Farmers and Their Languages: The First Expansions

Jared Diamond[1] and Peter Bellwood[2]

The largest movements and replacements of human populations since the end of the Ice Ages resulted from the geographically uneven rise of food production around the world. The first farming societies thereby gained great advantages over hunter-gatherer societies. But most of those resulting shifts of populations and languages are complex, controversial, or both. We discuss the main complications and specific examples involving 15 language families. Further progress will depend on interdisciplinary research that combines archaeology, crop and livestock studies, physical anthropology, genetics, and linguistics.

Until the end of the Pleistocene, all people on all continents lived as hunter-gatherers. Then, at different subsequent times between about 8500 and 2500 B.C., food production based on domestication of relatively few wild plant and animal species arose independently in at most nine homelands of agriculture and herding, scattered over all inhabited continents except Australia (Fig. 1) (*1*–*11*). Because food production conferred enormous advantages to farmers compared with hunter-gatherers living outside those homelands, it triggered outward dispersals of farming populations, bearing their languages and lifestyles (*12*–*14*). Those dispersals constitute collectively the most important process in Holocene human history.

The agricultural expansions ultimately resulted from three advantages that farmers gained over hunter-gatherers. First, because of far higher food yields per area of productive land, food production can support far higher population densities than can the hunter-gatherer lifestyle. Second, whereas most hunter-gatherer societies are mobile, most food-producing societies are sedentary and can thus accumulate stored food surpluses, which were a prerequisite for the development of complex technology, social stratification, centralized states, and professional armies. Third, epidemic infectious diseases of social domestic animals evolved into epidemic infectious diseases of crowded farming populations, such as smallpox and measles—diseases to which the farmers evolved or acquired some resistance, but to which unexposed hunter-gatherers had none. These advantages enabled early farmers to replace languages and societies of hunter-gatherers living in their main paths of expansion.

Whereas recently expanding Europeans described their conquests in writing, most of the major pre-Columbian expansions of agricultural populations occurred in preliterate times. Hence the evidence for them comes from five other independent sources: archaeology, records of plant and animal domestication, human skeletal remains, modern human genes (and sometimes ancient DNA), and dispersal histories of existing or extinct but attested languages. Thus, study of the agricultural expansions is preeminently interdisciplinary. To synthesize evidence from disparate fields is exciting but also challenging: Few scientists possess technical competence in all of these fields, and the different types of evidence may seem to yield conflicting conclusions.

This review begins by introducing the basic hypothesis and by explaining six complications sometimes raised as objections. We then discuss 2 general issues and 11 specific examples involving linked spreads of prehistoric farmers and language families outward from agricultural homelands, proceeding from relatively unequivocal examples to uncertain ones. Finally, we call attention to new types of evidence required to settle the many controversies in this field.

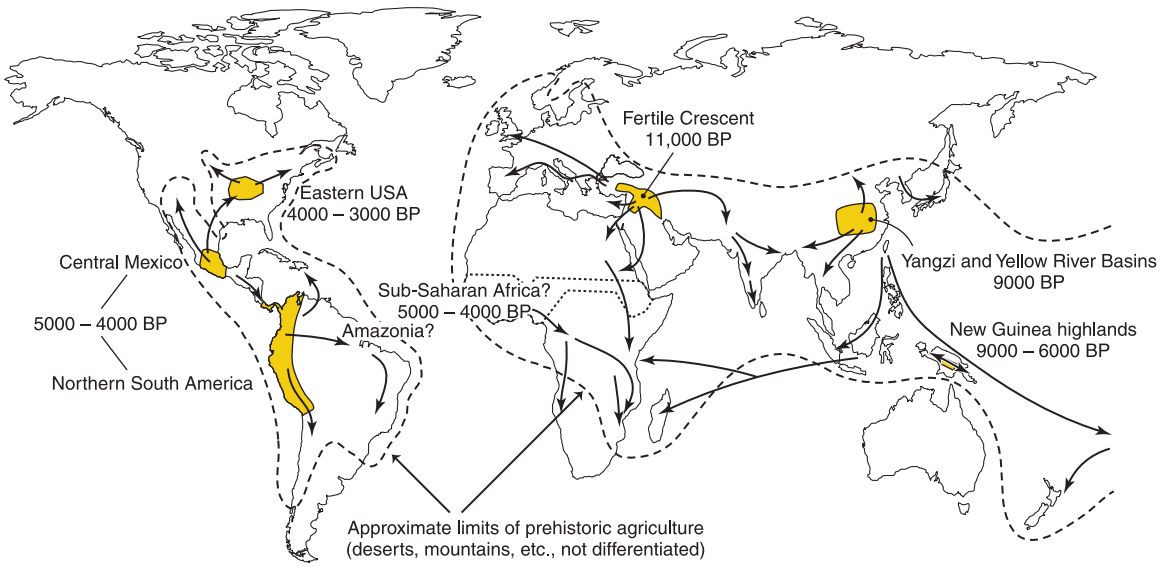

Fig. 1. Archaeological map of agricultural homelands and spreads of Neolithic/Formative cultures, with approximate radiocarbon dates.

[1]Department of Geography, University of California, Los Angeles, CA 90095-1524, USA. [2]School of Archaeology and Anthropology and Research School of Pacific Studies, Australian National University, Canberra ACT, 0200, Australia. E-mail: jdiamond@geog.ucla.edu (J.D.); peter.bellwood@anu.edu.au (P.B.)

The Basic Hypothesis and Six Complications

The simplest form of the basic hypothesis—that prehistoric agriculture dispersed hand-in-hand with human genes and languages—is that farmers and their culture replace neighboring hunter-gatherers and the latter's culture. This hypothesis would be supported if all five independent types of evidence coincided in attesting the replacement of local hunter-gatherers by expanding farmers bearing their own archaeologically visible culture, domesticates, skeletal types, genes, and languages, and if all those indicators were traceable back to the farmers' homeland of origin. Our two clearest examples of such concordance of evidence are the colonizations of previously uninhabited Polynesia and Micronesia by Neolithic populations speaking Austronesian languages (Fig. 2, no. 8), and the expansion of farmers speaking Bantu languages out of their tropical West African agricultural homeland after 1000 B.C. over most summer-rainfall regions of sub-equatorial Africa (Fig. 2, no. 1).

