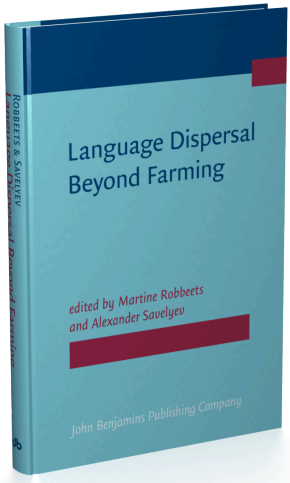
CHAPTER 8

The domestications and the domesticators of Asian rice

George van Driem | Institut für Sprachwissenschaft, Universität Bern, Switzerland

https://doi.org/10.1075/z.215.08van



Pages 183–214 of
Language Dispersal Beyond Farming
Edited by Martine Robbeets and Alexander Savelyev
2017. xiii, 324 pp.

CHAPTER 8

The domestications and the domesticators of Asian rice

George van Driem
Institut für Sprachwissenschaft, Universität Bern, Switzerland

Rice genetics has now provided molecular evidence for three distinct domestications of Asian rice, giving rise to *ahu*, *indica* and *japonica* rice and subsequently involving the multidirectional introgression of favoured alleles between these three families of *Oryza sativa* cultivars. The phylogeography of Asian wild and cultivated rice species also permits inferences with regard to the likely geographical range within which these three domestication processes involving Asian cultivated rice unfolded. Evidence from linguistic palaeontology permits the identification of two language families whose linguistic ancestors pose the likeliest candidates for the earliest rice domesticators, Austroasiatic and Hmong-Mien. The linguistic arguments and population genetic evidence on Asian rice are assessed. Recent advances in palaeobotany as well as a number of currently prevalent misunderstandings in rice archaeology are discussed. Another set of evidence from linguistic palaeontology involving reconstructible etyma denoting megafauna in light of the early Holocene distribution of these megafaunal species provides a geographical indication for the location of the early Austroasiatic homeland. Furthermore, the molecular genetics of human populations are discussed in order to shed light on the prehistory and geography of the Austroasiatic, Hmong-Mien and other language families. Finally, a synthesis of the disparate sets of evidence is presented.

Keywords: rice (*Oryza sativa*), Hmong-Mien, Austroasiatic, phylogeography, preservation bias

1. Rice genetics and rice domestications

In 1883, the director of the botanical garden in Geneva, Alphonse-Louis-Pierre Pyrame de Candolle, argued that the origin of cultivated rice lay in China and that rice was introduced to India from China (1883: 285, 309–311). Later, Nikolai̯ Ivanovič Vavilov (1926) argued against a Chinese origin for rice and contended

DOI 10.1075/z.215.08van

instead that the origin of Asian rice lay in India, whence the crop had spread to China and Japan. The old controversy about the original homeland of cultivated rice persisted well into the early years of the new millennium. In the Himalayan handbook, I have recounted how this controversy has influenced historical linguistic discourse over the years (van Driem 2001: 324–327 et passim). One might like to think that the old polarisation of arguments had been rendered obsolete ever since the evidence of molecular genetics has been brought to bear on the resolution of the question.

Three principal populations of cultivated rice *Oryza sativa* are distinguished, comprising the families of cultivars known as *ahu*, *indica* and *japonica* rice. Whereas the latter two varieties are characterised by wet cultivation, *ahu* rice is cultivated on dry fields and terraces and is sometimes referred to imprecisely as "upland rice". This dry land cultivar is known in Assamese as আহু *ahu*, in Nepali घैया as *ghaiyā* and in Bengali as আউশ *āuś*. The Assamese name *ahu* arguably provides the most apt candidate for an English name for this cultivar, both because this family of cultivars is most widespread throughout Assam and because the Assamese name *ahu* lends itself readily to being pronounced well in English. Neither the Nepali nor the Bengali name remain quite intact once uttered by someone who subjects the words to an English phonology. The Bengali name *āuś,* in particular, has the tendency to get unrecognisably transmogrified in the mouths of English speakers.

In the older literature before the turn of the millennium, *japonica* rice was often held to come from a wild precursor *Oryza rufipogon*, whereas *indica* rice was thought to derive from a wild precursor *Oryza nivara*. New research has not rendered this view entirely obsolete, but has instead refined our understanding of wild *Oryza rufipogon* as a highly diverse species that has long been undergoing a prolonged process of speciation. Rather, wild *nivara* rice can most accurately be considered to be an annual self-pollinating ecotype or subspecies of *rufipogon*, since these wild rice populations interbreed to a limited extent and therefore constitute a single internally diverse species complex. In the noughties, population genetic research based on the genome of wild and cultivated varieties of rice supported the novel hypothesis that Asian rice had been domesticated twice (Kovach et al. 2007; Sweeney & McCouch 2007; Kovach et al. 2009).

At one point, the mutation coding for a whiter grain pericarp *(rc)* changed the reddish seed of wild rice into the white seeds of modern rice. This gene is shared by the majority of rice cultivars, and the trait was held to have introgressed from *japonica* into both *indica* and *ahu* rice (Sweeney et al. 2007). Soon other parts of the tangled tale of rice domestication were unravelled. Although the *japonica* and *indica* cultivar families essentially derive from a single species of wild rice, the time of divergence of about 100,000 years calculated for the two distinct ancestral *rufipogon* subspecies from which the two cultivars had derived indicated

independent domestications. At the same time, although *ahu* rice was found to be genetically more closely affiliated to *indica* than to *japonica* rice, *ahu* rice appeared to have resulted from yet a third distinct domestication process (Londo et al. 2006). Subsequent genetic studies on Asian rice have corroborated these findings and identified the molecular footprints of three independent domestication events in different parts of Asia. Moreover, molecular evidence has demonstrated that the introgression of domesticated traits had occurred not just unidirectionally from *japonica* into *ahu* and *indica* rice, but multidirectionally from *ahu* and *indica* into *japonica* as well (McNally et al. 2009; Civáň et al. 2015).

The prehistory of rice cultivation and rice domestication is convoluted as a direct consequence of the role played by human rice cultivators. The prehistory of rice cultivation involved three distinct domestications as well as the sustained cultural exchange of rice cultivar knowledge over time between the populations of early rice cultivators. The cultivation and domestication of the annual self-pollinating *nivara* ecotype of *Oryza rufipogon* led to the development of the *indica* cultivar of *Oryza sativa*, and for various reasons it is likely that this process may have transpired in the Brahmaputra river basin. In this area, Asian rice was long cultivated before it was domesticated through selective breeding by humans, and grain shattering cultivars are still cultivated to this day. Various rice species other than *Oryza sativa* that have generally been deemed to be wild likewise continue to be cultivated in Assam. An eastern domestication of a perennial swamp subspecies of *Oryza rufipogon* gave rise to the *japonica* cultivar of *Oryza sativa*. The mutation *sh4* led to the partial development of the abcission zone where the mature grain detaches from the pedicle, and the reduced brittleness of the rachides reduced grain shattering. Subsequently, human domestication also favoured genes coding for a whiter grain pericarp *(rc)* and erecter stalks *(Prog1)*.

Several stages in the domestication of *indica* rice entailed the introduction of the traits *sh4, rc* and *Prog1* into the *nivara* gene pool through introgressive hybridisation, involving backcrossing with the *japonica* cultivar. The hill tracts surrounding the Brahmaputra river basin may have been where the domestication of *ahu* rice took place. The three domestication events which gave rise to modern rice cultivars took place long ago, and the relative popularity of many *japonica* strains today is likely to represent a secondary development on the grander time scale. Even subsequent to early cultivation, the exchange of rice cultivar knowledge between rice cultivating peoples persisted over time. The *javanica* cultivar has been demonstrated to represent a tropical variety of *japonica,* whereas a number of famous long-grained aromatic varieties, such as Indian *bāsmatī* rice, have likewise been shown to derive from *japonica* (Parsons et al. 1999; Garris et al. 2005). By contrast, Thai jasmine rice, for instance, has been shown to represent an *indica* variety, with the

fragrant allele of the betaine aldehyde dehydrogenase gene BADH2 introduced by introgression (Kovach et al. 2009).

Once, a team of geneticists ventured to conjecture that the introgression of the white grain pericarp *(rc)* allele might be construed as possible evidence for a single domestication of rice between 13,500 and 8,200 years ago, which they ventured to situate in the Yangtze river basin (Molina et al. 2011). Remarkably, this conjecture was not supported by the team's own phylogenetic data. Rather, the geneticists in question explicitly deferred to arguments advanced by archaeologists anxious to see the lower Yangtze recognised as the unique home of rice domestication (Fuller & Qin 2010; Fuller et al. 2010). On the basis of their own molecular findings, the geneticists were unable to exclude that *indica* and *japonica* had been independently domesticated, highlighting the possibility "that both *indica* and *japonica* originated from highly differentiated *Oryza rufipogon* gene pools that were not sampled" (Molina et al. 2011: 5). In fact, their evidence suggested that the wild *rufipogon* populations of the Indian subcontinent and mainland Southeast Asia, or some now extinct *rufipogon* population, may have been ancestral to all domesticated rice.

When not prejudiced *a priori* by an adamantly articulated archaeological opinion, rice geneticists have explained instead that the widespread transfer of the whiter grain pericarp *(rc)* gene more immediately "implies contact among the people who cultivated the different subspecies" (Sweeney et al. 2007: 1419). Evidence from both linguistic palaeontology and human population genetics inspired a reconstruction that involved precisely such an intense interaction between the early Yangtzeans, who were ancestral to the Hmong-Mien, and the ancient Austroasiatics (van Driem 2011, 2012). We shall recapitulate the evidence for this reconstruction and examine several of the principal implications of this model below.

By contrast, the simplistic model of a single rice domestication in the lower Yangtze advocated by some archaeologists who happen to work in that particular region not only flies in the face of the molecular genetic findings on Asian rice, this single domestication model also overlooks the human cultivators, who served not as unwitting mediators, but acted as knowledgeable agents during the three prolonged rice domestications. In their enthusiasm for the lower Yangtze basin, the archaeologists in question once allowed their reasoning to be clouded by denial of the preservation bias and consequently strayed beyond what I have called "the epistemological event horizon in archaeology" (van Driem 2017).

2. Linguistic palaeontology and the early rice cultivators

In 1830, Julius von Klaproth (1830: 112–113) became the first to discuss the prehistorical implications of the occurrence of phonologically regular reflexes in the

languages of a particular family for reconstructible etyma denoting particular plant and animal species with well-defined geographical ranges. Inspired by von Klaproth's pioneering work, Adolphe Pictet (1859) introduced the term "linguistic palaeontology" to denote an attempt to understand the ancient material culture of a language family or geographically to locate its *Urheimat* on the basis of the lexical items which can be reliably reconstructed for the common ancestral language. The reflexes of reconstructed roots should be attested across the various branches of the family, and their phonological development should be historically regular. With regard to rice, the two language families which reflect rice agriculture terminology most robustly are Austroasiatic and Hmong-Mien.

