

1 The Eastern Himalaya and the Mongoloid myth

George van Driem

1. Defining the Eastern Himalaya

In the west, the Himalayas are punctuated by the Tirič Mīr in the Hindu Kush at 7,708 m and by the K2 in the Qarāqoram at 8,661 m. In the east, the Himalayas are punctuated by the Ganš dKar-po [k^hǣ k̄a:pō] in eastern Tibet at 6,740 m and the Hkakabo Razi in northern Burma at 5,881 m. The Himalayan massif runs a vast length of over 3,600 km from the Hazārahjāt highlands in the west to the Liángshān in the east. The Eastern Himalaya can be said to encompass the eastern half of the Himalayas, beginning from the Dhaulāgiri 8,167 m in central Nepal on eastward. The Kālī Gaṇḍakī River, which flows just at the foot of the Dhaulāgiri, bisects the great Himalayan range into two halves of roughly equal length.

Although the Himalayas are the highest mountain range on our planet, they form no watershed, since many of the rivers are of greater antiquity than the mountains themselves. The Himalayas only began to rise long after the Tethys Sea shrivelled up some 35 million years ago and the once insular habitat of the Indian subcontinent had fused with the Eurasian mainland. Like a number of other prominent Himalayan rivers, the Kālī Gaṇḍakī runs right through the Himalayas, originating on the Tibetan plateau and coursing down through the mountains onto the Gangetic plain. This dramatic invagination at the very centre of the Himalayan range is prominently visible to any airplane passenger flying across the Gangetic plain. For ethnolinguistic phylogeography, the Kālī Gaṇḍakī demarcates a vast region known as the Eastern Himalaya, which extends eastward all the way into the Indo-Burmese borderlands and the Chinese provinces of Yúnnán and Sìchuān and constitutes an area of pivotal importance to population prehistory.

2. The Mongoloid myth

As a species, we have always been obsessed with how we look and in which ways we appear to be similar or different from one another. The ancient Hindu caste system and the apartheid system of South Africa were just two of many systems based on our perceptions of caste, tribe and race. Even before the Portuguese first set foot in Japan in 1542, Europeans were trying to come to grips with the human phenotypical diversity which they observed in the peoples whom they met on their voyages across the seas. Today we understand that in scientific terms, there is actually no such thing as race (Cavalli-Sforza *et al.* 1994). We are all members of one large human family. The relationship between genes, their phenotypical expression and their pleiotropic interplay is inordinately complex, and our individual differences often tend to be larger than the differences between groups.

Historically, long before the discovery of the molecular mechanisms underlying genetics, scholars resorted to superficial classifications in their attempts to understand human diversity. Classification was conducted on the basis of somatology, which involved crude observations about external appearance. In 1758, in the famous tenth edition of his *Systema Naturæ*, Carl Linnæus distinguished between four geographical subspecies of *Homo sapiens*, i.e. *europæus*, *afër*, *asiaticus* and *americanus*. Later, Johann Friedrich Blumenbach, in a dissertation which he defended at Göttingen in 1775, distinguished between what he imagined were five human races, namely the ‘white’ *Caucasiae*, the ‘yellow’ *Mongolicae*, the ‘black’ *Aethiopicae*, the ‘red’ *Americanae* and the ‘brown’ *Malaiacae* (1776 [1795]: xxiii, xxiv). With his coinages, Blumenbach single-handedly invented the ‘Mongoloid’ and ‘Caucasoid’ races. With regard to his *Varietas Caucasia*, Blumenbach opined:

The name of this variety is taken from the Caucasus mountains, as well as, indeed, most of the southern flank thereof, in the Georgian area, where the most beautiful race of men is to be found and in whom all the physiological reasons converge so that it may be presumed that the first human beings are likely to have been native to this region.

(1795/1776: 303)¹

Later, Johann Christian Erxleben recognised four of the same races as Blumenbach but under different names, with his *Homo sapiens europæus*, *asiaticus*, *afër* and *americanus* (1777: 1, 2) corresponding

to Blumenbach's *Varietas Caucasia*, *Mongolica*, *Aethiopica* and *Americana* (1795 [1776]: 304, 307, 310, 319) respectively. As opposed to Blumenbach's *Varietas Malaica*, Erxleben distinguished no separate Malay race, but he made finer distinctions in northern Asia, distinguishing a more northerly *Homo sapiens tatarus* from the Chinese phenotype, which he termed *Homo sapiens asiaticus*, and he grouped Lapps, Samoyeds and other Uralic peoples under a distinct heading named *Homo sapiens lappo*. In France in 1801, Julien-Joseph Virey basically followed Blumenbach in recognising five races, but he outdid Erxleben in his attempts further to subclassify within these races.²

Taking his inspiration from Blumenbach, the German scholar Christoph Meiners (1747–1810), on the basis of the descriptions in Dutch and Russian accounts of the peoples encountered in other parts of the world, set up a classification of races based on what he imagined where the *uralte Stammvölker* or racial prototypes of mankind. His cogitations were published posthumously in three volumes. In the second volume, *der alte Mongolische Stamm* or 'the Mongoloid race' was designated by Meiners as one of the main races of mankind. He wrote:

In physiognomy and physique, the Mongol diverges as much from the usual form as does the Negro. If any nation merits being recognised as a racial prototype, then it should rightfully be the Mongol, who differs so markedly from all other Asian peoples in his physical and moral nature.

(1813, 2: 61)³

Meiners described the cruelty of the invading hordes led by Genghis Khan as being inherent to the 'moral nature' of the Mongoloid race, conveniently overlooking the historically well-documented cruelties of Western and other peoples. The serendipity of the nomenclatural choices made by Blumenbach (1795 [1776]) and Meiners (1813) gave rise to the Mongoloid myth. If the Mongols were the primordial tribe from which all peoples of the Mongoloid race descended, then it was logical to think that the homeland of all Mongoloids lay in Mongolia.

Jean Baptiste Bory de Saint-Vincent (1825: 323–325) subsequently introduced the term *Homo sapiens sinicus* for the Chinese, who he thought distinct from proper Mongoloids, but the 'Chinese race' would later vanish from subsequent classificatory schemes because the Chinese came to be seen by such early physical anthropologists as a

mixture of the Northeast Asian 'Tungids' and the 'Palaeomongoloids' of the Himalayas and Southeast Asia.⁴

I have often been told by people in Nepal and northeastern India that their ancestors came from Mongolia. Some even adorn their lorries, cars and motorcycles with captions like 'Mongol' or 'Mongolian'. When I ask them why they think so, they tell me that they are members of the Mongoloid race or मंगोल जाति *Maṅgol jāti*, which, as the name tells us, must have originated in Mongolia. I do not have the heart to tell them that the very idea was dreamt up by a German scholar in Göttingen in the early 1770s, who was just imaginatively trying to make sense of human diversity, though he had no expertise or specialist knowledge to do so.

People in the West suffer from the same obsolete ideas. A friend of mine from Abkhazia, who happens to be a renowned linguist, was travelling in the United States of America with a colleague of his from the Republic of Georgia. Whilst driving a rented car, they were pulled over by a police officer. The obese and heavily armed man in uniform demanded to see my friend's driving licence and then asked them, 'Are you folks Arabs?' The policeman spoke with a heavy American accent and pronounced the word Arabs as ['eɪræ:bz]. Since Abkhazia and Georgia both lie in the Caucasus, my friend responded, 'No, Sir, we are both Caucasians'. This response somehow displeased the police officer, who asserted, 'I am a Caucasian!'. My friend coolly responded, 'No, Sir, you are not a Caucasian, and you do not look particularly like a Caucasian. We are Caucasians.' The exasperated policeman spluttered, '... but ... but I am white!'

In the aftermath, my friend had to explain to the American policeman where the Caucasus Mountains lay and who the Caucasians were. However, he did not go as far as to explain that the idea that Europeans were purportedly Caucasian originated with Blumenbach in the early 1770s. Like the Mongoloid, the Caucasoid was another one of his racial prototypes. Americans who apply for a driving licence, take a Scholastic Aptitude Test or fill in any number of other official forms are often asked to specify their race. A person of European ancestry often checks a box saying that he or she is a 'Caucasian'. Some people from Asia and Africa are baffled by these racial questions and by the choices of race on offer, which differ from one form to another, and then end up having to decide whether they are 'coloured' or belong to some other 'race'. Although the topic of race is taboo in America, American society is both riddled with antique modes of thinking about race and very much in denial about widely held racist assumptions.

By contrast, indigenous peoples of Nepal, Sikkim, Bhutan and north-eastern India have a legitimate interest in their ancestry. Native people of the Eastern Himalaya share a natural and logical curiosity about why they appear to be different from the Brahmins and Chetris of Nepal and from the majority of Indians in India. We are all interested in where we came from, and both historical linguistics and population genetics can shed some light on this question. Before we examine some of the new insights from the field of ethnolinguistic phylogeography, a number of caveats should be noted.

