

## Chapter 2

# Genetic Basis of Health Disparity

**Abstract** People of the African diaspora often have a high prevalence of certain common and complex diseases, and genetic predisposition has been proposed as one of the causes. Genetic studies have identified variation or mutations in genes that are commonly associated with both common and complex diseases. The completion of the human genome sequence and technological advances, including the so-called genome-wide association studies, have identified genetic variants that are associated with increased disease risks/resistance in individuals of the African diaspora. The patterns of genetic variations in different populations could shed some light on the biological mechanisms underlying health disparities.

### 2.1 Introduction: Genetic Basis of Human Diseases

Our understanding of the important role of genes in determining physical traits and susceptibility to diseases can be traced back to the pioneer work carried out by Gregor Mendel (Mendel 1866) on the transmission of certain traits in the common pea plant (*Pisum sativum*). Mendel studied the inheritance of simple and obvious traits beginning with pure-breeding populations of yellow seeds that produced only plants with yellow seeds and green seeds that produced plants with only green seeds, over several generations. When he crossed a yellow seed with a green seed, all he saw was green seeds in the first generation, but then, when he crossed the green seeds to themselves or with each other, they produced yellow and green seeded plants. Mendel reasoned that these traits (now known as genes) were transmitted with precision from one generation to the next, a discovery that formed the foundations of the genetics of heredity. Mendel's work clearly demonstrated that a single trait (as in the case of pea-plant color) can be determined by a single gene, and we now know that most genes in eukaryotic organisms follow this Mendelian pattern of inheritance. Gregor Mendel's work lead to two important principles governing heredity, namely the laws of segregation and independent assortment also referred to as the Mendelian laws of heredity (Strachan and Read 1999).

How the genetic information is stored or the mechanism of transmitting this information from one generation to the next was not known until the discovery of deoxyribonucleic acid (DNA) as the genetic material by Watson and Crick in 1943. This milestone scientific breakthrough set the stage for understanding how genetic information could be successfully passed down from one generation to the next, how changes in the DNA structure could give rise to variation in gene expression, differences in physical phenotypes, and changes in metabolic pathways, as well as give rise to disease susceptibility or resistance.

DNA consists of four sugar (nucleotide) bases: Adenine (A), Cytosine (C), Guanine (G), and Thymine (T) that can be linked together in any number of combinations as a linear polymer by a phosphodiester bond. With the exception of some viruses and prokaryotic organisms, the DNA structure in most organisms exists as double-stranded molecules in which complementary strands pair together, such that the As on one strand will always pair with the Ts on the second strand, whereas the Cs will always pair with the Gs forming base pairs. Based on the completion of the human genome sequence, it is now known that the human genome contains approximately three billion base pairs of DNA arranged on 46 chromosomes (22 autosome pairs plus the X and Y chromosomes).

The Watson and Crick discovery paved the way for a detailed understanding of how DNA is transcribed into RNA, which is in turn translated into protein that perform all the biological functions for normal physiological processes in living organisms. Thus changes in DNA structure, such as mutations that affect protein encoding genes, could affect gene expression and ultimately lead to heritable changes in organisms, as well as contribute to disease susceptibility. Although DNA is made up of only four bases, there are remarkably different combinations of base nucleotides that encode for all the information needed to make all proteins inside a cell. Thus, the structure and hence the properties of all the enzymes and proteins, in essence all the components that make up a living organism including human beings, are contained in genes. It is estimated that there are between 20,000 and 25,000 protein-coding genes in the human genome necessary for all physiological functions.

The DNA molecule encodes for all the proteins expressed inside cells, making this molecule very important for normal cellular processes. The genes encoded by DNA are involved in determining normal phenotypic characteristics such as skin color, hair, body form, and sex. Alterations in the DNA molecule can lead to changes at the protein expression level, such as insufficient expression, absence of expression, or over-expression of proteins, which consequently alter their normal biological functions, and this is one way that genetic alteration can lead to the development of disease. For example, a single nucleotide change in a beta-globin gene creates a mutant form of beta-globin protein, which, under an environmental situation such as oxygen stress, creates a condition known as sickle cell anemia. This condition causes the shape of normal blood cells to “sickle,” impeding blood flow and causing severe pain and, in some extreme cases, even death.

Some genetic diseases are directly caused by mutations in a single gene (monogenic), and this type of disease follows the Mendelian pattern of inheritance. This Mendelian pattern of inheritance may be the result of an affected individual having a mutation in one copy of the gene (autosomal dominant mutation), and, alternatively, mutations in both copies of the gene may be necessary for the disease to manifest itself (autosomal recessive). In some instances, there is no dominant or recessive phenotype, but both alleles are present and expressed equally (co-dominant). Other genetic disorders that follow this Mendelian pattern of inheritance are associated with the X (female) or Y (male) chromosome and known as sex-linked mutations. An example is the classical hemophilia (X-linked dominant mutation) condition whereby mutated genes associated with this condition in an affected male (because he has only one X chromosome) can cause him to bleed uncontrollably, although such conditions are relatively rare.

Studies in humans with these inheritable defects have shed light on the normal function of the protein product encoded by these genes and the pathways that are implicated in genetic disorders. Another classic example of such an inherited disease is cystic fibrosis. The cystic fibrosis transmembrane conductance regulator gene (CFTR) encodes for a protein that functions in channeling chloride ions across cell membranes in cells that produce mucus, sweat, saliva, or digestive enzymes. A common mutation, a three nucleotide deletion in cystic fibrosis Delta-F508 that removes a phenylalanine codon (Goossens 1991), results in patients being unable to produce this essential protein and causes severe problems in lung and pancreatic function, a condition known as cystic fibrosis. Another genetic disease that follows the Mendelian law of inheritance is Duchenne muscular dystrophy (DMD). The DMD encodes for a protein called dystrophin whose biological function is maintaining muscle integrity. Mutations in the DMD gene cause the most common type of hereditary muscle-wasting diseases, collectively called muscular dystrophies (Tidball and Wehling-Henricks 2004).

There are a large number of other diseases that do not follow the Mendelian pattern of inheritance. These are considered polygenic or complex diseases in the sense that they tend to cluster in families but do not follow Mendelian inheritance patterns, and in some instances there are multiple gene–gene interactions that underlie one particular disease phenotype. For instance, it was discovered that genes on the same chromosome (linkage) are passed along in tandem unless meiotic crossover occurs such that some diseases tend to be caused by two or more genetic factors that contribute to a trait, and there is also a strong influence by the environment.

Therefore, complex diseases can arise as a result of multiple gene–gene interactions and/or gene–environmental interactions in susceptibility genes that act together and contribute to both the occurrence and the severity of a disease in an individual. Examples of complex diseases include diabetes, obesity, heart disease, some types of cancer, a variety of mental disorders that have at least some heritable properties, and many more. There are also instances where mutations in any one of multiple different genes can cause the same disease. This is exemplified by retinitis pigmentosa, a disease that causes degeneration of the retina that, if not treated,

can ultimately lead to blindness. More than 40 different genes with varied biological functions have been associated with retinitis pigmentosa (Ferrari et al. 2011). Several of the genes are expressed in the epithelium of the retinal pigment or encode for photoreceptors.

## 2.2 What Gives Rise to Genetic Variation?

All human beings, regardless of race, ethnicity, or sex, share 99.5% similarity in their DNA sequence. Thus, the 0.5% difference in the DNA sequence in individuals belonging to the same or different ethnic population is what contributes to the surprising diversity in human phenotypic differences, including hair and skin color, height, and body shape. In addition, this sequence difference is responsible for variations in disease susceptibility and resistance and accounts for differential responses to therapeutic interventions.

With the completion of the human genome project (discussed in Sect. 2.3), it is now known that there are between ten million to 15 million variants in common sequences with sufficient frequency in one or more populations to be polymorphic in humans. The majority of these sequence variants are single nucleotide polymorphisms (SNPs), whilst others are insertions or deletions of short sequence fragments and copy-number variations. Sequence variations, and in particular SNPs, have been around for a long time as a result of natural selection. Some of these sequence variations are natural variations in the human genome that are associated with the phenotypic differences observed in individuals such as hair, eye, and skin color, as well as height or body shape, but other variations contribute to disease susceptibility and drug resistance, and these could be of biomedical and/or therapeutic interest. The challenge is how to distinguish between which genetic variants are associated with natural selection and which contribute to disease risk and susceptibility.

Single nucleotide polymorphisms are the most-frequently inherited genetic variations and usually serve as biological markers rather than as underlying causes of disease, and many of these SNPs do not have any obvious biological role. Studies carried out by the international SNP consortium on the entire human genome have identified more than four million SNPs,  $\sim 1$  per 1–2 kb (Lander et al. 2001). However, most SNPs studies have been carried out in the EA population, and very few such studies have been reported for the African or other racial and ethnic populations.

It is important to identify SNPs across different ethnic populations because the frequency of SNPs can vary substantially between different populations, as well as their usefulness as markers for gene-mapping studies (Chakravarti 2001). Significant differences in SNP frequencies could uncover a putative disease marker that may eventually be used in diagnosis. It is estimated that it takes about 40,000 generations for a new SNP to form. It is also estimated that ancient pieces of DNA sequence segments ( $\sim 5 \times 10^6$  bp long) called haplotypes, such as maternal or

paternal haplotypes, are inherited intact for several hundreds of generations and can serve as useful tools to study ancestry. Thus, a small DNA segment of about 10 kB may contain approximately 40 SNPs that is enough to distinguish it from a haplotype of a different person with the same-sized fragment, and this would indicate that the second person got his/her segment of haplotype from a different ancestor (different sets of alleles have enough characteristic differences). Supposedly, we all have segments of Neanderthal DNAs in our genome, which means that we can detect ancient ancestral haplotypes. The ability to use genetic markers to predict ancestry is driving ancestry mapping using either mitochondria DNA (mtDNA; the genome within mitochondria that is passed on through the maternal lineage) or the Y-chromosome DNA that is passed on from father to son.

Human DNA sequence analysis shows that various populations differ in the frequency of both mutations of biomedical interest and polymorphisms of biological interest. There is, therefore, a need to understand the causes of these genetic variations in various populations in order to understand the biological basis of health disparities. Genetic variations or mutations in DNA occur constantly, and this can arise during DNA synthesis in replication, chromosomal crossing over as in sexual reproduction, or as a result of natural selection due to environmental pressures, such as disease, diet, or parasitic infection and even climatic changes, to create certain phenotypes. These genetic variations do not result in the creation of new species or organisms; rather, they result in the adaptation of existing organisms and their progenies to new environments through natural selection. Over time, the small alterations in cellular structures and functions may prove to be advantageous. Thus human genetic variation is the end result of diverse forces in nature.

Genetic variations that arise as a result of human migration cause adaptation and influence genetic susceptibility to diseases. An advantage to identifying the unique patterns of genes across various populations and disease susceptibility will contribute to understanding complex diseases, and that will ultimately lead to the design of novel therapies targeted to the populations that are most likely to respond. Others argue against the use of genetic variation as the basis for categorizing various populations into ethnic and racial groups and the disease prevalence and health outcomes in various groups (Cooper et al. 2003), noting that the vast majority of genetic variation (90–95%) occurs within, not across, human populations (Braun 2002) and that the *Homo sapiens* species consists of a single population. Therefore, biologically distinct human races do not exist, as genetic studies clearly demonstrate that human beings share 99.5% of their DNA in common.

## 2.3 Tools for Studying Genetic Variants

Genetic studies over the past century have been centered on finding traits and, in particular, disease traits and the genes that are responsible for causing these traits in populations where a particular disease is rampant. Such studies have been carried out in families that transmit the disease traits from one generation to the next.

Various molecular approaches are now used to identify chromosomes that carry the trait and then to map in detail of the chromosomal region until the gene is identified. Linkage studies have been instrumental in identifying chromosomal location of disease-associated candidate genes. This is made possible by the observation that genes that reside close together on a chromosome are linked together during meiosis. By definition, genes are said to be in linkage equilibrium if two genes/traits/loci are inherited completely independently of each generation. On the other hand, two genes are said to be in linkage disequilibrium (LD) if the frequency of association of different genes are inherited together more often than would be expected by chance. In linkage studies, data from multiple families with the same disease are compared or combined to determine whether a statistically significant linkage exists between the disease gene and other known molecular markers for the condition.

Once disease candidate genes are identified, then linkage analysis is carried out in human pedigrees. Typical protocols involve a few to large number of pedigrees, typically using many families (e.g., sibling pairs) such as twin studies, where concordance rates are compared between monozygotic (MZ) and dizygotic (DZ) twins. Twins are commonly used in these studies because it is usually assumed that twins share a common environment that lessens the impact of environmental influence (although this may not be true for studies of adult twins). For example, one study found a higher concordance for cellular immune responses to mycobacteria and other antigens in MZ compared to DZ twins, suggesting that genetic factors are important regulators of this immune response (Newport et al. 2004). There are other approaches, such as case-control association studies and family-based association, as well as linkage studies where the co-segregation of a marker with the disease phenotype are tested in families (Hill 2006).

However, genetic heterogeneity such as that exhibited by retinitis pigmentosa can confound such an approach because any statistical trend in the linkage data from one family or group tends to be canceled out by another set of data obtained from a different family with an unrelated causative gene(s). Furthermore, in the epigenetic phenomena, which are discussed in Chap. 3, environmental and numerous other factors come into play. The take-home message is that genetic factors that underlie disease often work within complicated, poorly understood, and highly nonlinear relationships. This explains why the search for signature disease genes (those with allelic variants thought to be primarily responsible for disease) has largely been a failure when the underlying gene(s) is known because, while the gene(s) may confer an increased risk, it does not directly cause the disease.

Much of the early genetic studies have led to the discovery of the alterations in genes that underlie or contribute to disease etiology and/or progression, albeit such studies were focused on finding one disease gene at a time. Beginning in the 1970s with Frederick Sanger, sequencing technology that enables reading off a DNA sequence, one nucleotide at a time, has made it possible for researchers to sequence one gene, many genes, and eventually an entire organism. This technique enabled other novel techniques for DNA, RNA, and protein analysis to follow rapidly, culminating in the era of genetic-recombination technology.

The human genome project has revolutionized the approach for searching for disease genes. The complete list of three-billion bases that make up the human genome was published in 2003. It took more than a decade and at a cost of US\$3 billion in federal government funds to accomplish this feat. Nowadays, it cost about US\$1000 to sequence an individual's genome. Advances in cutting-edge technologies including dense genotyping arrays and genome sequencing are driving down the cost of whole genome sequencing, making it more robust and affordable. It is estimated that in the near future, every person can have their genome sequenced for much less than US\$1000.

The completion of the human genome sequence has immense scientific benefits: The speed with which the genetic basis of disease is being unraveled has largely been aided by the completion of the reference human genome sequence. This has enabled cataloging of sequence variations in individuals belonging to the same or different racial or ethnic groups (including both healthy individuals and those with various diseases). This, in turn, is providing new insight into understanding molecular processes that underlie disease and disease susceptibility. The challenge still remains as to how to distinguish sequence variations of biomedical interest (i.e., mutations that contribute to disease) from natural variations and to fully elucidate the significance and impact of sequence variations in physiological processes.

One of the potential benefits from the completed human genome sequence project is the design of genome-wide association studies (GWAS), which, in conjunction with powerful statistical and other computational tools, have become **useful for assessing genetic variants and association with traits or disease** in various individuals or populations. GWAS are useful for the study of both Mendelian patterns of disease inheritance and complex genetic diseases (many genetic and environmental factors acting together). The GWAS approach is instrumental in elucidating genetic variants associated with increased disease risk/resistance in one racial or ethnic population in comparison to another group and has the potential for being used to develop better markers to detect, treat, and prevent disease. As a result of GWAS, there is increasing insight into and the discovery of more genetic variances that are associated with the high prevalence of certain common and complex diseases in the African population and how this may contribute to disease disparities in this population when compared to other racial and ethnic populations. GWAS are increasingly being utilized in pharmacogenomics, which is driving the future of precision medicine.

### ***2.3.1 Genome-Wide Association Studies (GWAS)***

DNA sequencing (candidate gene, exome, or whole genome) and linkage analysis have been successful in the identification of germline and somatic mutations. The GWAS approach is another level of DNA sequencing that is useful in the identification of susceptible genetic variants or alleles for many diseases and traits. Unlike the classical approach of investigating genetic variation one gene at a time,

the GWAS approach enables comprehensive analysis of the genetic variation across the entire human genome. GWAS take advantage of linkage disequilibrium (LD), i.e., non-random association of alleles at adjacent loci, and focus on common variations (>5% frequency) not limited to coding regions and not limited to prior biological knowledge. This approach has led to the identification of rare variants and unidentified common variants and structural variations that has eluded previous investigations.

Briefly, GWAS rely on hybridization of genomic DNA extracted from human tissues or blood obtained from patients with a particular disease (case) and individuals without the disease (control) in glass chips containing hundreds to thousands of SNP markers across the entire genome using microarrays in order to identify differential genetic variants in the case versus control groups. The approach is to scan several hundreds to thousands of genetic variants for a particular disease gene by comparing thousands of individuals with the disease with those who do not have the disease in order to link variations to particular traits and diseases in a manner that was not possible before. The data obtained from such analysis is compared to a computational database that contains the reference human genome sequence. This approach makes possible examining a large number of DNA markers in populations of individuals with a particular disease, as well as in control populations of individuals without the disease, to find statistical correlations between group phenotype and polymorphic markers. Thus, if certain variants are found to be significantly more frequent in the disease population in comparison with the non-disease population, the variants/loci are said to be associated with the disease. The associated genetic variations can serve as markers or powerful pointers to disease loci on the human genome.

This approach has been used to identify multifactorial susceptibility alleles associated with diseases. For example, GWAS have identified over 200 multigenic allelic variants to be associated with diabetes- disease susceptibility, physiological, and behavioral traits. Another GWAS has identified 18 risk loci in the EA and Asian populations to be associated with germline pancreatic cancer (Amundadottir 2016). Furthermore, GWAS have the potential for the identification of disease susceptibility variants that are associated with survival, therapeutic dosage responses, gene–gene interactions, or the environmental factors that act in concert with genes.

Various organizations, government, and private sectors including the National Institutes of Health (NIH), Pfizer, and others have formed public-private partnerships to fund GWAS. Although GWAS can be a powerful tool to identify candidate disease genes, additional studies are needed in order to determine how an individual carrying a particular genetic variant might be predisposed to the disease. Nowadays, GWAS are used in association with meta-analysis to reveal associations of heritable variations with disease risks. However, sensitivity can be a concern with the undetected heritability reported in most GWAS studies (Eichler et al. 2010). As the field advances, more complex and improved statistical tools will help in the identification of heritable variances with low penetrance. One recent statistical model that took into consideration epistasis, i.e., the interaction of multiple genetic



variants in influencing a trait, found that in some studies the total frequency of heritability is much higher than reported (Zuk et al. 2012).

Some of the notable success stories of the GWAS studies are reported in a landmark paper from the Wellcome Trust Case Control Consortium, which carried out GWAS on a cohort of 2000 individuals for seven major diseases: coronary artery disease, type-1 and type-2 diabetes, hypertension, rheumatoid arthritis, bipolar disorder, and Crohn's disease (The Wellcome Trust Case Control Consortium 2007). They identified 24 independent disease risk associated loci: one in coronary artery disease, seven in type-1 diabetes and three in type-2, three in rheumatoid arthritis, one in bipolar, and nine loci in Crohn's disease. In addition, some of the loci were associated with the risk of more than one of the diseases studied.

## 2.4 Genetic Variation in the African Population

Research in human evolutionary genetics and the identification of fossil remains with resemblance to modern humans in East Africa support the existence of human life in Africa about 200,000 years ago, which then gradually spread across the rest of the world within the past ~100,000 years (Campbell and Tishkoff 2008). Therefore, modern humans have continuously existed on the African continent longer than any other part of the world. The African continent has a wide range of environments, including deserts, tropical rainforests, savannas, swamps, mountain highlands, and coastal plains (Campbell and Tishkoff 2008). Climatic changes mean that some of these environments have undergone dramatic changes over the course of time (Campbell and Tishkoff 2008; Scholz et al. 2007; Trauth et al. 2007). Because of this considerable diversity, the environment has given rise to an African population with a diverse range of linguistics, cultures, and diet. Furthermore, differences in exposure to pathogens have given rise in the African populations to selection pressures that have also contributed to the most genetic adaptation and phenotypic variation (Campbell and Tishkoff 2008; Reed and Tishkoff 2006) that is not found in non-African populations. Greater genetic diversity in the African population is also reflected in higher levels of nucleotide and haplotype diversity in African populations, also not found in non-African populations, in both nuclear and mitochondrial DNA (Tishkoff et al. 2003).

### 2.4.1 *Causes of Genetic Variation in the Human Population—A Case for Disease*

The introduction of pathogens into a population with no previous exposure plays a very important role in natural selection. This can result in infectious disease epidemics where a large proportion of susceptible individuals dies, with a remnant of

the surviving population, who are now adapted to the pathogenic exposure and equipped with protective allelic variants.

