

Etyma, shouldered adzes and molecular variants

George van Driem

1. Promise and peril in prehistory

The prospect of reconstructing lost chapters of prehistory through interdisciplinary research thrills and tantalises. The epistemological limitations of combining insights from linguistics, archaeology, palaeobotany, paleoclimatology and genetics are occasionally pointed out. In the absence of writing, potsherds do not speak any particular language, nor does any molecular variant on the genome represent an absolute marker that its bearers in prehistory spoke a language belonging to some particular linguistic phylum. The links between language and palaeobotany and between linguistics and palaeoclimatology appear to be even more tenuous.

Yet scholars who have ventured gingerly to compare and correlate findings from disparate disciplines have come up with bold inferences and fascinating reconstructions of prehistoric pasts. Instead of a contrite sermon on the methodological dangers of constructing arguments on the basis of apparent correlations between heterogeneous sets of data which may in fact not be related at all, let us showcase instead a few interesting examples of where interdisciplinary research in one particular region of Asia has led to the construction of competing models of the past, some of which at times skirt the epistemological abyss.

2. A centre of gravity which now carries more weight

The Austroasiatic language family has over 100 million speakers today, predominantly thanks to just a few languages with numerous speakers such as Vietnamese, Khmer, Santhali and Khasi. Most of the over two hundred languages of the language family, however, are spoken by small and usually dwindling language communities strewn across a vast area stretching from eastern India across the Indo-Burmese borderlands to the Nicobars, the Malay peninsula, the Mekong delta as far east as Vietnam and adjacent parts of the Chinese province of Yúnnán.

Robert von Heine-Geldern (1917, 1928, 1932) was perhaps the first to argue that the Austroasiatic homeland may have lain on mainland Southeast

Asia, where most Mon-Khmer groups are still found. He argued that the Munda by contrast were racially South Asian indigenes who had been linguistically assimilated by incursive Austroasiatics from mainland Southeast Asia. He also interpreted the spread of what he called the *Schulterbeilkultur* in this light. Archaeology has progressed in the century since this model was first proposed. The shouldered adze or shouldered celt is now known to be just one artefact of a widely diffused technology known as the Hòabinhian industry, and the shouldered adze, taken just by itself, is no diagnostic for any particular Asian cultural complex.

Heine-Geldern's thinking on the spread of Austroasiatic into the Indian subcontinent, supported as it was by language and physical anthropology, may very well have been on the mark. His interpretation of the shouldered adze, however, may not have been correct. The Hòabinhian industry remains important to our understanding of lithic industries in Southeast Asia after the Last Glacial Maximum. Yet the shouldered adze is emblematic for the nature of archaeological argumentation more generally, which, though not inherently epistemologically flawed, generally yields only circumstantial evidence for reconstructing ethnolinguistic population prehistory.

More recently, Gérard Diffloth (2001, 2005) has looked toward the Bay of Bengal littoral and the area surrounding the mouths of the Ganges and the Brahmaputra for a possible Austroasiatic homeland. The most obvious argument in favour of Diffloth's homeland is that this area lies at the family's geographical centre of gravity, based on the distribution of modern Austroasiatic language communities. Recently, Diffloth has strengthened the centre of gravity argument, however, with evidence from linguistic phylogeny. After a decade of entertaining a tripartite division of the family, Diffloth's historical research has brought him back to a fundamental bifurcation of Austroasiatic into a western and an eastern branch. His drastically revised phylogeny of the eastern branch, now renamed Khasi-Aslian, no longer represents the 'Mon-Khmer' of yore. Mon-Khmer is but one branch within Khasi-Aslian.

Furthermore, Diffloth has adduced two additional sets of evidence from the realm of what Adolphe Pictet in 1859 called linguistic palaeontology. The flora and fauna reflected in the reconstructible Austroasiatic lexicon paints the picture of a tropical ecology inhabited by peacocks, tree monitors and binturongs. His reconstructions include the etyma *mra:k 'peacock *Pavo muticus*', *tərkuət 'tree monitor *Varanus nebulosus* or *bengalensis*', *tənyu:ʔ 'binturong *Arctitis binturong*', *(bən)jɔ:l ~ *j(ərm)o:l 'pangolin *Manis javanica*', *dəkan 'bamboo rat *Rhizomys sumatrensis*' (an Aus-

troasiatic root which has found its way into Malay as a loan), *kaciaŋ 'the Asian elephant *Elephas maximus*', *kiaə 'mountain goat *Capricornis sumatrensis*', *rəma:s 'rhinoceros *Dicerorhinus sumatrensis*' and *tənriak 'bubfalo *Bubalus bubalus*' (Diffloth 2005: 78).

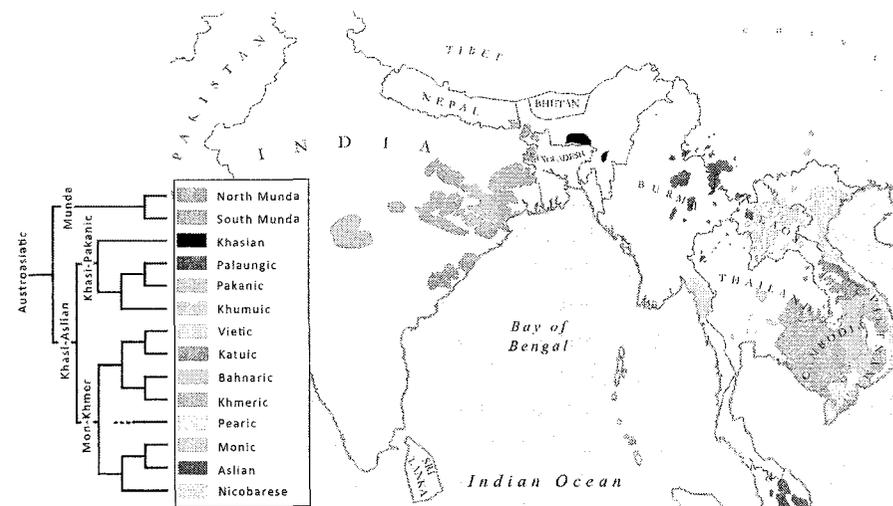


Figure 1. Gérard Diffloth's revised (2009) Austroasiatic phylogeny and the geographical distribution of Austroasiatic languages, adapted from Chaubey et al. (2010). The two trunks of the Austroasiatic phylum are Munda, in eastern, northeastern and central India, and Khasi-Aslian, which stretches from the Meghālaya in the northeast of the subcontinent to the Nicobars, the Malay peninsula and the Mekong delta in Southeast Asia. The precise phylogenetic position of Pearic within Mon-Khmer remains uncertain.

As I have argued elsewhere (van Driem 2012), the Holocene ranges of species for which Austroasiatic has reconstructible etyma covers a vaster area than would be suggested *prima facie* by the ranges of species named by Diffloth. In fact, some of these faunal habitats are compatible with both Robert von Heine-Geldern's Austroasiatic homeland and Diffloth's Austroasiatic homeland. For the Austroasiatic etymon *tərkuət, for example, Diffloth offers both the Bengal tree monitor *Varanus bengalensis* and the clouded tree monitor *Varanus nebulosus* as possible referents. By the same token, the ancient Austroasiatics may have used an etymon such as *rəma:s just as readily to refer to the Indian rhinoceros *Rhinoceros unicornis* or

Javan rhinoceros *Rhinoceros sondaicus* as to the Sumatran rhinoceros *Dicerorhinus sumatrensis*.

When Diffloth's reconstructible Austroasiatic fauna etyma are reassessed in this light, the area of overlap of the possible species denoted by the reconstructible fauna etyma more specifically turns out to cover north-eastern India, the Indo-Burmese borderlands, Burma and Yúnnán. The implication is that the Austroasiatic homeland may not have lain in India proper, but more towards Southeast Asia. On the basis of the reconstructible faunal etyma, the contiguous area comprising upper Burma, the Chinese province of Yúnnán and adjacent portions of northeastern India, northern Thailand and western Laos looks likelier to represent a possible Austroasiatic homeland than either the areas straddling the lower course of the Mekong and surrounding the Mekong delta or the areas straddling the lower course of the Brahmaputra and surrounding the Brahmaputra delta (van Driem 2012).

