome-wide association study of alcohol dependence. Proc Natl Acad Sci USA 2010;107:5082–5087.
71 Adeyemo A, Gerry N, Chen G, Herbert A, Doumatey A, Huang H, Zhou J, Lashley K, Chen Y, Christman M, Rotimi C: A genome-wide association study of hypertension and blood pressure in African Americans. PLoS Genet 2009;5:e1000564.
72 Fox ER, Young JH, Li Y, Dreisbach AW, Keating BJ, Musani SK, Liu K, Morrison AC, Ganesh S, Kutlar A, Ramachandran VS, Polak JF, Fabsitz RR, Dries DL, Farlow DN, Redline S, Adeyemo A, Hirschorn JN, Sun YV, Wyatt SB, Penman AD, Palmas W, Rotter JI, Townsend RR, Doumatey AP, Tayo BO, Mosley TH, Lyon HN, Kang SJ, Rotimi CN, Cooper RS, Franceschini N, Curb JD, Martin LW, Eaton CB, Kardia SLR, Taylor HA, Caulfield MJ, Ehret GB, Johnson T; International Consortium for Blood Pressure Genome-wide Association Studies (ICBP-GWAS), Chakravarti A, Zhu X, Levy D: Association of genetic variation with systolic and diastolic blood pressure among African Americans: the Candidate Gene Association Resource study. Hum Mol Genet 2011;20:2273–2284.
73 Arnett DK, Li N, Tang W, Rao DC, Devereux RB, Claas SA, Kraemer R, Broeckel U: Genome-wide association study identifies single-nucleotide polymorphism in *KCNB1* associated with left ventricular mass in humans: the HyperGEN study. BMC Med Genet 2009;10:43.

74 Ikram MA, Seshadri S, Bis JC, Fornage M, DeStefano AL, Aulchenko YS, Debette S, Lumley T, Folsom AR, van den Herik EG, Bos MJ, Beiser A, Cushman M, Launer LJ, Shahar E, Struchalin M, Du Y, Glazer NL, Rosamond WD, Rivadeneira F, Kelly-Hayes M, Lopez OL, Coresh J, Hofman A, DeCarli C, Heckbert SR, Koudstaal PJ, Yang Q, Smith NL, Kase CS, Rice K, Haritunians T, Roks G, de Kort PL, Taylor KD, de Lau LM, Oostra BA, Uitterlinden AG, Rotter JI, Boerwinkle E, Psaty BM, Mosley TH, van Duijn CM, Breteler MM, Longstreth WT Jr, Wolf PA: Genomewide association studies of stroke. N Engl J Med 2009;360:1718–1728.
75 Silverstein RL, Febbraio M: CD36, a scavenger receptor involved in immunity, metabolism, angiogenesis, and behavior. Sci Signal 2009;2:re3.
76 Jallow M, Teo YY, Small KS, Rockett KA, Deloukas P, Clark TG, Kivinen K, Bojang KA, Conway DJ, Pinder M, Sirugo G, Sisay-Joof F, Usen S, Auburn S, Bumpstead SJ, Campino S, Coffey A, Dunham A, Fry AE, Green A, Gwilliam R, Hunt SE, Inouye M, Jeffreys AE, Mendy A, Palotie A, Potter S, Ragoussis J, Rogers J, Rowlands K, Somaskantharajah E, Whittaker P, Widden C, Donnelly P, Howie B, Marchini J, Morris A, Sanjoaquin M, Achidi EA, Agbenyega T, Allen A, Amodu O, Corran P, Djimde A, Dolo A, Doumbo OK, Drakeley C, Dunstan S, Evans J, Farrar J, Fernando D, Hien TT, Horstmann RD, Ibrahim M, Karunaweera N, Kokwaro G, Koram KA, Lemnge M, Makani J, Marsh K, Michon P, Modiano D, Molyneux ME, Mueller I, Parker M, Peshu N, Plowe CV, Puijalon O, Reeder J, Reyburn H, Riley EM, Sakuntabhai A, Singhasivanon P, Sirima S, Tall A, Taylor TE, Thera M, Troye-Blomberg M, Williams TN, Wilson M, Kwiatkowski DP; Wellcome Trust Case Control Consortium; Malaria Genomic Epidemiology Network: Genome-wide and fine-resolution association analysis of malaria in West Africa. Nat Genet 2009;41:657–665.