Natural selection can have either positive or negative consequences. A positive natural selection such as increased fitness is a positive adaptive response that gives rise to individuals with an advantageous genetic variant that makes them more likely to have more offspring than individuals with another genotype or variant, i.e., fitness benefit allele. A negative natural selection, on the other hand, leads to deleterious mutations that reduce fitness. Individuals with deleterious mutations are less likely to produce offspring, and the lineage is more likely to die off eventually. Some of the best described natural selections are restricted to infectious diseases (e.g., resistance to malaria), or nutritional adaptations (e.g., lactose tolerance). For instance, genetic variation in the Duffy gene and receptor is a classic example of a positive genetic selection and is a marker for protection against malaria (Oliveira et al. 2012).

Other studies have shown selection for genetic variation in the NADSYN1 gene on human chromosome 11q13 that may play a role in prevention of pellagra, whereas differential geographic variation in the toll-like receptor at chromosome 4p14 has been linked with differential susceptibility to autoimmune diseases, such as tuberculosis and leprosy (Todd et al. 2007). It is increasingly becoming apparent that varying frequencies of heritability and somatic genetic mutations (or variants), in addition to lifestyle and environmental factors, underlie disease etiology, as well as the disparities associated with several complex diseases. It is also increasingly becoming clear that genetic variation may be advantageous for individuals in one particular environment but can become disadvantageous in a totally different environment. Further studies of these selected variants may help to elucidate the correlation between genotype and phenotype for genes that play a role in important diseases.

A classic example of the effect of natural selection on a population is described by the work of a Danish physiologist Dr. P.L. Panum. In Dr. Panum's published report entitled "Observations made during the epidemic of measles on the Faroe Islands in the year 1846" (Poland 1998), he described an epidemic of measles on the Faroe Islands where 6000 of the 7800 people of that island developed measles and more than 200 died from the disease. A previous measles epidemic had occurred in 1781, and it was discovered that none of the people previously exposed in the 1781 epidemic (many still living in 1846) developed the disease. Some of the questions that arose from Dr. Panum's studies pertain to why some people survived the measles epidemic while others died, why the mortality was so high, and why the protective effect from the past epidemic was so high. Some of these issues can be resolved by understanding the role of innate susceptibility and resistance, as well as the acquired immunity conferred by previous exposure. The human leukocyte antigen (HLA) is important in disease susceptibility, resistance, and progression (Bodmer 1996).

The HLA gene complex encodes for HLA antigens that are expressed as white blood cells, antigen-presenting cells, and other tissues. The HLA proteins play a very important role in binding and presenting pathogens to T cells for destruction.

Therefore, the ability of an individual to respond to viral infection (or vaccination) by antibody production is partly under the control of immune-response genes within the HLA complex. Genetic variations in HLA genes and their interactions with other factors (such as host or environmental pathogens) are increasingly recognized as playing a significant role in disease susceptibility and resistance, as well as disease progression (Bodmer 1996). Genetic variation in HLA isoforms is not only crucial in immune responses but also linked to various diseases including diabetes, Hepatitis C virus (HCV) or HIV infection, cancer, and spontaneous abortion, as well as the outcomes of organ transplantation.

Immune adaptation to infectious disease as a result of previous exposure in combination with genetic susceptibility or resistance underlies the disparities associated with infectious diseases. Natural selection may give rise to protective alleles in individuals or populations exposed to adverse environmental exposures, such as infectious epidemics for the progeny of those that survive; however, this process may take decades to several hundreds of years of constant and high exposure (Ramsay 2012). Genetic variations that are associated with disease risk could be geographically restricted as a result of new mutations, genetic drift, or region-specific pressures (Campbell and Tishkoff 2008). All we know is the pattern of variation we see today, which we can use to trace what happened in the past.

## 2.5 Infectious Diseases and Inherited Conditions in African Populations

Historically, infectious diseases have been the most important cause of morbidity and mortality globally until recently when non-communicable diseases started to rival or even exceed that of infectious diseases in many parts of the world. For instance, infectious diseases were the number one killer in the US until the early twenty-first century, when scientific breakthroughs in immunization significantly reduced mortality rates from infectious diseases. However, mortality rates as a result of infectious disease continue to disproportionately affect individuals on the African continent as compared to any other part of the world. Africans have an extremely high mortality rate from infectious diseases, with malaria and HIV alone responsible for millions of deaths per year (<http://www.cdc.gov/malaria/index.htm>) (Anon 2006a).

Malaria is a major killer of children in Africa and also in other parts of the world. Malaria is also perhaps the most important evolutionary driving force for natural selection of other commonly known genetic diseases that have evolved as a by-product to protect against malaria. For instance, sickle cell, glucose-6-phosphatase deficiency, and other erythrocyte defects are all diseases that may have been selected evolutionarily to protect against malaria (Weatherall and Clegg 2001). Given that malaria exerts such a powerful evolutionary force on human genetic variation, malaria protective variants are likely to be in high frequencies in

affected populations. Thus, the patterns of genetic variations and association with neighboring genetic markers (such as strong LD if recently selected) could contribute to understanding any immunologic, inflammatory, or chronic diseases that are more prevalent in the African population, African diaspora, or other racial and ethnic populations.

### 2.5.1 Malaria

Malaria accounts for one in five childhood deaths in Africa and one million deaths per year globally, making it a serious global health issue. The World Health Organization (WHO) estimated that almost 74% of the African population live in areas endemic with malaria, with about 19% in epidemic-prone and only 7% in malaria-free areas (Anon 2006b). Malaria is a complex disease whereby the malaria parasite infects the host red blood cells (RBCs) and induces changes that can ultimately destroy the RBCs (discussed below). Clinically, malaria can be either uncomplicated or severe. In the severe cases, various organs and tissues can be affected, and this can lead to death. Genetic studies have identified polymorphisms or variants in the disease pathways that contribute to protection against malaria as mentioned in Sect. 2.4.1.

The molecular mechanisms of erythrocyte invasion by malaria parasites is central to the disease process. The signal transmission involves binding of the malaria parasite to Duffy glycoproteins, cytokine receptors that are expressed on RBCs. The importance of Duffy signaling is demonstrated by the observation that RBCs that do not express the Duffy antigens (encoded by the FY gene) are resistant to invasion by the malaria *Plasmodium vivax* (*P. vivax*) parasite. The absence of Duffy expression is the result of a SNP variant in the Duffy promoter that alters the binding interaction and transcription from the GATA-1 transcription factor (Tournamille et al. 1995). Thus, individuals that are Duffy (negative; homozygous mutant) are fully protected against invasion by this parasite. However, other reports indicate that individuals who are heterozygous carriers of the Duffy-negative allele have some susceptibility to *P. vivax* infection. The *P. vivax* is one of many malaria parasite species and the non-lethal form of malaria parasites. To complicate matters further, other reports indicate that *P. vivax* can infect RBCs and cause clinical manifestation of malaria in Duffy-negative individuals, suggesting that the relationship between *P. vivax* and Duffy antigen are more complex than previously described (Zimmerman 2013).

It appears that, in some populations, alpha-thalassemia, another common genetic disorder of erythrocytes, has been naturally selected for protection against malaria. Here again a complication is the observation that, through cross-immunity, alpha-thalassemia may increase susceptibility to *P. vivax* that could be beneficial to humans because it confers some natural immunity by protection against infection of the more severe *Plasmodium falciparum* (*P. falciparum*) (Williams et al. 1996). However, one recent report challenges this notion of cross-immunological

protection between *P. vivax* and *P. falciparum* against the more severe *P. falciparum* in alpha-thalassemia patients (Rosanas-Urgell 2012). Additional studies are warranted to clarify these inconsistencies between the cross-immunological protections of one parasite and another.

Knowledge of the genetic basis for disease resistance or susceptibility to infectious disease in the oldest human population can guide research and medical intervention, and these benefits are not limited to African populations. The following section will describe some of the genetic disorders that have resulted from selection by, and protection against, malaria, a common and perhaps the single most important public health concern in Africa. Understanding the biology of genetic variations and association with susceptibility to infectious diseases is crucial in improving the overall health of people on the African continent, but it is also important in understanding the role of genetic susceptibility and the higher disease prevalence among the US AA population.

### 2.5.2 Sickle Cell Disease

Sickle cell disease has been extensively studied as a classic single gene disease that fits the Mendelian pattern of inheritance. Sickle cell disease is predominantly seen in individuals of African descent; however, sickle cell disease is common in other geographic areas where malaria is also common, such as Arabian Peninsula. The disease is caused by point mutation in the beta-globin gene, an important subunit of hemoglobin, the protein molecule in RBC that plays a role in binding oxygen and the transport of oxygen throughout the body. The disease is inherited in an autosomal recessive pattern, in that individuals who are homozygous for the mutant gene have the full-blown disease, whereas individuals heterozygous for the mutation do not have the disease but instead are referred to as “carriers.” Individuals with sickle cell disease have an abnormal variant of hemoglobin called HbS in their red blood cells that clump together, unable to bind oxygen, and change from a smooth circular disc-like shape to sickle shape. The clinical symptoms of HbS include hemolysis and microvascular occlusion that destroy RBC and cause tissue damage and bone pain. Despite intense research in this field, there is no cure for sickle cell disease.

Another disease that is caused by mutation in hemoglobin C (HbC) is thalassemia. Interestingly, a variation in the HbC gene is associated with the protection of RBCs from infection by the malaria parasite. For example, one report suggests that a variant of HbC has a protective effect against the severe *P. falciparum* malaria in W. African population and the protection from this variant may be greater than that from the HbS variant (Agarwal 2000). The beta thalassemia (HbB) sickle-cell trait or HbA/HbS heterozygosity is associated with a ten-fold reduction in malaria risk. Variations in HbC and alpha thalassemia (HbA) genes have also been reported to confer protection against the severe form of malaria, as well as malarial anemia (Williams et al. 1996). The HbB gene has also been reported to have some protective effect against malaria; however, this protective effect has only been reported

in isolated pockets of W. Africa population (Willcox et al. 1983). Clearly, differential genetic variants in the hemoglobin genes can influence malaria risk, but the currently known gene variants vary significantly among various populations and do not explain all the disease risks or protection. Further studies are required to investigate the distribution of genetic defects that are associated with the manifestations of the severe form of malaria. Other contributing factors to the differences in susceptibility or resistance to malaria disease includes genetic variants in the immune response genes which are associated with the parasitic invasion of malaria, or some environmental factors (Torcia et al. 2008).

### 2.5.3 *Glucose-6-Phosphate Dehydrogenase Deficiency*

Glucose 6-phosphate dehydrogenase (G6PD) is an enzyme that catalyzes the first step in the pentose phosphate pathway for the production of nicotinamide adenine dinucleotide phosphate (NADPH). In RBCs, NADPH is required for maintaining the integrity of RBC shape. Deficiency of G6PD protein, which is the result of X-linked recessive mutation, is commonly observed in men of African ancestry and reported in about 11% of the AA male population. However, deficiency in G6PD is protective against *P. falciparum* malaria infection, a mutation that has risen to high frequency in populations exposed to malarial infection, despite the negative consequences associated with this deficiency (Sabeti et al. 2002).

Deficiency in G6PD plays another role in the protection of RBCs. Infection of RBCs by the malaria parasites induces oxidative stress in the infected cell; this compromises the pentose pathway and ultimately the cells and parasites die from oxidative damage. Thus, G6PD mutation is a natural selection mechanism of protection against malaria, which accounts for its high frequency among individuals of African descent. The ability of G6PD deficiency to protect against malaria does, however, create a public-health conundrum in that the use of primaquine to treat malaria has an adverse hemolysis effect in patients with G6PD mutation.

### 2.5.4 *Tuberculosis*

A relatively recent infectious disease in comparison to malaria is tuberculosis (TB). It is estimated that about eight million people develop TB disease and about three million patients die from TB every year. The TB causative agent, *Mycobacterium tuberculosis*, has infected over two billion people worldwide, without any clinical symptoms. In Africa alone, there were ~3.8 million cases of TB in 2005, and this was accompanied by almost 550,000 deaths (<http://www.who.int/mediacentre/factsheets/fs104/en/>).

However, the molecular mechanisms causing TB infection is largely unknown. Twin studies in the African population, comparing MZ and DZ twins, have

provided evidence of a significant role for host genetic factors in the innate immune-response pathways that may contribute to TB susceptibility (Jepson et al. 2001). Other reports have demonstrated that variants in several genes including SLC11A1 (NRAMP1) (Awomoyi et al. 2002), IL1B (Awomoyi et al. 2005), and vitamin D receptor (Bornman et al. 2004) with diverse cellular functions are associated with TB susceptibility. Candidate gene studies have also provided some evidence for major TB susceptibility loci in specific populations, but few candidate genes have consistently shown strong association of TB in Africans, suggesting that several loci may determine or modulate susceptibility to TB. Elucidating the biological mechanisms underlying TB infection is further complicated by co-morbid conditions including HIV and HBC infections.

A GWAS study of 96 Moroccan multiplex families including 227 siblings identified a single region on human chromosome 8q12-q13 to be significantly associated with TB. The linkage was strong if one parent was also affected with TB, suggesting an autosomal dominant TB susceptibility loci on human chromosome 8q12-13 (Baghdadi et al. 2006). Several malaria and TB GWAS, such as the MalariaGEN initiative funded by the Gates Foundation ([www.malariagen.net](http://www.malariagen.net)) and the African TB Genetics Consortium and the Wellcome Trust Case Control Consortium (University of Oxford, UK; [www.well.ox.ac.uk/aug-10-tb-susceptibility-locus](http://www.well.ox.ac.uk/aug-10-tb-susceptibility-locus)), are ongoing to determine the etiological and associated genetic risk variants for these diseases.

### 2.5.5 *HIV/AIDS*

Approximately 10% of the world's population resides in sub-Saharan Africa, and 68% of adults and nearly 90% of children infected with HIV-1 virus live in this region. So, approximately 22.5 million Africans are estimated to be infected with the HIV-1 virus, and about 12 million AIDS orphans live on the continent, making Africa the worst affected region in the AIDS pandemic ([www.whois/database/core/core\\_select.cfm](http://www.whois/database/core/core_select.cfm)). Without antiretroviral treatment, the great majority of infected individuals will progress to full-blown AIDS and die following HIV infection. Although the asymptomatic period averages ten years and ranges from a few years to 20 years, there are rare instances of infected individuals who do not progress to AIDS at all. For instance, a group of female sex workers in Nairobi, Kenya who are at risk of HIV infection (Kimani et al. 2008). Such cases demonstrate the complexities in viral infections involving multiple genetic and environmental factors and their interactions, resulting in altered disease susceptibility or resistance (Lama and Planelles 2007).

One report identified variations in interferon regulatory factor 1 (IRF1), which causes deficiency in the gene expression in a cohort of the Kenyan sex workers, is associated with resistance to HIV-1 infection, whereas other variations of the IRF1 gene were found to be associated with disease progression (Ball et al. 2007). These results suggest that the key to protection from infection lies in gene variants that do not support viral transcription and replication. Other studies have identified genetic

variants in genes that play roles in the viral replication pathway including cytidine deaminase enzymes APOBEC3F, APOBEC3G, CUL5, and TRIM5 alpha that are associated with protection in AA against HIV infection, accelerated loss of CD4 lymphocytes, and progression to full blown AIDs (Lama and Planelles 2007).

Several other infectious diseases are relatively common in Africa and present serious public-health issues including leishmaniasis, leprosy, schistosomiasis, and trachoma. However, unlike malaria and HIV/AIDs research, few resources have been devoted to these types of diseases. Overall, HIV/AIDs, TB, and malaria are the three major infectious diseases that continue to devastate sub-Saharan Africans. Malaria is a relatively ancient infectious disease in comparison with HIV/AIDs and TB. The host genetic diversity as a result of natural selection against malaria is well known; however, the genetic diversities associated with HIV/AIDs and TB are in their infancies, and evolutionary forces are still at play. The genetic diversity associated with this trio is not known; whether having the one disease increases protection or susceptibility to the others remains to be established. Because of the high morbidity and mortality rates from infectious diseases on the African continent, there is an urgent need for safe and effective therapeutic interventions.

## 2.6 Non-Communicable Diseases

Although infectious diseases are the most important public-health concerns in Africa to date, the health pattern is rapidly changing as a result of economic development and urbanization. For instance, a recent report from S. Africa indicates that over 75% of black South Africans have at least one major risk factor for heart disease (Tibazarwa et al. 2009). Other common diseases of the Western world such as obesity, diabetes, and hypertension are increasingly becoming prevalent on the African continent as the general population transitions from a rural to an urban lifestyle (Abubakari et al. 2008). The increasing incidences of complex diseases on the African continent correlate with known environmental risk factors, such as urbanization and sedentary lifestyle. Genomic studies including GWAS for elucidating genetic diversity are limited in the African population, although they represent the human population with the highest genetic diversity. It is, however, important to carry out genomic studies in the sub-Saharan African population because their genetic diversity is crucial to understanding the genetic origins to many diseases.

## 2.7 GWAS in the African Population

GWAS approaches have been used extensively to study the genetic variations associated with complex traits in individuals of European decent, but very few studies have been carried out in non-admixed African populations.



Conducting GWAS in the nascent African populations is important because the unique genome of the African populations has an important role to play in understanding human health and disease susceptibility. The challenges of doing such studies in the African population, as explained by Teo et al. (2010), include the diversity and structure of the African population, lack of funding, poor infrastructure and public-health systems, and perhaps the small pool of trained scientists to carry out this work. The International HapMap project ([www.1000genomes.org](http://www.1000genomes.org)) is a leader in GWAS in the African population.

Among the few GWAS that have been carried out have been conducted to identify genetic variants associated with susceptibility to malaria in West-African children. The results from this study highlighted the role of genetic variants in the HbS locus in the disease prevalence (Jallow et al. 2009). Other data from GWAS suggest higher levels of genetic diversity within the African population in comparison with non-African populations, confirming the results of previous mtDNA, X-chromosome, and Y-chromosome studies (Campbell and Tishkoff 2008; Conrad et al. 2006; Frazer et al. 2009). Recent studies have shown considerable variation in SNPs, INDELs, copy-number variations, translocations, and inversions (Scherer et al. 2007) within individuals of African descent, as well as other racial and ethnic populations, with African populations showing the highest levels within populations of genetic diversity relative to non-Africans (Tishkoff et al. 2009). The emerging observations from GWAS is that this approach can be used to roughly cluster diverse racial and ethnic populations by geographic region (Wall et al. 2008).

## 2.8 Genetic Variation in the African Diaspora—The African-American Population

As described in the introductory chapter and elsewhere, one of the largest migrations in modern human history involved the millions of Africans who were brought to the Americas as slaves over a 500-year period. However, the modern AA population has on average approximately 20% European ancestry (Shriver and Kittles 2004), making their genetic pool heterogeneous. The African diaspora in the US continues to share a common genetic heritage with nascent Africans, while, at the same time, being exposed to different environmental and lifestyle factors compared to Africans on the native African continent. Various populations of African descent, therefore, provide a great opportunity to compare and contrast the genetic variants underlying common disease susceptibility in these genetically related populations but in different environments.

Disease types and prevalence differ between the nascent African population in comparison with AA and other racial and ethnic populations in the Western (Europe and N. America) world. As discussed above in Sect. 2.5, infectious and communicable diseases are the most prevalent in the developing world, including in

sub-Saharan Africa, whereas non-communicable diseases are most prevalent in the Western world. Genetic variation has been shown to be associated with susceptibility to various communicable and non-communicable diseases that have a genetic component (Cooke and Hill 2001).

People of the African diaspora often have a high prevalence of certain complex diseases (discussed in the following sections), such as hypertension, diabetes, and cancer for which the genetic bases are poorly understood, and many of the environmental factors that contribute to these complex diseases are not as common in the African populations (Forrester et al. 1998; Kaufman et al. 1999). This makes it possible to differentiate the relative contribution of genetic and environmental risk factors of these diseases. The following sections will discuss the genetic alterations associated with the leading causes of mortality for US racial and ethnic groups, with some reference to the sub-Saharan Africans.

## 2.9 Cardiovascular Disease (CVD) Disparities

According to the World Health Organization (WHO) factsheets, cardiovascular diseases (CVDs) including heart disease and stroke are the leading cause of death globally. Annual statistics indicate that more people die from CVDs than from any other disease. In 2012 alone, an estimated 17.5 million people died from CVDs, representing 31% of total global mortality in that year (<http://www.who.int/mediacentre/factsheets/fs317/en/>). In the US, not only are CVDs the leading cause of mortality, but there also is a disproportionate impact of CVDs in the AA population in comparison with EA counterparts, as well as other racial and ethnic groups. The good news is that the mortality rates for CVDs are on the decline among the general US population as a result of improved cardiovascular procedures such as coronary arteriography and coronary-artery-bypass graft surgery. However, the decline in mortality rates have not been seen in AA populations, whose mortality rates continue to be high (Hertz 2005). One of the factors that is attributable to the disparities in CVD mortality rates is unequal access to cardiovascular treatment procedures.

Other contributing factors to the high mortality rate of CVDs in AA are the high prevalence of associated risk factors or co-morbidities such as diabetes, hypertension (Cooper et al. 1997; Kaufman et al. 1996), and obesity (Colilla et al. 2000; Rotimi et al. 1995). These risk factors can lead to changes in vascular structure, causing vascular dysfunction followed by tissue injury/wall stress resulting in organ dysfunction, end-stage heart failure, and ultimately death. Several scientific reports demonstrate that there is a complex interaction between genetic risk factors (including familial predisposition) and environmental risk factors (including psychosocial stress, smoking, metabolic syndrome, and hypercholesterolemia) in CVD incidence and mortality rates. Other risk factors include congenital cardiac defects that are associated with abnormal blood pressure (BP) level and aberrant expression of vasodilators such as nitric oxide.