The Austroasiatic language family boasts the most impressive reconstructible repertoire of rice agriculture terms. Gérard Diffloth (2005) has adduced the following eleven reconstructible Austroasiatic roots: *(kə)ɓaːʔ 'rice plant', *rəŋkoːʔ 'rice grain', *cəŋkaːm 'rice outer husk', *kəndək 'rice inner husk', *pheːʔ 'rice bran', *təmpal 'mortar', *jənreʔ 'pestle', *jəmpiər 'winnowing tray', *guːm 'to winnow', *jərmuəl 'dibbling stick' and *kəntuːʔ 'rice complement', i.e. accompanying cooked food other than rice. Diffloth has long been the most knowledgeable authority on the comparative study of Austroasiatic. The historical phonology and grammar of Austroasiatic are not as tractable as the comparative study of Hmong-Mien, since Austroasiatic exhibits far greater internal diversity than does Hmong-Mien. Strecker's (1987) Hmong-Mien phylogeny recognised the three branches Hmongic *(Miáo)*, Mienic *(Yáo)* and Ho Nte *(Shē)*, and treated the precise classification of the Na-e language as problematic. More recently, Ratliff (2010) presented an improved Hmong-Mien family tree. In terms of its internal diversity, the Hmong-Mien language family looks like a vestigial branch of what once may have been a greater linguistic phylum, which Starosta (2005) called "Yangtzean".

Martha Ratliff (2004, 2010) identified ten rice cultivation etyma as reconstructible to the Proto-Hmong-Mien level: *hnrəaŋH 'cooked rice', *hnɔn 'rice head, head of grain', *mbləu 'rice plant, paddy', *mphi̯ɛk 'chaff', *mblut 'glutinous', *ljiŋ 'paddy field', *ljim 'sickle', *ŋkjuəX 'rice cake', *tuX 'husk/pound rice' and *tsjɛŋH 'rice steamer'. Five rice agriculture terms are reconstructible to the Proto-Hmongic level: *S-phjæ^{C} 'chaff', *mbljæ^{C} 'have food with rice', *ʔrin^{A} 'dry (rice) in sun', *ntsuw^{C} 'husked rice' and *tshɛŋ^{B} 'husked rice or millet'. The two roots *hmei^{B} 'husked rice' and the rice measure etymon *hrau^{A} are reconstructible to the Proto-Mienic level. Six of the ten reconstructible Proto-Hmong-Mien etyma are also found in Old Chinese, where, however, they are more likely to represent ancient loans into Sinitic from Hmong-Mien rather than the other way around (*pace* Ratliff 2004: 158–159).

First of all, the Old Chinese forms 秫 *məlut *(shú)* 'glutinous millet' (i.e. not rice), 田 *lˤiŋ *(tián)* 'field', 鎌 *[r]em *(lián)* 'sickle', 粔 *[g](r)a(k)-s *(jù)* 'cakes', 擣 *tˤuʔ *(dǎo)* 'pound, thresh' and 甑 *s-təŋ-s *(zèng)* 'steamer' are not reconstructible

to Trans-Himalayan, of which Sinitic is but a subgroup (van Driem 2005, 2007, 2014a; Old Chinese forms as given by Baxter & Sagart 2014a, 2014b; cf. Ho 2016). Similarly, Ratliff relates Proto-Hmongic *ʔjɛŋ^{A} 'seedling' and Proto-Mienic *ʔjaŋ^{A} 'seedling' to Middle Chinese 秧 *ʔjang *(yāng)*, evidently due to a discrepancy in vocalism between the Hmongic and Mienic forms, and relates Proto-Hmongic *ljɛŋ^{A} 'rice measure' to Old Chinese 量 *[r]aŋ *(liàng, liáng),* but again neither etymon is well reflected in Trans-Himalayan outside of Sinitic. Moreover, not only are the earliest textual attestations of the Chinese forms 田 *lˤiŋ 'field', 粔 *[g](r)a(k)-s 'cakes', 甑 *s-təŋ-s 'steamer' and the measure word 量 *[r]aŋ absolutely ambiguous as to what kind of grain they refer to (though 田 *lˤiŋ 'field' may reflect a Hmong-Mien loan into Sino-Bodic rather than just into Sinitic), furthermore the form 粔 *[g](r)a(k)-s 'cakes' is not actually an Old Chinese form, for its first known attestation occurs in the poetry anthology of the feudal state of Chǔ, entitled 楚辭 *Chǔcí*, dating from the Hàn period, whereas 鎌 *[r]em 'sickle' likewise first occurs in the Hàn period as a western dialect word (Wolfgang Behr, p.c., 19 April 2011).

The Proto-Mienic *hmei^{B} 'husked rice' appears to correspond to Old Chinese 米 *[m]ˤijʔ *(mǐ)*, and rice terms containing a bilabial nasal initial appear in other languages in the east of the Trans-Himalayan area, e.g. Bái *me*^{33} 'husked rice', Jinuo *a*^{44}*mɛ*^{44} 'rice', Black Lahu *mi*^{33} 'paddy', Nusu *me*^{33}*me*^{31} 'rice', Garo *mi,* Dimasa *mai* 'rice', Tangkhul *ma* 'paddy', Sgaw Karen *me* 'boiled rice'. Yet the meanings of these forms are disparate, viz. paddy, hulled rice, boiled rice, and the forms may represent mere look-alikes, since no regular phonological correspondence is yet known to obtain between them. Paul Benedict "set up" a Bodo-Koch proto-form *mey or *may 'rice, paddy' (1972: 149), which Matisoff later inflated to "*ma ⇔ *may *or* *mey" (2003: 216, 231) by adding a "monophthongal allofam" and stressing the uncertainty of the rhyme. In fact, no rice agricultural terminology can be confidently reconstructed for the Trans-Himalayan phylum as a whole, an issue noted by Blench (2009).

Rice cultivation terminology is likely to have been borrowed into Sinitic from ancient Hmong-Mien rice cultivators at a time when Proto-Sinitic millet growers intensified their cultural exchange with their southern neighbours. The main split in the Hmong-Mien family is between Hmongic and Mienic. The scattered distributions of the modern language communities belonging to each of these two branches exhibit approximately the same geographical range, which is roughly bisected by the Pearl River. On the basis of the historical sources, it has long been mooted that the geographical centre of gravity of the family would originally have lain further north along the middle Yangtze (Cushman 1970). The historically attested distribution of the early Hmong-Mien tribes during the Eastern Zhōu (770–256 BC) is shown in Figure 1. There is currently no palaeobotanical evidence for the co-cultivation

of rice and foxtail millet along the middle Yangtze until around 3800 BC (Nasu et al. 2006).

Figure 1. The relative position of early Hmong-Mien (Miáo-Yáo) tribes and early Kradai (T'ai) tribes with respect to the Yangtze and to Old Chinese territory in late Zhōu times, with the hatched portion representing the imperial domain (reproduced from Forrest 1948: 129)

Population genetic findings indicate three distinct domestications of Asian rice. Linguistic palaeontology provides evidence that enables us to ascertain the likely ethnolinguistic identity of two of the three Asian rice domesticators, i.e. the ancient Austroasiatics and the ancient Hmong-Mien. It might appear parsimonious

to ascribe the domestication of the *japonica* cultivar putatively to the Hmong-Mien and the domestication of *indica* and perhaps also *ahu* rice to the ancient Austroasiatics, but the prehistorical reality may have been more intricate. A more interesting proposal emerging from a synthesis of the disparate sets of evidence is presented below. First, however, we shall address problems with the archaeology of rice agriculture and with the argumentation used by archaeobotanists advocating a single original domestication of Asian rice in the lower Yangtze basin.

3. Challenges to the archaeology of rice agriculture

The archaeology of rice agriculture is plagued by an empirical quandary commonly known in the field as a preservation bias. This empirical issue pertains to the archaeological recoverability of rice agriculture sites. The traces of ancient farming communities tend to have been better preserved in the hill tracts surrounding the Brahmaputra flood plains than on the fertile fields themselves. Likewise, in the Yangtze river basin, most salvageable rice agriculture sites are in the foothills or at the base of the foothills (Nakamura 2010). Yet the earliest rice-based cultures may first have developed on those very flood plains. Perhaps the remains of the first rice cultivating cultural assemblages lie buried forever deep beneath the silty sediments of the sinuous lower Brahmaputra basin. Maybe the palaeobotanical evidence for the earliest domestications of rice was washed out by the Brahmaputra long ago and now lies submerged in the depths of the Bay of Bengal.

Archaeologists have looked for the remains of early rice agriculture and indeed found them at some sites and not at some others. The recovered remains of early cultivated rice are of differing antiquity and reflect distinct stages of domestication. Unsurprisingly, archaeologists have not found the remains of early rice agriculture in those places where they have not yet bothered to look. Vast swathes of Asia covering the areas identified by rice geneticists (Londo et al. 2006; Molina et al. 2011; Civáň et al. 2015) as harbouring likely sites for the domestication of Asian rice have not been subjected to systematic archaeological and palaeobotanical investigation. The archaeology of northeastern India, the Indo-Burmese borderlands, Burma and the northern Bay of Bengal littoral is virtually unresearched. Political, cultural, geographical and logistic factors have conspired to impede intensive archaeological research in a vast area extending from the lower Brahmaputra basin to the Tenasserim.

Despite the molecular genetic evidence for three independent rice domestications and multidirectional introgression of alleles between the three families of cultivars *ahu*, *indica* and *japonica*, Fuller argued in several publications for a single domestication of Asian rice near the mouth of the Yangtze, where circumstances

and substrate conditions happen strongly to have favoured the preservation of the palaeobotanical remains of early agriculture (Fuller & Qin 2009, 2010; Fuller 2012). His team then resorted to modelling in an attempt to buttress their claim with their archaeological assumptions built into the model (Silva et al. 2015). The model yielded the intuitively satisfying result that the rate of exchange of alleles accelerates over time as domestication progresses, but the trouble with the simulation was that the data fed into the model were largely fortuitous in terms of their geography.

The epistemological problem here is fundamental in nature and, as the old saw has it, the absence of evidence does not constitute the evidence of absence. Fuller (2012), though cursorily acknowledging this problem, initially continued to stress the absence of palaeobotanical evidence in areas where archaeologists had not sought such evidence. The argument for a single domestication in the Lower Yangtze relied on a tacit denial of the ramifications of the preservation bias and on the conceit that the absence of evidence somehow represented the evidence of absence. Continued reliance on this conceit became untenable in face of the utter dearth of archaeobotanical research on rice agriculture in most of the relevant areas (van Driem 2011). The advice was evidently taken to heart, and the population genetic findings on rice were also heeded, inspiring an intended programme of archaeobotanical research that now fortunately envisages the targeting of these regions (Stevens et al. 2016; Fuller et al. 2016).

In consonance with previous rice genetic findings, Choi et al. (2017) conceded the molecular evidence for "significant gene flow in both directions" between the three families of cultivars *ahu*, *indica* and *japonica*. Yet once again on the basis of the entrenched archaeological argumentation, Choi et al. (2017) attempted to mitigate the observed introgression of alleles from *ahu* and *indica* into the *japonica* family of cultivars by speculating that the "introgression from *aus/indica* to *japonica*, however, may have occurred during the diversification phase of rice". Trying to reinterpret inconvenient and possibly contradictory molecular genetic findings for Asian rice in order to fit them into the mould of a single domestication in the lower Yangtze leads further afield from an interdisciplinary consilience on rice and has brought Choi et al. (2017) to what they have rather optimistically qualified as "a paradox". Similarly, several incongruous conclusions drawn by Huang et al. (2012) are debunked by Civáň et al. (2015).