But the basic hypothesis is more often controversial, because in most other cases the five types of evidence are less concordant. Some critics believe that these discordances refute the hypothesis and that farming and language families spread mainly by diffusion amongst existing populations of hunter-gatherers (*15*). We conclude that reality is much richer and more complex than the simple version of the hypothesis, for many obvious reasons. The main classes of discordance are as follows:

Clinal genetic admixture between hunter-gatherers and farmers. Usually, arriving farmers do not exterminate or drive out hunter-gatherers completely. Instead, there is some intermarriage, especially of hunter-gatherer women to farmer men, resulting in dilution of farmer genes with hunter-gatherer genes. If the farmers' expansion consisted of hundreds of successive such steps of intermarriage and gene dilution, the hybrid population at the most remote step would have only low frequencies of the original farmers' genes, even though the hybrid population at each step might have consisted of only 10% local hunter-gatherers at that step and 90% invading hybrid farmers from the previous step. This is the wave-of-advance model by which Ammerman and Cavalli-Sforza [(*16*), see also (*17*)] interpreted southeast-northwest gene gradients across Europe, which they attributed to the northwestward expansion of farmers from Anatolia. As a result, genes of the modern population of western Ireland (the northwest terminus of the advance) are estimated to be derived 99% from Europe's original hunter-gatherers and only 1% from Anatolian farmers, even though Ireland's language, crops, livestock, religion, and writing system as of 1492 A.D. were derived almost entirely from eastern Mediterranean prototypes, with little or no contribution from the culture of Ireland's original Mesolithic hunter-gatherers (*18*).

Adoption of farming by peripheral hunter-gatherers. Some hunter-gatherer populations in the path of farming expansions succeeded

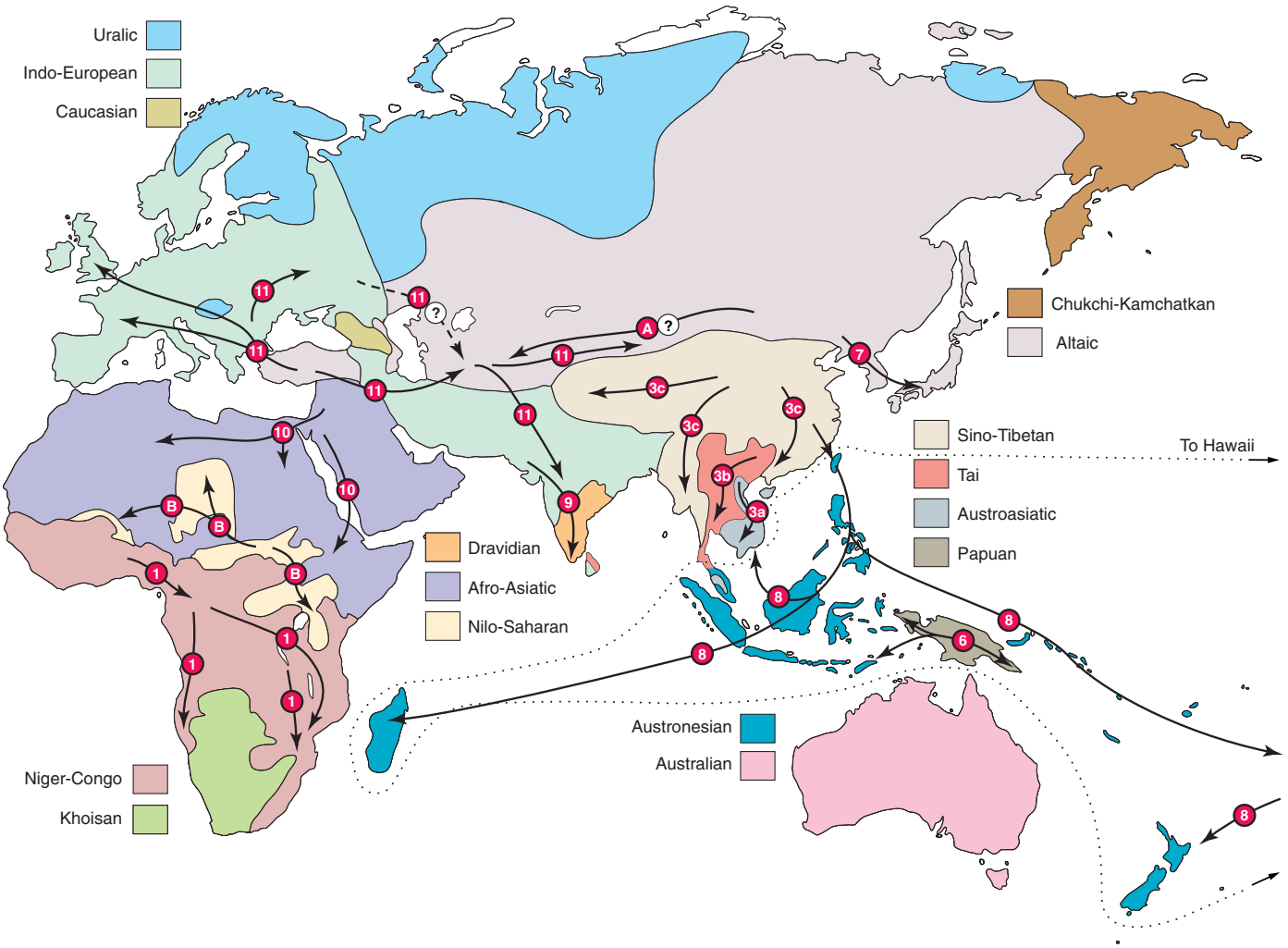

Fig. 2. Language families of the Old World and their suggested expansions. Map based on information in (*87*) and other sources. Numbered examples discussed in text are 1 (Bantu), 3a to 3c (Austro-Asiatic, Tai, and Sino-Tibetan, respectively), 6 (Trans New Guinea), 7 (Japanese), 8 (Austronesian), 9 (Dravidian), 10 (Afro-Asiatic), 11 (Indo-European). Other possible examples mentioned only briefly: A (Turkic), B (Nilo-Saharan).

in acquiring livestock and/or crops and some material culture (such as pottery) from food producers, while retaining their own languages, genes, and skeletons. The clearest example is the development of herding among some of southern Africa's indigenous Khoisan peoples (so-called Hottentots), who acquired livestock and pottery from food producers expanding southward in Africa about 2000 years ago (*19*, *20*). Other cases involve the Navajo in the southwestern United States adopting sheepherding between 1650 A.D. and 1700, and some Apache populations adopting casual maize farming (*21*, *22*).