### 2.9.1 Cardiovascular Tissue Remodeling and CVDs

Some studies have investigated the mechanical properties of cardiovascular tissue and described decreased vasodilation of the vascular lumen, as well as reduced expression of nitric oxide in AAs when compared to the EA population, and this may be associated with the increased risk of CVDs in the AA population (for a review see Taherzadeh 2010). Other studies have focused on mediators in cardiovascular physiology and disease, including two opposing mediators linking obesity, diabetes, and vascular dysfunction in CVDs.

Endocrine signaling plays a very important role in cardiovascular physiology, as well as disease state. One of the mediators in this signal pathway is angiotensin II, which is important in aldosterone synthesis via the so-called “renin-angiotensin-aldosterone” system (RAAS). Aldosterone plays a very important role in sodium reabsorption, potassium secretion, and stabilization of blood pressure. Abnormally high-level production of aldosterone (hyperaldosteronism) is caused by tumors in the adrenal gland, or in response to other diseases, and can lead to increased blood pressure, hypertension, cardiac remodeling, fibrosis, and impaired endothelial function, and is also implicated in left ventricular hypertrophy (Vlachopoulos and Terentes-Printzios 2015). Other conditions that are associated with hyperaldosteronism are decreased blood potassium level, myocardial infarction, and increased inflammation (Vlachopoulos C and Terentes-Printzios D 2015).

Another mediator in the endocrine-signal pathway is known as peroxisome proliferator activated-receptor gamma (PPAR-gamma), which plays a role in fat-cell differentiation and is expressed in endothelial and vascular smooth-muscle cells. PPAR-gamma is believed to play diverse physiological role in endothelial cells including lowering blood pressure and improving endothelial function and has growth-inhibition and anti-inflammatory properties (Gibbons 2004). Dominant negative mutation in PPAR-gamma is associated with severe insulin resistance and is a risk factor for early-age onset of diabetes and hypertension, suggesting that genetic variants in the PPAR-gamma gene can mediate increased susceptibility to CVDs (Gibbons 2004). Other reports have identified variations in the RAAS and PPAR-gamma genes as associated with differential susceptibility to hypertension and myocardial infarction in the AA, in comparison with the EA, population. Genetic variance associated with lower plasma renin-activity was higher in individuals of African descent with hypertension (Underwood et al. 2010). Other genetic alterations in the endocrine-signaling pathway such the beta-adrenergic receptors are associated with significant increased risk of heart failure in the AA, in comparison with the EA, population.

Several reports have implicated different genetic variants in key regulatory genes in disparities of vasoconstriction and CVDs. One study reported that genetic variation in leukotriene A4 hydrolase gene is associated with a three-fold increased risk of myocardial infarction in AA patients, while the relative risk is only 1.16 in Europeans (Helgadottir et al. 2006). Other reports demonstrated a relationship between genetic factors such as platelets-aggregation and blood-coagulation factors,

as well as specific polymorphisms in beta-adrenergic receptor genes, such as ADRB1, 2, 3 mediate physiological effects of epinephrine and neurotransmitter non-epinephrine, to be modestly associated with disparities in CVDs. Identification of genetic variance in genes associated with CVDs may contribute to understanding the genetic contribution to CVDs' disparity and help in developing novel therapies tailored for various populations.

Not many studies have investigated the genetic factors associated with CVDs' risk in the African population on their native continent, perhaps because this condition was uncommon until the 1970s. One South African study demonstrated that a family history of CVDs was significantly associated with myocardial infarction, hypertension, and type 2 diabetes in urban black South Africans (Look et al. 2006). Another study that investigated genetic variants in serpin peptidase inhibitor (plasminogen activator inhibitor type 1, PAI1) and tissue plasminogen activator (PLAT), important players in the blood coagulation process, reported increased risk of these genetic variants for thrombosis, an important risk factor in heart attacks (Williams et al. 2007). They observed significant genetic variation in PAI1 and PLAT by sex, suggesting a complex pattern of genetic regulation via gene–environment interaction (Schoenhard et al. 2008).

### **2.9.2 Heart Failure (HF)**

Heart failure is major public-health burden in the US population, particularly in individuals over the age of 20 years old. It is estimated that 5.1-million people in the US have been diagnosed and are living with the condition, and about 550,000 new cases are diagnosed each year (Dunlay 2014). In the US, the incidence and mortality rates of HF is highest in the AA population, followed by Hispanics and EAs, and then Chinese. Heart failure in AA patients occurs at an early age and has a worse prognosis compared to their EA counterparts, as well as having higher rates of hospitalization and advanced left-ventricular disease. Several risk factors are associated with this condition. For instance, excessive production of neurohormones such as vasopressin or norepinephrine, as well as nitric-oxide insufficiency, is a major risk factor for HF.

The good news is that there are current drugs on the market, such as mineralocorticoid blockers (antagonist beta blockers; renin angiotensin antagonists, aldosterone blockers), or, in the case of nitric oxide insufficiency, nitric oxide-enhancement can be orally administered to counteract the effect of neurohormonal excesses (Bansal 2009). Other risks factors that are associated with HF include oxidative stress and chronic inflammation, and there are significant racial differences in the levels of oxidative stress and inflammation in AA, EA, and Mexican populations. Although the mechanisms are unclear, diet, comorbidities (obesity, T2DM), lower blood levels of antioxidant nutrients, as well as higher levels of environmental stress, are all implicated in the increase risk of HF.

### 2.9.3 Hypertension

The prevalence of hypertension is increasing among the US population as a whole, but the AA population is disproportionately affected by this condition (Cooper et al. 1997). For example, AAs have a 1.5–2-fold increased risk of hypertension compared with European-descended Americans, with the largest difference being between AA and EA women (Eberhardt et al. 2001). This disparity is even more pronounced among women—African-American women have an age-adjusted incidence of 35.9% (based on data from 1988 to 1994) compared with 19.7% in female Americans of European descent. The reasons for the disparity in this disease is unknown, but several theories have been proposed, including a difference in genetic predisposition factors including differential genetic variants in susceptible disease genes, as well as differences in exposures to environmental risk factors (SES and diet).

Some of the associated risk factors include end-stage renal disease, heart failure, and stroke, risk factors that occur at higher frequencies in the AA population (Collins and Winkleby 2002; Hertz et al. 2005). Hypertension in AAs is potentially associated with high salt sensitivity, low levels of plasma renin, vascular dysfunction (e.g., vasoconstrictor hyperactivity and diminished vasodilatory function), as well higher frequencies of co-morbidities, such as diabetes mellitus and obesity. Other risk factors that are associated with hypertension are sedentary lifestyle, unhealthy diet, and family history of the condition.

Elucidating the genetic risk factors for hypertension has been a difficult task because of other co-morbidities that are associated with the disease. Studies of risk factors have focused on the molecular mechanisms involved in blood-pressure regulation, however identifying the candidate genes associated with this condition has met with little success. For instance, one report has identified a positive association of genetic variants in angiotensinogen (AGT which is involved in the renin-angiotensin pathway) that leads to vasoconstriction and increased blood pressure) with increased risk of hypertension. However, this result has not been replicated in other studies (Caulfield et al. 1995; Rotimi et al. 1994). The reasons for the conflicting observations of genetic studies in various populations are differences in the role of gene–gene and gene–environmental interactions in determining the hypertension phenotype (Kardia 2000; Moore and Williams 2002). A potential role for gene–gene interactions is supported by one study of the Ghanaian population that reported no evidence of single gene association with hypertension risk when analyzed independently. On the other hand, when multiple genes were analyzed together, they observed evidence of association with hypertension risk (Williams et al. 2000).

Despite the high prevalence of hypertension among AAs, hypertension was not considered a major health risk for individuals in sub-Saharan African until recently (Unwin et al. 2001). Studies of hypertension in both rural and urban settings in Nigeria and Cameroon estimate the incidence of hypertension at 15–20% (Cooper et al. 1997), and as high as 20–33% in Tanzanian population (Edwards et al. 2000).

In both studies, urban subjects and women had the highest prevalence, suggesting an increased incidence of hypertension as individuals convert from rural to urban living. Thus, in addition to the prevalence of infectious diseases, several complex diseases including obesity (discussed in Sect. 2.10) and associated hypertension are also becoming more common in Africa, presumably due to an increasingly urbanized Western lifestyle (Colilla et al. 2000; Morris et al. 2010).

A strong genetic component has been suggested for hypertension risk. Some reports indicate a genetic component in the etiology of hypertension with heritability that ranges 45–68% in people of sub-Saharan African descent (Gu et al. 1998; Rotimi et al. 1999a). Genetic variants in genes, e.g., angiotension-1 converting enzyme (ACE) and angiotensin II (AGTII) involved in the renin-angiotensin pathways, are highly heritable. However, there are differences in frequency of the variation in these traits in sub-Saharan Africans and AAs, as one report indicated heritability of 77% for AGT and 67% of ACE in Nigerians when compared to 18% heritability for AGT and ACE in AAs (Cooper 2000). The differences in the variants could be attributed to the admixture population of US AAs compared to the relatively homogenous Nigerian population, as well as different environmental lifestyle and dietary factors in the two different geographic populations. This observation supports a role for environmental variability across various populations that could influence disease risks and outcome.

Differential genetic variations that exist in various racial and ethnic populations can cause different responses to the same environmental risk as a result of differences in gene expression and phenotypic response. Thus, the genetic risk factors associated with hypertension underscore the difficulty associated with identifying genetic risk factors for complex diseases as allelic variants at several multiple genetic loci can independently, or as multivariant clusters, can interact to predispose individuals to the disease risk. This differs in various individuals belonging to the same population or different populations. Future studies in hypertension risks should focus on elucidating the underlying genetic architectures for various different populations.

## 2.10 Diabetes

Diabetes has become a global crisis, i.e., a worldwide epidemic, that requires both effective tools and therapeutic interventions to improve outcomes. Diabetes mellitus (DM) is a condition characterized by either the total lack of insulin (type 1) or the resistance of peripheral tissues to the effects of insulin (type 2, T2DM). Both diseases result from the absence of the signaling effect of insulin in the presence of normal or high glucagon and other metabolic signals. The disease DM is due to the imbalance in carbohydrate metabolism and its effects on other metabolic pathways. An estimated 246 million people world are affected by the T2DM condition, and this number is expected to increase to 380 million by 2025 ([www.bio-medicine.org](http://www.bio-medicine.org)), unless some drastic measures are taken to reduce the disease risk.

According to the American Diabetes Association, 29.1 million Americans were affected with diabetes in 2012 (<http://www.diabetes.org/diabetes-basics/statistics/>). The prevalence of diabetes is highest in the Native American population, followed by AAs, and the lowest incidence is among Alaskan Americans. Compared to the EA population, the prevalence of diabetes, especially among adults, was about twice that in the AA population, whereas the African population had the lowest comparative rate of diabetes, until recently when there is increased incidences in various parts of the continent, as mentioned in Sect. 2.1. It is estimated that one in every eight AAs over the age of 20 is living with diabetes (CDC 2008).

Type 1 DM (T1DM) is a condition linked to autoimmune destruction of the pancreatic beta-cells that can cause partial or complete loss of insulin production. Insulin is an important hormone that plays a role in signaling the body to absorb blood glucose, for instance, after the ingestion of food. Thus in the absence of insulin, gluconeogenesis is unrestrained, and there is an increase in blood glucose. However, cells such as muscle and fat cells cannot take up the available blood glucose, even when present at very high levels in the blood, and this can lead to health problems. The only available treatment is the injection of exogenous insulin into the body. However, even with optimal control, the damaging effects of elevated blood glucose eventually lead to medical complications including heart attack, stroke, blindness, kidney failure, and nerve damage.

Type 2 DM (T2DM) is distinct from T1DM and is characterized by a disorder whereby the pancreatic cells do not produce enough insulin. Type 2 DM can also be caused by improper insulin signaling, such as insulin not being able to bind to cell surface receptors and enable glucose to enter cells, creating a resistance to insulin's effects on target tissues and cells. In that case, the body acts as if there is no insulin, even when the hormone is present at sufficient levels. This condition shares similarity with T1DM. Gluconeogenesis is unrestrained, and muscle and fat cells do not take up glucose resulting in similar disease outcomes as seen for T1DM. The physiologic picture of DM is complex and involves multiple organ systems, including the adipose, muscle, pancreatic islets, immunity, central nervous system, and the liver.

The identification and functional characteristics of genetic variants (discussed in Sect. 2.10.1) that either cause or predispose one to diabetes have provided insight to help explain early pancreatic beta-cells dysfunction in children and adolescents (Hattersley et al. 2009). Hannon et al. (2008) demonstrated that healthy AAs adolescents have 14% lower insulin clearances and 63% higher early-stage insulin concentrations, despite having similar insulin sensitivity as do EAs. Insulin clearance is defined as the ratio of the insulin-degradation rate and insulin concentration that can be used to measure the rate of insulin metabolism. Accordingly, AA adolescents are more likely to have a lower insulin degradation and a higher insulin concentration, leading to the abnormalities in insulin metabolism as compared with their EA peers. Genetic variance is a primary contributor to insulin sensitivity, fasting serum insulin, insulin resistance, and hyperinsulinemia in AAs (Hannon et al. 2008).

Obesity, which is discussed in Sect. 2.11, is the major risk factor for T2DM and appears to drive tissue- insulin resistance in part via a gain of ectopic fat deposits in



abdominal regions, the liver, and the pancreas. However, ectopic fat in the pancreas may contribute to beta-cell dysfunction. It is estimated that individuals whose BMI is more than 35 kg/m<sup>2</sup> (obese individuals) have between a 50–80 fold increased risk of diabetes (T2DM) when compared with individuals whose BMI is less than 23 kg/m<sup>2</sup> (Chan et al. 1994). Indeed, the increasing prevalence of obesity is a key factor in increasing T2DM levels around the world, suggesting that measurement of adiposity (BMI, waist circumference or waist/hip ratios) can be a predictor of T2DM risk. Although diabetes is typically diagnosed in individuals with high BMI, the South Asian population with BMI of 22 kg/m<sup>2</sup> have an equivalent prevalence of T2DM as EAs with a BMI of 30 kg/m<sup>2</sup>. One postulated reason for the high susceptibility to diabetes in the South Asian population is that they have a lower storage capacity for fat (Sniderman et al. 2007), whereas the high prevalence of diabetes in the AA population is because of the high obesity prevalence.

Although differences in body weight or waist circumference are strongly associated with insulin resistance and diabetes, the highest rate of developing T2DM among AAs cannot simply be explained by obesity. Ethnic differences in triglycerides (TGs) and glucagon-like peptide-1 (GLP-1) levels have been found to be significantly associated with insulin resistance and T2DM (Seeleang 2011). Some studies have investigated whether intermuscular adipose tissue (IMAT) or the fat tissue within skeletal muscle is associated with an increased risk of developing insulin resistance and T2DM. One study found both male and female AAs have greater IMAT than EAs (Miljkovic-Gacic et al. 2008). Therefore, genetic studies of IMAT may help explain why some people with lean or normal body weight are at a higher risk for the development of hyperglycemia and T2DM.

Several studies have shown that AAs are more insulin resistant and have higher glycated hemoglobin (A1C) than other ethnic groups, either at or before diagnosis (Adams et al. 2008). Accordingly, AA adolescents seem to have the highest rate of propensity to develop T2DM in comparison with other racial or ethnic groups. Obese AAs have higher resting and stimulated insulin concentrations in comparison with their obese EA counterparts. One study reported that the glucagon-like peptide 1 (GLP-1), a potent metabolite that plays a role in lowering blood glucose by regulating insulin secretion, was observed to be expressed at higher concentrations in the AA obese than the EA obese (Velasquez-Mieyer et al. 2003), findings that might be important in the emergence of GLP-1 based diabetic medications and their use across various populations.

### ***2.10.1 Genetic Factors Associated with Type 1 Diabetes***

Genomic analysis such as GWAS to identify genetic susceptibility loci associated with T1DM has indicated over 50 allelic loci ([www.t1dbase.org](http://www.t1dbase.org)). A major locus effect is shown in genetic variants in the HLA complex that could affect the binding interactions with antigens, and their presentation to T-cells has been identified in perhaps one of the largest T1DM studies in the AA population. High allelic



frequencies of unique HLA-DRB1 genetic variants, which confers susceptibility for T1DM, was observed in AAs in comparison to EA (Howson et al. 2013), indicating a genetic basis for the increased risk of T1DM in the AA population.

### ***2.10.2 Genetics Factors Associated with Type 2 Diabetes***

Type 2 DM is currently the most common metabolic disorder in the world and among the many common diseases with a strong genetic component and strong familial aggregation. The risk of diabetes is estimated to be 50% for first-degree siblings and 100% risk for monozygotic twins among the families with history of diabetes (Rich 1990). Family history has been noted to double the risk of diabetes, which equivalent to the risk of obesity and which is also heritable. Furthermore, obesity and family history quadruple the risk of diabetes. Data in support of a strong familial component are marked by differences in T2DM prevalence across various different populations. Geographic regions of high T2DM incidence are found in the Pima and South Sea Island populations, where the incidence exceeds 50% of the population. On the other hand, the incidence of diabetes is reported to be lowest in European regions, reported to be about 5% in the European population. In the US, the incidence of T2DM in the AA population, with the highest incidence, is approaching 20% (Elbein 2007; Haffner 1998).

Complex diseases such as T2DM are influenced by a combination of genetic, lifestyle, and environmental risk factors. Several candidate gene studies has reported two genes, namely PPARG and the beta-cell potassium channel gene KCNJ11 (Das and Elbein 2006), to be associated with T2DM risk. GWAS, meant to identify obesity genetic variants that are associated with T2DM, identified only FTO, a gene whose protein product plays a role in energy metabolism (Do et al. 2008).

Other GWAS to identify susceptibility alleles with increased association with T2DM have identified more than 60 genetic variants in the EA, Asian, AA, and Hispanic populations. The genetic risk in the AA population persists even after adjusting for known environmental risk factors, such as body mass index (BMI), physical activity, and SES. One study by Keaton et al. (2014) showed that AAs carry a greater number of risk alleles at 46 established T2DM risk loci than EA patients. Cumulatively, these variants are strong predictors of diabetes risk in EAs, but poor predictors in AAs.

The prevalence of T2DM is relatively low in sub-Saharan Africans, reported to be lower than 5%. Although the prevalence of T2DM is on the rise across the continent, attributed mainly to changes in dietary and other environmental factors, obesity, which a major risk factor for T2DM, is still relatively uncommon, reflecting the high physical activity levels and a diet low in caloric intake. Obesity rates were reported to be 0% in Togo, between 4.8 and 8.0% in S. Africa, and 10% in Northern Sudan (Motala 2002).

The low incidence of T2DM in sub-Saharan African, an environment where dietary caloric intake is much lower than in the US, suggests that the incidence of T2DM might carry a proportionately greater genetic component. A multi-institutional, international collaboration was set up to create a genetic resource that can be used to find susceptibility genes for T2DM in the ancestral populations of African Americans in a African-American diabetes mellitus (AADM) study (Rotimi et al. 2001). The AADM study had the goal of collecting 800 sibling pairs affected by T2DM cases and 200 unaffected controls from families in both Ghana and Nigeria. One of the studies that utilized the useful AADM resource was a GWAS linkage analysis carried out to find genetic variants associated with three obesity-related phenotypes: body mass index (BMI); fat mass (FM); and percent of body fat (PBF) (Chen et al. 2005). Among the W. African population, PBF showed the strongest evidence of association with genetic variants chromosomes 2, 4, and 5. The FM showed evidence of association with genetic variants on chromosome 2, and BMI showed association with variants on chromosome 1 and 4. When all three phenotypes were combined, significant genetic variants on chromosomal regions 2p13, 4q23 and 5q14 were associated with obesity in the AADM patient population, providing genetic information of obesity in this less well-studied population.

Other GWAS to find susceptible chromosomal loci in AADM have focused on the downstream consequence of diabetes. Reduced renal function is associated with diabetes and hypertension, and renal function can be detected by measuring serum creatinine, creatinine clearance, and the glomerular filtration rate (GFR) (Chen et al. 2007). A GWAS to find susceptible loci in the genes encoding for serum creatinine, creatinine clearance, and GFR using the AADM resources identified variance in *CYBA*, *NOX1*, and *NOX3* genes that have been implicated in diabetic nephropathy and renal damage. Other studies have identified genetic variants in two positional candidate genes: the pituitary adenylate cyclase activating polypeptide (PACAP) on 18p11 and the peroxisome proliferator-activated receptor gamma coactivator 1 (PPARGC1) on 4p15, for increased diabetes risk in the W. African population. The diversity in genetic variations between different ethnic and racial groups and an association with risk to T2DM complicates drawing clear conclusions about the role of genetic risk factors and T2DM disparities.

## 2.11 Obesity

There is strong epidemiological data to demonstrate obesity is associated with increased risk of not only in T2DM (Eckel et al. 2011) but also cardiovascular diseases (Folsom et al. 1989) and metabolic syndrome, as well as certain cancers.

