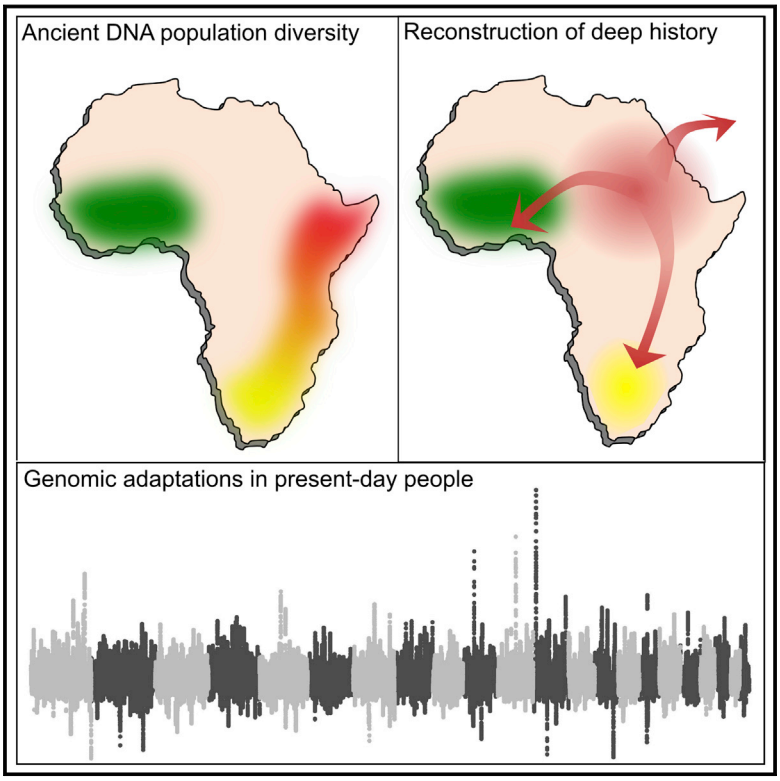
Article

Cell

Reconstructing Prehistoric African Population Structure

Graphical Abstract



Highlights

- Genome-wide analysis of 16 African individuals who lived up to 8,100 years ago
- Forager populations related to southern African San were once widespread in eastern Africa
- Comparison of ancient and modern Africans reveal recent genomic adaptations
- Evidence for a divergent human lineage contributing to western Africans

Authors

Pontus Skoglund, Jessica C. Thompson, Mary E. Prendergast, ..., Ron Pinhasi, Johannes Krause, David Reich

Correspondence

skoglund@genetics.med.harvard.edu (P.S.), reich@genetics.med.harvard.edu (D.R.)

In Brief

The prehistory of African populations is explored by genome-wide analysis of 16 human remains providing insight into ancestral lineages, admixture, and genomic adaptations.

Skoglund et al., 2017, Cell *171*, 59–71
September 21, 2017
http://dx.doi.org/10.1016/j.cell.2017.08.049

Cell

Reconstructing Prehistoric African Population Structure

Pontus Skoglund,[1,*] Jessica C. Thompson,[2] Mary E. Prendergast,[3] Alissa Mittnik,[4,5,33] Kendra Sirak,[2,6,33] Mateja Hajdinjak,[7,33] Tasneem Salie,[8,33] Nadin Rohland,[1] Swapan Mallick,[1,9] Alexander Peltzer,[4,10] Anja Heinze,[7] Iñigo Olalde,[1] Matthew Ferry,[1,11] Eadaoin Harney,[1,11,12] Megan Michel,[1,11] Kristin Stewardson,[1,11] Jessica I. Cerezo-Román,[13] Chrissy Chiumia,[14] Alison Crowther,[4,15] Elizabeth Gomani-Chindebvu,[14] Agness O. Gidna,[16] Katherine M. Grillo,[17] I. Taneli Helenius,[18] Garrett Hellenthal,[18] Richard Helm,[19] Mark Horton,[20] Saioa López,[18] Audax Z.P. Mabulla,[16] John Parkington,[21] Ceri Shipton,[22,23] Mark G. Thomas,[18] Ruth Tibesasa,[24] Menno Welling,[25,26] Vanessa M. Hayes,[27,28,29] Douglas J. Kennett,[30] Raj Ramesar,[8] Matthias Meyer,[7] Svante Pääbo,[7] Nick Patterson,[3,9] Alan G. Morris,[21] Nicole Boivin,[4] Ron Pinhasi,[6,31,32] Johannes Krause,[4,5,32] and David Reich[1,9,11,32,34,*]

[1]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
[2]Department of Anthropology, Emory University, Atlanta, GA 30322, USA
[3]Radcliffe Institute for Advanced Study, Harvard University, Cambridge, MA 02138, USA
[4]Max Planck Institute for the Science of Human History, Jena 07745, Germany
[5]Institute for Archeological Sciences, Eberhard-Karls-University, Tuebingen 72070, Germany
[6]School of Archaeology and Earth Institute, University College Dublin, Dublin 4, Ireland
[7]Max Planck Institute for Evolutionary Anthropology, Leipzig 04103, Germany
[8]Division of Human Genetics, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town 7925, South Africa
[9]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[10]Integrative Transcriptomics, Centre for Bioinformatics, University of Tuebingen, Tuebingen 72076, Germany
[11]Howard Hughes Medical Institute, Harvard Medical School, Boston, MA 02115, USA
[12]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA
[13]Department of Geography and Anthropology, California State Polytechnic University, Pomona, Pomona, CA 91768, USA
[14]Malawi Department of Museums and Monuments, Lilongwe 3, Malawi
[15]School of Social Science, The University of Queensland, Brisbane, Queensland 4072, Australia
[16]National Museums of Tanzania, Dar es Salaam, Tanzania
[17]Department of Archaeology and Anthropology, University of Wisconsin - La Crosse, La Crosse, WI 54601, USA
[18]Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, UK
[19]Canterbury Archaeological Trust, Canterbury CT1 2LU, UK
[20]Department Archaeology and Anthropology, University of Bristol, Bristol BS8 1UU, UK
[21]Department of Archaeology, University of Cape Town, Cape Town 7700, South Africa

(Affiliations continued on next page)

SUMMARY

We assembled genome-wide data from 16 prehistoric Africans. We show that the anciently divergent lineage that comprises the primary ancestry of the southern African San had a wider distribution in the past, contributing approximately two-thirds of the ancestry of Malawi hunter-gatherers ~8,100–2,500 years ago and approximately one-third of the ancestry of Tanzanian hunter-gatherers ~1,400 years ago. We document how the spread of farmers from western Africa involved complete replacement of local hunter-gatherers in some regions, and we track the spread of herders by showing that the population of a ~3,100-year-old pastoralist from Tanzania contributed ancestry to people from northeastern to southern Africa, including a ~1,200-year-old southern African pastoralist. The deepest diversifications of African lineages were complex, involving either repeated gene flow among geographically disparate groups or a lineage more deeply diverging than that of the San contributing more to some western African populations than to others. We finally leverage ancient genomes to document episodes of natural selection in southern African populations.

INTRODUCTION

Africa harbors more genetic diversity than any other part of the world (Cann et al., 1987; Tishkoff et al., 2009). This is reflected both in a higher average number of differences among sub-Saharan African genomes than among non-African genomes (Cann et al., 1987; Ramachandran et al., 2005) and in the fact that the ancestry found outside of Africa is largely a subset of that within it (Tishkoff et al., 2009). Today, some of the earliest-branching African lineages are present only in populations with relatively small census sizes, including the southern African Khoe-San (see STAR Methods for terminology), central African rainforest hunter-gatherers, and Hadza of Tanzania (Gronau et al., 2011; Schlebusch et al., 2012; Veeramah et al., 2012). However, the population structure of Africa prior to the

[22]McDonald Institute for Archaeological Research, Cambridge CB2 3ER, UK
[23]British Institute in Eastern Africa, Nairobi 30710, Kenya
[24]Department of Anthropology and Archaeology, University of Pretoria, Pretoria 0083, South Africa
[25]African Studies Centre Leiden, Leiden University, Leiden 2300 RB, Netherlands
[26]African Heritage Ltd, Zomba, Malawi
[27]Genomics and Epigenetics Division, Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia
[28]Central Clinical School, University of Sydney, Camperdown, NSW 2050, Australia
[29]School of Health Systems and Public Health, University of Pretoria, Gezina 0031, South Africa
[30]Department of Anthropology and Institutes for Energy and the Environment, Pennsylvania State University, University Park, PA 16802, USA
[31]Department of Anthropology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria
[32]Senior authors
[33]These authors contributed equally
[34]Lead Contact
*Correspondence: skoglund@genetics.med.harvard.edu (P.S.), reich@genetics.med.harvard.edu (D.R.)
http://dx.doi.org/10.1016/j.cell.2017.08.049

expansion of food producers (pastoralists and agriculturalists) remains unknown (Busby et al., 2016; Gurdasani et al., 2015; Patin et al., 2017). Bantu-speaking agriculturalists originating in western Africa are thought to have brought farming to eastern Africa by ~2,000 years BP (years before present, defined by convention as years before 1950 CE) and to southern Africa by ~1,500 BP, thereby spreading the largest single ancestry component to African genomes today (Russell et al., 2014; Tishkoff et al., 2009). Earlier migration(s), which brought ancestry related to the ancient Near East (Lazaridis et al., 2016; Pagani et al., 2012; Pickrell et al., 2014), brought herding to eastern Africa by ~4,000 BP (Marshall et al., 1984) and to southern Africa by ~2,000 BP (Sadr, 2015).

