

Санкт-Петербургский государственный университет

ПРОБЛЕМЫ КИТАЙСКОГО И ОБЩЕГО ЯЗЫКОЗНАНИЯ

К 90-летию С. Е. Яхонтова

Санкт-Петербург
2016

ББК 81(5Кит)
УДК 811.581
П78

Ответственный редактор:
канд. филол. наук Елена Николаевна Колпачкова

Редакционная коллегия:
доктор филол. наук О. И. Завьялова,
доктор филол. наук В. Б. Касевич,
канд. филол. наук Н. С. Яхонтова

П78 Проблемы китайского и общего языкознания. К 90-летию С. Е. Яхонтова / Отв. ред. Е. Н. Колпачкова. — СПб.: Изд-во «Студия «НП-Принт», 2016. — 656 с.

ISBN 978-5-901724-27-9

Сборник подготовлен в связи с 90-летием со дня рождения выдающегося лингвиста, одного из основателей отечественной школы китайского языкознания Сергея Евгеньевича Яхонтова. В издание вошли избранные статьи из научного наследия юбиляра, ранее разбросанные по различным изданиям, по большей части малотиражным, и практически недоступные до сих пор читателю, а также работы коллег, последователей и учеников, продолжающих и развивающих основные направления исследований С. Е. Яхонтова. Тематика статей отражает необычайно широкий круг научных интересов ученого, включающий проблемы фонетики, грамматики, диалектологии, сравнительно-исторического и общего языкознания, лингвистической типологии.

ББК 81(5Кит)
УДК 811.581

ISBN 978-5-901724-27-9

© Редколлегия, авторы, 2016
© СПбГУ, 2016

George van Driem

The Eastern Himalayan Corridor in Prehistory

The greater Himalayan region, including the Tibetan plateau in the north and the Gangetic plain in the south, served as the principal prehistoric thoroughfare for the peopling of East and Southeast Asia. The descendants of ancient migrants through this region ultimately settled lands as far away as New Zealand, Madagascar, Lapland and the Americas. Several of the keys to understanding the ethnogenesis of human diversity in Asia include the Father Tongue correlation, possible refugia during the Last Glacial Maximum and the hypothesis that language families may have arisen as the result of demographic bottlenecks in prehistory. Ethnolinguistically informed inferences based on Asian Y chromosomal phylogeography permit a reconstruction of episodes of ethnolinguistic prehistory which lie beyond the linguistic event horizon, i. e. beyond the time depth empirically accessible to historical linguistics. The origins of the language families which make up the hypothetical Uralo-Siberian and East Asian linguistic phyla are argued to have lain in the Eastern Himalayan corridor. Several other Asian language families are shown to be tied to the Indian subcontinent. The Centripetal Migration model, which assumes that migrations in quest of a better life unfolded in both centrifugal and centripetal directions with respect to technologically more advanced centres of civilisation, is opposed to the Farming Language Dispersal theory, which assumes that all linguistic dispersals were driven by agricultural centrifugal migration¹.

¹ This contribution in honour of Sergej Jevgen'evič Jakhontov is a reworked version of a piece which appeared in 2014 under the title 'A prehistoric thoroughfare between

1. Father tongues and our paternal ancestries

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 spoken and the most salient maternally inherited markers found in that speech community. This Father Tongue correlation was already described by [Poloni et al. 1997, 2000] before the appearance of the seminal articles on