Reversion of expanding farmers to the hunter-gatherer life-style. When expanding farmers reach areas unsuitable for farming with the domesticates available to them, they may survive by reverting to the hunter-gatherer life-style. Undoubted examples are the derivation of Polynesian hunter-gatherers on the Chatham Islands and New Zealand's South Island from ancestral Polynesian farmers (*23*–*25*), and of Punan hunter-gatherers in Borneo rainforests from other Austronesian farmers (*26*, *27*).

Language shift by indigenous populations. Discordance between languages and genes may arise when an expanding language is imposed on or adopted by a peripheral population, with only a minor contribution of expanding genes. This situation differs from the situation of clinal gene dilution, in which invaders constitute a majority at every step. A clear modern example is the increasing adoption of English as the language of government in Papua New Guinea, whose inhabitants nevertheless remain indigenous New Guineans with negligible admixture of European genes. Possible historical examples are the imposition of the Magyar and Turkish languages on medieval Hungarians and Anatolians, whose genes today are estimated to be derived only 10 and 30%, respectively, from their conquerors' genes (*28*). Likely prehistoric examples are the adoption of Austronesian languages by some former speakers of Papuan languages in the western islands of Melanesia and by Agta Negritos in the Philippines (*29*, *30*). However, though such replacements can be attested, under pre-state pre-literate conditions they were over only short distances compared with the spreads of many of the major agriculturalist language families over thousands of kilometers.

Replacement of the expanding farmers' language in the original homeland, after the expansion began. If this happened, modern language distributions might conflict with the combined evidence from genes, archaeology, skeletons, and domesticates. The original homeland might now either lack the original farmers' language family altogether, or else might support only one branch of the family compared with many branches in the periphery. Suggested examples of this tend to be controversial because they involve eradication of the original languages that can now, at best, only be reconstructed. Nevertheless, this seems to us the most plausible interpretation in some cases. For instance, one can suggest that languages closely related to Austronesian, Indo-European, and Japanese are no longer spoken in their putative ultimate homelands in South China, Anatolia, and Korea, respectively, because of the historical expansions of the Sinitic languages, Turkish, and Korean. The discovery of written documents attesting to the former existence of Hittite and other now-extinct Indo-European languages in Anatolia as well as the resulting big changes in our understanding of that language family confirm the reality of language loss in the potential homeland for that family (*31*, *32*).

Hunter-gatherer expansions. Not only farmers, but also sometimes hunter-gatherers, can expand at the expense of other hunter-gatherers, producing concordance of genes and languages without crops. Examples include the Inuit expansion eastwards across the Canadian Arctic and the Athabaskan expansion southward into the southwestern United States within the last millenium (*33*–*35*).

Language Family Origins and Spreads

Association of language family origins with agricultural homelands. If our basic hypothesis is correct, then a single agricultural homeland might have given rise to more than one language family radiating from it (*36*). This suggests that homelands will be areas where several major language families intersect geographically and where the methods of comparative linguistics suggest that those families originated. In contrast, most regions that lack independent agricultural origins should have a lesser diversity of language families. But these are not intended as universal generalizations, and regions with high numbers of language families need not always be regions of agricultural origin. For example, California in Native American times, certainly not a homeland for agriculture, is nevertheless famously diverse linguistically for other reasons including a long-term existence of partly sedentary hunter-gatherer populations. Each situation of linguistic diversity will need to be examined on its own terms.

Thus, the Fertile Crescent, or else nearby areas reached early by Fertile Crescent domesticates, is a zone of intersection for the Indo-European, Elamite (with Dravidian?), Afro-Asiatic, and Caucasian language families (*13*, *37*). In central and southern China and adjacent areas lies the intersection of the Sino-Tibetan, Tai-Kadai, Austro-Asiatic, and Hmong-Mien families, with Austronesian as an offshoot from Taiwan (Fig. 2) (*38*). Mesoamerica formed the homeland for the Uto-Aztecan, Oto-Manguean, Mayan, Mixe-Zoquean, and several other minor families. Highland New Guinea gave rise to the Trans-New-Guinea family, tropical West Africa to the Bantu subfamily and indeed to the whole Niger-Congo family of which Bantu languages form a subgroup. Linguistic relationships in the eastern United States and in the Andes-Amazon region are too uncertain to test this generalization, in part because of migrations and disease-caused language extinctions after European colonization. However, agriculture may have played a role in the spread of Iroquoian and Siouan languages in the eastern United States and of the Chibchan, Quechuan, and Aymaran languages in the Andes-Amazon (Fig. 3).

East-west versus north-south expansions. All other things being equal, crops and livestock, and people and the technologies and languages associated with them have spread more rapidly along east-west axes than along north-south axes (*4*). The reason is that day-length and seasonality (hence daylight-dependent plant germination schedules) depend only on latitude, so that sites at the same latitude but different longitudes are likely to share, and sites at the same longitude but different latitudes are likely to differ in, their domesticates, habitats, climates, diseases, and agricultural systems. As applied to linguistics, these differing average rates of spread along east-west and north-south axes may provide the underlying reason why there are three well-established language families (Indo-European, Afro-Asiatic, and Austronesian, and possibly a fourth if one includes the more controversial Altaic) with geographic ranges spanning 7000 to 14,000 km along the east-west axis of the Old World, but no language families spanning more than a few thousand kilometers along the north-south axis of the New World. These generalizations do not deny that a few language families spread rapidly north-south over much more modest distances (e.g., Bantu languages in subequatorial Africa, and Uto-Aztecan languages from Mexico into the southwestern United States).

