



Genomic perspectives on human dispersals during the Holocene

Mark Stoneking, Leonardo Arias, Dang Liu, Sandra Oliveira, Irina Pugach,
Jae Joseph Russell B Rodriguez

► To cite this version:

Mark Stoneking, Leonardo Arias, Dang Liu, Sandra Oliveira, Irina Pugach, et al.. Genomic perspectives on human dispersals during the Holocene. *Proceedings of the National Academy of Sciences of the United States of America*, 2023, 120 (4), 10.1073/pnas.2209475119 . hal-03952281

HAL Id: hal-03952281

<https://cnrs.hal.science/hal-03952281>

Submitted on 30 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License



Genomic perspectives on human dispersals during the Holocene

Mark Stoneking^{a,b,1,2}, Leonardo Arias^{a,c,1} , Dang Liu^{a,d,1} , Sandra Oliveira^{a,e,1} , Irina Pugach^{a,1}, and Jae Joseph Russell B. Rodriguez^{a,f,g,1}

Edited by Clark Larsen, The Ohio State University, Columbus, OH; received August 2, 2022; accepted October 25, 2022

Nearly 20 y ago, Jared Diamond and Peter Bellwood reviewed the evidence for the associated spread of farming and large language families by the demographic expansions of farmers. Since then, advances in obtaining and analyzing genomic data from modern and ancient populations have transformed our knowledge of human dispersals during the Holocene. Here, we provide an overview of Holocene dispersals in the light of genomic evidence and conclude that they have a complex history. Even when there is a demonstrated connection between a demographic expansion of people, the spread of agriculture, and the spread of a particular language family, the outcome in the results of contact between expanding and resident groups is highly variable. Further research is needed to identify the factors and social circumstances that have influenced this variation and complex history.

genomics | dispersal | Holocene | humans

The domestication of plants and animals was a crucial development in human evolution, as it allowed for unprecedented levels of population growth and expansion, as well as greatly increasing the burden of infectious diseases. Various plants and animals were independently domesticated in several locations around the world starting ~9–11 thousand years ago (kya), at the beginning of the Holocene. The transition from a hunter-gatherer to a farming lifestyle is known as the Neolithic, and archaeological investigations have documented the origins and spread of farming. Similarly, linguists have documented language families that have spread across large geographic areas (e.g., the Bantu, Austronesian, and Indo-European (IE) language families); these archaeological and linguistic investigations led to the proposal that the spread of these large language families was facilitated by the spread of farming (reviewed in refs. 1 and 2).

A critical question is how farming and languages spread: was it by migrations of farmers who brought their way of life and languages with them (i.e., demic diffusion), or was it via local hunter-gatherer groups adopting farming and language from neighboring farmers (i.e., cultural diffusion)? And if the answer is demic diffusion, what was the fate of the local hunter-gatherer groups: were they completely replaced, or were they at least partially assimilated into the expanding farmer groups? While various arguments based on archaeology, skeletal morphology, and linguistics have addressed this issue, ultimately, this is a question for genetics. Assuming that the homeland of the original farming group can be identified and that these people were genetically different from the groups into whose territory they expanded, then genetic investigations can reveal the extent to which modern groups

have farmer-derived vs. indigenous (pre-farming) ancestry (here we use the term *ancestry* to refer in a general way to the genetic contributions to populations; for a discussion of the use and misuse of this term, see refs. 3 and 4). Beginning with the seminal work of Cavalli-Sforza and colleagues on the role of cultural vs. demic expansions in the spread of agriculture and IE languages across Europe (5), genetic investigations have been used to address the same question for the spread of farming and/or language families elsewhere.

However, such genetic investigations are not without complications. As discussed in more detail by Diamond and Bellwood (2), these include clinal admixture between expanding farmers and hunter-gatherers, resulting in decreasing genetic contributions by farmers toward the periphery of expansions; some hunter-gatherer groups adopting farming by cultural diffusion; farmers reverting to hunting and gathering (if, for example, they enter areas that are unsuitable for the domesticates they are bringing); language shift by resident groups with little or no incorporation of genes from the expanding population; replacement of the original language spoken by farmers in their homeland after the expansion (leading to discrepancies between genes and languages); and expansions of hunter-gatherers. An additional complicating factor is colonialism and the associated genetic, demographic, and territorial impact on indigenous groups. Failure to take such complications into consideration can lead to incorrect conclusions concerning the role of demic vs. cultural diffusion processes in the spread of farming and/or languages.

Author affiliations: ^aDepartment of Evolutionary Genetics, Max Planck Institute for Evolutionary Anthropology, Leipzig D04103, Germany; ^bUniversité Lyon 1, CNRS, Laboratoire de Biométrie et Biologie Evolutive, UMR 5558, Villeurbanne, France; ^cLeiden University Centre for Linguistics, Leiden 2311 BE, The Netherlands; ^dHuman Evolutionary Genetics Unit, Institut Pasteur, UMR 2000, CNRS, Paris, France; ^eComputational and Molecular Population Genetics, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland; ^fGenetics and Molecular Biology Division, Institute of Biological Sciences, College of Arts and Sciences, University of the Philippines Los Baños, Laguna 4031, Philippines; and ^gDNA Analysis Laboratory, Natural Sciences Research Institute, College of Science, University of the Philippines Diliman, Quezon City 1101, Philippines

Author contributions: M.S., L.A., D.L., S.O., I.P., and J.J.R.B.R. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2023 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹M.S., L.A., D.L., S.O., I.P., and J.J.R.B.R. contributed equally to this work.

²To whom correspondence may be addressed. Email: stoneking@eva.mpg.de or mark.stoneking@univ-lyon1.fr.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2209475119/-DCSupplemental>.

Published January 17, 2023.

Here, we survey the genomic evidence for human dispersals during the Holocene in various geographic regions of the world, focusing on proposed expansions linked to agriculture. Developments in obtaining and analyzing genome-wide data in the nearly 20 y since Diamond and Bellwood's survey (2), particularly advances in ancient DNA analyses, are providing new insights into such dispersals and the various complicating factors mentioned above. Given the space constraints of this perspective, we do not include all dispersals or relevant literature; instead, we focus on what we believe are the most significant/interesting dispersal events for the points we wish to make. Maps depicting these dispersals are provided (*SI Appendix, Figs. S1–S5*), although we caution that arrows on maps generally imply much more certainty about routes of expansions than the data actually support.

Africa

Fossil, genetic, and archaeological evidence all strongly favor an African origin for our species; while very deep population relationships can still be found among present-day African hunter-gatherers, there are also high levels of genetic homogeneity across large areas that reflect expansions of food-producing groups (6–8). Domesticated animals from southwest Asia (Middle East) first spread into North Africa ~8,000 y ago, gradually moving south (9, 10) and entering eastern Africa ~5 kya and southern Africa ~2 kya (11–13). There appear to be several centers of crop domestication in western Africa, east Sudan, and the Ethiopian highlands, beginning ~4 kya (10, 14).

Spread of Pastoralism/Farming into North and East Africa.

The potential demographic changes associated with the spread of food production in North and East Africa were greatly clarified by the availability of ancient DNA from these regions. In addition to sub-Saharan African ancestry, by 15 kya, North Africans also have Near Eastern ancestry, indicating that back to Africa gene flow predates the introduction of pastoralism or farming (15). The pre-Neolithic ancestry continues in Early Neolithic groups (~7 kya), while Late Neolithic North Africans (~5 kya) received additional gene flow from Iberia (16). Therefore, it seems that the Neolithic transition involved both cultural and demic diffusion, but further ancient genomes from this vast area are necessary.

In East Africa, the first ancient genome obtained (dating to 4.5 kya) revealed no traces of Eurasian ancestry but confirmed the existence of later Eurasian gene flow into East Africa from a population closely related to Early Neolithic farmers from the Middle East (17), previously identified and dated to ~3 kya based on present-day populations (18). A recent study of 41 Eastern Africans from the Later Stone Age, Pastoral Neolithic, and Iron Age, further inferred two phases of admixture: ~6–5 kya in Northeastern Africa between groups carrying non-African genetic ancestry (related to groups from the Levant or North Africa) and local northeast African groups and ~4 kya between this admixed group and Eastern African hunter-gatherers (11). This work thus supported several movements of food producers and also showed that admixture with foragers was common.