77 Mackinnon MJ, Mwangi TW, Snow RW, Marsh K, Williams TN: Heritability of malaria in Africa. PLoS Med 2005;2:e340.
78 Liu Y, Rohrschneider LR: The gift of Gab. FEBS Lett 2002;515:1–7.
79 Sármay G, Angyal A, Kertész Á, Maus M, Medgyesi D: The multiple function of Grb2 associated binder (Gab) adaptor/scaffolding protein in immune cell signaling. Immunol Lett 2006;104:76–82.
80 Hodgetts RB, O'Keefe SL: Dopa decarboxylase: a model gene-enzyme system for studying development, behavior, and systematics. Annu Rev Entomol 2006;51:259–284.
81 Allison AC: Genetic control of resistance to human malaria. Curr Opin Immunol 2009;21:499–505.
82 Williams TN: Red blood cell defects and malaria. Mol Biochem Parasitol 2006;149:121–127.
83 Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ES: Positive natural selection in the human lineage. Science 2006;312:1614–1620.
84 Thye T, Vannberg FO, Wong SH, Owusu-Dabo E, Osei I, Gyapong J, Sirugo G, Sisay-Joof F, Enimil A, Chinbuah MA, Floyd S, Warndorff DK, Sichali L, Malema S, Crampin AC, Ngwira B, Teo YY, Small K, Rockett K, Kwiatkowski D, Fine PE, Hill PC, Newport M, Lienhardt C, Adegbola RA, Corrah T, Ziegler A, Morris AP, Meyer CG, Horstmann RD, Hill AV: Genome-wide association analyses identifies a susceptibility locus for tuberculosis on chromosome 18q11.2. Nat Genet 2010;42:739–741.
85 Shapira M, Hamlin BJ, Rong J, Chen K, Ronen M, Tan MW: A conserved role for a GATA transcription factor in regulating epithelial innate immune responses. Proc Natl Acad Sci USA 2006;103:14086–14091.
86 Bellusci S: Lung stem cells in the balance. Nat Genet 2008;40:822–824.
87 Zhang Y, Goss AM, Cohen ED, Kadzik R, Lepore JJ, Muthukumaraswamy K, Yang J, DeMayo FJ, Whitsett JA, Parmacek MS, Morrisey EE: A Gata6-Wnt pathway required for epithelial stem cell development and airway regeneration. Nat Genet 2008;40:862–870.
88 Fellay J, Shianna KV, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal S, Martinez-Picado J, Miro JM, Obel N, Smith JP, Wyniger J, Descombes P, Antonarakis SE, Letvin NL, McMichael AJ, Haynes BF, Telenti A, Goldstein DB: A whole-genome association study of major determinants for host control of HIV-1. Science 2007;317:944–947.
89 Pelak K, Goldstein DB, Walley NM, Fellay J, Ge D, Shianna KV, Gumbs C, Gao X, Maia JM, Cronin KD, Hussain SK, Carrington M, Michael NL, Weintrob AC; Infectious Disease Clinical Research Program HIV Working Group; National Institute of Allergy and Infectious Diseases Center for HIV/AIDS Vaccine Immunology (CHAVI): Host determinants of HIV-1 control in African Americans. J Infect Dis 2010;201:1141–1149.
90 Fellay J, Ge D, Shianna KV, Colombo S, Ledergerber B, Cirulli ET, Urban TJ, Zhang K, Gumbs CE, Smith JP, Castagna A, Cozzi-Lepri A, De Luca A, Easterbrook P, Gunthard HF, Mallal S, Mussini C, Dalmau J, Martinez-Picado J, Miro JM, Obel N, Wolinsky SM, Martinson JJ, Detels R, Margolick JB, Jacobson LP, Descombes P, Antonarakis SE, Beckmann JS, O'Brien SJ, Letvin NL, McMichael AJ, Haynes BF, Carrington M, Feng S, Telenti A, Goldstein DB: Common genetic variation and the control of HIV-1 in humans. PLoS Genet 2009;5:e1000791.
91 de Bakker PIW, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, Ke X, Monsuur AJ, Whittaker P, Delgado M, Morrison J, Richardson A, Walsh EC, Gao X, Galver L, Hart J, Hafler DA, Pericak-Vance M, Todd JA, Daly MJ, Trowsdale J, Wijmenga C, Vyse TJ, Beck S, Murray SS, Carrington M, Gregory S, Deloukas P, Rioux JD: A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat Genet 2006;38:1166–1172.