3. Language and genes

There is a long lineage of scholars from Julius von Klaproth and Friedrich Max Müller (1872) who, since the early nineteenth century, have stressed that language and biological ancestry are two different things. There have been others too, like Sir William Jones, who from time immemorial have confounded language and race. Generally, people throughout history have been inclined to speak the language spoken by their parents, but the language which we happen to speak today may very well not be the language of our parents. Since genes are invariably inherited by offspring from their biological parents, a probabilistic correlation may therefore exist between language and genes in human populations, though this need not necessarily be so.

Historical linguistics and human population genetics present two distinct windows on the past. At the same time, the time depth accessible to historical linguistics is an order of magnitude shallower than the time depth accessible to genetics. Language families represent the maximal time depth accessible to historical linguists because the relatedness of languages belonging to a recognised linguistic phylum represents the limit of what can be demonstrated by the comparative method. This epistemological barrier represents the linguistic event horizon. Languages and genes are independent, but correlations may exist between chromosomal markers and language. Yet, these relationships should not be confused with identity. The correlation of a particular genetic marker with the distribution of a certain language family must not be simplistically equated with populations speaking languages of a particular linguistic phylum.

Moreover, we must also take into account the potential skewing effects of natural selection, gene surfing, recurrent bottlenecks during range expansion and the sexually asymmetrical introgression of resident genes into incursive populations. Factors such as ancient population structure and possible ancient Y chromosomal introgression could also affect inferences and interpretations based on any single Y chromosomal locus when

attempting to reconstruct migrations and elucidate the geographical origins of populations (Mendez *et al.* 2013; van Driem 2012b). Even with all these caveats in place, we must be especially aware of all provisos and qualifications included in our inferences and working hypotheses when attempting to understand East Asian ethnolinguistic phylogeography.

4. Father tongues

When studying the distribution of maternally inherited markers in the mitochondrial DNA and paternally inherited markers on the Y chromosome, 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 had already been described by Poloni *et al.* (1997, 2000) before the appearance of the seminal articles on Y chromosomal phylogeography by Underhill *et al.* (2000, 2001). Subsequent work, e.g. Karafet *et al.* (2008), further refined the resolution of the Y chromosomal haplogroup tree.

The inference was made 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. It is reasonable to infer that some mechanisms of language change may be inherent to this pathway of transmission. Phylogenies of autosomal single nucleotide polymorphisms in whole genome studies are making headway (Li *et al.* 2008), but it is still too early to tell to what extent correlations of autosomal markers with language phyla will be identified that are as salient as the currently observed Father Tongue correlations.

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. By contrast, 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.

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 HàN Chinese during the sinification of southern China, recounted in detail in the Chinese chronicles, is just as faithfully 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). The saliency of Y chromosomal haplogroups in tribal and caste populations in India contrasts with the comparatively featureless nature and antiquity of the mitochondrial landscape (Thanseem *et al.* 2006; Thangaraj *et al.* 2006a).

Previously, it has been proposed that the spread of Y chromosomal R subclades is likely to be linked to the dispersal of Indo-European from an original homeland in the Pontic–Caspian steppe (van Driem 2007, 2012b), but the unfolding story of Y chromosomal R lineages will no doubt turn out to be complex (Underhill *et al.* 2010). In order to be conclusive, a fine-mesh study of populations inhabiting the Western Himalayan region should be undertaken. Similarly, it has been proposed that the Y chromosomal lineage L, which shows a great diversity of subclades on the Iranian highland, can be identified as the possible marker of a patrilingual dispersal of Elamo-Dravidian emanating from a region which included the Bactria and Margiana of later prehistory (van Driem 2012b), and that one of these Y chromosomal L subclades will appear to be correlated with the patrilingual spread of Dravidian languages from the Indus Valley into southern India (van Driem 2014b). I have also proposed that haplogroup Q, an offspring clade of Y chromosomal haplogroup P, could be a marker for the Greater Yenisseian linguistic phylum (van Driem 2008, 2014b).

Populations forming local exceptions to the Father Tongue correlation, such as the Hungarians and the Balti, have been discussed elsewhere (van Driem 2012b, 2014b). Even in areas of the globe where the Father Tongue correlation does not hold, such as Tibetan-speaking Baltistan, we can observe that the main Y chromosomal haplogroups are more recent arrivals than the main maternal lineages. In the following sections, the ancestry of the native peoples of the Eastern Himalaya is explained. In so doing, we focus on the identification of the paternal haplogroup O2a (M95) with the spread of Austroasiatic, haplogroup O3a3c (M134) with Trans-Himalayan, haplogroup O3a3b (M7) lineage with Hmong-Mien

and O1 (MSY2.2) with Austro-Tai (van Driem 2007, 2012b, 2014b). Against the background of the East Asian linguistic theory, linguistic ancestry will be seen to correlate well with paternal ancestry.

5. The Trans-Himalayan family and the Sino-Tibetan myth

The second most populous linguistic phylum on the planet is Trans-Himalayan or Tibeto-Burman. Most speakers of Trans-Himalayan languages today live to the north and east of the Himalayas (Figure 1.1), but



Figure 1.1 Geographical distribution of Trans-Himalayan languages.
Note: Maps in this chapter are not to scale.
Source: Dr. Chr. Enderle⁵

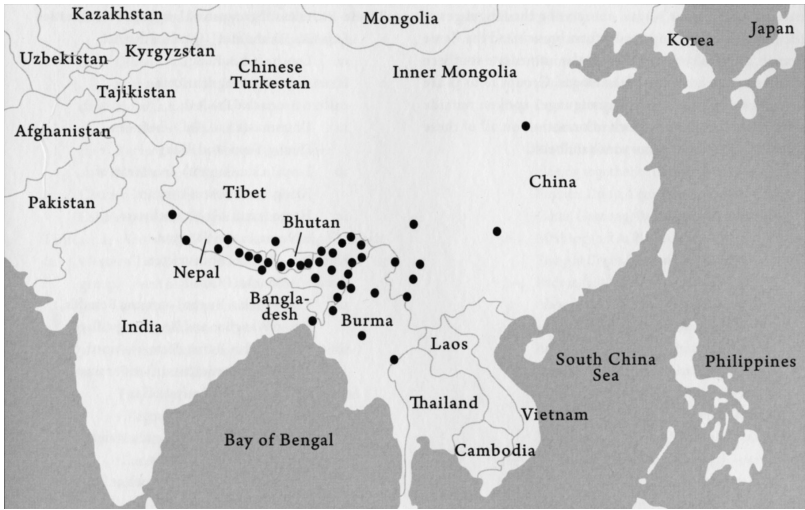


Figure 1.2 Geographical distribution of the major Trans-Himalayan subgroups. Each dot represents not just one language, but the putative historical geographical centre of each of forty-two major linguistic subgroups.

Source: Author

most of the over 300 different languages and three-fourths of the major Trans-Himalayan subgroups are located to the south of the Himalayan divide (Figure 1.2). The Trans-Himalayan linguistic phylum was first recognised by Klaproth in 1823, who identified the family as consisting of Tibetan, Chinese, Burmese and related languages. This linguistic phylum was called Tibeto-Burman by scholars in the British Isles, e.g. Hodgson (1857), Cust (1878), Forbes (1878), Houghton (1896).

In addition to the Mongoloid myth, another widespread myth which has only recently come to be dispelled is the Sino-Tibetan myth. Until 1924, Sino-Tibetan was called Indo-Chinese, a hypothetical language family containing all the languages of Asia and Oceania, including Japanese, the Polynesian languages and even all the languages of Papua New Guinea. The theory was dreamt up by a Scotsman called John Caspar Leyden, who made a meteoric career as a British civil servant in Asia during the Napoleonic wars but then died at the age of 35 in Indonesia. The idea that all Asian and Oceanic languages shared some ‘common mixed origin’ appealed to British colonial authorities.

The Indo-Chinese tree came to be whittled down in size over time but also became tinged with racist ideologies. The rebranding of the theory in 1924 as Sino-Tibetan helped to disguise the racist underpinnings of the model. Aside from its tainted history, the Sino-Tibetan family tree itself was false and consisted of two branches, one of which was Sino-Daic. When the Kradai or Daic languages were finally removed from Sino-Tibetan, the reduced Sino-Tibetan tree still represented a false phylogeny, uniting all non-Sinitic languages into a single subgroup which Sino-Tibetanists misleadingly labelled ‘Tibeto-Burman’. No Sino-Tibetanist has ever been able to adduce any historical linguistic evidence for this taxon and therefore for the family tree.⁶

Sino-Tibetan was assailed by scholars who proposed other models, e.g. Sino-Burman (Ramstedt 1957), Sino-Himalayan (Bodman 1973, 1980) and Sino-Kiranti (Starostin 1994). Finally, even Jim Matisoff, the retired Berkeley professor who once championed the model, has now publicly recanted the Sino-Tibetan phylogenetic model on three occasions.⁷ This step ought to be lauded as a noble act on his part because he had previously defended the Indo-Chinese family tree ever since, as a student at Columbia University in the 1960s, he inherited the antiquated model from his mentor Paul Benedict. The ability to change one’s mind in the face of evidence, or the lack thereof, is a defining trait of a scientist.

