Genetics and the African Past

Mary E. Prendergast, Department of Anthropology, Rice University, Elizabeth A. Sawchuk, Department of Anthropology, University of Alberta; Department of Anthropology, Stony Brook University, and Kendra A. Sirak, Department of Genetics, Harvard Medical School; Department of Evolutionary Biology, Harvard University

https://doi.org/10.1093/acrefore/9780190277734.013.143

Published online: 19 October 2022

Summary

Africa harbors the greatest human genetic diversity on the planet, a fact that has inspired extensive investigation of the population structure found across the continent and the demographic processes that shaped observed patterns of genetic variation. Since the 1980s, studies of the DNA of living people have repeatedly demonstrated that Africa is the cradle of human origins, in agreement with fossil and archaeological evidence. Since the first ancient human genome was published in 2010, ancient DNA (aDNA) has contributed additional possibilities for exploring population history, providing a direct window into genetic lineages that no longer exist or are barely discernible. Genetic data from both living and ancient people—when integrated with available archaeological, bioarchaeological, historical linguistic, and written or oral historical data—are important tools for contextualizing African genetic diversity and understanding the biological and cultural processes that have shaped it over time. While most studies to date have focused on humans, aDNA can also be obtained from plant and animal remains, sediments, and some artifacts, all of which can enable a more comprehensive understanding of human lives.

Genetic research on the African past often focuses on human origins and Pleistocene population structure, as well as on the origins, directionality, and tempo of demographic changes that accompanied Holocene transitions to herding and farming. The rise of cosmopolitan cities and states in the past two millennia has been examined with genetic evidence to a very limited extent, but this is a potentially rich vein of research. Increasingly, forced migrations of enslaved Africans and the development of the diaspora are the subjects of genetic study as well. Yet to date, Africa remains vastly understudied relative to parts of the world such as Eurasia, in terms of both ancient and present-day DNA. This shapes not only the study of the past but also medical innovations and public health.

While the bulk of published African genomes come from present-day people, there are problems with relying solely on this data to reconstruct the past, given the continent’s long and complex demographic history. Increasingly, aDNA is providing novel perspectives on a past largely invisible in the genomes of people living in the 20th and 21st centuries due to recent demographic shifts. A surge in African aDNA studies since 2015 has also renewed longstanding debates about the ethics of genetic research on people, both living and deceased. Researchers working in Africa today must consider ethical issues including stakeholder engagement, informed consent, and control of biological samples and data; in aDNA studies, descendant communities, museum curators, bioarchaeologists, and geneticists, among others, play critical roles in these discussions.

Keywords: archaeogenetics, ancient DNA, population structure, population history, research ethics, archaeology, interdisciplinary research

Subjects: Historiography and Methods
What is Ancient DNA and How is It Studied?

Deoxyribonucleic acid (DNA) is a molecule found in all living things. The majority of an individual’s DNA is found in the nucleus of cells. This DNA is called nuclear DNA, which is inherited from both parents and includes twenty-three pairs of chromosomes, including twenty-two pairs of autosomes (any chromosome that is not a sex chromosome) and one pair of sex chromosomes; males and females both inherit an X from their mother and either a second X (females) or Y (males) from their father. A small amount of DNA is also found in the mitochondria of cells. This DNA is referred to as “mtDNA” and is passed down from mother to offspring (that is, along the matriline). DNA is informative about biological traits that may be distinctive and traceable from one generation to the next; ancestral relationships among organisms; and systematic genetic differences at a population level (known as population structure). DNA can be extracted from the biological tissues of both living and long-dead organisms. When DNA is recovered from an organism that died anywhere from decades to hundreds of thousands of years ago, it is known as ancient DNA (aDNA). The study of aDNA is notably different from the study of DNA recovered from living people (sometimes referred to as “modern DNA”) because after death, DNA begins to degrade and is not repaired by the enzymatic processes that maintained the integrity of the molecules during life.1 Human aDNA is most frequently obtained from skeletal tissues including bone, tooth, and dental calculus (calcified plaque), but it can also be recovered from soft tissues, such as hair or mummified flesh. It may also be obtained from human “traces”—for example, fecal matter or decomposed biological material in sediments, or residues on objects used by people, although this has so far only been applied in parts of the world where DNA preservation is best.2

The preservation of aDNA varies depending on myriad factors, including temperature, humidity, and soil pH. Ancient DNA recovery has been particularly successful when the organisms that are the subject of study were interred in cold, dry environments and sites such as caves with stable environmental conditions (that is, places where the temperature does not fluctuate widely and where there may be less variation in humidity as well). In addition, aDNA preserves differently in different biological tissues and other substrates; a discovery key to the growth of ancient DNA research and its application to new contexts around the world was that the osseous inner ear, housed inside the petrous part of the temporal bone, is a skeletal element that contains up to more than a hundredfold more DNA than other bones from the same individual.3 Tooth cementum has been shown to be another optimal substrate for ancient DNA preservation, as have auditory ossicles.4

Once DNA is extracted from biological tissues or materials with human traces following standard methods, it is manipulated to prepare it for sequencing, a technique that “reads” the order of nucleic acids—adenine (A), thymine (T), guanine (G), and cytosine (C)—that comprise the molecule.5 Raw sequencing data are then processed, which involves removing any data that fall below quality control thresholds, aligning the data against the human reference genome, removing duplicate molecules, and evaluating the authenticity of the isolated aDNA.
Because aDNA data are used to reconstruct genetic landscapes of the past, it is essential for researchers to confirm that the DNA sequenced and studied is authentic endogenous aDNA—that is, DNA is from the individual of interest—and not exogenous contaminant DNA (DNA that is not derived from the organism of interest, much of it likely from modern sources such as the environment or the researchers themselves). Contamination is a particularly serious challenge in aDNA research, which ultimately begins during the excavation of human remains when the biological material from which DNA will be extracted is selected (guidelines have been published for archaeologists who plan to include an ancient DNA component in their work). The use of aDNA “clean rooms”—specialized facilities maintained as sterile environments through positive air pressure systems, decontaminated with ultraviolet (UV) light and chemical cleaning procedures, physically separated from spaces where DNA is studied following amplification (a process during which many identical copies of target DNA are created), and accessed only by trained technicians wearing personal protective equipment—further minimizes the risks of contamination.

Despite following stringent protocols to reduce contamination, it is likely that small amounts of exogenous DNA will be present when the DNA from an ancient individual is sequenced. Therefore, researchers must evaluate the amount of contaminant DNA prior to carrying out population genetics analysis. Importantly, aDNA can be differentiated from contaminant modern DNA by the presence of a certain type of damage that is characteristic of ancient molecules. Specifically, the “deamination of cytosine” (a process in which a cytosine nucleotide is converted to a uracil nucleotide and consequently read by sequencing technologies as a thymine) is reflected in aDNA as a high rate of “C-to-T” (that is, cytosine-to-thymine) transitions clustered at the ends of the DNA molecule, but lacking throughout the interior of the molecule. In addition to quantifying the rate of C-to-T damage that is characteristic of aDNA, there are other methods for evaluating authenticity and assessing rates of contamination. Methods involve assessing the consistency of mtDNA sequences or X chromosome sequences in males. All of an individual’s mtDNA sequences should be identical, as should all X chromosome sequences in males (who only have one copy of this chromosome), so any “mismatches” in these contexts can be used to quantify contamination. Additional methods to evaluate contamination rates include assessing the breakdown of linkage disequilibrium, the association of certain alleles (places in the genome where a substantial proportion of people do not carry the same nucleotide, but where there are instead two or more options for the nucleotide found at the position) more often than expected by chance because of their physical proximity within the genome.