Examples of Specific Language Families

We now examine how the basic hypothesis illuminates the dispersal histories of 15 language families, listed in an approximate sequence from unequivocal to controversial.

1. Bantu (Niger-Congo family). Beginning around 2000 B.C., farmers from the tropical West African agricultural homeland in eastern Nigeria and western Cameroon speaking early Bantu languages expanded east and then south over most of subequatorial Africa, replacing or intermarrying with most of the original inhabitants related to modern Pygmies and Khoisan people. The Bantu subgroup that covers most of this region is just one of the 177 subgroups of the whole Niger-Congo language family and comprises about

500 of the family's recorded 1436 languages. Of particular interest, as a model for what could be achieved elsewhere, is the exceptionally detailed integration of linguistic evidence with other types of evidence (from genetics, archaeology, and domesticated plants and animals) in this case (*20*, *39*–*42*).

2. Arawak (Taino). Around 400 BC, farmers from the Orinoco River of South America colonized the West Indies and eventually replaced most of the islands' earlier occupants, spreading up the Lesser Antillean chain to the Greater Antilles and Bahamas. They thereby became ancestral to the modern Taino people speaking Arawakan languages. The evidence from linguistics, pottery, and domesticates is detailed, but genetic evidence is slight because few Tainos survived European conquest. Linguistic relationships suggest that Arawakan languages had previously originated in and spread over much of the upper Amazon (*43*–*45*).

3. Austro-Asiatic, Tai ("Daic"), and Sino-Tibetan. Several independent sources of evidence suggest expansions of these three language families from agricultural homelands in China, at different times and over different geographic ranges. Austro-Asiatic spread west and south from southern China into the Indian subcontinent and Malay Peninsula (*46*), Sino-Tibetan spread from the Yellow River or Sichuan into Burma and the Himalayas (*47*, *48*). Much of the southward expansion of the Tai languages, like that of the Hmong-Mien (Miao-Yao) languages, has been within historic times. Austro-Asiatic languages are especially diverse, suggesting that they were the first of these three families to expand. They occur today almost entirely south of the political border of China, raising the possibility that they too originated in southern China, but were then largely replaced there by later expansions of Sino-Tibetan and Tai.

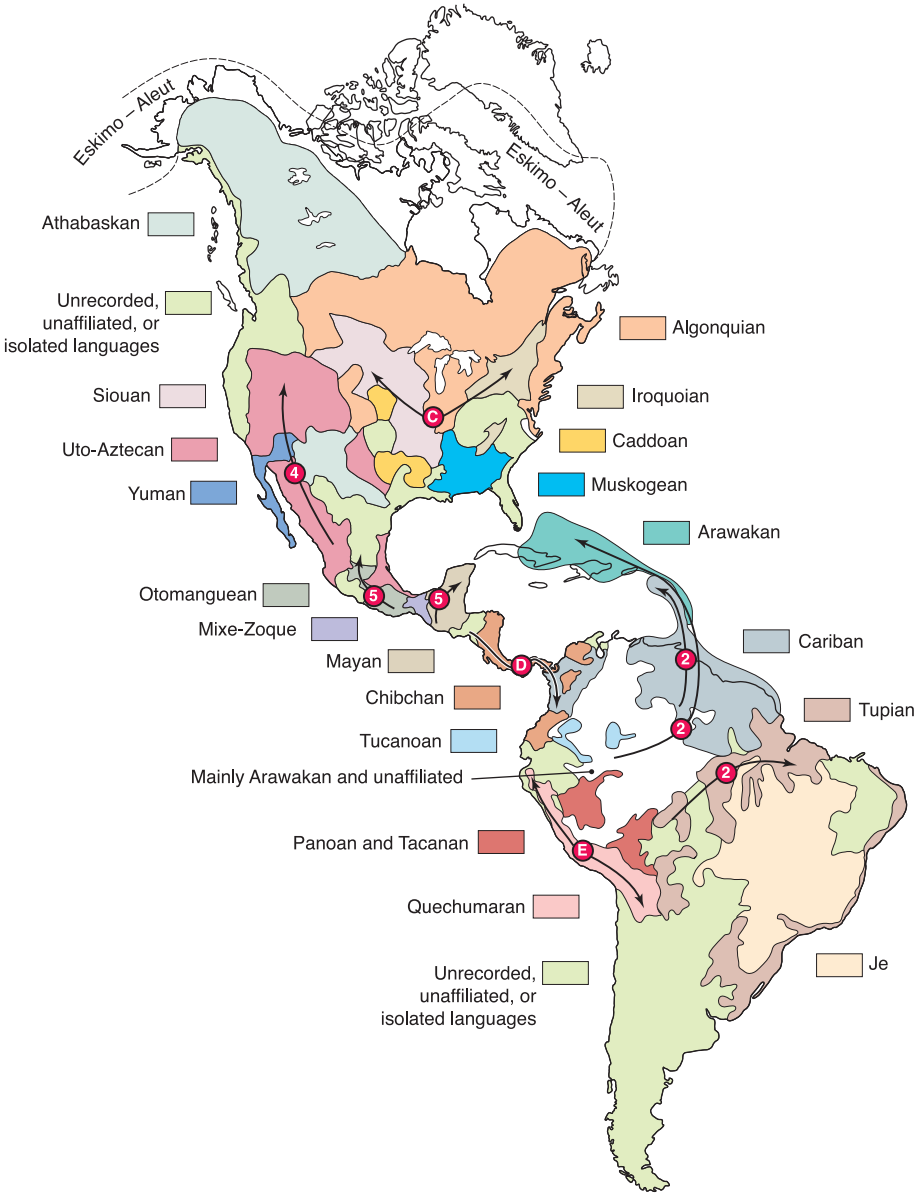

Fig. 3. Language families of the New World and their suggested expansions. Maps based on information in (*89*) and other sources. Numbered examples discussed in text are 2 (Arawakan, Cariban, and Tupian), 4 (Uto-Aztecan), 5 (Oto-Manguean and Mayan). Other possible examples mentioned only briefly: C (Iroquoian and Siouan, with maize after 500 A.D.), D (Chibchan), and E (Quechuan and Aymaran).