As previously mentioned, the prevalence of obesity is relatively low in sub-Saharan Africans in comparison with the US AA population. For instance, the incidence of obesity is reported to be 10–13% in AA and Afro-Caribbean population, compared to 1–2% in the W. African Nigerian population (Colilla et al. 2000; Rotimi et al. 1995, 2001). This is in spite of the high heritability levels of

obesity-related traits in many African-descended populations (Luke et al. 2001); thus, the difference could be a reflection of various environmental influences, such as a high-caloric Westernized diet and diminished physical activity.

However, over the past couple of decades, there is a rise in the incidence of obesity and associated diabetes in parts of the African continent. For instance, S. Africa and its neighboring countries, Namibia, Botswana, and Zimbabwe, have a high incidence of obesity similar to the levels to AAs, and this is correlated with an increase in socioeconomic status, urbanization, and adoption of a sedentary lifestyle (Walker et al. 2001). Similar observations of an increased incidence of obesity have been reported in the Tanzanian population, as there is a shift from a rural lifestyle to an urban environment (Aspray et al. 2000).

### ***2.11.1 Genetic Factors Associated with Obesity***

The genetic factors associated with obesity are not well characterized. One study has reported an association of genetic variants in a fat metabolism gene and obesity risk. Genetic-splice variants in exons 3 and 6 of the human uncoupling protein 3 (UCP3) gene have been shown to be associated with reduced fat oxidation and obesity in a group of African descendants who lived in the Gullah community of South Carolina in the US, with very low European admixture (Parra et al. 2001). Interestingly, similar splice variants were observed in an African population from Sierra Leon (W. Africa), whereas these variants were absent in the EA population. In addition, individuals who are heterozygote for exon 6 have basal fat oxidation rates reduced by 50%. The reduced fat oxidation would play a role in increased fat storage, suggesting that genetic variants in UCP3 could play a role in increased susceptibility to obesity by promoting fat storage.

In 2007, a GWAS analysis by Frayling et al. (2007) identified a common variant in the FTO (fat mass and obesity-associated protein) gene that is also associated with increased BMI, from childhood to adulthood across many generations. Because obesity is measured clinically with the surrogate measure of BMI, they reasoned that this common variant in the FTO gene could influence the development of obesity and predisposition to diabetes. In support of this hypothesis, they observed that 16% of the EA populations in their analysis are homozygous for high-risk variants of the FTO gene, and individuals with the homozygous trait weighed about three kilograms more and were 1.67 times more likely to be obese than individuals without this trait. The genetic variants in the FTO gene were also associated with increased T2DM (Frayling et al. 2007).

The FTO gene encodes for the nucleic acid demethylase, and the protein is highly expressed in the hypothalamus of the brain, where food intake is regulated, suggesting a role for FTO in energy homeostasis and body-weight regulation (Loos 2014). Ectopic fat leads to organ-specific insulin resistance via a process termed lipotoxicity (Cusi 2010). Individuals who are susceptible to T2DM appear to show a great tendency to accumulate visceral fat for a given BMI. On the other hand,

there are several individuals, in particular females, who, despite attaining very high BMIs (as high as 50–60 kg/m<sup>2</sup>), remain insulin-sensitive, normoglycemic, and normolipemic. Imaging studies show that these individuals have low levels of visceral and ectopic fat but high subcutaneous fat content (Stefan et al. 2008). Several reports have identified the association of genetic variants of the FTO gene and increased incidence of obesity. Studies shows significant differences in allele frequency in African ancestry compared to other populations. However there are no records of the frequency of genetic variants in the African population in their native countries.

Other studies searching for other genes that may be associated with increased risk of obesity in the AA population have led to the identification of a genetic variant in the G protein beta 3-subunit (GNB3), a gene whose protein product enhances signal transduction and play a role in adipogenesis. A high frequency of this genetic variant known as the 825T allele was found in the so-called old ethnicities, such as Australian aborigines, bushmen, pygmies, and all black populations, including AAs (Siffert et al. 1999). Individuals with this allele are prone to obesity. Patients with homozygous mutant allele demonstrated strong association to increased body-mass index, obesity, and the hypertension-related phenotype. Development of obesity in individuals with this allele was found to be also influenced by environmental-lifestyle factors.

Other genes that may directly be associated with obesity and increased risk of diabetes in the AA population have been identified, but their relationship is not well established. For instance, adiponectin, which is implicated in a number of metabolic disorders including T2DM and CVDs, is encoded by the ADIPOQ gene, and genetic variants of the ADIPOQ are associated with reduced plasma adiponectin protein levels (Sandy et al. 2013) in AA patients with diabetes. However, studies have not identified any significant association of ADIPOQ genetic variants with increased risk to T2DM in this population. This puts into question the role of adiponectin in T2DM pathogenesis. Whether low adiponectin levels are truly causal for T2DM in AAs or rather a consequence of T2DM remains to be determined.

## 2.12 Kidney Disease

According to the CDC 2014 factsheet, 10% of adults in the US have chronic kidney disease (CKD), making it the ninth leading cause of death in the USA. The condition is commonly diagnosed in individuals over the age of 50, and the risk factors associated with developing CKD includes diabetes, high blood pressure, CVDs, obesity and family history. ([http://www.cdc.gov/diabetes/pubs/pdf/kidney\\_factsheet](http://www.cdc.gov/diabetes/pubs/pdf/kidney_factsheet)). The incidence of CKD and mortality from end-stage renal disease (ESRD) is higher in AAs than in any other racial and ethnic group, with AAs having a three-fold higher incidence than their EA counterparts even when adjusted for age and gender. There is ongoing debate as to the reasons for the markedly higher frequency of chronic kidney disease (CKD) in AAs versus EAs, because the

condition is confounded by several interrelated risk factors, such as familial predisposition, socioeconomic status (SES), co-morbid conditions (diabetes, hypertension, and high blood pressure), and lifestyle risk factors such as smoking and excessive alcohol consumption. At the cellular level, CKD can be caused by numerous pathological assaults, such as chronic inflammation, oxidative stress, high lipid levels, viral infection, and hypertension, which affect the kidney glomerular and tubule-intestinal cells (Hu et al. 2012).

In search for susceptibility alleles that could be associated with ESRD, a GWAS was carried out on ESRD patients and control cases from both AA and EA populations (Kao 2008). The results of the studies revealed high frequencies of multiple genetic variants near the gene encoding for non-muscle myosin heavy chain type II isoform A (MYH9) on chromosome 22 in the AA cases in comparison with the EA population. The MYH9 gene may play a role in the cytoskeletal integrity of kidney glomerulus-capillary tubules, suggesting that this variant could be useful in screening for ESRD in the AA population.

Another study by Genovese et al. (2010) identified a strong association of two independent genetic variants within the apolipoproteinL1 (APOL1) gene with an increased risk of focal segmental glomerulosclerosis (FSGS) and hypertension-associated end-stage renal disease (ESRD). The APOL1 gene encodes for a serum protein that plays an important role in host defense and cellular homeostasis mechanisms, and alterations in gene expression is associated with other diseases, including African sleeping sickness, atherosclerosis, lipid disorders, obesity, schizophrenia, cancer, and chronic kidney disease (CKD), suggesting that disruption of this gene has important clinical implications.

Genetic variants in APOL1 are associated with a high risk of a spectrum of kidney diseases including FSGS and ESRD. The described variants are very common in individuals of African descent, including the W. African Nigerians and US AAs, but absent in individuals of European descent. The APOL1 variants most strongly associated with nondiabetic CKD were observed in 40% of the Yoruba population from Nigeria in W. African but not in European, Japanese, or Chinese populations. Statistical analysis showed that these variants might be the result of positive selection in Africa, which could be related to the fact that APOL1 is a trypanolytic factor in human serum (Perez-Morga et al. 2005; Vanhamme et al. 2003) that confers resistance to *T. brucei rhodesiense* (the parasite that causes trypanomiasis or sleeping sickness) which is transmitted by the tsetse flies. This protein is very important on the African continent because it confers resistance to sleeping sickness. The APOL1 gene provides another example of a genetic selection (discussed in Sects. 2.4 and 2.5) for resistance to one disease, in this case sleeping sickness, but is associated with increased susceptibility to another disease, CKD.

Only the disease-associated variants of APOL1 are capable of lysing trypanosome *brucei rhodesiense* parasite. Interestingly, there are two variants of the APOL1 gene located on haplotypes that contain signals of positive selection. This suggests that evolutionary APOL1 variants might have critical survival properties in certain parts of the African continent, where tsetse flies infections are rampant, to

protect against infection from this parasite, but that, however, in a different environment and perhaps in combination with some lifestyle choices, may contribute to the high incidence and prevalence of nondiabetic CKD in AAs.

The HIV viral infection is linked to APOL1-associated kidney failure. It is estimated that ~12% of the HIV-negative and >50% of HIV-positive AAs, who are homozygote carriers of the APOL1 allele, may be susceptible to kidney disease (Freedman et al. 2010). However, not every person who is homozygous for the APOL1 variants will develop ESRD, suggesting some other factor, such as gene–gene interaction or gene–environment, may contribute to the disease risk.

### 2.13 Genetic Variants in Dementia (Alzheimer’s)

Dementia and, in particular, Alzheimer disease (AD) afflict approximately 30 million of the aging population worldwide (Brookmeyer et al. 2007), making it a leading cause of morbidity and mortality among the elderly. There are differences in geographical prevalence: AD is very common in Western Europe and US and less prevalent in sub-Saharan Africa. In the US alone, it is estimated that one in nine persons over the age of 65 will be diagnosed with the disease. Late-onset Alzheimer’s disease (LOAD) is the most common cause of dementia, making it a very important public-health issue. African Americans and Hispanics are more than twice as likely to develop LOAD as EAs (Tang et al. 2001). Epidemiological studies have shown that AAs are more likely to be diagnosed with dementia (in general) and Alzheimer’s disease than EAs (Demirovic et al. 2003). Differences in AD prevalence may be caused by genetic variants, environmental factors, or both. African Americans and other minorities often receive delayed or inadequate health care services, especially for dementia, and this may contribute to the disparities in the disease incidence.

The molecular mechanisms underlying AD indicate that senile plaques (SP) and neurofibrillary tangles are key pathological hallmarks of AD, and genetic mutations in beta-amyloid gene (A $\beta$ ), SP, and presenilin (PSEN 1 and 2) have been shown to be associated with increased risk to Alzheimer disease (Reitz and Mayeux 2014). The differences in etiology of LOAD across populations are not well understood. Epidemiological findings have linked cardiovascular diseases (including elevated blood pressure in mid-life, between the ages of 40–60), T2DM, both low- and high-body weight, diabetes, coronary artery disease, and stroke, as well as traumatic brain injury and metabolic syndrome, with increased risk to AD (Reitz and Mayeux 2014). Because the US AA population is disproportionately affected by higher risks for CVD, diabetes, stroke, and many other conditions linked to AD risk, this could potentially explain the high prevalence of AD in the AA population in comparison with other US racial and ethnic groups.

In addition to the A $\beta$ , SP, and PSEN1 and PSEN2 genes, other genes such as FE65, A2 M, SOLR1, and APP genes have been associated with increased risk for AD. So far, the unequivocal risk gene is polymorphisms in the apolipoprotein E

(APOE) gene that is associated with increased risk of AD. The APOE-gene product is involved in diverse cellular functions, including lipid metabolism, as a transport of lipid from one tissue or cell to the other, inflammation, oxidative stress, and aging. Several reports have shown that APOE is associated with hyperlipidemia and hypercholesterolemia, conditions that are linked to increase atherosclerosis, coronary heart disease, strokes, and increased risk of AD. In animal studies, high levels of cholesterol are associated with increased A $\beta$  load (Refolo et al. 2000) and changes in amyloid precursor protein processing. There are three polymorphic alleles of APOE gene, namely  $\epsilon$ 2,  $\epsilon$ 3, and  $\epsilon$ 4 alleles that are associated with either increased susceptibility to, or protection from, AD in various populations (Liu et al. 2013). Various variants of the APOE gene are associated with either susceptibility or increased risk to AD. Individuals with the  $\epsilon$ 4 allele variants are at significantly higher risk of AD in comparison to individuals with the  $\epsilon$ 3 allele (neutral effect on AD risk), whereas the  $\epsilon$ 2 variant is associated with decreased risk of AD (Graff-Radford et al. 2002). Individuals, who are heterozygote for the APOE  $\epsilon$ 4 allele, are at two–three times higher risk of AD, and those homozygous for the allele are at 1–30 times higher risk. On the other hand, individuals homozygous for the APOE  $\epsilon$ 2 have a 25% reduction in the risk of AD. The APOE  $\epsilon$ 4 allele is well-established to be associated with increased risk to LOAD, as well as early onset of the disease in EA population (Graff-Radford et al. 2002). So far, the major findings have included apolipoprotein E e4 as a major risk factor for AD in most populations including AA, but the APOE e4 variant is not a risk factor for AD in some African populations such as the Yoruba people of Nigeria (Hendrie et al. 2006), suggesting that the differences in AD risk in Yorubas and AAs could be attributed to differences in either gene–gene or gene–environmental interactions.

Few epidemiological studies have been conducted in sub-Saharan African populations to look at the incidence of AD. One such study known as the Indianapolis-Ibadan Dementia Project (Hendrie et al. 2006) reported lower incidence of AD among the W. African Yoruba population in comparison with AAs. The AD incidence rates for AD in the Yoruba population was about half that of AAs for the same age group. With regards to genetic studies in the African diversity population, it is generally held that genetic analysis of complex diseases including AD may confer certain selective advantages. Another variant of APOE, called APOE e1Y, has been identified in the Yoruba population without AD, suggesting this could be the ancestral haplotype (Murrell et al. 2006), however this is the only report on this APOE variant in the Yoruba cohort.

Similar to the Yoruba study, another study from a Ghanaian cohort found that the allelic frequency of APOE e4 was lower compared to EAs and AAs, however there was a significant association with TOMM40\*523 (Translocase of Outer Mitochondria Membrane) alleles. The TOMM40 gene encodes for the principle mitochondrial protein import pore and is therefore critically important for mitochondrial biogenesis and function (Chang et al. 2005), and alteration in gene expression has been linked to neurodegenerative and neuropsychiatric diseases. This observation suggests the presence of a possible cross-over event occurring in the Ghanaians, a distant observation in W. Africans that has been introduced in the



AA population but rarely observed in EA cohorts (Hendrie et al. 2001). This W. African TOMM40'523 allele in linkage analysis to the APOE e4 is observed in the AA population but not in EAs, suggesting that a possible cross-over occurred in the Ghanaian population that has been introduced to the AA population (Roses et al. 2014).

Thus it is possible, for example that gene–gene and gene–environment interactions may be different in Yoruba than AAs, who may have various degrees of admixture with other populations, such as EA and Native Americans. The differences in AD risk between AAs and the nascent African population can also be attributed to different environmental or lifestyle factors that can interact with AD risk genes to cause higher incidence of AD in the AA population than in the African population.

Many of the conditions that are associated with increased risk to AD, such as high cholesterol and heart disease, are not common in the African population in their native countries, allowing for the identification of the role of gene–environmental interaction in the sub-Saharan African and AA populations and the risk to AD (Tishkoff et al. 1996). For instance, AAs have high incidences of diabetes and hypertension and have hypercholesterolemia risk factors for AD, but these conditions are less frequent in the Yoruba population. Perhaps because the Yorubas consume a low-calorie, low-fat diet, consisting mainly of grains, roots, and tubers supplemented by small amounts of fish (Hendrie et al. 2006), and this may explain the low serum cholesterol level in this population. One report indicates that having an African ancestry is highly protective against AD neuropathology (neurite plaques) (Schlesinger et al. 2013). However, caution should be used in generalizing these findings to the entire Nigerian population or, for that matter, the entire African population as only a small cohort of patient samples has been used in this study.

## 2.14 Septicemia-Sepsis

Sepsis or severe inflammatory response is a condition associated with infection of the lungs and urinary tract, or some microbial, bacteria, or viral infection that can arise through surgical procedures, making this condition a major challenge in intensive-care units in US hospital and across the globe. According to a recent JAMA report (2014), up to 50% of all hospital deaths in the USA are linked to sepsis infection (Liu et al. 2014). This condition is more prone in individuals who have a weakened immune system. In particular, infants and elderly are most vulnerable. Sepsis is ranked the tenth leading cause of mortality in the US. Some of the risk factors associated with sepsis include access to health care, gender, and race. The presence of co-morbidity conditions such as chronic obstructive pulmonary disease (COPD), diabetes, HIV/AIDs, ESRD, cancer, and lifestyle choices (such as excessive alcohol consumption) is associated with an increased risk of sepsis. Many of these co-morbidity conditions are observed at higher rates in the AA population than other racial and ethnic groups, and this may explain why AAs also have a



higher prevalence of sepsis than any other US racial or ethnic population. In addition, men are more likely than women to develop sepsis. Interestingly, one report indicates that, while there is an increased prevalence of arterial hypertension in AA patients with sepsis, mortality rates are lower, suggesting a possible protective effect of hypertension in the sepsis patients (Nunes 2003). Thus the high prevalence of hypertension in AAs may have evolved in a different environment as a protective mechanism in a population whose descendants are heavily exposed to infectious diseases.

Studies directed towards the identification of genetic risk factors associated with sepsis, the candidate microorganisms, and their interaction with the host's innate and adaptive immune system and downstream signals have focused on sepsis-associated diseases, such as acute lung injury that can occur in sepsis and other inflammatory disorders (Barnes 2005). Several sepsis-associated candidate-gene variants have been identified, including the IL-1 receptor antagonist (IL1RA), plasminogen activator inhibitor 1 (PAI1), Toll-like receptor1, 4 and 5 (TLR-1, 4, 5), tumor necrosis factor-(TNFA), TNFB, and lymphotaxin (LTA) (Barnes 2005). For instance, Mwantembe et al. (2001) found a high frequency of genetic variants in the interleukin 1 (IL-1) gene cluster associated with South African blacks with inflammatory bowel syndrome disease than in South African whites with the same condition. High TLR1 genetic variants that modulate the risk of bacterial infection were found to be associated with an increased risk of organ failure and a higher mortality rate in both AA and EA patients with sepsis (Wurfel et al. 2008). Plasminogen activator inhibitor (PAI) (Menges et al. 2001) and other cytokines, such as toll-like receptor 4 and tumor necrosis factor, are all associated with sepsis risk.

It is well documented that there are differences in genetic variants in candidate genes underlying inflammatory and immunologic responses to complex diseases such as hypertension, obesity, diabetes, myocardial infarction, Crohn's disease, asthma, and cancer. However, very few studies have looked at the role of ethnicity in the epidemiology of sepsis, therefore less is known about the frequency of the "high-risk" variants and how this differs across racial and ethnic populations. One report indicates polymorphism of IL1RA gene was significantly more frequent among black than white South Africans in a study of inflammatory bowel disease (Mwantembe et al. 2001). Ongoing studies are likely to identify more candidate genes.

## 2.15 Pneumonia and Influenza

Pneumonia and influenza infection were among the leading causes of mortality, particularly in infants at the beginning of the twentieth century. However, with improved immunization, the mortality rates have significantly been reduced, and they are now ranked as the eighth leading cause of death in the US (DeFrances 2008). US minority populations including AAs, Hispanics, and American Indians

have higher prevalence and mortality from pneumonia and influenza infection than the EA population. This is largely attributed to low immunization rates, low SES, which is associated with living in an unhealthy environment, and higher risks of certain co-morbidities such as diabetes, heart disease, and HIV infection (DeFrances 2008). The rise of the invasive drug-resistant *Streptococcus pneumonia* infection (primarily bacteria and meningitis) in the elderly population is a major public-health concern.

There are few studies that have investigated genetic variations in the host's immune response and association with pneumonia and influenza infection. One report has indicated a role for endothelial nitric oxide synthase (eNOS) in the pathophysiology of pneumonia infection, noting increased production of nitric oxide is a host defense mechanism against the bacterial infection to modulate apoptosis and inflammation. Genetic variants in eNOS and other host immune response genes, such as Tumor Necrosis Factor-alpha (TNF-alpha), and Interleukin-10 (IL-10), are reported to be associated with increased incidence of pneumonia in patients with the A/H1N1 influenza infection (Romanova 2013). Other studies have identified genetic variants in FCGR2A and C1QBP to be associated with increased susceptibility to pneumonia in A/H1N1 influenza infection (Zuniga 2012).

Very few studies have been carried out on pneumonia and influenza infection in the African population, except that outbreaks result in high fatality rates. A systematic review of the literature by Cohen (2015) has suggested that influenza infection is rather common on the continent and that co-morbid conditions, such as sickle cell anemia, dengue fever, and measles increase the severity of influenza diseases. Pneumonia is the next most common disease after malaria to be diagnosed in African children, and malnutrition, as well as HIV infection, are associated with pneumonia infection. Infection of *S. pneumonia* or *Haemophilus influenza* type B (HiB) are the common causes of pneumonia in the sub-Saharan African population, and other bacterial infections (*Staphylococcus aureus* and gram-negative bacteria) are known to be contributing factors to pneumonia (Muro 2015).