Bantu Expansion. The Bantu languages form a relatively homogeneous branch within the Niger-Congo language family and yet are spoken across a large part of sub-Saharan Africa by ~30% of

Africa's population (19). Even though the origin of these languages (in the Grassfields region around the border of eastern Nigeria and western Cameroon, where the highest linguistic diversity and early diverging branches are found) is broadly located where some agriculture domesticates arose (19, 20), the initial trigger of the Bantu expansion (~4–5 kya) could have been the climate-induced reduction of the rainforest in West Central Africa, rather than farming itself (21). The early phases of the Bantu expansion are associated with pottery and mixed subsistence economies; several plant and animal domesticates and iron were incorporated at later stages (21). The expansion was relatively rapid, reaching southern Africa by ~1.8 kya (12).

Genetic studies have demonstrated that Bantu-speaking groups from distant parts of the continent display a similar genetic profile and remarkable genetic homogeneity, as expected for a demographic expansion, while their close geographic neighbors who speak other languages show higher levels of differentiation (22–24). The ancestry of Bantu-speaking groups is largely derived from West Africa (23–25), although exceptions occur in areas still inhabited by hunter-gatherers, such as the Central African Rainforest and the Kalahari Desert and in eastern Africa, where Bantu speakers contacted Nilo-Saharan and Afro-Asiatic-speaking groups (6, 13).

Genetic evidence also informs on the routes of expansion of Bantu-speaking groups, showing that a “late split” model, in which the Bantu splits off into two migratory routes toward eastern and southern Africa following a passage through the rainforest, best explains the presently available data (26–28). And along the Indian Ocean Coast, genetic results support a rapid North–South dispersal (29) involving minimal or no admixture with resident populations in Mozambique (29) and Malawi (30).

By contrast, in southern Africa, Bantu-speaking groups show substantial amounts of autochthonous (Khoisan-related) ancestry (identified from foraging and pastoralist groups of the Kalahari), and Khoisan-speaking groups show substantial amounts of Bantu-related ancestry, indicating considerable interactions between them (6, 8)). These interactions were strongly sex biased, with primarily Khoisan-related maternal lineages found in Bantu-speaking groups and Bantu-related paternal lineages found in Khoisan-speaking groups (31). Moreover, the intensity of the sex bias increases from north to south in southern Africa, suggesting changing social circumstances influencing contact between groups (31).

In western Africa, archaeological and genetic evidence suggests that the Bantu expansion was not a single demic expansion but rather characterized by multiple expansion phases, with a population collapse inferred between 1.4–1.6 kya in the Congo rainforest (32). West Bantu groups have admixed extensively with hunter-gatherers in the Central African Rainforest, showing a similar signal of sex bias as in southern Africa (28, 33). In contrast, groups that presently occupy the savanna/grassland habitats to the south of the rainforest have primarily Bantu-related ancestry (28, 34).

Spread of Pastoralism from Eastern to Southern Africa.

Archaeological evidence indicates that pastoralism (sheep-herding) was introduced to southern Africa from eastern Africa ~2,000 y ago, before the arrival of farming (9, 12). Güldemann (35) additionally proposed that pastoralism spread together with the Khoe-Kwadi

language family, which might be related to the Sandawe language of East Africa. Yet, the Nama are the only Khoe-Kwadi group that currently practices pastoralism, suggesting substantial shifts in the lifestyle of Khoe-Kwadi groups.

Multiple genetic studies have found a signature of eastern African-related ancestry in southern African groups (8, 13, 18), which is usually attributed to the spread of pastoralism. However, Khoe-Kwadi groups have at most small amounts of East African ancestry, which is also found in other southern African groups (8, 13, 18), suggesting that the pastoral migrants extensively admixed with local hunter-gatherers, with genes flowing in both directions.

While ancient genomes are still scarce, the available data confirm inferences from modern populations suggesting that East African ancestry reached southern Africa and mixed with Khoisan-related ancestry before the arrival of Bantu-speaking groups (8, 36). In particular, East African ancestry is absent from 2 ky South African hunter-gatherers but is evident in a 1.2 ky individual found in a herder context in the Western Cape who lacks West African-related ancestry (30).

Europe

Europe was occupied by modern humans beginning at least 45 kya (37), and extensive investigations of both modern and ancient DNA from across Europe have revealed insights into the Neolithic transition and the impact of Bronze Age migrations of steppe pastoralists.

Neolithic Transition. European Neolithic sites are characterized by ceramics and various domesticated plants and animals, all probably domesticated starting ~11 kya in a region extending from Anatolia to the Near East (38, 39). While it is clear that the Neolithic in Europe has its source in this region, whether all aspects arrived as a complete “Neolithic package” or rather spread to/through Europe at different times and/or from different places is debated.

The Neolithic first appears in Europe ~9–10 kya, in Cyprus, Greece, and the Balkans (38, 39). Based on an extensive sample of radiocarbon dates (40, 41), farming likely spread via two main routes: along the Mediterranean south of the Alps and via the Danube corridor north of the Alps. Along the Mediterranean, the spread of agriculture is associated with Cardial Ware pottery, which reached the Iberian peninsula by ~7.5 kya (38, 39). North of the Alps, the Linearbandkeramik (LBK) culture spread rapidly between 6–7.5 kya from the Transdanubian region through central and western Europe, stopping before reaching the coast (38, 39). About 1,000 y later, the Neolithic reaches the British Isles (possibly via different entry points (42)) and southern Scandinavia (38, 39, 43), where it is associated with distinctive Funnel Beaker ceramics (TRB), while the Neolithic in Britain has both LBK and TRB elements. At around 4.8 kya, the TRB culture mostly disappears from the archaeological record of southern and western Sweden, and there is evidence of agriculture notably declining, being replaced by the more marine hunting economy of the Pitted Ware Culture (43). While there are traces of small-scale agriculture, fully agricultural societies do not turn up again until ~1,000 y later and do not completely occupy Finland until ~1 kya (38, 39). In the Baltic region, Ukraine, and eastern European plains, the initial spread of the Neolithic is associated with pottery rather than farming (44); farming shows up in these areas ~5–7 kya, with proposed sources

from southeastern Europe, Anatolia, and/or the Pontic-Caspian region.

There is a plethora of ancient DNA data from Europe (45), and these firmly support an appearance of Anatolian-related ancestry in Europe that coincides with the spread of farming technology, confirming that farming spread primarily via demic rather than cultural diffusion. While the Aegean region, encompassing northwestern Anatolia and northern Greece, is the source of farmer ancestry for most European populations (46), southern Greece shows a greater contribution from the Caucasus, which is also evident in the Greek Bronze age Minoan and Mycenaean cultures (47). Initial studies found little or no hunter-gatherer ancestry in the earliest farmers in the region, but recently, demographic modeling of ancient genomes revealed that Anatolian farmers entering Europe already harbored ancestry from European hunter-gatherers via earlier gene flow (48). The westward dispersal of farmers into Europe was accompanied by further gradual admixture over time with local populations, increasing the amount of hunter-gatherer ancestry in subsequent generations (45).

However, while the signal of admixture is pervasive, it varies regionally and through time. Some archeological sites show evidence of hunter-gatherer enclaves which coexisted with farmers without interaction for hundreds of years (e.g., “Blätterhöhle” in Germany (49), Sweden (50), and Poland (51)), while others show evidence of admixture even in the earliest farmer communities (52). In the British Isles, resident populations were completely replaced by populations with mixed farmer and hunter-gatherer ancestry, related to the Iberian Neolithic (53–55). The early Neolithic of Iberia is characterized by a greater genetic contribution of hunter-gatherers than elsewhere (56) that gradually increased over time (57–62). In Scandinavia, Neolithic farmers show appreciable genetic contributions from hunter-gatherers, but there are only low levels of farming-related ancestry in hunter-gatherers, suggesting that hunter-gatherers were mostly incorporated into expanding farming groups (63). And in the Eastern Baltic sea region, Ukraine, and Western Russia, the Neolithic progressed without substantial gene flow from Central European farmers until ~5 kya; here, hunter-gatherers persisted longer than elsewhere (63–66). Thus, there is substantial regional variation in the interactions between incoming farmers and local hunter-gatherers in Europe.

Bronze Age Migrations from the Steppes. Prior to the advent of ancient DNA studies, numerous studies based on modern populations tried to estimate the relative contributions of farmer vs. hunter-gatherer ancestry in European populations; using various datasets, approaches, and proxies for the farmer vs. hunter-gatherer ancestry, estimates of farmer ancestry ranged from less than 15% to more than 70% (67). Ancient DNA studies not only resolved this debate but additionally showed that a third source of ancestry, whose existence was generally not previously suspected, is present at frequencies of 10–50% in modern European populations (57, 68, 69). This ancestry, maximized in Yamnaya pastoralists of the Pontic-Caspian steppe, first appeared in Europe in the Baltic region (63, 65) ~5 kya, and took ~1 ky to spread to western Europe (54, 55). This massive migration may be associated with the expansion of the Corded Ware Complex (CWC) in central and eastern Europe; dating to ~4.4–4.9 kya, CWC sites have up to 75% of Yamnaya-like ancestry (57, 69). However, recent work has shown that following

its first appearance in the eastern Baltic, the massive expansion of steppe ancestry across Europe was mediated by a population with a substantial proportion of farmer ancestry; this ancestry was traced to the Late Neolithic Globular Amphora Culture, which preceded CWC in its central area (66).

