Ancient DNA has emerged as a powerful tool for investigating the human past and reconstructing the movements, mixtures, and adaptations that have structured genetic variation throughout human history. While the study of genome-wide ancient human DNA was initially restricted to regions with temperate climates, methodological breakthroughs have now extended the reach of ancient DNA analysis to parts of the world with hot and humid climates that are less conducive to biomolecular preservation. This includes Africa, where people harbor more genetic diversity than can be found anywhere else on the planet, reflecting deep and complex population histories. Since the first ancient African genome was published in 2015, the number of individuals with genome-wide data has increased to nearly 200, with greater coverage of diverse geographical, temporal, and cultural contexts. Ancient DNA sequences have revealed genetic variation in ancient African foragers that no longer exists in unadmixed form; illuminated how local-, regional-, and continental-scale demographic processes associated with the spread of food production and new technologies changed genetic landscapes; and discerned notable variation in interactions among people with distinct genetic ancestries, cultural practices, and, likely, languages. Despite an increasing number of studies focused on African ancient DNA, multiple regions and time periods have yet to be explored. Research to date has primarily focused on the past several thousand years in eastern and southern Africa, setting up northern, western, and central Africa, as well as deeper time periods, as key areas for future investigation.

As ancient DNA research becomes increasingly integrated with anthropology and archaeology, it is advantageous to understand the basic methodological and analytical techniques, the types of questions that can be investigated, and the ways in which the discipline may continue to grow and evolve. Critically, the growth and evolution of ancient DNA research must include attention to the ethics of this work, both in African contexts and globally. In particular, it is essential that research is conducted in a way that minimizes the potential of harm to both the living and the dead. Scientists conducting ancient DNA research in Africa especially must also contend with structural challenges, including a lack of ancient DNA facilities on the continent, the extensive fragmentation of African heritage (including ancient human remains) among curating institutions worldwide, and the complexities of identifying descendant groups and other stakeholders in the wake of colonial and postcolonial disruptions and displacements. Ancient DNA research projects should be designed in a way that contributes to capacity building and the reduction of inequities between the Global North and South to ensure that the research benefits the people and communities with connections to the ancient individuals studied. While ensuring that future studies are rooted in ethical and equitable practices will require considerable collective action, ancient DNA research has already become an integral part of our understanding of African population history and will continue to shape our understanding of the African past.

**Keywords:** aDNA, archaeogenetics, genetic diversity, African archaeology, Holocene, Later Stone Age, Neolithic, Iron Age, food production, research ethics
Ancient Human DNA and African Population History

Subjects:  Archaeology

A Primer to Ancient DNA

DNA (deoxyribonucleic acid) is the hereditary material present in humans and almost all other organisms. It provides a means for genetic information to be passed from one generation to the next and can be used to trace phylogenetic relationships among organisms. Ancient DNA (aDNA) is DNA from organisms that lived several decades to several hundreds of thousands of years ago, although there is no consensus on a date that serves as the threshold for “ancient.” DNA from ancient humans (or other hominins) can be recovered from diverse biological materials, including soft tissues (e.g., mummified skin, hair) and calcified tissues (e.g., bones, teeth, or dental calculus); it has also been recovered from sediment (Gelbert et al. 2021; Slon et al. 2017; Vernot et al. 2021; Zavala et al. 2021) and even from materials such as clay pipes and birch pitch mastics that were in contact with human biological fluids or tissue (Jensen et al. 2019; Kashuba et al. 2019; Schablitsky et al. 2019). The analysis of aDNA, sometimes called “paleogenetics” or “archaeogenetics” (or “paleogenomics” or “archaeogenomics” when referring to genome-level data), provides a snapshot of genetic variation at a known time and place in the past. Studying aDNA enables the direct evaluation of genetic relationships among past people (or between past and present-day people) and reveals the changes in genetic ancestry that occurred as humans moved to new places, reproduced with new people, and adapted to new environments. Observed genetic changes can then be connected to known processes—ranging from environmental phenomena, such as climate shifts, to human-driven events, such as shifts in subsistence strategies or long-distance trade—to aid in reconstructing the emergence of present-day patterns of human genetic variation.

DNA is a structurally formidable molecule arranged in the form of a spiraling ladder, with side pieces (“backbone”) formed by sugar and phosphate molecules and ladder rungs comprising the repeating chemical bases of adenine (A), thymine (T), cytosine (C), and guanine (G) paired in a specific way (A with T and C with G); however, it begins to degrade postmortem because enzymatic repair processes that maintained the integrity of the molecules during life cease to function (Lindahl 1993). Postmortem alterations—including molecule fragmentation and chemical modification of bases—are a hallmark of aDNA (Orlando et al. 2021). At the time of death, endogenous nucleases (enzymes that act as “molecular scissors” [Nishino and Morikawa 2002] and cleave bonds that make up the DNA backbone) begin to break the nucleotide chain into small pieces that can be millions of times shorter than their original length (some <50 base pairs [bp]; Hofman and Warinner 2019). Subsequently, hydrolytic and oxidation reactions further fragment the DNA backbone and chemically modify the bases (Brotherton et al. 2007; Höss et al. 1996; Lindahl 1993; O’Rourke, Hayes, and Carlyle 2000; Pääbo 1989). A particularly pervasive form of damage in aDNA is the deamination of cytosine (Briggs et al. 2007; Brotherton et al. 2007; Dabney, Meyer, and Pääbo 2013; Hofreiter et al. 2001; Pääbo 1989; Sawyer et al. 2012), a process in which a cytosine molecule is converted to uracil and consequently read by sequencing technologies as a thymine analogue (Briggs et al. 2007). These “C-to-T misincorporations” are drastically clustered at the ends of the DNA molecule relative to the interior of the molecule. This
is because single-stranded DNA overhangs that result from strand fragmentation occur mostly at the ends of DNA molecules and are more susceptible to damage from exposure to water than the double-stranded interior of the DNA molecule (Krause 2010; Overballe-Petersen, Orlando, and Willerslev 2012; Stoneking and Krause 2011). Overall, damage to the DNA molecule accumulates at a rate that is influenced by myriad environmental factors, including temperature (both absolute temperature and temperature stability), exposure to moisture, and the pH of the depositional environment (Kistler et al. 2017; Lindahl and Nyberg 1972; Smith et al. 2003). DNA is best preserved in cold and dry environments with stable temperatures and most poorly preserved in hot, humid places and those with extreme temperature fluctuations (Smith et al. 2003). Particularly important to aDNA analysis is that C-to-T misincorporations can be quantified (Jónsson et al. 2013; Peyrégne and Peter 2020) and used as a method to evaluate the authenticity of aDNA sequences (Krause 2010; Skoglund et al. 2014).

In addition to being invariably damaged, DNA is always present in much lower quantities in ancient biological material than in living organisms. Because there is only a single copy of the genome within a cell’s nucleus (nuclear DNA), a focus of aDNA research for a substantial part of the field’s history has been mitochondrial DNA (mtDNA), circular molecules around 16.5 kb (kilobases) in size that are present in multiple copies in each of the hundreds to thousands of mitochondria found in every cell (meaning that mtDNA is present at a higher copy number in each cell relative to nuclear DNA). mtDNA is passed from mother to offspring in a system of maternal inheritance without recombination (the exchange of genetic material, which makes a new sequence), meaning that an identical copy is passed down untransformed, barring de novo (new) mutations, from one generation to the next. While mtDNA mutates rapidly and is distributed in a geographically patterned way that allows the study of groups of humans (“populations”) in the past through the characterization of “haplotypes” (a group of mutations inherited together from a single parent) and “haplogroups” (a group of haplotypes inherited from a single parent that share a common ancestor, used in reference to both mtDNA and Y chromosome lineages), analyses of mtDNA illuminate only one line of maternal descent and, as a single genetic marker, can be used to reconstruct only one phylogenetic tree. This means that there is a limited amount of information about population history that can be garnered by studying mtDNA alone (Schlebusch et al. 2021). Herein lies the power of genome-wide DNA, where the parallel analysis of many genomic positions (and thus of independent phylogenetic trees) means that analyses gain several orders of magnitude in statistical power (Kelleher et al. 2019) and a more complete, detailed, and reliable account of population history is produced.

Of great importance for population studies that use genome-wide data are single positions throughout the genome where there is variation in the DNA sequence among individuals. At these positions, known as single nucleotide polymorphisms (SNPs, pronounced “snips”), a substantial proportion of people do not carry the same nucleotide, a term used to describe a component of DNA that contains a base (A, T, C, or G) and sugar and phosphate molecules; instead, there are two or more options (known as “alleles”) for the nucleotide found at this position. To be considered a SNP, the less frequent allele must be present at a rate of usually >1%. Despite all humans being ~99.9% genetically similar, SNPs are found at high frequency across the genome: there is on average at least one SNP for every several hundred to one thousand nucleotide bases
Ancient Human DNA and African Population History

(Cargill et al. 1999), making them the main type of genetic variation in humans. Because of the size of the human genome (~3.2 billion bp) and the low mutation rate of nucleotide base substitutions across the nuclear genome, it is unlikely that the same variant will arise independently in multiple populations, and so the presence of the same allele at any SNP locus in the genomes of two individuals is usually indicative of inheritance from a common ancestor (“identical by descent”). The frequency of each allele at a SNP locus within a population shifts over time and across space due to evolutionary forces, including mutation, genetic drift (which increases genetic diversity), and gene flow (which reduces genetic diversity), as well as natural selection. Using allele frequency-based analysis methods, the genetic distance (i.e., the rate of allele sharing) among individuals or populations can be quantified to investigate among whom and to what extent ancestry is shared and to reconstruct past demographic events—for example, changes in population size or the occurrence of admixture (the combination of DNA from people from genetically distinguishable groups)—that shaped patterns of genetic variation.