## 2.16 Cancer

Cancer is a global epidemic accounting for 8.2 million deaths in 2012 (<http://www.iarc.fr/en/publications/books/wcr/wcr-order.php>) and is estimated to reach 21.4 million cases and 13.2 million deaths by 2030. The increased incidence and mortality rates from cancer are attributed to demographic shift, an enlarged aging population, and adoption of unhealthy lifestyles, such as poor diet, smoking, and excessive alcohol. In the USA alone, it is estimated that 1,685,210 people will be diagnosed with cancer in 2016, and this will be accompanied by 595,690 deaths, making cancer the second leading cause of mortality in the USA (<http://www.cancer.gov/about-cancer/what-is-cancer/statistics>). Some of the common cancers diagnosed in the US population includes: breast, lung and bronchus, prostate, colon

and rectal, bladder, skin melanoma, lymphoma, thyroid, kidney and renal pelvis, leukemia, endometrial, and pancreatic cancer.

Recent cancer statistics for 1975–2011 by the North American Association of Central Cancer Registries (NAACCR) organization found that there is a decline in the number of new cases and mortality for several of the major cancer types: breast (women), prostate (men), lung, colon and rectal, throat, and brain cancer diagnosed in the US (Kohler 2015). The reduction parallels decreases in cancer risk factors such as smoking reduction in adults and improved and widespread screening and detection tools, as well as improved cancer-therapy treatment (Byers 2010). The downward trend in the incidence and mortality rates is reported for the US and other Western countries. On the other hand, several developing countries including many Africa countries have been observing upward trends in cancer incidence and mortality rates over the past two to three decades due to the adoption of unhealthy Western lifestyles, including smoking, sedentary living, and consumption of calorie-dense processed food.

Despite the decline in US cancer incidence and mortality rates, there remains persistence in cancer disparities as defined by differences in incidence, prevalence, and mortality rates in the US's racial and ethnic groups. According to the Center for Disease and Prevention Control (CDC) data and statistics on cancer ([www.cdc.gov/cancer/dpcp/data/race.htm](http://www.cdc.gov/cancer/dpcp/data/race.htm)), AA men have the highest cancer incidence, followed by EA men, and the lowest rates in reported for Asian/Pacific Island men. Among women, EAs have the highest incidence, followed by AA women, and the lowest incidence is reported for Asian/Pacific Island women. The statistics on mortality rates places both AA men and women at the top, followed by EAs, with lowest incidence in other groups. Although the incidence and mortality rates are lowest for Asians in comparison with EAs and AAs, they have a high prevalence of cancer caused by infection, such as cervical cancer (women), stomach, liver, and esophageal cancers (McCracken 2007). The following sections will discuss the genetic alterations associated with common cancers in different US racial and ethnic groups.

### **2.16.1 Breast Cancer**

Breast cancer is the most commonly diagnosed cancer in women in the US and the second leading cause of death after lung cancer. In 2016 alone, it is estimated that 246,660 women will be diagnosed with breast cancer, accompanied by 40,450 deaths (<http://seer.cancer.gov/statfacts/html/breast.html>). Breast cancer like prostate cancer (discussed in Sect. 2.16.4) tends to cluster in families, and it is estimated that 10–15% of all breast cancer cases may be due to familial predisposition (Pharoah 2002). Even though the incidence of breast cancer is highest among EA women in the US, AA women are more likely to present with late stages of the disease and die at a significantly high rates than any other racial and ethnic group in the US (Anon 2015). The reasons for the disparity in incidence and mortality rate is believed to be

a complex combination of socioeconomic status, which is reflected in AA's predominantly residing in low income neighborhoods with less access to mammography screening for breast cancer, and lifestyle exposures to poor diet, tobacco smoke, and alcohol, as well as genetic susceptibility (Russell et al. 2012). Groundbreaking work in breast-cancer research has identified various tumor subtypes among diverse racial and ethnic groups. Large-scale gene expression analysis has identified five subtypes of breast cancers. For instance, young AA women are predominantly diagnosed with basal-like breast tumors that tend to be more aggressive with metastatic potential in the lungs and brain and also to be more resistant to therapy than luminal tumor subtypes (Carey et al. 2006; Sorlie et al. 2001). Furthermore, AA women have a thrice higher risk of being diagnosed with triple-negative disease (ER, progesterone receptor [PR], and human epidermal growth receptor 2 [HER2] negative) than EA women (Carey et al. 2006). Because triple negative breast cancers are characterized by the absence of the necessary signaling receptors, treatment strategies and drugs that target estrogen, progesterone, and HER2 remain a challenge, and this might contribute to the higher mortality rates in the AA women (Chlebowski et al. 2005). Other reports indicate that the disparities associated with breast-cancer mortality may reflect the later stage of diagnosis for AA women than intrinsic biological differences (Adebamowo et al. 2008).

Tremendous progress has been made in understanding the molecular mechanisms and pathways underlying breast cancer, from gene-specific alterations or entire tumor signatures that hold the promise to develop targeted treatments. Mutations in two heritable DNA-repair genes BRCA1 and BRCA2 are commonly associated with breast-cancer risk in various racial and ethnic populations, and high frequencies of BRCA2 and BRCA1 mutations are reported for AA and EA breast-cancer patients, respectively (Malone et al. 2006). Clinical studies have led to the development of screening tests and preventive strategies in breast-cancer patients with BRCA2 and BRCA1 mutations (Maxwell 2012). Women with BRCA1 or BRCA2 mutation have a higher risk for breast and ovarian cancer because there is a general recommendation for individuals belonging to families with a high frequency in this somatic mutations to seek genetic testing and counselling in order to make decisions about surveillance or even surgical approaches to reduce the risk of developing breast cancer (Schwartz et al. 2008). As in Western societies, strong genetic factors contribute to a subset of breast-cancer cases in Africa, including a high frequency of heritable BRCA1 and BRCA2 mutations in African women with breast cancer (Awadelkarim et al. 2007). Even though breast-cancer diagnosis is less prevalent in Africa than Europe, due to late diagnosis and poor survival, similar mortality rates have been reported for breast cancer patients in Africa as in Western societies.

Despite the high risk of breast cancer in individuals with BRCA1 and BRCA2 mutations, these genetic variants only accounts for a fraction of familial breast-cancer cases, thus additional genetic alterations must contribute to breast-cancer susceptibility. Numerous genetic alterations that impact gene expression have been described, including p53 (Rose and Royak-Schaler 2001),

H-ras-1 mutations (Weston and Godbold 1997), and overexpression of cyclin D1 (Joe et al. 2001). Loo and coworkers reported significant differences in gene copy-number variations in triple negative tumors from AA and EA women (Loo et al. 2011), and others have demonstrated differences in gene expression patterns between AA and EA breast-cancer patients in pathways related to tumor angiogenesis and chemotaxis (Martin et al. 2009). More than 67 breast-cancer susceptibility loci have been identified through GWAS on various chromosomal regions and with various provenances of breast cancer susceptibility in various racial and ethnic groups; most of these studies were carried out in Europe and in the US (Maxwell 2013).

Many of these genetic studies of breast-cancer risk have included very few samples of AA patients. Recent GWAS that analyzed breast tumors from the AA population have discovered a high frequency of variants in vitamin D biosynthetic intermediates associated with increased breast-cancer risk in AAs in comparison with the EA population (Yao 2012), and genetic variants in hormone-signaling pathways show greater susceptibility loci for increased breast cancer in AAs than EAs (Haddad 2015). The plethora of genetic variants being identified suggests that comprehensive large-population studies are necessary to identify important genetic alterations that could predict the increased breast-cancer risks in the AA population.

### 2.16.2 *Colorectal Cancer*

Colorectal cancer is the third most-commonly diagnosed cancer in men and women in the US. The US cancer statistics on colorectal cancers indicate that, in 2016, 95,270 new colon cancer cases and 39,220 new rectal cancers will be diagnosed, and this is accompanied by 49,190 deaths (<http://www.cancer.org/cancer/colonandrectumcancer/detailedguide/colorectal-cancer-risk-factors>). Some of the risk factors associated with colorectal cancers include: an unhealthy diet that is deficient in essential nutrients; tobacco smoke; excessive alcohol consumption, and a sedentary lifestyle. The disease is also commonly diagnosed in adults over the age of 50 and having a family member, such as a first-degree relative, such as parent, sibling, or child, associated with increased risk of colorectal cancer, reflecting heterogeneity in disease risk from genetic component and/or shared environmental factors.

Hereditary predisposition to certain genetic mutations can cause a family cancer syndrome that contributes to 5–10% of colorectal-cancer incidence. The most common inherited syndromes are familial adenomatous polyposis (FAP) and Lynch syndrome (hereditary non-polyposis colorectal cancer or HNPCC) or other colon inflammations, such as ulcerative colitis and Crohn's disease, which increase colorectal-cancer risks. That up to 20% of CC occurs in individuals under 50 years of age appears to indicate a strong familial-predisposition component (hereditary nonpolytopic colorectal cancer, i.e. HNPCC, and familial adenomatous polyposis, or FAP mutations). Studies in the native African population have reported a familial

predisposition to colorectal cancer. One report identified HNPCC or Lynch syndrome in a small sample population of colorectal cancer patients in Nigeria (Adebamowo et al. 2000). Other reports indicate that a single founder mutation (g.1528C>T in the hMLH1 gene) was significantly associated with colorectal cancer in the Nama group of the far Northern Cape province of South Africa (Anderson et al. 2007). The hMLH1 gene encodes for mismatch repair and is frequently mutated in HNPCC. One report indicated mutations in glutathione S-transferase detoxifying genes: GSTM1 and GSTT1 in patients with an hMLH1 mutation and individuals with both null mutations having a three-fold increased risk of colorectal at an early age (Felix et al. 2006).

The African-American population has a significantly higher risk of colorectal cancer in comparison to the EA population. In addition, when the colorectal cancer occurs at a younger age in AAs, it turns out to be more aggressive. A certain proportion of disparity in the colorectal-cancer mortality rate can be explained by differences in screening rates and differences in stage-specific survival (Lansdorp-Vogelaar et al. 2012). There are also differences in anatomic site of disease presentation (Dimou et al. 2009): AAs are more likely to develop cancer in the colon, whereas EAs are more likely to develop cancer in the rectum, which is lower in the bowel and more easily detected by screening (Matanoski et al. 2006). Identification of precancerous lesions is more difficult for cancers of the colon (Matanoski et al. 2006). As a result, AAs are at increased risk of developing advanced colorectal cancer than EAs.

Evidence from genomic studies has identified widespread chromosomal instability with candidate genes including oncogenic *KRAS*, *APC*, and *DCC/MADH2/MADH4*, as well as tumor-suppressor *PTEN* and *p53* genes (Vogelstein and Kinzler 1993). Like most genetic analysis, a lot of the reported studies in colorectal cancer were conducted using EA samples and included few or no samples of AA patients. GWAS for genetic variants and association with colorectal cancer identified two major variants, Y179C and G396D (formerly known as Y165C and G382D) in the *MYH* gene. The *MYH* gene encodes for DNA glycosylase, which plays a role in DNA repair by removing mispairing of 8-hydroxyguanine caused by oxidative damage during DNA synthesis. These variants were predominant in EA patients, where 90% of the patient population carried at least one of these variants; however, these variants were not found in the AA population in this analysis (Nielsen et al. 2009).

Obesity is a risk factor for colorectal cancer, and insulin resistance is commonly associated with obesity. One report that investigated genetic alterations in the insulin-signaling pathway and colorectal cancer identified an association of genetic polymorphism in insulin-signaling pathway and increased risk of colorectal cancer in the EA population, but not in the AA population (Keku et al. 2012).

Several scientific reports have suggested an important role for vitamin D in physiological processes, such as calcium homeostasis and insulin secretion, as well as cellular immunity; and they also suggest that deficiency in vitamin D levels are linked to conditions such as multiple sclerosis, cardiovascular disease, infectious disease, and cancer including colorectal cancer. Thus, maintaining an adequate

level of vitamin D may have a protective effect against colorectal cancer. However, the AA population has lower serum levels of vitamin D than the EA population. One study by Pibiri et al. (2014) reported that genetic variants in genes involved in the biosynthetic pathway of vitamin D could modulate the serum level of vitamin D, noting significant association of genetic variants in 25-hydroxylase gene (CYP2R1) and increased risk of colorectal cancer in the AA population. The molecular mechanism whereby genetic variants in CYP2R1 could affect colorectal cancer risk is unknown; however, this observation indicates that genetic variants in the vitamin D biosynthetic pathway could influence susceptibility of the AA population to colorectal cancer. Other investigations have focused on the role of intestinal chronic inflammation and mediators, such as mannose-binding lectin 2 (MBL2). Genetic variation in the 3'-untranslated region of the MBL2 gene has been linked to increased susceptibility of colon cancer in AAs but not in EAs (Zanetti et al. 2012). Overall, it appears that various genetic alterations may contribute to the differential risk to colorectal cancers in the AA and EA populations.

### 2.16.3 Lung Cancer

Lung cancer is the leading cause of death worldwide across all income levels in both Westernized and developing countries, accounting for 1.6 million deaths per year (about 20% of all cancer mortality rates) (Ferlay 2015). According to the American Cancer Society statistics on cancers, lung cancer is the second most common cancer in both men and women, and, in 2016 alone, 224,390 new cases were diagnosed in both men and women combined, accompanied by 158,080 deaths ([www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/non-small-cell-lung-cancer-key-statistics](http://www.cancer.org/cancer/lungcancer-non-smallcell/detailedguide/non-small-cell-lung-cancer-key-statistics)). It is estimated that more people die from lung cancer than colon, breast, and prostate cancers combined. The prevalence of lung cancer is about 20% higher in AA men than EA men, whereas the prevalence is about 10% less in AA women than EA women.

Lung cancer has a strong environmental risk factor associated with the disease risk. Tobacco smoke is significantly associated with lung cancer. Successful public-health campaigns about the dangers of tobacco smoke are largely responsible for the decline in lung-cancer-related mortality in the US (Kohler 2015). Unfortunately, this is not true for other places such as China and India where tobacco smoking is on the increase. Other environmental risk factors include exposure to residential radon and asbestos, as well as radiation in the workplace. One study indicates that racial residential segregation and targeted tobacco advertising in such communities may contribute to the 10% higher lung-cancer mortality rate compared to heterogeneous neighborhoods, and this difference persists even after adjustment for socioeconomic status (SES) (Hayanga and D'Cunha 2013).

Extensive genomic analysis has revealed numerous chromosomal alterations and genomic instabilities and a high frequency of mutations in tumor suppressor p53, VHL, and Rb genes and known oncogenes including EGFR, KRAS, ALK, MYC,



and Her-2/NEU (Carbone and Minna 1992; Singh and Kathiresan 2014) with important prognostic significance. One report by Bauml et al. (2013) indicates a lower frequency of EGFR mutation in AA non-small-cell lung-cancer patients than in EA patients. Genetic variations in CYP1A1 and CYP2A6, which belongs to cytochrome P450 family of enzymes that functions in carcinogen detoxification, have been shown to be associated with lung-cancer risk. One report suggests high frequencies of these variants in AA lung-cancer patients (Wassenaar et al. 2015), whereas another report did not find a high prevalence of CYP1A1 and CYP2A6 genetic variance in AA patients (Shields et al. 1993). So, conclusive and well-replicated data are required before any association of CYP1A1 risk alleles in AA lung-cancer patients can be made.

#### **2.16.4 Prostate Cancer**

Prostate cancer is the most commonly diagnosed cancer among US men. The risk factors associated with the disease includes age, as most men diagnosed with the disease are over the age of 60 year. A family history of prostate disease, such as having a brother or father with the disease, increases the risk. Race is also an important risk factor as prostate cancer incidence and mortality rates are about twice as high in the AAs as in the EA population, suggesting a genetic component to the disease susceptibility. In support of a genetic basis to prostate cancer risk, men of African ancestry in various geographic regions of the world, including the Caribbean and S. American, share similar incidences and mortality rates comparable to AAs (Brawley 1998; Delongchamps et al. 2007). Similarly, data from sub-Saharan Africa estimates prostate cancer to be the third most-commonly diagnosed cancer in males, and incidence rates are increasing rapidly with urbanization and the adoption of a Western lifestyle (Magotha 2007; Parkin et al. 2003).

Prostate cancer tends to cluster within families, and there are increased rates of the disease in monozygotic twin pairs in comparison with dizygotic twin pairs, consistent with a strong genetic influence on disease risk (Johns 2001). Tremendous numbers of studies have been conducted over the past 30 years to identify the genetic basis of prostate cancer, and, as a result, multiple candidate genes that play diverse role in disease etiology and progress have been identified, including those involved in susceptibility to oxidative DNA damage, growth related pathways, androgen receptor signaling, chronic inflammatory response, and RNA processing (Rennert et al. 2005; Shand and Gelmann 2006; Zabaleta et al. 2008). Several of these candidate genes have been studied with regards to their allelic frequencies and prostate cancer risk in Africans, AA, and EA populations (Esteban et al. 2006; Kittles et al. 2001; Zeigler-Johnson et al. 2002).

Family-based linkage studies have been carried out to search for high penetrance genes, however these have met very little success. The few successful findings include rare variants in the prostate specific transcription factor HOXB13, found to be associated with increased risk of early-onset hereditary prostate cancer through



an unknown mechanism, although variants accounts for a small proportion of all prostate cancers (Ewing 2012).

A GWAS comparing men with prostate cancer and those without the disease in a case-control study has identified genetic variants in several genes including MYC, NKX3,1, MSMB, and PSA to be associated with prostate-cancer risk (Isaacs 2012). Other efforts have focused on identifying genetic variants associated with prostate-cancer aggressiveness using the Gleason score as an indicator for PCa aggression and have identified variants in chromosomes 5q31-32, 7q32, and 19q12-13 as harboring aggressive prostate cancer loci (Isaacs 2012).

Despite the high prevalence of prostate cancer in the AA population, many of these genetic studies have not included this population, even though AAs present very aggressive and late stages of disease. It is important to carry out these studies in AA populations because the current data shows that indolent and aggressive prostate cancer has totally different genomic patterns or signatures, and, thus, it is necessary to know the genomic signatures of prostate cancer in the AA population to understand how their ancestry contributes to disease prevalence and to distinguish the indolent from the aggressive disease phenotype. One important candidate gene that could be associated with aggressive prostate cancer is BRCA2 gene, a well-known DNA-repair gene whose mutation is associated with breast cancer. Germline mutation in BRCA2 gene is not only a risk factor for prostate cancer, but more likely to be associated with high-grade and non-organ confined disease (Edwards 2010).

GWAS and other cutting-edge analysis, including comparative genomic hybridization (CGH) arrays in prostate cancer studies using biospecimens from multiple racial and ethnic group studies, have identified increased risks of prostate cancer with high frequencies of chromosomal losses on 6q13-22, 8p21, 13q13-14, and 16q11-24 and gains of 7p21 and 8q24 (Castro et al. 2009). Several SNPs on 8q24 with association to PCa risk are perhaps the most significant findings of risk alleles and prostate cancer. Furthermore, seven of these SNPs on 8q24 have been documented by several groups to be associated with significantly high risk of PCa in the AA population (Xu et al. 2009). However, these significant variants do not map to known genes, and the biological mechanisms underlying this association are not known. Gene rearrangement is now the center stage in PCa genetic research. Recent studies have identified chromosomal fusion of TMPRSS2-ERG that results in androgen- regulated over-expression of ERG protein to be associated with PCa progression (Petrovics et al. 2005). The TMPRSS2-ERG fusion is found in more than 50% of EA men and in less than 30% of AA men with prostate cancer (Magi-Galluzzi et al. 2011; Petrovics et al. 2005), indicating a lower frequency of the ERG-positive phenotype in AA men with aggressive prostate cancer.

Other studies have focused on the genetic variation of genes involved androgen signaling and metabolic pathways in regards to prostate-cancer risks in the AA population. Allelic frequencies of several genes in the androgen-signaling pathways have been reported to differ in AA and EA populations and could potentially contribute to the differential risk to prostate cancer (Platz and Giovannucci 2004). One study investigated genetic variants in CYP3A4-V and association with increased risk

of prostate cancer in African (Nigerian) and AA and EA populations (Hainaut and Boyle 2008; Kittles et al. 2002). The CYP3A4 gene encodes for a protein which belongs to cytochrome P450 superfamily and plays a role in oxidative deactivation of testosterone (androgen signaling). The variant in CYP3A4-V that is an A–G promoter variant decreases CYP3A4 protein activity and thereby increases the availability of testosterone required for prostate growth. Kittles et al. (Hainaut and Boyle 2008; Kittles et al. 2002) found the highest frequency of this CYP3A4-V variant in Nigerian men, an intermediate frequency in AA, and low frequency in EA men. The high frequency of this variant was associated with prostate risk in AA in comparison with the EA counterpart. Interestingly, there was no association of this variant with prostate-cancer risk in the Nigerian population, although the variant was found at high frequency in this population. Because the AA population is genetically heterogeneous, with African ancestry and the later admixture with EA, some findings may be susceptible to spurious association, thus additional studies, such as the use of autosomal informatics markers (AIMs), are required to ensure that genetic-variant associations with disease risks are causal, and not just confounding, because of population stratification. Genetic variants in CYP3A4 has been reported in other cancers: lung, breast, kidney, and leukocytes (Lamba 2002), suggesting that genetic–environmental interactions as related to drug inactivation may play a role in cancer susceptibility.