In addition to being invariably damaged, aDNA is often present in low quantities. As such, most of the history of aDNA research was spent focusing on mtDNA, a uniparental marker that can be used to trace an individual’s matriline. Unlike nuclear DNA, which is present in a single copy in the nucleus of each cell, mtDNA molecules are present in multiple copies in each of the hundreds to thousands of mitochondria in every cell. As mtDNA is transferred from mother to offspring without recombination (the exchange of genetic material among two DNA molecules which makes a new sequence), an identical copy is passed down untransformed (barring new mutations) from generation to generation. The rapid mutation and geographically patterned distribution of mtDNA facilitates the analysis of past and present individuals’ “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). Complementary to mtDNA, the Y chromosome can be used to trace a male individual’s patriline, as it is passed from father to son, with the nonrecombining portion (NRY) comprising the majority (~95%) of the chromosome. Although both of these uniparental markers can be informative, they illuminate only a single line of descent and 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 uniparental markers alone. In contrast, genome–wide DNA—where the study of many independent genomic positions reflects many phylogenetic trees—results in an increase in statistical power and provides a more comprehensive account of population history.\textsuperscript{11}

As a discipline, archaeogenetic research is considered to be amid a revolution. Increasingly, summaries of aDNA analysis—including details of methodological and analytical tools, key findings from research since the 2010s, and ethical considerations that must be taken into account—are being made accessible to nonspecialists.\textsuperscript{12} This includes summaries focused specifically on aDNA from African contexts.\textsuperscript{13}

**What Can Ancient DNA Say About the African Past?**

Ancient DNA has rarely been recovered from African archaeological contexts, in part because of high rates of biomolecular degradation resulting from the heat and humidity across most of the continent. Instead, most of what is known about the African genetic past comes from the DNA of people living in the late 20th and 21st centuries. While these data are frequently collected in the framework of medical genetics—often focused on questions of disease heredity and public health—they are also informative about population structure in the past and present. Additionally, many studies of modern DNA have been conducted with explicitly evolutionary or archaeological questions in mind, and increasingly, ancient and modern DNA can be combined to enable more powerful large-scale analyses of population history.\textsuperscript{14}

Methodological advances since 2015 have enabled aDNA to be extracted under suboptimal preservation conditions while minimizing destruction, allowing aDNA research to be applied to contexts previously deemed too poor for DNA preservation. In Africa, these developments are quickly advancing understanding of the past.\textsuperscript{15} The growth of aDNA research in Africa makes an important contribution to the historical sciences because it is often not possible to reconstruct population structures from the DNA of living people alone. People are unlikely to live exactly where their ancestors did, and sweeping demographic changes—for example, those accompanying the spreads of herding and farming, the rise of cosmopolitan city–states, and forced relocations and enslavement—mean that DNA from African people living in the 20th and 21st centuries is not directly representative of the landscape of genetic variation that existed in the past. Furthermore, these demographic changes were often accompanied by admixture, when people from genetically distinguishable groups reproduce thus combining DNA, meaning that ancient genetic lineages, even if persistsing in present–day people, may no longer be present in unadmixed form. Human aDNA is critical to access this otherwise invisible genetic diversity and
reveal contributions from now-extinct “ghost lineages” (i.e., ancestries that cannot be traced to documented present-day people but that can be identified through analysis of individuals to whom they contributed DNA).

In addition to human DNA, nonhuman DNA is also a rich source of information. Genetic studies of present-day African livestock and crops, for example, can inform about the origins and spread of these domesticates, as well as about the accompanying processes of artificial selection and hybridization. Archaeologists have been particularly interested in the implications of such studies for the spread of pastoralism. Modern genetic studies are often conducted with the goal of understanding heirloom livestock and crop lineages—threatened by 20th- and 21st-century breed “improvement” efforts and industrialized farming—and underscore that these lineages are important African biological heritage, worthy of protection alongside cultural heritage and traditional ecological knowledge. Unfortunately, studies of nonhuman aDNA in Africa are rare at present. This is at least partly due to destruction of DNA in faunal or botanical remains that were heated as part of culinary preparation. As of 2021, aDNA has been sequenced from the remains of a few domestic and commensal animals, at a handful of sites. These studies have helped shed light on multiple questions, including the origins of the African donkey; introductions of non-African taxa such as chicken, black rat, domestic cat, and pig; and the distinction of skeletal remains of domestic livestock from morphologically similar wild bovids. Botanical aDNA studies remain limited to North Africa. Other informative nonhuman sources of aDNA include pathogens; for example, the bacterium Rickettsia felis and the parasite Toxoplasma gondii were both identified in the remains of a young boy buried in South Africa about 2,000 years ago, enabling a reconstruction of these pathogens’ pre-pastoral history.

Especially when carrying out research on ancient humans, it is important for researchers to consider that genetic research does not operate in a vacuum but instead is designed to be interpreted alongside and contextualized using other forms of data. In their interpretations, geneticists must consider other lines of evidence including bioarchaeology and skeletal morphology, archaeology, historical linguistics, oral histories, and—where available and relevant—written documentation. These lines of evidence can lead to interesting and productive correlations and contradictions among various lines of inquiry: for example, one aDNA study found relatively strong correspondence between independent lines of genetic and linguistic evidence for population histories but rejected prior arguments that distinct material cultural traditions could be linked to language and ancestry. Ancient DNA research can also support other lines of inquiry by, for example, refining chronologies (because new direct radiocarbon dates often accompany aDNA studies) or investigating previously forgotten and/or unpublished sites that may become loci of renewed archaeological focus. When done well, such interdisciplinarity is one of the strengths of this research and is reflected in the term archaeogenetics (the holistic study of aDNA from archaeological contexts).
Conducting Ethical Archaeogenetic Research in Africa

Ethical concerns are paramount to all human genetic research, whether DNA samples are obtained from living people, long-deceased individuals, or human-made cultural objects or residues; some concerns also extend to nonhuman archaeological materials including faunal and botanical remains. In African research contexts, two closely interrelated issues loom large in any discussion of ethics: legacies of colonialism and Global North–South inequalities. Colonial legacies impact all aspects of scientific research in Africa, ranging from where archaeological collections are curated and stored to how people approach biological research on human subjects. Global North–South inequalities mean that most of the funds, facilities, and training opportunities for research are located off the continent, and this exacerbates the longstanding problem of helicopter (or parachute) research whereby non-African scholars obtain samples for scientific analysis without generating long-lasting benefits for the scholars, institutions, and broader public in the countries from which those samples originate. Additional ethical concerns center around identifying and engaging with stakeholders including indigenous groups or others who identify as descendants or guardians of human remains, and the potential for political mobilization of historical narratives based on genetic research.

For geneticists who collect DNA samples from living people, there is a robust global bioethics literature, and many African countries have established guidelines and procedures for human-subjects research, though ethical problems persist. The Human Heredity and Health in Africa (H3 Africa) Consortium has played a key role in developing guidelines specific to African research contexts, including a framework for genomic research and biobanking (i.e., collection and storage of biological samples). Briefly, this framework is guided by four principles: sensitivity to and respect for African values and cultures; benefits for African people; intellectual participation of African stakeholders; and “respect, fairness, equity and reciprocity.” These core values then guide a detailed set of frameworks on issues such as African intellectual leadership, consent, capacity building, harm avoidance, and international collaboration, among others.

While issues like helicopter research and lack of capacity building cut across many research areas, several specific issues pertinent to archaeogenetic research are different from those encountered in medical and modern population genetics. For example, protocols designed for research with living people do not adequately address the destruction of ancient human tissues, treatment of human-made cultural objects, long-term curation of aDNA extracts and libraries, identification of indigenous or descendant groups for consultation, or repatriation of human remains. Ethical principles for aDNA research are being developed rapidly around the globe and especially in North America. However, research contexts vary widely, and in many African countries, as elsewhere in the world, regulation of ancient DNA research on human remains currently does not differ from other forms of destructive research on any archaeological materials, effectively treating human remains as artifacts. South Africa is a notable exception, where aDNA research is strictly regulated through the South African Heritage Resources Agency, as well as curating institutions, and includes consultation with representative community groups, thus applying many of the same principles as in genetic research on living people.
In the absence of a unified policy to guide aDNA research, Africanist researchers have begun calling for best practices that can be agreed upon by stakeholders including from the curatorial, archaeological, and genetic communities, and descendant groups where these can be identified. These best practices cover everything from sampling decisions—that is, which bone will be selected for destructive analysis—to data archiving and sharing. While implementation of best practices is still variable, occurring on a case-by-case basis through memoranda of understanding among local research institutions and agencies, pan-African or global agencies or professional associations may begin to play a role.

The State of Genetic Research on the African Past

African genetic data are strikingly understudied, especially given the present and ancient genetic diversity on the African continent, the potential for discoveries leading to lifesaving public-health advances, and the relevance to the study of human origins and African and African diaspora history. Data are spatially patchy across the continent, for both modern DNA and especially aDNA, and in the case of aDNA, coverage of diverse archaeological problems and time periods remains uneven (see figure 1). This data scarcity is particularly remarkable when compared against other global contexts, but the situation is also poised to change dramatically.