Importantly, genome-wide aDNA data from any one person is not just representative of that individual but effectively of hundreds or more of that person’s ancestors as well. This is because the variation in a person’s genome reflects contributions from a large number of lineal ancestors; the analysis of a single individual’s genome-wide data represents the theoretical analysis of a large effective sample size of individuals going back in time (Coutinho, Vicente, and Schlebusch 2020). Additionally, unlike archaeological remains which undergo a slow process of continued degradation once excavated from their depositional environments, the construction of DNA sequencing libraries (a “library” is a collection of DNA fragments that have undergone modifications to make them detectable by sequencing instruments) serves to “immortalize” the DNA molecules extracted from the remains of an ancient individual. Aliquots of these libraries can be used in future research potentially involving different approaches. Importantly, these DNA libraries can also be shared among research groups (if permission to do so has been granted), which greatly reduces the destruction of ancient human remains and enables the application of new technologies in the future.

Human Ancient DNA from African Contexts

Compared to more temperate parts of the world such as much of Eurasia, genome-wide aDNA recovered from African archaeological contexts remains limited, in part due to the challenges of climate-driven DNA degradation (fig. 1). The study of aDNA dates back to 1984 (Higuchi et al. 1984) when a team of researchers extracted DNA from dried the muscle tissue of a quagga. The first successful extraction of DNA from archaeological bone tissue took place in Europe and Japan in 1989–1990 (Hagelberg, Sykes, and Hedges 1989; Hänni et al. 1990; Horai et al. 1989) and mtDNA studies of ancient Africans emerging less than a decade later (Lalueza Fox 1997). However, it is only since the mid-2010s that methodological and technical improvements (reviewed by Orlando et al. 2021) enabled the study of aDNA preserved under suboptimal conditions, including much of tropical Africa.
Ancient DNA science requires a move from bones to base pairs, with all steps carried out in highly specialized aDNA facilities ("clean rooms") to minimize the risks of contamination during the extraction and manipulation of aDNA. Clean rooms are maintained as sterile environments through positive air pressure systems, decontaminated through UV light and chemical cleaning procedures, physically separated from spaces where DNA is amplified (a process by which many identical copies of target DNA are created), and accessed only by trained technicians wearing personal protective equipment (Coutinho, Vicente, and Schlebusch 2020; Fulton 2012; Fulton and Shapiro 2019; Knapp et al. 2012; Pääbo et al. 2004). The process of aDNA research, which ultimately begins during the excavation of human remains, requires the selection of biological material from which DNA will be extracted (guidelines for archaeologists who plan to include an aDNA component to their work are articulated by Bollongino, Tresset, and Vigne 2008; Llamas et al. 2017; Matisoo-Smith and Horsburgh 2012; and Yang and Watt 2005). The great majority of ancient human DNA—particularly where biomolecular preservation is likely to be poor—is recovered from bones and teeth. Particularly important advancements in aDNA research include the identification of the petrous part of the temporal bone (specifically, the cochlea) as a skeletal
element that contains up to >hundredfold more endogenous aDNA than other bones (Gamba et al. 2014; Pinhasi et al. 2015; see also Hansen et al. 2017; Parker et al. 2020) and the development of minimally invasive methods for accessing this skeletal element in a way that is compatible with preserving complete skulls (Sirak et al. 2017). In the absence of a petrous bone, tooth cementum has been shown to be an additional optimal substrate for aDNA preservation (Adler et al. 2011; Damgaard et al. 2015; Hansen et al. 2017), as have the auditory ossicles (Sirak et al. 2019).

Other important methodological developments that have allowed aDNA analysis of remains from African contexts include DNA extraction techniques that maximize the retrieval of short and damaged DNA fragments from highly degraded skeletal material (Essel, Korlević, and Meyer 2021; Gamba et al. 2016; Rohland et al. 2018), sometimes including a predigestion step (Damgaard et al. 2015; Korlević et al. 2015); the treatment of DNA extracts with uracil-DNA-glycosylase (“UDG”) in a way that restricts characteristic aDNA damage to the terminal nucleotides while eliminating it in the interior of the molecules, enabling the same molecules to be used in tests of aDNA authenticity and in population genetic analysis (Rohland et al. 2015); the implementation of target-enrichment approaches that focus sequencing efforts on generating data from the most informative parts of the genome, reducing the amount of sequencing required for adequately powerful population genetics analysis by up to thirtyfold (Fu et al. 2015; Fu, Meyer, et al. 2013; Haak et al. 2015; Mathieson et al. 2015); and the decreasing costs and increasing output of next-generation sequencing (NGS) technologies that generate many times more sequences compared to previous methods based on plasmid vectors and polymerase chain reaction (PCR) and consequently provide more data for analysis (Green et al. 2006; Orlando, Gilbert, and Willerslev 2015; Poinar et al. 2006).

Following the extraction and sequencing of aDNA, raw sequencing data undergoes processing, which involves filtering out data that fall below user-set quality control thresholds, mapping the data against the human reference genome, removing duplicate molecules, and evaluating the authenticity of the isolated DNA.¹ Methods for the latter include quantifying C→T substitutions at the terminal base of the molecule (Jónsson et al. 2013), making estimates of contamination using mtDNA (Fu, Mittnik, et al. 2013; Renaud et al. 2015) and the X chromosome in males (Korneliussen, Albrechtsen, and Nielsen 2014), and examining the breakdown of linkage disequilibrium (the nonrandom association of alleles at two or more SNP loci) to estimate contamination (Nakatsuka et al. 2020). It is essential for aDNA researchers to confirm that authentic endogenous DNA—that is, DNA of the individual of interest—is studied. Newly authenticated aDNA data are then analyzed alongside previously published data to facilitate population history inferences (population genetics analytical methods and commonly used software tools and programs are reviewed by Hofman and Warinner 2019; Orlando et al. 2021).

While new laboratory and computational methods developed over the past decade have opened new doors for aDNA research, there are still limitations associated with methods and tools. For example, while the petrous bone has been shown to be optimal for human aDNA research, the potential for the same element to be used for pathogen studies is limited in comparison to other skeletal elements like teeth (Margaryan et al. 2018). Additionally, aDNA data that show high evidence of modern DNA contamination may require that population genetics analyses are only performed on sequences that have evidence of damage patterns suggestive of authentic aDNA;
However, this reduces the amount of data available and may consequently limit the robustness of the conclusions that can be drawn. Researchers must ensure that the required amount of sequencing data required is available (e.g., at least several tens of thousands of SNPs are required for assessing intracontinental genetic relationships; see Günther and Jakobsson 2019) and must take caution that artifactual results or statistical noise does not skew their conclusions. Some limitations will be mitigated by future innovations (Orlando et al. 2021), and it is therefore essential for all experimental and computational procedures to be recorded in full detail and shared openly.

**Why Study Ancient Human DNA from Africa?**

The growing amount of aDNA data generated largely as a result of these methodological advancements has reshaped the genetic study of the African past, which has otherwise been mostly extrapolated from the DNA of people living in the late 20th and 21st centuries (e.g., Campbell and Tishkoff 2008; Fan et al. 2019; Henn et al. 2011; Hollfelder et al. 2017; Pagani et al. 2015; Pickrell et al. 2012; Petersen et al. 2013; Pickrell et al. 2014; Schlebusch et al. 2012). DNA from living people has been used as a tool for estimating ancient genetic population structures and specific details of transformative events in the deep past. For example, it has been leveraged in attempts to identify a specific African homeland for the origin of modern humans (e.g., Chan et al. 2015; Chan et al. 2019; Henn et al. 2011); however, there are notable pitfalls associated with attempting to resolve such complex questions with these data (Schlebusch et al. 2021), requiring the precise definition and rigorous testing of competing models (Henn, Steele, and Weaver 2018).

Ancient DNA data directly attest to long, complex, and dynamic population histories in Africa, characterized by extensive population movements in both prehistoric and historic times, intricate patterns of mixing and mingling that created cultural and biological mosaics, and adaptation to changing environments. Studies have also revealed the presence of uncharacterized deeply ancient “ghost” populations—populations that no longer exist but that can be identified through the analysis of individuals to whom they contributed DNA—that left a subtle signature in the genomes of their descendants (Durvasula and Sankararaman 2019, 2020; Hsieh et al. 2016; Lachance et al. 2012). Because of a complex demographic past, the patterns of African genetic diversity found in living people are inadequate to fully reconstruct ancient genetic landscapes. Simply put, people living in the late 20th and 21st centuries may not live exactly where their ancestors did, and demographic changes—ranging from population bottlenecks and expansions in response to changing environments to the forced relocations of marginalized people in the recent past—mean that the genomes of living people are not necessarily representative of the full extent of genetic variation that existed further back in time. Furthermore, while archaeological evidence has proven immensely informative about population movements, interactions, and adaptations in the African past, the data oftentimes cannot distinguish between changes in material culture associated with the physical movement of people versus the spread of ideas and cultural diffusion without an associated demographic component. The power of aDNA is in its ability to directly evaluate shifts in the genetic landscape across known periods of archaeologically evident or historically documented cultural change.
Despite the odds being stacked against African aDNA research (e.g., Campana, Bower, and Crabtree 2013), the number of genetic studies of archaeological human remains from Africa is rapidly growing. To further advance the incorporation of aDNA science into research projects based in Africa, it is essential to recognize the types of questions that can be addressed through aDNA research, especially when these results are considered in concert with data garnered from decades of research by anthropologists, archaeologists, linguists, and historians. Given the pace of discovery in the discipline of aDNA, the state of knowledge continues to change; however, there are key themes and long-standing questions crossing both space and time that are particularly pertinent to exploration of the African past using aDNA.

