The Evolution of Human Genetic and Phenotypic Variation in Africa

Review

Michael C. Campbell[1] and Sarah A. Tishkoff[1,2,*]

Africa is the birthplace of modern humans, and is the source of the geographic expansion of ancestral populations into other regions of the world. Indigenous Africans are characterized by high levels of genetic diversity within and between populations. The pattern of genetic variation in these populations has been shaped by demographic events occurring over the last 200,000 years. The dramatic variation in climate, diet, and exposure to infectious disease across the continent has also resulted in novel genetic and phenotypic adaptations in extant Africans. This review summarizes some recent advances in our understanding of the demographic history and selective pressures that have influenced levels and patterns of diversity in African populations.

Introduction

Modern humans evolved in Africa around 200 thousand years ago (kya), and have lived continuously on the African continent longer than in any other geographic region. Africa not only has the highest levels of human genetic variation in the world but also contains a considerable amount of linguistic, environmental and cultural diversity. For example, more than 2000 distinct ethno-linguistic groups, representing nearly a third of the world's languages, currently exist in Africa (http://www.ethnologue.com/) (Figure 1). Africans live in a wide range of environments, such as deserts, tropical rainforests, savannas, swamps, and mountain highlands [1,2]. Furthermore, some of these environments have undergone dramatic changes over the course of modern human evolution [1,3,4]. African populations also practice a wide array of subsistence strategies, including various forms of hunting-gathering, agriculture and pastoralism, across the continent perhaps in response to this environmental variability over time and geographic space.

African demographic history has consisted of fluctuations in population size, short- and long- range migration, admixture and extensive population structure which have resulted in complex patterns of variation in modern populations [1,5]. The timing and duration of some of these demographic events were often correlated with known major environmental changes and/or cultural developments in Africa [6]. A number of novel genetic and phenotypic adaptations have also evolved in Africans in response to dramatic variation in environment, diet, and exposure to infectious disease across the continent. In some cases, these adaptations have occurred in the last several thousand years, exemplifying the ongoing evolution of human populations. Thus, present-day patterns of variation in African genomes are a product of both demographic and selective events.

The characterization of extant genetic diversity in Africa will be critical for reconstructing modern human origins and African demographic history. In addition, this genetic information, together with phenotype data on variable traits, will be informative for identifying population-specific variants that play a role in gene function, phenotypic adaptation and complex disease susceptibility in Africans and populations of African descent.

Evolutionary History of Modern Humans in Africa

Current paleontological data suggest that the transition to anatomically modern *Homo sapiens* occurred in Africa, supporting the 'Recent African Origin' model of human evolution (Figure 2). The earliest known suite of derived traits associated with anatomically modern humans was identified in fossil remains from East Africa dated to around 195–150 kya [7–9]. Thus, the basic morphology of modern humans was established in Africa about 200 kya [10]. Other early anatomically modern humans, with a more full set of modern features, also appear in Africa before 100 kya and in the Near East around 100 kya [11–13], followed by the more recent expansion of anatomically modern humans into Eurasia within the past 40,000–80,000 years [1,2] (Figure 2). Although the mode of evolution is still unclear, it has been suggested that the emergence of modern humans was not a sudden event, but rather a continuous process of gradual morphological change from archaic to modern *H. sapiens* [11]. However, it has also been argued that modern human origins likely involved episodes of sudden morphological change, leading to the appearance of anatomically modern *H. sapiens* as a species distinct from archaic humans [14]. Regardless of the mode of evolution, current fossil and chronological evidence indicate that modern humans existed in Africa for a relatively long period of time before their migration across much of the globe.

Two main migratory routes out of Africa have been hypothesized for anatomically modern humans. The traditionally favored model involves a northern route of migration via North Africa and the Nile valley into the Levant with subsequent dispersal into both Europe and Asia [15]. Alternatively, an earlier southern coastal route has also been proposed in which modern humans first left Africa by crossing the Bab-el-Mandeb strait at the mouth of the Red Sea and then rapidly migrated along the South Asia coastline to Australia/Melanesia where evidence of human settlement dating to around 55 kya can be found [15–17]. A recent genetic study that correlated levels of microsatellite diversity and the geographic position of sampled populations inferred a waypoint of dispersal of anatomically modern humans out of Africa centered on the Red Sea, strongly supporting an East African origin of migration of modern humans [18]. Although this study was not able to rule out the possibility of multiple migrations out of Africa, prior analysis of autosomal haplotype variability suggests that migration events originating from multiple genetically distinct source populations in Africa are unlikely [19].

[1]Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA. [2]Department of Biology, University of Pennsylvania, School of Arts and Sciences, Philadelphia, PA 19104, USA.
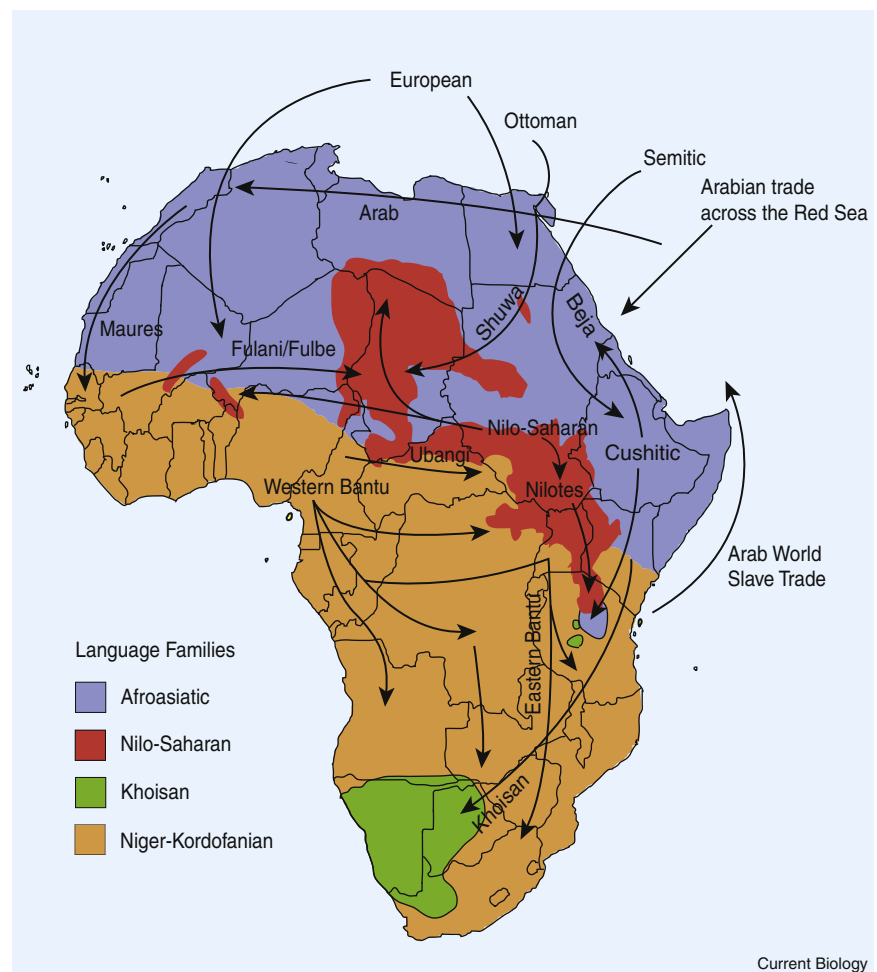*E-mail: tishkoff@mail.med.upenn.edu

Figure 1. Major language families and proposed historic and prehistoric migration events in Africa.