The second set of linguistic palaeontological data is the reconstructible Austroasiatic vocabulary which robustly reflects terminology pertaining to early rice agriculture, viz. *(kə)ʔa:ʔ 'rice plant', *rəŋko:ʔ 'rice grain', *cəŋka:m 'rice outer husk', *kəndək 'rice inner husk', *pʰe:ʔ 'rice bran', *təmpal 'mortar', *jənrəʔ 'pestle', *jəmpjəʔ 'winnowing tray', *gu:m 'to winnow', *jərmuəl 'dibbling stick' and *kəntu:ʔ 'rice complement', i.e. accompanying cooked food other than rice (Diffloth 2005). Rice agriculture must have been an early Austroasiatic technology if such terms are reconstructible for all major branches of the Austroasiatic family.¹

Moreover, Ferlus (2010) identifies Proto-Austroasiatic *C.rac as the primordial word for 'rice' and the ultimate source of ancient borrowings to other linguistic phyla, i.e. Proto-Austronesian *beRas, Tibeto-Burman languages, e.g. Tibetan *hbras* and Old Chinese 糲 *bərats, Indo-European, e.g. Sanskrit *vr̥hi*, Hindi *biryānī* (with a suffix), Latin *oryza*, and Dravidian e.g. Tamil *vir̥gi* and *arici*. Ferlus argues that originally Proto-Austroasiatic *C.rac need not have specifically denoted rice because the form is an ancient deverbative of *rac 'to collect by tearing off the grain along the stem with the hand'. Yet such linguistic evidence is best evaluated in light of where rice may have first been domesticated.

3. The absence of evidence is the absence of evidence

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). Nikolai Ivanovič Vavilov (1926) later argued against a Chinese origin for rice and contended instead that the origin of Asian rice lay in India, whence the crop had spread to China and Japan. The controversy continued until molecular genetics was applied to the study of rice. I have discussed the results of recent research in rice genetics and their implications for ethno-linguistic prehistory elsewhere (van Driem 2011). The story can be presented in a nutshell.

Molecular genetic studies have corroborated a view, which is not at all new, that Asian rice appears to have been domesticated twice, if not three times. An eastern domestication of the perennial swamp species *Oryza rufipogon* led to the development of the *japonica* cultivar of *Oryza sativa*. This domestication process involved the proliferation of the mutation *sh4*, which led to the partial development of the abscission zone where the mature grain detaches from the pedicle. 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*).

Further west, the domestication of *Oryza nivara* led to the development of the *indica* cultivar of *Oryza sativa*. However, *nivara* rice can be described as just an ecotype under a single *Oryza rufipogon* species complex, which encompasses both the annual self-pollinating *Oryza nivara*, adapted to disturbed shallow-water environments, and the perennial *Oryza rufipogon* in the strict sense, best adapted to stable deeper-water environments. Both wild species grow sympatrically and naturally hybridise with each other as well as with cultivated rice. Crucially, the domestication of *indica* rice was facilitated by the introduction of the domesticated traits *sh4*, *rc* and *Prog1* into the *nivara* gene pool through introgressive hybridisation involving backcrossing with the previously domesticated *japonica* cultivar.

To simplify matters, the famous *javanica* rice turns out to be just a tropical variety of *japonica* (Figure 2). To make matters more interesting, however, some long-grained aromatic varieties, such as the famous *bāsmati* rice of the Indian subcontinent, likewise derive from the *japonica* domestication event, not from *indica*, whereas Thai jasmine rice is actually an *indica* variety with the fragrant betaine aldehyde dehydrogenase gene *BADH2* introduced by introgression (Garriss et al. 2005, Kovach et al. 2009, Parsons et

al. 1999). By contrast, upland or dry-cultivated rice may represent a distinct, third domestication process. Upland rice is still widely cultivated in the sub-Himalayan hill tracts, where it is known in Nepali as *ghaiyā* and in Dzongkha as *kambjā* ‘dry paddy’. Upland rice has sometimes come to be referred to in the rice genetics literature as ‘aus’, a transmogrified version of Bengali *āus*.

Within the *rufipogon* species complex, the calculated time depth for the divergence of the *rufipogon* ancestor of *japonica* and the *nivara* ancestor of *indica* is in the order of magnitude of 100,000 years. Genome-wide analysis of single nucleotide polymorphisms yields a phylogeny that may suggest an eastern origin of Asian rice within the greater Himalayan region (Puruggannan 2010). The date, roughly as old as anatomically modern humans and so long before the invention of agriculture, has also been corroborated by the dating of the divergence of chloroplast DNA (Tang et al. 2004). Whether or not *ghaiyā* represents an independent domestication event, upland rice is genetically more closely affiliated to *indica* than to *japonica* rice. Figure 3 shows likely geographical ranges for the three Asian rice domestications based on the geographical distribution of genetic markers in the wild precursor *Oryza rufipogon*.

Recent phylogenetic evidence has been adduced, purportedly in support of a single domestication of Asian rice some time between 13,500 and 8,200 years ago (Molina et al. 2011). In fact, the presented findings do not exclude the possibility that *indica* and *japonica* rice originated from different *Oryza rufipogon* gene pools. Moreover, the phylogenetic evidence indicates that the wild *rufipogon* population in the region encompassing the Indian Subcontinent and mainland Southeast Asia was ancestral to all domesticated Asian rice, unless the ancestor was some now extinct *rufipogon* population. The evidence adduced by Molina et al. (2011) against multiple domestications of Asian rice is therefore neither complete nor conclusive.

Interpretations invariably hinge upon those key genetic adaptations to agricultural environments that define domestication and so represent salient traits in cultivated Asian rice today. These adaptations appear to have originated in *japonica* rice and been subsequently introduced to *indica* rice by introgression, but it cannot be excluded that wild rice may long have been collected in the vast region encompassing the eastern Himalayas, Yúnnán, northeastern India, the Indo-Burmese borderlands and Burma over long stretches of time before selection for domestication traits began. Moreover, ancient domestication genes native to the *indica* lineage could have gone extinct after the later introduction of *japonica* traits. The *indica* chloroplast

genome diverges enough from *nivara* to suggest that ancient *indica* lineages may have existed and subsequently gone extinct with the preferential westward introduction of *japonica* traits (Takahashi et al. 2008).

Hmong-Mien is a language family of East and Southeast Asia, for which historical linguists have been unable to adduce compelling comparative evidence that would unequivocally link these languages to any other Asian linguistic phylum. I identified the early Hmong-Mien as the original domesticators of *japonica* rice and the ancient Austroasiatics as the original domesticators of *indica* rice (van Driem 2011, 2012). The possibility of contact between the linguistic ancestors of the Hmong-Mien and the early Austroasiatics as well as the diffusion of rice agriculture from the ancient Austroasiatics to the early Hmong-Mien, or vice versa, are a possible implication of this reconstructed prehistoric scenario. The modern distribution of Hmong-Mien language communities is shown in Figure 4, but the ancient Hmong-Mien probably lived along the entire southern Yangtze river basin, where *japonica* rice may first have been domesticated. This model is provocative because the Hmong-Mien have in the course of known history been subject to, or subsumed within, more powerful Sinitic polities, and in recent centuries Hmong-Mien languages have often tended to borrow from Sinitic languages rather than the other way around.