92 Charles BA, Shriner D, Doumatey A, Chen G, Zhou J, Huang H, Herbert A, Gerry NP, Christman MF, Adeyemo A, Rotimi CN: A genome-wide association study of serum uric acid in African Americans. BMC Med Genomics 2011;4:17.
93 Spencer CC, Su Z, Donnelly P, Marchini J: Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. PLoS Genet 2009;5:e1000477.
94 Teo YY, Inouye M, Small KS, Gwilliam R, Deloukas P, Kwiatkowski DP, Clark TG: A genotype calling algorithm for the illumina BeadArray platform. Bioinformatics 2007;23:2741–2746.
95 H3Africa Consortium: Research capacity. Enabling the genomic revolution in Africa. Science 2014;344:1346–1348.
96 Pare G: Genome-wide association studies – data generation, storage, interpretation, and bioinformatics. J Cardiovasc Transl Res 2010;3:183–188.
97 Shriner D: Investigating population stratification and admixture using eigenanalysis of dense genotypes. Heredity (Edinb) 2011;107:413–420.
98 Shriner D, Adeyemo A, Rotimi CN: Joint ancestry and association testing in admixed individuals. PLoS Comput Biol 2011;7:e1002325.
99 Tang H, Siegmund DO, Johnson NA, Romieu I, London SJ: Joint testing of genotype and ancestry association in admixed families. Genet Epidemiol 2010;34:783–791.
100 de Vries J, Jallow M, Williams TN, Kwiatkowski D, Parker M, Fitzpatrick R: Investigating the potential for ethnic group harm in collaborative genomics research in Africa: is ethnic stigmatisation likely? Soc Sci Med 2012;75:1400–1407.
101 Tindana P, Bull S, Amenga-Etego L, de Vries J, Aborigo R, Koram K, Kwiatkowski D, Parker M: Seeking consent to genetic and genomic research in a rural Ghanaian setting: a qualitative study of the MalariaGEN experience. BMC Med Ethics 2012;13:15.
102 de Vries J, Bull SJ, Doumbo O, Ibrahim M, Mercereau-Puijalon O, Kwiatkowski D, Parker M: Ethical issues in human genomics research in developing countries. BMC Med Ethics 2011;12:5.

103 Walker ER, Nelson CR, Antoine-LaVigne D, Thigpen DT, Puggal MA, Sarpong DE, Smith AM: Research participants' opinions on genetic research and reasons for participation: a Jackson Heart Study focus group analysis. Ethin Dis 2014;24:290–297.
104 Halverson CM, Ross LF: Engaging African-Americans about biobanks and the return of research results. J Community Genet 2012;3: 275–283.
105 Yu JH, Crouch J, Jamal SM, Tabor HK, Bamshad MJ: Attitudes of African Americans toward return of results from exome and whole genome sequencing. Am J Med Genet A 2013;161A:1064–1072.
106 Wright GE, Koornhof PG, Adeyemo AA, Tiffin N: Ethical and legal implications of whole genome and whole exome sequencing in African populations. BMC Med Ethics 2013;14:21.
107 Parker M, Bull SJ, de Vries J, Agbenyega T, Doumbo OK, Kwiatkowski DP: Ethical data release in genome-wide association studies in developing countries. PLoS Med 2009; 6:e1000143.
108 Tindana PO, Rozmovits L, Boulanger RF, Bandewar SV, Aborigo RA, Hodgson AV, Kolopack P, Lavery JV: Aligning community engagement with traditional authority structures in global health research: a case study from northern Ghana. Am J Public Health 2011;101:1857–1867.
109 Kamuya DM, Marsh V, Kombe FK, Geissler PW, Molyneux SC: Engaging communities to strengthen research ethics in low-income settings: selection and perceptions of members of a network of representatives in coastal Kenya. Dev World Bioeth 2013;13:10–20.
110 Lemke AA, Wu JT, Waudby C, Pulley J, Somkin CP, Trinidad SB: Community engagement in biobanking: experiences from the eMERGE network. Genomics Soc Policy 2010;6:35–52.
111 Buseh AG, Stevens PE, Millon-Underwood S, Townsend L, Kelber ST: Community leaders' perspectives on engaging African Americans in biobanks and other human genetics initiatives. J Community Genet 2013;4:483–494.
112 Marsh VM, Kamuya DM, Mlamba AM, Williams TN, Molyneux SS: Experiences with community engagement and informed consent in a genetic cohort study of severe childhood diseases in Kenya. BMC Med Ethics 2010;11:13.