African languages have been classified into four major linguistic families: Niger-Kordofanian, spoken predominantly by agriculturalist populations across a broad geographic distribution in Africa; Afroasiatic, spoken predominantly by northern and eastern African pastoralists and agropastoralists; Nilo-Saharan, spoken predominantly by central and eastern African pastoralists; and Khoisan, a language containing click-consonants and spoken by eastern and southern African hunter-gatherers. This map also shows a number of key migration events that took place in Africa, most notably the expansion of Bantu-speakers from a homeland in Nigeria and Cameroon first into the equatorial rainforests and then into eastern and southern Africa, as well as the migration of Nilo-Saharan-speakers both westward and eastward from a Sudanese homeland. Additionally, Cushitic-speakers from the Ethiopian highlands migrated northward to the Red Sea coast of Sudan (where their modern-day descendants, the Beja, presently live) as well as southward into Kenya and Tanzania [6]. Modified from [2].

The geographic expansion of a small number of anatomically modern humans out of Africa resulted in a population bottleneck. The size of the ancestral population(s) that left Africa is estimated to be around 1000 effective founding males and females based on autosomal microsatellite loci [20] or around 1500 effective founding males and females based on combined mtDNA, Y-chromosome, and X- chromosome re-sequencing data [21]. Recent studies comparing X-chromosome and autosomal diversity also suggest that the effective population sizes of founding males and females were not equal due to sex-biased migration. However, whether the sex-ratio of migrating individuals was male-biased or female-biased is currently under debate [22–24].

Patterns of Genetic Variation in Africa

Genome-wide data have indicated higher levels of genetic diversity in Africans compared to non-Africans, confirming the results of previous mtDNA, X-chromosome, and Y-chromosome studies [1,25–33]. For example, a survey of 1327 nuclear microsatellite, insertion/deletion (INDEL) and single nucleotide polymorphism (SNP) markers showed that African populations, as well as African Americans, have the highest levels of within population genetic diversity relative to non-Africans [18]. In addition, more private alleles are present in Africa than in other geographic regions. The high level of genetic diversity in Africa is consistent with a larger long-term effective population size (N_e) which is estimated to be around 15000 for Africans (and around 7500 for non-Africans) based on re-sequencing data from a 10-kb autosomal non-coding region of the genome [34]. However, a recent study of 98 kb of sequence data from 20 loci on the X chromosome indicated that effective population sizes of individual populations may be smaller (ranging from 2300 to 9000 for African populations and from 300 to 3300 in non-African populations) [35]. Non-African populations also appear to have fewer private alleles and a subset of the genetic diversity present in sub-Saharan Africa, as expected under the 'Recent African Origin' model [1] (Figure 2).

A number of recent studies have also identified a considerable amount of structural variation in the human genome. 'Structural variation' commonly refers to genomic alterations that may be as small as a single nucleotide (excluding SNPs) or as large as millions of nucleotides in size, such as INDELs, copy number variation, translocations and inversions [36,37]. Although the population genetics of structural variation are still in their infancy, information regarding the distribution of such variation within and between populations is emerging [1,29,38–40]. For example, a recent phylogenetic analysis based on 396 non-singleton copy number variants demonstrated that globally diverse populations clustered roughly by geographic region [29], and a study of 67 common copy number variants in the HapMap populations also noted that average F_{ST}, a classical measure of population divergence, of these loci was 0.11, comparable to previous estimates based on nucleotide and haplotype data [1,38,40]. Overall, results from these different studies suggest that levels and patterns of copy number variation in global populations are influenced by demographic history. However, technical limitations associated with a number of
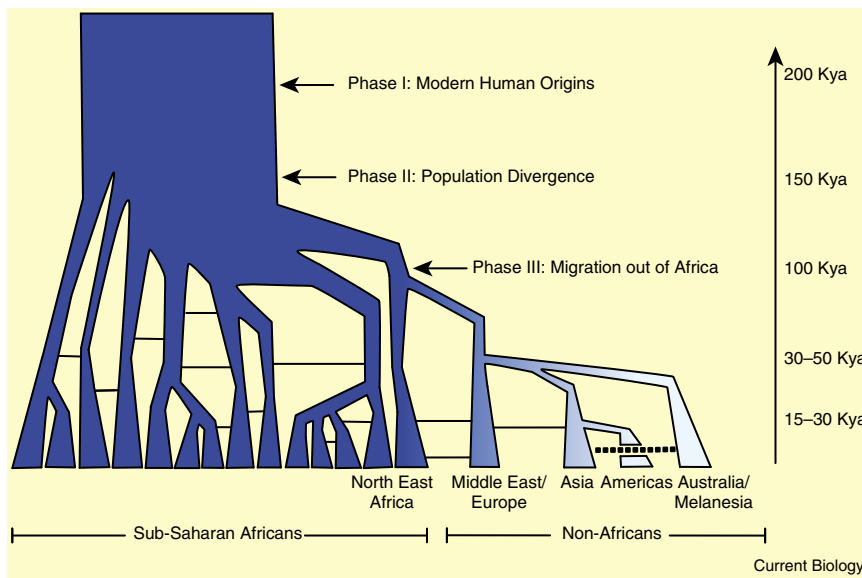
Figure 2. The Recent African Origin model of modern humans and population substructure in Africa.

A study of genome-wide polymorphic markers in 121 ethnically diverse African populations indicated the presence of 14 genetically distinct ancestral population clusters in Africa [18]. According to the Recent African Origin model of modern human origins, anatomically modern humans evolved in Africa around 200 kya, migrated to Eurasia within the last 40000–80000 years and then migrated to the Americas within the last 15000–30000 years. The geographic expansion from Africa is thought to have been accompanied by a population bottleneck and a concomitant loss of genetic diversity. Studies have also suggested that a serial founder model of migration occurred in the history of non-Africans in which the migration of populations across much of the globe occurred in many small steps, with each migration event involving a sampling of variation from the previous population [25,29,33,46]. In this figure, decreasing intensity of color represents the concomitant loss of genetic diversity as populations migrated in an eastward direction from Africa. Solid horizontal lines indicate gene-flow between ancestral human populations and the dashed horizontal line indicates recent gene-flow between Asian and Australian/Melanesian populations. Modified from [1].