Despite the archaeological work conducted in the Ganges and Yangtze basins, much of the archaeology of ancient rice agriculture simply remains unknown because little substantive work has been done in the most relevant areas, e.g. northeastern India, Bangladesh, the Indo-Burmese borderlands and Burma. The gargantuan lacunae in archaeological research highlight the impotence of argumentation in favour of a single domestication around the mouth of the Yangtze that denies the epistemological consequences of preservation bias, and even palliates

those molecular genetic findings that are inconvenient to the lower Yangtze unique rice cradle narrative. Future archaeological research will have to come to terms with both the reality and the ramifications of the strong preservation bias in rice agriculture archaeology. Many parts of northeastern India and the Indo-Burmese borderlands have maintained highly diverse rice cultures to the present day. One archaeologist of cereal cultivation in China has cogently argued the need for expanding the scope of archaeological research beyond the Yangtze river basin into these areas, i.e. Lu (2006, 2009).

At the same time, the absence of evidence for rice agriculture of great antiquity in mainland Southeast Asia, despite the relatively more well researched archaeology of the region, presently embarrasses those who have lately taken to espousing Robert von Heine-Geldern's (1917) homeland theory for Austroasiatic around the lower course of the Mekong, without acknowledging the original author of this hypothesis (Sidwell & Blench 2011). However, the fact that the archaeology of northeastern India, the Indo-Burmese borderlands, Burma and the northern Bay of Bengal littoral is virtually unresearched does not similarly compromise homeland proposals in this region. Moreover, the various rice cultivation methods practised in the Brahmaputra basin to this day and the nature of the substrate render it unlikely that palaeobotanical remains would ever be found, notwithstanding the long-term practice of rice agriculture in the region, as meticulously documented by Hazarika (2014, 2017). This incontrovertible given presents an additional epistemological challenge to archaeologists who propound that rice was domesticated around the mouth of the Yangtze.

Furthermore, the argumentation in favour of a single original rice domestication in the lower Yangtze basin also relies heavily on an exaggerated importance attributed to domestication in a highly restricted sense and on grain shattering. This undue emphasis stems inevitably from the archaeological focus on the micromorphological study of rice remains. Domestication in the restricted semantic sense of genetic modification by human agency was perhaps not in all places and at all times as pivotal as Fuller has made it out to be in his writings. It has been claimed that foxtail millet *Setaria italica* and broomcorn millet *Panicum miliaceum* were already collected in the middle Yellow River valley 23,000 ago and already cultivated 19,500 years ago, a full ten millennia anterior to domestication (Li 2015). Li's early dates are certainly questionable, however, and Hu et al. (2008) have argued that millet does not appear to have been a very important source of dietary protein until some time after domestication. Yet the fact remains that grain cultigens were gathered in the wild and subsequently cultivated for long stretches of time before the process of domestication began (Larsona et al. 2014). Moreover, some cultigens never or hardly undergo much domestication in the restricted sense of measurable microanatomical modifications by artificial genetic selection.

In a similar vein, the human domestication of Asian rice favoured the mutation *sh4*, which codes for the partial development of the abcission zone where the mature grain detaches from the pedicle so that the diminished brittleness of the rachides reduced grain shattering. It was human agency that facilitated the introgression of genes coding for a whiter grain pericarp *(rc)* and erecter stalks *(Prog1)* from one family of rice cultivars into another. However, domestication that can be measured in terms of morphological differences in microanatomical structure is not necessary for sustained cultivation over long spans of time.

A number of species of wild rice do not just commonly occur, but are also reportedly still cultivated in northeastern India, e.g. *Oryza rufipogon*, *Oryza nivara*, but especially *Oryza officinalis, Oryza meyeriana, Oryza perennis* and *Oryza granulata*. The shattering of the rice grains onto the field surface does not in practice impede the harvesting of such rice, which continues to be gathered both for human consumption and for use as animal feed (Hazarika 2005, 2006, 2013, 2017). In addition to such cultivated "wild" rice species, many hundreds of indigenous *Oryza sativa* cultivars are grown in this region. Cultivated Asian rice is harvested three times a year in most areas throughout the Brahmaputra basin, using different seasonal cultivation regimes.

The *ahu* family of cultivars is most usually sown directly onto rain-fed upland fields, mainly for swidden or ঝুম *jhūm* cultivation, but this group also exhibits considerable diversity. The usual growing season in lower areas extends from late March to early July, in the mid hills from late April to early October, and in the upper hills from late June to late December. An early harvest is also practised in some areas, with a growing season from February to May, in which case the rice seedlings are transplanted and irrigated. Some other *ahu* cultivars with a growing season from May to August may likewise employ transplanted seedlings, which may or may not be irrigated.

Another family of rice cultivars is known as শালি *śāli* [xali]. The growing season for these lowland rice cultivars usually stretches from late July to early December, and for some varieties a late growing season from late August to early January is observed. The rice seedlings are transplanted, and the rice is irrigated. Another family of rice cultivars is known as বড়ো *baḍo* [bɔɹɷ]. These wetland cultivars are sown in stagnant wetlands or in irrigated fields. The growing season is from late November to early May. It may be significant that the name of this set of rice cultivars in Assamese happens to be homophonous with the Assamese name for the indigenous Trans-Himalayan ethnic group dispersed throughout the Brahmaputra basin. Another family of rice cultivars is known as আচৰা *ācrā* [asɹa]. These shallow-water cultivars grow in water that is one to two feet deep. The growing season stretches from late March to early December. Yet another family of rice cultivars is বাও *bāo* [bao]. These deep-water cultivars grow in water that is two to five feet deep,

and can thrive in water that is more than twice that deep, and the growing season stretches from late March to early December (Hazarika 2014, 2017).

Despite weaknesses in the reasoning employed by archaeologists in their eagerness to gain recognition for the lower Yangtze basin as the unique cradle of rice domestication, the archaeology of rice agriculture has nonetheless produced important results. The domestication of *japonica* rice through genetic modification by selective breeding was possibly effectuated along the Yangtze by people, who previously relied far more heavily on the collecting of acorns, water chestnuts and foxnuts before becoming reliant on rice cultivation. In terms of measurable modifications to microanatomical morphology, the process of domestication appears to have begun in the middle of the sixth millennium and to have been largely completed by the end of the fifth millennium BC (Fuller et al. 2009; Nakamura 2010; Zhao 2010; Fuller & Qin 2009; Ruddiman et al. 2008; Fuller, Harvey & Qin 2007). Currently the oldest datable domesticated rice remains from the Pearl River delta date from ca. 3000 BC (Yang et al. 2016).

Rice cultivation reached the Yellow River basin during the third millennium BC (Crawford & Shen 1998) and Formosa and Vietnam between 2500 and 2000 BC (Higham & Lu 1998), but only spread throughout the Indochinese peninsula between 1500 and 500 BC (Weber et al. 2010; Oxenham et al. 2015). It has been claimed that rice may have been cultivated in the Gangetic basin as early as 7000 BC (Sharma et al. 1980; Pal 1990; Agrawal 2002), but the current datable evidence for the actual domestication of rice in the middle Ganges dates from no earlier than the second half of the third millennium BC. In line with the molecular genetics, archaeogenetic data from Asian rice remains found in sites in India and Thailand show hybridisation between *indica* and *japonica* cultivars of domesticated rice after their initial domestications (Castillo et al. 2016), even though the sterility of hybrids sometimes acts as a barrier that helps to keep the two cultivars distinct (Chen et al. 2008).

Both broomcorn and foxtail millet agriculture were practised in the high and arid hills of what today is Sìchuān province from ca. 4000 to 2500 BC. By 2700 BC, both rice and foxtail millet were cultivated by the inhabitants of the Bǎodūn culture (ca. 2700–1700 BC) in the Chéngdū plain in what today is west-central Sìchuān (d'Alpoim Guedes 2011; d'Alpoim Guedes et al. 2013). Based on the dating of the few known sites, such as མཁར་རོ་ mKhar-ro near ཆབ་མདོ་ Chab-mdo (van Driem 2001:430–431), it has been conjectured that the spread of agriculture to the Tibetan plateau was posterior to this date by archaeologists who envisage the agricultural colonisation of Sìchuān and eastern Tibet as proceeding from the middle Yangtze (d'Alpoim Guedes et al. 2014; d'Alpoim Guedes 2015). Although it appears likely that agriculture facilitated human habitation of the Tibetan plateau at around this time (Chen et al. 2015), various types of evidence indicate that the Tibetan plateau

was permanently occupied long beforehand (Xiang et al. 2013; Huerta-Sánchez et al. 2014; Lorenzo et al. 2014; van Driem 2015a; Lou et al. 2015; Hackinger et al. 2016; Lu et al. 2016). Indeed, eastern Tibet and modern Sìchuān lay beyond the periphery of the ancient rice corridor, which extended from the Brahmaputra basin to the Yangtze basin by way of Burma and Yúnnán.

4. Zooming in on the Austroasiatic and Hmong-Mien homelands

Scholars have sought to situate the Austroasiatic *Urheimat* as far west as the Indus valley and as far east as the Yangtze delta or insular Southeast Asia. Purely from the point of view of the current geographical distribution of Austroasiatic language communities, more logical contenders for the Austroasiatic homeland are the Indian subcontinent, the Bay of Bengal littoral, mainland Southeast Asia and the middle Yangtze. The gaping lacunae in palaeobotanical research are convenient to the argument in favour of the middle and lower Yangtze basin, where conditions happen to have favoured the preservation of archaeologically recoverable remains. Linguistically, the old hypothesis that proposed Old Chinese 江 *kˢroŋ (*jiāng*) 'Yangtze' to be a loan from Austroasiatic emboldened Pulleyblank (1983) to envision a major Austroasiatic presence all along the eastern seaboard from Việtnam to Shāndōng, and to impute an Austroasiatic ethnolinguistic identity to the Lóngshān horizon. This interpretation of the linguistic data has notably been challenged by Zhāng (1998).

Four types of evidence help us to zoom in on the possible geographical location of the Austroasiatic homeland. The first type of evidence, already mentioned, is linguistic and involves the current geographical distribution of Austroasiatic language communities, which is shown in Figure 2. Both the centre of gravity of the phylum on the basis on the geographical distribution of modern Austroasiatic language communities as well as the deepest phylogenetic divisions in the family tree point to the northern Bay of Bengal littoral. The deepest historical division in the family's phylogeny lies between Munda in the west and Khasi-Aslian in the east, which would put the homeland on either side of the Ganges and Brahmaputra delta. Even the deepest division within the Khasi-Aslian trunk, i.e. the split into Khasi-Pakanic and Mon-Khmer, would suggest a point of dispersal for Khasi-Aslian between South Asia proper and mainland Southeast Asia proper. The family tree of Austroasiatic, showing the correct phylogenetic position for Pearic, presented by Diffloth for the first time at Agay in 2012, is shown in Figure 3. The internal phylogeny of the Munda branch has not, however, been established.