Given the current patchiness of African aDNA data, a single ancient individual’s genome still has the potential to radically change what is known about the past. This review highlights some key questions that can be addressed through modern DNA and aDNA research in Africa, especially when considered in concert with the results of decades of research by archaeologists, linguists, and historians. While the state of knowledge will change as new studies emerge, it considers longstanding problems of interest to scholars of the African past, regardless of how fully or incompletely they have thus far been addressed via genetics.
Figure 1. A, map of all published ancient genomes in Africa as of 2021; B, all published ancient genomes as of 2021, by present-day country; C, approximate chronology of each of the directly dated individuals mapped in Figure 1A (individuals without direct dates are excluded). All maps created by the authors.

African Origins of *Homo sapiens*

Human evolution has been a central topic in African population genetics. A landmark 1987 study—based on mitochondrial DNA from 147 present-day individuals worldwide, including 20 Africans—firmly established what many paleoanthropologists and archaeologists had long argued: that *Homo sapiens* (*H. sapiens*) evolved in Africa in the past 200,000 years (200 ka).\(^{33}\) This genetic finding was confirmed through subsequent studies of DNA from living people, prompting rejection of earlier models that *H. sapiens* emerged in many parts of Eurasia.\(^{34}\) By the turn of the 21st century, fossil and archaeological data came to support a consensus for an African origin not only of *H. sapiens* but also of “modern behavior,” a concept that has since been problematized by reconsiderations of what it means to be “modern.”\(^{35}\) During the 2010s, multiple genetic studies attempted to identify the specific geographical area where *H. sapiens* first emerged, based on present-day distributions of DNA markers, but recent work has critiqued this approach.\(^{36}\)

While a growing body of genetic, archaeological, and fossil evidence has continued to support an African origin for *H. sapiens*, aDNA studies have further complicated the picture.\(^{37}\) For example, genetic studies have: (a) demonstrated admixture between *H. sapiens* and other members of the *Homo* genus (Neanderthals and Denisovans) occurring outside of Africa; (b) revealed small amounts of Neanderthal ancestry in some present-day African populations, resulting from later movements of people with Eurasian ancestry back into Africa; and (c) raised the more tentative
possibility of admixture within Africa between members of the genus *Homo* from an earlier diverged lineage (not yet sampled via aDNA) and *H. sapiens*, which might help account for late persistence of “archaic” morphological features observed in some fossil humans. Models for a recent African origin are being refined to allow the fundamentally African origin of humans while simultaneously recognizing that there may be no single place or time within Africa that represents *H. sapiens* emergence. Scholarship has also begun to disentangle what is meant by “modern” genetic ancestry, skeletal morphology, and behavior, recognizing these are unlikely to form a single package.

While a full understanding of African Late Pleistocene population structure is currently limited by the absence of aDNA more than 15 ka and from diverse localities, aDNA from Holocene individuals has enabled researchers to identify major lineages of human ancestry diverging in Africa roughly 200 to 300 ka. For example, a 2017 study analyzed the genomes of three ancient forager individuals buried about 2 ka in KwaZulu-Natal, South Africa, who bore no evidence for gene flow with herders or farmers. Researchers were able to leverage the high-coverage, unadmixed genome of one individual (Ballito Bay A) to re-estimate the date of the first human population divergence as roughly 350 to 260 ka. These findings pushed back prior estimates based on DNA from living peoples and are consistent with the current fossil evidence that suggests the appearance of *H. sapiens* about 315 ka. A subsequent study of ancient genomes from four children buried roughly 8 and 3 ka at the site of Shum Laka in Cameroon modeled four ancestral lineages diverging within Africa around the same time, leading to present-day central and southern African foragers and to all other living people in and outside of Africa; and a fourth “ghost” lineage, which no longer exists in unadmixed form.

Future archaegenetic research focused on Africa will likely reveal additional Pleistocene population structures and may be able to better identify specific genetic adaptations that enabled the success of early *H. sapiens*. Additionally, the potential exists for sequencing aDNA from archaic *Homo* fossils (as has been the case in Eurasia) or from fossil nonhuman apes, which may shed light on the hominids who inhabited the continent prior to the evolution of *Homo sapiens*. In temperate Europe, the earliest *H. sapiens* aDNA dates to roughly 45 ka, and even more ancient hominin aDNA dates to about 430 ka. By contrast, the earliest sequenced aDNA in Africa—as of 2021—is just about 15 ka, and an attempt to sequence much earlier human remains revealed a recent (~0.5 ka) origin for at least part of a putatively Late Pleistocene site. While the chances of aDNA preservation in skeletal remains buried in warm and tropical environments remain lower, advances in aDNA methods may yet push timeframes backward. Additionally, it is plausible that other methods—for example, the study of proteins that remain stable over millions of years (proteomics)—will be able to address questions at timescales aDNA cannot.

### Ancient African Foragers

Genetic studies have shown that people who identify as foragers (people who hunt, gather, and/or fish as their main form of subsistence; or whose ancestors have historically done so)—including eastern African foragers such as Hadza, southern African foragers such as Ju’hoansi, and central African foragers such as Mbuti—harbor very ancient, deeply diverging lineages. Increasingly,
aDNA from individuals found in archaeological contexts associated with foraging is playing an important role in illuminating population structures that existed prior to widespread pastoralism and farming. These archaeogenetic data reveal an African human landscape strikingly different from the one that exists today, which has been subsequently altered by major demographic changes (for example, population expansions, bottlenecks, and admixture events) during the late Pleistocene and Holocene, so that only parts of the genetic variation represented in the past are detectable in DNA today.

Archaeogenetic studies underscore different population histories for foragers living north and south of the Sahara. In northwestern Africa, aDNA data from Iberomaurusian foragers buried about 15 ka at the site of Taforalt (Morocco) demonstrates pre-Neolithic genetic ties between people living in this region and those associated with the Natufian culture in southwest Asia, a finding that can be linked to archaeological patterns. Subsequent ancient and modern DNA research demonstrated the persistence of this unique Paleolithic ancestry even to the present day, in small amounts and declining in frequency from northwestern to northeastern Africa.

The greater number of studied ancient individuals from more numerous sites south of the Sahara is catalyzing the development of a large-scale picture of African forager population structures from the Late Pleistocene through the Holocene. These data reveal the presence of a cline (or biological gradient) of relatedness among ancient foragers in eastern and southern Africa, stretching from the Horn of Africa to the Cape of South Africa. They also demonstrate that genetic branches leading to eastern, southern, and central African foragers diverged roughly around the same time, roughly 200 ka. Finally, these studies show that some ancient lineages are barely detectable or not detectable at all in the genomes of people living in the same regions today, attesting to major demographic shifts during the Holocene.

There is still much to learn about the population structures, interactions, and genetic adaptations of ancient African foragers. For example, questions remain about the connections among people descended from the deeply divergent eastern, southern, and central African lineages after about 200 ka. This is important because very little is known about human demography, movements across the continent, or genetic adaptations to environment during the Late Pleistocene, when paleoclimatic evidence attests to dramatic change, and archaeological evidence attests to transformations in technology, symbolic expression, and social organization associated with the development of the Middle Stone Age and the transition to the Later Stone Age. Archaeological, fossil, and modern genetic evidence point toward intermittent periods of contraction, isolation, and expansion in the Late Pleistocene. Yet determining when, where, and how these processes shaped modern African diversity requires a greater density of aDNA data spanning diverse temporal and geographical contexts. Such data are not available at present in part due to geographically variable research histories, limited research networks, and current political conditions and logistical challenges in some places, and are also reflective of the availability and accessibility of skeletal material and impact of environmental conditions, especially heat and humidity, on the preservation of human remains.
Later Holocene Demographic Transformations

Most of the published ancient genomes from Africa date to the later Holocene (see figure 1B), a period characterized by demographic events that fundamentally altered the present-day African genetic landscape. Consequently, these data have been primarily interpreted in terms of transformations associated with the emergence and spread of food production and, to a much lesser extent, those associated with the rise of states. Yet studied ancient individuals represent a small slice of ancient genetic diversity, with many regions and time periods still unexplored through aDNA. Here, a region–by–region overview is provided for the state of research and key questions, especially for the past 5000 years, noting that often genetic research is just beginning to scratch the surface of what is possible for this dynamic period.