The Hmong-Mien have a reconstructible lexicon of native rice agriculture terms almost as rich as that of the Austroasiatics. Forms taken from Martha Ratliff’s handbook on Hmong-Mien historical phonology² include Proto-Hmong-Mien *hnrəŋH ‘cooked rice’, *hnɔn ‘rice head, head of grain’, *TuX ‘husk/pound rice’, *mblut ‘glutinous’, *ljij ‘paddy field’, *ljim ‘sickle’, *ŋkjuəX ‘rice cake’, *tsjɛŋH ‘rice steamer’, *mbləu ‘rice plant, paddy’, *mphigk ‘chaff’, Proto-Hmongic *ntsuw^C ‘husked rice’, *S-phjæ^C ‘chaff’, *mbljæ^C ‘have food with rice’, *ʔrin^A ‘dry (rice) in sun’, *tshɛŋ^B ‘husked rice or millet’, *ljɛŋ^A ‘rice measure’, *ʔjɛŋ ‘seedling’ and Proto-Mienic *hrau^A ‘rice measure’, *ʔjaŋ ‘seedling’ and *hmei^B ‘husked rice’. Several of the Hmong-Mien rice etyma entered Sinitic as early loans, and the Proto-Mienic etymon *hmei^B ‘husked rice’ found its way into both Sinitic and Brahmaputran as well as into Austronesian, where it became the etymon now reconstructible as *Semay ‘cooked rice’.

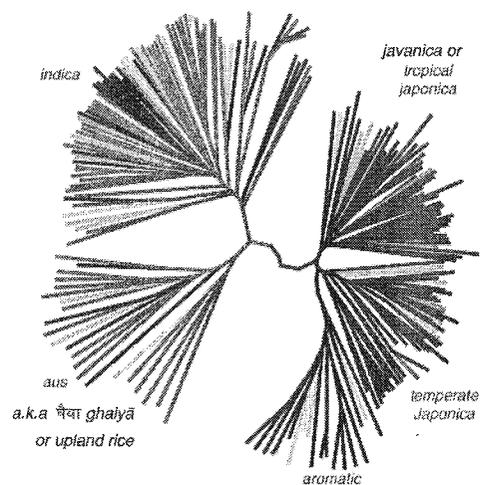


Figure 2. An unrooted phylogenetic tree of the diversity of *Oryza sativa* based on 169 nuclear simple sequence repeats (SSRs) and two chloroplast markers in 234 cultivars of *Oryza sativa*, adapted from Garris et al. (2005), Kovach et al. (2007).

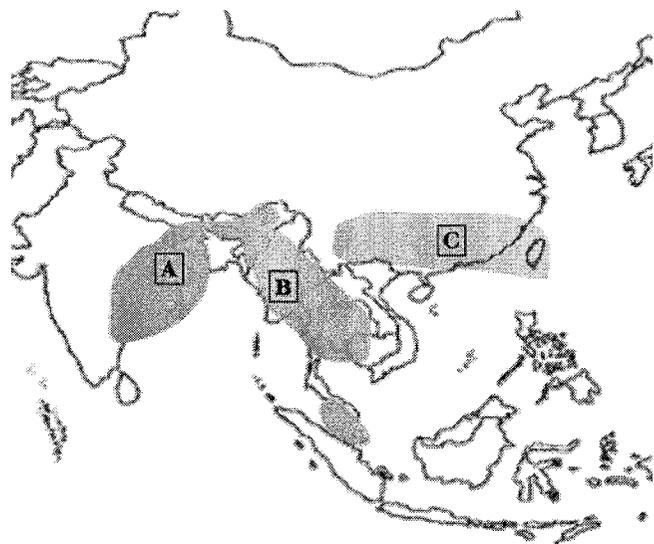


Figure 3. The geographical ranges for the possible domestication of (A) *ghaiyā* or upland rice, (B) wet *indica* rice and (C) the *japonica* cultivar, based on the geographical distribution of genetic markers in the wild precursor *Oryza rufipogon* (adapted from Londo et al. 2006).

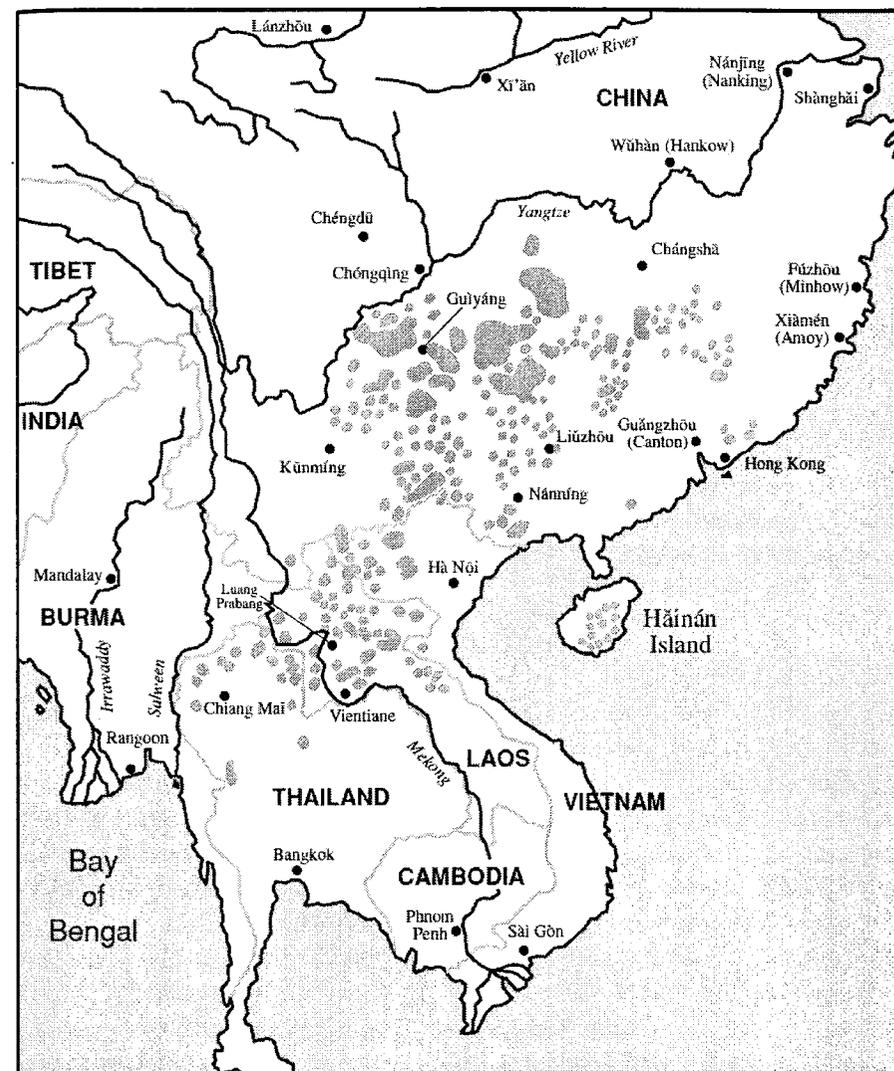


Figure 4. The geographical distribution of modern Hmong-Mien language communities (from van Driem 2001: 319).

All this brings us to the archaeology of rice and an associated empirical quandary. 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. It is hardly surprising, however, that archaeologists have not found the remains of early rice agriculture in those

places where they have not yet looked. A vast swathe of Asia covering the areas identified by rice geneticists as harbouring the possible sites of domestication of *indica* and upland rice has not been subjected to systematic archaeological and palaeobotanical investigation, particularly the areas surrounding the nexus where ranges A and B overlap (Figure 3).

Fuller concedes the possibility ‘perhaps that Northeast India will eventually prove to be a rice domestication area’, but he would like to see ‘more of an archaeological smoking gun’ (2011: 82). Smoking guns are unlikely to turn up as long as Burma, northeastern India, the eastern Himalayas and much of Yúnnán remain archaeologically uninvestigated. Aside from this vast geographical gap in the archaeology, there is the separate empirical issue of the archaeological recoverability of rice agriculture sites. Traces of ancient farming communities have been better preserved in the hill tracts surrounding the Brahmaputran flood plains than on the fertile fields themselves. In East Asia too, 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 have developed on the flood plains themselves. The remains of early rice cultivating cultural assemblages may lie buried forever in the silty sediments of the sinuous lower Brahmaputran basin. Alternatively, the palaeobotanical evidence for the earliest domestication of rice may have been washed out by the Brahmaputra long ago and now lie submerged in the depths of the Bay of Bengal.