The Sino-Tibetan myth must be ousted as a false theory because this model has continued to mislead a number of scholars even in recent years. Yet, dispelling myths can be an arduous task because of the tenacity with which such narratives can take hold of the human mind. Today the default model remains Klaproth’s original Tibeto-Burman linguistic family, augmented by all the linguistic subgroups which have come to be recognised by linguists since 1823 to the present day (Figure 1.3). Since 2004, the newer name Trans-Himalayan has been gaining currency for Tibeto-Burman because this neutral geographical name accurately reflects the pivotal concentration and distribution of main subgroups of the linguistic phylum.

6. The East Asian linguistic theory

Following in the footsteps of scholars such as Witsen (1692) and Hadrianus Relandus (1706, 1707, 1708), Klaproth challenged conventional wisdom in 1823 by proposing a polyphyletic view of Asian language families. In assailing the dominant biblically inspired paradigm of a single gargantuan language family encompassing all Asian languages, Klaproth was able to distinguish the contours of many of the known



Figure 1.3 Thirty out of forty-two Trans-Himalayan subgroups lie south of the Himalayan divide, seven to the north and east, and five (Tshangla, Bodish, Nungish, Lolo-Burmese and Kachinic) straddle both flanks of the Himalayas.

Source: Author

Asian linguistic phyla. The five major linguistic phyla recognised today which form part of the East Asian story are Trans-Himalayan, Hmong-Mien, Kra-dai, Austronesian and Austroasiatic (Figures 1.1, 1.4–1.7).

Once Klaproth’s polyphyletic view had been in place for nearly a century, scholars began to discern possible long-distance relationships between the recognised language families. We might say that for linguistic taxonomy, Klaproth’s centrifugal step was gradually followed by a series of centripetal steps. Gustave Schlegel (1901, 1902) agreed with Klaproth in assessing Kra-dai to be unrelated to Sinitic, merely replete with Sinitic loans, and argued instead that Kra-dai was related to Austronesian. Schlegel’s old theory was taken up by Benedict (1942, 1975, 1976, 1990) under the guise of ‘Austro-Thai’, though this putative genetic link constituted just an ingredient in Benedict’s grand and poorly supported ‘Japanese/Austro-Tai’.



Figure 1.4 Geographical distribution of Hmong-Mien.

Source: Dr. Chr. Enderle

Weera Ostapirat (2005, 2013) was the first to present methodologically sound and cogent historical comparative evidence that Kra-dai and Austronesian represent coordinate branches of an Austro-Tai family. The coordinate branches of Ostapirat's Austro-Tai represent an ancient migration from what today is southern China across the Taiwan Strait to Formosa, where the Austronesian linguistic phylum established itself, whilst the proto-language ancestral to today's Kra-dai language communities remained behind on the mainland. Much later, the Formosan exodus led to the spread of the Malayo-Polynesian



Figure 1.5 Geographical distribution of Kradaï.

Source: Dr. Chr. Enderle

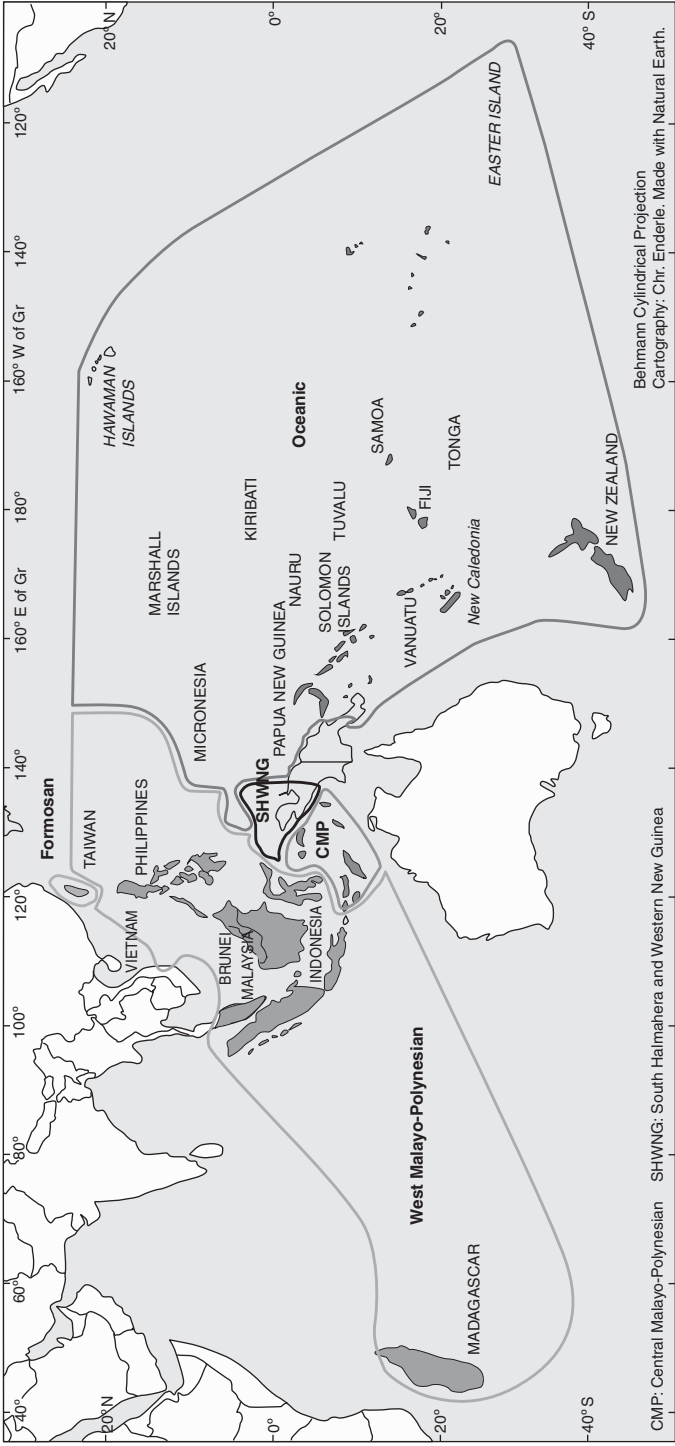


Figure 1.6 Geographical distribution of Austronesian.

Source: Dr. Chr. Enderle

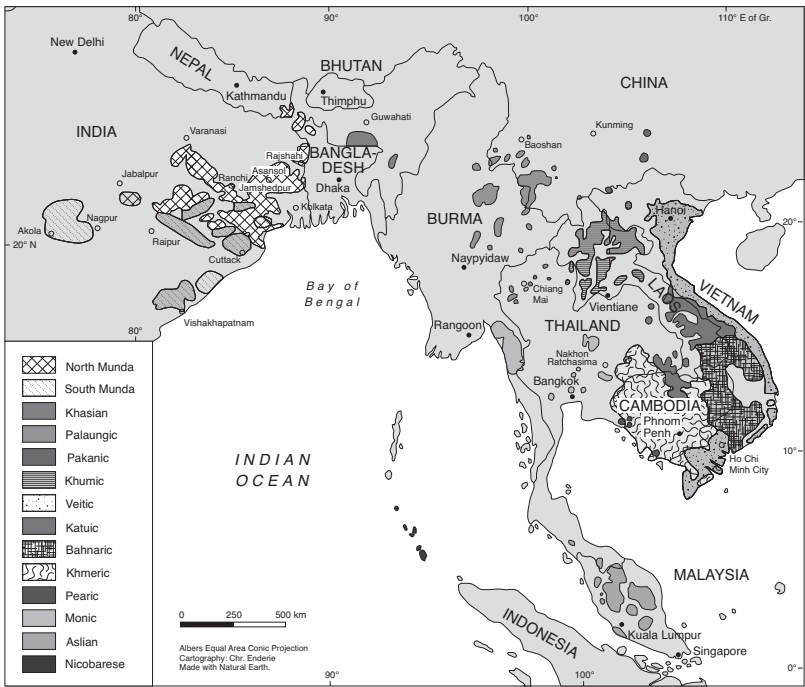


Figure 1.7 Geographical distribution of Austroasiatic.

Source: Dr. Chr. Enderle

branch throughout the Philippines, the Malay Peninsula, the Indonesian Archipelago, Madagascar and Oceania.