RESULTS

To reconstruct African population structure prior to the spread of food production, we generated new genome-wide data from 15 ancient sub-Saharan Africans (Table 1; Tables S1 and S2; STAR Methods). For three individuals from the western Cape of South Africa (~2,300–1,300 BP), we carried out direct shotgun sequencing to 0.7- to 2.0-fold coverage. For 12 individuals from eastern and south-central Africa, we used in-solution enrichment of ~1.2 million single nucleotide polymorphisms (SNPs). These included four individuals from the coastal region of Kenya and Tanzania (~1,400–400 BP), one from interior Tanzania (~3,100 BP), and seven from Malawi (ranging over ~8,100–2,500 BP) (Figure S1). All individuals had postmortem degradation characteristic of ancient DNA (Table 1), and we confirmed that key results are unlikely to be artifacts of contamination by restricting analysis to sequences with postmortem degradation (Skoglund et al., 2012; Skoglund et al., 2014a) (Figure S2). We merged the new ancient DNA data with previously reported shotgun sequence data from a ~4,500 BP Ethiopian highland individual (Llorente et al., 2015) and with SNP genotypes from 584 present-day African individuals from 59 diverse populations (including new data from 34 Malawi individuals; STAR Methods) (Lazaridis et al., 2014; Patterson et al., 2012), as well as 300 high-coverage genomes from 142 worldwide populations (Mallick et al., 2016). We also determined mitochondrial DNA and Y chromosome haplogroups for all newly reported samples (Tables S3 and S4).

An Ancient Cline of Southern and Eastern African Hunter-Gatherers

We used principal component analysis (PCA) (Patterson et al., 2006) and automated clustering (Alexander et al., 2009) to relate the 16 ancient individuals to present-day sub-Saharan Africans (Figures 1 and S3). Whereas the two individuals buried in ~2,000 BP hunter-gatherer contexts in South Africa share ancestry with southern African Khoe-San populations in the PCA, 11 of the 12 ancient individuals who lived in eastern and south-central Africa between ~8,100 and ~400 BP form a gradient of relatedness to the eastern African Hadza on the one hand and southern African Khoe-San on the other (Figure 1A). The genetic cline correlates to geography, running along a north-south axis with ancient individuals from Ethiopia (~4,500 BP), Kenya (~400 BP), Tanzania (both ~1,400 BP), and Malawi (~8,100–2,500 BP), showing increasing affinity to southern Africans (both ancient individuals and present-day Khoe-San). The seven individuals from Malawi show no clear heterogeneity, indicating a long-standing and distinctive population in ancient Malawi that persisted for at least ~5,000 years (the minimum span of our radiocarbon dates) but which no longer exists today.

We constructed a model where ancient and present-day African populations trace their ancestry to a putative set of nine ancestral populations. As proxies for these populations, three different ancient Near Eastern populations and six African populations were used that, according to our analyses, harbor substantial ancestry related to major lineages present in Africa today. The Mende from Sierra Leone are used in this model to represent a component of ancestry that exists in high proportions in western African populations, the ancient southern African genomes (South_Africa_2000BP) are used to represent the ancestry of southern Africa before agriculture, the Ethiopian individual (Ethiopia_4500BP) is used to represent northeastern African ancestry before agriculture, the Mbuti are used to represent central African rainforest hunter-gatherer ancestry, the individual from an eastern African pastoralist context (Tanzania_Luxmanda_3100BP) is used to represent an early pastoralist lineage from eastern Africa (see below; Figure 2), and the Dinka (from Sudan) are used to represent distinctive ancestry found in Nilotic speakers today. The ancient Near Eastern populations were representative of Anatolia, the Levant, and Iran,

Table 1. Summary of Ancient DNA from 15 Prehistoric Individuals Newly Reported in This Study

ID	Population label	Date (BP denotes calibrated years before present [defined as 1950], while "uncalibrated BP" denotes the radiocarbon age)*	Location	Lat.	Long.	Y chromosome haplogroup	mtDNA haplogroup	Damage rate at 5′ CpG dinucleotides	SNPs hit on 1.2M autosomal targets
I9028	South_Africa_2000BP	2241-1965 BP (2330 ± 25 uncalibrated BP, UGAMS-7255)	St. Helena, South Africa	−32.8	18.0	A1b1b2a	L0d2c1[$]	28%	731,098 (1.1X shotgun)
I9133	South_Africa_2000BP	2017-1748 BP (2000 ± 50 uncalibrated BP, Pta-5283)	Faraoskop Rock Shelter, South Africa	−32.0	18.5	A1b1b2a	L0d1b2b1b	33%	1,028,904 (2.3X shotgun)
I9134	South_Africa_1200BP	1282-1069 BP (1310 ± 50 uncalibrated BP, Pta-4373)	Kasteelberg, South Africa	−32.8	17.9	Female	L0d1a1a	21%	641,971 (0.8X shotgun)
I4427	Malawi_Fingira_6100BP	6175-5913 BP (5270 ± 25 uncalibrated BP, UCIAMS-186346)	Fingira, Malawi	−10.8	33.8	BT	L0d1b2b	37%	99,341
I4468	Malawi_Fingira_6100BP	6177-5923 BP (5290 ± 25 uncalibrated BP, UCIAMS-186347)	Fingira, Malawi	−10.8	33.8	BT	L0d1c	49%	30,257
I4421	Malawi_Chencherere_5200BP	5400-4800 BP (radiocarbon dating was unsuccessful; the date is based on context of other materials from the same site)	Chencherere, Malawi	−14.4	33.8	Female	L0k2	28%	59,470
I4422	Malawi_Chencherere_5200BP	5293-4979 BP (4525 ± 25 uncalibrated BP, UCIAMS-186348)	Chencherere, Malawi	−14.4	33.8	Female	L0k1	42%	9,355
I2966	Malawi_Hora_8100BP	10000-5000 BP (radiocarbon dating was unsuccessful; the date is based on context of other materials from the same site)	Hora, Malawi	−11.7	33.6	BT	L0k2 (PMDS > 3)	54%	610,605 (+0.26X shotgun)
I2967	Malawi_Hora_8100BP	8173-7957 BP (7230 ± 60 uncalibrated BP, PSUAMS-2535)	Hora, Malawi	−11.7	33.6	Female	L0a2	48%	65,686 (+0.03X shotgun)
I4426	Malawi_Fingira_2500BP	2676-2330 BP [2676-2343 BP (2425 ± 20 uncalibrated BP, PSUAMS-1734), 2483-2330 BP (2400 ± 20 uncalibrated BP, PSUAMS-1881)]	Fingira, Malawi	−10.8	33.8	Female	L0f	39%	635,427
I3726	Tanzania_Luxmanda_3100BP	3141-2890 BP (2925 ± 20 uncalibrated BP, ISGS-A3806)	Luxmanda, Tanzania	−4.3	35.3	Female	L2a1	65%	845,016
I0589	Tanzania_Zanzibar_1400BP	1370-1303 BP (1479 ± 23 uncalibrated BP, OxA-31427)	Kuumbi Cave, Zanzibar Island, Tanzania	−6.4	39.5	Female	L4b2a2c	22%	752,917
I1048	Tanzania_Pemba_1400BP	1421-1307 BP (1520 ± 30 uncalibrated BP, Beta-434912)	Makangale Cave, Pemba Island, Tanzania	−4.9	39.6	Female	L0a	42%	168,117
I2298	Tanzania_Pemba_600BP	639-544 BP (623 ± 20 uncalibrated BP, Wk-43308)	Makangale Cave, Pemba Island, Tanzania	−4.9	39.6	Female	L2a1a2	29%	695,242
I0595	Kenya_400BP	496-322 BP (388 ± 27 uncalibrated BP, OxA-30803)	Panga ya Saidi, Kenya	−3.7	39.7	E1b1b1b2	L4b2a2	30%	150,383

*Table S2 provides detailed information on the direct radiocarbon dating measurements.
[$]Consistent with previously published mtDNA sequence by Morris et al. (2014).

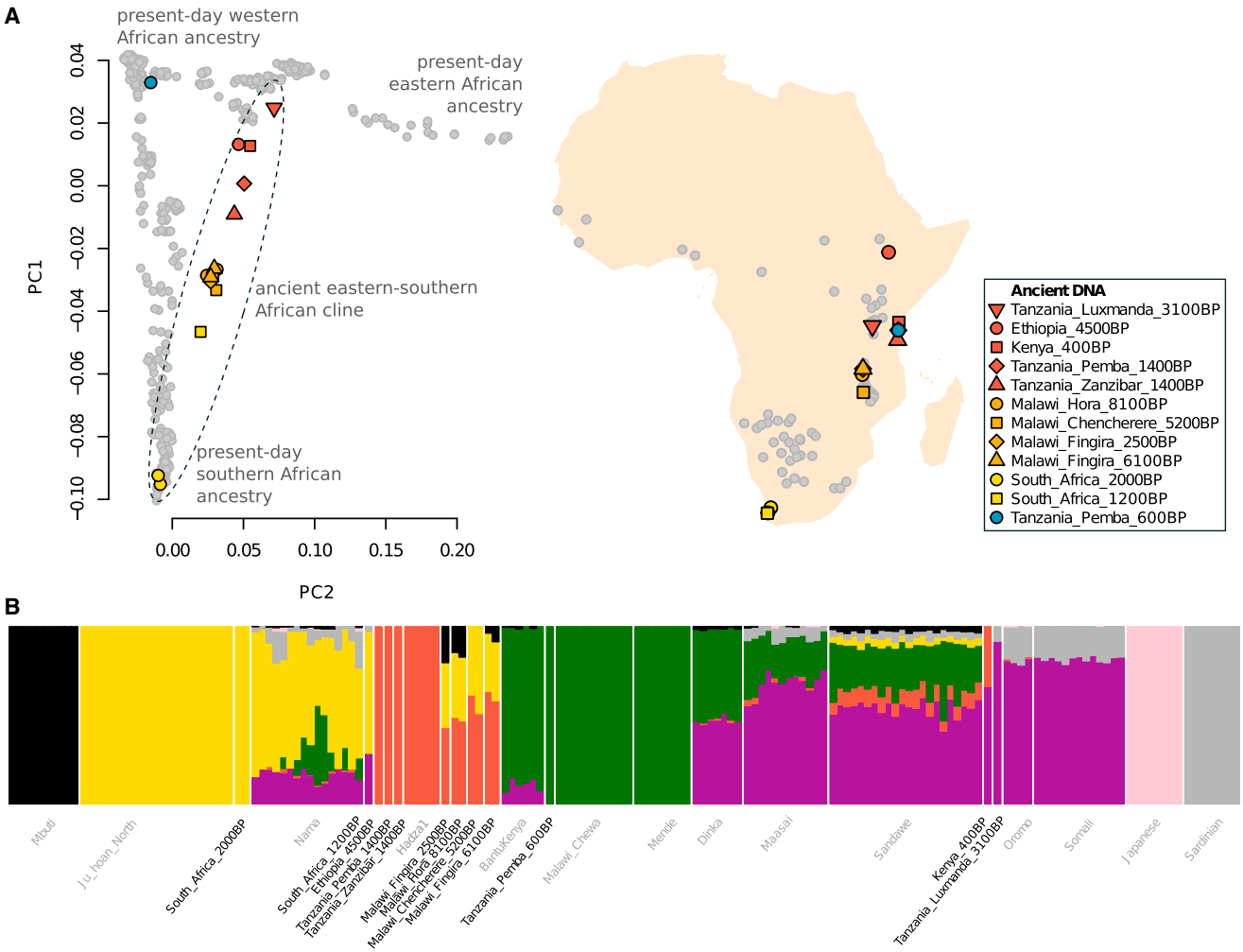


Figure 1. Overview of Ancient Genomes and African Population Structure

(A) Map of sampling locations in Africa and principal component analysis of all individuals. Present-day individuals are indicated with gray circles.