By contrast, the absence of evidence for early rice agriculture of great antiquity in meridional mainland Southeast Asia, despite the relatively well researched archaeology of the region, presently embarrasses those who have lately taken to espousing a homeland theory for Austroasiatic along the lower course of the Mekong and around the Mekong delta. By contrast, the fact that the archaeology of northeastern India, the Indo-Burmese borderlands and the northern Bay of Bengal littoral is poorly understood, because it is virtually unresearched, does not compromise Diffloth’s proposed homeland. As the old saw has it, the absence of evidence is not the evidence of absence. Political, cultural, geographical and logistic factors appear to have impeded intensive archaeological research in a vast swathe of territory extending from the lower Brahmaputran basin to the Tenasserim. The same factors have limited population genetic and rice genetic research in this region as well.

4. Languages are not people

Can the human population genetics of modern language communities be plausibly interpreted in light of the hypothesis that early Hmong-Mien were the first domesticators of *japonica* rice and ancient Austroasiatics the domesticators of *indica* rice, or vice versa? If, on the other hand, Asian rice were to be demonstrated to have been domesticated just once from a wild *rufipogon* population in the area which lies between India and Southeast Asia, then who were its domesticators? Such conjectures beg the question as to whether there need be any such correlation between language, people and cultigens at all.

There need not be. Nonetheless a tendency has repeatedly been observed, first recognised in the pioneering studies of Poloni et al. (1997, 2000), that a correlation often obtains between the most frequent Y chromosomal haplogroups of a community and the language which the people happen to speak. A correlation between a community’s language and that community’s most prevalent paternal ancestries is found often enough that I called this correlation the Father Tongue hypothesis (van Driem 2002).

There are a number of reasons why we might expect this outcome. Initial human colonisation of any 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. Theoretically, tribes of Amazons could have spread in a similar fashion. If so, then the tell-tale correspondences between mitochondrial lineages and the distribution of linguistic phyla would presumably have been detected by now, but the correlations between maternal lineages and linguistic phylogeography hitherto discerned 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 of these dispersals were predominantly later male-biased intrusions.

If we infer that a mother teaching her children their father’s tongue has been a recurrent, ubiquitous and prevalent pattern throughout linguistic history, then some of the mechanisms of language change over time are likely to be inherent to the dynamics of this pathway of transmission. Such correlations are observed worldwide. The correlation of Niger-Congo languages with Y chromosomal haplogroups is a striking example (Wood et al. 2005). Likewise, the martial and male-biased historical spread of Hà Chinese

during the sinification of southern China, recounted in painstaking detail in the Chinese chronicles, is clearly reflected in the genetic evidence (Wen et al. 2004). A recent common ancestry between native Americans and indigenous Altaians is also based preponderantly on the shared Y chromosomal heritage and is not quite as well reflected in the mitochondrial lineages (Dulik et al. 2012).

Whilst father tongues may predominate globally, mother tongues certainly do exist in the sense that there are areas on the planet where the linguistic affinity of a community corresponds more closely to the maternally transmitted mitochondrial lineage which the speakers share with other linguistically related communities. In this sense, in the north of today's Pakistan, the Balti speak a Tibetic mother tongue but profess a paternal religion that was first propagated in this area as early in the 8th century by men who came from the Near East. The most prevalent mitochondrial DNA lineages amongst the Baltis are shared with other Tibetan communities, whereas the prevalent Y chromosomal haplogroups probably entered Baltistan with the introduction of Islam (Zerjal et al. 1997, Quintana-Murci et al. 2001, Qamar et al. 2002).

At the same time, a jarring disconnect is sometimes seen between the occurrence of a highly salient genetic marker and the linguistic affinity of a community's language. Hungarians lack the TatC deletion defining the Y chromosomal haplogroup N1c,³ despite the sheer prevalence of this marker amongst all other Uralic language communities (Li et al. 1999). So, it deserves to be repeated that the linguistic ancestors of a language community were not necessarily the same people as the biological ancestors of that community. In fact, some of them could not have been the same people.

It also merits repeating that the time depth accessible to population geneticists studying polymorphisms on the genome is vastly greater than the reach of the linguistically reconstructible past. The wave of anatomically modern humans who introduced the proto-languages that were later to give rise to today's Asian linguistic phyla and language isolates can be dated to between 25,000 to 38,000 years ago (Rasmussen et al. 2011), and the antiquity of Y chromosomal haplogroups such as O1 or O2 has been calculated to be greater than 10,000 years (Yan et al. 2011). Historical linguists, on the other hand, generally estimate the linguistically reconstructible past to be shallower than 10,000 years. This temporal gap must temper and inform all speculations regarding correlations between linguistic and genetic affinity.

With such caveats in place, how can we address the question formulated at the beginning of this section? On the 28th of June 2006, at a symposium held at l'École Française d'Extrême-Orient at Siem Reap, I identified the Y chromosomal haplogroup O2a (M95) as the marker for the spread of Austroasiatic on the basis of the then available genetic data (later published in van Driem 2007). This view has been corroborated by subsequent genetic studies, e.g. Kumar et al. (2007), Chaubey et al. (2010). In the latter article, we concluded that Austroasiatic speakers in India today are derived from a dispersal from Southeast Asia, followed by extensive sex-specific admixture with local populations indigenous to the Subcontinent.

The autosomal data also reflect the distinction between two components in the genome, one represented by the predominantly indigenous maternal lineages and the other by the intrusive paternal O2a lineage that correlates with the linguistic affinity of the Austroasiatic language communities in the Indian subcontinent. These findings go well beyond Robert von Heine-Geldern's model of a Southeast Asian homeland and envisage a father tongue spread of Austroasiatic, borne to the Indian subcontinent by predominantly male speakers from mainland Southeast Asia, but also involving a complex sociolinguistic prehistory of bidirectional gene flow across the Bay of Bengal (Chaubey et al. 2010). In many parts of the world, the mitochondrial DNA lineages often appear preponderantly to reflect older resident maternal lineages.

The argument for the Father Tongue interpretation of the spread of major linguistic phyla in eastern Eurasia, such as Austroasiatic, is therefore not based solely on the frequencies of particular Y chromosomal haplogroups. The Father Tongue hypothesis is originally based on the differential correlation of Y chromosomal and mitochondrial lineages with the modern geographical distribution of language communities, i.e. the presence or absence of a strong correlation between linguistic affinity and genetic markers in the non-recombinant portions of the genome. As one might expect, a distinct provenance for the maternal and paternal lineages appears to be reflected by studies of autosomal markers as well (Chaubey et al. 2010). More importantly, a rooted topology of the Y chromosomal tree in its entirety and of the Y chromosomal haplogroup O in particular is central to the reconstruction of linguistic population prehistory in eastern Eurasia, operating on the assumption of the veracity of the Father Tongue hypothesis.

The available genetic data also enabled us to identify a correlation of the Y chromosomal haplogroup O3a3b (M7) with the spread of Hmong-Mien, whilst our genetic samplings throughout the Himalayan region had estab-

lished a correlation between Tibeto-Burman and the paternal lineage O3a3c (M134) (Parkin et al. 2006, 2007, Kraaijenbrink et al. 2007a, 2007b, 2009, van Driem 2011). The Y chromosomal haplogroup O is becoming ever more minutely mapped, and most recently the phylogenetic positions of mutations P164 and PK4 within the haplogroup have been revised (Yan et al. 2011). Yet the antiquity calculated for many of these mutations is generally greater than the time depth that most historical linguists are willing to ascribe to the major language phyla.