Population Structure

Several studies have indicated that ancestral populations were geographically structured before modern humans migrated out of Africa [1]. For example, it has been suggested that the deep coalescence times of mtDNA [41,42] and X-chromosome [43,44] lineages are consistent with a demographic scenario of ancient population structure in Africa. A recent analysis of cranial shape variability in anatomically modern human fossils (dating to 200–60 kya) from Africa and the Middle East also reported a high level of morphological divergence among these fossil hominids which was interpreted as evidence for ancestral population structure in Pleistocene Africa [45]. Thus, arguably a considerable amount of genetic and phenotypic diversity may have been present at an early stage of modern human evolution.

Studies of global population structure in samples from the Human Genome Diversity Panel (HGDP-CEPH) also identified substructure in Africa, particularly between hunter-gatherers and other Africans [29,46]. However, because the HGDP-CEPH contains a small number of African populations, some of which are likely to share recent common ancestry, these results do not reflect the full extent of population structure in Africa [1]. A recent genome-wide study of a much larger set of diverse Africans detected more extensive population structure within Africa than had been previously observed. Specifically, an analysis of 121 geographically diverse populations revealed the presence of 14 genetically distinct ancestral population clusters in Africa [18]. Each cluster consisted of populations that shared genetic

methods commonly used for structural variation detection [36,37], including bias in structural variation ascertainment, can possibly distort estimates of demographic parameters. Advances in next-generation sequencing and other technologies, as well as the identification of structural variation in ethnically diverse Africans, will be needed to better determine patterns of structural variation in Africa and their possible role in phenotypic variability.

similarity, as well as cultural and/or linguistic properties (for example, Pygmies, Khoesan-speaking hunter-gatherers, Bantu-speakers, Cushitic-speakers). Thus, there is a general correlation between genetic relatedness and linguistic and/or cultural similarity in Africa, although some exceptions exist. This study [18] also observed fine-scale genetic structure among populations speaking languages that belong to the same linguistic family, such as between western African and East African Bantu Niger-Kordofanian-speakers, among other examples [18]. East Africa was characterized by high levels of population differentiation, likely resulting from the historical migration of linguistically distinct populations into this geographic region and the long-term presence of indigenous East African Khoesan-speakers, namely the Hadza and Sandawe [18]. A pattern of strong genetic differentiation was also observed in East Africa based on mtDNA data [47]. Additionally, the above genome-wide analysis [18] in geographically diverse Africans showed that Central African Pygmies share common ancestry with several southern and eastern African Khoesan-speaking populations, who speak a language with click consonants, suggesting that these contemporary hunter-gatherers may have descended from a proto-Khoesan-Pygmy population of hunter-gatherers that diverged more than 35 kya [18]. This finding also raises the intriguing possibility that the original language spoken by African Pygmies, who are known to have lost their indigenous language, may have contained click consonants [18].

Several studies have also identified genetic structure within Central Africa, particularly between Pygmy and non-pygmy populations [18,31,48–51], as well as subtle substructure among Pygmies [18,48,50,51]. Specifically, data have shown that the inferred ancestors of modern Pygmy hunter-gatherers and Bantu-speaking agriculturalists could have diverged as long as 70 kya [49], and that ancestral western and eastern Pygmy populations separated more than 18 kya [48] with subsequent genetic differentiation among the western Pygmies within the past 2800 years [50]. The subtle structure among western Pygmies may be

due to recent geographic isolation, genetic drift, and differential levels of admixture between Pygmies and neighboring Bantu-speaking populations [18,49,50,52–54]. Overall, these patterns of genetic and phenotypic variation suggest that African populations have maintained a large and subdivided population structure throughout much of their evolutionary history. This population subdivision may have been facilitated by a number of factors, including ethnicity, culture, language, geography, as well as past fluctuations in geology and climate which may have affected population growth, contraction, fragmentation, and gene-flow in Africa [18,55].

Migration and Admixture

Both short- and long-range migration events, and subsequent admixture between migrating and indigenous populations have also influenced the current genetic landscape of sub-Saharan Africa. One of the most significant migration events in recent African history was the geographic expansion of the Bantu Niger-Kordofanian-speakers from Nigeria and Cameroon first into the rainforests of equatorial Africa, and then into eastern and southern Africa within the past 5000 years [18,56] (Figure 1). Indeed, the presence of Bantu Niger-Kordofanian ancestry in many African populations and the widespread distribution of Bantu-related languages are signatures of the historical migration of Bantu-speakers across Africa and their subsequent admixture with other indigenous populations [18,57]. Genetic evidence also indicates that independent waves of migration of western African and East African Bantu-speakers into southern Africa occurred, in agreement with previous linguistic and archaeological studies [18,58]. Although the reasons for the radiation of Bantu-speakers across sub-Saharan Africa are not entirely known, it has been suggested that a shift from a humid to a drier climate around 5000 years ago and the adoption of new crops suited to this drier environment may have contributed to the widespread movement of Bantu farmers throughout Africa [58,59].

Genetic signatures of both historic and prehistoric migration events are also observed in other regions of Africa [18,42,57,60–62]. For example, an analysis of microsatellite, INDEL and SNP polymorphisms in the nuclear genome showed that populations from central/southern Sudan, such as the Nuer and Dinka, have the highest proportion of Nilo-Saharan ancestry, with decreasing frequency observed in populations from northern Kenya to northern Tanzania in East Africa. These data suggest a Sudanese origin of Nilo-Saharan-speaking populations, with subsequent migration(s) southeastward to East Africa [18]. In addition, Nilo-Saharan-speakers from Sudan, Tanzania, Kenya and Chad also clustered closely with Afroasiatic Chadic-speaking populations from the southern Lake Chad Basin in genetic structure analyses, suggesting that these Chadic-speakers of Nilo-Saharan ancestry likely migrated westward from a Sudanese homeland to Lake Chad and adopted an Afroasiatic language at some point in their history without significant genetic exchange [18]. This shift in language may have occurred through interactions with proto-Chadic Afroasiatic-speakers who migrated from the central Sahara to the Lake Chad Basin around 8 kya [6,56,58,63]. These genetic data are in general agreement with archaeological and linguistic studies that advocate a common origin of Nilo-Saharan populations in eastern Sudan, with subsequent migration events northward to the eastern Sahara, westward to the Chad Basin, and southeastward into Kenya and Tanzania [6,64].