Northwestern Africa and the Canary Islands

The genetic histories of northwestern Africa and the Canary Islands are relatively well researched compared to many other parts of the continent, although genome–wide aDNA studies are few. Genome–wide aDNA data from two sites in Morocco highlight genetic transformations associated with Neolithization, demonstrating an influx of Eurasian ancestry attributed to trans-Gibraltar population movements between the Early and Late Neolithic periods (~7 ka to ~5 ka). This finding adds weight to archaeological evidence for trans-Gibraltar connections at this time, such as pottery with Iberian roots found in Morocco, and African ivory and ostrich eggshell recovered in the Iberian peninsula. Combining ancient and present–day DNA data, this Neolithic demographic transformation was shown to have a greater genetic impact on present–day North Africans than later processes such as Arabization. Genetic studies also show these trans-Gibraltar connections are not unidirectional: by about 3.6 ka, sub-Saharan African ancestry is detected in an individual buried at Cueva del Ángel in southern Spain, although more work is needed to better understand the timing and extent of this spread of African ancestry to Europe.

Connections between the African mainland and the pre–Hispanic Canary Islands, and demographic transformations associated with Castilian conquest of the islands, are well attested through both present–day and ancient DNA, making this one of the best–studied parts of Africa in terms of genetic history. DNA from present–day people with autochthonous Canarian ancestry, when taken together with both uniparental markers and genome–wide aDNA from pre–Hispanic burials, show a clear northwestern African origin for the first settlers of the Canary Islands about 2 ka, agreeing with biological anthropological and archaeological evidence. Collectively, these studies—together with the aforementioned study in Morocco—further show that this northwestern African ancestry was itself admixed from multiple sources; that there were multiple waves of people arriving from northwestern Africa to the islands; that individual islands have distinct population histories; and that genetic data help reject controversial hypotheses of Phoenician–Punic settlement of the islands.

Burials postdating the 15th century CE—as well as the genomes of present–day Canary Islanders—demonstrate the genetic impacts of conquest and colonization, lethal violence, and enslavement, and attest to increased interisland mobility. A large influx of European ancestry is
detected in ancient and present-day Canary Islanders and is sex-biased, showing replacement of autochthonous males by European males, contrasting with the longer-term survival of female lineages. However, genetic studies also demonstrate that autochthonous Canarian ancestry makes a greater contribution to the DNA of the present-day population than has been previously appreciated based on historical and ethnographic sources, challenging beliefs that the local population was almost entirely replaced by Europeans.

While North and Sub-Saharan African ancestry is well-documented in present-day Canary Islanders, aDNA is necessary to understand how and when these lineages contributed to present-day populations both prior to and following European colonization. Ancient DNA analyses of two historic burial grounds testify to the genetic legacies of Spanish imperialism. One, an 18th-century cemetery at a church in Santa Cruz, Tenerife, reflected this port town’s diverse population with individuals with European, African, and American ancestries. The other, dating to the 15th–to–17th century and associated with a sugar-cane field in Gran Canaria, revealed the northwest African, Senegambian, and local Canarian ancestries of probable enslaved laborers; this topic is discussed further in the section “African diasporas.”

Northeastern Africa

Northeastern Africa (here defined as Egypt, Sudan, and South Sudan, and the Horn of Africa), positioned at an intermediate location between the African and Eurasian continents, has been a nexus of population interaction for many millennia, serving as a conduit of trade and an interaction sphere involving goods, culture, and people, from multiple parts of Africa, Asia, and Europe. As such, it is unsurprising that the present-day people of this region harbor components of ancestry that are both African and non-African in origin. Like many other parts of the world, the genetic diversity in this part of Africa was initially studied using blood groups and other classical markers, as well as uniparental haplogroups and other genetic markers. More recently, it has been studied at a higher resolution through analysis of genome-level variation in living people including studies of physiological traits such as lactase persistence. Understanding of the ancient populations of this region is largely based on inferences drawn from present-day genetic diversity (as well as historical sources and archaeological data); 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, cannot be fully resolved without aDNA.

The first published genome-wide aDNA from Africa came from an adult male forager who lived in northeast Africa about 4.5 ka, buried at Mota Cave in the highlands of present-day southwest Ethiopia. While geneticists refer to this individual as “Mota,” archaeological work refers to him as “Bayira,” meaning “firstborn” in the Gamo language. This ancient individual has none of the West-Eurasian–related ancestry that is prevalent throughout northeastern Africa at varying proportions in many living groups after having been introduced via a “back-migration” to Africa by the descendants of people who initially left the continent possibly tens of thousands of years earlier. “Mota-related” ancestry (that is, ancestry of the type that characterized the individual from Mota Cave) has been identified in ancient individuals who lived between roughly 4.5 ka and the present and continues to be detected in present-day people. A 2021 study showed a striking
correlation between spatial distance from Mota Cave and the amount of Mota–related genetic ancestry in present–day people, suggesting notable preservation of population structure in parts of Ethiopia for more than four millennia. While the sequencing of genome–level information from the Mota Cave individual has proven highly useful in subsequent aDNA studies, there is still much to be learned; for example, one study revealed that this individual is likely partly descended from a “ghost” source that also contributed ancestry to some present–day people in west–central Africa but that has yet to be sequenced directly.

Other parts of northeastern Africa, despite receiving notable attention from archaeologists and in some studies of modern DNA (for example, examining Nile Valley population movements), have been largely passed over in aDNA research, likely due to the increased levels of biomolecular degradation in hot climates. In particular, the Nile River Valley—which bisects Egypt and Sudan—is a critical region for population genetics studies. Ancient Egypt in particular has been a 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. In 2017, the first genome–wide aDNA data from the Nile River Valley (Abusir el–Meleq, Egypt) was sequenced, providing new insight into the ancestry of two individuals who lived during the Pre–Ptolemaic Period (New Kingdom to Late Period) and one who lived during the Ptolemaic Period. Spanning more than 1300 years of ancient Egyptian history, aDNA data showed a close relationship among ancient Egyptians and people from the Near East, consistent with archaeological and historical evidence. Ancient Egyptians were genetically more similar to people from the Near East than to present–day Egyptians. The latter exhibit about 8 percent more sub–Saharan–African–related ancestry, reflecting an influx of such ancestry into the region after the Roman Period; this ancestry shift was previously suggested from analysis of present–day people.

Moving up the Nile and forward in time, an aDNA study of sixty–six individuals from Kulubnarti in Sudanese Nubia dating to the Christian Period (~650–1100 CE) attested to an especially close relationship between ancient Egypt and Nubia. Ancestry most like that found in Bronze Age and Iron Age people from the Levant made up the majority of the ancestry of the individuals from Kulubnarti (~57%) and was most likely introduced through Egypt, resonating with archaeological evidence. This ancestry was disproportionately associated with females, raising new questions about the impact of female mobility in this region. New studies of aDNA from northeast Africa—including from new regions and time periods—will continue to reveal the demographic processes that gave a unique character to this ancient genetic landscape.

**Western and Central Africa**

Western and central Africa represent vast parts of the continent with tremendous genetic and linguistic diversity, yet DNA sampling in these regions is uneven, and available aDNA is also limited. Given the important archaeological, linguistic, and ancestral connections between these regions, here they are combined with studies of present–day DNA, providing an outline of known Holocene population history and what can potentially be learned from future aDNA research. Research has highlighted the demographic histories of hunter–gatherers; expansions of
speakers of Bantu languages; and evidence for admixture between these groups and other genetically distinctive groups. The role of this region as the source for most of the ancestry found in Afro-descendant populations in the Americas is discussed in the “African Diasporas” section.

Central African rainforest hunter-gatherers have been the focus of extensive genetic research, including on population histories and on adaptations to the rainforest, such as small adult body size and resistance to malaria. As noted in the section “Foragers Past and Present,” this modern DNA data, together with aDNA from four children buried at Shum Laka in Cameroon about 8 ka and about 3 ka, have demonstrated deeply divergent lineages that split off roughly 300 to 200 ka from other major branches of human ancestry. Additional divergences occurred about 60 ka between rainforest hunter-gatherers and the ancestors of Bantu language speakers, and about 25 ka between rainforest hunter-gatherers living in western (Congo Basin) and eastern (Ituri Forest, Lake Victoria) parts of central Africa, who subsequently remained isolated from one another. Together these findings contribute to an emerging picture of complex population structure across Africa in the Late Pleistocene and into the Holocene. Ancient genomes will help clarify whether or not the central African rainforest acted as a barrier to gene flow in the Holocene. Central African forager ancestry found in a mid-Holocene individual buried near Lake Victoria, for example, suggests a level of connectivity that may be more extensive than is suggested from the archaeological records.