The Bell Beaker culture appears from ~4.8 kya (about the same time as CWC) in the Iberian Peninsula and later all across western Europe, extending eastward into Poland (and thus overlapping CWC Sites) and including Sicily, Sardinia, and North Africa. Individuals associated with Bell Beaker sites have remarkable regional variation in ancestry (55, 57, 69): in central Europe, Bell Beakers trace 50% of their ancestry to the steppe; the Bell Beakers of the British Isles are of largely central European ancestry, which replaced the existing Neolithic ancestry by up to 90%; and the early Bell Beakers of Iberia lack the steppe ancestry almost completely, with later samples showing only a modest amount of steppe-related ancestry (56, 59). These results indicate that the spread of the Bell Beaker culture was not mediated by a single migrating population but was also adopted by local groups by cultural diffusion. Moreover, the impact of the steppe migration was substantially smaller in southern Europe, i.e., the Balkans (47) and Mycenaean Greece (70).

On the surface, the Neolithic Anatolian-related and Bronze Age steppe-related migrations to and across Europe seem to fit the classic model of a demic expansion: both had a massive impact on the ancestry of Europe and thus involved a substantial number of people, and both occurred over relatively short time spans (~1 ky for the steppe-related migrations and ~3 ky for the Neolithic spread). However, both show regional patterns of variation in the timing and amount of ancestry contributed, as well as additional complications (such as the resurgence of hunter-gatherer ancestry after the initial spread of farmers, and the potential spread of steppe-related ancestry together with farmer-related ancestry), indicating that there is more to these migrations than simple models of demic diffusion would imply.

IE Languages. The origin and spread of IE languages across Europe is of considerable interest, and two main hypotheses have been proposed: the Anatolian hypothesis, according to which IE languages originated in Anatolia and then spread to Europe along with agriculture, beginning ~8–9.5 kya (71) and the steppe hypothesis, according to which IE languages originated in the steppes north of the Black and Caspian Seas and then spread to Europe as a consequence of the domestication of horses and the development of wheeled carts and wagons, beginning ~5.5–6.5 kya (72, 73).

Ancient DNA analyses of the Bronze Age in Eurasia seem to support the steppe hypothesis (57, 69, 74), but several issues remain. The highest proportion of steppe ancestry is found in northeast Europe, in populations that speak Uralic languages, while many IE-speaking regions of southern Europe have substantially less steppe ancestry (47, 55, 70) possibly reflecting dilution of steppe ancestry via subsequent migrations. Furthermore, ancient DNA data from domestic horses indicate that the expansion of steppe ancestry into Europe was not driven by horses (75), although there is evidence of horse milking by the Yamnaya (76). IE languages therefore appear to have a more complex history than can be explained by a simple model; perhaps some IE languages

were spread by farmers and others by steppe pastoralists, or perhaps some IE languages spread by demic and others by cultural diffusion.

Central and South Asia

Neolithic and the Spread of Farming. Despite the rich archaeobotanical remains which suggest that Iran was a key area for the development of agriculture (77), the spread of farming eastward into Central Asia and southward into South Asia has been explored far less than the spread of farming into Europe. Ancient DNA studies indicate that Neolithic Iran is genetically divergent from Neolithic Anatolia, but beginning ~6 kya, a substantial proportion of Anatolian farmer-related ancestry appears in Iran, and there is a genetic cline extending to Central Asia of decreasing Anatolian-related ancestry, suggesting an eastward migration of Anatolian farmers to the Iranian plateau and Central Asia (78–80). This migration coincides with the dispersal of domesticated goats (81) but cannot explain the presence of domesticated sheep in Kyrgyzstan ~8 kya, suggesting multiple and/or earlier expansions (82, 83).

Farming in South Asia first appears at the Neolithic site of Mehrgarh, in modern Pakistan to the west of the Indus valley, dating to ~8.5 kya (84). Other Neolithic settlements in Pakistan show that by 6–7 kya, farmers had started moving to the north and east, and farming-based permanent settlements characteristic of the Indus Valley Civilization (IVC), which flourished 3.9–4.6 kya, begin to appear. The discovery of thousands of sites across India documents diffusion of farming beyond the archaeological boundaries of the IVC (85).

Because it is so difficult to retrieve DNA from samples coming from hot and humid environments, like those found in South Asia, to date, ancient DNA is limited to a single IVC genome (86). This individual is indirectly dated to ~4.3–4.8 kya and shares ancestry with modern South Asians and some pre-Neolithic individuals from Iran and Central Asia but lacks Anatolian farmer-related ancestry (80). In contrast to Europe, farming apparently did not spread to South Asia via the migration of people from Anatolia, although one should be cautious about placing too much emphasis on the results from a single individual.

Steppe Migrations into Central and South Asia. As in Europe, steppe migrations into Central and South Asia show a complex history. The earliest eastward expansion from the steppe is the appearance in the Altai-Sayan region of the Afanasievo culture, dating to ~4.5–5.3 kya; Afanasievo people are genetically close to the Yamnaya (69). In Central Asia, the first evidence of steppe ancestry comes from the Bactria-Margiana Archaeological Complex beginning ~4 kya, and the spreading steppe ancestry seems to be associated with the Trans-Ural Sintashta archaeological complex (69, 80, 87). Thus, the ancient DNA data suggest (at least) two distinct eastward migrations. Furthermore, ~3.5–4 kya, a migration from Central Asia brought steppe ancestry to South Asia, mixing with the IVC people (86) and contributing up to 30% of the ancestral composition of modern South Asians (80). And, unlike in Europe, ancient DNA from domesticated horses implicates them in the eastward expansions from the steppes (75) and potentially the spread of Indo-Iranian languages (a major branch of the IE family, found in parts of Central and South Asia).

East Asia and Mainland Southeast Asia

East Asia has a long history of human occupation, dating back to at least 45 kya (4). During the Holocene, various regions around the Yellow, Huai, and Yangtze Rivers in China were important centers for the domestication of rice, broomcorn millet, and foxtail millet, beginning ~8–9 kya (88). Archaeological evidence indicates two main streams of agriculture southward, one into Mainland Southeast Asia (MSEA) and the other moving into Taiwan and ultimately continuing as the expansion of the Austronesian language family. In MSEA, the linguistic situation is more complicated, as there are five main language families that have spread and diversified throughout the region, and the challenge is to understand the forces and processes that influenced their spread.

Spread of Agriculture. Archaeologists have debated the extent to which agriculture in East Asia spread by the movement of farmers who admixed with and/or assimilated local hunter-gatherers (the “two-layer” hypothesis) vs. spreading via cultural diffusion and/or indigenous developments (the “regional continuity” hypothesis), with a growing consensus favoring the former (89, 90). Ancient DNA strongly supports the two-layer hypothesis: Hòabinhian hunter-gatherers from Laos and Malaysia, dated to ~7.8–8.0 kya and ~4.2–4.4 kya, respectively, are most closely related to modern indigenous groups from South and Southeast Asia, which could reflect the initial colonization, while Neolithic individuals from MSEA dating to ~4 kya can be modeled as a mixture of ancestry from Hòabinhian hunter-gatherers and early farmers from China (91, 92). These Neolithic individuals share ancestry with modern Austroasiatic-speaking groups from MSEA (91, 92), suggesting that farming may be linked to the spread of Austroasiatic languages. Ancient DNA analyses also indicate a westward migration of farmers from the region of the Yellow River, beginning ~6 kya, that contributed ancestry to both Tibetans and Han Chinese, and hence may be associated with the spread of Sino-Tibetan languages (93).

A provocative hypothesis argues that Korean, Japanese, Tungusic, Mongolic, and Turkic languages all have a common origin (the “Transeurasian” macrofamily) and spread via farmers migrating from northeast China during the Early Neolithic (94). However, other studies question the existence of this macrofamily (95, 96). In any case, overall, there is a strong signal of demic expansions associated with the spread of agriculture in East Asia.