The migration of Nilo-Saharan-speakers may have been associated with past changes in environmental conditions. For example, archeological data suggest that following a climatic shift from dry to more humid conditions in the early Holocene around 10.5 kya, several Nilo-Saharan-speaking populations expanded westward from the Middle Nile Basin to Lake Chad and southeastward to northern Kenya to exploit newly created aquatic food resources [6,65]. However, a small number of Nilo-Saharan-speakers, collectively referred to as Northern Sudanians, migrated northward to the eastern Sahara where they engaged in cattle domestication [6,65] (Figure 3). During the Mid-Holocene arid phase around 8.5–7.5 kya, aquatic food collecting Nilo-Saharans continued their mode of subsistence mainly along permanent rivers, such as the middle Niger and the Nile, and remaining lakes such as Lake Chad. However, during the return of more humid conditions after 7500 years ago, Northern Sudanic cattle raisers expanded westward across the Sahara [6,56,65] (Figure 3). A subset of Nilo-Saharans, namely Nilotic pastoralists, originating from eastern Sudan is also known to have migrated southeastward to Kenya and northern Tanzania within the past 3000 years [64].

Additionally, many Nilo-Saharan-speaking populations from East Africa have high levels of Afroasiatic Cushitic ancestry (likely of Ethiopian origin [64]), suggesting a long history of gene-flow between Nilo-Saharans and Cushites [18]. Archaeological studies have indicated a shared cultural practice of cattle herding between Nilo-Saharans and Cushites in northern Kenya which likely brought these linguistically distinct populations into repeated contact with one another over the past several thousand years [6,56]. Cushitic agropastoralists are also thought to have expanded into southern Kenya and central northern Tanzania, engaging in cultural exchange and inter-marriage with southern and eastern Nilotic pastoralists [6] (Figure 1). This archaeological and genetic evidence of gene-flow between Nilo-Saharans and Cushites is consistent with other genetic data that showed the shared presence of an East African-specific mutation associated with lactose tolerance in these linguistically distinct populations [66].

The reverse migration of non-Africans into Africa was also shown to contribute to the gene-pool of modern African populations. For example, high levels of both Middle Eastern/European and eastern African Cushitic ancestry were detected in the Saharan African Beja, indicative of possible gene-flow from non-African populations [18]. These genetic patterns correlate well with linguistic and archaeological data that suggest that modern-day Beja pastoralists descended from northern Cushitic-speakers who migrated from Ethiopia to the Red Sea coast of Sudan [56]. Furthermore, the Beja have also had more intensive contact with the Middle East through commercial trade across the Red Sea as early as the 9th century A.D. and with nomadic camel herders of Arab Bedouin origin who settled in Sudan beginning in the 14th century A.D. [6] (Figure 1). These studies demonstrate that migration in Africa occurred at different points in time and over a range of geographic areas, resulting in complex patterns of genetic variation.

Dietary Adaptations

Natural selection — the process by which favorable heritable traits become more common in successive generations and
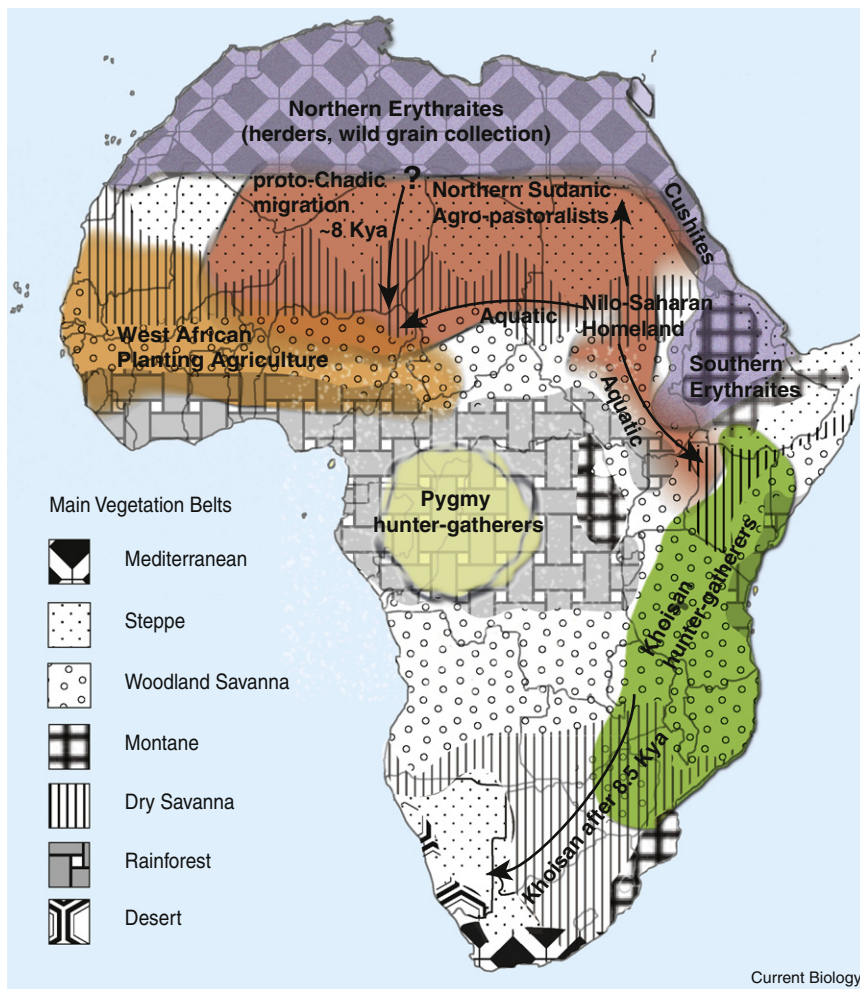
Figure 3. Map of African climate, vegetation and culture in the early Holocene.

Following a period of dramatic increase in rainfall around 10.5 kya, the expansion of both rainforest vegetation in the Congo Basin and woodland savanna farther north and south of the rainforest belt, as well as the expansion of tropical steppe and grassland vegetation into the Sahara region took place. New lakes also appeared, while old lakes grew in size along the southern edges of the Sahara [6]. Between 10.5 and 8.5 kya, several key demographic events occurred among culturally distinct ancestral African populations, including the migration of Nilo-Saharan-speakers from their homeland in Sudan northward to the dry tropical steppe vegetation of the eastern Sahara where they practiced agropastoralism [6]. Northern Erythraites (Afroasiatic-speakers) also lived mainly in the Mediterranean climate of the northern Sahara where they engaged in herding and wild grain collection [6]. The distributions of these cultural groups, among others depicted on this map, roughly correspond to the present-day distribution of major linguistic groups in Africa (Figure 1). The shift from dry to wetter climatic conditions in the early Holocene likely provided an opportunity for ancestral Africans to develop cultural traditions, such as agriculture and herding, in a more favorable environment leading to further technological developments in Africa [6]. Adapted from [6].

unfavorable heritable traits become less common [67] — operates to either increase or decrease the frequency of mutations that have an effect on an individual's fitness. Selectively advantageous mutations associated with diet have evolved in populations, which likely enabled ancestral humans to adapt to their environment, including changes in cultural practices. This section will focus on a few case studies of dietary adaptation primarily in African populations.