By uniting Kradai and Austronesian into Austro-Tai, Ostapirat reduced the five major linguistic phyla to just four: Austro-Tai, Trans-Himalayan, Hmong-Mien and Austroasiatic. Decades ago, transgressing the linguistic event horizon, Wilhelm Schmidt (1906) proposed an Austric macrofamily, uniting Austroasiatic and Austronesian, based on morphological evidence drawn especially from Nicobarese. Lawrence Reid became a proponent of Schmidt’s theory but also envisaged an even larger macrofamily, proposing that Austric ‘as a language family may eventually need to be abandoned in favour of a wider language family which can be shown to include both Austronesian and Austroasiatic languages but not necessarily as sisters of a common ancestor’ (Reid 2005: 150).

Conrady (1916, 1922) and Wulff (1934, 1942) each proposed a superfamily consisting of Austroasiatic, Austronesian, Kradai and Tibeto-Burman. Benedict (1942), Blust (1996) and Peiros (1998) proposed an

Austric superfamily comprising Austroasiatic, Austronesian, Kradai and possibly Hmong-Mien. Then in 2001 at Périgueux, a year before he died of congestive heart failure in Hawai'i, Stanley Starosta proposed the East Asian linguistic phylum encompassing Kradai, Austronesian, Tibeto-Burman, Hmong-Mien and Austroasiatic. Starosta's evidence was meagre, yet primarily morphological in nature. The ancient morphological processes shared by the families of this phylum, according to Starosta, were an agentive prefix *<m->, a patient suffix *<n->, an instrumental prefix <s-> and a perfective prefix *<n->. The East Asian word was ostensibly disyllabic and exhibited the canonical structure CVCVC.

Starosta's posthumously published East Asian phylogeny was marred by editorial errors (Starosta 2005: 183), which were later corrected (van Driem 2005: 322). A theory of linguistic relationship at this time depth lies at the frontier of what can be empirically demonstrated to the satisfaction of a methodologically rigorous historical linguist. This hypothesis will therefore remain an informed conjecture until solid historical linguistic evidence either further supports or overturns the model. At Benares in 2012, I presented the tweaked East Asian family tree depicted in Figure 1.8. The revised phylogeny is based on historical linguistic intuitions and other types of information about population prehistory (van Driem 2014b).

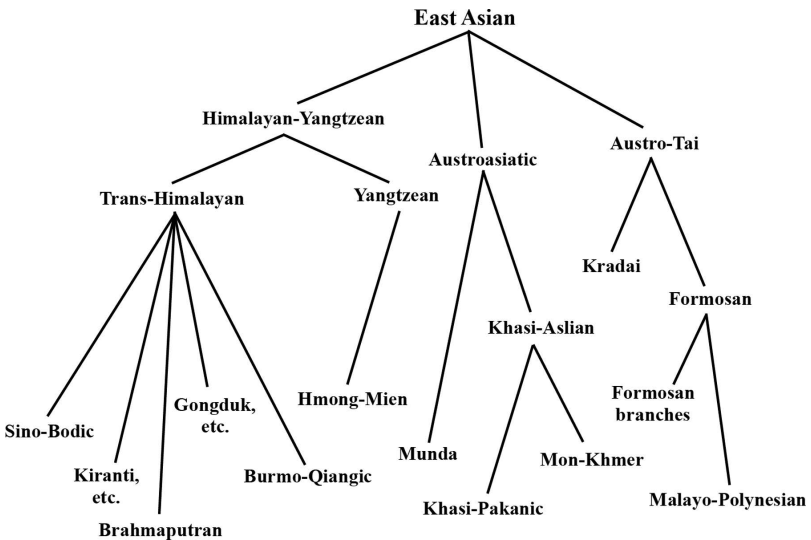


Figure 1.8 The 2012 Benares Recension: a revised East Asian phylogeny.

Source: Author

7. East Asian and the Eastern Himalayan homeland

The populations today speaking languages of the Trans-Himalayan, Hmong-Mien, Austroasiatic and Austro-Tai linguistic phyla are characterised by a preponderance of the Y-chromosomal haplogroup O. In fact, the four linguistic phyla are each characterised by a particular subclade of O, suggesting both a paternal spread of these language families as well as a time depth for the putative East Asian language family coeval with the antiquity of the paternal haplogroup O itself.

There is good reason to believe that the geographical locus of the ancestral haplogroup NO (M214) lay in the Eastern Himalaya. When the two paternal lineages N and O 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, gradually migrated across northern Eurasia as far west as Lapland (Rootsi *et al.* 2007; Derenko *et al.* 2007; Mirabal *et al.* 2009). I have identified this clade with the paternal spread of Michael Fortescue's Uralo-Siberian linguistic phylum (van Driem 2014b). The ancestral clade N* (M231) is still found in the highest frequency in northern Burma, Yúnnán and Sìchuān.

The fraternal clade O, which appears to be a marker for the linguistic ancestors of the hypothetical East Asian linguistic phylum, remained behind in the Eastern Himalaya. As temperature and humidity increased after the Last Glacial Maximum, the Y chromosomal haplogroup O (M175) split up into the subclades O1 (MSY2.2), O2 (M268) and O3 (M122). The three subclades can be putatively assigned to three geographical loci along an east–west axis for the sake of argument and without any claim to geographical precision. Whereas the haplogroup O1 (MSY2.2) moved to the drainage of the Pearl River and its tributaries in what today is Guǎngdōng, the bearers of haplogroup O2 (M268) moved to southern Yúnnán, whilst bearers of the O3 (M122) haplogroup remained in the southeastern Himalayas, expanding their range initially only into adjacent parts of northeastern India and northern Burma (Figure 1.9). The O2 (M268) clade split into O2a (M95) and O2b (M176), an event which took place just before the linguistic event horizon.

Asian rice, perhaps both *japonica* and *indica* rice, may have first been domesticated roughly in the area hypothetically imputed to O2 (M268), which would have included southern Yúnnán (van Driem 2011a, 2012a; Figure 1.9). The bearers of the subclade O2a (M95) became the *Stammväter* of the Austroasiatics (van Driem 2007; Chaubey *et al.* 2010). The Austroasiatics spread from this locus initially

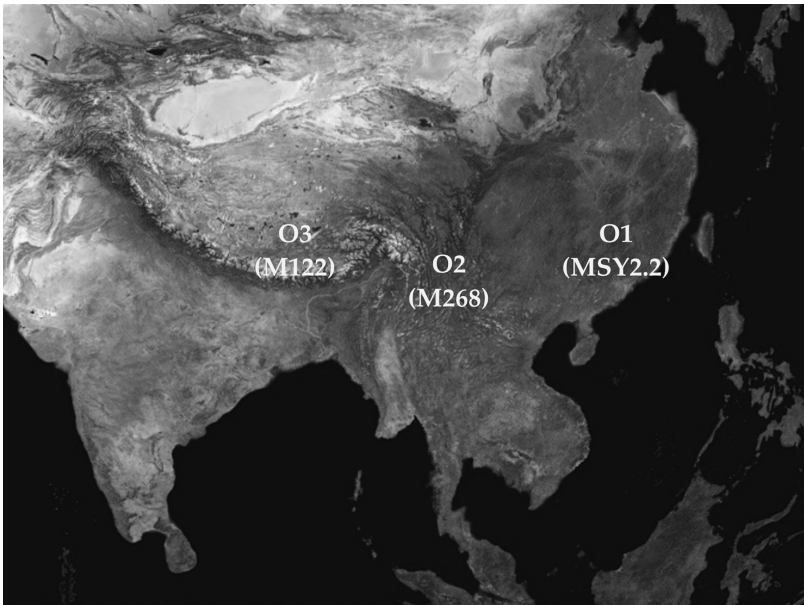


Figure 1.9 After the Last Glacial Maximum, the Y chromosomal haplogroup O (M175) split into the subclades O1 (MSY2.2), O2 (M268) and O3 (M122). Bearers of the O2 (M268) paternal lineage domesticated Asian rice.

Source: Author

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 introduce both their language and their paternal lineage to indigenous peoples of the subcontinent (Figure 1.10). Despite its prevalence in Munda populations, the topology of haplogroup O2a does not support a South Asian origin for this paternal lineage (Kumar *et al.* 2007; Chaubey *et al.* 2010). Once again the mitochondrial background is of greater antiquity, and the paternal lineage appears to be the signature for the spread of the language phylum and its adoption by resident populations (Thangaraj *et al.* 2006a; Kumar *et al.* 2006).