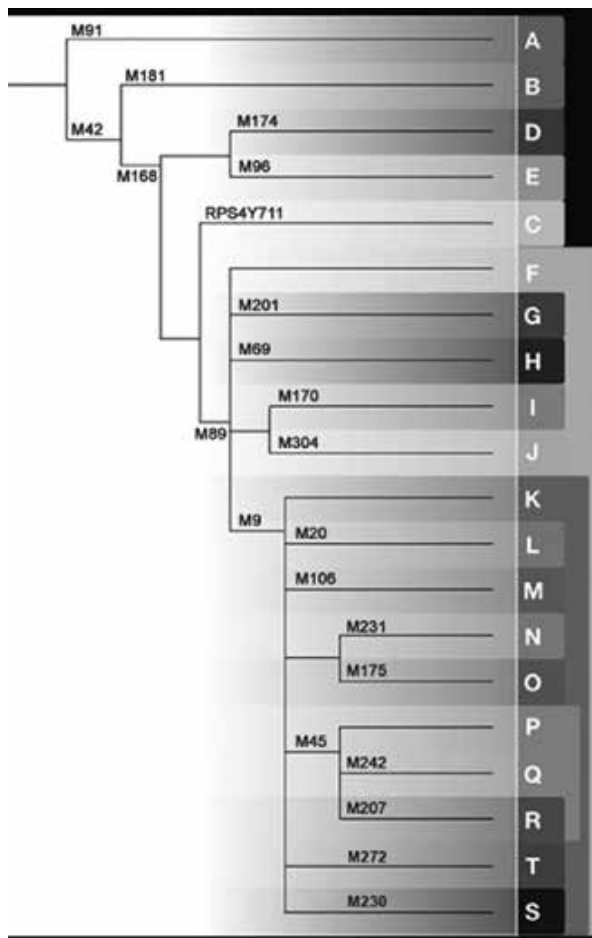


Figure 1. The Y chromosomal haplogroup tree based on [Underhill et al. 2000, 2001; Karafet et al. 2008] as it currently appears on the website www.familytreedna.com (15 May 2013)

Y chromosomal phylogeography by [Underhill et al. 2000, 2001]. Subsequent work, e. g. [Karafet et al. 2008], has further refined the resolution of the Y chromosomal haplogroup tree and led to the paternal phylogeny for mankind depicted in Figure 1.

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

the Ganges and the Himalayas' in Tiatoshi Jamir and Manjil Hazarika (eds.) 50 Years after Daojali-Hading: Emerging Perspectives in the Archaeology of Northeast India. New Delhi: Research India Press, pp. 60–98.

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. 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 correlations between maternal lineages and linguistic phylogeography discerned to date have been underwhelming. The Father Tongue correlation observed in many parts of the globe 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. A factor which may have played a role in many of these sexually asymmetrical migrations is what eminent Estonian geneticist Toomas Kivisild at Cambridge has described with grave jocularity as a sex-specific pathology linked to the Y chromosome, i. e. warfare.

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 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. 2006d].

Previously, it has been proposed that the subclades of the Y chromosomal haplogroup R (M207) are connected with the dispersal of the ancient Indo-Europeans, haplogroup O2a (M95) with the spread of Austroasiatic and O3a3c (M134) with Trans-Himalayan a.k.a. Tibeto-Burman [van Driem 2002, 2007, 2012b]. Molecular genetic findings shed light both on ethnolinguistic prehistory and its unrecorded sociolinguistic dimensions, and often

population geneticists find molecular corroboration of what linguists and ethnographers have been claiming for centuries. Yet correlations should not be confused with identity. The correlation of a particular genetic marker with the distribution of a certain language family should not be simplistically equated with populations speaking languages of a particular linguistic phylum.

Historical linguistics and human population genetics present two distinct windows on the past. 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. Yet the probabilistic basis for possible correlations between the genetic markers and the language of a speech community lies in the fact that genes are invariably inherited by offspring from their parents, whereas languages are in most cases, but not invariably, inherited by offspring from either or both of their parents.

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 have been discussed elsewhere [van Driem 2012b]. 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]. 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. 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].

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 appears to correspond 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 as the 8th century by men who came from the Near East, although the wholesale conversion of Baltistan to Islam is held to have begun only in the 14th century. The most prevalent mitochondrial DNA lineages amongst

the Baltis are shared with other Tibetan communities, whereas the prevalent Y chromosomal haplogroups probably entered Baltistan during the introduction of Islam [Zerjal et al. 1997; Quintana-Murci et al. 2001; Qamar et al. 2002]¹.

At the same time, a disconnect is sometimes observed between a highly salient genetic marker and the linguistic affinity of a community's language. Hungarians lack the TatC deletion defining the Y chromosomal haplogroup N3 (Tat) notwithstanding the prevalence of this marker amongst Uralic language communities [Li et al. 1999], suggesting that those who introduced the Hungarian language to Pannonia left no prominent genetic signature. This lineage died out, or a resident population adopted the language of newcomers without undergoing a replacement of their paternal lineages. In fact, the ancestrally preponderant Hungarian paternal lineage might already have been lost or greatly diluted by the time that the Hungarians reached Pannonia, for the case of the Hungarian language is geographically analogous to that of Ossetian. Hungarian forms the easternmost Uralic clade together with the closely related languages Khanti and Mansi, formerly known as Ostyak and Vogul. Yet in terms of its geographical position Hungarian is the most southwestern Uralic language, just as Ossetian, the westernmost Iranian language, is phylogenetically a member of the Eastern Iranian subgroup.

