Southern African Populations and the Search for the Genetic Basis of Disease Susceptibility and Drug Response

Tiffany Oliver¹ and Mongi Benjeddou*²

¹Department of Biology, Spelman College, 350 Spelman Lane, Atlanta, Georgia, 30310, USA
²Department of Biotechnology, University of the Western Cape, Robert Sobukwe Road, Bellville 7535, South Africa

The vast extent of inherited variation in the human genome has only become apparent since the complete DNA sequence of the human genome has become available [1]. This genomic variation has implications in a broad range of biological and medical disciplines. For this reason, the study of human genetic diversity is relevant to a variety of research areas including human and population genetics, molecular biology, evolutionary biology, biological anthropology, the health sciences and clinical medicine. Variation in the human genome is believed to be the most important cause of variable response to drugs and other xenobiotic and is implicated in susceptibility to almost all diseases [1].

Human genetic studies of complex traits have focused primarily on DNA Sequence Variants (DSVs) that contribute to disease susceptibility, clinical outcomes or response to therapy [2]. Variant genotypes, such as Single Nucleotide Polymorphisms (SNPs) play an important clinical role as they have been shown to alter enzyme activity, resulting in abnormally increased or decreased drug metabolism. For instance, VKORC1-1639 G>A (rs9923231), a SNP located in the promoter region of the gene vitamin K epoxide reductase, has been shown to diminish enzyme activity, causing a "low dose" phenotype in response to treatment with the blood thinning agent warfarin [3]. The approach to molecular genetic studies of complex phenotypes has evolved considerably during the recent years. The candidate gene approach, restricted to the analysis of a few Single Nucleotide Polymorphisms (SNPs) in a modest number of cases and controls, has been supplanted by the unbiased approach of Genome-Wide Association Studies (GWAS), wherein a large number of ‘tag SNPs’ are genotyped in a large number of individuals [2]. Genome-wide association studies, genome sequencing, epigenomics and gene expression are extremely valuable approaches for collecting data that will further our understanding of the pathophysiology of a variety of health-related conditions, however they are also useful for clinical assessments and testing purposes [4]. The understanding of the molecular underpinnings of disease will promote the development of screening and diagnostic tests that will allow us to predict disease outcomes as well as further the field of personalized medicine and facilitate post-treatment surveillance [4].

Public databases, such as The HapMap Project (www.HapMap.org), hold extensive genotype information pertain to genes that affect disease susceptibility and therapeutic outcomes [5]. This includes genome-wide data sets of SNP genotypes, copy-number variants and other forms of structural variations generated using a variety of platforms, including the Illumina HumanM-Duo and the Affymetrix GenomeWide Human SNP Array 6.0. High-throughput genotyping platforms, which can now genotype more than 5 million SNPs, are becoming more common in pharmacogenetic studies as they can be used to identify functional variants associated with disease and drug response [5].

Genomic diversity within sub-Saharan Africa, and for that matter the entire African continent, is relatively under-studied, despite the significant human genomic diversity represented by the people of this region [6,7]. There is thus much to be learnt from characterizing human genomic variation in this part of Africa, especially with regards to health applications [8]. The continent of Africa is the origin of all anatomically modern humans that dispersed across the planet during the past 100,000 years. As such, African populations are characterized by high genetic diversity and low levels of Linkage Disequilibrium (LD) among loci, when compared to populations from other continents [9]. Recent reports using genome-wide polymorphisms also suggested that: (i) genetic variation seen outside of Africa is generally a subset of the total genetic variation that exists within Africa, (ii) genetic diversity decreases with increased geographic distance from Africa, and (iii) linkage disequilibrium (LD) patterns increase proportionally to the distance from Africa [7,10,11]. Moreover, Rosenberg et al. [12] found that there was greater genetic diversity among African populations when compared to Caucasian or Asian populations [12]. African populations also possess a number of genetic adaptations that have evolved in response to the diverse climates, diets, geographic environments, and infectious agents that characterize the continent [9].