Other studies that have investigated gene-expression profiles of tumors obtained by microarray technology from AA and EA patients point to prominent differences in primary prostate-cancer immunology (Wallace et al. 2008). Results from these studies demonstrated consistent differences in immunobiological gene-expression patterns in tumors from the two patient groups. An analysis of 69 micro-dissected tumors from 33 AA and 36 EA patients that were matched for clinic-pathological characteristics (patient's age, Gleason score, disease stage, PSA at diagnosis, and largest tumor nodule) identified numerous differences in the expression of genes related to immunobiology. These genes clustered in pathways related to immune response, host defense, cytokine signaling and chemotaxis, and inflammation. Most of the immune-related genes were expressed more frequently in tumors from AA patients than EA patients. Previously, low-grade chronic inflammation in the noncancerous prostate gland is observed to be more prevalent in AA men (Eastham et al. 1998). Thus, it is possible that the observed gene-expression differences in the tumors are partly due to a low-grade chronic inflammation that is more prevalent in AA tumors.

Because prostate cancer has a very strong genetic component to the disease, it is important to carry out genetic analysis in the various racial and ethnic US populations in order to decipher the genetic basis of the disease. Overall, prostate cancer displays a heterogeneous array of genetic alterations in individuals belonging to one population or different populations, as well as in the prostate tumor itself, making it very difficult to identify predominant pathways that are altered in the disease pathway for efficacious therapeutic intervention.

In addition to genetic factors, the development of PCa is also influenced by environmental factors such as diet. The disease prevalence as evidence by second

and third generations of Chinese and Japanese migrants to the US, who have adopted the Western dietary lifestyle, have similar prostate cancer incidences to EA counterparts, whereas Chinese and Japanese living in their native countries have lower disease incidence and mortality rates (Brawley 1998). Other contributing factors includes access to screening and SES.

### 2.16.5 Other Cancers

Cancers that are associated with infection are more prevalent in sub-Saharan Africa than in the Western hemisphere. For instance, hepatocellular carcinoma is the second most common cancer in men from sub-Saharan Africa (Parkin 2003), but this type of cancer is ranked as only the fifth leading cause of cancer among US men. Hepatocellular cancer has a strong environmental component, such as the early infection with hepatitis B and C that interact with dietary exposure to aflatoxins from *Aspergillus* molds that is common in Africa (Hainaut and Boyle 2008). Bladder cancer occurs with high frequency particularly in N. Africa. Susceptibility to bladder cancer is modulated by variations in genes involved in pathways, such as metabolic detoxification, redox cycling, free radical injury, and metabolism of folate and methionine which are critical for DNA synthesis/repair and methylation (Ouerhani et al. 2007). Other studies have found some association of bladder cancer with GSTT1 and GSTM1 in Egypt (Saad et al. 2005) and Sudan (Ouerhani et al. 2006). Almost nothing is known about the influence of genetic factors on bladder cancer in sub-Saharan Africa.

## 2.17 Summary

Health disparities experienced by the African diaspora in the US is well documented. Whereas Africans in sub-Saharan Africa are plagued by infectious diseases, non-communicable diseases impose a huge burden on the US population and especially individuals of African descent, as well as other minority groups. The completion of the human genome sequence has made it possible for new genomic approaches, such as GWAS and next-generation sequencing, to unravel genomic variations and their association with increased risk to diseases such as CVDs, hypertension, obesity, diabetes, malaria, and cancer in various populations.

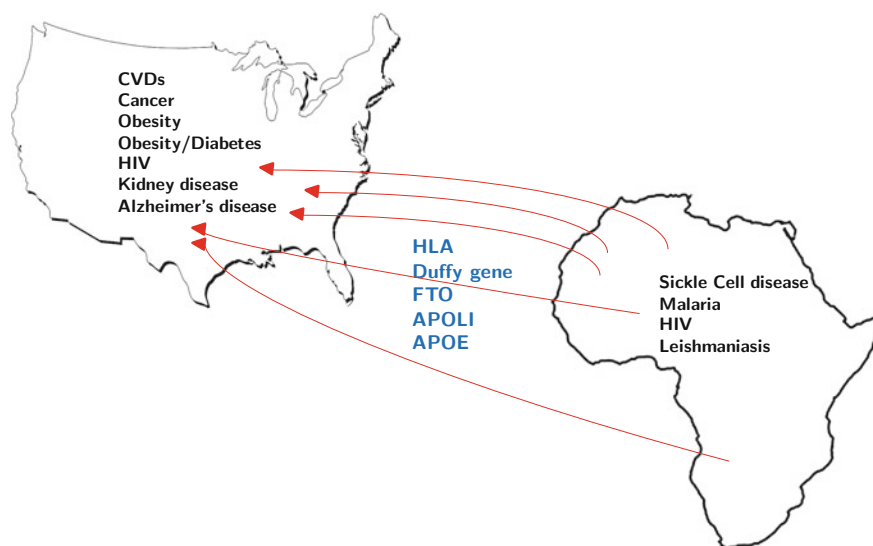
The people of continental Africa represent the oldest human populations and variations in their genetic sequence reflect influences of diet and climate and exposure to infectious pathogens that have been naturally selected to adapt to their environment. Africa is also the origin of most the ancestors of AAs, who left the continent about 400 years ago. However, while the native African population is under-represented in studies in human genetics, it is important to know the genetic variation in the African population in order to understand their contribution to the

health disparities of the African diaspora. On the other hand, the AA population represents a gene pool unique from native Africans because they have been subjected to a variety of environmental and racial influences, and have, on average, approximately 20% European ancestry (Shriver and Kittles 2004). We can predict that just as many European Americans are descendants of many different ethnic groups within Europe, so AAs are heterogeneous descendants from many different populations within Africa (and Europe).

Most GWAS have been carried out in EA populations, and few studies have been carried out in AA and other populations. Yet the inclusion of diverse populations in such studies is crucial for a more comprehensive appreciation of the role of human variation in complex diseases and for more accurate diagnosis and effective clinical management. GWAS is also a powerful tool to investigate the oldest human population, enabling accurate construction of ancestral haplotypes not found in non-African population. The few GWAS studies in this population have identified some of the genetic variants associated with susceptibility to infectious diseases such as HIV, malaria, and TB and have provided insight into the role of gene–gene and gene–environment interactions in disease susceptibility and resistance, as well as providing insights about how genetic variants have evolve to be protective against one type of disease resistance in one particular environment, but can contribute to an increase risk to another disease in a different environment.

GWAS holds promise to unravel genetic variations and disease risks and has the potential to help design more effective pharmaceutical treatments specifically targeted for various populations. However, there are several challenges to the full utilization of this approach:

- 1 The large population diversity means many populations would need to be studied, and one population cannot easily be used as proxy for another. Efforts to carry out such large-scale genomic research will be expensive, complex, challenging, and time consuming, but there are potential benefits for the large-scale genomic research of the African population.
- 2 Currently, GWAS has to be carried out on non-European populations using arrays designed for European populations that can only capture a proportion of the genomic variation in any non-European population leaving the other proportion of these variants not assayed (Spencer et al. 2009). There are efforts to develop African-diaspora arrays for GWAS. Furthermore, methods to capture gene–environment interactions are not well developed for complex diseases. This is more important for AAs and African ancestry because the AA population have evolved to live in a different environment than the continental African population. There are several natural genes, including HLA, Duffy, FTO, APOLI, and APOE, with variants that gives rise to the resistance to infectious diseases on the African continent, but, in a different environment such as in the US, is associated with increased risk to non-communicable diseases in the African diaspora (Fig. 2.1).



**Fig. 2.1** Schematic presentation of some natural selected genes (*blue*; HLA, Duffy, FTO, APOLI and APOE) with variants that are associated with resistance to infectious diseases on the African continent. In a different environment such as the migration to USA (indicated by *red arrows*), some of these genetic variants are now associated with increased risk to non-communicable diseases among the African diaspora

Overall, some researchers attribute all of the problem of health disparities to differences in genetic predisposition (Burchard et al. 2003; Hirsch et al. 2006), assuming human genetic variation can be differentiated into conventional racial clusters (Calafell 2003; Redon et al. 2006). They suggest that disease-causing alleles are likely among the variants that can be segregated out between different groups (Burchard et al. 2003). In support, genetic studies of population substructures, in which analysis of thousands of genetic loci simultaneously, have produced clusters of genetic information that can be used to correctly to identify individuals of self-described geographic ancestry (Redon et al. 2006).

Other investigators argue that social forces drive racial health disparities, pointing to social determinants such as SES, racial segregation, psychosocial stress, and institutional and interpersonal discrimination as causes of adverse health outcomes in racial and ethnic minority groups, such as the AA population (Troxel et al. 2003; Williams and Jackson 2005). Social, economic, and contextual factors can have significant impact on health and taken into account together can diminish health disparities between AAs and whites. The ongoing debate between these competing models is described as the perfect “storm” (Krieger 2005). Scientific research cannot downplay the contributory role played by non-genetic factors (Rotimi et al. 1999b) or the onset or severity of the disease results from a complex interaction of genetic and environmental factors (Coleman 1999; Sommer et al. 1991).

These opposing forces of genetics and non-genetic causes of health disparities are basically two sides of the same coin. Genetic studies are changing the way we think about human variation, and it is important that this change has a positive impact on medicine and public health, while complete answers to many questions are beyond our grasp, some questions such as why aggregation of variants within population groups are likely to be answered in the near future, whereas other effects such as racism, with social and economic consequences of the ideology, will only be eliminated through political processes.

## References

- Abubakari, A. R., Lauder, W., Agyemang, C., Jones, M., Kirk, A., & Bhopal, R. S. (2008). Prevalence and time trends in obesity among adult West African populations: A meta-analysis. *Obesity Reviews*, 9(4), 297–311. Available from: PM:18179616.
- Adams, A. S., Trinacty, C. M., Zhang, F., Kleinman, K., Grant, R. W., Meigs, J. B., et al. (2008). Medication adherence and racial differences in A1C control. *Diabetes Care*, 31(5), 916–921. Available from: PM:18235050.
- Adebamowo, C. A., Adeyi, O., Pyatt, R., Prior, T. W., Chadwick, R. B., & de la Chapelle, A. (2000). Case report on hereditary non-polyposis colon cancer (HNPCC) in Nigeria. *African Journal of Medicine and Medical Sciences*, 29(1), 71–73. Available from: PM:11379475.
- Adebamowo, C. A., Famooto, A., Ogundiran, T. O., Aniagwu, T., Nkwodimmah, C., & Akang, E. E. (2008). Immunohistochemical and molecular subtypes of breast cancer in Nigeria. *Breast Cancer Research and Treatment*, 110(1):183–188. Available from: PM:17674190.
- Agarwal, A. (2000). Hemoglobin C associated with protection from severe malaria in the Dogon of Mali, a West African population with a low prevalence of hemoglobin S.
- Amundadottir, L. T. (2016). Pancreatic cancer genetics.
- Anderson, D. W., Goldberg, P. A., Algar, U., Felix, R., & Ramesar, R. S. (2007). Mobile colonoscopic surveillance provides quality care for hereditary nonpolyposis colorectal carcinoma families in South Africa. *Colorectal Disease*, 9(6), 509–514. Available from: PM:17477847.
- Anon. (2006a). *Centers for Disease Control and Prevention (CDC): The global HIV/AIDS pandemic*. The Morbidity and Mortality Weekly Report (MMWR) 55:841–844.
- Anon. (2006b). [http://www.afro.who.int/malaria/publications/annual\\_reports/africa\\_malaria\\_report\\_2006.pdf](http://www.afro.who.int/malaria/publications/annual_reports/africa_malaria_report_2006.pdf)
- Anon. (2015). Center for Disease and Prevention Control. [www.cdc.gov/cancer/dcpc/data/race.htm](http://www.cdc.gov/cancer/dcpc/data/race.htm).
- Aspray, T. J., Mugusi, F., Rashid, S., Whiting, D., Edwards, R., Alberti, K. G., et al. (2000). Rural and urban differences in diabetes prevalence in Tanzania: The role of obesity, physical inactivity and urban living. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 94(6), 637–644. Available from: PM:11198647.
- Awadelkarim, K. D., Aceto, G., Veschi, S., Elhaj, A., Morgano, A., Mohamedani, A. A., et al. (2007). BRCA1 and BRCA2 status in a Central Sudanese series of breast cancer patients: interactions with genetic, ethnic and reproductive factors. *Breast Cancer Research and Treatment*, 102(2), 189–199. Available from: PM:17333343.
- Awomoyi, A. A., Charurat, M., Marchant, A., Miller, E. N., Blackwell, J. M., McAdam, K. P., et al. (2005). Polymorphism in IL1B: IL1B-511 association with tuberculosis and decreased lipopolysaccharide-induced IL-1beta in IFN-gamma primed ex-vivo whole blood assay. *Journal of Endotoxin Research*, 11(5), 281–286. Available from: PM:16263000.

- Awomoyi, A. A., Marchant, A., Howson, J. M., McAdam, K. P., Blackwell, J. M., & Newport, M. J. (2002). Interleukin-10, polymorphism in SLC11A1 (formerly NRAMP1), and susceptibility to tuberculosis. *Journal of Infectious Diseases*, 186(12), 1808–1814. Available from: PM:12447767.
- Baghdadi, J. E., Orlova, M., Alter, A., Ranque, B., Chentoufi, M., Lazrak, F., et al. (2006). An autosomal dominant major gene confers predisposition to pulmonary tuberculosis in adults. *Journal of Experimental Medicine*, 203(7), 1679–1684. Available from: PM:16801399.
- Ball, T. B., Ji, H., Kimani, J., McLaren, P., Marlin, C., Hill, A. V., et al. (2007). Polymorphisms in IRF-1 associated with resistance to HIV-1 infection in highly exposed uninfected Kenyan sex workers. *AIDS*, 21(9), 1091–1101. Available from: PM:17502719.
- Bansal, S. (2009). Sodium retention in heart failure and cirrhosis: Potential role of natriuretic doses of mineralocorticoid antagonist?
- Barnes, K. C. (2005). Genetic determinants and ethnic disparities in sepsis-associated acute lung injury. *Proceedings of the American Thoracic Society*, 2(3), 195–201. Available from: PM:16222037.
- Baumli, J., Mick, R., Zhang, Y., Watt, C. D., Vachani, A., Aggarwal, C., et al. (2013). Frequency of EGFR and KRAS mutations in patients with non small cell lung cancer by racial background: Do disparities exist? *Lung Cancer*, 81(3), 347–353. Available from: PM:23806795.
- Bodmer, J. (1996). World distribution of HLA alleles and implications for disease. *Ciba Foundation Symposium*, 197, 233–253. Available from: PM:8827377.
- Borman, L., Campbell, S. J., Fielding, K., Bah, B., Sillah, J., Gustafson, P., Manneh, K., et al. (2004). Vitamin D receptor polymorphisms and susceptibility to tuberculosis in West Africa: A case-control and family study. *Journal of Infectious Diseases*, 190(9), 1631–1641. Available from: PM:15478069.
- Braun, L. (2002). Race, ethnicity, and health: Can genetics explain disparities? *Perspectives in Biology and Medicine*, 45(2), 159–174. Available from: PM:11919376.
- Brawley, O. W. (1998). Prostate cancer and black men. *Seminars in Urology and Oncology*, 16(4), 184–186. Available from: PM:9858323.
- Brookmeyer, R., Johnson, E., Ziegler-Graham, K., & Arrighi, H. M. (2007). Forecasting the global burden of Alzheimer's disease. *Alzheimers & Dementia*, 3(3), 186–191. Available from: PM:19595937.
- Burchard, E. G., Ziv, E., Coyle, N., Gomez, S. L., Tang, H., Karter, A. J., et al. (2003). The importance of race and ethnic background in biomedical research and clinical practice. *The New England Journal of Medicine*, 348(12), 1170–1175. Available from: PM:12646676.
- Byers, T. (2010). Two decades of declining cancer mortality: Progress with disparity.
- Calafell, F. (2003). Classifying humans. *Nature Genetics*, 33(4), 435–436. Available from: PM:12665860.
- Campbell, M. C., & Tishkoff, S. A. (2008). African genetic diversity: Implications for human demographic history, modern human origins, and complex disease mapping. *Annual Review of Genomics and Human Genetics*, 9, 403–433. Available from: PM:18593304.
- Carbone, D. P., & Minna, J. D. (1992). The molecular genetics of lung cancer. *Journal of Advances in Internal Medicine*, 37, 153–171. Available from: PM:1348393.
- Carey, L. A., Perou, C. M., Livasy, C. A., Dressler, L. G., Cowan, D., Conway, K., et al. (2006). Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA*, 295(21), 2492–2502. Available from: PM:16757721.
- Castro, P., Creighton, C. J., Ozen, M., Berel, D., Mims, M. P., & Ittmann, M. (2009). Genomic profiling of prostate cancers from African American men. *Neoplasia*, 11(3), 305–312. Available from: PM:19242612.
- Caulfield, M., Lavender, P., Newell-Price, J., Farrall, M., Kamdar, S., Daniel, H., et al. (1995). Linkage of the angiotensinogen gene locus to human essential hypertension in African Caribbeans. *Journal of Clinical Investigation*, 96(2), 687–692. Available from: PM:7635961.
- Chakravarti, A. (2001). To a future of genetic medicine. *Nature*, 409(6822), 822–823. Available from: PM:11236997.

- Chan, J. M., Rimm, E. B., Colditz, G. A., Stampfer, M. J., & Willett, W. C. (1994). Obesity, fat distribution, and weight gain as risk factors for clinical diabetes in men. *Diabetes Care*, 17(9), 961–969. Available from: PM:7988316.
- Chang, S., ran Ma, T., Miranda, R. D., Balestra, M. E., Mahley, R. W., & Huang, Y. (2005). Lipid- and receptor-binding regions of apolipoprotein E4 fragments act in concert to cause mitochondrial dysfunction and neurotoxicity. *Proceedings of National Academic Science U S A*, 102(51), 18694–18699. Available from: PM:16344479.
- Chen, G., Adeyemo, A. A., Johnson, T., Zhou, J., Amoah, A., Owusu, S., et al. (2005). A genome-wide scan for quantitative trait loci linked to obesity phenotypes among West Africans. *International Journal of Obesity (London)*, 29(3), 255–259. Available from: PM:15611782.
- Chen, G., Adeyemo, A., Zhou, J., Chen, Y., Huang, H., Doumatey, A., et al. (2007). Genome-wide search for susceptibility genes to type 2 diabetes in West Africans: Potential role of C-peptide. *Diabetes Research and Clinical Practice*, 78(3), e1–e6. Available from: PM:17548123.
- Chlebowski, R. T., Chen, Z., Anderson, G. L., Rohan, T., Aragaki, A., Lane, D., et al. (2005). Ethnicity and breast cancer: Factors influencing differences in incidence and outcome. *Journal of the National Cancer Institute*, 97(6), 439–448. Available from: PM:15770008.
- Cohen, A. L., Mc Morrow, M., Walaza, S., Cohen, C., Tempia, S., Alexander-Scott, M., et al. (2015). Potential impact of co-infections and co-morbidities prevalent in Africa on influenza severity and frequency: A systematic review. *PLoS ONE*, 10, e128580.
- Coleman, A. L. (1999). Glaucoma. *Lancet*, 354(9192), 1803–1810. Available from: PM:10577657.
- Colilla, S., Rotimi, C., Cooper, R., Goldberg, J., & Cox, N. (2000). Genetic inheritance of body mass index in African-American and African families. *Genetic Epidemiology*, 18(4), 360–376. Available from: PM:10797595.
- Collins, R., & Winkleby, M. A. (2002). African American women and men at high and low risk for hypertension: A signal detection analysis of NHANES III, 1988–1994. *Preventive Medicine*, 35(4), 303–312. Available from: PM:12453706.
- Conrad, D. F., Jakobsson, M., Coop, G., Wen, X., Wall, J. D., Rosenberg, N. A., et al. (2006). A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nature Genetics*, 38(11), 1251–1260. Available from: PM:17057719.
- Cooke, G. S., & Hill, A. V. (2001). Genetics of susceptibility to human infectious disease. *Nature Reviews Genetics*, 2(12), 967–977. Available from: PM:11733749.
- Cooper, R. S. (2000). Heritability of angiotensin-converting enzyme and angiotensinogen: A comparison of US blacks and Nigerians.
- Cooper, R. S., Kaufman, J. S., & Ward, R. (2003). Race and genomics. *New England Journal of Medicine*, 348(12), 1166–1170. Available from: PM:12646675.
- Cooper, R., Rotimi, C., Ataman, S., McGee, D., Osotimehin, B., Kadiri, S., et al. (1997). The prevalence of hypertension in seven populations of west African origin. *American Journal of Public Health*, 87(2), 160–168. Available from: PM:9103091.
- Cusi, K. (2010). The role of adipose tissue and lipotoxicity in the pathogenesis of type 2 diabetes. *Current Diabetes Report*, 10(4), 306–315. Available from: PM:20556549.
- Das, S. K., & Elbein, S. C. (2006). The genetic basis of type 2 diabetes. *Cellscience*, 2(4), 100–131. Available from: PM:16892160.
- DeFrances, C. J. (2008). 2006 National Hospital Discharge Survey.
- Delongchamps, N. B., Singh, A., & Haas, G. P. (2007). Epidemiology of prostate cancer in Africa: Another step in the understanding of the disease? *Current Problems in Cancer*, 31(3), 226–236. Available from: PM:17543950.
- Demirovic, J., Prineas, R., Loewenstein, D., Bean, J., Duara, R., Sevush, S., et al. (2003). Prevalence of dementia in three ethnic groups: The South Florida program on aging and health. *Annals of Epidemiology*, 13(6), 472–478. Available from: PM:12875807.
- Dimou, A., Syrigos, K. N., & Saif, M. W. (2009). Disparities in colorectal cancer in African-Americans vs Whites: Before and after diagnosis. *World Journal of Gastroenterology*, 15(30), 3734–3743. Available from: PM:19673013.