Lactase persistence, the ability to digest fresh milk and other dairy products into adulthood, varies in frequency in different human populations. However, the lactase persistence trait is common in pastoralist and dairying populations, such as northern Europeans, and certain African and Arabic nomadic groups [68]. Previous studies identified the T allele at a C/T SNP located 13,910 bp upstream of the lactase gene (*LCT*) as the likely causal mutation of lactase persistence in Europeans and a few West African pastoralist populations, such as the Fulani. However, this mutation was not found to be a strong predictor of lactase persistence in other African populations that practice pastoralism [69,70]. A recent study [66] of 43 populations from Tanzania, Kenya, and the Sudan identified three polymorphisms (G/C-14010, common in Tanzania and Kenya, T/G-13915 and C/G-13907, common in northern Sudan and Kenya) located around 14 kb upstream of *LCT* that are significantly associated with lactase persistence in East African populations. Further analysis revealed evidence consistent with

a recent ongoing selective sweep over the past 3000–7000 years [66]. Recent re-sequencing studies have also found additional variants within this genomic region upstream of *LCT* in other African populations [71, 72]. However, the functional impact of these polymorphisms on lactose tolerance has not been firmly established.

Genetic data have indicated that past migration events may have resulted in the presence of shared genetic variation among pastoralist populations in different regions of Africa. For example, a survey of SNPs associated with lactase persistence in Africa found the C-14010 mutation at low frequency in several southern African Bantu-speaking pastoralists [73]. This study argued for the spread of this mutation into southern Africa by migrating Khoe pastoralists who may have admixed with migrating Nilotic or Cushitic herders from East Africa, with subsequent admixture occurring between Khoe and Bantu-speakers in southern Angola around 2 kya [73]. Similar results have been observed in other East African and southern African populations (A. Ranciaro, S.A. Tishkoff, unpublished data), [72]. Recent Y-chromosome evidence has also suggested that eastern African Nilotic pastoralists migrated to southern-central Africa and admixed with local populations within the last few thousand years, supporting the hypothesis of an East African origin of southern African pastoralism [60]. Overall, these findings are a striking example of convergent evolution, local adaptation due to strong selective pressure resulting from shared cultural practices in Europeans and Africans, and gene-culture co-evolution [66,74].

Similarly, starch consumption is variable among populations with different subsistence patterns, but is more notably elevated in agricultural and certain hunter-gatherer groups [75]. It has been suggested that a genetic adaptation arose in distinct populations in response to an increase in dietary starch intake. Perry and colleagues [75] examined the number of copies of the salivary amylase gene (*AMY1*), which plays a role in starch hydrolysis, in distinct populations with contrasting levels of starch consumption, including several African populations. This study observed a greater number of *AMY1* gene copies in populations with a high starch diet (consisting of European Americans, Japanese, Hadza (African)) compared to populations with a low-starch diet (consisting of Biaka and Mbuti Pygmies (African), Datog (African), and Yakut (Asian)), arguing that dietary differences rather than ancestry influenced gene copy number. Perry and coworkers [75] suggest that positive selection has been the dominant force affecting the divergence in gene copy number among populations at the *AMY1* locus. However, targeted re-sequencing and long-range haplotype analyses of the *AMY1* gene in additional populations with distinct subsistence patterns, including culturally diverse Africans, might be informative for confirming signatures of recent selection, providing further evidence for adaptive evolution at this locus.

Another important dietary adaptation is bitter taste perception which may have evolved in humans to prevent the ingestion of plant toxins. The ability to taste the bitter synthetic compound phenylthiocarbamide (PTC) is a highly variable trait in humans, and is correlated with both the ability to taste naturally bitter substances in food and food preference [76,77]. A large proportion of the phenotypic variance in PTC sensitivity has been attributed to genetic variability at *TAS2R38*, a bitter taste receptor gene, located on chromosome 7 [78]. Specifically, three amino acid substitutions at *TAS2R38* that form two common amino acid haplotypes have been shown to influence PTC sensitivity: a dominant taster haplotype, PAV and a non-taster haplotype, AVI [1]. A genetic analysis of *TAS2R38* in geographically diverse populations, including a small number of Africans, uncovered evidence of balancing selection at this locus, such as an excess of intermediate-frequency variants and an ancient divergence between the major taster and non-taster haplotypes (estimated to have taken place 1.5 million years ago) [79]. These patterns of diversity suggest that variants at this locus are selectively advantageous, likely playing a role in food choice [77,79–81], and may represent long-term genetic adaptations in humans. More recently, genetic data have also shown that additional haplotypes associated with a broader range of PTC sensitivity are present in culturally and linguistically distinct African populations (M.C. Campbell and S.A. Tishkoff, unpublished data), raising the possibility that a wider range of sensitivity to naturally bitter compounds among Africans might be advantageous in Africa. These studies provide further information regarding the evolution of genetic variation associated with bitter taste perception and the role that genetic/phenotypic variability may play in dietary preference in human populations.

Future Directions
To date, only a fraction of the many ethno-linguistic groups in Africa has been extensively studied for genome-wide variation. Analyses of genetic diversity in more geographically and ethnically distinct African populations will be critical for testing models of modern human origins and dispersal out of Africa, as well as for inferring African demographic history. With the recent development of both "next generation" sequencing technology and methods for targeted sequencing, together with their rapidly decreasing costs, it may be feasible to re-sequence large portions of the genome, and conceivably the entire genome [82], in diverse African populations to more accurately test evolutionary models and to infer demographic history. The integration of archaeological, linguistic, paleoclimatic, and geographical information with these genetic data will be particularly useful in this reconstruction, providing a more unified account of historic and prehistoric events in Africa. Indeed, the '1000 genomes' project, which aims to discover novel variants through targeted and whole-genome re-sequencing in a subset of the extended HapMap populations, represents a first step in identifying novel genome-wide variation in more diverse Africans. However, given the extensive population structure present in Africa, additional linguistically and culturally distinct Africans will likely be needed to more fully characterize levels and patterns of variation which can then be used to infer population history.

Finally, African populations have been severely underrepresented in studies of genetic adaptation. Given that Africans possess high levels of genetic and phenotypic diversity, and live in distinct environments, it is likely that ethnically and geographically diverse populations in Africa have undergone local adaptation. To gain a better understanding of genetic and phenotypic adaptations in human populations, it is imperative to include a wider range of ethnically diverse African populations living in distinct environments in genetic studies. Moreover, genome-wide re-sequencing in a number of African populations is likely to lead to the identification of novel variation associated with variable traits. The development of statistical and computational methods for detecting selection across the genome that distinguish the effects of selection and demography will also be critical for identifying the genetic basis of adaptation in Africa. Overall, studies of adaptation combined with information on African demographic history will provide a more robust and accurate view of the processes that have shaped the genomes of African populations.