Such cases underscore the fact that the linguistic ancestors of a language community are not precisely the same set of people as the biological

¹ Has female irony on the part of successive generations of Balti mothers been preserved in the Balti practice of calling their language བཤཀོང་ *phaskat* 'father tongue', their homeland བཤུལ་ *phayul* 'father-realm' and their birthplace བཤམ་ *phasa* 'fatherland'? [Sprigg 2002, Biellemeier, forthcoming]. Bettina Zeisler takes objection to the historically attested and recurrent phenomenon of an incursive population of male migrants availing themselves of the 'womenfolk in place'. Her fully understandable aversion leads her to construct a contrary narrative, whereby the Tibetic languages of Baltistan are viewed as being not as phonologically conservative as they, in point of fact, very much are. She furthermore suggests that instead the language spoken in Baltistan 'might have' been a 'Dardic' language, 'an Iranian language, Burushaski or perhaps even a Turkic language', but asserts that 'they certainly did not speak Tibetan'. One of the problems with Zeisler's 'alternative explanation' hinges upon the set of referents denoted by her use of the pronoun 'they' in the latter assertion and to the particular slice of time during which these referents are supposed to have lived, and where. The population genetic data show that, whether or not 'Balti is the original language of Baltistan', the current population of Baltistan share their mitochondrial ancestry with other Tibetic speaking populations on the Tibetan plateau, whilst the predominant Y chromosomal lineages in Baltistan are likely to be correlated with the historical introduction of Islam from the Near East and that these paternal lineages were not, contrary to the scenario proposed by Zeisler, introduced by 'Amdo speaking soldiers' [pace Zeisler 2005: 53–57, 2009: 88, 2016: 235–236].

ancestors of that community. Moreover, 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]. The antiquity of Y chromosomal haplogroups such as O1 or O2 has been estimated to be greater than 10,000 years old [Yan et al. 2011]. Yet historical linguists generally estimate the linguistically reconstructible past to be shallower than 10,000 years, and this temporal gap must temper and inform all speculations regarding correlations between linguistic and genetic affinity.

2. The linguistic event horizon and beyond

Early human populations outside of Africa no doubt must have interacted with each other in parapatric or sympatric modes at various junctures in their long prehistory. There were numerous waves of peopling, and the genetic evidence is compatible with sustained and multiple migrations out of Africa through the Levant over time. Evidence has been adduced of gene flow from the now extinct Denisovans into the ancestors of the people who ultimately settled Melanesia [Reich et al. 2010], and Neanderthals evidently introduced a minor paternal contribution into the ancestors of all non-Africans at the time that these populations had emerged from Africa but before Eurasian groups had diverged from each other, whilst modern humans have apparently retained no Neanderthal maternal lineages [Green et al. 2010; Rasmussen et al. 2011; Currat and Excoffier 2011]. Despite this paternal contribution, it has been suggested that our Cro-Magnon ancestors may have outcompeted the Neanderthals because of a more finely honed language aptitude, or because religious belief systems and a wrathful God may have made Cro-Magnon the more hostile adversary [van Driem 2001].

All Australian Y chromosomal lineages belong to either haplogroups C and F, both of which left Africa between 75,000 to 62,000 years ago. All Australian mitochondrial DNA lineages fall within the founder branches M and N. A study of M42 coding region sequences in the mitochondrial lineage of Indian and Australian aboriginal populations supported the colonisation of Australia via a southern littoral route at this time depth [Kumar et al. 2007, Hudjashov et al. 2007; Stanyon et al. 2009; Rasmussen et al. 2011]. Australian maternal lineages are most closely related to those of New Guinea and Melanesia and reflect the same Palaeolithic colonisation event some 50,000 years ago. It comes as no surprise that the deep branching of both the maternal and the paternal lineages of Australian populations vis-à-vis other popu-

lations around the Indian Ocean shows long-term isolation after initial settlement. Genomic evidence has also been detected for some secondary gene flow at the time of the Sahul land bridge between New Guinea and Australia some 8,000 years ago.