Let us venture into the twilight beyond the linguistically reconstructible past to a time just after the Last Glacial Maximum, when the Y chromosomal haplogroup O (M175) had split up into the subclades O1 (M119), O2 (M268) and O3 (M122). Based on what is known about linguistic phylogeny and about the geographical distribution of modern linguistic communities today, the three subclades can putatively be assigned to three geographical loci along an east-west axis. For the sake of argument and schematic representation, and without any claim to geographical precision or veracity, I shall assign the haplogroup O1 (M119) to the drainage of the Pearl River and its tributaries in what today is the Chinese province of Guangdong. I shall situate haplogroup O2 (M268) in southern Yunnan and O3 (M122) to the area where today's northeastern India, southeastern Tibet and northern Burma adjoin.

Since we have associated O2a (M95), which is a derivative clade of haplogroup O2 (M268), with the Austroasiatic language phylum, we might conjecture that Asian rice, perhaps both *japonica* and *indica* rice, was first domesticated roughly in the general area hypothetically imputed to O2 (M268) here. Whilst the bearers of the O2a (M95) haplogroup became the *Stammväter* of the Austroasiatics, the other derivative paternal subclade O2b (M176) spread eastward, where they introduced rice agriculture to the areas south of the Yangtze. Though the bearers of the O2b (M176) haplogroup continued to sow seed as they continued to move ever further eastward, they left little or no linguistic traces, except maybe an Austroasiatic name for the Yangtze river, as proposed by Pulleyblank (1993), reflected as the toponym borrowed by Old Chinese as 江 *k'roŋ (*jiāng*).

Meanwhile, back in southern Yunnan, the early Austroasiatics spread from this locus initially to the Salween drainage in northeastern Burma and to the area that today is northern Thailand and western Laos. In time, the Austroasiatics would spread as far as the Mekong delta, the Malay peninsula, the Nicobars and later even into eastern India, where they would intro-

duce both their language and their paternal lineage to indigenous peoples of the Subcontinent.

At the locus putatively assigned to the haplogroup O3 (M122), the bearers of this marker gave rise to the paternal lineages O3a3c (M134) and O3a3b (M7). Whilst the bearers of the polymorphism O3a3c (M134) stayed behind in the area comprising northeastern India, southeastern Tibet and northern Burma, the bearers of the O3a3b (M7) paternal 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 early 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 O2a (M95) in today's Hmong-Mien and of haplogroup O3a3b (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'. More precisely, the incidence of haplogroup O3a3b (M7) in Austroasiatic language communities of Southeast Asia appears to indicate a significant Hmong-Mien paternal contribution to the early Austroasiatic populations whose descendants settled in Southeast Asia, whereas the incidence of haplogroup O3a3b (M7) in Austroasiatic communities of the Indian subcontinent is undetectably low. The incidence of haplogroup O2a amongst the Hmong-Mien appears to indicate a slightly more modest Austroasiatic paternal contribution to Hmong-Mien populations than vice versa.

As the Hmong-Mien moved eastward, the bearers of haplogroup O2b (M176) likewise continued to move east. Even further east, the O1 (M119) paternal lineage gave rise to the O1a (M119) subclade, which moved from the Pearl River drainage eastward to the Mǐn river drainage in the hill tracts of Fújiàn province and across the strait to Formosa, which consequently became the *Urheimat* of the Austronesians. Back west in the easternmost spurs of the Himalayas, the bearers of Y chromosomal haplogroup O3a3c (M134) expanded eastward into Sichuān and Yunnan, north and northwest across the Tibetan plateau as well as westward into the Himalayas and southward into the Indo-Burmese borderlands. In the west and south, the early Tibeto-Burmans encountered Austroasiatics, who had preceded them.

The relative frequencies of the Y chromosomal haplogroup O2a (M95) in various Tibeto-Burman populations of the Indian subcontinent (Sahoo et

al. 2006, Reddy et al. 2007) suggest that a subset of the paternal ancestors of particular Tibeto-Burman populations in northeastern India, e.g. certain Bodo-Koch communities, may originally have been Austroasiatic speakers who married into Tibeto-Burman communities or were linguistically assimilated by ancient Tibeto-Burmans. At the same time, median-joining network analyses of haplogroup O2a (M95) microsatellites have suggested a division in the Indian subcontinent between Tibeto-Burmans vs. Austroasiatic and Dravidian language communities. Austroasiatics and Dravidians show greater Y chromosomal microsatellite diversification than Tibeto-Burman language communities, and the highest frequency of the O2a haplogroup is found in tribal populations in Orissa, Chattisgarh and Jharkhand (Sengupta *et al.* 2006).

We must bear in mind that Y haplogroups are subject to selection and that frequencies change over time. As stressed above, haplotype frequencies by themselves are not a sufficient criterion. A rooted topology of the Y chromosomal tree and its subsidiary clades provides key evidence. Moreover, the ethnolinguistic significance of paternal lineages becomes even more manifest when other portions of the genome are scoured for correlations with linguistic phylogeography. At the same time, our understanding of what constitutes neutral diversity has been tempered by mathematical modelling. Simulations have shown that a normally low-frequency allele could surf on a demic wave of advance and so attain high frequency across a vast area. Gene surfing during a spatial expansion is likely to result in distinct geographical sectors of low genetic diversity separated by sharp allele frequency gradients.

The result of recurrent bottleneck effects during range expansion into newly colonised territories can mimic complex phylogeographical patterns of adaptation and segregation into clades in post-glacial niche refugia. Likewise, the massive introgression of resident genes into the incursive population can also be misinterpreted as the result of a selective process (Excoffier and Ray 2008, Excoffier *et al.* 2009). Surfing on the crest of a demic wave of expansion confers a selective advantage when compared to alleles left behind in the core area (Klopfstein et al. 2006, Moreau et al. 2011). Both the dynamics of sex-biased dispersals as well as the process of the sexually asymmetrical introgression of resident alleles into incursive populations can be modelled in terms of hybridisation during range expansions (Petit and Excoffier 2009, Currat and Excoffier 2011).

An observed state of affairs for which a particular model of population prehistory has been advanced may in many cases very well be either the

result of demography or of selection on genome diversity (Fagundes et al. 2007). However, we must keep in mind that a scenario that has been computed to be the statistically more likely scenario may not necessarily correspond to the prehistorical reality. Though presumably paternal lineages may often preferentially enjoy the benefits of surfing, incursive Y chromosomal lineages can go entirely extinct, as the linguistic evidence⁴ would suggest may very well have happened with the Y chromosomal haplogroup N1c in Hungary.

We must also not lose sight of the fact that these speculations are based on correlations between language and Y chromosomal haplogroups and that these too are interpreted in the light of the assumed veracity of the Father Tongue hypothesis over a vast stretch of time. This assumption may not hold true for all times in the past. Furthermore, correlations may be due to different kinds of circumstances other than causation or direct relationship. So, whilst we are free cautiously to develop arguments which buttress a speculative model of ethnolinguistic prehistory, such as the one outlined here, we must not lose sight of the essential distinction between the facts and our assumptions and inferences as well as the precise nature and limitations of the empirical basis for our speculations.

5. The proof of the pudding depends on who is doing the eating

The Austric theory, introduced by the Austrian Jesuit Wilhelm Schmidt in 1906, posits a deep phylogenetic relationship between the Austronesian and the Austroasiatic language phyla. The theory with its potential implications for prehistory has continued to engage scholars, although its empirical basis has always remained decidedly scanty, e.g. Benedict (1991), Bengtson and Blažek (2000), Blust (1996), Diffloth (1994), van Driem (1999), von Hevesy (1930), Reid (1994).