Subsequent Dispersals into MSEA. Ancient DNA from ~2 kya from MSEA individuals associated with Bronze Age cultures, and continuing into the Iron Age and historical times, shows additional East Asian-related ancestry that is not present in Neolithic individuals (91, 92, 96). Much of the structure of present-day MSEA populations was formed as a result of these inferred migrations, as ancient individuals from this time start to more closely resemble genetically modern MSEA populations from the same areas.

MSEA today is characterized by extensive linguistic diversity with five major language families represented (Austroasiatic, Tai-Kadai, Sino-Tibetan, Hmong-Mien, and Austronesian). Austroasiatic languages, as discussed above, are associated with the Neolithic dispersal of rice and millet farming, and their scattered distribution today probably reflects later incursions of people speaking other languages. Linguistic links between Tai-Kadai and Austronesian languages have been suggested (97); genetic studies based on both ancient samples

and modern populations confirm a likely ancestral link between proto-Tai-Kadai and proto-Austronesian groups (96, 98, 99). Sino-Tibetan languages have their origins in northern China and probably spread to MSEA beginning ~3 kya, while Hmong-Mien languages probably arose in southern China and spread around the same time as Tai-Kadai (100). There are just a handful of Austronesian languages in MSEA (Malayic, Moken, and Chamic languages), and they probably originate from a migration from Borneo ~2–2.5 kya (101) that contributed little genetic ancestry to Austronesian-speaking MSEA groups (98).

Thus, compared to other regions of the world where single-language families have expanded over a wide geographic region and become dominant (e.g., Bantu, IE, Austronesian, etc.), MSEA is very different in that multiple-language families all originated and spread within a period of a few thousand years. While it has been suggested that food production was an important aspect (101), it seems that something more must have allowed these different language families to proliferate more or less simultaneously. Indeed, comprehensive studies of genome-wide data from modern populations of Vietnam (98) and Thailand (99) document a complex history involving expansions, extensive contact between groups speaking languages from different families, isolation, and likely cases of language shift. More detailed studies of ancient DNA from the region should shed further light on this complex history.

Island Southeast Asia (ISEA) and Oceania

Population movements in this region during the Holocene are largely driven by an agricultural expansion from Taiwan southward and eastward into ISEA and along the northern coast of New Guinea; the development of sophisticated sea navigation technology for long-distance voyaging into Remote Oceania; and the rise of maritime trade networks resulting in the emergence of sea nomadic groups and the settlement of Madagascar. Also, there was an early, independent domestication of plants in the New Guinea highlands, associated with the spread of Trans-New-Guinea languages.

Austronesian Expansion out of Taiwan. The Holocene settlement of ISEA is associated with the spread of peoples who brought agriculture and Austronesian languages from Taiwan (102–104). The Austronesian language family is one of the largest and most widespread in the world, with over 1,000 languages covering almost half the globe, including Taiwan, Southeast Asia, Near and Remote Oceania, and Madagascar (105). Austronesian languages clearly spread from Taiwan, as multiple divergent branches are found only in Taiwan, while all non-Taiwanese Austronesian languages belong to a single branch, Malayo-Polynesian (105). There is a strong signal in the languages (104), archaeology (102), and genetics (106–108) of people dispersing from Taiwan ~4–5 kya southward through the Philippines and into Indonesia, continuing westward to MSEA and eastward to New Guinea and nearby islands. This demic diffusion did not completely replace the indigenous groups, as attested by admixed modern and ancient genomes (107, 109, 110). Here is a classic signal of a farming-language dispersal via demic diffusion—but the story is more complicated; ancient DNA indicates the presence of MSEA-related ancestry in Wallacea (the islands of Eastern Indonesia) that

likely predates the Austronesian expansion (109) but for which there is no archaeological or linguistic evidence.

Long-Distance Sea Voyages into Oceania. Austronesian movement through ISEA and into Near Oceania (New Guinea and nearby islands) was likely achieved by island-hopping and intervisible water crossings, which would not have required sophisticated boating technology (although ancient DNA evidence from the Mariana Islands supports a direct migration ~3.5 kya from the Philippines across more than 2,000 km of open ocean (111, 112). Austronesians probably continued from eastern Indonesia along the New Guinea coast, eventually reaching the Bismarck Archipelago ~3.4 kya (113). Austronesian-associated ancestry in New Guinea is limited to the coast and offshore islands, with no evidence of Austronesians penetrating the New Guinea highlands (114, 115).

The further spread of Austronesians into Remote Oceania (islands to the north and east of New Guinea, including Micronesia and Polynesia, which could be reached only by long-distance voyaging) definitely required sophisticated boating technology. Austronesians moved rapidly through the islands of Remote Oceania, reaching Tonga and Samoa ~2.9 kya and the farthest islands (Hawaii and New Zealand) within the past 1 ky (116). Initial studies of modern populations indicated that Polynesians had ~80% Austronesian-related and 20% Papuan-related ancestry (117); moreover, this admixture was highly sex biased, with predominantly Austronesian maternal ancestry but mostly Papuan paternal ancestry (115). It seemed likely that Austronesians paused upon reaching New Guinea (most likely, the Bismarcks) and mixed extensively with Papuans. However, an initial ancient DNA study surprisingly found that individuals from Vanuatu and Tonga, dating to ~2.5–2.9 kya, possessed little or no Papuan-related ancestry (118). Subsequent studies showed that Papuan-related ancestry spread later, via mostly male-mediated, continuous migration (97, 120, 121). Other studies suggested back migrations from Polynesia (121) and native American ancestry that arrived in Polynesia before European contact (122), although the latter relies on analyses of modern samples and so far has not received any support from ancient DNA (123). Overall, instead of the stereotypical view that settlement of Remote Oceania involved one-time, “sweepstakes”-like successes amidst a sea of failures, genetic data confirm archaeological evidence for multiple dispersals and large-scale trade networks throughout Remote Oceania, including repeated contact with Near Oceania (124).

Madagascar and Sea Nomads. Austronesians not only settled vast expanses of the Pacific but they were also the first people to reach Madagascar, ~1.1–1.3 kya (125). Moreover, the Malagasy language groups with Greater Barito languages from southeastern Borneo (126), and genome-wide studies also point to a likely origin of proto-Malagasy speakers from southeast Borneo, who admixed with Bantu speakers arriving later from continental Africa (127, 128).

Interestingly, Sama-Bajaw languages spoken by certain groups in the Philippines, Malaysia, and Indonesia likewise belong to the Greater Barito group (129). These include populations described as sea nomads due to their boat-dwelling history and an intense maritime orientation. The distribution of Sama-Bajaw and the Malagasy languages may reflect eastward and westward movements, respectively, of peoples

from the Barito region, triggered by the expansion of the Srivijaya empire, a Malay-Indian thalassocracy, ~1 kya (129). Contrastingly, genome-wide studies of the Indonesian sea nomads indicate a Sulawesi origin (130), suggesting that there may be different origins for the language vs. genetic ancestry due to a complex history of interactions and cultural shifts between neighboring groups in the region. Further studies are needed on sea nomads and potential connections with the Malagasy, but these appear to be long-distance dispersals unrelated to farming.

Spread of Agriculture in Highland New Guinea. The New Guinea highlands were an independent and early site of plant domestication, with definite evidence for taro and banana cultivation by ~6.7–7 kya and indications of taro cultivation as early as 10 kya (131). The spread of agriculture in New Guinea may have been accompanied by the spread of Trans-New Guinea (TNG) languages, comprising nearly half of the ~850 languages spoken on the island (132). A comprehensive study of genome-wide variation across Papua New Guinea (PNG) found evidence for the formation of highland population structure and expansions ~10 kya, which could reflect the impact of agriculture and the spread of TNG languages (115). However, PNG highlanders show much higher levels of population differentiation (115) than is typical for regions with strong evidence of farming-related dispersals (e.g., involving Austronesian-speaking or Bantu-speaking populations), indicating either that the putative agricultural-related expansion was followed by extreme isolation, bottlenecks, and drift or that there was no significant expansion in the highlands. Ancient DNA data, as well as additional comparisons of modern PNG populations (in particular, TNG vs. non-TNG groups), would be helpful in further elucidating the impact of agriculture on the genetic structure of New Guinea.

The Americas

The Americas, comprising North and South America, were the last continents to be colonized by modern humans, with archaeological and genetic evidence converging on an initial colonization time of ~16–18 kya (133). Current evidence points to a rapid latitudinal spread from Alaska to southern Chile, along the Pacific coast, in a few thousand years (133–135). However, archaeological evidence of extensive human occupation across the Americas only begins with the Holocene. Plants were domesticated at various locations in Mesoamerica, the Andes, and Amazonia (136–138), but it seems that major dispersal events were not associated exclusively with farming or other technological or behavioral innovations.