[Huxley 1870] once proposed a link between populations of India and the ‘Australoid type’. However, the impressionism of early physical anthropology based on the phenotypical observations contrasts somewhat with the lack of uniquely shared haplogroups between India and Australia. Yet this is not to deny that the ancestors of the Australians must at one point in the distant past have passed through the Indian subcontinent and lived there for generations. Does the linguistic picture record for us where the maritime migration of the first Australians made landfall? Seven-eighths of the continent is covered by Pama-Nyungan languages, whilst the northwestern region of Australia shows a diversity of language families, marking this area as the probable hearth for the peopling of the Australian continent.

The set of ancestral Y chromosomal haplogroups CT (M168) encompasses a myriad of modern paternal lineages which first emerged from Africa in Palaeolithic times, first branching into the paternal lineages DE (YAP) and CF (P143). The paternal lineage DE (YAP) split into haplogroups D and E. The paternal lineage D might even have originated in the Himalayan region, where this haplogroup is still represented in the highest diversity, particularly in Nepal and Tibet. This paternal lineage migrated southward to the Andamans and eastward from the eastern Himalayas across the Tibetan plateau through what today is southern China, giving rise to the offspring clades D1 (M15), D2 (M55) and D3 (P47), and ultimately reaching the Japanese archipelago, where this paternal lineage is represented by the D2 (M55) subclade [Xue et al. 2006; Shi et al. 2008]. The mtDNA clades M31 and M32, specific to the Andaman Islands, also suggest a rapid Pleistocene dispersal along the Indian littoral with maritime expansion to the Andamans [Thangaraj et al. 2005, 2006b, 2006c; Tamang and Thangaraj 2012]. By contrast, the fraternal clade E remained principally in Africa, though this lineage also occurs at a very low frequency in western Eurasia as far as east as India and Central Asia (Figure 2)¹.

The other early Out of Africa paternal lineage CF (P143) gave rise to Y chromosomal haplogroup C (RPS4Y711) and a myriad of other haplogroups characterised by the shared innovation M89, i.e. the paternal lineages

¹ Such ancient African signatures must be distinguished from recent small migrations from Africa to the Indian subcontinent, such as the case of the Siddis [Shah et al. 2011].

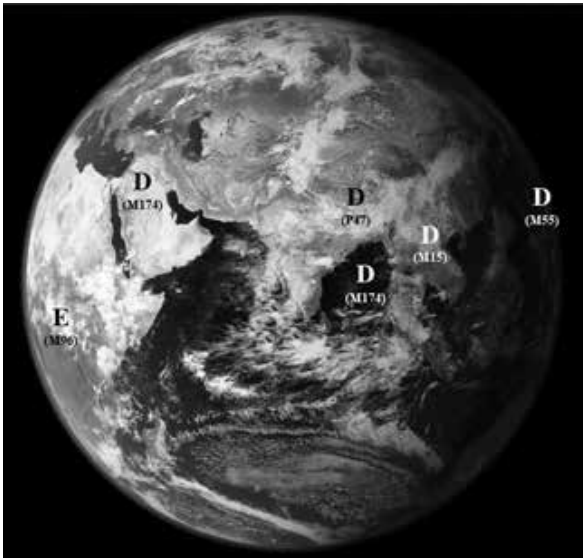
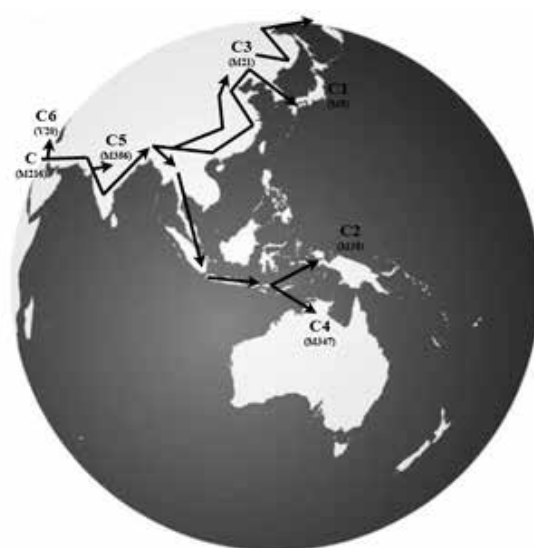
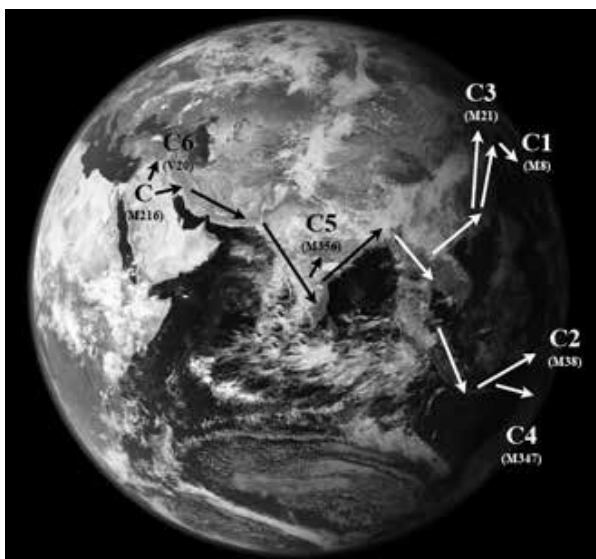