4. Uto-Aztecan. Maize, beans, and linguistic evidence suggest strongly that agriculture based on Mexican domesticates reached the southwestern United States from northern Mesoamerica with speakers of Uto-Aztecan languages (*49*–*51*). The principal discordance is that the northernmost groups of Uto-Aztecans (Numic, Tübatulabal, and Takic) are desert hunter-gatherers in the Great Basin and southern California, leading to the former assumption that Uto-Aztecan languages originated among northern desert hunter-gatherers who spread south and became farmers. More likely, as Hill has recently argued (*49*), northward-expanding Uto-Aztecan farmers reverted to hunting and gathering when they reached habitats either unsuitable for agriculture or rendered marginal by drought or agricultural over-exploitation, just as Austronesian farmers did in southern New Zealand when they entered a climate zone unsuitable for their main crop, the sweet potato.

5. Oto-Manguean, Mixe-Zoquean, Mayan. Oto-Manguean has the widest geographic range of any language family within the Mesoamerican agricultural homeland, spanning a distance of 1300 km from Mesoamerica's northwest boundary to its southeast boundary, although that range is still small by Old World standards. The reconstructed Proto-Oto-Manguean language, as well as Proto-Mayan and Proto-Mixe-Zoquean, includes terms for the major Mesoamerican crops, especially maize, supporting the hypothesis of an agricultural expansion. So-called glottochronological calculations of language family time depths (*43*, *52*) suggest that all four of these Mesoamerican families—Uto-Aztecan, Oto-Manguean, Mixe-Zoquean, and Mayan—spread initially between roughly 3000 and 1500 B.C. (*53*), a span that overlaps with the archaeological time span (~2500 to 1500 B.C.) for the beginnings of agricultural intensification and sedentary village life in Mesoamerica (*54*).

6. New Guinea Highlands. By far the greatest linguistic diversity in the modern world occurs on the island of New Guinea, with 1000 of the modern world's 6000 languages, and with dozens of language isolates or families that have no demonstrable relationship to each other or to any language outside New Guinea. Recent linguistic stud-

ies suggest that at least half of those 1000 languages belong to a family (the Trans-New-Guinea family) whose spread may have been driven by agricultural origins in the New Guinea Highlands (*55*). The principal uncertainties concern the age of agricultural origins in the Highlands (as early as 7000 B.C. or as late as 4000 B.C., perhaps) (*56*–*58*), the identity of the first staple crops, and the linguistic limits of the Trans-New-Guinea family.

7. Japanese. Around 400 B.C., intensive rice agriculture, new pottery styles, and new tools, all based on Korean models, appeared on the southwestmost Japanese island of Kyushu near Korea and spread northeast up the Japanese archipelago. Genes and skeletons of the modern Japanese suggest that they arose as a hybrid population between arriving Korean rice farmers and a prior Japanese population similar to the modern Ainu and responsible for Japan's earlier Jomon pottery. Modern southwest-to-northeast gene clines in Japan and DNA extracted from ancient skeletons support this interpretation (*59*, *60*). Japanese origins would thus rival Bantu origins as the most concordant and unequivocal example of an agricultural expansion, were it not for the flagrant discordance of the linguistic evidence. If Korean farmers really did become dominant in Japan as recently as 400 B.C., one might have expected the modern Japanese and Korean languages to be as closely similar as other languages that diverged at such a recent date (e.g., German and Swedish), whereas their relationship is in fact much more distant.

The likely explanation is language replacement in the Korean homeland. Early Korea consisted of three kingdoms with distinct languages. The modern Korean language is derived from that of the ancient Korean kingdom of Silla, the kingdom that unified Korea. However, the now-extinct language of one of the two ancient Korean kingdoms that Silla defeated, Koguryo, was much more similar to Old Japanese than is Sillan or modern Korean (*61*). Thus, a Koguryo-like language may have been spoken by the Korean farmers arriving in Japan, may have evolved into modern Japanese, and may have been replaced in Korea itself by Sillan that evolved into modern Korean.

8. Austronesian. Detailed archaeological evidence demonstrates the colonization of Taiwan by Neolithic pottery-making and rice-growing farmers from South China before 3000 B.C., followed by the spread of farming, pottery, and Neolithic tools to the Philippines (2000 to 1500 B.C.), then southwest to the Southeast Asian mainland and to Madagascar, and then east through Indonesia out across the Pacific to the furthest islands of Polynesia, eventually reaching New Zealand by about 1200 A.D. (*23*, *24*, *27*). In terms of distance covered, this was the world's largest pre-historic agricultural expansion, and it rivals the Bantu expansion in the degree of detail of its linguistic reconstruction (*62*, *63*). Austronesian primary subgroups are concentrated in the Taiwan homeland, just as Niger-Congo primary subgroups are concentrated in the West African homeland. These facts suggest that colonists derived ultimately from Austronesian-speaking farmers of coastal South China replaced or hybridized with the original population of the Philippines and Indonesia (related to modern Philippine Negritos, New Guineans, and Aboriginal Australians), and that language shift and hybridization were especially complex when spreading Austronesians encountered established dense populations in the New Guinea area and Northern Melanesia.

This interpretation faces three objections, at least with respect to Southeast Asia (the progression of human settlement into Remote Oceania—Polynesia, Micronesia and eastern Melanesia—is easiest to interpret because Austronesians were the first human population there). The most obvious objection is the absence of languages related to Austronesian in coastal South China today; an equally obvious explanation may be replacement there by expanding Sino-Tibetan speakers. A second issue is that rice of subtropical South Chinese origin was abandoned and tropical crops were domesticated or adopted when the expansion entered the equatorial tropics; this shift strikes us as an entirely expected development. Not a problem. A third issue is genetic studies interpreted to suggest genetic divergence between populations in what are now southern China and Island Austronesia long before 4000 B.C. (*64*), but the molecular clock calculations underlying that interpretation are controversial, and a Taiwan homeland for Austronesians is not ruled out by other genetic analyses (*65*–*67*).

9. Dravidian. Food production reached South India at about 3000 BC, partly through the spread of Fertile Crescent and Sahel domesticates via the Indus Valley and the northwestern Deccan, and partly through a simultaneous spread of rice cultivation from Southeast Asia with speakers of Austro-Asiatic (Mundaic) languages. In addition to these undoubted spreads of crops into India from elsewhere, Fuller (*68*) has recently argued for primary (independent) origins of rice, millet, and gram domestication in the Ganges Valley and South India.