The Genetic Landscapes of African Foragers

Foraging—a way of life organized around gathering plants, hunting or capturing animals from terrestrial or marine environments, and collecting animal products such as honey or eggs—has been the main form of subsistence for nearly all of the human past. In sub-Saharan Africa, foraging was the dominant economic strategy until the spread of food production beginning ~5000 years ago, with some groups continuing these practices until the present day (Barham and Mitchell 2008; Prendergast 2020). Genetic research on contemporary African groups who are presently or were historically foragers—which include some speakers of Khoe languages in southern Africa, Mbuti and Aka in central Africa, and Hadza in eastern Africa (among others)—has revealed that these genetically diverse people carry some of the most distinctive and deeply divergent human lineages in contrast to other Africans and all non–Africans, who trace the great majority of their ancestry to lineages that split more recently (Fan et al. 2019; Gopalan et al. 2019; Henn, Steele, and Weaver 2018; Lachance et al. 2012; Lorente-Galdos et al. 2019; Pickrell et al. 2012; Rito et al. 2019; Scheinfeldt et al. 2019; Schlebusch et al. 2012; Shriner et al. 2018; Tishkoff et al. 2009). While the study of recent and historic foraging groups has revealed deeply branching population structures and complex interactions in the past (including possible genetic contributions from archaic humans; Durvasula and Sankaraman 2020; Hammer et al. 2011; Hsieh et al. 2016; Lachance et al. 2012; Lorente-Galdos et al. 2019), the genetic landscape of ancient African foragers cannot be reconstructed from living people alone. This is because much of the genetic diversity among ancient foragers no longer exists in an unadmixed form: in other words, no studied living person or group is descended solely from any one of these ancient lineages, leaving only a breadcrumb trail of ancient genetic diversity. Fortunately, genome-wide African aDNA research beginning in 2015 established broad strokes patterns of genetic variation among ancient foragers and revealed genetic differences between ancient foragers and food producers, providing a foundation for subsequent work. New research continues to add detail to our understanding of the deep population structure of ancient African foragers and increasingly permits regional reconstructions of how foragers moved throughout the landscape and interacted with each other, how these patterns of movement and interaction were shaped by social, economic, environmental, and other factors, and how foragers engaged with food producers who introduced new technologies and distinct genetic signatures.
Ancient DNA from Foragers in Sub-Saharan Africa

The first published aDNA sequence from sub-Saharan Africa—a mtDNA sequence from a male individual who lived ~2300 years before the present (BP) and practiced marine-based foraging—revealed a previously undocumented lineage of haplogroup L0d2c and provided direct evidence that foragers in southern Africa carried the earliest diverged maternal modern human lineages (Morris et al. 2014). This study provided a glimpse into the power of aDNA to illuminate undiscovered African genetic landscapes and contribute to open debates rooted in archaeology. Only a year later, the first genome-wide aDNA data from Africa was published from an adult male forager who lived ~4500 BP in the highlands of present-day southwest Ethiopia; this individual is sometimes referred to as “Mota” after the location (Mota Cave) where his remains were discovered (Gallego-Llorente et al. 2015), although archaeological work refers to him as “Bayira” (“firstborn” in the Gamo language) (Arthur et al. 2019). This genome provided unprecedented insight into genetic ancestry that had never been documented directly. Unlike many present-day eastern African people, the Mota individual exhibited none of the West Eurasian–related ancestry that was introduced through a back migration to the African continent (“back migration” is used here because this movement back to Africa involved the descendants of people who initially left the continent, possibly tens of thousands of years earlier) that occurred after he lived. “Mota–related” genetic ancestry no longer exists in unadmixed form, although ancestry of this type has been identified in admixed form using aDNA in other parts of eastern Africa, as discussed in the section “Genetic Changes Associated with the Spread of Herding” (Prendergast et al. 2019; Wang et al. 2020). Such ancestry also persists in the genomes of a number of present-day eastern African groups (e.g., as shown by López et al. 2021) and constitutes an important point of comparison in genetic analyses of both past and present-day African people.

Subsequent studies of ancient foragers, particularly in eastern and southern Africa, revealed new details of deep population structures and population dynamics among geographically separated and genetically distinct ancient foragers. A 2017 study—the first to report the genomes of multiple ancient African individuals (sixteen individuals from Tanzania, Kenya, Malawi, and South Africa, the majority from foraging contexts)—showed that ancient African foragers were related in clinal patterns correlated with geography, with neighboring people sharing more genetic similarity to each other than to people who were separated by large geographic distances (Skoglund et al. 2017). This previously undocumented cline of geographically structured forager populations stretching from Ethiopia to South Africa is radically different from genetic landscapes documented after the spread of herding and farming. While a broad interpretation is that forager groups were related in an “isolation–by–distance” pattern influenced by geography, subsequent aDNA research will add greater nuance to our understanding of this pattern.

Eastern Africa in particular has been identified as a nexus of population–level interactions between people with ancestries differentially associated with groups of foragers whose habitation ranges expanded and contracted in relation to changing environmental conditions (Wang et al. 2020). A 2020 study that included three individuals from Kakapel Rockshelter in western Kenya revealed genetic signatures of Mbuti–related (central African forager) ancestry in a ~3900 BP
individual, which might be explained by ephemeral interactions among forager groups with different deeply divergent ancestries whose ranges overlapped when rainforest systems were more extensive during the early Holocene wet phase. The authors postulated the existence of a yet undetected forager population who contributed major amounts of ancestry to present-day San living in southern Africa and also to some eastern African foragers (as observed in a ~3500 BP individual from Nyarindhi, also in western Kenya). An alternate explanation is that San–related ancestry in eastern African foragers reflects an earlier and wider distribution of African foragers stretching from southern to eastern Africa. Additional research will further illuminate the dynamics of the ancient forager landscape and reveal whether ancestral components shared among foragers are the result of ongoing admixture throughout the Holocene.

It has become increasingly clear that there is much to learn about the deep population structure of Africa, and that aDNA analysis of sub-Saharan Africans who lived prior to the spread of food production is a tool well-suited to exploring this topic. Ancient DNA from southern Africa (Schlebusch et al. 2017) showed clear connections between three forager individuals from the sites of Ballito Bay A, Ballito Bay B, and Doonside and present-day southern Khoi–San people, who carry more unique genetic variants and more divergent lineages than any other living groups and plausibly represent the deepest population split seen among modern humans (Barbieri et al. 2013; Behar et al. 2008; Gronau et al. 2011; Henn et al. 2011; Pickrell et al. 2012; Schlebusch et al. 2012; Veeramah et al. 2011). However, all present–day Khoi–San groups have also been influenced by 9% to 30% genetic admixture from people—likely eastern African pastoralists who had ~31% Eurasian–related ancestry and ~69% eastern African ancestry—who migrated southward in the last two millennia (Schlebusch et al. 2017). In contrast, the individual from Ballito Bay A showed no evidence of recent gene flow from herders or farmers, and researchers were able to leverage this unadmixed genome to reestimate the date of the first modern human population divergence to 350,000–260,000 years ago (Schlebusch et al. 2017), coinciding with the identification of anatomically modern humans in the African fossil record (e.g., Grün et al. 1996; Richter et al. 2017). Previous dates of this divergence event—suggested to have occurred 100,000–160,000 BP based on the analysis of short sequence fragments or SNP data from present–day people (Gronau et al. 2011; Veeramah et al. 2011; Schlebusch et al. 2012) and 200,000–300,000 BP based on pedigree studies (Scally and Durbin 2012)—were underestimated due to the poorly understood magnitude and impact of gene flow between southern African foragers and other groups less than 2000 years ago. This demonstrates the power of a single ancient genome sequence to transform our views of human evolution in addition to illuminating aspects of that individual’s life and lived experiences (Pfeiffer, Harrington, and Lombard 2019).