The second and third type of evidence involve linguistic palaeontology. The Proto-Austroasiatic rice terms adduced above, reconstructed by Gérard Diffloth,

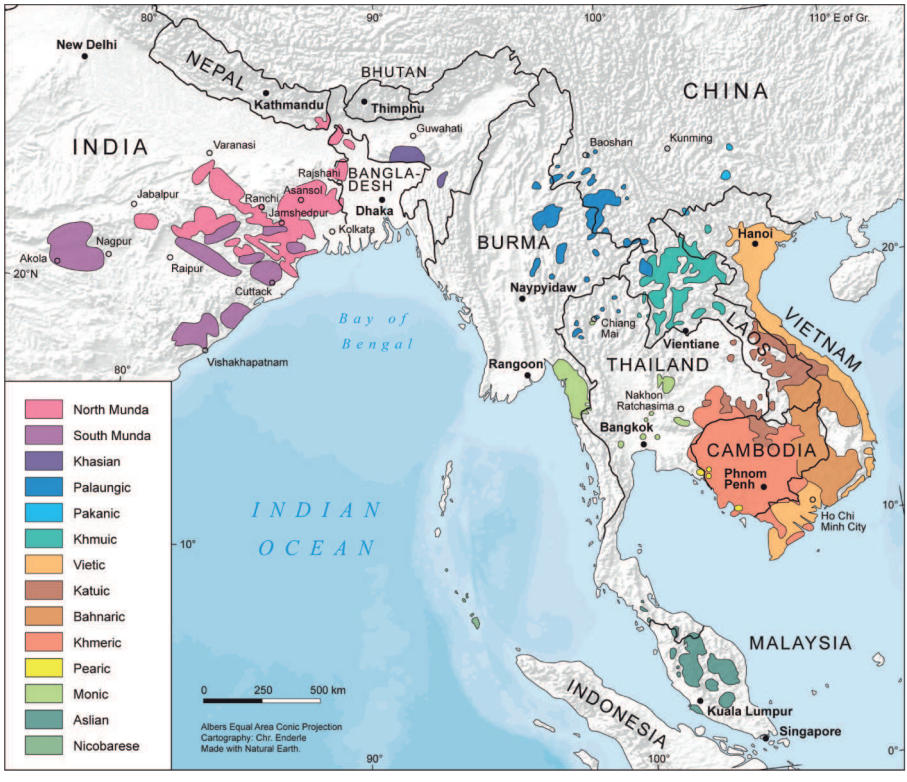


Figure 2. The geographical distribution of Austroasiatic language communities

constitute the second set of evidence. The suspected geographical ranges for the three rice domestications identified by Londo et al. (2006) on the basis of the geographical distribution of genetic markers in the wild precursor *Oryza rufipogon* are shown in Figure 4. The third set of evidence involves reconstructed roots denoting megafauna in the Proto-Austroasiatic lexicon in light of the attested geographical distribution of these species in the Holocene. This set of evidence formed the topic of an earlier study (van Driem 2012), for which Anne-Marie Bacon and Danièle Fouchier of the research unit *Dynamique de l'Évolution Humaine* at the *Centre National de la Recherche Scientifique* in Paris generously furnished the Holocene distribution maps. The Proto-Austroasiatic etyma reconstructed by Gérard Diffloth (2005: 78) evoke the fauna and ecology of a tropical humid homeland environment:

> *mraːk 'Indian peafowl *Pavo cristatus*' or 'Javan peafowl *Pavo muticus*'
> *tərkuət 'tree monitor *Varanus nebulosus* or *bengalensis*'
> *tənyuːʔ 'binturong *Arctitis binturong*'
> *(bən)joːl ~ *j(ərm)oːl 'Sunda pangolin *Manis javanica*' or 'Chinese pangolin *Manis pentadactyla*'

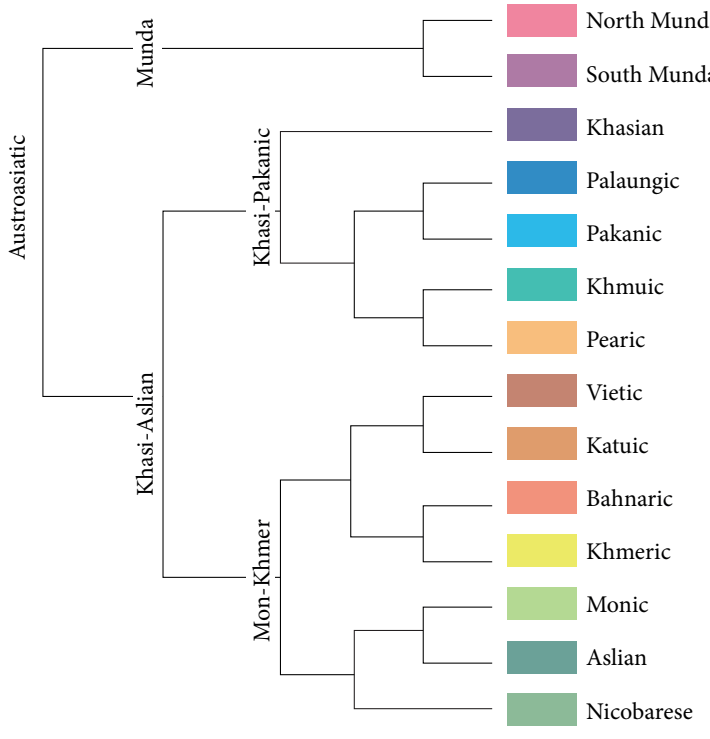


Figure 3. The family tree of Austroasiatic (Diffloth 2012). Unlike the Khasi-Aslian branch, the internal phylogeny of the Munda branch has not been established

> *dəkan 'Sumatran bamboo rat *Rhizomys sumatrensis*', 'Chinese bamboo rat *Rhizomys sinensis*', 'hoary bamboo rat *Rhizomys pruinosus*'
> *kaciaŋ 'the Asian elephant *Elephas maximus*'
> *kiaɕ 'mountain goat *Capricornis sumatrensis*'
> *rəmaːs 'Indian rhinoceros *Rhinoceros unicornis*', 'Javan rhinoceros *Rhinoceros sondaicus*' or 'Sumatran rhinoceros *Dicerorhinus sumatrensis*'

The Holocene distribution maps included in the 2012 study are not reproduced here. Instead, Figure 5 offers a synthesis of the mapped data by depicting the area where the ranges of the species for which the Proto-Austroasiatic lexicon has reconstructible etyma overlap in northeastern India, the Indo-Burmese borderlands and Burma. A comparison of Figures 4 and 5 shows that the areas suggested for an Austroasiatic homeland by the two sets of linguistic palaeontological evidence correspond to a large degree. The fourth and last set of evidence pertains to human populations genetics.

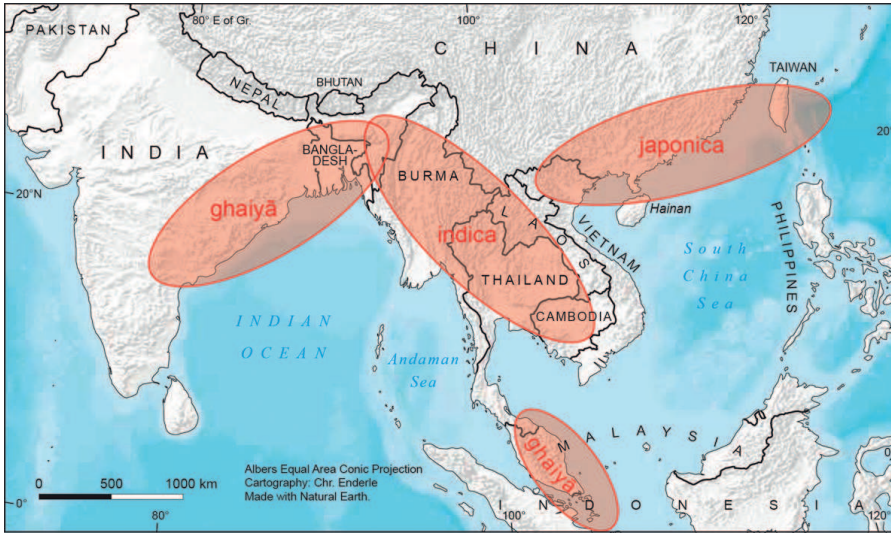


Figure 4. The geographical ranges for the possible domestication of (A) *ghaiyā* or upland rice, (B) wet *indica* rice and (C) the *japonica* cultivar, based on the geographical distribution of genetic markers in the wild precursor *Oryza rufipogon* (adapted from Londo et al. 2006)

Figure 5. The region of overlap of the geographical ranges of megafaunal species for which Proto-Austroasiatic etyma are reconstructible

5. The Father Tongue correlation and the East Asian linguistic phylum

Evidently, it cannot be repeated too often that a proto-language can only be reconstructed on the basis of linguistic evidence and that the linguistic ancestors of any modern language community were not necessarily the same people as the community's biological forebears. Although these points have long been reiterated from the time of Julius von Klaproth (1823) and Max Müller (1872), these lessons are often lost on some audiences. By the same token, each of us has countless ancestors via numerous lineages. There is no such thing as a pure race. In fact, in molecular genetic terms there is no such thing as race (Cavalli-Sforza, Menozzi and Piazza 1994). We are all members of one large human family. Moreover, even when languages and genes happen to exhibit a correlation, such a marker relationship should not be confused with identity. The correlation of a particular chromosomal marker with the distribution of a certain language family must not be simplistically equated with populations speaking languages of a particular linguistic phylum. Rather, molecular markers on the Y chromosome serve as proxies or tracers for the movements of paternal ancestors.

When studying the distribution of maternally inherited markers in the mitochondrial DNA and paternally inherited markers on the Y chromosome, a Swiss-Italian team of population geneticists soon found that it was easier to find statistically relevant correlations between the language of a particular community and the paternally inherited markers prevalent in that community than between the language and the most salient maternally inherited markers found in that speech community. This Father Tongue correlation was first described by Poloni et al. (1997, 2000). On the basis of this finding, it was inferred that paternally inherited polymorphisms may serve as markers for linguistic dispersals in the past, and that a correlation of Y chromosomal markers with language may point towards male-biased linguistic intrusions. The Father Tongue correlation is ubiquitous but not universal. Its preponderance allows us to deduce that a mother teaching her children their father's tongue must have been a prevalent and recurrent pattern in linguistic prehistory.

There are a number of reasons why we might expect this outcome. The Y chromosome underwent a global bottleneck towards the end of the last ice age, when certain paternal clades started eradicating or out-competing other clades (Karmin et al. 2015). The founding dispersals of many major language families appear to be related to the robust spread and reproductive success of the bearers of a subset of Y chromosomal haplogroups that survived this bottleneck. As a consequence, the global phylogeography of Y chromosomal haplogroups is shallower in terms of time depth than the worldwide mitochondrial landscape. The initial human colonisation of any virgin part of the planet must have involved both sexes in order for

a population of progeny to establish itself. Once a population is in place, however, subsequent migrations could have been heavily gender-biased. Subsequently, male intruders could impose their language whilst availing themselves of the womenfolk already in place. In this regard, population geneticist Toomas Kivisild (2014) has wryly characterised warfare as a sex-specific pathology linked to the Y chromosome. Whereas the landscape of paternal lineages often appears to correlate with language at the comparatively shallower time depth of the linguistically reconstructible past, correlations between maternal lineages and linguistic phylogeography discerned to date have been underwhelming. The Father Tongue hypothesis suggests that linguistic dispersals were, at least in most parts of the world, posterior to initial human colonisation and that many linguistic dispersals were predominantly later male-biased intrusions. Such patterns are observed worldwide.