On the 11th of January 2003 at the 9th International Conference on Austronesian Linguistics at Australian National University in Canberra, I presented a talk entitled 'The Austric Problem: Issues, Solutions, Ramifications'. This concise account of the history of thinking about the Austric theory included an assessment of the various types of evidence mustered in support of this theory by its various proponents and the criticisms levelled against the theory by its assailants. I strove to keep the overview as impartial and as comprehensive as possible.

At the conclusion of my talk, people approached me to air their views. Two scholars congratulated me enthusiastically for utterly demolishing the Austric theory, and one even said that he now expected never to see anything about this theory in print ever again. Moments later, two other members of the audience thanked me wholeheartedly for establishing the verisimilitude of Austric (one said ‘proving Austric’) and definitively relegating the criticism of skeptics and detractors to the dustbin of history.

The same set of linguistic evidence and the same discussion of the history of thinking on this hypothetical linguistic relationship had convinced some that Austric had been conclusively demonstrated, whilst others felt that the theory had now been decisively discarded. Although these reactions reassured me that I might have succeeded in making my presentation of the Austric case as impartial as possible, I was left wondering to what extent proof could be a matter of taste. How selectively do we filter out what does not corroborate the views that we already hold? Personal inclinations and matters of taste certainly play a role in shaping the history of science, and paradigm shifts, as Kuhn (1969) and Murray (1994) have shown, are also social phenomena. However, proof in science cannot be just a matter of taste. Rather, it can often be observed that some scholars are highly selective about how they interpret evidence.

Confronted with the overwhelming growing body of evidence in support of the Father Tongue hypothesis, Forster and Renfrew (2011: 1391) impute the spread of language families to ‘emigrating agriculturalists’ who ‘took local wives’. This interpretation is a transparent attempt to succour Bellwood and Renfrew’s embattled First Farmers hypothesis, which seeks to ascribe the founding dispersals of language families to the spread of agriculture (Bellwood and Renfrew 2002). At the same time, in order to buttress Renfrew’s widely doubted hypothesis of an Indo-European homeland in Asia Minor, Forster and Renfrew also propose a correlation of Indo-European with the Y chromosomal haplogroup J2a. In fact, it remains moot whether any part of Y chromosomal phylogeography correlates well with the spread of the Neolithic horizon.

Not every population movement led to the spread of a language phylum, and population movements are not uniform in nature. Whether during the exodus of anatomically modern humans out of Africa or at the shallow time depth of the colonisation of Oceania by Austronesian populations, the colonisation of previously uninhabited lands invariably involved both sexes and the introduction of a language phylum. During the Neolithic horizon, the spread of farming was necessarily a sedentary and incremental process,

which likewise must mostly have involved both sexes. Early farmers might only have been able to spread their language at times of great surplus and concomitant population growth, perhaps sometimes involving the establishment of agricultural colonies elsewhere. By contrast, the modern ethnolinguistic composition of Asian populations must be understood, at least in part, as having resulted from male-biased linguistic intrusions, whether motivated by conquest, land grab or the urge to seek out new habitats.

In my argument against the premises and the reasoning behind the hypothesis of the founding dispersals of language phyla having been mediated by the spread of farming, I proposed the telic and more complex Centripetal Migration theory (van Driem 2007). I shall not repeat that exposition here, but, with reference to Forster and Renfrew’s wilful interpretation of the Y chromosomal haplogroup J2, I shall reiterate that, in the context of the Indian subcontinent, ‘the J2 haplogroup... appears to emanate from the Arabian Peninsula and, unlike haplogroups N and R1a, attains no high frequency in Ceylon’ and ‘probably reflects the historically attested male-borne eastward spread of Islam’, whereas Y chromosomal haplogroups of the R subclades spread to the Subcontinent ‘from the northwest along with Indo-Aryan language across northern India and to Ceylon’ (van Driem 2007: 5). The spread of various Y chromosomal R subclades is likely to be linked to the dispersal of Indo-European from an original homeland in the Pontic-Caspian steppe, whilst the current geographical distribution of the Y chromosomal lineage L provides the likeliest candidate for a vestige of an earlier patrilineal dispersal of Elamo-Dravidian emanating from a region which encompassed the Bactria and Margiana of later prehistory.

Notes

1. Ferlus (1996, 2010), however, contests the reconstructibility of *(kə)ḃa:ʔ ‘rice plant’ and *pʰe:ʔ ‘rice bran’ and contends that the etymon *rəŋko:ʔ originally reflects the word for taro, arguing that this semantic shift occurred because rice was originally a wild grass which grew in taro fields. However, these hypotheses are assessed by Diffloth (2011).
2. The linguistic and other evidence, outlined here and discussed in detail elsewhere (van Driem 2011, 2012), was first presented at the 4th International Conference on Austroasiatic Linguistics (ICAAL 4) at Mahidol University on 29 October 2009. I should like to thank Martha Ratliff for sharing the index portion of her then still forthcoming handbook on Hmong-Mien historical pho-

nology. Meanwhile, Ratliff's handbook (2010) has now been published by Pacific Linguistics in Canberra.

3. The 2008 Y Chromosome Consortium haplogroup labels are used here.
4. The presence of the Hungarian language in the region that was once Pannonia represents incontrovertible linguistic evidence of the advent of Uralic linguistic ancestors, a fact which is historically attested at any rate, but the hypothetical correlation of the Y chromosomal haplogroup N1c with the Uralic linguistic phylum, of course, remains entirely conjectural.

References

- Benedict, Paul King
1991 Austric: an "extinct" proto-language. In *Austroasiatic Languages: Essays in honour of H.L. Shorto*, Jeremy H.C.S. Davidson (ed.), 7–11. London: School of Oriental and African Studies, University of London.
- Bellwood, Peter, and Colin Renfrew (eds.)
2002 *Examining the Farming/Language Dispersal Hypothesis*. Cambridge: McDonald Institute for Archaeological Research.
- Bengtson, John D., and Václav Blažek
2000 Lexical parallel between Ainu and Austric, and their implications. *Archiv Orientální* 68: 237–258.
- 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. (Transactions of the American Philosophical Society, Vol. 86, Pt. 5) Philadelphia: American Philosophical Society.
- Cai Xiaoyun, Zhendong Qin, Bo Wen, Shuhua Xu, Yi Wang, Yan Lu, Lanhai Wei, Chuanchao Wang, Shilin Li, Xingqiu Huang, Huí Lǐ, and the Geographic Consortium
2011 Human migration through bottlenecks from Southeast Asia into East Asia during Last Glacial Maximum 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.
- Chaubey, Gyaneshwer, Mait Metspalu, Ying Choi, Reedik Mägi, Irene Gallego Romero, Siiri Rootsi, Pedro Soares, Mannis van Oven, Doron M. Behar, Siiri Rootsi, Georgi Hudjashov, Chandana Basu Mallick, Monika Karmin, Mari Nelis, Jüri Parik, Alla Goverdhana Reddy, Ene Metspalu, George van Driem, Yali Xue, Chris Tyler-Smith, Kumarasamy Thangaraj, Lalji Singh, Maito Remm, Martin B. Richards, Marta Mirazon Lahr, Manfred Kayser, Richard Villems, and Toomas Kivisild
2010 Population genetic structure in Indian Austroasiatic speakers: The role of landscape barriers and sex-specific admixture. *Molecular Biology and Evolution*, doi: 10.1093/molbev/msq28.
- Currat, Mathias, and Laurent Excoffier
2011 Strong reproductive isolation between humans and Neanderthals inferred from observed patterns of introgression. *Proceedings of the National Academy of Sciences* 108 (37): 15129–15134.
- Diffloth, Gérard
1994 The lexical evidence for Austric, so far. *Oceanic Linguistics* 33 (2): 309–321.
2001 Tentative calibration of time depths in Austroasiatic branches. paper presented at the Colloque «Perspectives sur la Phylogénie des Langues d'Asie Orientales» at Périgueux, 30 August 2001.
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, and Alicia Sanchez-Mazas (eds.), 77–80. London: Routledge Curzon.
2009 More on Dvaravati Old Mon. paper presented at the Fourth International Conference on Austroasiatic Linguistics, Mahidol University at Salaya, 29 October 2009.
2011 Austroasiatic word histories: boat, husked rice and taro. In *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*, Nick J. Enfield (ed.), 295–313. Canberra: Pacific Linguistics.
- van Driem, George
1999 Four Austric Theories. *Mother Tongue* V: 23–27.
2001 *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region, containing an Introduction to the Symbiotic Theory of Language* (2 volumes). Leiden: Brill.
2002 The Father Tongue hypothesis: Sexually dimorphic dispersals. In prehistory. paper presented at the 17th conference of the Indo-Pacific Prehistory Association, held at the Academia Sinica, Taipei, 9–15 September 2002.
2003 The Austric Problem: Issues, Solutions, Ramifications. paper presented at the 9th International Conference on Austronesian Linguistics at Australian National University in Canberra, 11 January 2003.
2007 Austroasiatic phylogeny and the Austroasiatic homeland in light of recent population genetic studies. *Mon-Khmer Studies* 37: 1–14.