While some DNA research attests to the deepest population split having occurred in southern Africa, deep population structures among ancient African foragers are not completely resolved. Coanalysis of ancient and modern DNA is aiding in revealing the complexity of the deepest diversifications of African lineages, which could have involved repeated gene flow among geographically separate groups. One possible scenario is that gene flow connected ancient southern and eastern Africa to some groups in central and western Africa, such as the ancestors of the Yoruba, more than to others, such as the ancestors of the Mende in West Africa (Skoglund et al. 2017). Alternatively, a “basal West African” lineage—which could represent the earliest known divergence of a modern human lineage that contributed a major proportion of ancestry to
living human groups—may have contributed more ancestry to some western African groups (e.g., Mende) than to others (e.g., Yoruba) (Skoglund et al. 2017). Both models address the finding that genetic differences between the Mende and Yoruba are inconsistent with descent from a homogenous ancestral population isolated from ancient southern Africans. In 2020, sequences from four children buried at Shum Laka in present-day Cameroon further revised perspectives on the genetic history of west-central African foragers (Lipson et al. 2020). Two of the individuals—one of whom carries the most deeply divergent Y chromosome haplogroup documented in living people (A00)—lived ~8000 years ago, while the other two lived ~5000 years later. Their genome-wide ancestry profiles are more like present-day foragers from central Africa (e.g., Baka, Bakola, Bedzan, Aka) than people speaking Niger–Congo languages in present-day western Cameroon, indicating that people presently living in the region do not harbor a large proportion of their ancestry from the population represented by the Shum Laka individuals. Modeling of deep ancestry in this work suggested the presence of four deeply splitting branches: three contribute the primary ancestry to present-day central, southern, and eastern African foragers, while another previously unidentified modern human “ghost” source contributed some ancestry to the Mota individual, with the same lineage (or perhaps a fifth deeply splitting lineage) also contributing some ancestry to present-day people in western Africa. All of these lineages were proposed to have diverged within a very short time span around 250,000–200,000 years ago. Using this model, the Shum Laka individuals harbored ~64% basal West African ancestry (defined here as a separate lineage within the western African clade that diverged before extant western African groups) after this clade received ~10% ancestry from a modern human component that diverged at almost the same point as central and southern African hunter-gatherers (and also contributed ancestry to the Mota individual) and ~2% ancestry from an archaic component that diverged close to the split between Neanderthals and modern humans (Lipson et al. 2020). While some research using present-day populations suggests that the split of southern African foragers from other populations was the earliest divergence event (e.g., Chan et al. 2019; Gronau et al. 2011; Mallick et al. 2016), the analysis of aDNA from the Shum Laka individuals supports the possibility that lineages leading to central African foragers split earlier than (or at least close in time to) those of southern African foragers (Lipson et al. 2020).

Additional data from ancient people, especially from sites in central and western Africa, will help refine our understanding of deep human evolutionary history and enable testing of hypotheses that, to date, are based on relatively limited archaeological and skeletal data (reviewed by Scerri et al. 2018).

**Ancient DNA from North African Foragers**

Until 2018, the ancient genetic history of the region north of the Sahara Desert remained poorly understood; however, aDNA has now also contributed new insight into the ancestry of foragers identified in North Africa’s archaeological record. Key research questions about the human past in this region, though different from the questions that guide research in sub-Saharan Africa, are equally pressing. While many present-day North Africans are genetically more similar to present-day people from the Near East than to those from sub-Saharan Africa, the time depth of this Eurasian genetic connection—although estimated by studies of modern mtDNA lineages to
have been established far in the past (González et al. 2007; Olivieri et al. 2006)—has been a topic of debate. Was this genetic connection established during the back-to-Africa migrations of the Paleolithic 10,000 years ago, or was it the result of more recent migrations of farmers into northeast Africa and westward across the continent during the Neolithic (Barbujani et al. 1994; Henn et al. 2012)?

Genome-wide aDNA data from seven individuals from the Later Stone Age (LSA) Iberomaurusian culture who lived ~15,000 years ago and were buried at Taforalt in Morocco allowed the investigation of when Eurasian–related ancestry entered North Africa. Results illuminated connections between Africa and the Near East that existed prior to the Neolithic transition (van de Loosdrecht et al. 2018), documenting genetic affinity between the Iberomaurusian people at Taforalt and Natufians (Epipaleolithic hunter–gatherers from the Levant). Individuals from Taforalt belong to mtDNA haplogroups U6 and M1, associated with autochthonous northwestern African (Maghrebi) ancestry and consistent with a pre-Holocene back-to-Africa movement (González et al. 2007; Olivieri et al. 2006). In addition to sharing the majority (~63.5% on average) of their ancestry with Natufians, the ancient foragers from Taforalt also received approximately one-third of their ancestry from sub-Saharan Africans, best approximated by components found in both present-day western and eastern Africans. This suggests that this contribution represents ancestry shared by these populations before they diverged >200,000 years ago (van de Loosdrecht et al. 2018).

After demonstrating that gene flow from Eurasia predated the Holocene transition ~12,000–10,000 BP and the development of farming practices shortly thereafter, the authors connected this genetic pattern to an archaeologically known event, speculating that the Natufian–related ancestral population was plausibly the people associated with the Iberomaurusian microlithic bladelet industry that spread across northern Africa and the Near East, beginning by at least 25,000 BP (Barton et al. 2013). The foragers from Taforalt contributed ancestry to Early Neolithic people who lived ~7000 years ago in the same region in present-day Morocco, and an endemic element of ancestry is still detected in present-day Maghrebi populations (and is restricted to present-day populations in this region). Ancient DNA analysis thereby provides outstanding resolution into the long-term genetic continuity in this region, revealing that human populations have been isolated to some extent in the Maghreb since Upper Paleolithic times (Fregel et al. 2018).

**Genetic Changes Associated with the Spread of Herding**

Beginning in the early Holocene, the cultural and genetic landscape of ancient African foragers was drastically altered by the spread of people who practiced herding and farming. Unlike many other parts of the world, the first form of food production in much of Africa was mobile pastoralism, a way of life organized around herding domesticated animals such as cattle, sheep, and goats. The earliest evidence for herding is found ~8000 years ago in the Sahara (Honegger and Williams 2015), with domestic animals spreading throughout sub-Saharan Africa beginning ~5000 BP and reaching the tip of Southern Africa by ~2000 BP (Lander and Russell 2018; Marshall and Hildebrand 2002). The introduction of food-producing technologies had a significant impact...
not only on the cultural traditions associated with foraging lifeways (Prendergast 2020), but also on the forager genetic landscape, as these technologies were spread through the movement of genetically distinct people who variably admixed with local forager groups (Skoglund et al. 2017; Prendergast et al. 2019; Wang et al. 2020). The timing, extent, and nature of interactions between foragers and food producers is a key research question in archaeology that can be explored with aDNA.

Ancient DNA data revealed a complex mosaic landscape in which herders spread into new environments and interacted in diverse ways with autochthonous foragers. A study of ancient individuals from forager, Pastoral Neolithic (PN; ~5000–1200 BP), and Iron Age (~1200 BP to recent years) contexts across Kenya and Tanzania leveraged genome-wide aDNA data to document a multistep spread of food producers into sub-Saharan Africa (Prendergast et al. 2019). This work revealed that the spread of pastoralism involved at least two phases of genetic admixture. The first likely occurred ~6000–5000 years ago in northeastern Africa and involved local groups with ancestry similar to that found in present-day Dinka people and groups with ancestry similar to that of ancient people from the Levant, but which has also been present in northeastern Africa for thousands of years; here, we refer to this latter ancestry as “West Eurasian–related” because we do not have ancient genetic data from an appropriate phylogenetically adjacent reference group from Africa at present. The second phase of admixture may have occurred ~4000 years ago in eastern Africa, between this already admixed group and local foragers associated with LSA foraging traditions in Kenya who had ancestry similar to the Mota individual (Prendergast et al. 2019). In contrast, during the peak development of early herding in eastern Africa (~3300–1200 BP), there was minimal gene flow between foragers and herders, potentially reflecting a shift in social barriers as specialized pastoralism became entrenched.

Subsequent research by Wang et al. (2020) documented additional and previously unknown variation in the amount of forager–related ancestry and Dinka–related ancestry among PN individuals from southern Kenya. Such variation in ancestry proportions may reflect recent or ongoing admixture that prevented genetic homogenization. This finding raises the possibility that periodic admixture between herders and foragers (or between populations with majority proportions of ancestry derived from these groups) may have continued, albeit rarely, well into the PN era. This is best illustrated by two individuals from Molo Cave in Kenya dated to the late PN (~1500 cal BP) who exhibit ~60% forager ancestry with admixture date estimates ranging from a few hundred to a few thousand years before the deaths of these people, suggesting repeated and recent admixture between foragers and herders (Wang et al. 2020). The suggestion that admixture between foragers and herders occurred periodically for thousands of years is consistent with archaeological evidence for ongoing herder–forager interaction (e.g., Prendergast and Mutundu 2009) and forces careful consideration of how population dynamics played out during the PN among genetically and culturally diverse groups. Ancient DNA data attest not only to multiple admixture events, but also to geographic variation; there is less evidence for admixture of pastoralist–related ancestry into forager individuals near lake and ocean coasts, while in the Central Rift, such ancestry spread much more pervasively (Wang et al. 2020).
Ancient DNA has also contributed to advancing long-standing debates of whether the PN herders who left behind different types of material culture (characterized as Elmenteitan and Savanna Pastoral Neolithic traditions) were genetically distinct. It had been proposed on the basis of their distinctive lithic and ceramic traditions, as well as other notable differences such as mortuary practices, that these contemporaneous herders occupying the Rift Valley in southern Kenya and Tanzania represented separate migrations of genetically and potentially linguistically distinct lineages into eastern Africa (Ambrose 2001; Ambrose 1982; Ehret 1984). However, aDNA revealed no genetic differentiation among people associated with either tradition, suggesting that these classifications represent cultural, but not ancestral, differences (Prendergast et al. 2019). All individuals from PN contexts studied to date show the greatest genetic affinity to present-day Afro-Asiatic speakers, a finding that supports the hypothesis linking the initial expansion of pastoralism into eastern Africa and the spread of Afro-Asiatic languages (Ehret 1984), further illustrating the power of aDNA to contribute to the testing of hypotheses generated from other lines of evidence.