The Dravidian language family is concentrated in South India, with one distinctive outlier (Brahui) far to the northwest in Pakistan, and perhaps an even more distinctive extinct outlier (Elamite) much further to the northwest in southwest Iran (Elamite's relation to Dravidian languages is debated) (*69*). Either Dravidian languages are the original languages of much of the Indian subcontinent or they arose to the west and spread at about 3000 B.C. with Fertile Crescent domesticates into the Indian subcontinent, subsequently becoming extinct in their homeland. If the latter interpretation were correct, then one would have to assume either clinal gene dilution or else language shift to explain why South Indians today are phenotypically and genetically so unlike peoples of the Fertile Crescent. Thus, for South Asian early agriculture, both the archaeological and the linguistic records remain equivocal.

10. Afro-Asiatic. This language family consists of six branches, five (including Ancient Egyptian) confined to North Africa, one (Semitic) also extending in ancient times to Southwest Asia. That distribution suggests an African origin for the family, whose Semitic branch might then have spread into Southwest Asia. But the overwhelming archaeologically attested flow of domesticated crops and animals from Neolithic times onward, into Egypt and through the Arabian Peninsula into Ethiopia, is out of Southwest Asia rather than out of Africa. That would make it surprising for Semitic languages to have spread against that stream.

There are two principal competing hypotheses for the origin of Afro-Asiatic. One, based on reconstruction of early vocabulary for cultural and environmental referents, places the homeland in the Levant during the earliest Neolithic (the late Natufian culture, 9500 B.C.) (*32*, *70*, *71*), with a subsequent two-pronged spread by 5000 B.C. that is well documented archaeologically: mixed farming across the Nile into Egypt and North Africa, giving rise to the Egyptian and Berber branches of Afro-Asiatic languages, and sheep- and goat-based pastoralism from western Arabia across the Red Sea into Ethiopia and Sudan, giving rise to the Cushitic, Omotic, and Chadic branches (Semitic spread into Ethiopia much later). That Southwest Asian origin would now be masked by language replacement in the homeland, including the spread of the Semitic branch of Afro-Asiatic languages (including Akkadian or Babylonian, Aramaic, and Arabic) in historic times. The other hypothesis, reflecting Afro-Asiatic language subgrouping but with no clear archaeological support, favors a homeland in northeastern Africa (*72*, *73*). That African origin would imply a preagricultural spread for Afro-Asiatic, perhaps with population movement into a wetter early Holocene Sahara.

11. Indo-European. We have saved for last the most intensively studied, yet still the most recalcitrant, problem of historical linguistics: the origin of the Indo-European language family, distributed before 1492 A.D. from Ireland east to the Indian subcontinent and (as represented by the extinct Tocharian

languages) western China. Unlike the Niger-Congo and Austronesian language families, each consisting of about a thousand languages that sometimes intergrade geographically, the Indo-European language family contains only 144 languages divided among 11 markedly distinct branches. These and other facts suggest that the task of reconstructing Indo-European origins is complicated by massive extinctions of Indo-European languages in the past, resulting from the expansions of a few highly successful subgroups (Germanic, Romance, Slavic, and Indo-Iranian).

The two main competing hypotheses of Indo-European origins both face severe difficulties. One hypothesis views Proto-Indo-European as having been spoken in the steppes north of the Black Sea by horse-riding nomadic pastoralists, whose supposed domestication of the horse and invention of the wheel around 4000 B.C. enabled them to expand militarily (*74*–*76*). But objections include that horse domestication and riding may not have begun until thousands of years later (*77*); that it is hard to understand (perhaps even inconceivable) how steppe pastoralists could have imposed their language on so much of Europe west of the steppes (*78*); and that even linguists who reject glottochronology agree that Indo-European languages (including Anatolian) are so different from one another that their divergence probably began before 4000 B.C. (*31*).

The other hypothesis, based on the recognition that the extinct Anatolian languages (including Hittite and Luvian, the probable language of Troy) represent the most distinctive branch and hence the earliest documented branching in the family tree, views Proto-Indo-European (or, more strictly speaking, Proto-Indo-Hittite) as a language of Neolithic Anatolian farmers who carried Fertile Crescent domesticates west into Europe, east to the Indus Valley, and north and then east across the Central Asian steppes beginning around 7000 to 6000 B.C. (*32*, *78*–*80*). But objections include that the reconstructed Proto-Indo-European lexicon has a strong bias toward domesticated animals rather than crops (*81*, *82*) and that reconstructed Proto-Indo-European words relating to wheels and wheeled vehicles suggest (some would say "prove") late Indo-European origins around the time of the invention of the wheel (~4000 B.C.) (*76*). Even we two authors of this paper have differing views on this issue.

Future Directions

We mentioned at the outset that study of the first agricultural expansions requires interdisciplinary research synthesizing five types of evidence. To resolve the many controversies concerning agricultural expansions, we need new evidence of all five types. Archaeological evidence of the first farmers is basic to many questions, such as the distinction between primary and secondary homelands of agriculture (*83*): Did agriculture really arise independently in Ethiopia, the Sahel, tropical West Africa, South India, and Amazonia? We need a more balanced record of the earliest crops and livestock, their wild relatives, and their dates and places of domestication, so that regions poorly known in this respect can take their place alongside the better understood Fertile Crescent (*9*). Detailed studies of gene gradients similar to those available for Europe and the Pacific are needed for other parts of the world; that information will be especially difficult to obtain in the New World, because so many Native American populations have disappeared or shifted in the last 500 years. Studies of ancient skeletons and of ancient DNA may give us snapshots of human distributions at known times in the past.

Most importantly, given that genetic data can, in theory, provide an ultimate test of whether farming systems and languages spread with or without people, it is essential to understand more about how genetic data should be interpreted in terms of human history. Problems with natural selection, differing rates of mutation, the effects of population bottlenecks on stochastic loss of lineages, and the modeling of clinal variation across geographical space are behind many of the current debates between geneticists, particularly in Europe (*17*, *84*, *85*) and Austronesia (*64*–*67*). Conflict in these fields is the main reason why this review is focused on languages and archaeology, rather than genetics. Early farming systems and languages have clearly spread, probably on many occasions in historical association, and we believe that human populations have spread too, but detailed documentation of this inference still lies in the future.