- 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.
- 2012 The ethnolinguistic identity of the domesticators of Asian rice. *Comptes Rendus Palevol* 11 (2): 117–132.
- Dulik, Matthew C., Sergey I. Zhadanov, Ludmila P. Osipova, Ayken Askapuli, Lydia Gau, Omer Gokcumen, Samara Rubinstein, and Theodore G. Schurr
- 2012 Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between native Americans and indigenous Altaians. *American Journal of Human Genetics* 90: 1–18.
- Excoffier, Laurent, Matthieu Foll, and Rémy J. Petit
- 2009 Genetic consequences of range expansions. *Annual Review of Ecology, Evolution and Systematics* 40: 481–501.
- Excoffier, Laurent, and Nicolas Ray
- 2008 Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology and Evolution* 23 (7): 347–351.
- Fagundes, Nelson J.R., Nicolas Ray, Mark Beaumont, Samuel Neuenschwander, Francisco M. Salzano, Sandro L. Bonatto, and Laurent Excoffier
- 2007 Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences* 104 (45): 17614–17619.
- Ferlus, Michel
- 1996 Du taro au riz en Asie du Sud-Est: petite histoire d'un glissement sémantique. *Mon-Khmer Studies* 25: 39–49.
- 2010 The Austroasiatic vocabulary for rice: its origin and expansion. *Journal of the Southeast Asian Linguistic Society* 3 (2): 61–76.
- Forster, Peter, and Colin Renfrew
- 2011 Mother tongue and Y chromosomes. *Science* 333: 13901–1391.
- Fuller, Dorian Q.
- 2011 Pathways to Asian civilisations: Tracing the origins and spread of rice and rice agricultures. *Rice* 4 (3–4): 78–92.
- Garris, Amanda J., Thomas H. Tai, Jason Coburn, Steve Kresovich, and Susan McCouch
- 2005 Genetic structure and diversity in *Oryza sativa* L.. *Genetics* 169: 1631–1638.
- 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.
- 1928 Ein Beitrag zur Chronologie des Neolithikums in Südostasien. In *Festschrift Pater Wilhelm Schmidt: sprachwissenschaftliche, ethnologische, religionswissenschaftliche, prähistorische und andere Studien*, Wilhelm Koppers (ed.), 809–843. Wien: Mechitaristen-Congregations-Buchdruckerei.
- 1932 Urheimat und früheste Wanderungen der Austronesier. *Anthropos* XXVII: 543–619.
- von Hevesy, Wilhelm F.
- 1930 On Wilhelm Schmidt's Munda-Mon-Khmer comparisons (Does an "Austic" family of languages exist?). *Bulletin of the School of Oriental Studies* VI (1): 187–200.
- Klopfstein, Seraina, Mathias Currat, and Laurent Excoffier
- 2006 The fate of mutations surfing on the wave of a range expansion. *Molecular Biology and Evolution* 23 (3): 482–490.
- Kovach, Michael J., Megan T. Sweeney, and Susan McCouch
- 2007 New insights into the history of rice domestication. *Trends in Genetics* 23 (11): 578–587.
- Kovach, Michael J., Mariafe N. Calingacion, Melissa A. Fitzgerald, and Susan R. McCough
- 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.
- Kraaijenbrink, Thirsa, Emma J. Parkin, Denise R. Carvalho-Silva, George van Driem, Guido Barbuji, Chris Tyler-Smith, Mark A. Jobling, and Peter de Knijff
- 2009 Genetic and linguistic borders in the Himalayan region. In *Becoming Eloquent: Advances in the Emergence of Language, Human Cognition and Modern Cultures*, Francesco d'Errico, and Jean-Marie Hombert (eds.), 181–201. Amsterdam: John Benjamins.
- Kraaijenbrink, Thirsa, Karma Tshering of Gaselô, and Peter de Knijff
- 2007 Allele frequency distribution for 21 autosomal STR loci in Bhutan. *Forensic Science International* 170: 68–72.
- Kraaijenbrink, Thirsa, Jean Robert M.L. Opgenort, Nirmal Man Tuladhar, George van Driem, and Peter de Knijff
- 2007 Allele frequency distribution for 21 autosomal STR loci in Nepal. *Forensic Science International* 168 (2–3): 227–231.
- Kuhn, Thomas S.
- 1969 *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- Kumar, Vikrant, Arimanda N.S. Reddy, Jagedeesh P. Babu, Tiprisetti N. Rao, Banrida T. Langstieh, Kumarasamy Thangaraj, Alla G. Reddy, Lalji Singh, and Battini M. Reddy
- 2007 Y-chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations. *BioMed Central Evolutionary Biology*, 7: 47 (14) www.biomedcentral.com/1471-2148/7/47