(B) Automated clustering of key ancient and present-day populations (for K = 7 cluster components). Present-day populations are labeled in gray.

respectively (Lazaridis et al., 2016; Mathieson et al., 2015). We used *qpAdm* (Haak et al., 2015), a generalization of f_4 symmetry statistics, to successively test 1-source, 2-source, or 3-source models and admixture proportions for all other ancient and present-day African populations, with a set of 10 non-African populations as outgroups (STAR Methods).

We find that ancestry closely related to the ancient southern Africans was present much farther north and east in the past than is apparent today. This ancient southern African ancestry comprises up to 91% of the ancestry of Khoe-San groups today (Table S5), and also 31% ± 3% of the ancestry of Tanzania_Zanzibar_1400BP, 60% ± 6% of the ancestry of Malawi_Fingira_6100BP, and 65% ± 3% of the ancestry of Malawi_Fingira_2500BP (Figure 2A). Notably, the Khoe-San-related ancestry in ancient individuals from Malawi and Tanzania is symmetrically related to the two previously identified lineages present in the San ($Z < 2$; Figure S2), estimated to have diverged at least 20,000 years ago (Mallick et al., 2016; Pickrell et al., 2012; Schlebusch et al., 2012), implying that this was an ancient divergent branch of this group that lived in eastern Africa at least until 1,400 BP. However, it was not present in all eastern Africans, as we do not detect it in the ~400-year-old individual from coastal Kenya nor in the present-day Hadza.

Displacement of Forager Populations in Eastern Africa

Both unsupervised clustering (Figure 1B) and formal ancestry estimation (Figure 2B) suggest that individuals from the Hadza group in Tanzania can be modeled as deriving all their ancestry from a lineage related deeply to ancient eastern Africans such as the Ethiopia_4500BP individual (Figure 3A; Table S5). However, this lineage appears to have contributed little ancestry to present-day Bantu speakers in eastern Africa, who instead trace their ancestry to a lineage related to present-day western Africans, with additional components related to the Nilotic-speaking Dinka and to the Tanzania_Luxmanda_3100BP

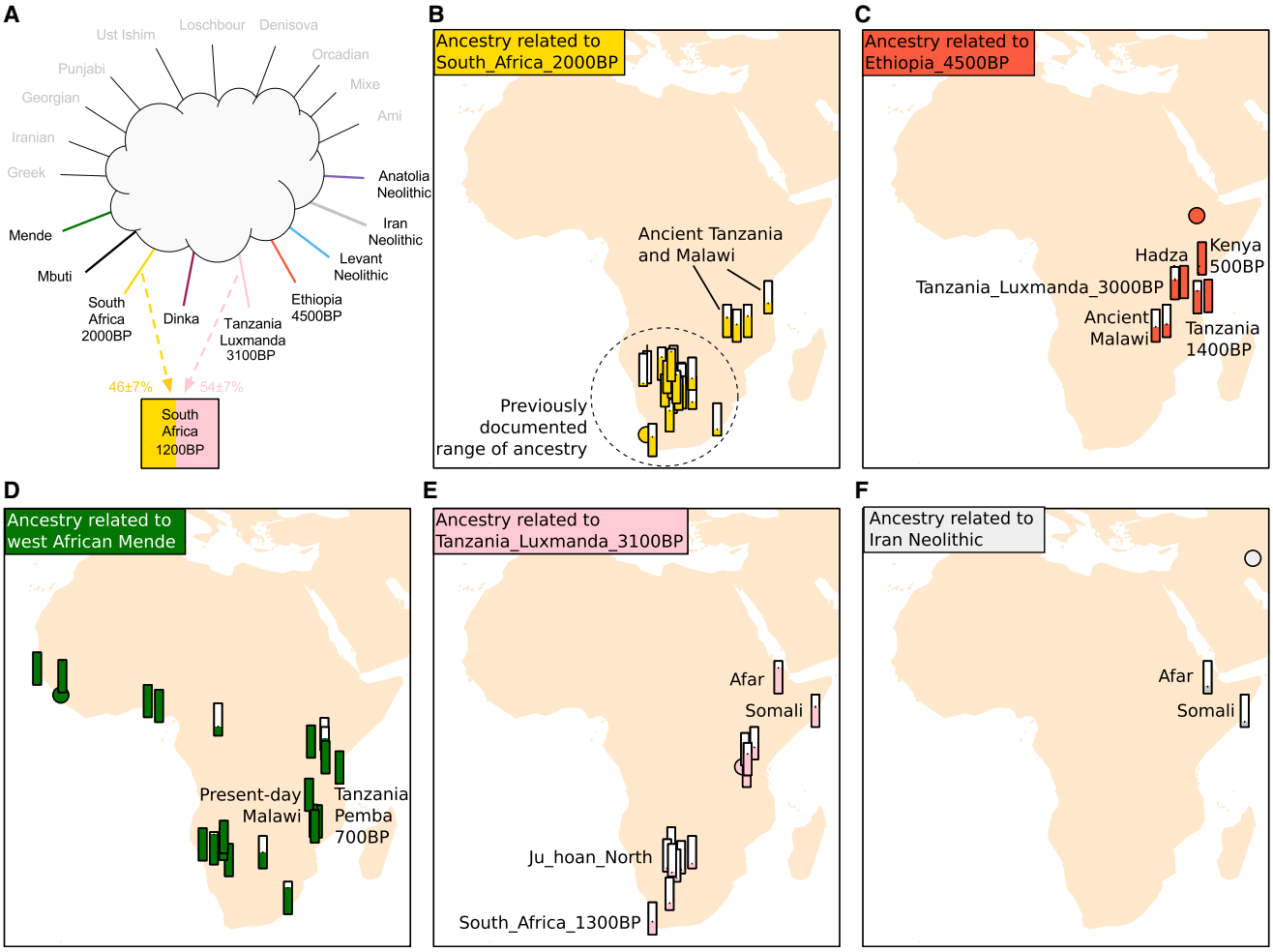


Figure 2. Ancestral Components in Eastern and Southern Africa

We show bar plots with the proportions inferred for the best model for each target population. We used a model that inferred the ancestry of each target population as 1-source, 2-source, or 3-source mixture of a set of potential source populations.

(A) We show an example of the inferred model for South_Africa_1200BP, an early pastoralist. A filled circle symbol in each panel indicates the geographic location of the sample that we use as a representative of the source population. We show five sources:

(B) South_Africa_2000BP, representing forager populations in southern Africa and a component of prehistoric Malawi and Tanzania that is no longer extant;

(C) Ethiopia_4500BP, which is today found in the Hadza but in the past was characteristic of eastern African hunter-gatherers;

(D) The Mende from Sierra Leone, which is related deeply to the western African ancestry that was spread with the Bantu expansion of agriculturalists;

(E) The Savanna Pastoral Neolithic sample Tanzania_Luxmanda_3100BP, a missing-link pastoralist population that almost certainly is responsible for having spread ancestry most closely related to the ancient Levant to southern Africa and which is also closely but not exclusively related to present-day Cushitic speakers;

(F) Ancestry more closely related to early Iranian farmers (Iran Neolithic) than what is found in Tanzania_Luxmanda_3100BP and which may have entered the Horn of Africa in later migrations.

pastoralist (see below; Figure 2). The Sandawe, another population that like the Hadza uses click consonants in their spoken language, are modeled as having ancestry similar to the Hadza but also admixture related to that of neighboring populations (Figure 3A; Table S5) consistent with previous findings (Henn et al., 2011; Tishkoff et al., 2009). Population replacement by incoming food producers appears to have been nearly complete in Malawi, where we detect little if any ancestry from the ancient individuals who lived ~8,100–2,500 BP. Instead, present-day Malawian individuals are consistent with deriving all their ancestry from the Bantu expansion of ultimate western African origin (Figure 3).

Among the ancient individuals analyzed here, only a ~600 BP individual from the Zanzibar archipelago has a genetic profile similar to present-day Bantu speakers (Figure 1). Notably, this individual has even more western-African-related ancestry than the present-day Bantu speakers we analyzed from Kenya, who also derive some of their ancestry from lineages related to Dinka and Tanzania_Luxmanda_3100BP (Figure 1B). Using linkage disequilibrium, we estimate that this admixture between western- and eastern-African-related lineages occurred an average of 800–400 years ago (STAR Methods). This suggests a scenario of genetic isolation between early farmers and