In line with previous inferences made using modern DNA (e.g., Henn et al. 2008; Pickrell et al. 2012, 2014; Schlebusch et al. 2012), aDNA data have also contributed to further clarifying the spread of eastern African and West Eurasian–related ancestry into southern Africa in the past ~2000 years, speaking to a decades-long debate about the extent to which herders moved with livestock (summarized by Orton 2015). Ancient DNA analysis of individuals buried in a pastoralist context at Kasteelberg in South Africa revealed admixed eastern African forager and West Eurasian–related ancestry components closely related to a PN-era individual from Luxmanda in Tanzania (Skoglund et al. 2017). Over the long term, this ancestry profoundly changed the southern African gene pool: as discussed in the section “Ancient DNA from Foragers in Sub-Saharan Africa,” all present-day Khoe–San groups have 9% to 30% admixture from Eurasian and eastern African groups introduced after 2000 BP (Schlebusch et al. 2017).

Expanding the geographic reach of aDNA analysis in Africa is critical to improving our understanding of the extent and scale of the genetic change catalyzed by the movements of people associated with shifts toward food production. For example, ancestry modeling of an individual buried in an Iron Age context (~750 BP) in the Democratic Republic of the Congo (DRC) revealed a mixture of Mbuti– or Mota–related ancestry along with PN–related ancestry, the latter documented for the first time in a region previously unsampled for aDNA (Wang et al. 2020). This may reflect the continued expansion of groups with eastern African PN–related ancestry into Central Africa during the Iron Age, possibly following displacement by groups carrying Nilotic– and West African–related ancestry that moved into the Rift Valley after ~2500 years ago. While this hypothesis is only supported by data from a single individual, it demonstrates the potential of aDNA data from new geographic areas to change our understanding of the genetic impacts of the spread of food production.
The Spread of Farming and Its Impact on the African Genetic Landscape

One of the most profound transitions in human history was the shift from a foraging lifestyle to a reliance on farming, a transition accompanied by changes in human health, demography, mobility, social organization, and material culture, to name but a few. In Africa, plant cultivation followed by crop domestication developed independently during the last 5000 years in at least five centers (Fuller and Hildebrand 2013): three centers in distinct ecological zones of western Africa, one in the eastern Sudanic grasslands, and one in the Ethiopian highlands; other centers remain debated. Additionally, crops of southwest Asian origin were introduced to the Egyptian Nile Valley and subsequently to other parts of North Africa after ~7000 BP (Haaland and Haaland 2013). Farming spread throughout Africa from these centers.

Genetic studies have been a valuable tool for investigating the spread of farming across and among continents; this event is particularly well-studied in Eurasia (reviewed by Skoglund and Mathieson 2018; Lazaridis 2018). As in other parts of the world, a key question in Africa is whether farming spread from various centers of domestication with or without the movement of people and if demic diffusion was involved, at what scale and tempo. Additional questions that can be addressed by aDNA include biological adaptations to agricultural lifestyles and changes to social organization, such as marriage patterns. To date, however, the number of studies able to address these questions in Africa has been limited, offering only broad views of demographic changes associated with farming.

North Africa

Compared to other parts of the world, the Neolithic transition in North Africa (and in the “Maghreb” western region in particular) remains largely unresolved. In the Maghreb, there is long-standing debate about the extent to which Neolithic innovations—including ceramics resembling those of Iberia and crops of southwest Asian origin—involved migrating groups from Iberia or elsewhere in the Mediterranean or whether they developed largely within autochthonous populations, with only minimal migration from elsewhere (reviewed by Barton and Bouzouggar 2013). Ancient DNA is an ideal tool to explore this question of demic versus cultural diffusion, with the former most often resulting in detectable changes in the genetic makeup of a population that correlate with a transition in culture and the latter recognizable by long-term genetic continuity spanning cultural transitions.

A 2018 study (Fregel et al. 2018) leveraged aDNA data from people who lived at the Early Neolithic (~7000 BP) site of Ifri n-Amr or Moussa and the Late Neolithic (~5000 BP) site of Kelif el Boroud, both in Morocco, and investigated these data in the context of previous knowledge about the genetic makeup of Iberomaurusian foragers from Taforalt (discussed in the section “Ancient DNA from North African Foragers”; van de Loosdrecht et al. 2018) as well as new data from farmers from the Early Neolithic (~7000 BP) site of El Toro in southern Spain. By combining these data, Fregel et al. (2018) were able to directly test whether the introduction of farming technology was associated with population continuity in North Africa or the movement of people across the Strait.
of Gibraltar. The results highlighted a multidimensional process that involved both cultural transmission and demographic shifts: while genetic continuity between LSA and Early Neolithic populations in the Maghreb indicated that the initial spread of farming occurred through a process of cultural transmission, mtDNA and Y chromosome haplogroups as well as autosomal data revealed either population replacement or at least a transformative genetic influx of people into the Maghreb between the Early and Late Neolithic. In contrast to the Early Neolithic inhabitants of the Maghreb, Late Neolithic individuals harbored both autochthonous Maghrebi and European ancestry components. Specifically, genetic similarities were observed with the Neolithic culture that was present in the Iberian Peninsula >7000 BP. This supports a scenario of demic movement from Europe into North Africa between the Early and Late Neolithic that resulted in genetic variation among groups of people living at these times.

Ancient DNA data indicate a complex process of Neolithization in North Africa that involved both the spread of ideas and the movement of people. That this process is multifaceted is not surprising based on data from studies of present-day North African people in the Maghreb region. These studies identify ancestral components contributed from the Near East, sub-Saharan Africa, and Europe alongside an endemic Maghrebi ancestry that has been present since the LSA and is still detected in some present-day groups (Fadhlaoui-Zid et al. 2013; Font-Porterias et al. 2018; Henn et al. 2012; van de Loosdrecht et al. 2018). In this case, robust aDNA data from time periods spanning both sides of a known change in culture were able to further resolve the nature of the processes and timing of the admixture that resulted in admixed populations who now live in North Africa having a starkly different ancestry makeup than people residing south of the Sahara Desert. However, the integration of archaeological and genetic data spanning millennia of complex demographic processes in northern Africa is still in early days.

Sub-Saharan Africa

Farming was introduced to most of sub-Saharan Africa as a result of the so-called “Bantu expansion,” which began ~5000–3000 years ago in west-central Africa and is associated with the spreads (not necessarily in a single package) of West African–related ancestry, key African crops such as pearl millet, and languages in the Bantu language group, a subset of the Benue-Congo branch of the Niger-Congo language phylum, spoken by almost one-third of the continent’s population (Bostoen 2018). It is widely accepted today that was a demic expansion—that is, it involved the movement of people (Busby et al. 2016; Li, Schlebusch, and Jakobsson 2014; Pakendorf, Bostoen, and de Filippo 2011; Patin et al. 2017). It was arguably the most momentous migration in African prehistory, set apart from other ancient dispersal events by the scale and rapid tempo of human movement, coupled with adaptation to new ecologies.

Although linguists generally accept the Grassfields area of northwestern Cameroon as the homeland of Bantu languages, the aDNA study of individuals buried at Shum Laka in this region, discussed in the section “Ancient DNA from Foragers in Sub-Saharan Africa,” did not offer any further resolution on the genetic origins of Bantu language speakers (Lipson et al. 2020). Furthermore, there is still little consensus regarding the timing and geographical routes followed by people with West African–related ancestry as they expanded through the southern part of the
continent (de Filippo et al. 2012). This process was characterized by parallel and interacting changes that are represented archaeologically by the appearance of metals and metalworking technologies, new types of ceramics, and in some limited cases, archaeobotanical evidence for crops (Crowther et al. 2018; de Maret 2013; Mapunda 2013).

Given the geographic scale and time span involved, scholars studying the Bantu expansion generally consider genetic, linguistic, archaeological, and chronological evidence both independently and jointly (de Maret 2013). Genetic evidence from present-day people, and more recently, from aDNA, is one of the newest tools being used to explore the details of the spread of Bantu-speaking people who carried West African-related ancestry across sub-Saharan Africa. Linguistic data, archaeological information from well-dated sites, and present-day DNA evidence point to a “spread-over-spread” scenario for these migrations (in contrast to a single, long-term, continuous migratory event), with an initial spread of Bantu speakers truncated by a mid-first-millennium CE population collapse, which was subsequently followed by renewed population growth and expansions into eastern and southern Africa (de Filippo et al. 2012; Schlebusch 2019; Seidensticker et al. 2021). The distribution of Bantu-speaking groups in the 21st century reflects this gradual and repeated dispersal of populations from West Africa that displaced or absorbed communities along the way (Berniell–Lee et al. 2009; de Filippo et al. 2012; Pakendorf, Bostoen, and de Filippo 2011; Quintana-Murci et al. 2008; Schlebusch et al. 2012, among many others; although see Sikora et al. 2011). Research has also revealed robust fine-scale genetic structure among Bantu-speaking groups in South Africa that corresponds with linguistic divisions and reflects geography (Sengupta et al. 2021). An open question is the extent to which aDNA data in particular attest to the demographic nature of this expansion and what additional detail the data can provide about the scale and tempo of this spread as well as about possible forager–farmer interactions.