In two previous studies, I have shown that the geographical distribution and phylogeography of subclades of the Y chromosomal haplogroup O appear to be correlated with the dissemination of four recognised language families, viz. Austroasiatic, Trans-Himalayan, Hmong-Mien and Austro-Tai (van Driem 2014b, 2015b). These four language families were united into a single East Asian linguistic phylum in a hypothesis proposed by Starosta (2005). In presenting my own tweaked recension of Starosta's East Asian family tree in 2012 in Benares (van Driem 2014b), shown in Figure 6, I pointed out that Starosta was the most recent exponent of a long tradition of linguists who had attempted to unite one or more of these language families into a grander linguistic phylum and, in so doing, ventured beyond the epistemological constraints of what I call the "linguistic event horizon". This horizon is the maximal time depth accessible through methodologically sound linguistic reconstruction and the boundary beyond which any reconstructions are at one point reduced to sheer speculation. Scholars who have proposed earlier renditions of the East Asian linguistic phylum have ranged from methodologically rigorous historical linguists such as Blust (1996) to megalocomparativists such as Benedict (1942), and from those offering just unsupported conjecture, e.g. Schlegel (1901, 1902), to those providing sound evidence in the form of phonologically regular correspondences, e.g. Ostapirat (2005, 2013).

The shared morphological vestiges adduced by Starosta in support of his East Asian linguistic phylum comprised the agentive prefix *<m->, the patient suffix *<-n>, what he called the instrumental prefix *<s-> and what he termed the perfective prefix *<n->. A discussion of the merits of the evidence advanced by Starosta for this linguistic phylum strikes me as being of little utility, since I consider the phylum to lie at the linguistic event horizon and therefore doubt whether this issue can ever be conclusively resolved on the basis of firmly reconstructible linguistic evidence. Rather, Starosta himself proposed that the "potential utility" of his hypothesis lay "in helping to focus scholars' efforts on particular specific questions,

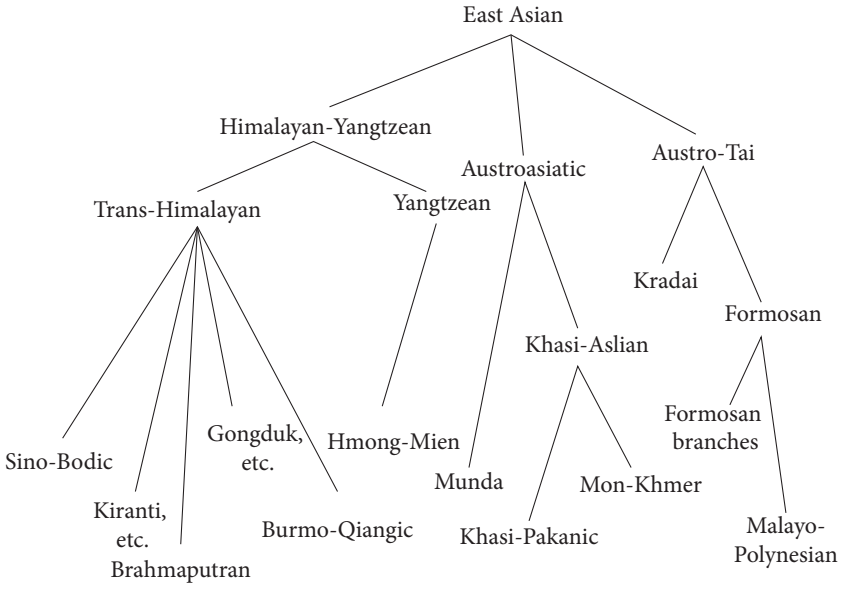


Figure 6. The 2012 Benares recension of Stanley Starosta's 2001 Périgueux East Asian linguistic phylum (Starosta 2005; van Driem 2014b)

resulting in the replacement of parts of this hypothesis with better supported arguments" (2005: 194).

The resolution of the Y chromosomal tree is constantly being enhanced. Haplogroup labels are updated to reflect our improved understanding of the phylogeny. Mutations numbers tend to remain unchanged, provided that the markers in question prove to be reliable in defining haplogroups. Conventional haplogroup labels of the Y Chromosome Consortium are still widely in use, but have been replaced here with the newer labels of the International Society of Genetic Genealogy, reflecting refinements incorporated up to the 12th of May 2017. In my two previous studies, I noted that the paternal haplogroup O1b1a1a (M95) was correlated with populations speaking languages belonging to the Austroasiatic language family, the haplogroup O2a2b1 (M134) with the Trans-Himalayan language family, the haplogroup O2a2a1a2 (M7) with Hmong-Mien and the haplogroup O1 (F265, M1354) with the Austro-Tai language family.

The complex history of Sinitic populations featured successive constellations of dynastic empires governed from geographically ever shifting capitals, whereby subjugated and neighbouring populations as well as immigrants were absorbed. Not surprisingly therefore, Hàn Chinese populations tend to represent an amalgam of East Asian paternal lineages. Yet even in Hàn Chinese populations, the molecular marker associated with the spread of a Trans-Himalayan father tongue, i.e.

haplogroup O2a2b1 (M134), taken together with its subclade O2a2b1a1 (M117), occurs in a much higher frequency than any other O subclade, and approximately twice as frequently as the next most frequent fraternal subclade O2a1c (002611) (Yan et al. 2011; Wang et al. 2013; Yao et al. 2017).

In observing the non-random correlation of these four recognised language families with subclades of the paternal haplogroup O, I speculated that the four major East Asian language families were the result of prehistoric bottlenecks. Palaeolithic populations were small, and the effective founder population sizes of the major modern paternal subclades must have been quite small, whilst new populations arise from the small surviving subsets that have passed through bottlenecks. The four language families Austroasiatic, Trans-Himalayan, Hmong-Mien and Austro-Tai appear to have arisen in this way in correlation with specific paternal lineages.

In another study, we showed that the Munda branch of Austroasiatic had arisen as the result of a sexually biased linguistic intrusion into the Indian subcontinent from the region to the north of the Bay of Bengal (Chaubey et al. 2010). As a consequence of the comparatively younger date and the nearly absolute gender asymmetry of this linguistic intrusion, it appears that the deepest division within the Khasi-Aslian trunk of Austroasiatic, i.e. the split between Khasi-Pakanic and Mon-Khmer, might perhaps be more indicative of the geographical location of the Austroasiatic homeland than the split between Munda and Khasi-Aslian. If we accept this line of reasoning, then the point of dispersal for Khasi-Aslian would appear to have lain in the area between South Asia proper and mainland Southeast Asia proper.

6. Rice and the East Asian dispersal

Long before the linguistically reconstructible past, at a time that lay well beyond the linguistic event horizon, the paternal haplogroup K (M9) was centred in the area between South Asia and Southeast Asia, where the ancestral K* appears to have been situated. This clade spawned many successful paternal lineages, some of which moved into insular Southeast Asia, i.e. the haplogroups S (M69) and M (M304), whereas other clades moved back westward into South Asia and beyond, viz. the haplogroups Q (M242), R (M201), T (M89) and L (M429) (Karafet et al. 2015). The geographical locus of yet another descendant subclade lay in the Eastern Himalaya, i.e. the ancestral haplogroup NO (M214). Millennia after the two paternal lineages N and O had split up, the bearers of haplogroup N set out for East Asia just after the last glacial maximum, braving ice and tundra, and – in a grand counterclockwise sweep – migrated across northern Eurasia as far as west as Lappland, whilst the

ancestral form *N appears to have been situated in northern Burma (Rootsi et al. 2007; Derenko et al. 2007; Mirabal et al. 2009; Ilumäe et al. 2016).

The paternal clade O is a marker that was overwhelmingly shared by the linguistic ancestors of what Starosta (2005) called the East Asian linguistic phylum. The non-random correlation of the subclades of this particular Y chromosomal haplogroup with the four recognised language families enables us to infer the following sequence of events. Millennia before the end of the last glacial maximum, the paternal lineage O (M175) split into the subclades O2 (M122) and O1 (F265, M1354), as shown in Figure 7. The two subclades can be putatively assigned to two geographical loci, with the haplogroup O1 (F265, M1354) moving eastward into East Asia south of the Yangtze, whilst bearers of the O2 (M122) haplogroup settled in the general region of the Eastern Himalaya.

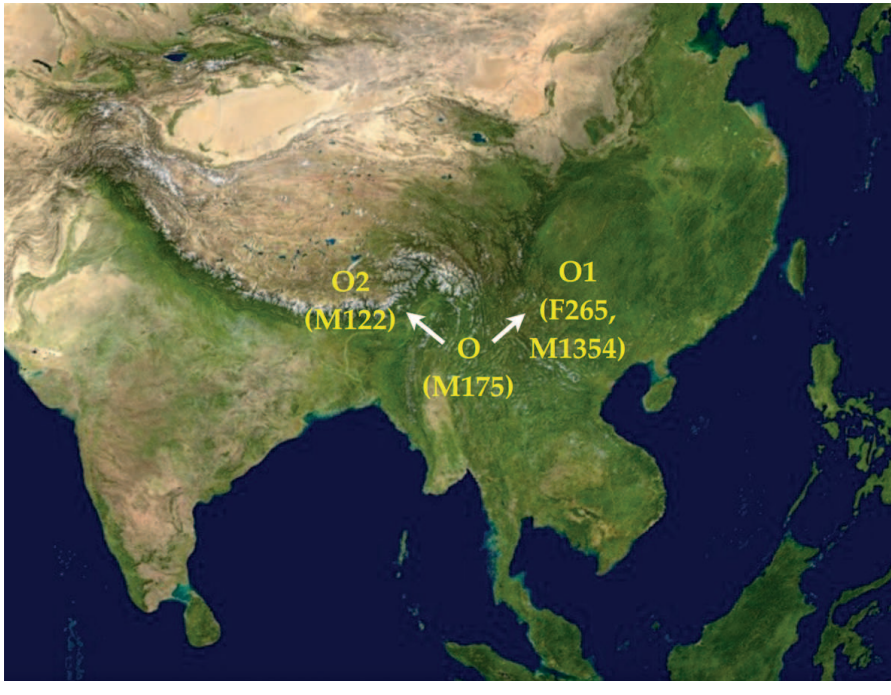


Figure 7. After the last glacial maximum, the Y chromosomal haplogroup O (M175) split into the subclades O1 (F265, M1354) and O2 (M122)

Subsequently, as temperature and humidity increased after the last glacial maximum, haplogroup O split further into the paternal lineages that serve as tracers for the spread of Trans-Himalayan, Hmong-Mien, Austroasiatic and Austro-Tai. The O1 (F265, M1354) lineage south of the Yangtze split into the subclades O1b (M268) and O1a (M119), with the latter moving eastward to the Fújiàn hill tracts and across

the strait to settle on Formosa, which so became the *Urheimat* of the Austronesians (cf. Abdulla et al. 2009). Subsequently, the subclade O1b (M268) gave rise to the filial subclades O1b2 (M176) and O1b1a1a (M95). The bearers of haplogroup O1b1a1a (M95) became the progenitors of the Austroasiatics (van Driem 2007; Chaubey et al. 2010). The Austroasiatics spread throughout the Salween drainage and thence to southern Yúnnán, northern Thailand and western Laos. In time, the Austroasiatics would spread as far as the Mekong delta, the Malay peninsula and the Nicobars. Secondarily, bands of male Austroasiatics would introduce both their language and their paternal lineage, O1b1a1a (M95), to the indigenous peoples of the Choṭā Nāgpur, as shown in Figure 8.