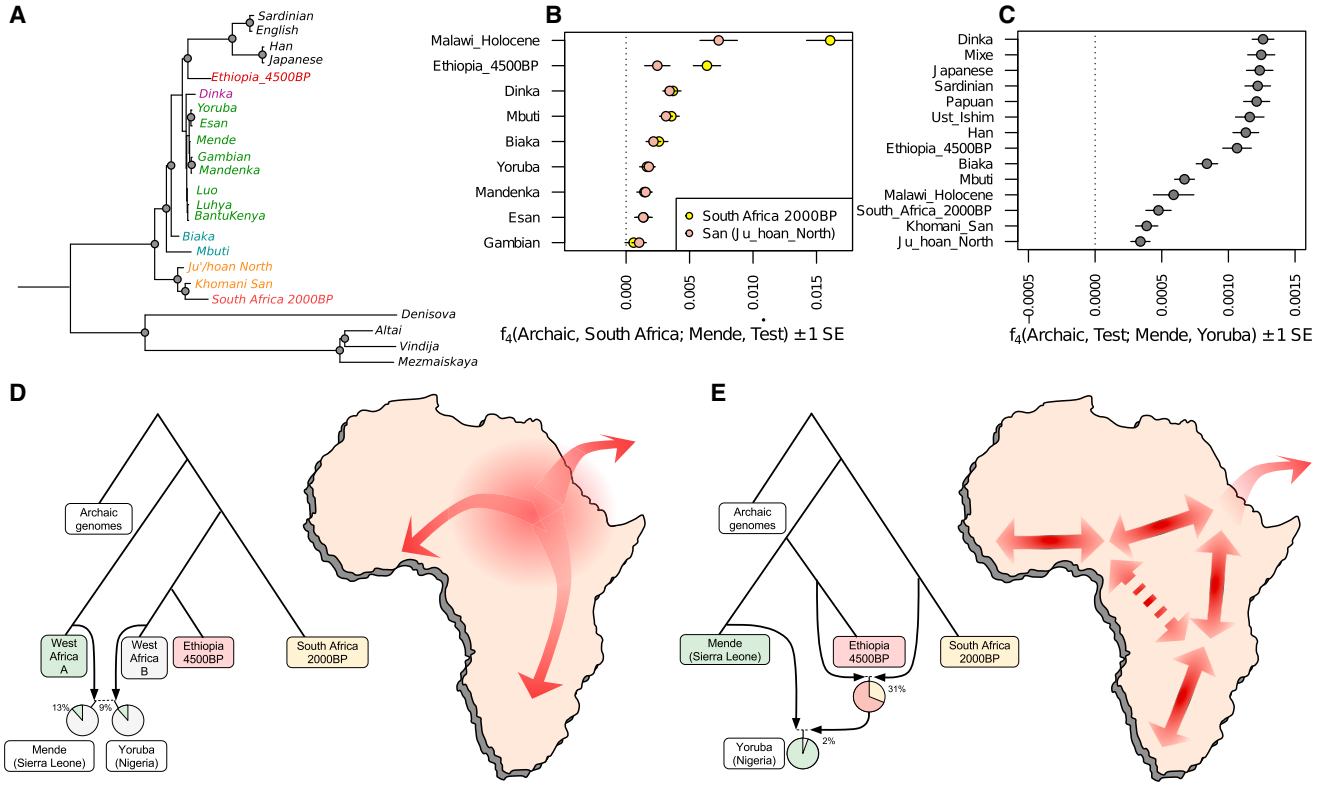


Figure 3. Mixture Events in the Deeper Population History of Continental African Lineages

(A) Maximum likelihood tree of genome sequences from present-day and ancient populations, excluding populations with evidence of asymmetrical allele sharing with non-Africans indicative of recent gene flow (Table S6). Nodes with bootstrap support >95% are indicated with a circle.

(B) A symmetry test of the hypothesis that ancient southern Africans are an outgroup lineage to other African populations, which can be rejected for most pairs.

(C) Asymmetry between western African Mende and Yoruba in the 1000 Genomes Project data is maximized in the Yoruba's excess affinity to eastern Africans and non-Africans but highly significant also for groups as distant as southern Africans.

(D) Admixture graph solution where Mende from Sierra Leone and Yoruba from Nigeria have ancestry from a basal western African lineage. The other source of western African ancestry is most closely related to eastern Africans and non-Africans (Figure S5D), which could be consistent with an expansion from eastern Africa. Note that the exact proportion "West Africa A" ancestry is not well constrained by the model, but the difference between Yoruba and Mende is highly significant (C).

(E) Admixture graph solution where the Yoruba have gene flow from a population related to both southern and eastern Africa, which could be consistent with a more complex pattern of isolation-by-distance on the continent.

previously established foragers during the initial phase of the Bantu expansion into eastern Africa (Crowther et al., 2017; Ribot et al., 2010), a barrier that broke down over time as mixture occurred. This parallels the patterns previously observed in genomic analyses of the Neolithic expansion into Europe (Haak et al., 2015; Skoglund et al., 2012) and the East Asian farming expansion into Remote Oceania (Skoglund et al., 2016). However, this process of delayed admixture did not always apply in Africa, as is evident in the absence of admixture from previously established hunter-gatherers in present-day Malawians.

Early Levantine Farmer-Related Admixture in an ~3,100-Year-Old Pastoralist from Tanzania

Western-Eurasian-related ancestry is pervasive in eastern Africa today (Pagani et al., 2012; Tishkoff et al., 2009), and the timing of this admixture has been estimated to be ~3,000 BP on average (Pickrell et al., 2014). We found that the ~3,100 BP individual (Tanzania_Luxmanda_3100BP), associated with a Savanna Pastoral Neolithic archeological tradition, could be modeled as having 38% ± 1% of her ancestry related to the nearly 10,000-year-old pre-pottery farmers of the Levant (Lazaridis et al., 2016), and we can exclude source populations related to early farmer populations in Iran and Anatolia. These results could be explained by migration into Africa from descendants of pre-pottery Levantine farmers or alternatively by a scenario in which both pre-pottery Levantine farmers and Tanzania_Luxmanda_3100BP descend from a common ancestral population that lived thousands of years earlier in Africa or the Near East. We fit the remaining approximately two-thirds of Tanzania_Luxmanda_3100BP as most closely related to the Ethiopia_4500BP ($p = 0.029$) or, allowing for three-way mixture, also from a source closely related to the Dinka ($p = 0.18$; the Levantine-related ancestry in this case was 39% ± 1%) (Table S4).

While these findings show that a Levant-Neolithic-related population made a critical contribution to the ancestry of present-day eastern Africans (Lazaridis et al., 2016), present-day Cushitic speakers such as the Somali cannot be fit simply as having Tanzania_Luxmanda_3100BP ancestry. The best fitting model for the Somali includes Tanzania_Luxmanda_3100BP ancestry, Dinka-related ancestry, and 16% ± 3% Iranian-Neolithic-related ancestry ($p = 0.015$). This suggests that ancestry related to the Iranian Neolithic appeared in eastern Africa after earlier gene flow related to Levant Neolithic populations, a scenario that is made more plausible by the genetic evidence of admixture of Iranian-Neolithic-related ancestry throughout the Levant by the time of the Bronze Age (Lazaridis et al., 2016) and in ancient Egypt by the Iron Age (Schuenemann et al., 2017).

Direct Evidence of Migration Bringing Pastoralism to Eastern and Southern Africa

In contrast to the Malawi and Zanzibar individuals, all three ancient southern Africans show affinities to the ancestry predominant in present-day Tuu speakers in the southern Kalahari more than to present-day Ju|'hoan speakers in the northern Kalahari (Figures S2B and S2C). However, the ~1,200 BP sample from the western Cape that is found in a pastoralist context has a specific similarity in clustering analyses to present-day Khoe-Khoe-speaking pastoralist populations such as the Nama (Figure 1B), and like them it has affinity to three groups: Khoe-San, western Eurasians, and eastern Africans. This supports the hypothesis that a non-Bantu-related population carried eastern African and Levantine ancestry to southern Africa by at least around 1,200 BP, providing direct evidence for claims previously made based on analysis of present-day populations (Pickrell et al., 2014).

We used our modeling framework to show that the South_Africa_1200BP pastoralist individual from the western Cape is consistent with being a mixture of just two streams of ancestry relative to non-southern African populations, with 40.3% ± 2.3% ancestry related to the Tanzania_Luxmanda_3100BP individual (54% ± 7% when restricting analysis to sequences with postmortem damage) and the remainder being related to the South_Africa_2000BP hunter-gatherers (Table S5). This supports the hypothesis that the Savanna Pastoral Neolithic archaeological tradition in eastern Africa is a plausible source for the spread of herding to southern Africa. Even the Ju|'hoan San group with the least genetic affinity to eastern Africans (Ju_hoan_North), have 9% ± 1% of their ancestry most closely related to Tanzania_Luxmanda_3100BP, consistent with previous findings that the ancestries of all present-day San and Khoe were affected by agropastoralist migrations in the last two millennia (Pickrell et al., 2014).

The Earliest Divergences among Modern Human Populations

Previous studies have suggested that the primary ancestry in the San is from a lineage that separated from all other lineages represented in modern humans today, before the latter separated from each other (Gronau et al., 2011; Veeramah et al., 2012). Such a model emerges when we automatically fit a tree without admixture to the data (Figure 3A), but we also find that a tree-like representation is a poor fit (Figure S4A), in the sense that ancient southern Africans who lived ~2,000 BP were not strictly an outgroup to extant lineages in other parts of sub-Saharan Africa. In particular, we find that ancient southern Africans, who have none of the eastern African admixture that is ubiquitous today, share significantly more alleles with present-day and ancient eastern Africans (including Dinka, Hadza, and Ethiopia_4500BP) than they do with present-day western Africans (Figure 3B; Table S6). Even within present-day western Africans, the genetic differences between Yoruba from Nigeria and the Mende from Sierra Leone are inconsistent with descent from a homogeneous ancestral population isolated from ancient southern Africans. The asymmetry between Yoruba and Mende is also observed with non-Africans but is no stronger than in eastern Africans (the most closely related Africans to the ancestral out-of-Africa population), and thus these signals are not driven by admixture from outside Africa and instead likely reflect demographic events entirely within Africa (Figure 3C; Table S6).

We carried out admixture graph modeling of the allele frequency correlations and found two parsimonious models that fit the data. The first posits that present-day western Africans harbor ancestry from a basal African lineage that contributed more to the Mende than it did to the Yoruba, with the other source of western African ancestry being related to eastern Africans and non-Africans (Figures 3D, S4, and S5; Table S7). The second model posits that long-range and long-standing gene flow has connected southern and eastern Africa to some groups in western Africa (e.g., the ancestors of the Yoruba) to a greater extent than to other groups in western Africa (e.g., the ancestors of the Mende) (Figure 3E) (Pleurdeau et al., 2012). The possible basal western African population lineage would represent the earliest known divergence of a modern human lineage that contributed a major proportion of ancestry to present-day humans. Such a lineage must have separated before the divergence of San ancestors, which is estimated to have begun on the order of 200–300 thousand years ago (Scally and Durbin, 2012). Such a model of basal western African ancestry might support the hypothesis that there has been ancient structure in the ancestry of present-day Africans, using a line of evidence independent from previous findings based on long haplotypes with deep divergences from other human haplotypes (Hammer et al., 2011; Lachance et al., 2012; Plagnol and Wall, 2006). One scenario consistent with this result could involve ancestry related to eastern Africans (and the out-of-Africa population) expanding into western Africa and mixing there with more basal lineages. Our genetic data do not support the theory that this putative basal lineage diverged prior to the ancestors of Neanderthals, since the African populations we analyze here are approximately symmetrically related to Neanderthals (Mallick et al., 2016; Prüfer et al., 2014).