We also need more studies of languages themselves. Hundreds of historically important languages remain poorly described, efforts to trace so-called deep language relationships (i.e., relations between languages that diverged long ago) remain highly controversial, and relationships of New World languages are especially controversial (*43*, *86*, *87*), due in part to the scale of loss and replacement since 1500. Fortunately, linguists are today concerned with modeling the formation of linguistic diversity in time and through space (*88*), and this is a development to be applauded.

Environmental issues also arise, for not all early farmers underwent territorial expansion. The Mixe-Zoqueans essentially stayed at home, as did speakers of languages of the Caucasus and also the Hmong-Mien, before their recent movements into Southeast Asia. Evidently, the early spreads of farmers and their languages depended on being in the right place at the right time. Relevant factors would have included availability of new lands nearby into which to spread (pull factors) and perhaps resource shortfalls at home due to climate changes or prior resource mismanagement (push factors).

To extract reliable conclusions from all this evidence will require comparative research on a worldwide scale within multiple disciplines. It is quite a challenge, but a uniquely fascinating one.

References and Notes

1. B. Smith, *The Emergence of Agriculture* (Scientific American, Washington D.C., 1995).
2. D. Harris, Ed., *The Origins and Spread of Agriculture and Pastoralism in Eurasia* (Univ. College of London Press, London, 1996).
3. A. Damania *et al.*, Eds, *The Origins of Agriculture and Crop Domestication* (International Center for Agricultural Research in the Dry Areas, Aleppo, Syria, 1998).
4. J. Diamond, *Guns, Germs and Steel* (Norton, New York, 1997).
5. *The Transition to Agriculture in the Old World*, O. Bar-Yosef, Ed., special issue of *Rev. Archaeol.*, **19** (no. 2), 1998).
6. D. Piperno, D. Pearsall, *The Origins of Agriculture in the Lowland Neotropics* (Academic Press, San Diego, CA, 1998).
7. B. Smith, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1324 (1998).
8. M. van der Veen, Ed., *The Exploitation of Plant Resources in Ancient Africa* (Kluwer Academic, New York, 1999).
9. D. Zohary, M. Hopf, *Domestication of Plants in the Old World* (Clarendon, Oxford, ed. 3, 2000).
10. P. Bellwood, *Ann. Rev. Anthropol.* **30**, 181 (2001).
11. P. Bellwood, C. Renfrew, Eds., *Examining the Language/Farming Dispersal Hypothesis* (McDonald Institute for Archaeological Research, Cambridge, 2003).
12. P. Bellwood, in (*2*), pp. 465–498.
13. C. Renfrew, in (*2*), pp. 70–92. Renfrew has been one of the most active researchers in favor of strong linkages between the spreading of early farmers and language families, as discussed here, especially for Indo-European and other western Eurasian langage families. See (*11*, *36*, *78*, *80*) and further references to Renfrew's work therein.
14. L. L. Cavalli-Sforza, F. Cavalli-Sforza, *The Great Human Diasporas* (Addison-Wesley, Reading, PA, 1995).
15. A. Whittle, *Europe in the Neolithic* (Cambridge Univ. Press, Cambridge, 1996).
16. A. J. Ammerman, L. L. Cavalli-Sforza, *The Neolithic Transition and the Genetics of Populations in Europe* (Princeton Univ. Press, Princeton, NJ, 1984).
17. L. Chikhi *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 11008 (2002).
18. E. W. Hill, M. A. Jobling, D. G. Bradley, *Nature* **404**, 351 (2000).
19. C. B. Bousman, *African Archaeol. Rev.* **15**, 133 (1998).
20. C. Ehret, *An African Classical Age* (Univ. Press of Virginia, Charlottesville, VA, 1998).
21. J. F. Downs, *Animal Husbandry in Navajo Society and Culture* (Univ. of California Press, Berkeley, CA, 1964).
22. D. H. Snow in *Farmers, Hunters and Colonists*, K. A. Spielmann Ed. (Univ. of Arizona Press, Tucson, AZ, 1991), pp. 71–88.
23. P. V. Kirch, *On the Road of the Winds* (Univ. of California Press, Berkeley, CA, 2000).
24. P. Bellwood *The Polynesians* (Thames and Hudson, London, 1987).
25. A. Anderson, *Prodigious Birds* (Cambridge Univ. Press, Cambridge, 1989).
26. V. T. King, *The Peoples of Borneo* (Blackwell, Oxford, 1993).
27. P. Bellwood, *Prehistory of the Indo-Malaysian Archipelago* (Univ. of Hawaii Press, Honolulu, rev. ed., 1997).
28. L. L. Cavalli-Sforza, P. Menozzi, A. Piazza, *The History and Geography of Human Genes,* (Princeton Univ. Press, Princeton, NJ, 1994).
29. M. Ross, in *Archaeology and Language I*, R. Blench and M. Spriggs, Eds. (Routledge, London, 1997), pp. 209–261.