References
1. Campbell, M.C., and Tishkoff, S.A. (2008). African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. Annu. Rev. Genomics Hum. Genet. 9, 403–433.
2. Reed, F.A., and Tishkoff, S.A. (2006). African human diversity, origins and migrations. Curr. Opin. Genet. Dev. 16, 597–605.
3. Scholz, C.A., Johnson, T.C., Cohen, A.S., King, J.W., Peck, J.A., Overpeck, J.T., Talbot, M.R., Brown, E.T., Kalindekafe, L., Amoako, P.Y., et al. (2007). East African megadroughts between 135 and 75 thousand years ago and bearing on early-modern human origins. Proc. Natl. Acad. Sci. USA 104, 16416–16421.
4. Trauth, M.H., Maslin, M.A., Deino, A.L., Strecker, M.R., Bergner, A.G., and Duhnforth, M. (2007). High- and low-latitude forcing of Plio-Pleistocene East African climate and human evolution. J. Hum. Evol. 53, 475–486.

5. Manica, A., Amos, W., Balloux, F., and Hanihara, T. (2007). The effect of ancient population bottlenecks on human phenotypic variation. Nature *448*, 346–348.

6. Ehret, C. (2002). The Civilizations of Africa: A History to 1800 (Charlottesville: University of Virginia Press).

7. McDougall, I.B.F., and Fleagle, J.G. (2005). Stratigraphic placement and age of modern humans from Kibish, Ethiopia. Nature *433*, 733–736.

8. Stringer, C. (2003). Human evolution: Out of Ethiopia. Nature *423*, 692–693, 695.

9. White, T., Asfaw, B., DeGusta, D., Gilbert, H., Richards, G., Suwa, G., and Howell, F. (2003). Pleistocene Homo sapiens from Middle Awash, Ethiopia. Nature *423*, 742–747.

10. Tattersall, I. (2009). Becoming Modern Homo sapiens. Evolution: Education and Outreach, *Epub Ahead of print*.

11. Brauer, G. (2008). The origin of modern anatomy: by speciation or intraspecific evolution. Evolutionary Anthropol. *17*, 22–37.

12. Niewoehner, W.A. (2001). Behavioral inferences from the Skhul/Qafzeh early modern human hand remains. Proc. Natl. Acad. Sci. USA *98*, 2979–2984.

13. Rightmire, G.P. (2009). Out of Africa: modern human origins special feature: middle and later Pleistocene hominins in Africa and Southwest Asia. Proc. Natl. Acad. Sci. *106*, 16046–16050.

14. Lieberman, D.E., McBratney, B.M., and Krovitz, G. (2002). The evolution and development of cranial form in Homosapiens. Proc. Natl. Acad. Sci. USA *99*, 1134–1139.

15. Mellars, P. (2006). Why did modern human populations disperse from Africa ca. 60,000 years ago? A new model. Proc. Natl. Acad. Sci. USA *103*, 9381–9386.

16. Stanyon, R., Sazzini, M., and Luiselli, D. (2009). Timing the first human migration into eastern Asia. J. Biol. *8*, 18.

17. Quintana-Murci, L., Semino, O., Bandelt, H.J., Passarino, G., McElreavey, K., and Santachiara-Benerecetti, A.S. (1999). Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa. Nat. Genet. *23*, 437–441.

18. Tishkoff, S.A., Reed, F.A., Friedlaender, F.R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J.B., Awomoyi, A.A., Bodo, J.M., Doumbo, O., *et al.* (2009). The genetic structure and history of Africans and African Americans. Science *324*, 1035–1044.

19. Tishkoff, S.A., Dietzsch, E., Speed, W., Pakstis, A.J., Kidd, J.R., Cheung, K., Bonne-Tamir, B., Santachiara-Benerecetti, A.S., Moral, P., and Krings, M. (1996). Global patterns of linkage disequilibrium at the *CD4* locus and modern human origins. Science *271*, 1380–1387.

20. Liu, H., Prugnolle, F., Manica, A., and Balloux, F. (2006). A geographically explicit genetic model of worldwide human-settlement history. Am. J. Hum. Genet. *79*, 230–237.

21. Garrigan, D., Kingan, S.B., Pilkington, M.M., Wilder, J.A., Cox, M.P., Soodyall, H., Strassmann, B., Destro-Bisol, G., de Knijff, P., Novelletto, A., *et al.* (2007). Inferring human population sizes, divergence times and rates of gene flow from mitochondrial, X and Y chromosome resequencing data. Genetics *177*, 2195–2207.

22. Bustamante, C.D., and Ramachandran, S. (2009). Evaluating signatures of sex-specific processes in the human genome. Nat. Genet. *41*, 8–10.

23. Hammer, M.F., Mendez, F.L., Cox, M.P., Woerner, A.E., and Wall, J.D. (2008). Sex-biased evolutionary forces shape genomic patterns of human diversity. Plos Genet. *4*, e1000202.

24. Keinan, A., Mullikin, J.C., Patterson, N., and Reich, D. (2009). Accelerated genetic drift on chromosome X during the human dispersal out of Africa. Nat. Genet. *41*, 66–70.

25. Conrad, D.F., Jakobsson, M., Coop, G., Wen, X., Wall, J.D., Rosenberg, N.A., and Pritchard, J.K. (2006). A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. Nat. Genet. *38*, 1251–1260.

26. Guthery, S.L., Salisbury, B.A., Pungliya, M.S., Stephens, J.C., and Bamshad, M. (2007). The structure of common genetic variation in United States populations. Am. J. Hum. Genet. *81*, 1221–1231.

27. Hinds, D.A., Stuve, L.L., Nilsen, G.B., Halperin, E., Eskin, E., Ballinger, D.G., Frazer, K.A., and Cox, D.R. (2005). Whole-genome patterns of common DNA variation in three human populations. Science *307*, 1072–1079.

28. Hughes, A.L., Welch, R., Puri, V., Matthews, C., Haque, K., Chanock, S.J., and Yeager, M. (2008). Genome-wide SNP typing reveals signatures of population history. Genomics *92*, 1–8.

29. Jakobsson, M., Scholz, S.W., Scheet, P., Gibbs, J.R., VanLiere, J.M., Fung, H.C., Szpiech, Z.A., Degnan, J.H., Wang, K., Guerreiro, R., *et al.* (2008). Genotype, haplotype and copy-number variation in worldwide human populations. Nature *451*, 998–1003.

30. Wall, J.D., Cox, M.P., Mendez, F.L., Woerner, A., Severson, T., and Hammer, M.F. (2008). A novel DNA sequence database for analyzing human demographic history. Genome Res. *18*, 1354–1361.

31. Long, J.C., Li, J., and Healy, M.E. (2009). Human DNA sequences: more variation and less race. Am. J. Phys. Anthropol. *139*, 23–34.