North America. Archaeological evidence supports the existence of two major Arctic-wide human dispersals, where farming obviously played no role. The first involved Paleo-Inuit people, occurring ~5 kya, associated with several archaeological cultures such as Pre-Dorset, Saqqaq, and Independence I (139). Genomic data from a Paleo-Inuit who lived ~4 kya shows that this individual stemmed from a population dispersal from Siberia, independent of previous dispersals, and not related to any present-day Native American population (135, 140). The Paleo-Inuit disappear from the archaeological record ~1.5 kya and were replaced by people from the Thule culture, the genetic and cultural ancestors of modern-day Inuit and Iñupiat (135, 141). The Thule people first appear in the archaeological record by ~1 kya along the shores of Alaska and rapidly reached Greenland, with the aid of dogsleds and umiaks

(large open-skin boats) (139). Genomic evidence from Thule individuals shows that they admixed extensively with other northern Native American groups (135).

Mesoamerica. Maize was a staple food for many societies at the time Europeans arrived in the Americas; archaeological evidence suggests that after its domestication in south-central Mexico at least 8.7 kya, it spread widely, reaching the southwestern USA by ~4.5 kya and coastal South America as early as ~7 kya (137, 142). However, there is disagreement as to whether maize agriculture spread northward from Mesoamerica primarily as a process of cultural diffusion, or if it was introduced to the southwestern USA through the long-distance migration of Mesoamerican farmers speaking Proto-Uto-Aztecan (PUA) languages (137). The disagreements revolve around the proposed homeland of PUA (2, 143) and whether the PUAs were northern foragers who adopted maize agriculture while dispersing to the south or instead were early southern maize farmers who expanded north. To complicate matters, some agricultural groups may have reverted to foraging as they spread into environments not suitable for farming (137).

Genetic studies have so far not been able to resolve these different views, as they are limited largely to modern populations due to the poor preservation of ancient DNA. Even with modern populations, there are significant sampling gaps, due to the understandable reluctance of some contemporary native American groups to participate in such studies. The largest study to date of genome-wide data from Mexican populations shows that the genetic structure of present-day groups has been influenced by the interplay of demographic, cultural, and geographic events (144). For instance, the divergence time between Aridoamerican (i.e., from northwest Mexico and the southwestern United States) and Mesoamerican populations was estimated between 4–9.9 kya, around the time when sedentary farming started in Mesoamerica. However, genome-wide diversity patterns and genetic structure reflect the influence of geography rather than linguistic affiliation (144). Additionally, contrasting patterns for mtDNA and Y chromosome variation across Aridoamerica and Mesoamerica suggest different demographic histories for females and males (145, 146). There is thus no strong signature of a genetic expansion associated either with the spread of maize agriculture or of UA languages.

South America. South America harbors a great deal of cultural and linguistic diversity, containing several language families that have dispersed over large geographic areas, namely Quechuan along the Andes and Arawakan, Tupian, and Cariban across the Amazonian lowlands (147, 148). However, these far-reaching dispersals did not lead to a complete replacement of small, local-language families, resulting in a mosaic-like linguistic landscape. The Andes and Amazonia are considered important centers of plant domestication that started ~8.5–10 kya (136, 138), but the expansion and diversification of the widespread South American language families are assumed to have occurred only ~1–4 kya (149, 150). Hence, early farming was not responsible for the large-scale dispersal of the major South American language families. It took several millennia until plant domesticates became a significant part of the human diet, perhaps because demographic expansions required further technological innovations and gains in crop productivity (151). Climatic changes might have also influenced such

dispersals; during the transition from the middle to the late Holocene (~4.2 kya), South America experienced increased rainfall, and tropical rainforest expanded at the expense of savannas (152–154).

Genetic studies, based solely on modern populations to date, are beginning to provide some insights into some language family dispersals. For example, Arawakan is the most widespread language family in the Americas; at the time of European arrival, Arawakan languages were present in Central America, from the Caribbean islands south to northern Argentina and from the Andean foothills to eastern South America (155). Arawakan societies are traditionally riverine horticulturalists, associated with the Saladoid-Barrancoid ceramic tradition, and played a central role in the exchange networks that connected vast areas of Amazonia, the Andes, and the Caribbean. However, a study of uniparental markers indicates that Arawakan groups in Northwestern Amazonia tend to be more closely related to neighboring non-Arawakans than to more distant Arawakan groups (156), indicating a discordance between genetic and linguistic relationships.

Similarly, the Tupi family is the most diverse in South America linguistically and is nearly as widespread geographically as the Arawakan family (148). The proposed homeland is in southwestern Amazonia, from where it expanded to the east and north (157). However, as with Arawakan, Tupi groups are genetically closer to neighboring non-Tupis than to more distant Tupi groups (158). The same holds true for Quechuan speakers from the highlands and lowlands of the Andes (159), although there are also some genetic connections between Quechuan groups from these two regions.

Overall, it seems that the spread of Arawakan and Tupi languages in the Amazon, and the spread of Quechua languages from the Andean highlands toward the near eastern lowlands, was either via cultural diffusion or extensive recent admixture between dispersing people and their neighboring groups that has diminished the genetic signal of the original demic expansion. Indeed, there has been extensive admixture between groups speaking different languages, likely reflecting the strongly patrilocal social structure of most South American groups (156, 159, 160). The present-day picture is further complicated by the impact of European colonization; ancient DNA studies would be very informative for disentangling signals of expansion and admixture.

Conclusions

This brief survey highlights the complexities of Holocene human dispersals. Although strong genetic signals of expansion can be identified that link the spread of farming and some language families (e.g., Bantu, Austronesian), even here there is significant heterogeneity in the genetic outcome of the interactions between expanding farmers and indigenous hunter-gatherers (e.g., complete or near-complete replacement of the pre-farming groups in Malawi and Mozambique by expanding Bantu groups vs. extensive admixture between Bantu and Khoisan-speaking groups in southern Africa). In other areas of the world, the links between expansions and agriculture are more tenuous (e.g., the Americas), possibly because of extensive postexpansion admixture or other complications. Ancient DNA investigations have greatly facilitated the identification of some of these complicating factors in some parts of the world

as well as expansions that were not detected in surveys of modern populations, most notably the impact of Bronze Age migrations from the steppes that may have brought IE languages to Europe and/or Central/South Asia. MSEA is a particularly complex region, with no fewer than five different language families having spread into the region during a time span of a few thousand years; what were the circumstances that allowed these diverse language families to survive and proliferate? Although ancient DNA analyses in many parts of the world are hampered by issues related to DNA survival, we can hope that further technological advances will bring novel insights into Holocene dispersals. In the meantime, more

comprehensive studies of modern populations—and further developments in computational analyses of genomic data—would be beneficial. In addition, there is a clear need for investigations into the sociocultural circumstances that underlie the rich complexity of the history of Holocene dispersals, in order to understand why there were such variable outcomes in the expansions of farming and languages.

Data, Materials, and Software Availability. There are no data underlying this work.

ACKNOWLEDGMENTS. We thank Lara Rubio Arauna, Chiara Barbieri, Hsiao-Chun Hung, Nina Marchi, and two anonymous reviewers for their helpful comments.