Since we have associated the paternal lineage O2a (M95), which is a derivative clade of haplogroup O2 (M268), with the Austroasiatic

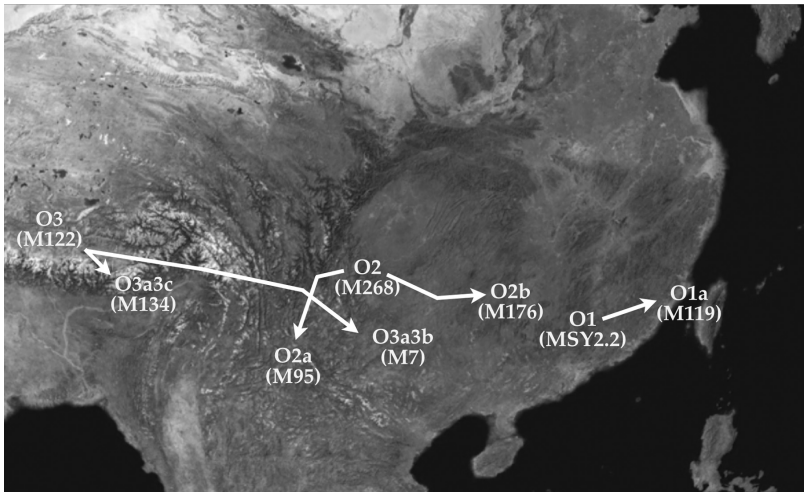


Figure 1.10 Paternal lineages branching into new subclades. Each event involved a linguistic bottleneck leading to language families that today are reconstructible as distinct linguistic phyla. The O1 (MSY2.2) lineage gave rise to the O1a (M119) subclade, which moved eastward to the Fújiàn hill tracts and across the strait to Formosa, which so became the *Urheimat* of the Austronesians. Bearers of O3a3b (M7) became the Proto-Hmong-Mien. In the Eastern Himalaya, the bearers of haplogroup O3a3c (M134) expanded and became the Trans-Himalayans. Haplogroup O2a (M95) is the Proto-Austroasiatic paternal lineage. The para-Austroasiatic fraternal clade O2b (M176) spread eastward, sowing seed along the way.

Source: Author

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). 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 perhaps an Austroasiatic name for the Yangtze River in Chinese (Pulleyblank 1983). This para-Austroasiatic paternal lineage moved as far as the Korean Peninsula and represents

the second major wave of peopling attested in the Japanese genome (Jin *et al.* 2009; Karafet *et al.* 2009).

We can identify the O2b (M176) lineage with the Yayoi people, who introduced rice agriculture to Japan, perhaps as early as the second millennium BC, during the final phase of the Jōmon period (Tanaka *et al.* 2004; Hammer *et al.* 2006). The Yayoi appear also to have introduced other crops of continental inspiration to the Japanese archipelago such as millet, wheat and melons. The gracile Yayoi immigrants soon outnumbered the more robust and less populous Jōmon, who had been the first anatomically modern humans to populate Japan. The Y chromosomal haplogroup O2b and other O haplogroups in Japan are later arrivals but account for more than half of all Japanese paternal lineages, with their highest frequencies in Kyūshū. A Father Tongue theory for Altaic which assumes no close affinity between Altaic and Uralo-Siberian entails that an antique C haplogroup, perhaps C3, represents an early trace of a paternally disseminated linguistic phylum at a great time depth. Much of this old linguistic stratum was lost long ago. The remnants of this Father Tongue survive in Japan as Japanese and elsewhere in Asia as the other languages of the Altaic language family, i.e. Korean, Tungusic, Mongolic and Turkic.⁸ Another Father Tongue, anciently introduced to Japan by the bearers of the Y chromosomal haplogroup D2 (M55), survives today as Ainu.

At the dawn of the Holocene in the Eastern Himalaya, haplogroup O3 (M122) gave rise to the ancestral Trans-Himalayan paternal lineage O3a3c (M134) and the original Hmong-Mien paternal lineage O3a3b (M7). The bearers of the polymorphism O3a3c (M134) stayed behind in the area comprising Nepal, Sikkim, Bhutan, northeastern India, southeastern Tibet and northern Burma, whilst the bearers of the O3a3b (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 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'. It would be more precise to infer that the incidence of haplogroup O3a3b (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 O3a3b (M7) in Austroasiatic communities of the Indian subcontinent is undetectably low. On the other hand, the incidence of Y chromosomal 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 (MYS2.2) paternal lineage gave rise to the O1a (M119) subclade, which moved from the Pearl River drainage eastward to the Mìn River drainage in the hill tracts of Fújiàn and across the strait to Formosa, which consequently became the *Urheimat* of the Austronesians (cf. Abdulla *et al.* 2009). Back west in the Eastern Himalaya, the bearers of Y chromosomal haplogroup O3a3c (M134) expanded further 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. In the southwest on the Brahmaputra Plain, the early Tibeto-Burmans encountered Austroasiatics, who had preceded them.

If we assume a linguistic dispersal in which languages were spread by populations in which a particular paternal lineage was dominant, then the Malayo-Polynesian expansion via the Philippines into insular Southeast Asia must have entailed the introduction of Austronesian by bearers of the Y chromosomal haplogroup O1a (M119) to resident communities, in which an originally Austroasiatic paternal lineage O2a (M95) was and would remain dominant even after linguistic assimilation, and other older paternal lineages also persisted (Karafet *et al.* 2005; Li *et al.* 2008). Similarly, Malagasy is linguistically clearly Austronesian, but genetically the Malagasy trace both their maternal and paternal ancestries equally to Borneo and to the African mainland (Hurles *et al.* 2005).

The ancestral Trans-Himalayan paternal lineage O3a3c (M134) spread from the Eastern Himalaya in a northeasterly direction across East Asia to the North China Plain. Subsequently, at a far shallower time depth, the Tibeto-Burman paternal lineage O3a3c (M134) spread from the Yellow River basin into what today is southern China, beginning with the Hàn expansion southward during the Qín dynasty in the third century BC. The ancestral Tibeto-Burman paternal lineage O3a3c (M134) is intrusively present in the Korean Peninsula and beyond, although Uralo-Siberian populations such as the Evenki predominantly retain the paternal lineage N. The distribution map of major

Trans-Himalayan linguistic subgroups shows the centre of linguistic phylogenetic diversity to be rooted squarely in the Eastern Himalaya, with outliers trailing off towards the Loess plains of the Yellow River basin in the northeast. This geographical projection of Trans-Himalayan linguistic diversity appears to reflect the spread of the paternal O3a3c (M134) lineage putatively associated with this linguistic dispersal.

Molecular genetic findings shed light both on ethnolinguistic prehistory and its unrecorded sociolinguistic dimensions, and often population geneticists find molecular corroboration of what some linguists and ethnographers have been claiming for centuries. Although paternal ancestry only represents a very small segment of our ancestry, emerging autosomal findings appear, at least in part, to corroborate the reconstruction presented here for meridional East Asia (Chaubey *et al.* 2010; Jinam *et al.* 2013). Correlations between linguistic, archaeology and genetics must inform a chronologically layered view of ethnolinguistic prehistory (Bellwood *et al.* 2011; van Driem 2011b).

The Eastern Himalaya from the Dhaulāgiri to the Liángshān and more particularly the region comprising Nepal, Sikkim, Bhutan, southeastern Tibet and northeastern India furnished the cradle for the ethnogenesis of all East Asian language families: Trans-Himalayan, Hmong-Mien, Austroasiatic and Austro-Tai. At even greater time depths, the Uralo-Siberian and Altaic linguistic phyla too may have ultimately originated in the Eastern Himalaya. In the hoary past, when our anatomically modern ancestors emerged from Africa on their way to East Asia, Southeast Asia, Oceania, Siberia, the Americas and even Lappland, many of these ancestors must at one point have passed through the Eastern Himalayan region and crossed the mighty Brahmaputra.

Notes

- 1 Nomen huic varietati a Caucaso monte, tum quos vicinia eius et maxime quidem australis plaga pulcherrimam hominum stirpem, Georgianam foveat; tum quod et omnes physiologicae rationes in eo conspirent, in eandem regionem, si uspiam, primos humani generis avtochthones verisimillime ponendes esse.
- 2 Julien-Joseph Virey distinguished ‘cinq races principales’ (all in Tome I: 124). The first race is ‘la celtique’, which has various types ranging from ‘les scandinaves’ in the northwest to the ‘scythes, persans, arabes, maures . . . et même les indous cisangetiques’ in the southeast (pp. 129–131), whereas ‘la racine originelle des mongols se partagent en trois branches’ (pp. 131, 132), i.e. those ‘qui embrasse presque toute la circonférence du pole arctique . . . esquimaux. tschutchis, kamtschadales . . . koriaques, ostiaques, gakates, yukagres, samoïedes . . . lapons’. ‘La seconde division’ comprised ‘les éléuths et calmouks . . . les tunguses, baskirks, kosaques vrais . . . kirguis, tschouvaches,