Figure 8. A male-biased linguistic intrusion introduced both Austroasiatic language and a paternal lineage, haplogroup O1b1a1a (M95), into the indigenous population of the Choṭā Nāgpur

The linguistic palaeontological evidence adduced above shows that the ancestral Austroasiatics practised rice agriculture, whilst the geographical distribution of haplogroup O1b1a1a (M95) correlates neatly with populations speaking Austroasiatic languages. The inference can therefore be made that Asian rice was cultivated by the ancestral bearers of haplogroup O1b1a1a (M95). The fraternal clade O1b2 (M176), which we may call "para-Austroasiatic", spread eastward, where they disseminated

rice agriculture to the lower Yangtze. Although the genetic legacy of the eastward migration of the bearers of the O1b2 (M176) persists residually today in mainland East Asia, these ancestral fathers left no linguistic trace of the father tongue that they once spoke, except for perhaps an old name for the Yangtze river that was ultimately borrowed by Old Chinese as 江 *kˤroŋ (*jiāng*), as proposed by Pulleyblank (1983).

This para-Austroasiatic paternal lineage O1b2 (M176) advanced as far as the Korean peninsula and also represents a major wave of immigration recorded in the Japanese genome. We can identify the O1b2 (M176) lineage with the Yayoi people, who introduced rice agriculture to Japan, as early as the second millennium BC, during the final phase of the Jōmon period. In addition to rice, the Yayoi also introduced other crops of continental origin to Japan such as millet, wheat and melons. The gracile Yayoi immigrants soon outnumbered the more robust and less populous Jōmon people, who were Palaeolithic hunters and foragers and the descendants of earlier waves of peopling, including the first anatomically modern humans to populate the Japanese archipelago.

About twelve thousand years ago, at the dawn of the Holocene, in the southeastern Himalayas and eastern slopes of the Tibetan Plateau, haplogroup O2 (M122) gave rise to the ancestral Trans-Himalayan paternal lineage O2a2b1 (M134) and the "Yangtzean" or Hmong-Mien paternal lineage O2a2a1a2 (M7), as shown in Figure 9. It is a reasonable conjecture that the bearers of the polymorphism O2a2b1 (M134) at first remained in the Eastern Himalaya, which today continues to represent the centre of phylogenetic diversity of the Trans-Himalayan language family based on the geographical distribution of primary linguistic subgroups. Only later would early Trans-Himalayan language communities spread into northeastern India, southeastern Tibet and northern Burma, but first the bearers of the O2a2a1a2 (M7) lineage migrated eastward to settle in the areas south of the Yangtze. On their way, the early Hmong-Mien encountered the ancient Austroasiatics, from whom they adopted rice agriculture. The intimate interaction between ancient Austroasiatics and the ancestral Hmong-Mien not only involved the sharing of knowledge about rice agriculture technology, but also left a genetic trace in the high frequencies of haplogroup O1b1a1a (M95) in today's Hmong-Mien and of haplogroup O2a2a1a2 (M7) in today's Austroasiatic populations.

On the basis of these Y chromosomal haplogroup frequencies, Cai et al. (2011: 8) observed that Austroasiatics and Hmong-Mien are "closely related genetically" and ventured to speculate about "a Mon-Khmer origin of Hmong-Mien populations". It would be more accurate to infer that the incidence of haplogroup O2a2a1a2 (M7) in Austroasiatic language communities of Southeast Asia indicates a significant Hmong-Mien paternal contribution to the early Austroasiatic populations whose descendants settled in Southeast Asia, whereas the incidence of haplogroup O2a2a1a2 (M7) in Austroasiatic communities of the Indian subcontinent

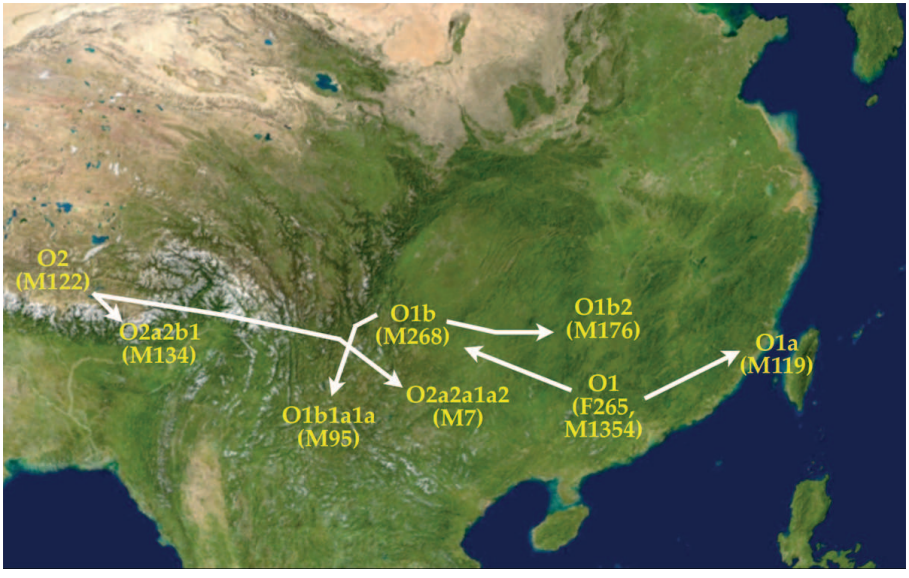


Figure 9. At a more recent time depth, paternal lineages branched into new subclades, and each event involved a linguistic bottleneck leading to language families that today are reconstructible as distinct linguistic phyla. The O1 (F265, M1354) lineage gave rise to the O1a (M119) and O1b (M268) subclades. The former moved eastward to the Fújiàn hill tracts and across the strait to Formosa, which so became the *Urheimat* of the Austronesians. Bearers of the paternal lineage O1b (M268) domesticated Asian rice and spawned the paternal subclades O1b1a1a (M95) and O1b2 (M176). Haplogroup O1b1a1a (M95) is the Proto-Austroasiatic paternal lineage, whereas the para-Austroasiatic fraternal clade O1b2 (M176) spread eastward, sowing seed along the way. The haplogroup O2 (M122) gave rise to the paternal subclades O2a2b1 (M134) and O2a2a1a2 (M7). The spread of the molecular marker O2a2b1 (M134) from the Eastern Himalaya serves as a tracer for the dissemination of people speaking languages of the Trans-Himalayan family, whereas the paternal lineage O2a2a1a2 (M7) serves as a tracer for the spread of people speaking languages of the Hmong-Mien family.

is undetectably low. The incidence of the Y chromosomal haplogroup O1b1a1a (M95) amongst the Hmong-Mien appears to indicate a slightly lower Austroasiatic paternal contribution to Hmong-Mien populations than vice versa. As the Hmong-Mien moved eastward, the bearers of para-Austroasiatic haplogroup O1b2 (M176) likewise continued to move east.

Three domestications of Asian rice *Oryza sativa*, involving the cultivar families *ahu*, *indica* and *japonica*, took place through the agency of ancient rice cultivators who bore three distinct paternal lineages, i.e. the Austroasiatic paternal subclade

O1b1a1a (M95), the para-Austroasiatic paternal lineage O1b2 (M176) and the "Yangtzean" or Hmong-Mien paternal lineage O2a2a1a2 (M7). The region between the Brahmaputra river basin and the Yangtze river basin runs through Burma and southern Yúnnán and harbours numerous ecotypes and topographies. In this area, the domestication of three different families of Asian rice cultivars took place, each suited to a different ecology.

The three populations involved not only exchanged paternal lineages but also rice knowledge which enabled the introgression of favoured traits between the three families of cultivars *ahu*, *indica* and *japonica*. I propose that the cultivar families *ahu* and *indica* were first cultivated by the ancient Austroasiatics and by the ancient Hmong-Mien or Yangtzeans, whereas the domestication of *japonica* rice was conducted by the bearers of the para-Austroasiatic paternal lineage O1b2 (M176), who left no linguistic trace other than perhaps an old para-Austroasiatic toponym for the Yangtze, but whose descendants surfaced in the archaeological record of the Japanese archipelago as the people behind the Yayoi culture.

Meanwhile, the bearers of Y chromosomal haplogroup O2a2b1 (M134) in the eastern Himalayan region expanded further eastward throughout Sìchuān and Yúnnán, north and northwest across the Tibetan plateau as well as further westward across the Himalayas and southward into the Indo-Burmese borderlands. On the Brahmaputra plain, the early Trans-Himalayans encountered the Austroasiatics, who had preceded them. The relative frequencies of the Y chromosomal haplogroup O1b1a1a (M95) in Trans-Himalayan speaking populations of the Indian subcontinent (Sahoo et al. 2006; Reddy et al. 2007) suggest that a subset of the paternal ancestors of some Trans-Himalayan populations in northeastern India, e.g. certain Bodo-Koch communities, may originally have been Austroasiatic speakers who were linguistically assimilated by Trans-Himalayans.

Finally, the ancestral Trans-Himalayan paternal lineage O2a2b1 (M134) spread from the Eastern Himalaya in a northeasterly direction to the North China plain. At a much later and shallower time depth, the Trans-Himalayan paternal lineage O2a2b1 (M134) spread in tandem with early Sinitic speaking populations southward expansion from the Yellow River basin into southern China during the Qín dynasty in the third century BC. The ancestral Trans-Himalayan paternal lineage O2a2b1 (M134) is intrusively present in the Korean peninsula and beyond.

References

Abdulla, Mahmood Ameen *et al.* 2009. Mapping human genetic diversity in Asia. *Science* 326: 1541–1545. doi:10.1126/science.1177074

Agrawal, Dharma P. 2002. The earliest pottery and agriculture in South Asia. In *The Origins of Pottery and Agriculture*, Yoshinori Yasuda (ed.), 81–88. New Delhi: Lusre Press and Roli Books, for the International Center for Japanese Studies.

d'Alpoim Guedes, Jade. 2011. Millets, rice, social complexity and the spread of agriculture to the Chengdu plain and southwest China. *Rice* 4(3): 104–113. doi:10.1007/s12284-011-9071-1

d'Alpoim Guedes, Jade. 2015. Rethinking the spread of agriculture to the Tibetan plateau. *Holocene* 25(9): 1–13.

d'Alpoim Guedes, Jade *et al.* 2013. Site of Baodun yields earliest evidence for the spread of rice and foxtail millet agriculture to south-west China. *Antiquity* 87: 758–771. doi:10.1017/S0003598X00049449

d'Alpoim Guedes, Jade *et al.* 2014. Moving agriculture onto the Tibetan plateau: the archaeobotanical evidence. *Archaeological and Anthropological Sciences* 6(3): 255–269. doi:10.1007/s12520-013-0153-4

Baxter, William Hubbard & Sagart, Laurent. 2014a. Baxter-Sagart Old Chinese reconstruction, version 1.1. <http://ocbaxtersagart.lsait.lsa.umich.edu> (20 September 2014). doi:10.1093/acprof:oso/9780199945375.001.0001

Baxter, William Hubbard & Sagart, Laurent. 2014b. *Old Chinese: A New Reconstruction*. Oxford: OUP. doi:10.1093/acprof:oso/9780199945375.001.0001

Benedict, Paul King. 1942. Thai, Kadai, and Indonesia: A new alignment in southeastern Asia. *American Anthropologist* 44: 576–601. doi:10.1525/aa.1942.44.4.02a00040

Benedict, Paul King. 1972. *Sino-Tibetan: A Conspectus*. Cambridge: CUP. doi:10.1017/CBO9780511753541

Blench, Roger Marsh. 2009. If agriculture cannot be reconstructed for Proto-Sino-Tibetan, what are the consequences? Paper presented at the 42nd Conference on Sino-Tibetan Languages and Linguistics at Chiang Mai, 3 November.

Blust, Robert. 1996. Beyond the Austronesian homeland: The Austric hypothesis and its implications for archaeology. In *Prehistoric Settlement of the Pacific*, Ward H. Goodenough (ed.), 117–160. Philadelphia PA: American Philosophical Society.