A Selective Sweep Targeting a Taste Receptor Locus in Southern Africa

The availability of ancient African genomes provides an opportunity to search for genomic footprints of natural selection, manifested as regions of greater allele frequency differentiation between ancient and present-day populations than predicted by the genome-wide background. We compared the two ancient

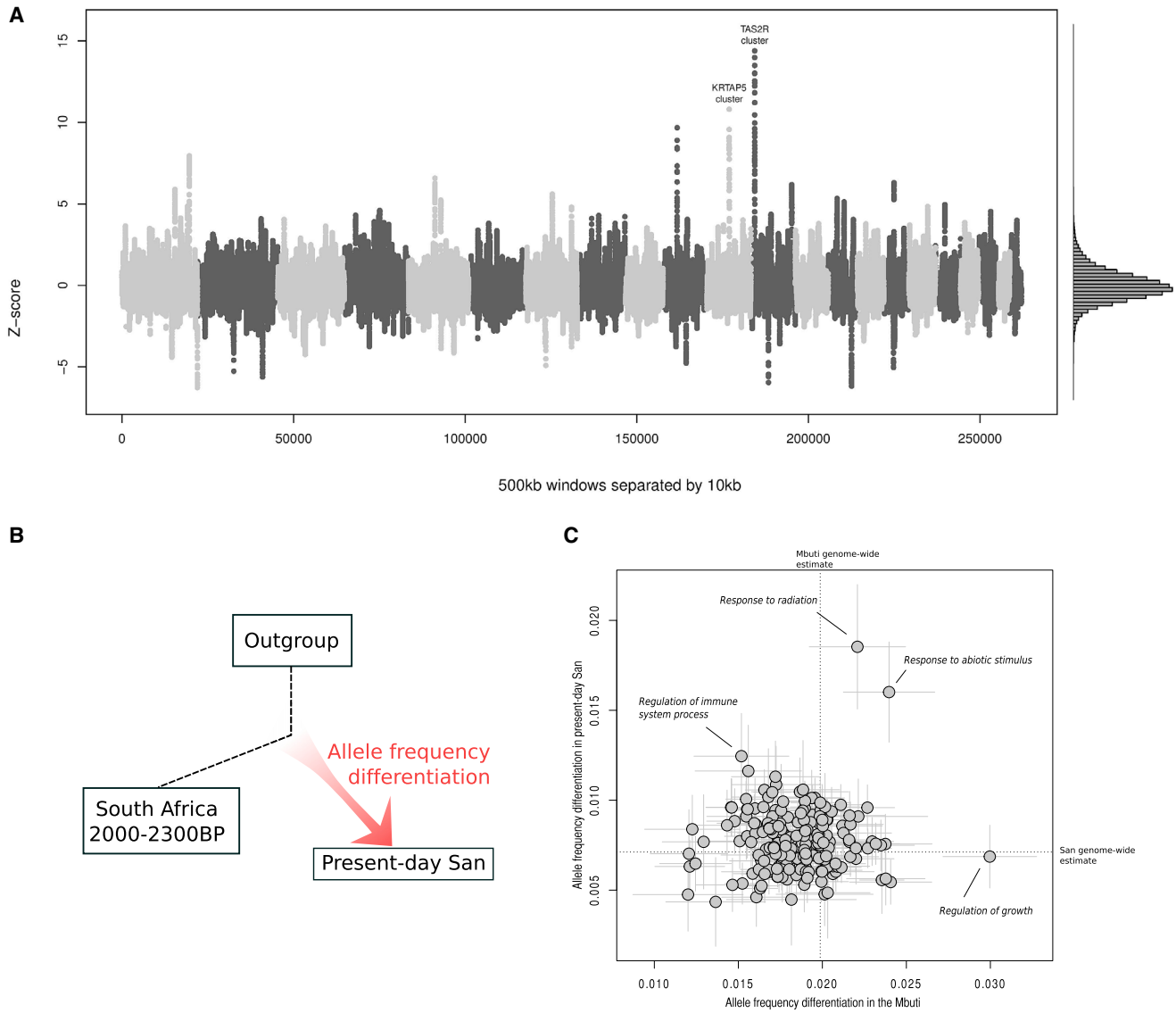


Figure 4. Ancient Genomes Provide Evidence of Natural Selection in Present-Day Southern African San Populations

We computed branch-specific allele frequency differentiation in six present-day high-coverage San genomes compared to a pool of two ~2,000 BP South African genomes as an outgroup using two approaches.

(A) We computed the statistic in windows of 500 kb separated by 10 kb. We also estimated genome-wide average and standard deviation of the statistic using windows separated by at least 5 Mb, and we transformed the genome-wide distribution of the sliding windows to be approximately normal (right). We observe outliers 15 standard deviations from the mean in a taste-receptor gene cluster on chromosome 12 and a secondary peak in the Keratin Associated Protein 4 gene cluster. The outgroup used was 4 Central African Mbuti genomes. See Table 2 for details on all major outlying regions.

(B) Illustration of the branch-specific allele frequency differentiation approach.

(C) y axis: We computed the statistic and block jackknife standard errors for 208 gene ontology categories with at least 50 genes each. The outgroup used was western Africans. x axis: As a control to confirm that outlier categories do not show larger magnitudes of allele frequency differentiation across populations, we replaced the present-day San with the central African Mbuti.

southern African ~2,000 BP shotgun sequence genomes to six present-day high-coverage San genomes with minimal recent mixture. The small number of ancient individuals does not permit inference of changing allele frequencies at single loci, so we performed a scan for high allele frequency differentiation in 500 kb windows with a step size of 10 kb. Using ~500 windows spaced at least 5 million base pairs apart as a null distribution, we found that the most differentiated locus was 15 standard deviations from the observed genome-wide mean and overlapped a cluster of eight taste-receptor genes on chromosome 12 (Figure 4A; Table 2). Taste-receptor genes have previously been identified as targets of natural selection in humans, as they modulate the ability to detect poisonous compounds in plants (Campbell et al., 2012).

Table 2. Top Five Candidate Regions Identified in Genome-wide Scan for Selective Sweeps in Present-Day San Populations in Southern Africa Compared to Ancient Genomes

Rank	Chrom.	start-end	f_3-statistic	Z-score	Genes in top 500 kb window in peak region
1	12	11,123,548-11,623,548	0.163	14.3	TAS2R43, PRH1-PRR4, TAS2R20, TAS2R50, TAS2R42, TAS2R46, TAS2R30, TAS2R31, PRB1, PRB2, PRB3, PRB4, LOC100129361, TAS2R19
2	11	71,208,258-71,708,258	0.125	10.8	LOC100129216, KRTAP5-7, DEFB108B, KRTAP5-8, NADSYN1, KRTAP5-9, FAM86C1, RNF121, ALG1L9P, LOC100133315, KRTAP5-10, KRTAP5-11
3	10	46,069,893-46,569,893	0.113	9.7	DQ577099, PTPN20B, PTPN20A, ZFAND4, AGAP4, DQ588224, FAM21C
4	1	224,960,062-225,460,062	0.095	8.0	DNAH14
5	5	82,375,629-82,875,629	0.08	6.6	VCAN, XRCC4

Polygenic Adaptation

Natural selection on phenotypic traits in humans is expected to only occasionally take the form of sweeps on a single locus, instead acting on multiple genes simultaneously to drive phenotypic adaptation (Coop et al., 2009). While a lack of genome-wide association studies in eastern and southern Africans has left the genetic basis of phenotypic traits far less well documented than it is for other populations, a variety of studies have linked broad functional classes of genes to phenotypic traits. To test for evidence of selection on specific functional categories of genes in present-day San since the divergence of the two ancient genomes from southern Africa (Figure 4B), we estimated allele frequency differentiation for 208 gene ontology categories with 50 or more genes in each, and we computed weighted block jackknife standard errors. The functional category that displays the most extreme allele frequency differentiation between present-day San and ancient southern Africans is "response to radiation" (Z = 3.3 compared to the genome-wide average). To control for the possibility that genes in this category show an inflated allele frequency differentiation in general, we computed the same statistic for the Mbuti central African rainforest hunter-gatherer group but found no evidence for selection affecting the response to radiation category (Figure 4C). Instead, the top category for the Mbuti is "response to growth," suggesting the possibility that the small stature of rainforest hunter-gatherer populations such as the Mbuti may be an acquired adaptation (although we have no ancient central African genome and thus no information about the time frame of selection). We speculate that the signal for selection in the response to radiation category in the San could be due to exposure to sunlight associated with the life of the ‡Khomani and Ju|'hoan North people in the Kalahari Basin, which has become a refuge for hunter-gatherer populations in the last millenia due to encroachment by pastoralist and agriculturalist groups (Morris, 2002).