- burattes, soongarés . . . tous les peuples mantcheoux du nord de la Chine, et les tribus tangutiques du Thibet' (pp. 133, 134). 'Les mogols méridionaux' comprised 'les chinois, les japonais, les corésiens, tonquinois, cochinchinois, les habitans d'Iesso, plusieurs thibetains, siamois, etc.' (p. 135). The third race comprised 'les tribus malaies' throughout insular Southeast Asia and Madagascar (pp. 136, 137). The fourth race was 'l'espèce nègre', which comprised sub-Saharan Africa as well as the people 'de la nouvelle Hollande [i.e. Australia] . . . et la nouvelle Calédonie'. Virey named the fifth race 'caraïbe', which was disseminated throughout the Americas.
- 3 Die Gesichts – und Körperbildung der Mongolen steht von der gewöhnlichen Form eben so sehr ab, als die der Neger. Und wenn irgend eine Nation verdient, als uraltes Stammvolk betrachtet zu werden; so kommt dieser Name mit recht den von allen anderen Asiatischen Völkern, der körperlichen und moralischen Beschaffenheit nach so sehr verschiedenen Mongolen zu.
 - 4 Jean Baptiste Bory de Saint-Vincent (1825: 297) distinguished the 'Espèce Sinique, *Homo sinicus*. Presque toujours, mais improprement confondue avec la précédente sous le nom Mongole' (comprising 'Coréens, Japonais, Chinois, Tonkinois, Cochinchinois, Siamois, et des Hommes qui peuplent l'empire du Birman') as distinct from the Tungid type, which he inappropriately labelled 'Espèce Scythique, *Homo scythicus*' (p. 296), comprising 'Turcomans, Kirguises, Cosaques, Tartares, Kalmouks, Mongols et Mantchoux' (p. 294). Bory de Saint-Vincent also distinguished 'Espèce Hindoue, *Homo indicus*' (p. 300) and 'Espèce Hyperboréenne, *Homo hyperboreus*', comprising '. . . les Ostiaks, les Tonguses et les Jakoutes . . . les Jukaghires, les Tchoutchis, les Kouriaques, et quelques hordes de Kamtschadales . . .'.
 - 5 The figures drawn by Dr. Chr. Enderle are reproduced from George van Driem's contribution to *Nepal: An Introduction to the Natural History, Ecology and Human Environment in the Himalayas* (2015) with the gracious permission of the editors Colin Pendry and Georg Miede.
 - 6 This long episode in linguistic history has been recounted elsewhere (van Driem 2014a).
 - 7 The first such pronouncement took place on 29 October 2009 at the 4th International Conference on Austroasiatic Linguistics at Mahidol University, the second on 24 February 2012 in a talk entitled 'The present state of Sino-Tibetan studies: Progress and outstanding issues' at a special seminar for the Hakubi Project and Centre for Southeast Asian Studies at Kyōto University and the third on 26 October 2012 at the 45th International Conference for Sino-Tibetan Languages and Linguistics at Nanyang Technological University in Singapore.
 - 8 Martine Robbeets (2014) applies 'Altaic', the traditional name of this linguistic phylum, just to the language family comprising only Tungusic, Mongolic and Turkic, and she has introduced the new label 'Trans-Eurasian' for the linguistic phylum encompassing Japonic, Koreanic and 'Altaic' *sensu* Robbeets.

References

- Abdulla, M. A., I. Ahmed, A. Assawamakin, J. Bhak, S. K. Brahmachari, G. C. Calacal, A. Chaurasia, C.-H. Chen, J. Chen, Y.-T. Chen, J. Chu, E. M. C. Cutionco-de la Paz, M. C. A. de Ungria, F. C. Delfin, J. Edo, S. Fuchareon, H.

- Ghang, T. Gojobori, J. Han, S.-F. Ho, B. P. Hoh, W. Huang, H. Inoko, P. Jha, T. A. Jinam, J. Li, J. Jung, D. Kangwanpong, J. Kampuansai, G. C. Kennedy, P. Khurana, H.-L. Kim, K. Kim, S. Kim, W.-Y. Kim, K. Kimm, R. Kimura, T. Koike, S. Kulawonganunchai, V. Kumar, P. S. Lai, J.-Y. Lee, S. Lee, E. T. Liu, P. P. Majumder, K. K. Mandapati, S. Marzuki, W. Mitchell, M. Mukerji, K. Naritomi, C. Ngamphiw, N. Niikawa, N. Nishida, B. Oh, S. Oh, J. Ohashi, A. Oka, R. Ong, C. D. Padilla, P. Palittapongarnpim, H. B. Perdigon, M. E. Phipps, E. Png, Y. Sakaki, J. M. Salvador, Y. Sandraling, V. Scaria, M. Seielstad, M. R. Sidek, A. Sinha, M. Srikummool, H. Sudoyo, S. Sugano, H. P. Suryadi, Y. Suzuki, K. A. Tabbada, A. Tan, K. Tokunaga, S. Tongsimma, L. P. Villamor, E. Wang, Y. Wang, H. Wang, J.-Y. Wu, H. Xiao, S. Xu, J. O. Yang, Y. Y. Shugart, H.-S. Yoo, W. Yuan, G. Zhao, B. A. Zilfalil and The Indian Genome Variation Consortium. 2009. 'Mapping human genetic diversity in Asia', *Science*, 326: 1541–1545.
- Bellwood, P., G. Chambers, M. Ross and H.-C. Hung. 2011. 'Are "cultures" inherited: Multidisciplinary perspectives on the origins and migrations of Austronesian-speaking peoples prior to 1000 BC', in B. W. Roberts and M. van der Linden (eds.), *Investigating Archaeological Cultures: Material Culture, Variability and Transmission*. Dordrecht: Springer, pp. 321–354.
- Benedict, P. K. 1942. 'Thai, Kadai, and Indonesia: A new alignment in south-eastern Asia', *American Anthropologist*, 44: 576–601.
- Benedict, P. K. 1975. *Austro-Thai: Language and Culture, with a Glossary of Roots*. New Haven: Human Relations Area Files.
- Benedict, P. K. 1976. 'Sino-Tibetan: Another look', *Journal of the American Oriental Society*, 96 (2): 167–197.
- Benedict, P. K. 1990. *Japanese/Austro-Tai*. Ann Arbor: Karoma Publishers.
- Blumenbach, J. F. 1795 [originally 1776]. *De generis humani varietate nativa. Editio tertia*. Göttingen: Vandenhoeck & Ruprecht.
- Blust, R. 1996. 'Beyond the Austronesian homeland: The Austric hypothesis and its implications for archaeology', in W. H. Goodenough (ed.), *Prehistoric Settlement of the Pacific*. Philadelphia: American Philosophical Society, pp. 117–160.
- Bodman, N. C. 1973. 'Some Chinese reflexes of Sino-Tibetan s-clusters', *Journal of Chinese Linguistics*, 1 (3): 383–396.
- Bodman, N. C. 1980. 'Proto-Chinese and Sino-Tibetan: Data towards establishing the nature of the relationship', in F. van Coetsem and L. R. Waugh (eds.), *Contributions to Historical Linguistics: Issues and Materials*. Leiden: E.J. Brill, pp. 34–199.
- Bory de Saint-Vincent, Jean Baptiste. 1825. *Dictionnaire classique d'histoire naturelle. Tome Huitieme*. Paris: Rey et Gravier.
- Cai, X., Z. Qin, B. Wen, S. Xu, Y. Wang, Y. Lu, L. Wei, C. Wang, S. Li, X. Huang, H. Li? and the Genographic 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.

- Cavalli-Sforza, L. L., P. Menozzi and A. Piazza. 1994. *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.
- Chaubey, G., M. Metspalu, Y. Choi, R. Mägi, I. G. Romero, S. Rootsi, P. Soares, M. van Oven, D. M. Behar, S. Rootsi, G. Hudjashov, C. B. Mallick, M. Karmin, M. Nelis, J. Parik, A. G. Reddy, E. Metspalu, G. van Driem, Y. Xue, C. Tyler-Smith, K. Thangaraj, L. Singh, M. Remm, M. B. Richards, M. M. Lahr, M. Kayser, R. Villems and T. Kivisild. 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.
- Conrady, A. 1916. 'Eine merkwürdige Beziehung zwischen den austrischen und den indochinesischen Sprachen', in *Aufsätze zur Kultur – und Sprachgeschichte vornehmlich des Orients: Ernst Kubn zum 70. Geburtstag am 7. Februar 1916 gewidmet von Freunden und Schülern*. München: Verlag von M. & H. Marcus, pp. 475–504.
- Conrady, A. 1922. 'Neue austrisch-indochinesische Parallelen', in *Asia Major: Hirth Anniversary Volume*. London: Robsthan and Company, pp. 23–66.
- Cust, R. N. 1878. *A Sketch of the Modern Languages of East India*. London: Trübner and Company.
- Derenko, M., B. Malyarchuk, G. Denisova, M. Wozniak, T. Grzybowski, I. Dambueva and I. Zakharov. 2007. 'Y-chromosome haplogroup N dispersal from south Siberia to Europe', *Journal of Human Genetics*, 52: 763–770.
- Dulik, M. C., S. I. Zhadanov, L. P. Osipova, A. Askapuli, L. Gau, O. Gokcumen, S. Rubinstein and T. 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 (2): 229–246.
- Erxleben, J. C. P. 1777. *Systema regni animalis per Classes, ordines, Genera, Species, Varietates cum Synonymia et Historia Animalium*. Leipzig: Weygand.
- Forbes, C. J. F. S. 1878. 'On Tibeto-Burman languages', *Journal of the Royal Asiatic Society of Great Britain and Ireland*, 10: 210–227.
- Hammer, M. F., T. M. Karafet, H. Park, K. Omoto, S. Harihara, M. Stoneking and S. Horai. 2006. 'Dual origins of the Japanese: Common ground for hunter-gatherer and farmer Y chromosomes', *Journal of Human Genetics*, 51: 47–58.
- Hodgson, B. H. 1857. 'Comparative vocabulary of the languages of the broken tribes of Népal', *Journal of the Asiatic Society of Bengal*, 26: 317–371.
- Houghton, B. 1896. 'Outlines of Tibeto-Burman linguistic palæontology', *Journal of the Royal Asiatic Society*, January: 23–55.
- Hurles, M. E., B. C. Sykes, M. A. Jobling and P. Forster. 2005. 'The dual origin of the Malagasy in Island Southeast Asia and East Africa: Evidence from maternal and paternal lineages', *American Journal of Human Genetics*, 76: 894–901.
- Jin, H.-J., C. Tyler-Smith and W. Kim. 2009. 'The peopling of Korea revealed by analyses of mitochondrial DNA and Y-chromosomal markers', *Public Library of Science PLoS One*, 4 (1): e4210.