Most studies involving African populations have examined regions on the Y-chromosome [13-17] and mitochondrial genomes [16,18-21] in order to characterize the relatedness of individual African subpopulations and the migration of people out of Africa. These studies, suggest that from a genetics standpoint, there is no single "representative" African population. Tishkoff et al. [7] performed a genome-wide analysis of substructure based on DNA from 2432 Africans from 121 genetically diverse populations. The authors analyzed patterns of variation at 1327 nuclear microsatellites and insertion/deletion markers and identified 14 ancestral population clusters that correlated well with self-reported ethnicity and shared cultural or linguistic properties of the populations examined. The results of this study suggest that African populations may have maintained a large and subdivided population structure throughout much of their evolutionary history [9]. Investigations into the population-specific genetic causes underlying communicable and non-communicable diseases have, however, largely relied on the HapMap reference data for Yoruba and Luhya populations to guide study design [22]. The accuracy of this approach remains in doubt, as it is still unclear if conclusions made based on the examination tag SNPs from the Yoruba and the Luhy are applicable to other African populations [22]. The lack of local genetic information with robust allele frequency distributions, particularly for the southern region of Africa, currently serves as a significant hurdle to designing biomedical research and may have important medical implications [22]. This region is inhabited

*Corresponding author: Mongi Benjeddou, Department of Biotechnology, University of the Western Cape, Robert Sobukwe Road, Bellville 7535, South Africa, Tel: +27 (21) 959 2080; E-mail: mongi.benjeddou@gmail.com

Received April 24, 2016; Accepted April 28, 2016; Published April 28, 2016

Citation: Oliver T, Benjeddou M (2016) Southern African Populations and the Search for the Genetic Basis of Disease Susceptibility and Drug Response. J Pharmacokinet Exp Ther 1: e002. doi:10.4172/jpet.1000e002

Copyright: © 2016 Oliver T, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
predominantly by southeastern Bantu-speakers, and is currently suffering under the dual burden of infectious and non-communicable diseases. Limited reference data for this region hampers medical research and impedes our understanding of the underlying population substructure [22].

According to the United Nations Geoscheme, the southern region of Africa is defined as the collection of Botswana, Lesotho, Swaziland, Namibia and South Africa. It is home to a predominant population of Bantu-speakers; a sub-group of the Niger-Kordofanian (NK) linguistic group that expanded southwards from Nigeria and Cameroon, beginning approximately five thousand years ago, reaching South Africa ~ 1500 to 1000 years ago [22]. Southern Africans are geographically distant from other African populations, such as the Yoruba and the Luo, resulting in genetic differentiation due to genetic drift, different selection pressures and admixture with different indigenous groups (such as the Khoe and San) [22]. This region of southeastern Bantu-speakers constitutes one of the African continent’s largest health burdens, and understanding their susceptibility to disease, both communicable and non-communicable, grows increasingly important [22]. Progress, however, is hampered by a paucity of genetic data that necessitates the use of proxy populations; an approach with obvious limitations. An appropriate reference dataset would thus greatly improve local research capabilities and eliminate the need for proxy genetic data [22]. The examination of genetic information from people in the Southern region of Africa would therefore present several key benefits, one being that it would allow us to determine if the variation detected in populations such as the Yoruba and the Luo is representative of the variation present in people from the Southern region of Africa. It would also provide a more accurate reference foundation on which to support future disease research [22]. It is possible that populations from the Southern region of Africa display widely varying genetic allele frequencies for clinically relevant SNPs [3]. Genotyping studies that examine local populations are needed to provide the best medical care to all individuals from this part of the world. In the context of South Africa, with its diverse population groups, the limited studies that have been conducted suggest that South African populations display unique genetic profiles which include novel and rare variants, with allele frequencies that differ between population groups, the latter being a uniquely admixed population of immigrant Europeans, Asians and the indigenous populations [8]. Admixed groups, such as Latinos, African Americans, or Cape Coloureds from South Africa, share varying proportions of different ancestral populations and their genetic complexity can potentially complicate biomedical research studies [28]. Their mixed ancestry, however, can provide the intrinsic variability needed to untangle complex gene-environment interactions, which may help to explain population differences in the epidemiology of complex diseases [28].

References