- Li Jīn, Peter A. Underhill, Vishal Doctor, Ronald W. Davis, Peidong Shen, L. Luca Cavalli-Sforza, and Peter J. Oefner
1999 Distribution of haplotypes from a chromosome 21 region distinguishes multiple prehistoric human migrations. *Proceedings of the National Academy of Sciences of the United States of America* 96: 3796–3800.
- Londo, Jason P., Yu-chung Chiang, Kuo-Hsiang Hung, Tzen-Yuh Chiang, and Barbara A. Schaal.
2006 Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proceedings of the National Academy of Sciences of the United States of America* 103 (25): 9578–9583.
- Molina, Jeanmaire, Martin Sikora, Nandita Garud, Jonathan M. Flowers, Samara Rubinstein, Andy Reynolds, Huang Pu, Scott Jackson, Barbara A. Schaal, Carlos D. Bustamante, Adam R. Boyko, and Michael D. Purugganan
2011 Molecular evidence for a single evolutionary origin of domesticated rice. *Proceedings of the National Academy of Sciences of the United States of America*, www.pnas.org/cgi/doi/10.1073/pnas.1104686108.
- Moreau, Claudia, Claude Bhérier, Hélène Vézina, Michèle Jomphe, Damian Labuda, and Laurent Excoffier
2011 Deep human genalogies reveal a selective advantage to being on an expanding wave front. *Science* 334: 1148–1150.
- Murray, Stephen O.
1994 *Theory Groups and the Study of Language in North America: A Social History*. Amsterdam: John Benjamins.
- Nakamura Shin-ichi
2010 The origin of rice cultivation in the Lower Yangtze region, China. *Archaeological and Anthropological Sciences*, 2 (2): 107–113.
- Parkin, Emma J., Thirsa Kraaijenbrink, Jean Robert M.L. Opgenort, Nirmal Man Tuladhar, George van Driem, Peter de Knijff, and Mark Jobling
2007 Diversity of 26 Y-STR haplotypes in a Nepalese population sample: Isolation and drift in the Himalayas. *Forensic Science International* 166 (2-3): 176–181.
- Parkin, Emma J., Thirsa Kraaijenbrink, Karma Tshering of Gaselô, Peter de Knijff, and Mark Jobling
2006 26-locus Y-STR typing in a Bhutanese population sample. *Forensic Science International* 161 (1): 1–7.
- Parsons, Beverley J., H. John Newbury, Michael T. Jackson and Brian V. Ford-Lloyd
1999 The genetic structure and conservation of aus, aman and boro rices from Bangladesh. *Genetic Resources and Crop Evolution* 46 (6): 587–598.
- Petit, Rémy J., and Laurent Excoffier
2009 Gene flow and species delimitation. *Trends in Ecology and Evolution* 24 (7): 386–393.
- Pictet, Adolphe
1859 *Les origines indo-européennes ou les aryas primitifs: Essai de paléontologie linguistique (première partie)*. Paris: Joël Cherbuliez.
- Poloni, Estella Simone, Ornella Semino, Giuseppe Passarino, A.S. Santachiara-Benerecetti, I. Dupanloup, André Langaney, and Laurent Excoffier
1997 Human genetic affinities for Y chromosome P49a,f/TaqI haplotypes 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).
- Poloni, Estella Simone, Nicolas Ray, Stefan Schneider, and André Langaney
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, and 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: University of California Press.
- Purugganan, Michael D.
2010 The evolution of rice: molecular vignettes on its origins and spread. *Archaeological and Anthropological Sciences* 2 (2): 61–68.
- Qamar, Raheel, Qasim Ayub, Aisha Mohyuddin, Agnar Helgason, Kehkashan Mazhar, Atika Manoor, Tatiana Zerjal, Chris Tyler-Smith, and S. Qasim Mehdi
2002 Y-chromosomal variation in Pakistan. *American Journal of Human Genetics* 70: 1107–1124.
- Quintana-Murci, Lluís, C. Krausz, Tatiana Zerjal, S. Hamid Sayar, Michael F. Hammer, S. Qasim Mehdi, Qasim Ayub, Raheel Qamar, Aisha Mohyuddin, U. Radhakrishna, Mark A. Jobling Chris Tyler-Smith, and Ken McElreavey
2001 Y-chromosome lineages trace diffusion of people and languages in southwestern Asia. *American Journal of Human Genetics* 68: 537–542.
- Rasmussen, Morten, Xiaosen Guo, Yong Wang, Kirk E. Lohmueller, Simon Rasmussen, Anders Albrechtsen, Line Skotte, Stinus Lindgreen, Mait Metspalu, Thibaut Jombart, Toomas Kivisild, Weiwei Zhai, Anders Eriksson, Andrea Manica, Ludovic Orlando, Francisco de la Vega, Silvano Tridico, Ene Metspalu, Kasper Nielsen, María C. Ávila-Arcos, J. Víctor Moreno-Mayar, Craig Muller, Joe Dortch, M.

- Thomas P. Gilbert, Ole Lund, Agata Wesolowska, Monika Karmin, Lucy A. Weinert, Bo Wang, Jun Li, Shuaishuai Tai, Fei Xiao, Tsunehiko Hanihara, George van Driem, Aashish R. Jha, François-Xavier Ricaut, Peter de Knijff, Andrea B. Migliano, Irene Gallego-Romero, Karsten Kristiansen, David Lambert, Søren Brunak, Peter Forster, Bernd Brinkmann, Olaf Nehlich, Michael Bunce, Michael Richards, Ramneek Gupta, Carlos Bustamante, Anders Krogh, Robert A. Foley, Marta Mirazón Lahr, François Balloux, Thomas Sicheritz-Pontén, Richard Villems, Rasmus Nielsen, Wang Jun, Eske Willerslev
- 2011 An aboriginal Australian genome reveals separate human dispersals into Asia. *Science* 334 (6052): 94–98.
- Ratliff, Martha
- 2010 *Hmong-Mien Language History*. Canberra: Pacific Linguistics.
- Reddy, B. Mohan, B.T. Langstieh, Vikrant Kumar, T. Nagaraja, A.N.S. Reddy, M. Aruna, K. Thangaraj, A.G. Reddy, and Lalji Singh
- 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
- Reid, Lawrence A.
- 1994 Morphological evidence for Austric. *Oceanic Linguistics* 33 (2): 323–344.
- Sahoo, Sanghamitra, Anamika Singh, G. Himabindu, Jheeman Banerjee, T. Sitalaximi, Sonali Gaikwad, R. Trivedi, Phillip Endicott, Toomas Kivisild, Mait Metspalu, Richard Villems, and V. K. Kashyap
- 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.
- Schmidt, Wilhelm
- 1906 Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens. *Archiv für Anthropologie, Neue Folge*, V: 59–109.
- Sengupta, Sanghamitra, Lev A. Zhivotovsky, Roy King, S.Q. Mehdi, Christopher A. Edmonds, Cheryl-Emiliane T. Chow, Alice A. Lin, Mitashree Mitra, Samir K. Sil, A. Ramesh, M.V. Usha Rani, Chitra M. Thakur, L. Luca Cavalli-Sforza, Partha P. Majumder and Peter A. Underhill
- 2006 Polarity and temporality of high-resolution Y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralists. *The American Journal of Human Genetics* 78: 202–221.
- Takahashi Hiroko, Sato Yo-ichiro, and Nakamura Ikuo
- 2008 Evolutionary analysis of two plastid DNA sequences in cultivated and wild species of *Oryza*. *Breeding Science* 58 (3): 225–233.
- Tang Jiabin, Xia Hong'ai, Cao Mengliang, Zhang Xiuqing, Zeng Wanyong, Hu Songnian, Tong Wei, Wang Jun, Wang Jian, Yu Jun, Yang Huanming and Zhu Lihuang
- 2004 A comparison of rice chloroplast genomes. *Plant Physiology* 135: 412–420.
- Vavilov, Nikolaï Ivanovič
- 1926 Studies on the origin of cultivated plants. *Trudy po Prikladnoi Botanike, Genetike i Selekcii (Bulletin of Applied Botany, Genetics and Plant Breeding)* 16: 1–248.
- Wen Bo, Li Hui, Lu Daru, Song Xiufeng, Zhang Feng, He Yungang, Li Feng, Gao Yang, Mao Xianyun, Zhang Liang, Qian Ji, Tan Jingze, Jin Jianzhong, Huang Wei, Ranjan Deka, Sù Bīng, Ranajit Chakraborty, and Jin Li
- 2004 Genetic evidence supports demic diffusion of Han culture. *Nature* 431: 302–305.
- Wood, Elizabeth T., Daryn A. Stover, Christopher Ehret, Giovanni Destro-Bisol, Gabriella Spedini, Howard McLeod, Leslie Louie, Mike Bamshad, Beverly I. Strassmann, Himla Soodyall, and Michael F. Hammer
- 2005 Contrasting patterns of Y chromosome and mtDNA variation in Africa: Evidence for sex-biased demographic processes. *European Journal of Human Genetics* 13: 867–876.
- Yan Si, Chuan-Chao Wang, Huí Lǐ, Shi-Lin Li, Jīn Lì, and the Genographic Consortium
- 2011 An updated tree of Y-chromosome haplogroup O and revised phylogenetic positions of mutations P164 and PK4. *European Journal of Human Genetics* 19: 1013–1015.
- Zerjal, Tatjana, B. Dashnyam, A. Pandya, Manfred Kayser, Lutz Roewer, F.R. Santos, W. Schiefenhover, N. Fretwell, M.A. Jobling, S. Harihara, K. Shimuzu, D. Semjidaa, Antti Sajantila, P. Salo, M.H. Crawford, E.K. Ginter, Oleg V. Efgrafov, and Chris Tyler-Smith
- 1997 Genetic relationships of Asian and Northern Europeans, revealed by Y chromosomal DNA analysis. *American Journal of Human Genetics* 62: 1174–1183.