- Do, R., Bailey, S. D., Desbiens, K., Belisle, A., Montpetit, A., Bouchard, C., et al. (2008). Genetic variants of FTO influence adiposity, insulin sensitivity, leptin levels, and resting metabolic rate in the Quebec Family Study. *Diabetes*, 57(4), 1147–1150. Available from: PM:18316358.
- Dunlay, S. M. (2014). Understanding the epidemic of heart failure: Past, present, and future.
- Eastham, J. A., May, R. A., Whatley, T., Crow, A., Venable, D. D., & Sartor, O. (1998). Clinical characteristics and biopsy specimen features in African-American and White men without prostate cancer. *Journal of National Cancer Institute*, 90(10), 756–760. Available from: PM:9605645.
- Eberhardt, M. S., Ingram, D. D., Makuc, D. M., et al. (2001) Health United States, 2001 with urban and rural health chartbook. Hyattsville, MD.
- Eckel, R. H., Kahn, S. E., Ferrannini, E., Goldfine, A. B., Nathan, D. M., Schwartz, M. W., et al. (2011). Obesity and type 2 diabetes: What can be unified and what needs to be individualized? *Diabetes Care*, 34(6), 1424–1430. Available from: PM:21602431.
- Edwards, S. M. (2010). Prostate cancer in BRCA2 germline mutation carriers is associated with poorer prognosis.
- Edwards, R., Unwin, N., Mugusi, F., Whiting, D., Rashid, S., Kissima, J., et al. (2000). Hypertension prevalence and care in an urban and rural area of Tanzania. *Journal of Hypertension*, 18(2), 145–152. Available from: PM:10694181.
- Eichler, E. E., Flint, J., Gibson, G., Kong, A., Leal, S. M., Moore, J. H., et al. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews Genetics*, 11(6), 446–450. Available from: PM:20479774.
- Elbein, S. C. (2007). Evaluation of polymorphisms known to contribute to risk for diabetes in African and African-American populations. *Current Opinion in Clinical Nutrition and Metabolic Care*, 10(4), 415–419. Available from: PM:17563458.
- Esteban, E., Rodon, N., Via, M., Gonzalez-Perez, E., Santamaria, J., Dugoujon, J. M., et al. (2006). Androgen receptor CAG and GGC polymorphisms in Mediterraneans: Repeat dynamics and population relationships. *Journal of Human Genetics*, 51(2), 129–136. Available from: PM:16365681.
- Ewing, C. M. (2012). Germline mutations in HOXB13 and prostate-cancer risk.
- Felix, R., Bodmer, W., Fearnhead, N. S., van der Merwe, L., Goldberg, P., & Ramesar, R. S. (2006). GSTM1 and GSTT1 polymorphisms as modifiers of age at diagnosis of hereditary nonpolyposis colorectal cancer (HNPCC) in a homogeneous cohort of individuals carrying a single predisposing mutation. *Mutation Research*, 602(1–2), 175–181. Available from: PM:17087981.
- Ferlay, J. (2015). Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012.
- Ferrari, S., Di Iorio, E., Barbaro, V., Ponzin, D., Sorrentino, F. S., & Parmeggiani F. (2011). Retinitis pigmentosa: genes and disease mechanisms. *Current Genomics*, 12, 238–249.
- Folsom, A. R., Kaye, S. A., Potter, J. D., & Prineas, R. J. (1989). Association of incident carcinoma of the endometrium with body weight and fat distribution in older women: Early findings of the Iowa Women's Health Study. *Cancer Research*, 49(23), 6828–6831. Available from: PM:2819722.
- Forrester, T., Cooper, R. S., & Weatherall, D. (1998). Emergence of Western diseases in the tropical world: The experience with chronic cardiovascular diseases. *British Medical Bulletin*, 54(2), 463–473. Available from: PM:9830210.
- Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*, 316(5826), 889–894. Available from: PM:17434869.
- Frazer, K. A., Murray, S. S., Schork, N. J., & Topol, E. J. (2009). Human genetic variation and its contribution to complex traits. *Nature Reviews Genetics*, 10(4), 241–251. Available from: PM:19293820.
- Freedman, B. I., Kopp, J. B., Langefeld, C. D., Genovese, G., Friedman, D. J., Nelson, G. W., et al. (2010). The apolipoprotein L1 (APOL1) gene and nondiabetic nephropathy in African

- Americans. *Journal of American Society of Nephrology*, 21(9), 1422–1426. Available from: PM:20688934.
- Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., et al. (2010). Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science*, 329(5993), 841–845. Available from: PM:20647424.
- Gibbons, G. H. (2004). Physiology, genetics, and cardiovascular disease: Focus on African Americans.
- Goossens, M. (1991). The cystic fibrosis gene: Mutation and the function of CFTR protein. *Annals of Pediatrics (Paris)*, 38(9), 591–594. Available from: PM:1721508.
- Graff-Radford, N. R., Green, R. C., Go, R. C., Hutton, M. L., Edeki, T., Bachman, D., et al. (2002). Association between apolipoprotein E genotype and Alzheimer disease in African American subjects. *Archives of Neurology*, 59(4), 594–600. Available from: PM:11939894.
- Gu, C., Borecki, I., Gagnon, J., Bouchard, C., Leon, A. S., Skinner, J. S., et al. (1998). Familial resemblance for resting blood pressure with particular reference to racial differences: Preliminary analyses from the HERITAGE Family Study. *Human Biology*, 70(1), 77–90. Available from: PM:9489236.
- Haddad, S. A. (2015). Hormone-related pathways and risk of breast cancer subtypes in African American women.
- Haffner, S. M. (1998). Epidemiology of type 2 diabetes: Risk factors. *Diabetes Care*, 21(Suppl 3), C3–C6. Available from: PM:9850478.
- Hainaut, P., & Boyle, P. (2008). Curbing the liver cancer epidemic in Africa. *Lancet*, 371(9610), 367–368. Available from: PM:18242399.
- Hannon, T. S., Bacha, F., Lin, Y., & Arslanian, S. A. (2008). Hyperinsulinemia in African-American adolescents compared with their American white peers despite similar insulin sensitivity: A reflection of upregulated beta-cell function? *Diabetes Care*, 31(7), 1445–1447. Available from: PM:18417751.
- Hattersley, A., Bruining, J., Shield, J., Njolstad, P., & Donaghue, K. C. (2009). The diagnosis and management of monogenic diabetes in children and adolescents. *Pediatrics Diabetes*, 10(Suppl 12), 33–42. Available from: PM:19754616.
- Hayanga, A. J., & D’Cunha, J. (2013). Racial disparities and lung cancer care: Why the unequal playing field? *Seminars in Thoracic Cardiovascular Surgery*, 25(1), 2–3. Available from: PM:23800522.
- Helgadóttir, A., Manolescu, A., Helgason, A., Thorleifsson, G., Thorsteinsdóttir, U., Gudbjartsson, D. F., et al. (2006). A variant of the gene encoding leukotriene A4 hydrolase confers ethnicity-specific risk of myocardial infarction. *Nature Genetics*, 38(1), 68–74. Available from: PM:16282974.
- Hendrie, H. C., Murrell, J., Gao, S., Unverzagt, F. W., Ogunniyi, A., & Hall, K. S. (2006). International studies in dementia with particular emphasis on populations of African origin. *Alzheimer Disease and Associated Disorders*, 20(3 Suppl 2), S42–S46. Available from: PM:16917194.
- Hendrie, H. C., Ogunniyi, A., Hall, K. S., Baiyewu, O., Unverzagt, F. W., Gureje, O., et al. (2001). Incidence of dementia and Alzheimer disease in 2 communities: Yoruba residing in Ibadan, Nigeria, and African Americans residing in Indianapolis, Indiana. *JAMA*, 285(6), 739–747. Available from: PM:11176911.
- Hertz, R. P. (2005). Racial disparities in hypertension prevalence, awareness, and management.
- Hertz, R. P., Unger, A. N., Cornell, J. A., & Saunders, E. (2005). Racial disparities in hypertension prevalence, awareness, and management. *Archives of Internal Medicine*, 165(18), 2098–2104. Available from: PM:16216999.
- Hill, A. V. (2006). Aspects of genetic susceptibility to human infectious diseases. *Annual Reviews of Genetics*, 40, 469–486. Available from: PM:17094741.
- Hirsch, C., Anderson, M. L., Newman, A., Kop, W., Jackson, S., Gottdiener, J., et al. (2006). The association of race with frailty: The cardiovascular health study. *Annals of Epidemiology*, 16(7), 545–553. Available from: PM:16388967.

- Howson, J. M., Roy, M. S., Zeitels, L., Stevens, H., & Todd, J. A. (2013). HLA class II gene associations in African American type 1 diabetes reveal a protective HLA-DRB1\*03 haplotype. *Diabetics Medicine*, 30(6), 710–716. Available from: PM:23398374.
- Hu, C. A., Klopfer, E. I., & Ray, P. E. (2012). Human apolipoprotein L1 (ApoL1) in cancer and chronic kidney disease. *FEBS Letters*, 586(7), 947–955. Available from: PM:22569246.
- Isaacs, W. B. (2012). Inherited susceptibility for aggressive prostate cancer.
- Jallow, M., Teo, Y. Y., Small, K. S., Rockett, K. A., Deloukas, P., Clark, T. G., et al. (2009). Genome-wide and fine-resolution association analysis of malaria in West Africa. *Nature Genetics*, 41(6), 657–665. Available from: PM:19465909.
- Jepson, A., Fowler, A., Banya, W., Singh, M., Bennett, S., Whittle, H., et al. (2001). Genetic regulation of acquired immune responses to antigens of *Mycobacterium tuberculosis*: A study of twins in West Africa. *Infection and Immunity*, 69(6), 3989–3994. Available from: PM:11349068.
- Joe, A. K., Arber, N., Bose, S., Heitjan, D., Zhang, Y., Weinstein, I. B., et al. (2001). Cyclin D1 overexpression is more prevalent in non-Caucasian breast cancer. *Anticancer Research*, 21(5), 3535–3539. Available from: PM:11848520.
- Johns, L. E. (2001). A systematic review and meta-analysis of familial colorectal cancer risk.
- Kao, W. H. (2008). MYH9 is associated with nondiabetic end-stage renal disease in African Americans.
- Kardia, S. L. (2000). Context-dependent genetic effects in hypertension. *Current Hypertension Report*, 2(1), 32–38. Available from: PM:10981124.
- Kaufman, J. S., Durazo-Arvizu, R. A., Rotimi, C. N., McGee, D. L., & Cooper, R. S. (1996). Obesity and hypertension prevalence in populations of African origin. The investigators of the international collaborative study on hypertension in blacks. *Epidemiology*, 7(4), 398–405. Available from: PM:8793366.
- Kaufman, J. S., Owoaje, E. E., Rotimi, C. N., & Cooper, R. S. (1999). Blood pressure change in Africa: Case study from Nigeria. *Human Biology*, 71(4), 641–657. Available from: PM:10453105.
- Keaton, J. M., Cooke Bailey, J. N., Palmer, N. D., Freedman, B. I., Langefeld, C. D., Ng, M. C., et al. (2014). A comparison of type 2 diabetes risk allele load between African Americans and European Americans. *Human Genetics*, 133(12), 1487–1495. Available from: PM:25273842.
- Keku, T. O., Vidal, A., Oliver, S., Hoyo, C., Hall, I. J., Omofoye, O., et al. (2012). Genetic variants in IGF-I, IGF-II, IGFBP-3, and adiponectin genes and colon cancer risk in African Americans and Whites. *Cancer Causes & Control*, 23(7), 1127–1138. Available from: PM:22565227.
- Kimani, J., Kaul, R., Nagelkerke, N. J., Luo, M., MacDonald, K. S., Ngugi, E., et al. (2008). Reduced rates of HIV acquisition during unprotected sex by Kenyan female sex workers predating population declines in HIV prevalence. *AIDS*, 22(1), 131–137. Available from: PM:18090401.
- Kittles, R. A., Chen, W., Panguluri, R. K., Ahaghotu, C., Jackson, A., Adebamowo, C. A., et al. (2002). CYP3A4-V and prostate cancer in African Americans: Causal or confounding association because of population stratification? *Human Genetics*, 110(6), 553–560. Available from: PM:12107441.
- Kittles, R. A., Young, D., Weinrich, S., Hudson, J., Argyropoulos, G., Ukoli, F., et al. (2001). Extent of linkage disequilibrium between the androgen receptor gene CAG and GGC repeats in human populations: Implications for prostate cancer risk. *Human Genetics*, 109(3), 253–261. Available from: PM:11702204.
- Kohler, B. A. (2015). Annual report to the nation on the status of cancer, 1975–2011, featuring incidence of breast cancer subtypes by race/ethnicity, poverty, and state.
- Krieger, N. (2005). Stormy weather: Race, gene expression, and the science of health disparities. *American Journal of Public Health*, 95, 2155–2160.
- Lama, J., & Planelles, V. (2007). Host factors influencing susceptibility to HIV infection and AIDS progression. *Retrovirology*, 4, 52. Available from: PM:17651505.
- Lamba, J. K. (2002). Genetic contribution to variable human CYP3A-mediated metabolism.

- Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), 860–921. Available from: PM:11237011.
- Lansdorp-Vogelaar, I., Kuntz, K. M., Knudsen, A. B., van, B. M., Zauber, A. G., & Jemal, A. (2012). Contribution of screening and survival differences to racial disparities in colorectal cancer rates. *Cancer Epidemiology, Biomarkers & Prevention*, 21(5), 728–736. Available from: PM:22514249.
- Liu, V., Escobar, G. J., Greene, J. D., Soule, J., Whippy, A., Angus, D. C., et al. (2014). Hospital deaths in patients with sepsis from 2 independent cohorts. *JAMA*, 312(1), 90–92. Available from: PM:24838355.
- Liu, C. C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: Risk, mechanisms and therapy. *Nature Reviews Neurology*, 9(2), 106–118. Available from: PM:23296339.
- Loo, L. W., Wang, Y., Flynn, E. M., Lund, M. J., Bowles, E. J., Buist, D. S., et al. (2011). Genome-wide copy number alterations in subtypes of invasive breast cancers in young white and African American women. *Breast Cancer Research & Treatment*, 127(1), 297–308. Available from: PM:21264507.
- Loock, M., Steyn, K., Becker, P., & Fourie, J. (2006). Coronary heart disease and risk factors in Black South Africans: A case-control study. *Ethnicity & Disease*, 16(4), 872–879. Available from: PM:17061740.
- Loos, R. J. (2014). The bigger picture of FTO: The first GWAS-identified obesity gene.
- Luke, A., Guo, X., Adeyemo, A. A., Wilks, R., Forrester, T., Lowe, W., Jr., et al. (2001). Heritability of obesity-related traits among Nigerians, Jamaicans and US black people. *International Journal of Obesity and Related Metabolic Disorders*, 25(7), 1034–1041. Available from: PM:11443503.
- Magi-Galluzzi, C., Tsusuki, T., Elson, P., Simmerman, K., LaFargue, C., Esgueva, R., et al. (2011). TMPRSS2-ERG gene fusion prevalence and class are significantly different in prostate cancer of Caucasian, African-American and Japanese patients. *Prostate*, 71(5), 489–497. Available from: PM:20878952.
- Magooha, G. A. (2007). Overview of prostate cancer in indigenous black Africans and blacks of African ancestry in diaspora 1935–2007. *East African Medical Journal*, 84(9 Suppl), S3–S11. Available from: PM:18154197.
- Malone, K. E., Daling, J. R., Doody, D. R., Hsu, L., Bernstein, L., Coates, R. J., et al. (2006). Prevalence and predictors of BRCA1 and BRCA2 mutations in a population-based study of breast cancer in white and black American women ages 35 to 64 years. *Cancer Research*, 66(16), 8297–8308. Available from: PM:16912212.
- Martin, D. N., Boersma, B. J., Yi, M., Reimers, M., Howe, T. M., Yfantis, H. G., et al. (2009). Differences in the tumor microenvironment between African-American and European-American breast cancer patients. *PLoS One*, 4(2), e4531. Available from: PM:19225562.
- Matanoski, G., Tao, X., Almon, L., Adade, A. A., & Davies-Cole, J. O. (2006). Demographics and tumor characteristics of colorectal cancers in the United States, 1998–2001. *Cancer*, 107(5 Suppl), 1112–1120. Available from: PM:16838314.
- Maxwell, K. N. (2012). Cancer treatment according to BRCA1 and BRCA2 mutations.
- Maxwell, K. N. (2013). Common breast cancer risk variants in the post-COGS era: A comprehensive review.
- McCracken, M. (2007). Cancer incidence, mortality, and associated risk factors among Asian Americans of Chinese, Filipino, Vietnamese, Korean, and Japanese ethnicities.
- Mendel, G. (1866). Versuche über Pflanzen-Hybriden. *Verhandlungen des naturforschenden Vereines, Abhandlungen. Brunn*, 4, 3–47.
- Menges, T., Hermans, P. W., Little, S. G., Langefeld, T., Boning, O., Engel, J., et al. (2001). Plasminogen-activator-inhibitor-1 4G/5G promoter polymorphism and prognosis of severely injured patients. *Lancet*, 357(9262), 1096–1097. Available from: PM:11297964.

- Miljkovic-Gacic, I., Gordon, C. L., Goodpaster, B. H., Bunker, C. H., Patrick, A. L., Kuller, L. H., et al. (2008). Adipose tissue infiltration in skeletal muscle: Age patterns and association with diabetes among men of African ancestry. *American Journal of Clinical Nutrition*, 87(6), 1590–1595. Available from: PM:18541544.
- Moore, J. H., & Williams, S. M. (2002). New strategies for identifying gene-gene interactions in hypertension. *Annals of Medicine*, 34(2), 88–95. Available from: PM:12108579.
- Morris, M. R., Ricketts, C., Gentle, D., Abdulrahman, M., Clarke, N., Brown, M., et al. (2010). Identification of candidate tumour suppressor genes frequently methylated in renal cell carcinoma. *Oncogene*, 29(14), 2104–2117. Available from: PM:20154727.
- Motala, A. A. (2002). Diabetes trends in Africa. *Diabetes Metabolism Research and Reviews*, 18 (Suppl 3), S14–S20. Available from: PM:12324980.
- Muro, F. (2015). Acute respiratory infection and bacteraemia as causes of non-malarial febrile illness in African children: A narrative review.
- Murrell, J. R., Price, B. M., Baiyewu, O., Gureje, O., Deeg, M., Hendrie, H., et al. (2006). The fourth apolipoprotein E haplotype found in the Yoruba of Ibadan. *American Journal of Medical Genetics B Neuropsychiatric Genetics*, 141B(4), 426–427. Available from: PM:16583434.
- Mwntembe, O., Gaillard, M. C., Barkhuizen, M., Pillay, V., Berry, S. D., Dewar, J. B., et al. (2001). Ethnic differences in allelic associations of the interleukin-1 gene cluster in South African patients with inflammatory bowel disease (IBD) and in control individuals. *Immunogenetics*, 52(3–4), 249–254. Available from: PM:11220627.
- Newport, M. J., Allen, A., Awomoyi, A. A., Dunstan, S. J., McKinney, E., Marchant, A., et al. (2004). The toll-like receptor 4 Asp299Gly variant: No influence on LPS responsiveness or susceptibility to pulmonary tuberculosis in The Gambia. *Tuberculosis (Edinburgh)*, 84(6), 347–352. Available from: PM:15525557.
- Nielsen, M., Joerink-van de Beld, M.C., Jones, N., Vogt, S., Tops, C. M., Vasen, H. F., et al. (2009). Analysis of MUTYH genotypes and colorectal phenotypes in patients with MUTYH-associated polyposis. *Gastroenterology*, 136(2), 471–476. Available from: PM:19032956.
- Nunes, J. P. (2003). Arterial hypertension and sepsis. *Revista Portuguesa de Cardiologia*, 22(11), 1375–1379. Available from: PM:14768492.
- Oliveira, T. Y., Harris, E. E., Meyer, D., Jue, C. K., & Silva, W. A., Jr. (2012). Molecular evolution of a malaria resistance gene (DARC) in primates. *Immunogenetics*, 64(7), 497–505. Available from: PM:22395823.
- Ouerhani, S., Oliveira, E., Marrakchi, R., Ben Slama, M. R., Sfaxi, M., Ayed, M., et al. (2007). Methylenetetrahydrofolate reductase and methionine synthase polymorphisms and risk of bladder cancer in a Tunisian population. *Cancer Genetics and Cytogenetics*, 176(1), 48–53. Available from: PM:17574963.
- Ouerhani, S., Tebourski, F., Slama, M. R., Marrakchi, R., Rabeh, M., Hassine, L. B., et al. (2006). The role of glutathione transferases M1 and T1 in individual susceptibility to bladder cancer in a Tunisian population. *Annals of Human Biology*, 33(5–6), 529–535. Available from: PM:17381051.
- Parkin, D. M., Ferlay, J., Hamdi-Chérif, M., Sitas, F., Thomas, J., Wabinga, H., et al. (2003). *Cancer in Africa: Epidemiology and prevention*. Lyon: IARC Press.
- Parra, E. J., Kittles, R. A., Argyropoulos, G., Pfaff, C. L., Hiester, K., Bonilla, C., et al. (2001). Ancestral proportions and admixture dynamics in geographically defined African Americans living in South Carolina. *American Journal of Physical Anthropology*, 114(1), 18–29. Available from: PM:11150049.
- Perez-Morga, D., Vanhollenbeke, B., Paturiaux-Hanocq, F., Nolan, D. P., Lins, L., Homble, F., et al. (2005). Apolipoprotein L-I promotes trypanosome lysis by forming pores in lysosomal membranes. *Science*, 309(5733), 469–472. Available from: PM:16020735.
- Petrovics, G., Liu, A., Shaheduzzaman, S., Furusato, B., Sun, C., Chen, Y., et al. (2005). Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. *Oncogene*, 24 (23), 3847–3852. Available from: PM:15750627.