Cai Xiaoyun *et al.* 2011. Human migration through bottlenecks from Southeast Asia into East Asia during Last Glacial Maxiumum revealed by Y chromosomes. *Public Library of Science* 6(8): e24282.

de Candolle, Alphonse-Louis-Pierre Pyrame. 1883. *Origine des plantes cultivées*. Paris: Librairie Germer Ballière et Compagnie.

Castillo, Cristina Cobo *et al.* 2016. Archaeogenetic study of prehistoric rice remains from Thailand and India: evidence of early japonica in South and Southeast Asia. *Archaeological and Anthropological Sciences* 8(3): 523–543. doi:10.1007/s12520-015-0236-5

Cavalli-Sforza, Luigi Luca, Menozzi, Paolo & Piazza, Alberto. 1994. *The History and Geography of Human Genes*. Princeton NJ: Princeton University Press.

Chaubey, Gyaneshwer *et al.* 2010. Population genetic structure in Indian Austroasiatic speakers: The role of landscape barriers and sex-specific admixture. *Molecular Biology and Evolution* 28(2): 1013–1024. doi:10.1093/molbev/msq288

Chen, Jiongjiong *et al.* 2008. A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of *indica-japonica* hybrids in rice. *Proceedings of the National Academy of Sciences* 105(32): 11436–11441. doi:10.1073/pnas.0804761105

Chen, Fahu *et al.* 2015. Agriculture facilitated permanent human occupation of the Tibetan Plateau after 3600 B.P. *Science* 347 (6219): 248–250. doi:10.1126/science.1259172

Choi, Jae Young *et al.* 2017. The rice paradox: Multiple origins but single domestication in asian rice. *Molecular Biology and Evolution* <https://doi.org/10.1093/molbev/msx049>. doi:10.1093/molbev/msx049

Civáň, Peter, Craig, Hayley, Cox, Cymon J. & Brown, Terence A. 2015. Three geographically separate domestications of Asian rice. *Nature Plants*, 1, article No. 15164 <doi:10.1038/nplants.2015.164>. See van Driem (2011: 131–132) for additional references on rice and the population genetics of rice.

Crawford, Gary W. & Shen, Chen. 1998. The origins of rice agriculture: Recent progress in East Asia. *Antiquity* 72: 858–866. doi:10.1017/S0003598X00087494

Cushman, Richard D. 1970. Rebel Haunts and Lotus Huts: Problems in the Ethnohistory of the Yao. PhD dissertation, Cornell University.

Derenko, Miroslava *et al.* 2007. Y-chromosome haplogroup N dispersal from south Siberia to Europe. *Journal of Human Genetics* 52: 763–770. doi:10.1007/s10038-007-0179-5

Diffloth, Gérard. 2005. The contribution of linguistic palaeontology to the homeland of Austroasiatic. In *The Peopling of East Asia: Putting Together the Archaeology, Linguistics and Genetics*, Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds), 77–80. London: Routledge Curzon. doi:10.4324/9780203343685_chapter_5

Diffloth, Gérard. 2012. The four registers of Pearic. Paper presented at the *22nd Meeting of the Southeast Asian Linguistics Society*, Agay, 2 June 2012.

van Driem, George. 2001. *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region, containing an Introduction to the Symbiotic Theory of Language*, 2 Vols. Leiden: Brill.

van Driem, George. 2005. Tibeto-Burman vs. Indo-Chinese: Implications for population geneticists, archaeologists and prehistorians. In *The Peopling of East Asia: Putting Together the Archaeology, Linguistics and Genetics*, Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds), 81–106. London: Routledge Curzon. doi:10.4324/9780203343685_chapter_6

van Driem, George. 2007. The diversity of the Tibeto-Burman language family and the linguistic ancestry of Chinese. *Bulletin of Chinese Linguistics* 1(2): 211–270. doi:10.1163/2405478X-90000023

van Driem, George. 2011. Rice and the Austroasiatic and Hmong-Mien homelands. In *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*, Nick J. Enfield (ed.), 361–389. Canberra: Pacific Linguistics.

van Driem, George. 2012. The ethnolinguistic identity of the domesticators of Asian rice. *Comptes Rendus Palévol* 11(2): 117–132. doi:10.1016/j.crpv.2011.07.004

van Driem, George. 2014a. Trans-Himalayan. In *Trans-Himalayan Linguistics*, Nathan Hill & Thomas Owen-Smith (eds), 11–40. Berlin: Mouton de Gruyter.

van Driem, George. 2014b. A prehistoric thoroughfare between the Ganges and the Himalayas. In *50 Years after Daojali-Hading: Emerging Perspectives in the Archaeology of Northeast India*, Tiatoshi Jamir & Manjil Hazarika (eds), 60–98. New Delhi: Research India Press.

van Driem, George. 2015a. Health in the Himalayas and the Himalayan homelands. In *Tibetan and Himalayan Healing*, Charles Ramble & Ulrike Roesler (eds), 161–194. Kathmandu: Vajra Publications.

van Driem, George. 2015b. The Himalayas as a prehistoric corridor for the peopling of East and Southeast Asia. In *Nepal: An Introduction to the Natural History, Ecology and Human Environment in the Himalayas*, Georg Miehe & Colin Pendry (eds), 318–325. Edinburgh: Royal Botanic Garden Edinburgh.

van Driem, George. 2017. The epistemological event horizon in archaeology. In *Prehistory and Archaeology of Northeast India*, Manjil Hazarika (ed.), xv–xvii. New Delhi: OUP.

Forrest, Robert Andrew Dermod. 1948. *The Chinese Language*. London: Faber and Faber.

Fuller, Dorian Q., Harvey, Emma & Qin, Ling. 2007. Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity* 81: 316–331. doi:10.1017/S0003598X0009520X

Fuller, Dorian Q. & Qin, Ling. 2009. Water management and labour in the origins and dispersal of Asian rice. *World Archaeology* 41 (1): 88–111. doi:10.1080/00438240802668321

Fuller, Dorian Q. *et al.* 2009. The domestication process and domestication rate in rice: Spikelet bases from the Lower Yangtze. *Science* 323: 1607–1610. doi:10.1126/science.1166605

Fuller, Dorian Q. & Qin, Ling. 2010. Declining oaks, increasing artistry, and cultivating rice: The environmental and social context of the emergence of farming in the Lower Yangtze Region. *Environmental Archaeology* 15 (2): 139–159. doi:10.1179/146141010X12640787648531

Fuller, Dorian Q. *et al.* 2010. Consilience of genetics and archaeobotany in the entangled history of rice. *Archaeological and Anthropologiocal Sciences* 2(2): 115–131. doi:10.1007/s12520-010-0035-y

Fuller, Dorian Q. 2012. Pathways to Asian civilizations: Tracing the origins and spread of rice and rice cultures. *Rice* 4 (3–4): 78–92.

Fuller, Dorian Q., Weisskopf, Alison R. & Castillo, Cristina Cobo. 2016. Pathways of rice diversification across Asia. *Archaeology International* 19: 84–96. doi:10.5334/ai.1915

Garris, Amanda J., Tai, Thomas H., Coburn, Jason, Kresovich, Steve & McCouch, Susan. 2005. Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169: 1631–1638. doi:10.1534/genetics.104.035642

Hackinger, Sophie *et al.* 2016. Wide distribution and altitude correlation of an archaic high-altitude adaptive EPAS1 haplotype in the Himalayas. *Human Genetics* 135 (4): 393–402. doi:10.1007/s00439-016-1641-2

Hazarika, Manjil. 2005. Neolithic Culture of Northeast India with Special Reference to the Origins of Agriculture and Pottery. MA thesis, Deccan College, Pune.

Hazarika, Manjil. 2006. Neolithic culture of northeast India: A recent perspective on the origins of pottery and agriculture. *Ancient Asia* 1: 25–43. doi:10.5334/aa.06104

Hazarika, Manjil. 2013. Prehistoric cultural affinities between Southeast Asia, East Asia and Northeast India: An exploration. In *Unearthing Southeast Asia's Past*, Vol. 1, Marijke J. Klokke & Véronique Degroot (eds), 16–25. Singapore: National University.

Hazarika, Manjil. 2014. Austroasiatic and Tibeto-Burman Linguistic Dispersals in Northeast India: An Archaeological Perspective. PhD dissertation, University of Bern.

Hazarika, Manjil. 2017. *Prehistory and Archaeology of Northeast India*. New Delhi: OUP. doi:10.1093/oso/9780199474660.001.0001

von Heine-Geldern, Robert. 1917. Kopfjagd und Menschenopfer in Assam und Birma und ihre Ausstrahlungen nach Vorderindien. *Mitteilungen der Anthropologischen Gesellschaft in Wien* XXXVII: 1–65.

Higham, Charles & Lu, Tracey. 1998. The origins and dispersal of rice cultivation. *Antiquity* 72: 867–877. doi:10.1017/S0003598X00087500

Ho Dah-an (何大安). 2016. Such errors could have been avoided: Review of *Old Chinese: A New Reconstruction* by William H. Baxter and Laurent Sagart. *Journal of Chinese Linguistics* 44(1): 175–230. doi: 10.1353/jcl.2016.0004

Hu, Yaowu, Wang, Shougong, Luan, Fengshi, Wang, Changsui & Richards, Michael P. 2008. Stable isotope analysis of humans from Xiaojingshan site: Implications for understanding the origin of millet agriculture in China. *Journal of Archaeological Science* 35: 2960–2965. doi: 10.1016/j.jas.2008.06.002

Huang, Xuehui *et al.* 2012. A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490: 497–503. doi: 10.1038/nature11532

Huerta-Sánchez, Emilia *et al.* 2014. Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512: 194–197. doi: 10.1038/nature13408

Ilumäe, Anne-Mai *et al.* 2016. Human Y chromosome haplogroup N: A non-trivial time-resolved phylogeography that cuts across language families. *American Journal of Human Genetics* 99: 163–173. doi: 10.1016/j.ajhg.2016.05.025

Karafet, Tatiana M., Fernando L. Mendez, Herawati Sudoyo, J. Stephen Lansing & Michael F. Hammer. 2015. Improved phylogenetic resolution and rapid diversification of Y-chromosome haplogroup K-M526 in Southeast Asia. *European Journal of Human Genetics* 23: 369–373.

Karmin, Monika *et al.* 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome Research* 25 (4): 459–466. doi: 10.1101/gr.186684.114

Kivisild, Toomas. 2014. Mapping chromosome diversity in Eurasia. Paper presented at the conference Migrations and Transfers in Prehistory Asian and Oceanic Ethnolinguistic Phylogeography, University of Bern, 28 to 30 July.

von Klaproth, Julius Heinrich. 1823. *Asia Polyglotta*. Paris: A. Schubart.

von Klaproth, Julius Heinrich. 1830. Réponse à quelques passages de la préface du roman chinois intitulé: Hao khieou tchhouan, traduit par M. J. F. Davis. *Journal Asiatique* V: 97–122.

Kovach, Michael J., Sweeney, Megan T. & McCouch, Susan. 2007. New insights into the history of rice domestication. *Trends in Genetics* 23(11): 578–587. doi: 10.1016/j.tig.2007.08.012

Kovach, Michael J., Calingacion, Mariafe N., Fitzgerald, Melissa A. & McCough, Susan R. 2009. The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proceedings of the National Academy of Sciences of the United States of America* 106 (34): 14444–14449. doi: 10.1073/pnas.0904077106

Larsona, Greger *et al.* 2014. Current perspectives and the future of domestication studies. *Proceedings of the National Academy of Sciences* 111(17): 6139–6146. doi: 10.1073/pnas.1323964111

Li Liu. 2015. A long process towards agriculture in the middle Yellow River valley, China: Evidence from macro- and micro-botanical remains. *Journal of Indo-Pacific Archaeology* 35: 3–14.