DISCUSSION

This study, which multiplies by 16-fold the number of individuals with genome-wide ancient DNA data from sub-Saharan Africa, highlights the power of ancient African genomes to provide insights into prehistoric events that are difficult to discern based solely on analysis of present-day genomes. We reveal the presence of a hitherto unknown cline of geographically structured hunter-gatherer populations stretching from Ethiopia to South Africa, which we show existed prior to the great population transformations that occurred in the last few thousand years in association with the spread of herders and farmers. We also document deeper structure in western Africa, possibly predating the divergence of the ancestors of southern African hunter-gatherers from other population lineages. We finally provide case examples of how populations in eastern and southern Africa were transformed by the spread of food producers and show how the process gave rise to interactions with the previously established hunter-gatherers, with the outcomes ranging from no detectable mixture in present-day populations to substantial mixture. Our documentation of a radically different landscape of human populations before and after the spread of food producers highlights the difficulty of reconstructing the African past based solely on analysis of present-day populations and the importance of using ancient DNA to study deep African population history in an era in which technological improvements have now made this feasible. It is clear that ancient DNA studies with larger sample sizes and covering a broader chronological and geographic range have the potential to make major progress in improving our understanding of African prehistory.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Terminology
  - Panga ya Saidi Cave, Kilifi County, Kenya (n = 1)
  - Makangale Cave, Pemba, Tanzania (n = 2)
  - Kuumbi Cave, Zanzibar, Tanzania (n = 1)
  - Luxmanda, Babati District, Tanzania (n = 1)
  - Hora, Malawi (n = 2)
  - Fingira, Malawi (n = 3)
  - Chencherere II, Malawi (n = 2)
  - St. Helena Bay, South Africa (n = 1)
  - Faraoskop, South Africa (n = 1)
  - Kasteelberg, South Africa (n = 1)

- METHOD DETAILS
  - Direct AMS ^{14}C Bone Dates
  - Ancient DNA sample processing in Leipzig: St. Helena Bay sample
  - Ancient DNA sample processing in Tübingen: Faraoskop and Kasteelberg samples
  - Ancient DNA sample processing in Dublin: Malawi samples
  - Ancient DNA sample processing in Boston: Tanzania samples, Kenya samples, and Malawi sample powder
  - Shotgun genome sequencing
  - In-solution nuclear target enrichment
  - Genotyping and initial processing of 34 present-day individuals from Malawi
  - Data processing and preparation
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Principal component analysis and ADMIXTURE clustering analyses
  - Symmetry statistics and admixture tests
  - Y chromosome and mitochondrial haplogroups
  - Ancestry model and estimates with *qpAdm*
  - Maximum likelihood tree model
  - Testing a tree-like model of African population history
  - Testing admixture graph models of African population history
  - Automated grafting of populations onto a skeleton admixture graph
  - Support for a single out-of-Africa founding population
  - Date of admixture between expanding agriculturalists and previously established foragers
  - Evidence for selective sweeps in the ancestry of present-day San
  - Evidence for polygenic selection
- DATA AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2017.08.049.

An audio PaperClip is available at http://dx.doi.org/10.1016/j.cell.2017.08.049#mmc8.

AUTHOR CONTRIBUTIONS

Conceptualization, P.S., J.C.T., M.E.P., V.M.H., R.R., A.G.M., N.B., R.P., J.K., and D.R.; Formal Analysis, P.S., S.M., A.P., and I.O.; Investigation, P.S., A.M., K.Si., M.Ha., T.S., N.R., A.H., M.F., E.H., M.Mi., K.St., and J.I.C.-R.; Resources, J.C.T., M.E.P., C.C., A.C., E.G.-C., A.O.G., K.M.G., I.T.H., G.H., R.H., M.Ho., S.L., A.Z.P.M., J.P., C.S., M.G.T., R.T., M.W., V.M.H., A.G.M., N.B., and N.P.; Data Curation, P.S., M.Ha., N.R., S.M., A.P., I.O., M.F., E.H., M.Me., K.St., D.J.K., N.P., and D.R.; Writing, P.S. and D.R.; Supervision, V.M.H., M.Me., S.P., N.P., N.N., N.B., R.P., J.K., and D.R.

ACKNOWLEDGMENTS

Permission to analyze the remains from Kenya and Tanzania was granted by the National Museums of Kenya; the Antiquities Division of the Ministry of Natural Resources and Tourism, Tanzania; and the Zanzibar Department of Museums and Antiquities. Permission to analyze the remains from Malawi was granted by the Malawi Department of Museums and Monuments and the University of Cape Town. We thank I. Lazaridis, M. Lipson, I. Mathieson, and S. Tishkoff for discussions and I. Kucukkalipci and K. Majander for laboratory support. P.S. was supported by the Wenner-Gren Foundation and the Swedish Research Council (VR grant 2014-453). J.C.T was supported by a grant from the Program for the Enhancement of Research at Emory University. J.K. and A.M. were supported by the DFG grant KR 4015/1-1 and the Max Planck Society. K.Si. was supported by NSF grant BCS-1613577. M.Ha., A.H., M.Me., and S.P. were supported by the Max Planck Society. A.G.M. and J.P. were supported by the National Research Foundation of South Africa. R.R. was supported by the South African Medical Research Council. N.B. was supported by ERC starting grant SEALINKS (206148), and R.P. was supported by ERC starting grant ADNABIOARC (263441). M.G.T. was supported by Wellcome Trust Senior Investigator Award (grant number 100719/Z/12/Z). D.R. was supported by NIH grant GM100233 and by NSF HOMINID BCS-1032255 and is a Howard Hughes Medical Institute investigator. The laboratory at Penn State was supported by the NSF Archaeometry program (BCS-1460369, D.J.K.).

Received: December 12, 2016
Revised: July 1, 2017
Accepted: August 29, 2017
Published: September 21, 2017

REFERENCES

Alexander, D.H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Res. *19*, 1655–1664.

Behar, D.M., van Oven, M., Rosset, S., Metspalu, M., Loogväli, E.-L., Silva, N.M., Kivisild, T., Torroni, A., and Villems, R. (2012). A "Copernican" reassessment of the human mitochondrial DNA tree from its root. Am. J. Hum. Genet. *90*, 675–684.

Briggs, A.W., and Heyn, P. (2012). Preparation of next-generation sequencing libraries from damaged DNA. Methods Mol Biol. *840*, 143–154.

Briggs, A.W., Stenzel, U., Meyer, M., Krause, J., Kircher, M., and Pääbo, S. (2010). Removal of deaminated cytosines and detection of in vivo methylation in ancient DNA. Nucleic Acids Res. *38*, e87.

Bronk Ramsey, C. (2009). Bayesian analysis of radiocarbon dates. Radiocarbon *51*, 337–360.

Brown, T.A., Nelson, D.E., Vogel, J.S., and Southon, J.R. (1988). Improved collagen extraction by modified Longin method. Radiocarbon *30*, 171–177.

Busby, G.B.J., Band, G., Si Le, Q., Jallow, M., Bougama, E., Mangano, V.D., Amenga-Etego, L.N., Enimil, A., Apinjoh, T., Ndila, C.M., et al.; Malaria Genomic Epidemiology Network (2016). Admixture into and within sub-Saharan Africa. eLife *5*, e15266.

Campbell, M.C., Ranciaro, A., Froment, A., Hirbo, J., Omar, S., Bodo, J.-M., Nyambo, T., Lema, G., Zinshteyn, D., and Drayna, D. (2012). Evolution of functionally diverse alleles associated with PTC bitter taste sensitivity in Africa. Mol. Biol. Evol. *29*, 1141–1153.

Cann, R.L., Stoneking, M., and Wilson, A.C. (1987). Mitochondrial DNA and human evolution. Nature *325*, 31–36.

Chami, F. (2009). Zanzibar and the Swahili coast from c. 30,000 years ago (E&D Vision Pub.).

Chami, F., Khator, J., and Hamis Ali, A. (2009). The excavation of Mapangani cave, Pemba island, Zanzibar. In Studies in the African Past, F.A. Chami and C. Radimilahy, eds. (University of Dar es Salaam Press), pp. 74–79.

Clark, J. (1972). Prehistoric origins (London: The Early History of Malawi Longman), pp. 17–27.

Clark, J.D. (1956). Prehistory in Nyasaland. The Nyasaland Journal *9*, 92–119.

Clark, J.D. (1959). The prehistory of southern Africa (Penguin Books).

Clark, J.D. (1973). Archaeological investigation of a painted rock shelter at Mwana wa Chencherere, north of Dedza, Central Malaŵi, In July to September, 1972. Soc. Malawi J. *26*, 28–46.

Cole-King, P. (1973). Kukumba Mbiri Mu Malaŵi: A Summary of Archaeological Research to March, 1973 (Government Press).

Coop, G., Pickrell, J.K., Novembre, J., Kudaravalli, S., Li, J., Absher, D., Myers, R.M., Cavalli-Sforza, L.L., Feldman, M.W., and Pritchard, J.K. (2009). The role of geography in human adaptation. PLoS Genet. *5*, e1000500.

Crader, D.C. (1984a). Faunal remains from Chencherere II rock shelter, Malawi. S. Afr. Archaeol. Bull. *39*, 37–52.

Crader, D.C. (1984b). Hunters in Iron Age Malaẇi: The Zooarchaeology of Chencherere Rockshelter (Lilongwe: Malaẇi Government Ministry of Education and Culture).

Crowther, A., Prendergast, M.E., Fuller, D.Q., and Boivin, N. (2017). Subsistence mosaics, forager-farmer interactions and the transition to food production in eastern Africa. Quat. Int. http://dx.doi.org/10.1016/j.quaint.2017.01.014.

Crowther, A., Veall, M.-A., Boivin, N., Horton, M., Kotarba-Morley, A., Fuller, D.Q., Fenn, T., Haji, O., and Matheson, C.D. (2015). Use of Zanzibar copal (*Hymenaea verrucosa* Gaertn.) as incense at Unguja Ukuu, Tanzania in the 7–8th century CE: chemical insights into trade and Indian Ocean interactions. J. Archaeol. Sci. *53*, 374–390.

Dabney, J., Knapp, M., Glocke, I., Gansauge, M.-T., Weihmann, A., Nickel, B., Valdiosera, C., García, N., Pääbo, S., Arsuaga, J.-L., and Meyer, M. (2013). Complete mitochondrial genome sequence of a Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. Proc. Natl. Acad. Sci. USA *110*, 15758–15763.

Dabney, J., and Meyer, M. (2012). Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. Biotechniques *52*, 87–94.

DeBusk, G.H. (1997). The distribution of pollen in the surface sediments of Lake Malawi, Africa, and the transport of pollen in large lakes. Rev. Palaeobot. Palynol. *97*, 123–153.

Fleisher, J., and LaViolette, A. (2013). The early Swahili trade village of Tumbe, Pemba Island, Tanzania, AD 600–950. Antiquity *87*, 1151–1168.

Fu, Q., Hajdinjak, M., Moldovan, O.T., Constantin, S., Mallick, S., Skoglund, P., Patterson, N., Rohland, N., Lazaridis, I., Nickel, B., et al. (2015). An early modern human from Romania with a recent Neanderthal ancestor. Nature *524*, 216–219.