The second major focus of genetic research in this region is on the origins and spread of speakers of Bantu languages, one of most defining demographic events in African history. Bantu languages—a group within the Niger–Kordofanian branch of the Niger–Congo family—are spoken by nearly one in three Africans today, across a vast part of the southern half of the continent. Linguists pinpoint their origin in the Grassfields region of the Cameroon/Nigeria borderlands, from which they expanded beginning roughly 5 to 4 ka. Although no one speaks “Bantu,” here “Bantu speakers” is used as shorthand to refer collectively to people who speak languages from this group. Bantu speakers in eastern, central, and southern Africa are weakly genetically differentiated, reflecting a recent shared ancestry. This demonstrates that the distribution of Bantu speakers today is the result of demic diffusion, that is, movement of people along with words and ideas, even if archaeological evidence suggests that languages, farming, and new material cultural traditions did not move as a single package. Geneticists have critically examined two hypotheses for this expansion put forward by linguists: the early-split model (in which Bantu speakers would have spread toward eastern and southern Africa simultaneously) and the late-split model (in which Bantu speakers would have moved south through the rainforest before splitting into eastern and southern streams). These studies have consistently found support for the late-split model, in agreement with recent linguistic studies, and in conflict with at least some archaeological data.

As Bantu speakers moved south through the rainforest, they admixed with autochthonous peoples in ways that reflect certain social dynamics. For example, analyses of the uniparental markers (mtDNA and the Y chromosome) revealed sex-specific patterns of mixture in modern Bantu-speaking populations. Specifically, ancient female hunter-gatherers had children with Bantu-speaking agriculturalist men far more often than the reverse.
agriculturalist admixture also conferred advantageous genetic mutations to admixed populations, resulting in many Bantu speakers across Africa today having malarial resistance inherited from their rainforest hunter-gatherer ancestors.

Beyond the rainforests of central Africa, Bantu speakers expanded into eastern and southern Africa via spatially and temporally variable processes involving multiple migration and admixture events. Data from present-day DNA, linguistics, and well-dated archaeological sites suggest a spread-over-spread scenario, with the initial spread of Bantu speakers truncated by a mid–first-millennium CE population collapse, followed by renewed population growth and expansions. The West–African–related ancestry found in Bantu speakers in eastern and southern Africa today is admixed—in variable proportions—with lineages associated with rainforest hunter-gatherers, southern African Khoe–San foragers, and eastern African pastoralists speaking Afro-Asiatic and Nilo-Saharan languages. Large-scale studies are beginning to shed light on routes and timing of these gene flow events, and critical areas for further study in south–central Africa. Ancient DNA will hopefully assist by providing greater spatial and temporal resolution. West–African–related ancestry has been identified in the ancient genomes of more than a dozen individuals buried at Iron Age sites spanning the past 1200 years in eastern, south–central, and southern Africa, but it is conspicuously absent in others. Greater data density is needed to understand these patterns, as well as the nature of population replacements hinted at by aDNA from Malawi and South Africa.

Additional genetic research in western Africa has probed present-day linguistically and ethnically diverse groups in the Chad Basin and, more broadly, across the Sahel. These studies demonstrate important east–west and north–south connections, linking peoples between the Nile Valley and the Atlantic Coast of Africa, and from the Sahara to sub-Saharan savannas. Population histories can be linked to well-documented paleoclimatic and archaeological data—for example, mobility across the shrinking “Green Sahara” during the early to middle Holocene—and also to less well-known connections between eastern Africa and the Chad Basin, indicating fruitful lines for archaeological research. Research in the Sahel has also focused on Fulani population history and the role of North African admixture in conferring lactase persistence, a key dietary adaptation.

Future lines of research in western and central Africa may examine the origins of powerful polities, such as those which emerged during the first to second millennia CE in the Middle Niger Delta, Congo Basin, Upemba Depression, and Great Lakes. While virtually no aDNA is yet available from these areas, one recent study leveraged present-day DNA from residents of Kananga (Democratic Republic of the Congo) to uncover the genetic impacts of 17th–century Kuba state formation, showing the people who identified as Kuba today were highly genetically diverse and were genetically more similar to their non–Kuba neighbors than to other Kuba. This was interpreted as evidence that the Kuba state subsumed many ancestrally distinctive groups into a single identity, in general agreement with oral histories. This study paves the way forward for genetically derived population histories of western and central Africa’s precolonial states.
Eastern Africa

Within Africa, eastern Africa is considered especially diverse, with high levels of present–day genetic diversity correlating with ethnolinguistic heterogeneity. This is because eastern Africa has been a continental crossroads for most of human (and earlier hominin) history, providing a backdrop for multiple population movements on and off the continent stretching back several million years. However, it is the past 5 ka that has had the most profound impact on the region’s current genetic landscape. Multiple population movements into and through eastern Africa associated with the diffusion of herding, farming, and iron working, as well as growing connections with the Indian Ocean world, produced a mosaic of people with different cultural, linguistic, and genetic backgrounds living side by side. Not only is it challenging to disentangle relationships within this mosaic across space and time, but limited DNA sampling of contemporary eastern African groups means fewer points for comparison for ancient lineages. While the region is currently the best sampled part of Africa in terms of aDNA research, this still only amounts to a few dozen sequenced individuals, all of whom date to within the past several millennia. Because the full extent of both past and present eastern Africa genetic diversity is unknown, scholars’ understanding of population history will continue to evolve as more genomes are sequenced.

One of the best studied phenomena in eastern African archaeology is the spread of food production into the region beginning roughly 5 ka. This multi-step process differed from Neolithic transitions in many other parts of the world because: a) key innovations such as domesticated animals, domesticated plants, and pottery appeared in different places at different times as opposed to spreading as a unified package; b) mobile pastoralism (a way of life organized around herding domesticated animals) preceded farming by several thousand years; c) foraging lifeways persisted (and still persist) alongside herding and farming in many places. The earliest evidence for food production in the form of herding appeared in the Sahara about 8 ka, with cattle, sheep, and goats spreading southward, arriving in eastern Africa about 5 ka and southern Africa by about 2 ka. Within this overall trajectory, there is considerable debate regarding timing of livestock introduction, routes taken, and the extent to which people were moving with animals versus trading them, making this a key area where aDNA methods can contribute to furthering archaeological and historical discourse.

The largest aDNA study in Africa to date sequenced forty–one individuals from forager, Pastoral Neolithic (PN, ~5–1.2 ka) and Iron Age and Pastoral Iron Age (IA/PIA, ~1.2 ka to recent) contexts from across Kenya and Tanzania, with results pointing toward a multi–step spread of food production that entailed admixture between migrant herders and farmers and autochthonous forager groups. The initial spread of pastoralism into eastern Africa involved at least two phases of admixture: one that likely occurred in northeastern Africa roughly 6–5 ka between people with ancestry related to present–day Dinka people in Sudan and ancient individuals from the Near East (who may have also been present in northeastern Africa), and a second phase roughly 4 ka in eastern Africa between this admixed group and local foragers. The IA/PIA entailed additional movements of people from northeastern Africa and western Africa after about 1.2 ka, reflected in the ancestries of later individuals descended from these groups and PN herders. While
gene flow between herders and foragers was common in the early phases of herding’s spread, evidence for genetic isolation between these groups later in the PN implies hardening of social barriers as pastoralism became entrenched.

When combined with other lines of archaeological and skeletal evidence, aDNA is a powerful tool for understanding how diverse peoples interacted within this mosaic and formed relationships. For example, PN herders associated with distinct material cultural traditions (characterized as Elmenteitan and Savanna Pastoral Neolithic) who were previously proposed to represent linguistically and ancestrally distinct groups were found to form a tight genetic cluster, indicating cultural differentiation occurred among genetically similar neighbors. Subsequent research documented additional variation in the amounts of forager- and Dinka-related ancestries among PN individuals from southern Kenya, raising the possibility of continued, albeit rare, gene flow between foragers and herders well into the PN. This is best illustrated by two individuals dated to about 1.5 ka from Molo Cave in Kenya with about 60 percent forager ancestry, supporting a scenario of ongoing connections in some places, consistent with archaeological evidence. Ancient DNA has also revealed geographic variation in the patterning of such relationships, with less evidence of herder–forager gene flow near lake and ocean coasts, as well as previously undocumented interregional connections with central and southern Africa.