In eastern Africa, archaeological evidence suggests agriculture may have appeared by ~2500 BP, although archaeobotanical evidence is rare and suggests a later introduction (Crowther et al. 2018). Ancient genomes from a ~600–300 BP individual from the site of Munsa, Uganda (Wang et al. 2020) and a ~600 BP individual from the Zanzibar archipelago in Tanzania (Skoglund et al. 2017) exhibited genetic profiles similar to present-day Bantu speakers, with the latter exhibiting even more West African-related ancestry than some present-day Bantu-speaking people from Kenya. Another individual from Deloraine Farm—the earliest Iron Age agricultural site in Kenya’s Rift Valley, dated to ~1200–1000 BP—exhibits shared ancestry with West Africans in both genome-wide ancestry as well as through his Y haplogroup, E1b1a1a1a1a, which is predominantly West African-associated (Prendergast et al. 2019). In South Africa, four individuals from Iron Age contexts dating to ~300–500 years ago derive most of their ancestry from a source related to present-day West Africans (Schlebusch et al. 2017), consistent with archaeological evidence for Iron Age farmers arriving in the eastern part of Southern Africa by ~1700 years ago (Lander and Russell 2018; Mitchell 2013) and confirming that a large-scale population replacement occurred in this part of Africa.
Cultural changes in Iron Age eastern Africa are not exclusively associated with the Bantu expansion. Attesting to this are individuals associated with Pastoral Iron Age (~1200 BP–recent) contexts in Kenya’s Rift Valley who show greater genetic affinity to present-day Nilotic speakers than to individuals from nearby (but earlier) PN contexts (Prendergast et al. 2019). It is postulated that an influx of Nilotic–related ancestry into Kenya around the same time that West African ancestry and farming were introduced (~1200 BP) reflects a separate introduction of iron metallurgy into the region from northeastern Africa, potentially around Sudan and South Sudan. The impact of these demographic movements did not result in a homogenous process of replacement of autochthonous populations. Ancient DNA data directly attest to this in ways that other lines of evidence cannot; for example, documenting the absence of West African–related ancestry in an individual who lived ~900 BP at Kakapel (Wang et al. 2020). This individual—despite living after West African–related ancestry is first documented in the broader region (Prendergast et al. 2019)—has majority ancestry (~88%) from Nilotic–related sources along with the remaining small proportion of ancestry from West Eurasian–related sources such as is found in early PN herders, indicating admixture between PN–related herders and iron–using pastoralists carrying Nilotic–related ancestry. Furthermore, an individual dating to ~400 BP from the coastal cave site of Panga ya Saidi, Kenya can be modeled as having 100% Mota–related ancestry (Skoglund et al. 2017), demonstrating that patterns of farmer dispersal and admixture during the Iron Age resulted in a complex landscape of people with diverse ancestries as opposed to a comprehensive population replacement. Recognizing that there is great heterogeneity in the nature of population change during this time in eastern Africa, additional aDNA studies are needed.

Ancient DNA has the potential to shed light on the extent to which people with West Africa–related ancestry immediately or eventually displaced or absorbed local communities as they spread throughout sub-Saharan Africa. In South Africa, aDNA revealed that ~16% of the ancestry in ancient Iron Age individuals was derived from forager–related admixture (represented by Khoe–San as a proxy), compared to ~19% in present–day southeast Bantu language speakers from South Africa (Schlebusch et al. 2017). Ancient genomes from three individuals from the sites of Xaro and Nqoma in the Okavango Delta region of northwestern Botswana reveal an even greater proportion of forager ancestry (~30%–40%), with these individuals also modeled as having some (~14%–22%) PN–related genome–wide ancestry and some uniparental haplogroups also supporting admixed ancestry (Wang et al. 2020). Importantly, no present–day group has been found to have the same ancestry mix as the two individuals from Xaro, who show evidence of having components of South African forager–, eastern African PN–, and West African–related ancestry, providing the first evidence of a population that no longer exists following replacement by the unadmixed Bantu–speaking populations who presently inhabit the region (Wang et al. 2020). So far, aDNA data are largely consistent with archaeological data that indicate multiple trajectories of interaction and integration among people living in diverse temporal and geographic settings as food production spread into southern Africa (Mitchell 2013).
Evidence of admixture between foragers and farmers is also seen in some eastern and southern African contexts. While modern DNA data have revealed fine-scale variation in interactions between foragers and farmers both temporally and geographically (e.g., Sengupta et al. 2021), the combination of ancient and modern DNA attests to remarkable variation in this regard. Data from present-day Bantu speakers from Kenya (who also derive some ancestry from Nilotic-related and PN-related lineages) revealed that admixture occurred there an average of 800–400 years ago (Skoglund et al. 2017), suggesting that a scenario of initial genetic isolation between Bantu-speaking farmers and autochthonous foragers during the Bantu expansion eventually broke down, paralleling patterns observed in Neolithic Europe (Haak et al. 2015). By contrast, data from ancient and present-day individuals in Malawi reveal a starkly different pattern of interaction. Barriers to admixture were apparently never removed, as evidenced by the absence of any genetic signature of forager populations in present-day Malawians (Skoglund et al. 2017). Indeed, present-day Malawians from the Chewa, Ngoni, Tumbuka, and Yao groups are consistent with having ~100% West African-related ancestry, indicating the complete replacement of forager lineages in Malawi that existed as recently as 2500 years ago and reflecting the highly localized genetic change in this region driven by the initial Bantu expansion and subsequent population movements and demographic changes.

Further study with more precision and added nuance is needed to determine the complex demographic dynamics associated with the expansions of cultural changes, Bantu languages, and West African-related ancestry. Greater emphasis can also be placed on exploring this process at a regional and even local level, as is increasingly done in Europe where data-dense analyses of single cemeteries have been able to point to, for example, kinship and postmarital residence patterns (e.g., Knipper et al. 2017; Mittnik et al. 2019). This is a key theme in studies of southern Africa where the Iron Age origins of a “matrilineal belt” have been much debated (e.g., reviewed by de Luna 2016). Especially given the limited number of ancient genomes from Africa so far, there is danger of using an oversimplified Bantu expansion model uncritically, so great care must be taken to test existing hypotheses with genetic, linguistic, and archaeological data rather than merely reinforcing them (de Maret 2013).

Expanding Ancient DNA Research into New African Frontiers

Since the first genome-level aDNA data from Africa were published in 2015, ancient human remains from some parts of Africa—especially the eastern and southern parts of the continent—have been the subject of multiple aDNA studies, while other regions have garnered far less attention (fig. 1). While geographically variable research histories, research networks, current political conditions, and the availability and accessibility of skeletal material play large roles in these biases, another important factor is environmental conditions, especially heat and humidity. For example, skeletal remains buried in the Sahara Desert and Central African rainforests may experience greater degradation of aDNA molecules relative to other parts of Africa and certainly to other parts of the world. However, methodological improvements that are designed for highly fragmented and damaged aDNA molecules and computational techniques that facilitate the identification and analysis of low amounts of authentic aDNA may yet extend aDNA research to severely understudied parts of Africa. This is an essential step forward in generating a
comprehensive understanding of ancient genetic variation, as many of the regions where DNA preservation is poorest are those regions that are critical for a comprehensive understanding of ancient African genetic diversity.

The Nile Valley

The Nile River Valley in present-day Egypt and Sudan has received an outstanding amount of attention from archaeologists and anthropologists dating as far back as the 19th century, but there remains a paucity of genome-wide aDNA studies focused on the people who once inhabited this region (although studies of uniparental markers in ancient individuals have been carried out; see, e.g., Lalueza Fox 1997; Schuenemann et al. 2017). The study of ancient Egyptian mummies was a key focus during the incipient days of aDNA research (a logical target of study given that mumification often preserves both hard and soft tissue, making wide-ranging analyses possible), with the first reported retrieval of human aDNA from the skin tissue of a mummified child who lived ~2400 years ago (Pääbo 1985). However, the reliability of this analysis has since been questioned due to high levels of modern DNA contamination introduced by the enormous amplifying power of PCR (Willerslev and Cooper 2005). The authenticity of the first aDNA sequences from Egyptian mummies generated using next-generation sequencing technologies (Khairat et al. 2013) was also called into question after being found to be unsupported by authenticity and contamination tests. Challenges associated with sequencing authentic aDNA from mummified remains from this region cast a shadow over the entire field of aDNA and resulted in many researchers nearly abandoning the genetic study of Egyptian human remains altogether (Gilbert et al. 2005; Loreille et al. 2018). The Nile Valley, however, is a critical region for population genetics studies due to its intermediate location between the African and Eurasian continents and its long-standing role as a conduit of trade and an interaction sphere involving goods, and potentially people, from multiple parts of Africa, Asia, and Europe. Until recently, the study of the ancient populations of this region was based primarily on historical sources, archaeological data, and inferences drawn from present-day genetic diversity; however, the landscape of genetic diversity prior to events that led to drastic demographic change, such as the trans-Saharan slave trade and the Islamic expansion, remains poorly understood. There are key questions of pertinence in this region—both on a broad scale (e.g., are present-day people directly descended from people who lived in the same place in the past?) and a specific scale (e.g., how do people within a site relate to each other?)—that are ideal for exploration with aDNA.