Fu, Q., Li, H., Moorjani, P., Jay, F., Slepchenko, S.M., Bondarev, A.A., Johnson, P.L.F., Aximu-Petri, A., Prüfer, K., de Filippo, C., et al. (2014). Genome sequence of a 45,000-year-old modern human from western Siberia. Nature *514*, 445–449.

Gansauge, M.-T., and Meyer, M. (2013). Single-stranded DNA library preparation for the sequencing of ancient or damaged DNA. Nat. Protoc. *8*, 737–748.

Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M.H.Y., et al. (2010). A draft sequence of the Neandertal genome. Science *328*, 710–722.

Gronau, I., Hubisz, M.J., Gulko, B., Danko, C.G., and Siepel, A. (2011). Bayesian inference of ancient human demography from individual genome sequences. Nat. Genet. *43*, 1031–1034.

Gurdasani, D., Carstensen, T., Tekola-Ayele, F., Pagani, L., Tachmazidou, I., Hatzikotoulas, K., Karthikeyan, S., Iles, L., Pollard, M.O., Choudhury, A., et al. (2015). The African Genome Variation Project shapes medical genetics in Africa. Nature *517*, 327–332.

Haak, W., Lazaridis, I., Patterson, N., Rohland, N., Mallick, S., Llamas, B., Brandt, G., Nordenfelt, S., Harney, E., and Stewardson, K. (2015). Massive migration from the steppe is a source for Indo-European languages in Europe. arXiv preprint arXiv:150202783.

Hammer, M.F., Woerner, A.E., Mendez, F.L., Watkins, J.C., and Wall, J.D. (2011). Genetic evidence for archaic admixture in Africa. Proc. Natl. Acad. Sci. USA *108*, 15123–15128.

Helm, R., Crowther, A., Shipton, C., Tengeza, A., Fuller, D., and Boivin, N. (2012). Exploring agriculture, interaction and trade on the eastern African littoral: preliminary results from Kenya. Azania *47*, 39–63.

Helm, R.M. (2000). Conflicting histories: the archaeology of the iron-working, farming communities in the central and southern coast region of Kenya (University of Bristol).

Henn, B.M., Gignoux, C.R., Jobin, M., Granka, J.M., Macpherson, J.M., Kidd, J.M., Rodríguez-Botigué, L., Ramachandran, S., Hon, L., Brisbin, A., et al. (2011). Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. Proc. Natl. Acad. Sci. USA *108*, 5154–5162.

Hogg, A.G., Hua, Q., Blackwell, P.G., Niu, M., Buck, C.E., Guilderson, T.P., Heaton, T.J., Palmer, J.G., Reimer, P.J., and Reimer, R.W. (2013). SHCal13 Southern Hemisphere calibration, 0–50,000 years cal BP. Radiocarbon *55*, 1889–1903.

Jostins, L., Xu, Y., McCarthy, S., Ayub, Q., Durbin, R., Barrett, J., and Tyler-Smith, C. (2014). YFitter: Maximum likelihood assignment of Y chromosome haplogroups from low-coverage sequence data. arXiv preprint arXiv:14077988.

Juma, A. (2004). Unguja Ukuu on Zanzibar: An archaeological study of early urbanism (Uppsala University).

Kennett, D.J., Plog, S., George, R.J., Culleton, B.J., Watson, A.S., Skoglund, P., Rohland, N., Mallick, S., Stewardson, K., Kistler, L., et al. (2017). Archaeogenomic evidence reveals prehistoric matrilineal dynasty. Nat Commun. *8*, 14115.

Kircher, M. (2012). Analysis of high-throughput ancient DNA sequencing data. In Ancient DNA, B. Shapiro and M. Hofreiter, eds. (Springer), pp. 197–228.

Kircher, M., Sawyer, S., and Meyer, M. (2011). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. *40*, e3.

Korlević, P., Gerber, T., Gansauge, M.-T., Hajdinjak, M., Nagel, S., Aximu-Petri, A., and Meyer, M. (2015). Reducing microbial and human contamination in DNA extractions from ancient bones and teeth. Biotechniques *59*, 87–93.

Lachance, J., Vernot, B., Elbers, C.C., Ferwerda, B., Froment, A., Bodo, J.M., Lema, G., Fu, W., Nyambo, T.B., Rebbeck, T.R., et al. (2012). Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse African hunter-gatherers. Cell *150*, 457–469.

Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D.C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., et al. (2016). Genomic insights into the origin of farming in the ancient Near East. Nature *536*, 419–424.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., Sudmant, P.H., Schraiber, J.G., Castellano, S., Lipson, M., et al. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature *513*, 409–413.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics *25*, 2078–2079.

Llorente, M.G., Jones, E.R., Eriksson, A., Siska, V., Arthur, K.W., Arthur, J.W., Curtis, M.C., Stock, J.T., Coltorti, M., Pieruccini, P., et al. (2015). Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern Africa. Science *350*, 820–822.

Loh, P.-R., Lipson, M., Patterson, N., Moorjani, P., Pickrell, J.K., Reich, D., and Berger, B. (2013). Inference of admixture parameters in human populations using weighted linkage disequilibrium. Genetics *193*, 1233–1254.

Lohse, J.C., Madsen, D.B., Culleton, B.J., and Kennett, D.J. (2014). Isotope paleoecology of episodic mid-to-late Holocene bison population expansions in the Southern Plains, USA. Quat. Sci. Rev. *102*, 14–26.

Longin, R. (1971). New method of collagen extraction for radiocarbon dating. Nature *230*, 241–242.

Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., et al. (2016). The Simons Genome Diversity Project: 300 genomes from 142 diverse populations. Nature *538*, 201–206.

Manhire, A. (1993). A report on the excavations at Faraoskop Rock Shelter in the Graafwater district of the south-western Cape. Southern African Field Archaeology *2*, 3–23.

Maricic, T., Whitten, M., and Pääbo, S. (2010). Multiplexed DNA sequence capture of mitochondrial genomes using PCR products. PLoS ONE *5*, e14004.

Marshall, F., Stewart, K., and Barthelme, J. (1984). Early domestic stock at Dongodien in northern Kenya. Azania *19*, 120–127.

Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A., Harney, E., Stewardson, K., Fernandes, D., Novak, M., et al. (2015). Genome-wide patterns of selection in 230 ancient Eurasians. Nature *528*, 499–503.

Meyer, M., Fu, Q., Aximu-Petri, A., Glocke, I., Nickel, B., Arsuaga, J.-L., Martínez, I., Gracia, A., de Castro, J.M.B., Carbonell, E., and Pääbo, S. (2014). A mitochondrial genome sequence of a hominin from Sima de los Huesos. Nature *505*, 403–406.

Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harb Protoc *2010*, pdb.prot5448.

Meyer, M., Kircher, M., Gansauge, M.-T., Li, H., Racimo, F., Mallick, S., Schraiber, J.G., Jay, F., Prüfer, K., de Filippo, C., et al. (2012). A high-coverage genome sequence from an archaic Denisovan individual. Science *338*, 222–226.

Miller, S.F. (1971). The age of Nachikufan industries in Zambia. S. Afr. Archaeol. Bull. *26*, 143–146.

Moorjani, P., Patterson, N., Hirschhorn, J.N., Keinan, A., Hao, L., Atzmon, G., Burns, E., Ostrer, H., Price, A.L., and Reich, D. (2011). The history of African gene flow into Southern Europeans, Levantines, and Jews. PLoS Genet. *7*, e1001373.

Morris, A.G. (2002). Isolation and the origin of the Khoisan: Late Pleistocene and Early Holocene human evolution at the southern end of Africa. Hum. Evol. *17*, 231–240.

Morris, A.G., Heinze, A., Chan, E.K., Smith, A.B., and Hayes, V.M. (2014). First ancient mitochondrial human genome from a prepastoralist southern African. Genome Biol. Evol. *6*, 2647–2653.

Morris, A.G., and Ribot, I. (2006). Morphometric cranial identity of prehistoric Malawians in the light of sub-Saharan African diversity. Am. J. Phys. Anthropol. *130*, 10–25.

Odner, K. (1972). Excavations at Narosura, a Stone Bowl Site in the Southern Kenya Highlands. Azania *7*, 25–92.

Pagani, L., Kivisild, T., Tarekegn, A., Ekong, R., Plaster, C., Gallego Romero, I., Ayub, Q., Mehdi, S.Q., Thomas, M.G., Luiselli, D., et al. (2012). Ethiopian genetic diversity reveals linguistic stratification and complex influences on the Ethiopian gene pool. Am. J. Hum. Genet. *91*, 83–96.

Parkington, J., and Dlamini, N. (2015). First People: Ancestors of the San (Creda Communications).

Patin, E., Lopez, M., Grollemund, R., Verdu, P., Harmant, C., Quach, H., Laval, G., Perry, G.H., Barreiro, L.B., Froment, A., et al. (2017). Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. Science *356*, 543–546.

Patterson, N., Moorjani, P., Luo, Y., Mallick, S., Rohland, N., Zhan, Y., Genschoreck, T., Webster, T., and Reich, D. (2012). Ancient admixture in human history. Genetics *192*, 1065–1093.

Patterson, N., Price, A.L., and Reich, D. (2006). Population structure and eigenanalysis. PLoS Genet. *2*, e190.

Peltzer, A., Jäger, G., Herbig, A., Seitz, A., Kniep, C., Krause, J., and Nieselt, K. (2016). EAGER: efficient ancient genome reconstruction. Genome Biol. *17*, 60.

Phillipson, D.W. (1976). The Early Iron Age in eastern and southern Africa: a critical re-appraisal. Azania *11*, 1–23.

Pickrell, J.K., Patterson, N., Barbieri, C., Berthold, F., Gerlach, L., Güldemann, T., Kure, B., Mpoloka, S.W., Nakagawa, H., Naumann, C., et al. (2012). The genetic prehistory of southern Africa. Nat. Commun. *3*, 1143.

Pickrell, J.K., Patterson, N., Loh, P.-R., Lipson, M., Berger, B., Stoneking, M., Pakendorf, B., and Reich, D. (2014). Ancient west Eurasian ancestry in southern and eastern Africa. Proc. Natl. Acad. Sci. USA *111*, 2632–2637.

Pickrell, J.K., and Pritchard, J.K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. PLoS Genet. *8*, e1002967.

Plagnol, V., and Wall, J.D. (2006). Possible ancestral structure in human populations. PLoS Genet. *2*, e105.