Fruitful areas for future archaeogenetic research in eastern Africa include population structure and interactions among forager groups prior to the spread of food production, the emergence and spread of novel adaptations such as lactase persistence, and demographic change during the Iron Age and rise of states, topics now only hinted at through ancient genomes. Additional research may focus on coastal communities, where present–day genetic data have been used to estimate the timing and routes of dispersal of Bantu speakers toward southern Africa, and other studies consider the genetic input of people with Asian ancestry in otherwise West–Africa–related, Bantu–speaking populations in medieval Swahili communities. Beyond the coast, the settlement and population histories of the Comoros and Madagascar have been investigated through modern population genetics, but these are entirely unexplored via aDNA. This is a particularly fruitful vein for future research given the role of the western Indian Ocean in a far-reaching interaction sphere and the historically, archaeologically, and linguistically documented contributions of people with southeast Asian, southwest Asian, and African origins.

Southern and South-Central Africa

Southern Africa is a crossroads for groups with deeply divergent genetic ancestries, reflecting a deep population history overlaid by complex population dynamics especially in the past 2000 years. A major focus of aDNA research in southern Africa has been on deeper history and human origins. This is because present–day Khoe–San groups in southern Africa (see note for discussion of terminology) carry the greatest number of unique genetic variants and most deeply divergent lineages on the planet, indicating they split off from other human groups soon after the emergence of H. sapiens. However, more recent processes have had a significant impact on the genetic landscape too: for example, living Khoe–San peoples can trace 9 to 30 percent of their ancestry to genetic admixture with groups who migrated into the region within the past two
millennia. While some genetic studies of present-day groups in southern Africa assume a degree of population continuity and isolation, the potential movements of ancient foraging groups, and admixture with genetically distinct groups especially in the past roughly 2 ka, render aDNA one of the only means of exploring population structure in this region prior to the spread of food production. As noted in the sections “African Origins of Homo sapiens” and “Foragers Past and Present,” the study of ancient southern African foragers has revealed new perspectives on deep-time population structure. One exceptionally informative genome, from a young boy buried at Ballito Bay in KwaZulu-Natal, pushed back the timing of human origins; meanwhile aDNA from another individual, buried at St. Helena’s Bay in the western Cape, revealed a cline of ancestry among ancient and some present-day foragers, stretching between eastern and southern Africa. These individuals’ genomes serve as important context to understand genetic changes associated with the later appearance of food production.

As in eastern Africa, aDNA research in southern Africa has yielded information that can help resolve decades-long debates about how herding spread—whether through demic diffusion or cultural transmission—and putative connections between eastern and southern African pastoralists. An individual buried in a pastoralist context at Kasteelberg, South Africa roughly 1.2 ka was shown to have southern-African-forager–related ancestry admixed with eastern-African–or–West-Eurasian–related ancestry components closely related to a roughly 3.1 ka PN-era individual from Luxmanda, Tanzania. PN–related ancestry was also inferred for three individuals buried about 1.4–0.9 ka in the Okavango Delta in Botswana, suggesting admixture of southern African forager and eastern African pastoralist ancestries predated the appearance of ancestry related to Bantu speakers, which was also identified in the Okavango Delta individuals. These findings are consistent with modern DNA evidence for gene flow between eastern and southern Africa within the past roughly 2 ka, linguistic evidence linking eastern and southern African herders, and archaeological evidence that the first livestock and ceramics appear in southern Africa well prior to the mixed, agropastoralist “Iron Age package” associated with Bantu speakers.

Together, DNA and aDNA data strongly support a scenario wherein livestock were introduced to southern Africa by migrants from eastern Africa who then had children with local foragers, explaining the roughly 9 to 30 percent admixed ancestry observed in present-day KhoeSan peoples. Yet, while a direct interregional connection is important for understanding the mechanism for herding’s spread, there is still much to learn about how interactions unfolded as diverse peoples came together amid economic and social change. Questions remains about the scale of population movement involved and whether there were multiple infiltration events, as has been hypothesized based on archaeological evidence. Finally, the genetic findings do not preclude the possibility that local foragers adopted herding practices in many contexts, and cannot inform about fluidity between herding and foraging lifeways. Instead, aDNA offers an important tool for investigating what were likely multiple trajectories of herding’s spread and for exploring the diverse genetic outcomes of this event across space and time.

Ancient DNA has also illuminated complex dynamics associated with the introductions of farming and iron working into southern Africa in the first millennium CE by people with West-African–related ancestry who likely spoke Bantu languages. One study found about 16 percent forager...
ancestry in four Iron Age individuals buried in KwaZulu-Natal (all dated to within the past 500 years), compared to about 19 percent in present-day speakers of southeast Bantu languages in South Africa. As noted above, three individuals from the Okavango Delta indicate variable patterns of gene flow within a mosaic of foragers, herders, and farmers, having ancestry related to ancient pastoralists (who were themselves admixed with ancient foragers) and to Bantu speakers. Interestingly, no living group can be modeled as having the same ancestry mix as the two individuals from Xaro, offering a window into a population that has since been replaced by Bantu-speaking groups who inhabit the region today. Ancient DNA data also attest to population replacement in south-central Africa, where there is no trace of the ancient forager ancestry documented in individuals from Hora 1, Fingira, and Chencherere II rock shelters in present-day people. Instead, DNA data from Malawian groups such as the Chewa, Ngoni, Tumbuka, and Yao are consistent with their having about 100 percent West-African-related ancestry, indicating complete replacement of forager lineages that existed in the region as recently as roughly 2.5 ka. At the same time, other aDNA studies have demonstrated the persistence of groups unadmixed with Bantu speakers, nor with European colonists, despite living at the southernmost tip of southern Africa, an area of high contact, some two hundred years ago.

Relatively few individuals from southern and south-central Africa have been sampled for aDNA, and those sequences provide snapshots of very different periods in the region’s history, just as present-day sampling is biased toward specific groups, such as Khoe-San. Further study is needed to untangle demographic processes underlying the spreads of herding and farming, the expansion of Bantu languages and West-African-related ancestry, and eventually the rise of centers of power such as Mapungubwe and Great Zimbabwe with extended political influence and long-distance trade networks. Ancient DNA may be a particularly useful tool for exploring regional and local processes associated with social complexity in the Iron Age, such as the origins of the much-debated “matrilineal belt” and mobility and kinship within emerging class-based societies engaged in long-distance trade. Genetic data also enable examination of the formation of present-day southern African societies as products of European colonization, and forced migrations and segregations, bringing together of people with diverse African, Asian and European ancestries whose admixture patterns can be connected to known historical processes.

The African Diaspora and Formation of the Atlantic World

New data are beginning to reveal the genetic legacies of the Trans-Atlantic Slave Trade (TAST) which forcibly moved an estimated twelve million enslaved Africans to the Americas and Europe. These data complement abundant, albeit incomplete, historical and archaeological records for the TAST and the formation of the Atlantic World. To a lesser extent, studies are also revealing the genetic impacts of the trans-Saharan and Indian Ocean slave trade networks.

Population genetics provides insights into the diverse geographic origins of enslaved people, Maroon communities, and their descendants; dynamics of forced migrations to and throughout the Americas; and spatially and temporally variable and often sex-biased patterns of admixture with people of other African, Indigenous American, and European ancestries during and after the
TAST, which reflect specific political conditions and legal and other social barriers to gene flow. When taken together with insights from history, archaeology, and bioarchaeology, genetics can be a powerful tool for understanding specific histories of Afro-descendant individuals and groups, in addition to improving medical genetics by generating genetic data for presently underrepresented groups. This research is also prompting renewed discussion of treatment of Afro-descendant human remains, calls to address the patchy and biased nature of DNA sampling of Africans and people of African descent, and consideration of the social implications of and professional responsibilities invoked by studying ancestry.

While the African diaspora is a major focus of population genetics, ancient genomes are only just beginning to play a role, following initial studies using uniparental markers. Recent studies underscore that enslavement brought people of diverse ethnic and linguistic backgrounds together: for example, three enslaved African individuals buried in a 16th-century cemetery in Mexico City had distinct geographic origins in Africa, and this is also true of three enslaved Africans buried together in a 17th-century context on Saint Martin. Bioarchaeological and mtDNA analysis of a 17th-to-18th-century cemetery in the eastern United States showed stratification of individuals of European and African descent, and it contrasted relatedness among Euro-descendant individuals with diverse origins for the enslaved individuals.