1. P. Bellwood, C. Renfrew, Eds., *Examining the Farming/Language Dispersal Hypothesis* (McDonald Institute for Archaeological Research, Cambridge, 2002).
2. J. Diamond, P. Bellwood, Farmers and their languages: The first expansions. *Science* **300**, 597–603 (2003).
3. I. Mathieson, A. Scally, What is ancestry? *PLoS Genet.* **16**, e1008624 (2020).
4. M. A. Yang, "A genetic history of migration, diversification, and admixture in Asia" in *Human Population Genetics and Genomics* (Elsevier Science, 2022), pp. 1–32, 10.47248/hpgg2202010001.
5. A. J. Ammerman, L. L. Cavalli-Sforza, *The Neolithic Transition and the Genetics of Populations in Europe* (Princeton University Press, Princeton, New Jersey, 1984).
6. A. Choudhury, D. Sengupta, M. Ramsay, C. Schlebusch, Bantu-speaker migration and admixture in southern Africa. *Hum. Mol. Genet.* **30**, R56–R63 (2021).
7. N. Hollfelder, G. Breton, P. Sjödin, M. Jakobsson, The deep population history in Africa. *Hum. Mol. Genet.* **30**, R2–R10 (2021).
8. B. Pakendorf, M. Stoneking, The genomic prehistory of peoples speaking Khoisan languages. *Hum. Mol. Genet.* **30**, R49–R55 (2021).
9. D. Gifford-Gonzalez, "Pastoralism in sub-Saharan Africa: Emergence and ramifications" in *The Oxford Handbook of Zooarchaeology*, U. Albarella, M. Rizzetto, H. Russ, K. Vickers, S. Viner-Daniels, Eds. (Oxford University Press, Oxford, 2017). 10.1093/oxfordhb/9780199686476.013.27.
10. P. J. Lane, "Early agriculture in sub-Saharan Africa to c. 500 ce" in *The Cambridge World History* (Cambridge University Press, Cambridge, 2015), pp. 736–773.
11. M. E. Prendergast *et al.*, Ancient DNA reveals a multistep spread of the first herders into sub-Saharan Africa. *Science* **365**, eaaw6275 (2019). 10.1126/science.aaw6275.
12. F. Lander, T. Russell, The archaeological evidence for the appearance of pastoralism and farming in southern Africa. *PLoS One* **13**, e0198941 (2018).
13. C. M. Schlebusch, M. Jakobsson, Tales of human migration, admixture, and selection in Africa. *Annu. Rev. Genomics Hum. Genet.* **19**, 405–428 (2018).
14. M. D. Purugganan, D. Q. Fuller, The nature of selection during plant domestication. *Nature* **457**, 843–848 (2009).
15. M. van de Loosdrecht *et al.*, Pleistocene north african genomes link near eastern and sub-saharan african human populations. *Science* **360**, 548–552 (2018).
16. R. Fregel *et al.*, Ancient genomes from North Africa evidence prehistoric migrations to the Maghreb from both the Levant and Europe. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6774–6779 (2018).
17. M. Gallego Llorente *et al.*, Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent. *Science* **350**, 820–822 (2015).
18. J. K. Pickrell *et al.*, Ancient west eurasian ancestry in southern and eastern Africa. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2632–2637 (2014).
19. R. Grollemund *et al.*, Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13296–13301 (2015).
20. D. Q. Fuller, E. Hildebrand, "Domesticating plants in Africa" in *The Oxford Handbook of African Archaeology*, P. Mitchell, P. J. Lane, Eds. (Oxford University Press, Oxford, 2013). 10.1093/oxfordhb/9780199569885.013.003.
21. K. Bostoen, "The Bantu expansion" in *Oxford Research Encyclopedia of African History*, T. Spear, Ed. (Oxford University Press, Oxford, 2018). 10.1093/acrefore/9780190277734.013.191.
22. C. de Filippo *et al.*, Y-chromosomal variation in sub-Saharan Africa: Insights into the history of Niger-Congo groups. *Mol. Biol. Evol.* **28**, 1255–1269 (2011).
23. J. Rocha, A. M. Fehn, "Genetics and demographic history of the Bantu" in *eLS* (John Wiley & Sons Ltd, Chichester, 2016). 10.1002/9780470015902.a0022892.
24. S. A. Tishkoff *et al.*, The genetic structure and history of Africans and African Americans. *Science* **324**, 1035–1044 (2009).
25. B. Pakendorf, C. de Filippo, K. Bostoen, Molecular perspectives on the bantu expansion: A synthesis. *Lang. Dyn. Change* **1**, 50–88 (2011).
26. G. B. Busby *et al.*, Admixture into and within sub-Saharan Africa. *Elife* **5**, e15266 (2016). 10.7554/eLife.15266.
27. C. de Filippo, K. Bostoen, M. Stoneking, B. Pakendorf, Bringing together linguistic and genetic evidence to test the Bantu expansion. *Proc. Biol. Sci.* **279**, 3256–3263 (2012).
28. E. Patin *et al.*, Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. *Science* **356**, 543–546 (2017).
29. A. Semo *et al.*, Along the indian ocean coast: Genomic variation in mozambique provides new insights into the bantu expansion. *Mol. Biol. Evol.* **37**, 406–416 (2020).
30. P. Skoglund *et al.*, Reconstructing prehistoric african population structure. *Cell* **171**, 59–71.e21 (2017).
31. V. Bajic *et al.*, Genetic structure and sex-biased gene flow in the history of southern African populations. *Am. J. Phys. Anthropol.* **167**, 656–671 (2018).
32. D. Seidensticker *et al.*, Population collapse in Congo rainforest from 400 CE urges reassessment of the Bantu expansion. *Sci. Adv.* **7**, 8352 (2021). 10.1126/sciadv.abd8352.
33. E. Patin *et al.*, The impact of agricultural emergence on the genetic history of African rainforest hunter-gatherers and agriculturalists. *Nat. Commun.* **5**, 3163 (2014).
34. S. Beleza, L. Gusmao, A. Amorim, A. Carracedo, A. Salas, The genetic legacy of western Bantu migrations. *Hum. Genet.* **117**, 366–375 (2005).
35. T. Güldemann, A linguist's view: Khoe-Kwadi speakers as the earliest food-producers of southern Africa. *South. Afr. Humanit.* **20**, 93–132 (2008).
36. M. Vicente, C. M. Schlebusch, African population history: An ancient DNA perspective. *Curr. Opin. Genet. Dev.* **62**, 8–15 (2020).
37. J. J. Hublin *et al.*, Initial upper palaeolithic homo sapiens from bacho kiro cave, bulgaria. *Nature* **581**, 299–302 (2020).
38. T. D. Price, Ed., *Europe's First Farmers* (Cambridge University Press, Cambridge, 2000).
39. C. Renfrew, P. Bahn, Eds., *The Cambridge World Prehistor* (Cambridge University Press, Cambridge, 2014), vol **3**: **West and Central Asia and Europe**.
40. N. Isern, J. Fort, M. Vander Linden, Space competition and time delays in human range expansions. Application to the neolithic transition. *PLoS One* **7**, e11106 (2012).
41. R. Pinhasi, J. Fort, A. J. Ammerman, Tracing the origin and spread of agriculture in Europe. *PLoS Biol* **3**, e410 (2005).
42. A. Sheridan, "The neolithization of Britain and Ireland: The 'Big Picture'" in *Landscapes in Transition*, B. Finlayson, G. Warren, Eds. (Oxbow Books, Oxford, 2010), pp. 89–105.
43. Å. M. Larsson, "Breaking and making bodies and pots: Material and ritual practices in Sweden in the Third Millenium BC", Ph.D. Thesis, Department of Archaeology and Ancient History, University of Uppsala, Uppsala, Sweden (2009), p. 455.
44. P. Dolukhanov *et al.*, The chronology of neolithic dispersal in central and eastern europe. *J. Archaeol. Sci.* **32**, 1441–1458 (2005).
45. I. Lazaridis, The evolutionary history of human populations in Europe. *Curr. Opin. Genet. Dev.* **53**, 21–27 (2018).
46. Z. Hofmanova *et al.*, Early farmers from across europe directly descended from neolithic aegeans. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6886–6891 (2016).
47. I. Mathieson *et al.*, The genomic history of southeastern Europe. *Nature* **555**, 197–203 (2018).
48. N. Marchi *et al.*, The genomic origins of the world's first farmers. *Cell* **185**, 1842–1859.e1818 (2022).
49. R. Bollongino *et al.*, 2000 years of parallel societies in Stone Age Central Europe. *Science* **342**, 479–481 (2013).
50. P. Skoglund *et al.*, Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science* **336**, 466–469 (2012).
51. D. M. Fernandes *et al.*, A genomic Neolithic time transect of hunter-farmer admixture in central Poland. *Sci. Rep.* **8**, 14879 (2018).
52. A. G. Nikitin *et al.*, Interactions between earliest linearbandkeramik farmers and central european hunter gatherers at the dawn of european neolithization. *Sci. Rep.* **9**, 19544 (2019).
53. L. M. Cassidy *et al.*, Neolithic and bronze age migration to ireland and establishment of the insular atlantic genome. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 368–373 (2016).
54. S. Brace *et al.*, Ancient genomes indicate population replacement in early neolithic britain. *Nat. Ecol. Evol.* **3**, 765–771 (2019).
55. I. Olalde *et al.*, The Beaker phenomenon and the genomic transformation of northwest Europe. *Nature* **555**, 190–196 (2018).
56. A. Szecsenyi-Nagy *et al.*, The maternal genetic make-up of the iberian peninsula between the neolithic and the early bronze age. *Sci. Rep.* **7**, 15644 (2017).
57. W. Haak *et al.*, Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature* **522**, 207–211 (2015).
58. M. Lipson *et al.*, Parallel palaeogenomic transects reveal complex genetic history of early European farmers. *Nature* **551**, 368–372 (2017).
59. R. Martiniano *et al.*, The population genomics of archaeological transition in west iberia: Investigation of ancient substructure using imputation and haplotype-based methods. *PLoS Genet.* **13**, e1006852 (2017).
60. I. Olalde *et al.*, The genomic history of the Iberian Peninsula over the past 8000 years. *Science* **363**, 1230–1234 (2019).
61. C. Valdiosera *et al.*, Four millennia of Iberian biomolecular prehistory illustrate the impact of prehistoric migrations at the far end of Eurasia. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 3428–3433 (2018).
62. V. Villalba-Mouco *et al.*, Survival of late pleistocene hunter-gatherer ancestry in the iberian peninsula. *Curr. Biol.* **29**, 1169–1177.e1167 (2019).
63. A. Mittnik *et al.*, The genetic prehistory of the Baltic Sea region. *Nat. Commun.* **9**, 442 (2018).