- Jinam, T. A., L.-C. Hong, M. E. Phipps, M. Stoneking, M. Ameen, J. Edo and HUGO Pan-Asian SNP Consortium and Naruya Satou. 2013. 'Evolutionary history of continental Southeast Asian: "Early train" hypothesis based on genetic analysis of mitochondrial and autosomal DNA data', *Molecular Biology and Evolution*, 29 (11): 3513–3527.
- Karafet, T. M., J. S. Lansing, A. J. Redd and J. C. Watkins. 2005. 'Balinese Y-chromosome perspective on the peopling of Indonesia: Genetic contributions from pre-Neolithic hunter-gatherers, Austronesian farmers and Indian traders', *Human Biology*, 77 (1): 93–114.
- Karafet, T. M., F. L. Mendez, M. B. Meilerman, P. A. Underhill, S. L. Zegura and M. F. Hammer. 2008. 'New binary polymorphisms reshape and increase resolution of the human Y-chromosomal haplogroup tree', *Genome Research*, 18: 830–838.
- Karafet, T. M., S. L. Zegura and M. F. Hammer. 2009. 'Y-chromosome Japanese roots', in P. N. Peregrine, I. Peiros and M. Feldman (eds.), *Ancient Human Migrations: A Multidisciplinary Approach*. Salt Lake City: University of Utah Press, pp. 137–148.
- Kumar, V., B. T. Langstieh, S. Biswas, J. P. Babu, T. N. Rao, K. Thangaraj, A. G. Reddy, L. Singh and B. M. Reddy. 2006. 'Asian and non-Asian origins of Mon-Khmer and Mundari speaking Austroasiatic populations of India', *American Journal of Human Biology*, 18: 461–469.
- Kumar, V., A. N. S. Reddy, J. P. Babu, T. N. Rao, B. T. Langstieh, K. Thangaraj, A. G. Reddy, L. Singh and B. M. Reddy. 2007. 'Y-chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations', *BioMed Central Evolutionary Biology*, 7: 47, doi:10.1186/1471-2148-7-47.
- Li, J. Z., D. M. Absher, H. Tang, A. M. Southwick, A. M. Casto, S. Ramachandran, H. M. Cann, G. S. Barsh, M. Feldman, L. L. Cavalli-Sforza and R. M. Myers. 2008. 'Worldwide human relationships inferred from genome-wide patterns of variation', *Science*, 319: 1100–1104.
- Lí Huī, B. W., S.-J. Chen, B. Su, P. Pramoongjago, Y. Liu, S. Pan, Z. Qin, W. Liu, X. Cheng, N. Yang, X. Li, D. Tran, D. Lu, M.-T. Hsu, R. Deka, S. Marzuki, C.-C. Tan and J. Li. 2008. 'Paternal genetic affinity between western Austronesians and Daic populations', *BMC Evolutionary Biology*, 8: 146, doi:10.1186/1471-2148-8-146.
- Linnæus, C. 1758. *Systema Naturæ per Regna tria Naturæ, secundum Classes, Ordines, Genera, Species, cum characteristibus, differentiis synonymis, locis*. (Editio Decima). Holmiæ: Laurentius Salvius.
- Meiners, C. 1813 [posthumous]. *Untersuchungen über die Verschiedenheiten der Menschennaturen (die verschiedenen Menschenarten) in Asien und in den Südländern, in den Ostindischen und Südseeinseln, nebst einer historischen Vergleichung der vormahligen und gegenwärtigen Bewohner dieser Continente und Eylande*. Tübingen: in der J.G. Cotta'schen Buchhandlung.
- Mendez, F. L., T. Krahn, B. Schrack, A.-M. Krahn, K. R. Veeramah, A. E. Woerner, F. L. M. Fomine, N. Bradman, M. G. Thomas, T. M. Karafet and M. F. Hammer. 2013. 'An African American paternal lineage adds an extremely

- ancient root to the human Y chromosome phylogenetic tree', *American Journal of Human Genetics*, 92: 454–459.
- Mirabal, S., M. Reguiero, A. M. Cadenas, L. L. Cavalli-Sforza, P. A. Underhill, D. A. Verbenko, S. A. Limborska and R. J. Herrera. 2009. 'Y chromosome distribution within the geo-linguistic landscape of northwestern Russia', *European Journal of Human Genetics*, 17: 1260–1273.
- Müller, F. M. 1872. *Über die Resultate der Sprachwissenschaft: Vorlesung gehalten in der kaiserlichen Universität zu Strassburg am XXIII. Mai MDCCCLXXII*. Strassburg: Karl J. Trübner, und London: Trübner & Co.
- Ostapirat, W. 2005. 'Kra-Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution', in L. Sagart, R. Blench and A. Sanchez-Mazas (eds.), *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon, pp. 107–131.
- Ostapirat, W. 2013. 'Austro-Tai revisited', *23rd Annual Meeting of the South-east Asian Linguistic Society*, 29 May 2013, Chulalongkorn University, Bangkok, Thailand.
- Peiros, I. 1998. *Comparative Linguistics in Southeast Asia*. Canberra: Pacific Linguistics.
- Poloni, E. S., N. Ray, S. Schneider and A. Langaney. 2000. 'Languages and genes: Modes of transmission observed through the analysis of male-specific and female-specific genes', in J.-L. Dessalles and L. Ghadakpour (eds.), *Proceedings: Evolution of Language, 3rd International Conference 3–6 April 2000*. Paris: École Nationale Supérieure des Télécommunications, pp. 185–186.
- Poloni, E. S., O. Semino, G. Passarino, A. S. Santachiara-Benerecetti, I. Dupanloup, A. Langaney and L. 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).
- Pulleyblank, E. G. 1983. 'The Chinese and their neighbours in prehistoric and early historic times', in D. N. Keightley (ed.), *The Origins of Chinese Civilization*. Berkeley: University of California Press, pp. 411–466.
- Ramstedt, G. J. 1957 [posthumous]. *Einführung in die altaische Sprachwissenschaft: Lautlehre* (Suomalais-Ugrilaisen Seuran Toimituksia, Mémoires de la Société Finno-Ougrienne, 104, 1). Helsinki: Suomalais-Ugrilaisen Seura.
- Reid, L. A. 2005. 'The current status of Austric: A review and evaluation of the lexical and morphosyntactic evidence', in L. Sagart, R. Blench and A. Sanchez-Mazas (eds.), *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon, pp. 132–160.
- Relandus, H. [Adriaan van Reeland]. 1706, 1707, 1708. *Dissertationum Miscellanearum, Pars Prima, Pars Altera, Pars Tertia et Ultima*. Trajecti ad Rhenum: Gulielmus Broedelet.
- Robbeets, M. 2014. 'The Japanese inflectional paradigm in a Trans-Eurasian perspective', in M. Robbeets and W. Bisang (eds.), *Paradigm Change in the Trans-Eurasian Languages and Beyond*. Amsterdam: John Benjamins, pp. 197–232.