Ancient Egypt has long been a particular source of fascination for archaeologists and historians who have sought to trace the affinities of and interactions among Egyptians and people in West Eurasia, as well as with groups south of the Sahara Desert. The first genome-wide aDNA data from the Nile River Valley (from Abusir el-Meleq, Egypt) were generated from two individuals dating to the Pre-Ptolemaic Period (New Kingdom to Late Period) and one dating to the Ptolemaic Period (spanning ~1300 years of Egyptian history) (Schuenemann et al. 2017). Corroborating archaeological discoveries and historical documents, aDNA data attest to a close relationship among ancient Egyptians and people from the Near East. From all sampled ancient individuals worldwide, ancient Egyptians were most closely related to Neolithic and Bronze Age people from the Levant as well as Neolithic Anatolian and European populations. In fact, ancient
Egyptians were genetically more like people from the Near East than from present-day Egyptians, who exhibit ~8% more sub-Saharan African ancestry than the ancient Egyptians studied, indicating an influx of sub-Saharan African ancestry into the region after the Roman Period (previously suggested from analysis of present-day people; Henn et al. 2012). Ancient DNA data suggest that the African gene flow observed in present-day Egyptians occurred predominantly within the last 2000 years. It is plausible that growth in long-distance trade between sub-Saharan Africa and Egypt and accompanying mobility along the Nile, and particularly the trans-Saharan slave trade—which is known to have moved 6–7 million individuals from sub-Saharan Africa to northern Africa over a period of ~1250 years—could have contributed to the increase in sub-Saharan African ancestry observed in living people (Lydon 2009; Wright 2007).

Connections among populations in Upper Egypt and Nubia—a region of the Nile Valley stretching between present-day Aswan in Egypt and Khartoum in Sudan—are attested by archaeological, anthropological, and historical evidence to have been established more than 5000 years ago with political dynamics between Egyptian and Nubian entities ranging from peaceful coexistence to variably successful attempts at occupation (Adams 1977; Edwards 2007). Like Egyptians, present-day Nubian people have been impacted by relatively recent demographic movements; for example, genetic analyses of genotyped Nubian people point to an influx of West Eurasian-related ancestry resulting from the introduction of Islam from the Arabian Peninsula starting in ~1300 BP (Hollfelder et al. 2017). Ancient DNA is therefore essential in revealing a genetic landscape that no longer exists in unadmixed form in the present day.

An aDNA study of sixty-six individuals from the Christian Period site of Kulubnarti in Sudanese Nubia (~650–1100 CE) brought to light the genetic impact of the close relationship between Egypt and Nubia. Ancestry of West Eurasian origin (most closely related to that found in the gene pool of Bronze Age and Iron Age people from the Levant, and comprising ~57% of the ancestry of the ancient Nubians studied) was most likely introduced into Kulubnarti through Egypt as an intermediary (Sirak et al. 2021), resonating with archaeological evidence (Adams 1977). This is also consistent with the presence of Levantine-related ancestry at an earlier date in ancient Egyptians (Schuenemann et al. 2017). Previously unknown from other lines of evidence, aDNA showed that this ancestry was disproportionately associated with females, raising new questions about the impact of female mobility in this region (Sirak et al. 2021). Speaking to the ability of DNA to answer broad questions about regional genetic diversity and interpopulation interactions as well as site- or group-specific questions, this study also showed that there were no significant differences in ancestry among individuals from two plausibly socially stratified cemeteries at Kulubnarti, supporting existing hypotheses based on archaeological evidence that people living on the riverine island and on the adjacent west bank may have been socially divided but were not genetically distinct (Adams et al. 1999). As in Egypt, present-day Nubians are not descended from the Christian Period people from Kulubnarti, attesting to the influence of demographic shifts that occurred after Christian times.

To date, all studies of genome-wide aDNA in the Nile Valley have included individuals who lived at a single archaeological site, and the representativeness of these individuals outside of their own site cannot necessarily be evaluated with confidence. Additional genetic studies on ancient
human remains from this region will help illuminate the genetic landscape and refine our understanding of the past African people who inhabited this important crossroads between continents.

**Central Africa**

Another notably underexplored area for aDNA research is central Africa. This region is home to some of the most genetically diverse people in the world, who harbor ancestry from deeply divergent ancient lineages (Lipson et al. 2020). As of 2021, published ancient genomes are limited and capture only snapshots of this vast region: four individuals from Shum Laka in Cameroon (Lipson et al. 2020; see the section “Ancient DNA from Foragers in Sub-Saharan Africa”); three individuals from Kindoki and one from Ngongo Mbata, both within the historic Kongo kingdom in western DRC (Wang et al. 2020); and one individual from Matangai Turu Northwest, in the Ituri Forest of eastern DRC (Wang et al. 2020). Though the individuals from DRC are relatively recent, the remarkable insight garnered from each of these genomes reinforces the potential of each ancient individual studied to have a transformative effect on how we understand the past. For example, individuals dated to ~230 and ~220 BP from Kindoki and Ngogo Mbata, respectively, show greater genetic affinity to ancient individuals associated with the eastern African Iron Age (see the section “Sub-Saharan Africa”), including the ~1200 BP Iron Age individual from Deloraine Farm in Kenya and the ~600 BP individual from Pemba Island in Tanzania, than to present-day Bantu-speaking groups. That these temporally and spatially separated ancient individuals with ancestry related to Bantu speakers are so closely related to each other raises the possibility that additional gene flow introduced through very recent migrations has genetically distinguished present-day Bantu-speaking populations from those that lived only a few hundred years ago in the same regions. Once again, aDNA hints at another genetic landscape that no longer exists.

An individual who was buried ~150 BP at Kindoki presents yet another unique genetic profile, with notable difference from an individual who lived around the same place only 80 years prior: the more recent individual exhibits ~15% West Eurasian-related ancestry that is absent in the individual who lived earlier (Wang et al. 2020). This is consistent with the region’s colonial history and Portuguese presence, as well as with this individual’s Christian-style burial. These findings caution against generalization from single individuals, since even people living relatively close together in space and time may be genetically unique, shaped by historical circumstances. Though logistical and preservation challenges have made the generation of aDNA data from Central Africa rare to date, this region remains a critical area for future archaeogenetic research.

**African Islands**

A further frontier of great interest are Africa’s islands. Because of their geography, islands are an ideal topic of exploration, in part because expansion dynamics and demographic processes that occurred on and shaped the genetic landscape of the African continent may (or may not) have
extended into the islands off the coast. The islands around Africa—which may have acted as zones for settling and mixing of genetically distinct groups and may have experienced geographically driven drift—have great potential for future aDNA research.

Presently, the only aDNA data available for African islands come from two archipelagos. The first is Zanzibar in Tanzania: two individuals at Kuumbi Cave on Unguja Island and Makangale Cave on Pemba Island, both dating ~1400 BP and genetically closest to ancient and present-day eastern African foragers (discussed in the section “Ancient DNA from Foragers in Sub-Saharan Africa”); and another individual from Makangale Cave dating to ~600 BP with West African-related ancestry (discussed in the section “Sub-Saharan Africa”) (Skoglund et al. 2017). The second archipelago is the Canary Islands, located off the coast of northwestern Africa but part of Spain. Here, aDNA has the potential to resolve long-standing questions about the origins of the Indigenous Guanche people. Given the genetic legacy of the Spanish conquest of these islands, these questions cannot be resolved from genetic data from living people alone.

A 2017 study examined genome-wide data from individuals from Gran Canaria and Tenerife islands who lived in the 7th–11th centuries CE, prior to the 15th-century Spanish conquest (Rodríguez-Varela et al. 2017). Previous aDNA studies of uniparental markers had shown that the Guanches carried common North African Y chromosome and mtDNA lineages as well as European ones (Fregel et al. 2009a, 2009b; Maca-Meyer et al. 2003; Ordóñez et al. 2017) and suggested genetic connections between Guanches and a North African Berber-related population (i.e., the islands may have been initially populated by such a source population). By leveraging genome-wide aDNA data, Rodríguez-Varela et al. (2017) assessed the ancient Guanches as being genetically most similar to present-day Berber groups from northwestern Africa, with an additional signal of early European farmer-related ancestry; the latter component is largely absent in some genotyped present-day Berber populations, though present in other modern North African groups. This component is associated with early Neolithic farmers from Anatolia and Europe, hinting at the possibility of Neolithic or post-Neolithic gene flow among North African groups. Overall, results are consistent with an origin from a single ancestral North African population, although both archaeological and genetic evidence attest to contact with nonlocal people prior to the demographic overhaul that occurred because of European colonization. Leveraging both ancient and modern data, as well as genome-wide data and frequencies of uniparental markers, it is further suggested that Guanche ancestry not only persists in present-day Canary Islanders (comprising 16% and 31% of ancestry in two present-day individuals) but is predominantly derived from female ancestors (Fregel et al. 2009b; Rodríguez-Varela et al. 2017). Additional work in the Canary Islands as well as other islands around Africa will shed new light on the demographic processes that shaped present-day genetic diversity in these areas and connected or separated their inhabitants from those who lived on the African continent.
Looking Forward in African Ancient DNA Research: Increasing Attention to Ethical and Engaged Research and Capacity Building

As ancient human DNA research in Africa and around the world continues to grow in scope and scale through methodological and technical advancements, conversations about how to carry out ethical and engaged research and contribute to capacity building are becoming increasingly necessary (e.g., Alpaslan-Roodenberg et al. 2021; Austin et al. 2019; Bardill et al. 2018; Crellin and Harris 2020; Gibbon 2020; Morris 2017; Prendergast and Sawchuk 2018; Sirak and Sedig 2019). Although in many ways we are still in the midst of an aDNA “revolution,” such conversations are an important signal of how the discipline is maturing as it continues along a trajectory of growth.