Londo, Jason P., Chiang, Yu-chung, Hung, Kuo-Hsiang, Chiang, Tzen-Yuh & Schaal, Barbara A. 2006. Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultuvated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America* 103(25): 9578–9583. doi: 10.1073/pnas.0603152103

Lorenzo, Felipe R. *et al.* 2014. A genetic mechanism for Tibetan high-altitude adaptation. *Nature Genetics* 46(9): 951–958. doi: 10.1038/ng.3067

Lou, Haiyi *et. al.* 2015. A 3.4-kb copy-number deletion near *EPAS1* is significantly enriched in high-altitude Tibetans but absent from the Denisovan sequence. *American Journal of Human Genetics* 97 (1): 54–66. doi: 10.1016/j.ajhg.2015.05.005

Lu, Dongsheng *et al.* 2016. Ancestral origins and genetic history of Tibetan highlanders. *American Journal of Human Genetics* 99 (3): 580–594. doi: 10.1016/j.ajhg.2016.07.002

Lu, Tracey (= Lǔ Lièdān). 2006. The occurrence of cereal cultivation in China. *Asian Perspectives* 45 (2): 129–158. doi:10.1353/asi.2006.0022

Lu, Tracey (= Lǔ Lièdān). 2009. Prehistoric coexistence: The expansion of farming society from the Yangtze River Valley to western South China. *Asian Perspectives (Senri Ethnological Studies)* 74: 47–52.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman*. Berkeley CA: University of California Press.

McNally, Kenneth L. *et al.* 2009. Genome-wide SNP variation reveals relationships among landraces and modern varieties of rice. *Proceedings of the National Academy of Sciences of the United States of America* 106(30): 12273–12278. doi:10.1073/pnas.0900992106

Mirabal, Sheyla *et al.* 2009. Y chromosome distribution within the geolinguistic landscape of northwestern Russia. *European Journal of Human Genetics* 2009: 1–14.

Molina, Jeanmaire *et al.* 2011. Molecular evidence for a single evolutionary origin of domesticated rice. *Proceedings of the National Academy of Sciences of the United States of America* 108(20): 8351–8356. doi:10.1073/pnas.1104686108

Müller, Friedrich Max. 1872. *Über die Resultate der Sprachwissenschaft: Vorlesung gehalten in der kaiserlichen Universitæt zu Strassburg am XXIII. Mai MCCCCLXXII*. Strassburg: Karl J. Trübner & London: Trübner & Co.

Nakamura Shin-ichi. 2010. The origin of rice cultivation in the Lower Yangtze region. *China. Archaeological and Anthropological Sciences* 2(2): 107–113. doi:10.1007/s12520-010-0033-0

Nasu, Hirō, Momohara, Arata, Yasuda, Yoshinori & He, Jiejun. 2006. The occurrence and identification of *Setaria italica* (L.) P. Beauv. (foxtail millet) grains from the Chengtoushan site (ca. 5800 cal B.P.) in central China, with reference to the domestication centre in Asia. *Vegetation History and Archaeobotany* 16(6): 481–494. doi:10.1007/s00334-006-0068-4

Ostapirat, Weera. 2005. Kra-Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution. In *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds), 107–131. London: Routledge Curzon. doi:10.4324/9780203343685_chapter_7

Ostapirat, Weera. 2013. Austro-Tai revisited. Paper presented at the 23rd Annual Meeting of the Southeast Asian Linguistic Society, Chulalongkorn University, 29 May.

Oxenham, Marc F. *et al.* 2015. Emergence and diversification of the Neolithic in southern Vietnam: Insights from coastal Rach Nui. *Journal of Island and Coastal Archaeology* 10(3): 309–338. doi:10.1080/15564894.2014.980473

Pal, J. N. 1990. The early farming culture of northern India. *Bulletin of Deccan College Post-Graduate and Research Institute* 49: 297–304.

Parsons, B. J., Newbury, H. J. & Jackson, M. T. 1999. The genetic structure and conservation of aus, aman and boro rices from Bangladesh. *Genetic Resources and Crop Evolution* 46(6): 587–598. doi:10.1023/A:1008749532171

Pictet, Adolphe. 1859, 1863. *Les origines indo-européennes ou les aryas primitifs: Essai de paléontologie linguistique (première partie, seconde partie)*. Paris: Joël Cherbuliez.

Poloni, Estella Simone *et al.* 1997. Human genetic affinities for Y chromosome P49a,f/*Taq*I haploptypes show strong correspondence with linguistics. *American Journal of Human Genetics* 61: 1015–1035 (cf. the erratum published in 1998 in the *American Journal of Human Genetics* 62: 1267). doi:10.1086/301602

Poloni, Estella Simone, Ray, Nicolas, Schneider, Stefan & Langaney, André. 2000. Languages and genes: Modes of transmission observed through the analysis of male-specific and female-specific genes. In *Proceedings: Evolution of Language, 3rd International Conference 3–6 April 2000*, Jean-Louis Dessalles & Laleh Ghadakpour (eds), 185–186. Paris: École Nationale Supérieure des Télécommunications.

Pulleyblank, Edwin George. 1983. The Chinese and their neighbours in prehistoric and early historic times. In *The Origins of Chinese Civilization*, David N. Keightley (ed.), 411–466. Berkeley CA: University of California Press.

Ratliff, Martha. 2004. Vocabulary of environment and subsistence in the Hmong-Mien protolanguage. In *Hmong/Miao in Asia*, Nicholas Tapp, Jean Michaud, Christian Culas & Gary Yia Lee (eds), 147–165. Chiang Mai: Silkworm Books.

Ratliff, Martha. 2010. *Hmong-Mien Language History*. Canberra: Pacific Linguistics.

Reddy, B. Mohan *et al.* 2007. Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia. *Public Library of Science One* 2 (11): e1141. doi:10.1371/journal.pone.0001141.

Rootsi, Siiri *et al.* 2007. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe. *European Journal of Human Genetics* 15: 204–211. doi:10.1038/sj.ejhg.5201748

Ruddiman, William F., Guo, Zhengtang, Zhou, Xin, Wu, Hanbin & Yu, Yanyan. 2008. Early rice farming and anomalous methane trends. *Quaternary Science Reviews* 27: 1291–1295. doi:10.1016/j.quascirev.2008.03.007

Sahoo, Sanghamitra *et al.* 2006. A prehistory of Indian Y chromosomes: Evaluating demic diffusion scenarios. *Proceedings of the National Academy of Sciences of the United States of America* 103(4): 843–848. doi:10.1073/pnas.0507714103

Schlegel, Gustave. 1901. Review: *Elements of Siamese Grammar* by O. Frankfurter, PhD. Bangkok: Printed at the American Presbyterian Mission Press, Leipzig, Karl W. Hiersemann, 1900, *T'oung Pao* (Série II) II: 76–87.

Schlegel, Gustave. 1902. *Siamese Studies* [T'oung Pao, New Series II, Volume II, Supplement]. Leiden.

Sharma, Govardhan R., Misra, V. D., Mandal, D., Misra, B. B. & Pal, J. N. 1980. *Beginnings of Agriculture: From Hunting and Food Gathering to Domestication of Plants and Animals*. Allahabad: Abhinav Prakashan.

Sidwell, Paul & Blench, Roger. 2011. The Austroasiatic *Urheimat*: The southeastern riverine hypothesis. In *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*, Nick J. Enfield (ed.), 317–345. Canberra: Pacific Linguistics.

Silva, Fabio *et al.* 2015. Modelling the geographical origin of rice cultivation in Asia using the rice archaeological database. *PLOS One* 21.<doi:10.1371/journal.pone.0137024>.

Starosta, Stanley. 2005 (posthumous). Proto-East-Asian and the origin and dispersal of languages of East and Southeast Asia and the Pacific. In *The Peopling of East Asia: Putting Together Archaeology, Lingtuistics and Genetics*, Laurent Sagart, Roger Blench & Alicia Sanchez-Mazas (eds), 182–197. London: Routledge Curzon (Editorial misrepresentations that had crept into Starosta's posthumously published East Asian phylogeny were rectified that same year (cf. van Driem 2005)). doi:10.4324/9780203343685_chapter_11

Stevens, Chris J. *et al.* 2016. Between China and South Asia: A Middle Asian corridor of crop dispersal and agricultural innovation in the Bronze Age. *The Holocene* 26(10): 1541–1555. doi:10.1177/0959683616650268

Strecker, David. 1987. The Hmong-Mien languages. *Linguistics of the Tibeto-Burman Area* 10(2): 1–11.

Sweeney, Megan T. & McCouch, Susan R. 2007. The complex history of the domestication of rice. *Annals of Botany* 100: 951–957. doi:10.1093/aob/mcm128

Sweeney, Megan T. *et al.* 2007. Global dissemination of a single mutation conferring white pericarp in rice. *Public Library of Science – Genetics* 3(8): 1418–1424.

Vavilov, Nikolaj Ivanovič. 1926. Centry proisxoždenija kulturnyx rastenij (Studies on the origin of cultivated plants). *Trudy po Prikladnoj Botanike, Genetike i Selekcii* (*Bulletin of Applied Botany, Genetics and Plant Breeding*) 16(2): 1–248.

Wang, Chuan-Chao, Yan, Shi, Qin, Zhen-Dong, Lu, Yan & Ding, Qi-Liang. 2013. Late Neolithic expansion of ancient Chinese revealed by Y chromosome haplogroup O3a1c-002611. *Journal of Systematics and Evolution* 51(3): 280–286. doi:10.1111/j.1759-6831.2012.00244.x

Weber, Steve, Lehman, Heather, Barela, Timothy, Hawks, Sean & Harriman, David. 2010. Rice or millets: Early farming strategies in prehistoric central Thailand. *Archaeological and Anthropological Sciences* 2(2): 79–88. doi:10.1007/s12520-010-0030-3

Xiang, Kun *et al.* 2013. Identification of a Tibetan-specific mutation in the hypoxic gene EGLN1 and its contribution to high-altitude adaptation. *Molecular Biology and Evolution* 30(8): 1889–1898. doi:10.1093/molbev/mst090

Yan, Shi *et al.* 2011. An updated tree of Y-chromosome haplogroup O and revised phylogenetic positions of mutations P164 and PK4. *European Journal of Human Genetics* 19(9): 1013–1015. doi:10.1038/ejhg.2011.64

Yang, Xiaoyan *et al.* 2016. New radiocarbon evidence on early rice consumption and farming in South China. *The Holocene* 7.<doi:10.1177/0959683616678465>.

Yao, Xiaotian *et al.* 2017. Improved phylogenetic resolution for Y-chromosome haplogroup O2a1c-002611. *Scientific Reports* 7: 1146. doi:10.1038/s41598-017-01340-z

Zhao Zhijun. 2010. New data and new issues for the study of the origin of rice agriculture in China. *Archaeological and Anthropological Sciences* 2(2): 99–105. doi:10.1007/s12520-010-0028-x

Zhāng Hóngmíng. 1998. Chinese etyma for river. *Journal of Chinese Linguistics* 26(1): 1–47.