32. Frazer, K.A., Murray, S.S., Schork, N.J., and Topol, E.J. (2009). Human genetic variation and its contribution to complex traits. Nat. Rev. Genet. *10*, 241–251.

33. Ramachandran, S., Deshpande, O., Roseman, C.C., Rosenberg, N.A., Feldman, M.W., and Cavalli-Sforza, L.L. (2005). Support from the relationship of genetic and geographic distance in human populations for a serial founder effect originating in Africa. Proc. Natl. Acad. Sci. USA *102*, 15942–15947.

34. Zhao, Z., Yu, N., Fu, Y.X., and Li, W.H. (2006). Nucleotide variation and haplotype diversity in a 10-kb noncoding region in three continental human populations. Genetics *174*, 399–409.

35. Cox, M.P., Woerner, A.E., Wall, J.D., and Hammer, M.F. (2008). Intergenic DNA sequences from the human X chromosome reveal high rates of global gene flow. BMC Genet. *9*, 76.

36. Scherer, S.W., Lee, C., Birney, E., Altshuler, D.M., Eichler, E.E., Carter, N.P., Hurles, M.E., and Feuk, L. (2007). Challenges and standards in integrating surveys of structural variation. Nat. Genet. *39*, S7–15.

37. Wu, V., and Xiao, H. (2009). Progress in the detection of human genome structural variations. Sci. China C. Life Sci. *52*, 560–567.

38. Consortium, T.I.H. (2005). A haplotype map of the human genome. Nature *437*, 1299–1320.

39. McCarroll, S.A., Kuruvilla, F.G., Korn, J.M., Cawley, S., Nemesh, J., Wysoker, A., Shapero, M.H., de Bakker, P.I., Maller, J.B., Kirby, A., *et al.* (2008). Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat. Genet. *40*, 1166–1174.

40. Redon, R., Ishikawa, S., Fitch, K.R., Feuk, L., Perry, G.H., Andrews, T.D., Fiegler, H., Shapero, M.H., Carson, A.R., Chen, W.W., *et al.* (2006). Global variation in copy number in the human genome. Nature *444*, 444–454.

41. Behar, D.M., Villems, R., Soodyall, H., Blue-Smith, J., Pereira, L., Metspalu, E., Scozzari, R., Makkan, H., Tzur, S., Comas, D., *et al.* (2008). The dawn of human matrilineal diversity. Am. J. Hum. Genet. *82*, 1130–1140.

42. Gonder, M.K., Mortensen, H.M., Reed, F.A., de Sousa, A., and Tishkoff, S.A. (2007). Whole-mtDNA genome sequence analysis of ancient African lineages. Mol. Biol. Evol. *24*, 757–768.

43. Shimada, M.K., Panchapakesan, K., Tishkoff, S.A., Nato, A.Q., Jr., and Hey, J. (2007). Divergent haplotypes and human history as revealed in a worldwide survey of X-linked DNA sequence variation. Mol. Biol. Evol. *24*, 687–698.

44. Yotova, V., Lefebvre, J.F., Kohany, O., Jurka, J., Michalski, R., Modiano, D., Utermann, G., Williams, S.M., and Labuda, D. (2007). Tracing genetic history of modern humans using X-chromosome lineages. Hum. Genet. *122*, 431–443.

45. Gunz, P., Bookstein, F.L., Mitteroecker, P., Stadlmayr, A., Seidler, H., and Weber, G.W. (2009). Early modern human diversity suggests subdivided population structure and a complex out-of-Africa scenario. Proc. Natl. Acad. Sci. USA *106*, 6094–6098.

46. Li, J.Z., Absher, D.M., Tang, H., Southwick, A.M., Casto, A.M., Ramachandran, S., Cann, H.M., Barsh, G.S., Feldman, M., Cavalli-Sforza, L.L., *et al.* (2008). Worldwide human relationships inferred from genome-wide patterns of variation. Science *319*, 1100–1104.

47. Poloni, E., Naciri, Y., Bucho, R., Niba, R., Kervaire, B., Excoffier, L., Langaney, A., and Sanchez-Mazas, A. (2009). Genetic Evidence for Complexity in Ethnic Differentiation and History in East Africa. Ann. Hum. Genet. *73*, 582–600.

48. Patin, E., Laval, G., Barreiro, L.B., Salas, A., Semino, O., Santachiara-Benerecetti, S., Kidd, K.K., Kidd, J.R., Van der Veen, L., Hombert, J.M., *et al.* (2009). Inferring the demographic history of African farmers and pygmy hunter-gatherers using a multilocus resequencing data set. Plos Genet. *5*, e1000448.

49. Quintana-Murci, L., Quach, H., Harmant, C., Luca, F., Massonnet, B., Patin, E., Sica, L., Mouguiama-Daouda, P., Comas, D., Tzur, S., *et al.* (2008). Maternal traces of deep common ancestry and asymmetric gene flow between Pygmy hunter-gatherers and Bantu-speaking farmers. Proc. Natl. Acad. Sci. USA *105*, 1596–1601.

50. Verdu, P., Austerlitz, F., Estoup, A., Vitalis, R., Georges, M., Thery, S., Froment, A., Le Bomin, S., Gessain, A., Hombert, J.M., *et al.* (2009). Origins and genetic diversity of pygmy hunter-gatherers from Western Central Africa. Curr. Biol. *19*, 312–318.

51. Batini, C., Coia, V., Battaggia, C., Rocha, J., Pilkington, M.M., Spedini, G., Comas, D., Destro-Bisol, G., and Calafell, F. (2007). Phylogeography of the human mitochondrial L1c haplogroup: genetic signatures of the prehistory of Central Africa. Mol. Phylogenet. Evol. *43*, 635–644.

52. Coia, V., Caglia, A., Arredi, B., Donati, F., Santos, F.R., Pandya, A., Taglioli, L., Paoli, G., Pascali, V., Spedini, G., *et al.* (2004). Binary and microsatellite polymorphisms of the Y-chromosome in the Mbenzele pygmies from the Central African Republic. Am. J. Hum. Biol. *16*, 57–67.

53. Destro-Bisol, G., Coia, V., Boschi, I., Verginelli, F., Caglia, A., Pascali, V., Spedini, G., and Calafell, F. (2004). The analysis of variation of mtDNA hypervariable region 1 suggests that Eastern and Western Pygmies diverged before the Bantu expansion. Am. Nat. *163*, 212–226.

54. Destro-Bisol, G., Donati, F., Coia, V., Boschi, I., Verginelli, F., Caglia, A., Tofanelli, S., Spedini, G., and Capelli, C. (2004). Variation of female and male lineages in sub-Saharan populations: the importance of sociocultural factors. Mol. Biol. Evol. *21*, 1673–1682.

55. Premo, L.S., and Hublin, J.J. (2009). Culture, population structure, and low genetic diversity in Pleistocene hominins. Proc. Natl. Acad. Sci. USA *106*, 33–37.