Ethics and Engagement

Ethical concerns must take priority in aDNA research for several reasons. First, the rapid growth of the field has led to simultaneous development of sampling and analytical methods and standards for ethical and engaged work, meaning that sometimes ethical standards are catching up to methodological developments; second, the social and political implications of studying ancestry include the potential for aDNA data to be misused in ways that may be harmful to present-day groups; and third, aDNA analysis requires the destruction and study of the remains of once-living people who cannot consent. While there is near-universal agreement among scholars that ethics and engagement with diverse stakeholder groups (people who have a connection to a study, including descendant and guardian communities, custodians of human remains, and researchers) must be foregrounded in aDNA research, it remains challenging to articulate what this means in practical terms and how it should be implemented in different research contexts. However, we are beginning to see the publication of detailed recommendations and guidelines for ethical and engaged aDNA research that are globally applicable (Alpaslan-Roodenberg et al. 2021), as well as those that are focused more specifically on research carried out in Africa (e.g., Gibbon 2020; Morris 2015, 2017; Prendergast and Sawchuk 2018).

The ethics of conducting aDNA research in African contexts in particular—including dealing with topics such as sampling human remains, community engagement, and permission-seeking—has been explored by a number of Africanist biological anthropologists and archaeologists. Some of the key issues brought to light are unique to Africa while others apply around the world. Issues particularly pertinent to research in Africa include the potential consequences of competition between aDNA laboratories for a limited number of skeletal samples relative to other parts of the world, given decades of underfunded archaeological research in Africa and the danger of disengaged and inequitable “parachute” or “helicopter” research, where scholars (often from the Global North) “drop in” to a country in order to obtain samples for scientific research but fail to generate any meaningful benefits for the scholars, institutions, and broader public in the countries from which those samples originate (often located within the Global South); and issues that are highly relevant globally, including a lack of communication and collaboration among archaeologists, biological anthropologists, and geneticists, and the inability to communicate
efficiently and effectively across disciplines, as well as the range of highly specialized methods used by different aDNA labs, which are not well understood by nonspecialists and may generate data that are difficult to compare and critically reassess (see Morris 2017).

There has been notable progress toward more transparent and detailed definitions of ethical and engaged research in Africa and beyond. Many recommendations for “best practices” build upon existing frameworks, including Indigenous-driven policies such as the San Code of Ethics (SASI 2017) and those designed for modern African DNA data and led by the Human Heredity and Health in Africa (H3 Africa) initiative (e.g., Yakubu et al. 2018), while recognizing that it is also essential for archaeological individuals and their living descendants to be treated with consideration and respect (Prendergast and Sawchuk 2018; Gibbon 2020). Though analogous to recommendations for DNA research on living people, it is important to have aDNA-specific ethics guidelines because some of the key issues inherent in aDNA research—for example, the inability of the deceased to give research consent and the need to identify appropriate stakeholders and descendants who can make decisions on their behalf—are not encountered during DNA research on the living. Improving the quality and quantity of communication and collaboration among geneticists, archaeologists, anthropologists, curators, and other stakeholders, including descendant and guardian communities, has proven to be an effective mechanism for stimulating aDNA-specific research guidance. Workshops, conferences, and professional meetings are also increasingly providing forums for discussing guiding principles for ethical and engaged aDNA research, an encouraging trend that is likely to continue as the discipline continues to grow. However, it is essential that any such guiding principles acknowledge that research contexts vary widely from one part of the world to another, making it particularly challenging to articulate “one-size-fits-all” principles for ethical and engaged aDNA research (Alpaslan-Roodenberg et al. 2021). What is appropriate and effective in one part of the world may not be so in another.

Guidelines for ethical research must also consider issues specific to regional contexts. One example in Africa is the ongoing impact of colonial legacies in research. Many skeletal collections from Africa remain located outside their country of origin, having been sent to Europe or North America for scientific analysis and consequently kept long-term. In practice, this lasting reflection of a long, complex, and oftentimes painful history of colonialism means that African archaeological collections (including human remains) are often fragmented across institutions, countries, and even continents with material from the same regions and sometimes even the same sites separated based on when they were recovered and by whom. This is then further exacerbated by piecemeal distribution of aDNA samples to laboratories around the world, which may result in the loss of important contextual information over time, making it difficult to identify communities who should be consulted and engaged as stakeholders and which governmental or institutional bodies should provide official permissions for scientific study. Researchers with projects predicated on archaeological remains curated outside of their country of origin must take particular care to ensure necessary permissions from the correct government and institution are in place and that appropriate stakeholders are identified and included in the research process. Public calls to repatriate skeletal material to Africa—though still rare—have occurred but tend to focus either on early human fossils (Musonda 2013) or individuals whose remains were taken during the colonial era (e.g., Förster 2020). This may become more frequent
as conversations progress about the repatriation of other forms of African heritage and the need to address the curation of historic African and Afro-descendant human remains (Ogbechie 2019; Dunnavant, Justinvil, and Colwell 2021). Ancient DNA researchers should not repeat the errors of their scientific predecessors but rather should be active in recognizing the power and autonomy of African authorities over their cultural and biological heritage and advocating for the return of skeletal material when appropriate and desired.

Even in cases where the original provenance of the remains is known, identifying the most appropriate stakeholder communities to involve in aDNA research can be a difficult process requiring an in-depth knowledge of local history and population dynamics. Across much of Africa, connections between living people and the land on which they reside are complex, shaped by histories of elective and forced migrations. The people who currently occupy an area may not identify as the descendants of people comprising past local populations based on histories of migrations from elsewhere or due to belief systems that differ from those exercised by past people, among other reasons. This adds a layer of complexity to the identification of stakeholder groups who should be consulted and engaged as part of the aDNA research process. It is the responsibility of researchers to make sure stakeholder groups are actively involved in the study and research questions in appropriate ways. Researchers also have a professional ethical obligation to understand whether reporting a result in a particular way is likely to cause harm.

Finally, as practitioners of aDNA research become increasingly conscientious of the need for ethical and engaged research, many are including statements detailing how they navigated specific ethical issues (e.g., Prendergast et al. 2019: Materials and Methods; Wang et al. 2020: Text S1; Sirak et al. 2021: Main Text). Although at present institutional and governmental guidelines for obtaining permission to analyze ancient human remains vary widely and do not always ensure ethical and engaged research, in some cases progress is being made at higher levels. For example, in South Africa, aDNA research is strictly regulated through the South African Heritage Resources Agency (SAHRA) and through additional consultation with representative community groups following defined guidelines (Gibbon 2020). A powerful solution would be to declare ancient human remains as sources of biological information that require the same level of ethical consideration as research carried out in living people (Gibbon 2020); indeed, this has been done in places such as the University of Cape Town and may be applied in wider contexts going forward.

**Capacity Building and Concluding Thoughts**

A pervasive issue in aDNA research in Africa and many other parts of the world is the potential for this work to exacerbate inequalities between the Global North (where aDNA labs are often located) and the Global South (where aDNA studies are increasingly focused). This has resulted in growing calls for long-term collaborations and capacity building to be considered an integral part of aDNA research projects so that this imbalance can be rectified. As there are currently no aDNA facilities in Africa, all analysis is carried out abroad, a process that parallels historical justifications for exporting African cultural heritage material to foreign institutions where much
of it remains curated. Ancient DNA research in Africa and beyond must reduce such extractive behaviors that reproduce colonial dynamics and should determine effective means of capacity building, depending on individual research contexts.

Capacity building can take many forms, such as generating opportunities for training and continuing education for under-resourced colleagues. To avoid the pitfalls of parachute research (Morris 2017), researchers should be committed to establishing and maintaining long-term collaborations that have benefits—both tangible and intangible—for collaborating African institutions and scholars. Preferably, such efforts should include the transfer of resources and skills (e.g., providing lab equipment or training in laboratory techniques for collecting or analyzing data). Such efforts should be accompanied by a shift over time toward more equitable partnerships that include more work being done in-country. African collaborators must be given opportunities to contribute to all academic output produced from data generated at labs outside of Africa. Wherever possible, financial support should be offered to ensure that African collaborators can participate fully in knowledge generation and attend professional meetings or training events that are most often held in the Global North.

Additional issues of structural inequality—including a lack of funding for African scholars and institutions, logistic and financial issues associated with securing visas for African nationals to travel abroad, and insurmountable costs associated with publishing and access to scientific journals—will require profound shifts in how science is structured, funded, awarded, and shared. However, change can also begin at the individual level. It is essential for aDNA researchers to commit to capacity building when undertaking research projects in Africa. Ensuring that ethical and engaged research and capacity building are foundational in ongoing and future aDNA research initiatives in Africa requires intention and commitment. Researchers must hold each other accountable for meeting high ethical standards.

Future aDNA research promises to reveal more details about the dynamic and complex nature of human history in Africa—the birthplace of all humanity—and ethics must be a foundational part of that work. If we can successfully combine a desire to explore some of the most pressing questions about the human past with improving methodologies and prioritizing ethical, equitable, and engaged research involving collaborators from across the African continent and around the world, the future of aDNA research in Africa is undeniably bright.

**Links to Digital Material**

Assessment and Reflection on the Ethical Dimensions of Archaeogenetics Research from the Max Planck Institute for Evolutionary Anthropology [https://www.eva.mpg.de/archaeogenetics/ethics].

European Nucleotide Archive—publicly available sequences [https://www.ebi.ac.uk/ena/browser/home].
Further Reading


References


**Notes**


**Related Articles**

Hominin Taxic Diversity