64. E. R. Jones *et al.*, The neolithic transition in the baltic was not driven by admixture with early european farmers. *Curr. Biol.* **27**, 576–582 (2017).
65. L. Saag *et al.*, Extensive farming in estonia started through a sex-biased migration from the steppe. *Curr. Biol.* **27**, 2185–2193.e2186 (2017).
66. M. E. Allentoft, Population genomics of stone age eurasia. *bioRxiv* [Preprint]. <https://doi.org/10.1101/2022.05.04.490594> (Accessed 5 May 2022).
67. M. Stoneking, *An Introduction to Molecular Anthropology* (John Wiley and Sons Inc, Hoboken, New Jersey, 2017).
68. I. Lazaridis *et al.*, Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
69. M. E. Allentoft *et al.*, Population genomics of Bronze Age Eurasia. *Nature* **522**, 167–172 (2015).
70. I. Lazaridis *et al.*, Genetic origins of the Minoans and Mycenaeans. *Nature* **548**, 214–218 (2017).
71. C. Renfrew, *Archaeology and Language: The Puzzle of Indo-European Origins* (Jonathan Cape, London, 1987).
72. D. W. Anthony, D. Ringe, The Indo-European homeland from linguistic and archaeological perspectives. *Annu. Rev. Linguist.* **1**, 199–219 (2015).
73. J. P. Mallory, *In Search of the Indo-Europeans: Language, Archaeology, and Myth* (Thames & Hudson, London, 1989).
74. I. Lazaridis *et al.*, The genetic history of the Southern Arc: A bridge between West Asia and Europe. *Science* **377**, eabm4247 (2022).
75. P. Librado *et al.*, The origins and spread of domestic horses from the Western Eurasian steppes. *Nature* **598**, 634–640 (2021).
76. S. Wilkin *et al.*, Dairying enabled early bronze age yamnaya steppe expansions. *Nature* **598**, 629–633 (2021).
77. S. Riehl, M. Zeidi, N. J. Conard, Emergence of agriculture in the foothills of the zagros mountains of Iran. *Science* **341**, 65–67 (2013).
78. F. Broushaki *et al.*, Early neolithic genomes from the eastern fertile crescent. *Science* **353**, 499–503 (2016).
79. M. Gallego-Llorente *et al.*, The genetics of an early Neolithic pastoralist from the Zagros, Iran. *Sci. Rep.* **6**, 31326 (2016).
80. V. M. Narasimhan *et al.*, The formation of human populations in South and Central Asia. *Science* **365**, eaat7487 (2019).
81. L. A. Frantz, D. G. Bradley, G. Larson, L. Orlando, Animal domestication in the era of ancient genomics. *Nat. Rev. Genet.* **21**, 449–460 (2020).
82. W. T. T. Taylor *et al.*, Evidence for early dispersal of domestic sheep into Central Asia. *Nat. Hum. Behav.* **5**, 1169–1179 (2021).
83. E. Yurtman *et al.*, Archaeogenetic analysis of Neolithic sheep from Anatolia suggests a complex demographic history since domestication. *Commun. Biol.* **4**, 1279 (2021).
84. E. Kingwell-Banham, C. A. Petrie, D. Q. Fuller, "Early agriculture in South Asia" in *The Cambridge World History*, G. Barker, C. Goucher, Eds. (Cambridge University Press, Cambridge, 2015), pp. 261–288.
85. V. N. Misra, Prehistoric human colonization of India. *J. Biosci.* **26**, 491–531 (2001).
86. V. Shinde *et al.*, An ancient harappan genome lacks ancestry from steppe pastoralists or iranian farmers. *Cell* **179**, 729–735.e710 (2019).
87. P. de Barros Damgaard *et al.*, The first horse herders and the impact of early Bronze Age steppe expansions into Asia. *Science* **360**, 6583 (2018).
88. C. J. Stevens, D. Q. Fuller, The spread of agriculture in eastern Asia: Archaeological bases for hypothetical farmer/language dispersals. *Lang. Dyn. Change* **7**, 152–186 (2017).
89. P. Bellwood, The search for ancient DNA heads east. *Science* **361**, 31–32 (2018).
90. C. F. W. Higham, "The neolithic occupation of Southeast Asia" in *The Languages and Linguistics of Mainland Southeast Asia*, P. Sidwell, M. Jenny, Eds. (Walter de Gruyter GmbH, Berlin, 2021), pp. 21–32.
91. M. Lipson *et al.*, Ancient genomes document multiple waves of migration in Southeast Asian prehistory. *Science* **361**, 92–95 (2018).
92. H. McColl *et al.*, The prehistoric peopling of Southeast Asia. *Science* **361**, 88–92 (2018).
93. M. A. Yang *et al.*, Ancient DNA indicates human population shifts and admixture in northern and southern China. *Science* **369**, 282–288 (2020).
94. M. Robbeets *et al.*, Triangulation supports agricultural spread of the Transeurasian languages. *Nature* **599**, 616–621 (2021).
95. Z. Tian, Triangulation fails when neither linguistic, genetic, nor archaeological data support the Transeurasian narrative. *bioRxiv* [Preprint]. <https://doi.org/10.1101/2022.06.09.495471> (Accessed 12 June 2022).
96. C. C. Wang *et al.*, Genomic insights into the formation of human populations in East Asia. *Nature* **591**, 413–419 (2021).
97. L. Sagart, The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* **43**, 411–444 (2004).
98. D. Liu *et al.*, Extensive ethnolinguistic diversity in vietnam reflects multiple sources of genetic diversity. *Mol. Biol. Evol.* **37**, 2503–2519 (2020).
99. W. Kutanan *et al.*, Reconstructing the human genetic history of mainland southeast asia: Insights from genome-wide data from thailand and laos. *Mol. Biol. Evol.* **38**, 3459–3477 (2021).
100. P. Sidwell, "Southeast Asian mainland: Linguistic history" in *The Global Prehistory of Human Migration*, P. Bellwood, Ed. (Wiley Blackwell, West Sussex, 2015), pp. 259–268.
101. P. Bellwood, "Homelands and dispersal histories of Mainland Southeast Asian language families: A multidisciplinary perspective" in *The Languages and Linguistics of Mainland Southeast Asia*, P. Sidwell, M. Jenny, Eds. (Walter de Gruyter GmbH, Berlin, 2021), pp. 33–44.
102. P. Bellwood, *Prehistory of the Indo-Malaysian Archipelago* (Australian National University Press, Canberra, 1997).
103. R. Blust, The prehistory of the Austronesian-speaking peoples: A view from language. *J. World Prehistory* **9**, 453–510 (1995).
104. R. D. Gray, A. J. Drummond, S. J. Greenhill, Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* **323**, 479–483 (2009).
105. R. Blust, *The Austronesian Languages* (The Australian National University, Canberra, 2013).
106. A. M. Ko *et al.*, Early Austronesians: Into and out of Taiwan. *Am. J. Hum. Genet.* **94**, 426–436 (2014).
107. M. Lipson *et al.*, Reconstructing austronesian population history in island southeast asia. *Nat. Commun.* **5**, 4689 (2014).
108. J. Choin *et al.*, Genomic insights into population history and biological adaptation in Oceania. *Nature* **592**, 583–589 (2021).
109. S. Oliveira *et al.*, Ancient genomes from the last three millennia support multiple human dispersals into Wallacea. *Nat. Ecol. Evol.* **6**, 1024–1034 (2022).
110. S. Xu *et al.*, Genetic dating indicates that the Asian-Papuan admixture through eastern indonesia corresponds to the austronesian expansion. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4574–4579 (2012).
111. Y. C. Liu *et al.*, Ancient DNA reveals five streams of migration into Micronesia and multilocality in early Pacific seafarers. *Science* **377**, 72–79 (2022).
112. I. Pugach *et al.*, Ancient DNA from Guam and the peopling of the Pacific. *Proc. Natl. Acad. Sci. U.S.A.* **118**, (2021).
113. P. V. Kirch, Peopling of the Pacific: A holistic anthropological perspective. *Annu. Rev. Anthropol.* **39**, 131–148 (2010).
114. A. Bergstrom *et al.*, A Neolithic expansion, but strong genetic structure, in the independent history of New Guinea. *Science* **357**, 1160–1163 (2017).
115. M. Kayser, The human genetic history of Oceania: Near and remote views of dispersal. *Curr. Biol.* **20**, R194–201 (2010).
116. J. M. Wilmshurst, T. L. Hunt, C. P. Lipo, A. J. Anderson, High-precision radiocarbon dating shows recent and rapid initial human colonization of East Polynesia. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 1815–1820 (2011).
117. A. Wollstein *et al.*, Demographic history of Oceania inferred from genome-wide data. *Curr. Biol.* **20**, 1983–1992 (2010).
118. P. Skoglund *et al.*, Genomic insights into the peopling of the Southwest Pacific. *Nature* **538**, 510–513 (2016).
119. C. Posth *et al.*, Language continuity despite population replacement in Remote Oceania. *Nat. Ecol. Evol.* **2**, 731–740 (2018).
120. M. Lipson *et al.*, Population turnover in remote oceania shortly after initial settlement. *Curr. Biol.* **28**, 1157–1165.e1157 (2018).
121. L. R. Arauna *et al.*, The genomic landscape of contemporary western remote Oceanians. *Curr. Biol.* **32**, 4565–4575.e6 (2022).
122. A. G. Ioannidis *et al.*, Paths and timings of the peopling of Polynesia inferred from genomic networks. *Nature* **597**, 522–526 (2021).
123. L. Fehren-Schmitz *et al.*, Genetic ancestry of Rapanui before and after European contact. *Curr. Biol.* **27**, 3209–3215.e3206 (2017).
124. B. Shaw *et al.*, Frontier Lapita interaction with resident Papuan populations set the stage for initial peopling of the Pacific. *Nat. Ecol. Evol.* **6**, 802–812 (2022).
125. A. Anderson *et al.*, New evidence of megafaunal bone damage indicates late colonization of Madagascar. *PLoS One* **13**, e0204368 (2018).
126. M. Serva, The settlement of Madagascar: What dialects and languages can tell us. *PLoS One* **7**, e30666 (2012).
127. N. Brucato *et al.*, Malagasy genetic ancestry comes from an historical malay trading post in Southeast Borneo. *Mol. Biol. Evol.* **33**, 2396–2400 (2016).
128. D. Pierron *et al.*, Genomic landscape of human diversity across Madagascar. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6498–E6506 (2017).
129. R. Blust, The linguistic position of Sama-Bajaw. *Stud. Philippine Lang. Cult.* **15**, 73–114 (2007).
130. P. Kusuma *et al.*, The last sea nomads of the Indonesian archipelago: Genomic origins and dispersal. *Eur. J. Hum. Genet.* **25**, 1004–1010 (2017).
131. T. Denham, Early agriculture and plant domestication in New Guinea and Island Southeast Asia. *Curr. Anthropol.* **52**, S379–S395 (2011).
132. A. Pawley "The chequered career of the Trans New Guinea hypothesis: Recent research and its implications" in *Papuan Past: Cultural, Linguistic, and Biological Histories of Papuan-speaking Peoples*, A. Pawley, R. Attenborough, J. Golson, R. Hide, Eds. (Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University, Canberra, 2005), pp. 67–107.
133. M. R. Waters, Late Pleistocene exploration and settlement of the Americas by modern humans. *Science* **365**, 6562 (2019).
134. T. D. Dillehay *et al.*, Monte Verde: Seaweed, food, medicine, and the peopling of South America. *Science* **320**, 784–786 (2008).
135. E. Willerslev, D. J. Meltzer, Peopling of the Americas as inferred from ancient genomics. *Nature* **594**, 356–364 (2021).
136. U. Lombardo *et al.*, Early Holocene crop cultivation and landscape modification in Amazonia. *Nature* **581**, 190–193 (2020).
137. W. L. Merrill *et al.*, The diffusion of maize to the southwestern United States and its impact. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21019–21026 (2009).
138. D. R. Piperno, The origins of plant cultivation and domestication in the new world tropics. *Curr. Anthropol.* **52**, S453–S470 (2011).
139. M. Friesen, O. Mason, M. Friesen, "Pan-Arctic population movements: The early Paleo-Inuit and Thule Inuit migrations" in *The Oxford Handbook of the Prehistoric Arctic*, M. Friesen, O. Mason, Eds. (Oxford University Press, Oxford, 2016), 10.1093/oxfordhb/9780199766956.013.40.
140. M. Rasmussen *et al.*, Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463**, 757–762 (2010).
141. M. Raghavan *et al.*, The genetic prehistory of the New World Arctic. *Science* **345**, 1255832 (2014).
142. L. Kistler *et al.*, Multiproxy evidence highlights a complex evolutionary legacy of maize in South America. *Science* **362**, 1309–1313 (2018).
143. L. Campbell, *American Indian Languages: The Historical Linguistics of Native America* (Oxford University Press, Oxford, 1997).
144. H. Garcia-Ortiz *et al.*, The genomic landscape of Mexican Indigenous populations brings insights into the peopling of the Americas. *Nat. Commun.* **12**, 5942 (2021).
145. B. M. Kemp *et al.*, Evaluating the farming/language dispersal hypothesis with genetic variation exhibited by populations in the Southwest and Mesoamerica. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6759–6764 (2010).
146. K. Sandoval *et al.*, Linguistic and maternal genetic diversity are not correlated in Native Mexicans. *Hum. Genet.* **126**, 521–531 (2009).
147. P. Epps, Language classification, language contact, and Amazonian prehistory. *Lang. Linguist. Compass* **3**, 581–606 (2009).