Pleurdeau, D., Imalwa, E., Détroit, F., Lesur, J., Veldman, A., Bahain, J.-J., and Marais, E. (2012). "Of sheep and men": earliest direct evidence of caprine domestication in southern Africa at Leopard Cave (Erongo, Namibia). PLoS ONE *7*, e40340.

Prendergast, M.E., Mabulla, A.Z.P., Grillo, K.M., Broderick, L.G., Seitsonen, O., Gidna, A.O., and Gifford-Gonzalez, D. (2013). Pastoral Neolithic sites on the southern Mbulu Plateau, Tanzania. Azania *48*, 498–520.

Prendergast, M.E., Rouby, H., Punnwong, P., Marchant, R., Crowther, A., Kourampas, N., Shipton, C., Walsh, M., Lambeck, K., and Boivin, N.L. (2016). Continental island formation and the archaeology of defaunation on Zanzibar, eastern Africa. PLoS ONE *11*, e0149565.

Prüfer, K., Racimo, F., Patterson, N., Jay, F., Sankararaman, S., Sawyer, S., Heinze, A., Renaud, G., Sudmant, P.H., de Filippo, C., et al. (2014). The complete genome sequence of a Neanderthal from the Altai Mountains. Nature *505*, 43–49.

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

Reich, D., Green, R.E., Kircher, M., Krause, J., Patterson, N., Durand, E.Y., Viola, B., Briggs, A.W., Stenzel, U., Johnson, P.L.F., et al. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. Nature *468*, 1053–1060.

Reich, D., Patterson, N., Campbell, D., Tandon, A., Mazieres, S., Ray, N., Parra, M.V., Rojas, W., Duque, C., Mesa, N., et al. (2012). Reconstructing Native American population history. Nature *488*, 370–374.

Reich, D., Thangaraj, K., Patterson, N., Price, A.L., and Singh, L. (2009). Reconstructing Indian population history. Nature *461*, 489–494.

Reimer, P.J., Bard, E., Bayliss, A., Beck, J.W., Blackwell, P.G., Bronk Ramsey, C., Buck, C.E., Cheng, H., Edwards, R.L., Friedrich, M., et al. (2013). IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP. Radiocarbon *55*, 1869–1887.

Renaud, G., Kircher, M., Stenzel, U., and Kelso, J. (2013). freelbis: an efficient basecaller with calibrated quality scores for Illumina sequencers. Bioinformatics *29*, 1208–1209.

Renaud, G., Slon, V., Duggan, A.T., and Kelso, J. (2015). Schmutzi: estimation of contamination and endogenous mitochondrial consensus calling for ancient DNA. Genome Biol. *16*, 224.

Ribot, I., Morris, A.G., Sealy, J., and Maggs, T. (2010). Population history and economic change in the last 2000 years in KwaZulu-Natal, RSA. Southern African Humanities *22*, 89–112.

Robinson, K., and Sandelowsky, B. (1968). The Iron Age of northern Malawi: recent work. AZANIA. Journal of the British Institute in Eastern Africa *3*, 107–146.

Robinson, K.S.R. (1982). Iron Age of northern Malaẇi: an archaeological reconnaissance (Malaẇi Govt. Ministry of Education and Culture).

Rohland, N., Harney, E., Mallick, S., Nordenfelt, S., and Reich, D. (2015). Partial uracil-DNA-glycosylase treatment for screening of ancient DNA. Philos. Trans. R. Soc. Lond. B Biol. Sci. *370*, 20130624.

Rohland, N., and Hofreiter, M. (2007). Comparison and optimization of ancient DNA extraction. Biotechniques *42*, 343–352.

Russell, T., Silva, F., and Steele, J. (2014). Modelling the spread of farming in the Bantu-speaking regions of Africa: an archaeology-based phylogeography. PLoS ONE *9*, e87854.

Sadr, K. (2015). Livestock First Reached Southern Africa in Two Separate Events. PLoS ONE *10*, e0134215.

Sadr, K., Smith, A., Plug, I., Orton, J., and Mütti, B. (2003). Herders and foragers on Kasteelberg: interim report of excavations 1999-2002. S. Afr. Archaeol. Bull. *58*, 27–32.

Sandelowsky, B. (1972). Later Stone Age assemblages from Malawi and their technologies. PhD thesis (University of California, Berkeley).

Scally, A., and Durbin, R. (2012). Revising the human mutation rate: implications for understanding human evolution. Nat. Rev. Genet. *13*, 745–753.

Schlebusch, C. (2010). Issues raised by use of ethnic-group names in genome study. Nature *464*, 487.

Schlebusch, C.M., Skoglund, P., Sjödin, P., Gattepaille, L.M., Hernandez, D., Jay, F., Li, S., De Jongh, M., Singleton, A., Blum, M.G.B., et al. (2012). Genomic variation in seven Khoe-San groups reveals adaptation and complex African history. Science *338*, 374–379.

Schuenemann, V.J., Peltzer, A., Welte, B., van Pelt, W.P., Molak, M., Wang, C.-C., Furtwängler, A., Urban, C., Reiter, E., Nieselt, K., et al. (2017). Ancient Egyptian mummy genomes suggest an increase of Sub-Saharan African ancestry in post-Roman periods. Nat Commun *8*, 15694.

Sealy, J.C., Patrick, M.K., Morris, A.G., and Alder, D. (1992). Diet and dental caries among later stone age inhabitants of the Cape Province, South Africa. Am. J. Phys. Anthropol. *88*, 123–134.

Shipton, C., Crowther, A., Kourampas, N., Prendergast, M.E., Horton, M., Douka, K., Schwenninger, J.-L., Faulkner, P., Quintana Morales, E.M., Langley, M.C., et al. (2016). Reinvestigation of Kuumbi Cave, Zanzibar, reveals Later Stone Age coastal habitation, early Holocene abandonment and Iron Age reoccupation. Azania *51*, 197–233.

Shriver, M.D., Kennedy, G.C., Parra, E.J., Lawson, H.A., Sonpar, V., Huang, J., Akey, J.M., and Jones, K.W. (2004). The genomic distribution of population substructure in four populations using 8,525 autosomal SNPs. Hum. Genomics *1*, 274–286.

Sinclair, P., Juma, A., and Chami, F. (2006). Excavations at Kuumbi Cave on Zanzibar 2005. In The African Archaeology Network: Research in Progress, J. Kinahan and J.H.A. Kinahan, eds. (Dar es Salaam University Press), pp. 95–106.

Skoglund, P., Mallick, S., Bortolini, M.C., Chennagiri, N., Hünemeier, T., Petzl-Erler, M.L., Salzano, F.M., Patterson, N., and Reich, D. (2015). Genetic evidence for two founding populations of the Americas. Nature *525*, 104–108.

Skoglund, P., Malmström, H., Raghavan, M., Storå, J., Hall, P., Willerslev, E., Gilbert, M.T.P., Götherström, A., and Jakobsson, M. (2012). Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. Science *336*, 466–469.

Skoglund, P., Northoff, B.H., Shunkov, M.V., Derevianko, A.P., Pääbo, S., Krause, J., and Jakobsson, M. (2014a). In Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal, R.G. Klein, ed. (Proceedings of the National Academy of Sciences), pp. 2229–2234.

Skoglund, P., Posth, C., Sirak, K., Spriggs, M., Valentin, F., Bedford, S., Clark, G.R., Reepmeyer, C., Petchey, F., Fernandes, D., et al. (2016). Genomic insights into the peopling of the Southwest Pacific. Nature *538*, 510–513.

Skoglund, P., Sjödin, P., Skoglund, T., Lascoux, M., and Jakobsson, M. (2014b). Investigating population history using temporal genetic differentiation. Mol. Biol. Evol. *31*, 2516–2527.

Smith, A.B. (1992a). Kasteelberg. In Guide to Archaeological Sites in the South-western Cape, A.B. Smith and B. Mutti, eds. (Cape Town: South African Association of Archaeologists Conference), pp. 28–30.

Smith, A.B. (1992b). Pastoralism in Africa (Johannesburg: Witwatersrand University Press).

Smith, B.W. (1995). Rock art in south-central Africa: a study based on the pictographs of Dedza District, Malawi and Kasam District (Zambia: University of Cambridge).

Stafford, T.W., Hare, P.E., Currie, L., Jull, A.J.T., and Donahue, D.J. (1991). Accelerator radiocarbon dating at the molecular level. J. Archaeol. Sci. *18*, 35–72.

Stuiver, M., and Polach, H.A. (1977). Discussion: reporting of 14 C data. Radiocarbon *19*, 355–363.

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

Van Klinken, G.J. (1999). Bone collagen quality indicators for palaeodietary and radiocarbon measurements. J. Archaeol. Sci. *26*, 687–695.

van Oven, M., and Kayser, M. (2009). Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. Hum. Mutat. *30*, E386–E394.

Veeramah, K.R., Wegmann, D., Woerner, A., Mendez, F.L., Watkins, J.C., Destro-Bisol, G., Soodyall, H., Louie, L., and Hammer, M.F. (2012). An early divergence of KhoeSan ancestors from those of other modern humans is supported by an ABC-based analysis of autosomal resequencing data. Mol. Biol. Evol. *29*, 617–630.

Wang, Y., and Nielsen, R. (2012). Estimating population divergence time and phylogeny from single-nucleotide polymorphisms data with outgroup ascertainment bias. Mol. Ecol. *21*, 974–986.

Weissensteiner, H., Pacher, D., Kloss-Brandstätter, A., Forer, L., Specht, G., Bandelt, H.-J., Kronenberg, F., Salas, A., and Schönherr, S. (2016). HaploGrep 2: mitochondrial haplogroup classification in the era of high-throughput sequencing. Nucleic Acids Res. *44*, W58–63.

Yi, X., Liang, Y., Huerta-Sanchez, E., Jin, X., Cuo, Z.X.P., Pool, J.E., Xu, X., Jiang, H., Vinckenbosch, N., Korneliussen, T.S., et al. (2010). Sequencing of 50 human exomes reveals adaptation to high altitude. Science *329*, 75–78.

Zubieta, L.F. (2016). Learning through practise: Chewa women's roles and the use of rock art in passing on cultural knowledge. J. Anthropol. Archaeol. *43*, 13–28.