- Pharoah, P. D. (2002). Polygenic susceptibility to breast cancer and implications for prevention.
- Pibiri, F., Kittles, R. A., Sandler, R. S., Keku, T. O., Kupfer, S. S., Xicola, R. M., et al. (2014). Genetic variation in vitamin D-related genes and risk of colorectal cancer in African Americans. *Cancer Causes & Control*, 25(5), 561–570. Available from: PM:24562971.
- Platz, E. A., & Giovannucci, E. (2004). The epidemiology of sex steroid hormones and their signaling and metabolic pathways in the etiology of prostate cancer. *Journal of Steroid Biochemistry and Molecular Biology*, 92(4), 237–253. Available from: PM:15663987.
- Poland, G. A. (1998). Variability in immune response to pathogens: Using measles vaccine to probe immunogenetic determinants of response. *American Journal of Human Genetics*, 62(2), 215–220. Available from: PM:9463343.
- Ramsay, M. (2012). Africa: Continent of genome contrasts with implications for biomedical research and health. *FEBS Letters*, 586(18), 2813–2819. Available from: PM:22858376.
- Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D., et al. (2006). Global variation in copy number in the human genome. *Nature*, 444(7118), 444–454. Available from: PM:17122850.
- Reed, F. A., & Tishkoff, S. A. (2006). African human diversity, origins and migrations. *Current Opinion in Genetics & Development*, 16(6), 597–605. Available from: PM:17056248.
- Refolo, L. M., Malester, B., Lafrancois, J., Bryant-Thomas, T., Wang, R., Tint, G. S., et al. (2000). Hypercholesterolemia accelerates the Alzheimer's amyloid pathology in a transgenic mouse model. *Neurobiology of Disease*, 7(4), 321–331. Available from: PM:10964604.
- Reitz, C., & Mayeux, R. (2014). Genetics of Alzheimer's disease in Caribbean Hispanic and African American populations. *Biological Psychiatry*, 75(7), 534–541. Available from: PM:23890735.
- Rennert, H., Zeigler-Johnson, C. M., Addya, K., Finley, M. J., Walker, A. H., Spangler, E., et al. (2005). Association of susceptibility alleles in ELAC2/HPC2, RNASEL/HPC1, and MSR1 with prostate cancer severity in European American and African American men. *Cancer Epidemiology, Biomarkers & Prevention*, 14(4), 949–957. Available from: PM:15824169.
- Rich, S. S. (1990). Mapping genes in diabetes. Genetic epidemiological perspective. *Diabetes*, 39(11), 1315–1319. Available from: PM:2227105.
- Romanova, E. N. (2013). TNF-alpha, IL-10, and eNOS gene polymorphisms in patients with influenza A/H1N1 complicated by pneumonia.
- Rosanas-Urgell, A. (2012). Lack of associations of alpha(+)-thalassemia with the risk of *Plasmodium falciparum* and *Plasmodium vivax* infection and disease in a cohort of children aged 3–21 months from Papua New Guinea.
- Rose, D. P., & Royak-Schaler, R. (2001). Tumor biology and prognosis in black breast cancer patients: A review. *Cancer Detection and Prevention*, 25(1), 16–31. Available from: PM:11270418.
- Roses, A. D., Lutz, M. W., Saunders, A. M., Goldgaber, D., Saul, R., Sundseth, S. S., et al. (2014). African-American TOMM40\*523-APOE haplotypes are admixture of West African and Caucasian alleles. *Alzheimers.Dementia*, 10(6), 592–601. Available from: PM:25260913.
- Rotimi, C. N., Cooper, R. S., Ataman, S. L., Osotimehin, B., Kadiri, S., Muna, W., et al. (1995). Distribution of anthropometric variables and the prevalence of obesity in populations of west African origin: The International Collaborative Study on Hypertension in Blacks (ICSHIB). *Obesity Research*, 3(Suppl 2), 95s–105s. Available from: PM:8581794.
- Rotimi, C. N., Cooper, R. S., Cao, G., Ogunbiyi, O., Ladipo, M., Owoaje, E., et al. (1999a). Maximum-likelihood generalized heritability estimate for blood pressure in Nigerian families. *Hypertension*, 33(3), 874–878. Available from: PM:10082502.
- Rotimi, C. N., Cooper, R. S., Okosun, I. S., Olatunbosun, S. T., Bella, A. F., Wilks, R., et al. (1999b). Prevalence of diabetes and impaired glucose tolerance in Nigerians, Jamaicans and US blacks. *Ethnicity & Disease*, 9(2), 190–200. Available from: PM:10421081.
- Rotimi, C. N., Dunston, G. M., Berg, K., Akinsete, O., Amoah, A., Owusu, S., et al. (2001). In search of susceptibility genes for type 2 diabetes in West Africa: The design and results of the first phase of the AADM study. *Annals of Epidemiology*, 11(1), 51–58. Available from: PM:11164120.

- Rotimi, C., Morrison, L., Cooper, R., Oyejide, C., Effiong, E., Ladipo, M., et al. (1994). Angiotensinogen gene in human hypertension. Lack of an association of the 235T allele among African Americans. *Hypertension*, 24(5), 591–594. Available from: PM:7960018.
- Russell, E. F., Kramer, M. R., Cooper, H. L., Gabram-Mendola, S., Senior-Crosby, D., & Jacob Arriola, K. R. (2012). Metropolitan area racial residential segregation, neighborhood racial composition, and breast cancer mortality. *Cancer Causes & Control*, 23(9), 1519–1527. Available from: PM:22825071.
- Saad, A. A., O'Connor, P. J., Mostafa, M. H., Metwalli, N. E., Cooper, D. P., Povey, A. C., et al. (2005). Glutathione S-transferase M1, T1 and P1 polymorphisms and bladder cancer risk in Egyptians. *International Journal of Biological Markers*, 20(1), 69–72. Available from: PM:15832776.
- Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z., Richter, D. J., Schaffner, S. F., et al. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419(6909), 832–837. Available from: PM:12397357.
- Sandy, A. S., Palmer, N. D., Hanley, A. J., Ziegler, J. T., Mark, B. W., Freedman, B. I., et al. (2013). Genetic analysis of adiponectin variation and its association with type 2 diabetes in African Americans. *Obesity (Silver Spring)*, 21(12), E721–E729. Available from: PM:23512866.
- Scherer, S. W., Lee, C., Birney, E., Altshuler, D. M., Eichler, E. E., Carter, N. P., et al. (2007). Challenges and standards in integrating surveys of structural variation. *Nature Genetics*, 39(7 Suppl), S7–S15. Available from: PM:17597783.
- Schlesinger, D., Grinberg, L. T., Alba, J. G., Naslavsky, M. S., Licinio, L., Farfel, J. M., et al. (2013). African ancestry protects against Alzheimer's disease-related neuropathology. *Molecular Psychiatry*, 18(1), 79–85. Available from: PM:22064377.
- Schoenhard, J. A., Asselbergs, F. W., Poku, K. A., Stocki, S. A., Gordon, S., Vaughan, D. E., et al. (2008). Male-female differences in the genetic regulation of t-PA and PAI-1 levels in a Ghanaian population. *Human Genetics*, 124(5), 479–488. Available from: PM:18953568.
- Scholz, C. A., Johnson, T. C., Cohen, A. S., King, J. W., Peck, J. A., Overpeck, J. T., et al. (2007). East African megadroughts between 135 and 75 thousand years ago and bearing on early-modern human origins. *Proceedings of National Academy Science U S A*, 104(42), 16416–16421. Available from: PM:17785420.
- Schwartz, G. F., Hughes, K. S., Lynch, H. T., Fabian, C. J., Fentiman, I. S., Robson, M. E., et al. (2008). Proceedings of the international consensus conference on breast cancer risk, genetics, & risk management, April, 2007. *Cancer*, 113(10), 2627–2637. Available from: PM:18853415.
- Seeleang, K. (2011). Genetic disparities in the development of type 2 diabetes among African Americans. *Journal of American Academy of Nurse Practitioners*, 23(9), 473–478. Available from: PM:21899642.
- Shand, R. L., & Gelmann, E. P. (2006). Molecular biology of prostate-cancer pathogenesis. *Current Opinion in Urology*, 16(3), 123–131. Available from: PM:16679847.
- Shields, P. G., Caporaso, N. E., Falk, R. T., Sugimura, H., Trivers, G. E., Trump, B. F., et al. (1993). Lung cancer, race, and a CYP1A1 Genetic polymorphism. *Cancer Epidemiology, Biomarkers & Prevention*, 2(5), 481–485. Available from: PM:8220094.
- Shriver, M. D., & Kittles, R. A. (2004). Genetic ancestry and the search for personalized genetic histories. *Nature Reviews Genetics*, 5(8), 611–618. Available from: PM:15266343.
- Siffert, W., Naber, C., Walla, M., & Ritz, E. (1999). G protein beta3 subunit 825T allele and its potential association with obesity in hypertensive individuals. *Journal of Hypertension*, 17(8), 1095–1098. Available from: PM:10466464.
- Singh, C. R., & Kathiresan, K. (2014). Molecular understanding of lung cancers—A review. *Asian Pacific Journal of Tropical Biomedicine*, 4(Suppl 1), S35–S41. Available from: PM:25183110.
- Sniderman, A. D., Bhopal, R., Prabhakaran, D., Sarrafzadegan, N., & Tchernof, A. (2007). Why might South Asians be so susceptible to central obesity and its atherogenic consequences? The adipose tissue overflow hypothesis. *International Journal of Epidemiology*, 36(1), 220–225. Available from: PM:17510078.



- Sommer, A., Tielsch, J. M., Katz, J., Quigley, H. A., Gottsch, J. D., Javitt, J. C., et al. (1991). Racial differences in the cause-specific prevalence of blindness in east Baltimore. *New England Journal of Medicine*, 325(20), 1412–1417. Available from: PM:1922252.
- Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., et al. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proceedings of National Academic Science U S A*, 98(19), 10869–10874. Available from: PM:11553815.
- Spencer, C. C., Su, Z., Donnelly, P., & Marchini, J. (2009). Designing genome-wide association studies: Sample size, power, imputation, and the choice of genotyping chip. *PLoS Genetics*, 5 (5), e1000477. Available from: PM:19492015.
- Stefan, N., Kantartzis, K., Machann, J., Schick, F., Thamer, C., Rittig, K., et al. (2008). Identification and characterization of metabolically benign obesity in humans. *Archives of Internal Medicine*, 168(15), 1609–1616. Available from: PM:18695074.
- Strachan, T., & Read, A. P. (1999). Mendelian pedigree patterns. *Human Molecular Genetics*, 2 (Garland Science).
- Taherzadeh, Z. (2010). Function and structure of resistance vessels in black and white people.
- Tang, M. X., Cross, P., Andrews, H., Jacobs, D. M., Small, S., Bell, K., et al. (2001). Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. *Neurology*, 56(1), 49–56. Available from: PM:11148235.
- Teo, Y. Y., Small, K. S., & Kwiatkowski, D. P. (2010). Methodological challenges of genome-wide association analysis in Africa. *Nature Reviews Genetics*, 11(2), 149–160. Available from: PM:20084087.
- The Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14000 cases of seven common diseases and 3000 shared controls. *Nature*, 447, 661–678.
- Tibazarwa, K., Ntyintyane, L., Sliwa, K., Gernholtz, T., Carrington, M., Wilkinson, D., et al. (2009). A time bomb of cardiovascular risk factors in South Africa: Results from the Heart of Soweto Study “Heart Awareness Days”. *International Journal of Cardiology*, 132(2), 233–239. Available from: PM:18237791.
- Tidball, J. G., & Wehling-Henricks, M. (2004). Evolving therapeutic strategies for Duchenne muscular dystrophy: Targeting downstream events. *Pediatric Research*, 56(6), 831–841. Available from: PM:15531741.
- Tishkoff, S. A., Dietzsch, E., Speed, W., Pakstis, A. J., Kidd, J. R., Cheung, K., et al. (1996). Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science*, 271(5254), 1380–1387. Available from: PM:8596909.
- Tishkoff, S. A., Gonder, K., Hirbo, J., Mortensen, H., Powell, K., Knight, A. et al. (2003). *The genetic history of linguistically diverse Tanzanian populations: A multilocus analysis*. Tempe: American Association of Physical Anthropology.
- Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., et al. (2009). The genetic structure and history of Africans and African Americans. *Science*, 324(5930), 1035–1044. Available from: PM:19407144.
- Todd, J. A., Walker, N. M., Cooper, J. D., Smyth, D. J., Downes, K., Plagnol, V., et al. (2007). Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. *Nature Genetics*, 39(7), 857–864. Available from: PM:17554260.
- Torcia, M. G., Santarlasci, V., Cosmi, L., Clemente, A., Maggi, L., Mangano, V. D., et al. (2008). Functional deficit of T regulatory cells in Fulani, an ethnic group with low susceptibility to *Plasmodium falciparum* malaria. *Proceedings of National Academic Science U S A*, 105(2), 646–651. Available from: PM:18174328.
- Tournamille, C., Colin, Y., Cartron, J. P., & Le Van, K. C. (1995). Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nature Genetics*, 10(2), 224–228. Available from: PM:7663520.
- Trauth, M. H., Maslin, M. A., Deino, A. L., Strecker, M. R., Bergner, A. G., & Duhnforth, M. (2007). High- and low-latitude forcing of Plio-Pleistocene East African climate and human evolution. *Journal of Human Evolution*, 53(5), 475–486. Available from: PM:17959230.

- Troxel, W. M., Matthews, K. A., Bromberger, J. T., & Sutton-Tyrrell, K. (2003). Chronic stress burden, discrimination, and subclinical carotid artery disease in African American and Caucasian women. *Health Psychology, 22*(3), 300–309. Available from: PM:12790258.
- Underwood, P. C., Sun, B., Williams, J. S., Pojoga, L. H., Chamarthi, B., Lasky-Su, J., et al. (2010). The relationship between peroxisome proliferator-activated receptor-gamma and renin: a human genetics study. *Journal of Clinical Endocrinology & Metabolism, 95*, E75–E95.
- Unwin, N., Setel, P., Rashid, S., Mugusi, F., Mbanja, J. C., Kitange, H., et al. (2001). Noncommunicable diseases in sub-Saharan Africa: Where do they feature in the health research agenda? *Bulletin of World Health Organisation, 79*(10), 947–953. Available from: PM:11693977.
- Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D. P., Lins, L., Van Den Abbeele, J., et al. (2003). Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature, 422* (6927), 83–87. Available from: PM:12621437.
- Velasquez-Miery, P. A., Cowan, P. A., Umpierrez, G. E., Lustig, R. H., Cashion, A. K., & Burghen, G. A. (2003). Racial differences in glucagon-like peptide-1 (GLP-1) concentrations and insulin dynamics during oral glucose tolerance test in obese subjects. *International Journal of Obesity and Related Metabolic Disorders, 27*(11), 1359–1364. Available from: PM:14574347.
- Vlachopoulos, C., & Terentes-Printzios, D. (2015). Thyroid, aldosterone, and cardiovascular disease.
- Vogelstein, B., & Kinzler, K. W. (1993). The multistep nature of cancer. *Trends in Genetics, 9*(4), 138–141. Available from: PM:8516849.
- Walker, A. R., Adam, F., & Walker, B. F. (2001). World pandemic of obesity: The situation in Southern African populations. *Public Health, 115*(6), 368–372. Available from: PM:11781845.
- Wall, J. D., Cox, M. P., Mendez, F. L., Woerner, A., Severson, T., & Hammer, M. F. (2008). A novel DNA sequence database for analyzing human demographic history. *Genome Research, 18*(8), 1354–1361. Available from: PM:18493019.
- Wallace, T. A., Prueitt, R. L., Yi, M., Howe, T. M., Gillespie, J. W., Yfantis, H. G., et al. (2008). Tumor immunobiological differences in prostate cancer between African-American and European-American men. *Cancer Research, 68*(3), 927–936. Available from: PM:18245496.
- Wassenaar, C. A., Ye, Y., Cai, Q., Aldrich, M. C., Knight, J., Spitz, M. R., et al. (2015). CYP2A6 reduced activity gene variants confer reduction in lung cancer risk in African American smokers—Findings from two independent populations. *Carcinogenesis, 36*(1), 99–103. Available from: PM:25416559.
- Weatherall, D. J., & Clegg, J. B. (2001). Inherited haemoglobin disorders: An increasing global health problem. *Bulletin of the World Health Organisation, 79*(8), 704–712. Available from: PM:11545326.
- Weston, A., & Godbold, J. H. (1997). Polymorphisms of H-ras-1 and p 53 in breast cancer and lung cancer: A meta-analysis. *Environmental Health Perspectives, 105*(Suppl 4), 919–926. Available from: PM:9255581.
- Willcox, M., Bjorkman, A., & Brohult, J. (1983). Falciparum malaria and beta-thalassaemia trait in northern Liberia. *Annals of Tropical Medicine and Parasitology, 77*(4), 335–347. Available from: PM:6357119.
- Williams, S. M., Addy, J. H., Phillips, J. A., III, Dai, M., Kpodonu, J., Afful, J., et al. (2000). Combinations of variations in multiple genes are associated with hypertension. *Hypertension, 36*(1), 2–6. Available from: PM:10904004.
- Williams, D. R., & Jackson, P. B. (2005). Social sources of racial disparities in health. *Health Affairs (Millwood), 24*(2), 325–334. Available from: PM:15757915.
- Williams, T. N., Maitland, K., Bennett, S., Ganczakowski, M., Peto, T. E., Newbold, C. I., et al. (1996). High incidence of malaria in alpha-thalassaemic children. *Nature, 383*(6600), 522–525. Available from: PM:8849722.
- Williams, S. M., Stocki, S., Jiang, L., Brew, K., Gordon, S., Vaughan, D. E., et al. (2007). A population-based study in Ghana to investigate inter-individual variation in plasma t-PA and PAI-1. *Ethnicity & Diseases, 17*(3), 492–497. Available from: PM:17985503.

- Wurfel, M. M., Gordon, A. C., Holden, T. D., Radella, F., Strout, J., Kajikawa, O., et al. (2008). Toll-like receptor 1 polymorphisms affect innate immune responses and outcomes in sepsis. *American Journal of Respiratory and Critical Care Medicine*, 178(7), 710–720. Available from: PM:18635889.
- Xu, J., Kibel, A. S., Hu, J. J., Turner, A. R., Pruett, K., Zheng, S. L., et al. (2009). Prostate cancer risk associated loci in African Americans. *Cancer Epidemiology, Biomarkers & Prevention*, 18(7), 2145–2149. Available from: PM:19549807.
- Yao, S. (2012). Variants in the vitamin D pathway, serum levels of vitamin D, and estrogen receptor negative breast cancer among African-American women: A case-control study.
- Zabaleta, J., Lin, H. Y., Sierra, R. A., Hall, M. C., Clark, P. E., Sartor, O. A., et al. (2008). Interactions of cytokine gene polymorphisms in prostate cancer risk. *Carcinogenesis*, 29(3), 573–578. Available from: PM:18174250.
- Zanetti, K. A., Haznadar, M., Welsh, J. A., Robles, A. I., Ryan, B. M., McClary, A. C., et al. (2012). 3'-UTR and functional secretor haplotypes in mannose-binding lectin 2 are associated with increased colon cancer risk in African Americans. *Cancer Research*, 72(6), 1467–1477. Available from: PM:22282660.
- Zeigler-Johnson, C. M., Walker, A. H., Mancke, B., Spangler, E., Jalloh, M., McBride, S., et al. (2002). Ethnic differences in the frequency of prostate cancer susceptibility alleles at SRD5A2 and CYP3A4. *Human Heredity*, 54(1), 13–21. Available from: PM:12446983.
- Zimmerman, P. A. (2013). Red blood cell polymorphism and susceptibility to *Plasmodium vivax*.
- Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of National Academic Science U S A*, 109(4), 1193–1198. Available from: PM:22223662.
- Zuniga, J. (2012). Genetic variants associated with severe pneumonia in A/H1N1 influenza infection.

Health Outcomes in a Foreign Land  
A Role for Epigenomic and Environmental Interaction  
Kwabi-Addo, B.  
2017, XXVI, 324 p. 16 illus., Hardcover  
ISBN: 978-3-319-55864-6