While most aDNA studies of the Atlantic World have focused on the Americas, studies in early modern Canary Islands and Portugal have also contributed information on peoples of African origin who were likely enslaved, as discussed in the section “Northwestern Africa and the Canary Islands.” A study of ancient genomes from St. Helena Island in the South Atlantic, where the British navy left Africans captured from slave ships in the mid-19th century revealed these individuals’ had mainly west-central African ancestry, in line with modern DNA and historical studies.

Future archaeogenetic studies may be able to further illuminate the impacts upon African societies as the Atlantic World took shape. For example, the genome of a person buried about 150 years ago in the western (near-Atlantic) Congo showed about 85 percent ancestry related to Bantu speakers and about 15 percent ancestry related to Europeans, likely Portuguese. Combining aDNA and bioarchaeological studies can also provide insights into health, such as the identification of pathogens in the aforementioned Mexico City burials. Finally, aDNA holds potential to link material culture and genetic identity, as in the case of a tobacco pipe stem recorded in 19th-century slave quarters in the eastern United States, enabling the genetic profiling of the woman who used it and providing information of interest to descendant communities.

The Future of African Archaeogenetic Research

A challenge in reviewing the state of the field is the fast pace of change in African population genetics and especially in aDNA studies, where each newly sequenced genome has the potential to upend prior narratives. Yet even as new studies accumulate, longstanding debates—some dating...
to the earliest population genetics studies of the 1980s—persist. The following section outlines some thorny issues that problematize archaeogenetic research, whose resolution will require disciplinary specialists’ close communication and a desire to find common ground.

What Good is Ancient DNA?

A common sentiment among some archaeologists, linguists, and historians is that archaeogenetics merely confirms what is already known from other lines of evidence. While this is sometimes the case—one example being the close correspondence between historical and genetic records for the role of west-central Africa in the TAST—aDNA can also help resolve debates, provide unexpected results that contradict prior research, and incite new questions. For example, genetic research tips the scales in favor of the late-split model for the Bantu expansion, despite ambiguity on this issue in the archaeological record. In eastern Africa, unexpectedly high genetic homogeneity among makers of different PN archaeological traditions has upended hypotheses that they represent groups with distinctive geographic origins and arrival times in the region. Finally, aDNA discoveries—and the direct radiocarbon dates that often accompany them—are pointing to new lines of archaeological research, for example on early Holocene links between central and eastern African foragers. While aDNA is just one tool for reconstructing the past, it has the potential to add new information and build upon decades of work in other fields. There may be times, however, when researchers find that aDNA cannot add much value beyond what is known from robust historical or archaeological records, and it is important that this research remains question-driven rather than simply “doing the DNA” for its own sake.

Problems of Scale, Perspective, and Nomenclature

A central point of tension among scholars in genetics, linguistics, archaeology, and history is differing approaches to their shared curiosity about the past. Some of this has to do with scale and the kinds of questions that can be asked at different scales. Geneticists working in Africa tend to do so at regional or even continental scales. While part of this has to do with the topics that interest geneticists—for example, Pleistocene human population structure or the Bantu expansion—it is also a reflection of the patchy nature of both modern and ancient DNA sampling, which thwarts finer-grained analyses and leads to the mistaken impression that archaeogenetics is exclusively concerned with “grand narratives.” In this sense, archaeogenetics stands where African archaeology did in the mid-to–late 20th century, when practitioners did the basic work of defining culture–histories through lithics, pottery, and other types of material culture. Similarly, much early linguistic work examined continental–scale change. It was only once basic frameworks were understood that new questions became accessible, allowing both archaeologists and linguists to investigate at smaller scales.

For a model of what might be possible in the future with greater data density, one can look to other case studies—mainly in Europe (but see note for a case study of a Christian Period cemetery in Sudanese Nubia)—where numerous ancient individuals from the same region or even in the
same cemetery are sequenced.\textsuperscript{161} This approach permits social questions such as status inequality, kinship, postmarital residence, and health to be investigated in greater detail, in conjunction with archaeological and bioarchaeological evidence.\textsuperscript{162} Using different kinds of datasets is another point of potentially productive tension, as they can lead to divergent perspectives on the same problems. Ancient DNA might inform us, for example, that a large-scale migration occurred in the past, and statistical modeling can help indicate origin places and times for that migration. But archaeology and radiocarbon dating are critical to providing ground truth for these inferences with well-dated sites; archaeology, bioarchaeology, linguistics, and even ethnography may help determine the social conditions under which such events took place, a perspective genetics cannot offer.

As scales and datasets vary among disciplines, so too do vocabularies. Geneticists study ancestry, which is not the same as ethnic or linguistic identity, nor as a grouping based upon material culture, though all of these can come into play in nomenclatures. One must be careful, however, not to conflate different kinds of identity, a topic considered carefully by practitioners.\textsuperscript{163} For example, here the phase “ancestry related to Bantu speakers” has been used, rather than “Bantu ancestry.” Similar problems arise when using “farmer” or “pastoralist” as shorthand for genetic groupings, given the flexibility with which people move among subsistence strategies on individual, seasonal, and other bases.\textsuperscript{164} Finally, categories based upon archaeological cultures (e.g., “PN-related ancestry”) also pose challenges, given decades of archaeological debate about defining material culture traditions. While shorthand terms are preferred for ease of presentation—especially when analyzing and comparing multiple ancient genomes—they fail to capture important distinctions among kinds of identity, something geneticists increasingly discuss explicitly in publications.

### Integrating Bioarchaeology and Archaeogenetics

Bioarchaeology—the study of human remains in their archaeological contexts—provides different but complementary information to genetics regarding life in the past.\textsuperscript{165} While aDNA can inform on an individual’s lineage, bioarchaeological assessment is well-poised to yield insights into the lived experiences of the deceased. In many ways, bioarchaeological data (in conjunction with other methods such as radiocarbon dating) “give life” to aDNA samples by providing context on when, where, how, and how long the person lived, and whether they experienced skeletally evident disease or trauma. Although aDNA studies have had a tendency to disregard or underutilize bioarchaeological data, recent years have seen a notable change in this regard.\textsuperscript{166} In southern Africa in particular, research is focusing on the “people behind the samples,” a process which involves integrating genetic, skeletal, and archaeological data to paint a more holistic picture of what life was like for the studied individuals.\textsuperscript{167} Such efforts will become increasingly important as African aDNA studies shift from continental-scale surveys toward exploration of regional and even local processes, enabling the investigation of local-or site-level research questions that may be based in—but unable to be fully answered by—archaeological and bioarchaeological research.\textsuperscript{168} Bioarchaeological data is also playing a role in the emerging study of ancient diseases in Africa, both through the identification and paleopathological study of
affected individuals (e.g., the boy from Ballito Bay who likely suffered from schistosomiasis) and by interpreting the context of genetic adaptations for disease resistance (e.g., the Duffy null allele that confers protection against malaria).  

As is often the case when different lines of research first begin to intersect, there is ample room for greater synergy between bioarchaeological and archaeogenetic research. Especially since aDNA research requires destructive sampling, bioarchaeological assessment should be a required first step to determine how many individuals are present, collect basic contextual data (e.g., age at death, estimated sex, cultural associations), select elements that will minimally compromise future research, and, for commingled or isolated collections, ensure the same person is not sampled twice. Bioarchaeologists should be actively involved in shaping research design and identifying the best individuals for study depending on research questions, for example, those who may have been exposed to certain pathogens like tuberculosis, or regional contexts where morphological evidence indicates demographic change. As scholars with deep working knowledge of archaeological human remains, bioarchaeologists are and will continue to be instrumental in providing insight on individuals and collections that cannot or should not be studied using aDNA, and shaping research approaches for those that can.

**Rethinking Interdisciplinarity**

Population genetics has been usefully disruptive to studies of African history for more than a half century, from the first studies of classical markers through the Human Genome Project to the aDNA revolution. The challenges explored in this article are, in some ways, nothing new. Just as the pendulum swings back and forth between larger- and smaller-scale narratives in African history, so too do the disciplines discussed here merge and diverge more or less productively over time. Given the pace at which genetic research is moving, specialists must work past differences in vocabularies, perspectives, disciplinary training, and academic cultures in order to forge new ways of collaborating. This means moving beyond a “confirmation paradigm” in which specialists seek out verification of existing viewpoints from other disciplines. This tendency can lead to what has been called the “Last Paragraphs Problem,” where scholars cherry-pick and uncritically cite work from other disciplines to offer an interpretive summary of their own data. Fortunately, in studies of the African past, archaeologists, linguists, historians, and geneticists are discussing how they can better work together. Future training in the historical sciences—which must include capacity building in African institutions as a key component—should expose students to these different areas of expertise and vocabularies, to ensure that scholars continue to strive toward common ground.