30. L. Reid, in *Language Contact and Change in the Austronesian World*, T. Dutton, D. Tryon, Eds. (Mouton de Gruyter, Berlin, 1994), pp. 443–76.
31. R. Drews, Ed., *J. Indo-European Studies Mono*. **38**, "Greater Anatolia and the Indo-Hittite Language Family" (Institute for the Study of Man, Washington, DC, 2001).
32. A. Dolgopolsky, *Mediterranean Lang. Rev.* **6**, 230 (1993).
33. D. Dumond, *The Eskimos and Aleuts* (Thames and Hudson, London, 1987).
34. M. Fortescue, *Language Relations across Bering Strait* (Cassell, London, 1998).
35. R. G. Matson, in (*11*), chap. 27.
36. A. Sherratt, S. Sherratt, *Antiquity* **62**, 584 (1988).
37. C. Renfrew, *Cambridge Archaeol. J.* **10**, 7 (2000).
38. P. Bellwood, in *Archaeology in Southeast Asia*, Y. Chun-tong, L. Wai-ling, Eds. (University Museum and Art Gallery, Hong Kong, 1995), pp. 11–22.
39. C. Ehret, M. Posnansky, Eds. *The Archaeological and Linguistic Reconstruction of African History* (Univ. of California Press, Berkeley, CA, 1982).
40. D. Phillipson, in (*11*), chap. 15.
41. K. Williamson, R. Blench in *African Languages*, B. Heine, D. Nurse, Eds. (Cambridge Univ. Press, Cambridge, 2000), pp. 11–42.
42. P. Underhill, *Ann. Hum. Genet.* **65**, 43 (2001).
43. L. Campbell, *American Indian Languages* (Oxford Univ. Press, New York, 1997).
44. I. Rouse, *The Tainos* (Yale Univ. Press, New Haven, CT, 1992).
45. A. Aikhenvald, in *The Amazonian Languages*, R. Dixon, A. Aikhenvald, Eds. (Cambridge Univ. Press, Cambridge, 1999), p. 75.
46. C. Higham, *The Bronze Age of Southeast Asia* (Cambridge Univ. Press, Cambridge, 1996).
47. G. van Driem, *Bull. Indo-Pacific Prehistory Assn.* **18**, 43 (1999).
48. J. Janhunen, *Manchuria: An Ethnic History* (Suomalais-Ugrilainen Seura, Helsinki, Finland, 1996).
49. J. Hill, in (*11*), chap. 26.
50. P. Bellwood, in *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, A. Aikhenvald, R. Dixon, Eds. (Oxford Univ. Press, Oxford, 2001), pp. 27–43.
51. J. Hill, *Am. Anthropol.* **103**, 913 (2001).
52. P. Bellwood, in *Time Depth in Historical Linguistics*, C. Renfrew, A. McMahon, L. Trask, Eds. (McDonald Institute for Archaeological Research, Cambridge, 2000), vol. 1, pp. 109–40.
53. T. Kaufman, in *Amazonian Linguistics*, D. L. Payne, Ed. (Univ. of Texas Press, Austin, TX, 1990), pp. 13–73.
54. R. Adams, M. J. MacLeod, Eds, *The Cambridge History of the Native Peoples of the Americas* (Cambridge Univ. Press, Cambridge, 2000), vol. II, part I.
55. A. K. Pawley, in *Perspectives on the Bird's Head of Irian Jaya, Indonesia*, J. Miedema, C. Ode, R. Dam, Eds. (Rodopi, Amsterdam, 1998), p. 684.
56. J. Golson, in *Sunda and Sahul*, J. Allen, J. Golson, R. Jones, Eds. (Academic Press, London, 1977), pp. 601–638.
57. M. Spriggs, in (*2*), pp. 524–537.
58. S. G. Haberle, A. Chepstow-Lusty, *Environ. Hist.* **6**, 349 (2000).
59. S. Horai *et al*., *Am. J. Hum. Genet.* **59**, 579 (1996).
60. K. Omoto, N. Saito, *Am. J. Phys. Anthropol.* **102**, 437 (1997).
61. M. Hudson, *Ruins of Identity* (Univ. of Hawaii Press, Honolulu, 1999).
62. R. Blust, *J. World Prehistory* **9**, 453 (1995).
63. A. K. Pawley, in (11), chap. 20.
64. S. Oppenheimer, M. Richards, *Nature* **410**, 166 (2001).
65. T. Melton *et al*. *Am. J. Hum. Genet*. **63**, 1807 (1998).
66. E. Hagelberg, in *The Archaeology of Lapita Dispersal in Oceania*, G. R. Clark *et al*., Eds. (Pandanus, Canberra, Australia, 2001), pp. 167.
67. M. Hurles *et al*., *Genetics* **160**, 289 (2002).
68. D. Fuller, in (*11*), chap. 16.
69. V. Blazek, in *Archaeology and Language IV*, R. Blench, M. Spriggs, Eds. (Routledge, London, 1998), pp. 48–78.
70. I. Diakonoff, *J. Semitic Studies* **43**, 219–217, (1998).
71. A. Militarev, in (*11*), chap. 12.
72. C. Ehret, *J. African Hist.* **20**, 161 (1979).
73. C. Ehret, *Reconstructing Proto-Afro-Asiatic* (Univ. of California Press, Berkeley, CA, 1995).
74. M. Gimbutas, *J. Indo-European Studies* **13**, 185 (1985).
75. J. Mallory, *In Search of the Indo-Europeans* (Thames and Hudson, London, 1989).
76. D. Anthony, *Antiquity* **69**, 554 (1995).
77. M. Levine, in *Late Prehistoric Exploitation of the Eurasian Steppe* (McDonald Institute for Archaeological Research, Cambridge, 1999), pp. 5–58.
78. C. Renfrew, *Archaeology and Language* (Jonathan Cape, London, 1987).
79. G. Krantz, *Geographical Development of European Languages* (Lang, New York, 1988).
80. C. Renfrew *J. Indo-European Studies* **27**, 257 (1999).
81. C. Watkins, *The American Heritage Dictionary of Indo-European Roots* (Houghton Mifflin, Boston, 1985).
82. J. Mallory, in *Archaeology and Language I*, R. Blench, M. Spriggs, Eds. (Routledge, London, 1997), pp. 93–121.
83. J. Diamond, *Nature* **418**, 34 (2002).
84. M. Richards *et al*., *Am J. Hum. Genet.* **67**, 1251 (2000).
85. L. Simoni *et al*., *Am. J. Hum. Genet.* **66**, 262 (2000).
86. J. Greenberg, *Language in the Americas* (Stanford Univ. Press, Stanford, CA, 1987).
87. M. Ruhlen, *A Guide to the World's Languages* (Stanford Univ. Press, Stanford, CA, 1987), vol. 1.
88. D. Nettle, *Linguistic Diversity* (Oxford Univ. Press, Oxford, 1999).
89. *Encyclopaedia Britannica* (Encyclopedia Britannica, Chicago, IL, ed. 15, 1987), vol. 22

90. We thank C. Hilliker for drafting the illustrations.