- Rootsi, S., L. A. Zhivotovsky, M. Baldovič, M. Kayser, I. A. Kutuev, R. Khusainova, M. A. Bermisheva, M. Gubina, S. A. Federova, A.-M. Ilumäe, E. K. Khusnutdinova, M. I. Voevoda, L. P. Osipova, M. Stoneking, A. A. Lin, V. Ferak, J. Parik, T. Kivisild, P. A. Underhill and R. Villems. 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.
- Schlegel, G. 1901. Review: 'Elements of Siamese Grammar by O. Frankfurter, Ph.D., Bangkok: Printed at the American Presbyterian Mission Press, Leipzig, Karl W. Hiersemann, 1900', *T'oung Pao* (Série II), 2: 76–87.
- Schlegel, G. 1902. *Siamese Studies* (T'oung Pao, New Series II, Volume II, Supplement). Leiden: Henri Cordier and Gustave Schlegel.
- Schmidt, W. 1906. 'Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens', *Archiv für Anthropologie, Neue Folge*, 5: 59–109.
- Starosta, S. 2005 [posthumous]. 'Proto-East-Asian and the origin and dispersal of languages of East and Southeast Asia and the Pacific', in L. Sagart, R. Blench and A. Sanchez-Mazas (eds.), *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon, pp. 182–197.
- Starostin, S. A. 1994. 'The reconstruction of Proto-Kiranti', *Paper Presented at the 27ème Congrès International sur les Langues et la Linguistique Sino-Tibétaines*, 14 octobre 1994, Centre International d'Études Pédagogiques à Sèvres, Sèvres, France.
- Tanaka, M., V. M. Cabrera, A. M. González, J. M. Larruga, T. Takeyasu, N. Fuku, L.-J. Guo, R. Hirose, Y. Fujita, M. Kurata, K.-I. Shinoda, K. Umetso, Y. Yamada, Y. Oshida, Y. Sato, N. Hattori, Y. Mizuno, Y. Arai, N. Hirose, S. Ohta, O. Ogawa, Y. Tanaka, R. Kawamori, M. Shamoto-Nagai, W. Maruyama, H. Shimokata, R. Suzuki and H. Shimodaira. 2004. 'Mitochondrial genome variation in eastern Asia and the peopling of Japan', *Genome Research*, 10A: 1832–1850.
- Thangaraj, K., G. Chaubey, V. Kumar Singh, A. Vanniarajan, I. Thanseem, A. G. Reddy and L. Singh. 2006a. 'In situ origin of deep rooting lineages of mitochondrial macrohaplogroup M in India', *BMC Genomics*, 7: 151, doi:10.1186/1471-2164-7-151.
- Thangaraj, K., V. Sridhar, T. Kivisild, A. G. Reddy, G. Chaubey, V. Kumar Singh, S. Kaur, P. Agrawal, A. Rai, J. Gupta, C. B. Mallick, N. Kumar, T. P. Velavan, R. Suganthan, D. Udaykumar, R. Kumar, R. Mishra, A. Khan, C. Annapurna and L. Singh. 2006b. 'Different population histories of the Mundari- and Mon-Khmer-speaking Austro-Asiatic tribes inferred from the mtDNA 9-bp deletion/insertion polymorphism in Indian populations', *Human Genetics*, 119 (1–2): 223, 224.
- Thanseem, I., K. Thangaraj, G. Chaubey, V. Kumar Singh, L. V. K. S. Bhaskar, B. M. Reddy, A. G. Reddy and L. Singh. 2006. 'Genetic affinities among the lower castes and tribal groups of India: Inference from Y chromosome and mitochondrial DNA', *BMC Genetiocs*, 7: 42, doi:10.1186/1471-2156-7-42.

- Underhill, P. A., N. M. Myres, S. Rootsi, M. Metspalu, L. A. Zhivotovsky, R. J. King, A. A. Lin, C.-E. T. Chow, O. Semino, V. Battaglia, I. Kutuev, M. Järve, G. Chaubey, Q. Ayub, A. Mohyuddin, S. Q. Mehdi, S. Sengupta, E. I. Rogaev, E. K. Khusnutdinova, A. Pshenichnov, O. Balanovsky, E. Balanovska, N. Jeran, D. H. Augustin, M. Baldovič, R. J. Herrera, K. Thangaraj, V. Singh, L. Singh, P. Majumder, P. Rudan, D. Primorac, R. Villems and T. Kivisild. 2010. 'Separating the post-glacial coancestry of European and Asian Y chromosomes within haplogroup R1a', *European Journal of Human Genetics*, 18: 479–484.
- Underhill, P. A., G. Passarino, A. A. Lin, P. Shen, M. Mirazón-Lahr, R. A. Foley, P. J. Oefner and L. L. Cavalli-Sforza. 2001. 'The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations', *Annals of Human Genetics*, 65: 4–62.
- Underhill, P. A., P. Shen, A. A. Lin, J. Li, G. Passarino, W. H. Yang, E. Kauffman, B. Bonnè-Tamir, J. Bertranpetit, P. Francalacci, M. Ibrahim, T. Jenkins, J. R. Kidd, S. Q. Mehdi, M. T. Seielstad, R. S. Wells, A. Piazza, R. W. Davis, M. W. Feldman, L. L. Cavalli-Sforza and P. J. Oefner. 2000. 'Y chromosome sequence variation and the history of human populations', *Nature Genetics*, 26: 358–361.
- van Driem, G. 2005. 'Sino-Austronesian vs. Sino-Caucasian, Sino-Bodic vs. Sino-Tibetan, and Tibeto-Burman as default theory', in Y. P. Yadava, G. Bhattarai, R. R. Lohani, B. Prasain and K. Parajuli (eds.), *Contemporary Issues in Nepalese Linguistics*. Kathmandu: Linguistic Society of Nepal, pp. 285–338.
- van Driem, G. 2007. 'Austroasiatic phylogeny and the Austroasiatic homeland in light of recent population genetic studies', *Mon-Khmer Studies*, 37: 1–14.
- van Driem, G. 2008. 'Reflections on the ethnolinguistic prehistory of the greater Himalayan region', in B. Huber, M. Volkart and P. Widmer (eds.), *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag* (2 vols., Vol. 1). Halle: International Institute for Tibetan and Buddhist Studies, pp. 39–59.
- van Driem, G. 2011a. 'Rice and the Austroasiatic and Hmong-Mien homelands', in N. J. Enfield (ed.), *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*. Canberra: Pacific Linguistics, pp. 361–389.
- van Driem, G. 2011b. 'The Trans-Himalayan phylum and its implications for population prehistory', *Communication on Contemporary Anthropology*, 5: 135–142.
- van Driem, G. 2012a. 'The ethnolinguistic identity of the domesticators of Asian rice', *Comptes Rendus Palevol*, 11 (2): 117–132.
- van Driem, G. 2012b. 'Etyma, shouldered adzes and molecular variants', in A. Ender, A. Leemann and B. Wälchli (eds.), *Methods in Contemporary Linguistics*. Berlin: Mouton de Gruyter, pp. 335–361.
- van Driem, G. 2014a. 'Trans-Himalayan', in N. Hill and T. Owen-Smith (eds.), *Trans-Himalayan Linguistics*. Berlin: Mouton de Gruyter, pp. 11–40.

- van Driem, G. 2014b. 'A prehistoric thoroughfare between the Ganges and the Himalayas', in T. Jamir and M. Hazarika (eds.), *50 Years after Daojali-Hading: Emerging Perspectives in the Archaeology of Northeast India*. New Delhi: Research India Press, pp. 60–98.
- Virey, J.-J. 1801 [An IX]. *Histoire naturelle du genre humain, ou Recherches sur ses principaux Fondemens physiques et moraux; précédées d'un Discours sur la nature des êtres organiques, et sur l'ensemble de leur physiologie. On y a joint une dissertation sur le sauvages de l'Aveyron* (deux tomes). Paris: F. Dufart.
- von Klaproth, J. H. 1823. *Asia Polyglotta*. Paris: A. Schubart.
- Wen, B., L. Hui, L. Daru, S. Xiufeng, Z. Feng, H. Yungang, L. Feng, G. Yang, M. Xianyun, Z. Liang, Q. Ji, T. Jingze, J. Jianzhong, H. Wei, R. Deka, S. Bīng, R. Chakroborty and J. Lì. 2004. 'Genetic evidence supports demic diffusion of Han culture', *Nature*, 431 (7006): 302–305.
- Witsen, N. 1692. *Noord en Oost Tartarye, ofte Bondig Ontwerp van eenige dier Landen en Volken, welke voormaels bekend zijn geweest, beneffens verscheide tot noch toe onbekende, en meest nooit voorheen beschreven Tartersche en Nabuurige Gewesten, Landstreeken, Steden, Rivieren, en Platen, in de Noorder en Oostelykste Gedeelten van Asia en Europa* (2 vol.). Amsterdam: François Halma.
- Wood, E. T., D. A. Stover, C. Ehret, G. Destro-Bisol, G. Spedini, H. McLeoid, L. Louie, M. Bamshad, B. I. Strassmann, H. Soodyall and M. 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.
- Wulff, K. 1934. *Chinesisch und Tai: Sprachvergleichende Untersuchungen*. Copenhagen: Levin & Munksgaard.
- Wulff, K. 1942 [posthumous]. *Über das Verhältnis des Malay-Polynesischen zum Indochinesischen*. Copenhagen: Munksgaard.