**Discussion of the Literature**

African genomics and especially paleogenomics are young fields, with most scholarship dating since the early 2000s. However, more than a half century of research has explored questions of human population structures pertaining to deep human origins in Africa, as well as to demographic transformations that can be linked, with varying degrees of strength, to linguistic,
archaeological, and historical records. As a broad generalization, the study of African population histories has moved from study of classical markers from the 1960s through 1980s (e.g., the study of blood types, Rh factors, or immunological markers), to uniparental markers (mtDNA and Y-chromosome DNA) from the 1980s through 2000s, and finally progressed to genome–wide data with the birth of the Human Genome Project in the early 21st century. Since 2015, the shift has been toward sequencing genome–wide ancient DNA data, beginning with the study at Mota Cave in Ethiopia. The literature on African population history is no longer divided between present–day DNA and ancient DNA; rather, these different sources of data are increasingly integrated.

In parallel with methodological advances, archaeologists, linguists, anthropologists, and historians have increasingly engaged with the approaches and findings of geneticists in their study of the African past. MacEachern’s landmark paper “Genes, Tribes, and African History,” published in 2000 partly as a response to Cavalli-Sforza and colleagues’ 1994 opus The History and Geography of Human Genes, problematizes use of identity–based categories in genetic research, and more broadly examines the value and pitfalls of genetic research in light of decades of archaeological, historical, ethnographic, and linguistic scholarship on the African past. While this essay is not alone in its critical assessment of genetics, it continues to resonate today. The genomic revolution of the early 2000s provoked additional consideration of the integration of archaeology, historical linguistics, and genetics, and the benefits and perils of interdisciplinarity. The “ancient DNA revolution,” gaining pace in Africa since 2015, has provoked a third wave of thought pieces on the genomic turn in African studies. Many of the concerns are the same as raised by MacEachern in 2000. Others focus on the ethics of destructive sampling of human skeletal remains and the need for meaningful community engagement. Although Africanist scholars are increasingly seeing these issues through the lens of decolonization and as part of broader calls to make the historical sciences more equitable, critiques of DNA and aDNA research in Africa continue to reflect unresolved issues that remain from the earliest days of the field.

**Primary Sources**

Unlike other historical sciences that rely on documents as their primary sources, the primary sources for human DNA studies are people. In the case of modern DNA studies, this includes not only genetic data but also the information that participants provide, such as ethnic self-identification, languages spoken, and so forth. In the case of paleogenetic studies, human aDNA studies primarily rely upon archaeological human remains. Additional sources of ancient DNA include faunal and botanical remains, and cultural materials that have organic content and thus may preserve DNA (e.g., stone tools with adhering plant-based mastic; parchment made from animal skins; or human saliva residues on the stem of a tobacco pipe). Archaeological collections are commonly cared for by museums and universities, both across the African continent and, often, in Europe and North America. Under colonial rule, many African archaeological collections—including human skeletal remains—were removed from their place of origin and taken either to European museums and universities or to another location within European-controlled territory: for example, African cultural heritage and human remains from multiple former British colonies are today found in South Africa. While repatriation is more openly discussed in the 21st century, it remains far from a reality in most cases, and relatively little attention has focused on archaeological (as opposed to historical) human remains.
Museums often maintain databases of their collections, which are sometimes publicly accessible. These collections can normally only be accessed with permissions from the relevant government agencies, curatorial staff, or other stakeholders such as indigenous communities. Since these stakeholders are specific to each national context and may vary depending on the location and timeframe of an archaeological site, they are impossible to list exhaustively here. A good starting point is the directory of [https://icom.museum/en/network/committees-directory/](https://icom.museum/en/network/committees-directory/) maintained by the International Council of Museums (ICOM).

Once sampled from archaeological remains, derived molecular products (DNA extracts and libraries) are typically curated for the long term in the laboratory where the work took place and where temperature and other conditions are controlled for optimal preservation. This is an essential step that effectively immortalizes this information, so it can be re-examined and used in other genetic studies without requiring further sampling of archaeological remains, which can be returned to the curating institution. Eventually, curation of DNA extracts and libraries may also shift to the samples’ countries of origin; however, this will require dedicated lab infrastructure that does not yet exist in most African countries.

Genomic and paleogenomic research has been cited as a model for open science. As a condition of publication, both ancient and modern DNA sequences are made publicly accessible, making it possible for other researchers to independently confirm results and incorporate these data into their own analyses. Three major archives are currently active and sharing data with one another and with the public: the European Nucleotide Archive, GenBank, and the DNA DataBank of Japan. Many laboratories also maintain a website hosting publicly accessible DNA data produced by their studies, with Internet links often published in the relevant paper. Ongoing discussions regarding the need to balance open science with indigenous data sovereignty are highly relevant to aDNA research in some contexts and will likely play a major role in shaping practices in the coming years.

A key and sometimes-underappreciated aspect of paleogenetics studies is that the supplementary information often contains much of the detailed primary data that historians, archaeologists, historical linguists, and other scholars of the past may be seeking. While DNA studies are often published in scientific journals that do not allow for expansive manuscripts, the supplement has become a venue in which to publish—often for the first time—key details of archaeological sites, burials, or radiocarbon dates, or data on ethnic self-identity and languages spoken by living people whose DNA is in the study.

**Links to Digital Materials**

Human Heredity and Health in Africa consortium (H3Africa): professional association invested in genomics and public health, aiming to support African researchers in genomics [https://h3Africa.org/](https://h3Africa.org/).

Publicly available data for selected papers in African paleogenomics:

David Reich Lab [https://reich.hms.harvard.edu/datasets](https://reich.hms.harvard.edu/datasets).

Jakobsson Lab [http://jakobssonlab.iob.uu.se/data/](http://jakobssonlab.iob.uu.se/data/).

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


Notes


Genetics and the African Past


42. Schlebusch et al., “Southern African Ancient Genomes Estimate Modern Human Divergence to 350,000 to 260,000 Years Ago.”


58. Tryon, “The Middle/Later Stone Age Transition and Cultural Dynamics of Late Pleistocene East Africa”; and Scerri et al., “Did Our Species Evolve in Subdivided Populations across Africa, and Why Does It Matter?”


60. Fregel et al., “Ancient Genomes from North Africa Evidence Prehistoric Migrations to the Maghreb from Both the Levant and Europe.”


Genetics and the African Past


75. Gallego Llorente et al., “Ancient Ethiopian Genome Reveals Extensive Eurasian Admixture in Eastern Africa.”


78. López et al., “Evidence of the Interplay of Genetics and Culture in Ethiopia.”


114. Prendergast et al., “Ancient DNA Reveals a Multistep Spread of the First Herders into Sub-Saharan Africa.”


Genomes Estimate Modern Human Divergence to 350,000 to 260,000 Years Ago.


135. Schlebusch et al., “Southern African Ancient Genomes Estimate Modern Human Divergence to 350,000 to 260,000 Years Ago.”


155. Schablitsky et al., “Ancient DNA Analysis of a Nineteenth Century Tobacco Pipe from a Maryland Slave Quarter.”

156. Fortes-Lima and Verdu, “Anthropological Genetics Perspectives on the Transatlantic Slave Trade.”

157. de Filippo et al., “Bringing Together Linguistic and Genetic Evidence to Test the Bantu Expansion.”

158. Prendergast et al., “Ancient DNA Reveals a Multistep Spread of the First Herders into Sub-Saharan Africa.”


---

*Printed from Oxford Research Encyclopedias, African History. Under the terms of the licence agreement, an individual user may print out a single article for personal use (for details see Privacy Policy and Legal Notice). date: 18 April 2023*

161. Sirak et al., “Social Stratification without Genetic Differentiation at the Site of Kulubnarti in Christian Period Nubia.”


164. de Luna, *Collecting Food, Collecting People: Subsistence and Society in Central Africa*.


166. Morris, “Ancient DNA Comes of Age, but Still Has Some Teenage Problems.”


168. Sirak et al., “Social Stratification without Genetic Differentiation at the Site of Kulubnarti in Christian Period Nubia.”


173. De Luna, Fleisher, and McIntosh.

174. MacEachern, “Genetics and Archaeology.”


176. MacEachern, “Genetics and Archaeology.”


Slave Trades and Diaspora in the Middle East, 700 to 1900 CE