56. Blench, R. (2006). Archaeology, Language, and the African Past (New York: Rowman & Littlefield Publishers, Inc.).

57. Berniell-Lee, G., Calafell, F., Bosch, E., Heyer, E., Sica, L., Mouguiama-Daouda, P., van der Veen, L., Hombert, J.M., Quintana-Murci, L., and Comas, D. (2009). Genetic and demographic implications of the Bantu expansion: insights from human paternal lineages. Mol. Biol. Evol. *26*, 1581–1589.

58. Ehret, C. (1998). An African Classical Age: Eastern and Southern Africa in World History, 1000 B.C. to A.D. 400 (Charlottesville: University Press of Virginia).

59. Marret, F., Maley, J., and Scourse, J. (2006). Climatic instability in West Equatorial Africa during the Mid- and late Holocene. Quat. Int. 150, 71–81.

60. Henn, B.M., Gignoux, C., Lin, A.A., Oefner, P.J., Shen, P., Scozzari, R., Cruciani, F., Tishkoff, S.A., Mountain, J.L., and Underhill, P.A. (2008). Y-chromosomal evidence of a pastoralist migration through Tanzania to southern Africa. Proc. Natl. Acad. Sci. USA *105*, 10693–10698.

61. Tishkoff, S.A., Gonder, M.K., Henn, B.M., Mortensen, H., Knight, A., Gignoux, C., Fernandopulle, N., Lema, G., Nyambo, T.B., Ramakrishnan, U., *et al*. (2007). History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. Mol. Biol. Evol. *24*, 2180–2195.

62. Wood, E.T., Stover, D.A., Ehret, C., Destro-Bisol, G., Spedini, G., McLeod, H., Louie, L., Bamshad, M., Strassmann, B.I., Soodyall, H., *et al*. (2005). Contrasting patterns of Y chromosome and mtDNA variation in Africa: evidence for sex-biased demographic processes. Eur. J. Hum. Genet. *13*, 867–876.

63. Cerny, V., Fernandes, V., Costa, M.D., Hajek, M., Mulligan, C.J., and Pereira, L. (2009). Migration of Chadic speaking pastoralists within Africa based on population structure of Chad Basin and phylogeography of mitochondrial L3f haplogroup. BMC Evol. Biol. *9*, 63.

64. Ehret, C. (1982). Population Movement and Culture Contact in the Southern Sudan, c. 3000 B.C. to A.D. 1000 (Nairobi: British Institute in Eastern Africa).

65. Ehret, C. (2006). Linguistic Stratigraphies and Holocene History in Northeastern Africa (Posnan: Posnan Archaeological Museum).

66. Tishkoff, S.A., Reed, F.A., Ranciaro, A., Voight, B.F., Babbitt, C.C., Silverman, J.S., Powell, K., Mortensen, H.M., Hirbo, J.B., Osman, M., *et al*. (2007). Convergent adaptation of human lactase persistence in Africa and Europe. Nat. Genet. *39*, 31–40.

67. Hancock, A., and DiRienzo, A. (2008). Detecting the Genetic Signature of Natural Selection in Human Populations: Models, Methods, and Data. Annu. Rev. Anthropol. *37*, 197–217.

68. Hollox, E.J., Poulter, M., Zvarik, M., Ferak, V., Krause, A., Jenkins, T., Saha, N., Kozlov, A.I., and Swallow, D.M. (2001). Lactase haplotype diversity in the Old World. Am. J. Hum. Genet. *68*, 160–172.

69. Enattah, N.S., Sahi, T., Savilahti, E., Terwilliger, J.D., Peltonen, L., and Jarvela, I. (2002). Identification of a variant associated with adult-type hypolactasia. Nat. Genet. *30*, 233–237.

70. Poulter, M., Hollox, E., Harvey, C.B., Mulcare, C., Peuhkuri, K., Kajander, K., Sarner, M., Korpela, R., and Swallow, D.M. (2003). The causal element for the lactase persistence/non-persistence polymorphism is located in a 1 Mb region of linkage disequilibrium in Europeans. Ann. Hum. Genet. *67*, 298–311.

71. Ingram, C.J., Raga, T.O., Tarekegn, A., Browning, S.L., Elamin, M.F., Bekele, E., Thomas, M.G., Weale, M.E., Bradman, N., and Swallow, D.M. (2009). Multiple rare variants as a cause of a common phenotype: several different lactase persistence associated alleles in a single ethnic group. J. Mol. Evol., *Epub Ahead of print*.

72. Torniainen, S., Parker, M.I., Holmberg, V., Lahtela, E., Dandara, C., and Jarvela, I. (2009). Screening of variants for lactase persistence/non-persistence in populations from South Africa and Ghana. BMC Genet. *10*, 31.

73. Coelho, M., Sequeira, F., Luiselli, D., Beleza, S., and Rocha, J. (2009). On the edge of Bantu expansions: mtDNA, Y chromosome and lactase persistence genetic variation in southwestern Angola. BMC Evol. Biol. *9*, 80.

74. Gerbault, P., Moret, C., Currat, M., and Sanchez-Mazas, A. (2009). Impact of selection and demography on the diffusion of lactase persistence. PLoS One *4*, e6369.

75. Perry, G.H., Dominy, N.J., Claw, K.G., Lee, A.S., Fiegler, H., Redon, R., Werner, J., Villanea, F.A., Mountain, J.L., Misra, R., *et al*. (2007). Diet and the evolution of human amylase gene copy number variation. Nat. Genet. *39*, 1256–1260.

76. Garcia-Bailo, B., Toguri, C., Eny, K.M., and El-Sohemy, A. (2009). Genetic variation in taste and its influence on food selection. Omics *13*, 69–80.

77. Sandell, M.A., and Breslin, P.A. (2006). Variability in a taste-receptor gene determines whether we taste toxins in food. Curr. Biol. *16*, R792–R794.

78. Drayna, D. (2005). Human taste genetics. Annu. Rev. Genomics Hum. Genet. *6*, 217–235.

79. Wooding, S., Kim, U.K., Bamshad, M.J., Larsen, J., Jorde, L.B., and Drayna, D. (2004). Natural selection and molecular evolution in PTC, a bitter-taste receptor gene. Am. J. Hum. Genet. *74*, 637–646.

80. Kim, U., and Drayna, D. (2004). Genetics of individual differences in bitter taste perception: lessons from the PTC gene. Clin. Genet. *67*, 275–280.

81. des Gachons, C.P., Beauchamp, G.K., and Breslin, P.A. (2009). The genetics of bitterness and pungency detection and its impact on phytonutrient evaluation. Ann. N Y Acad. Sci. *1170*, 140–144.

82. Mardis, E.R. (2008). Next-generation DNA sequencing methods. Annu. Rev. Genomics Hum. Genet. *9*, 387–402.