148. L. Michael, The classification of South American languages. *Ann. Rev. Linguist.* **7**, 329–349 (2021).
149. A. Hornborg, Ethnogenesis, regional integration, and ecology in prehistoric Amazonia: Toward a system perspective. *Curr. Anthropol.* **46**, 589–620 (2005).
150. F. S. Noelli, The Tupi: Explaining origin and expansions in terms of archaeology and of historical linguistics. *Antiquity* **72**, 648–663 (2015).
151. P. Heggarty, D. Beresford-Jones, "Farming-language dispersals: Principles" in *Encyclopedia of Global Archaeology*, C. Smith, Ed. (Springer, New York, NY, 2014), pp. 2739–2749, 10.1007/978-1-4419-0465-2_2415.
152. S. L. Cross, P. A. Baker, G. O. Seltzer, S. C. Fritz, R. B. Dunbar, A new estimate of the Holocene lowstand level of Lake Titicaca, central Andes, and implications for tropical palaeohydrology. *The Holocene* **10**, 21–32 (2000).
153. J. Gregorio de Souza, F. S. Noelli, M. Madella, Reassessing the role of climate change in the Tupi expansion (South America, 5000–500 BP). *J. R. Soc. Interface* **18**, 20210499 (2021).
154. R. Marchant, H. Hooghiemstra, Rapid environmental change in African and South American tropics around 4000 years before present: A review. *Earth-Science Reviews* **66**, 217–260 (2004).
155. A. Y. Aikenvald, "The Arawak language family" in *The Amazonian Languages*, R. W. Dixon, A. Y. Aikenvald, Eds. (Cambridge University Press, Cambridge, 1999), pp. 65–106.
156. L. Arias *et al.*, Cultural innovations influence patterns of genetic diversity in Northwestern Amazonia. *Mol. Biol. Evol.* **35**, 2719–2735 (2018).
157. F. S. Noelli, The Tupi: Explaining origin and expansions in terms of archaeology and of historical linguistics. *Antiquity* **72**, 648–663 (1998).
158. E. S. M. A. Castro *et al.*, Genomic insight into the origins and dispersal of the Brazilian coastal natives. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2372–2377 (2020).
159. C. Barbieri *et al.*, The current genomic landscape of Western South America: Andes, Amazonia, and Pacific Coast. *Mol. Biol. Evol.* **36**, 2698–2713 (2019).
160. L. Arias, C. Barbieri, G. Barreto, M. Stoneking, B. Pakendorf, High-resolution mitochondrial DNA analysis sheds light on human diversity, cultural interactions, and population mobility in Northwestern Amazonia. *Am. J. Phys. Anthropol.* **165**, 238–255 (2018).