Figure 2. The paternal lineage DE (YAP) split into haplogroups D and E. Whereas Y chromosomal haplogroup D is found mainly outside of Africa, haplogroup E remained in Africa

F through T. The paternal lineage C first entered the Indian subcontinent, where the lineage C5 (M356) may have branched off. The spread of the paternal lineage C may to some extent have moved in tandem with the spread of the paternal lineage D. Ancient populations bearing the Y chromosomal haplogroup C colonised Australia and New Guinea, where the lineages C2 (M38), C4 (P309, M347) and C6 (P55) are found, and moved into northeast Asia and the Americas, where the lineages C1 (P122) and C3 (P44, M217) are found. Both the paternal haplogroups C (RPS4Y711) and D (M174) represent paternal lineages which colonised East Asia before the

Ice Age, which lasted from 21,000 to 18,000 years ago. These ancient populations underwent a Palaeolithic expansion between 34,000 and 22,000 years ago and are therefore believed to have exploited the megafauna of Siberia and East Asia as a food source (Figures 3 and 4).

At a later time depth, but still at a time inaccessible to historical linguistics, the first anatomically modern humans to populate Japan bore the



Figures 3 and 4. The paternal lineage C (RPS4Y711) ventured outside of Africa

mitochondrial haplogroup R30 and the Y chromosomal haplogroup D2 to the archipelago. The paternal subclade D2 is specific to Japan, but immediately related to this clade are the ancient D lineages preserved on the Andaman Islands and in the Himalayan region. The highest frequency of D is retained in Japan amongst the Ainu and the Ryūkyūan populations [Hammer et al. 2006], and this paternal lineage accounts for over a third of Japanese paternal lineages. Both the mitochondrial lineage R30 and the Y chromosomal haplogroup D2 indicate that this first wave of peopling of Japan originated in the Indian subcontinent at a time depth of perhaps 25,000 years ago. The paternal haplogroup C is represented as a minor lineage in Japan in a frequency of over 8%. These ancient lineages appear to represent the first wave of peopling of Japan, and the culture of their bearers later surfaced in the archaeological record from the tenth millennium BC onward as the mesolithic Jōmon. The Jōmon people were adept ancient potters who subsisted on hunting and coastal foraging and may have practised rudimentary forms of plant husbandry. The Ainu language probably represents a linguistic legacy of the original Jōmon population. The paternal lineages C and D, representing vestiges of this early wave of Palaeolithic hunter-gatherers, have also been preserved on the Korean peninsula [Jin et al. 2009]

The distinct waves of peopling reaching Japan are equally reflected in the maternal lineages. Mitochondrial haplogroup M7 has a southern distribution in East Asia, especially in the Yellow Sea littoral. Its daughter groups M7a and M7b2, specific to Japanese and Korean populations, attest to an ancient contribution to the modern Japanese mitochondrial DNA pool. The estimated coalescence times for the subclades M7a, M7b, and M7c range between 6,000 and 18,000 years. This date suggests either that these star-like clades reflect a resettlement process around the Sea of Japan from the south after the Last Glacial Maximum, contemporary with the spread of microblades of the Suyanggae type and before the onset of the Jōmon culture, or that M7a and M7b entered Japan during initial settlement over 30,000 years ago and underwent a genetic bottleneck at the time of the Last Glacial Maximum. By contrast, the mitochondrial haplogroups A5, B5, C, F1a, N9a, and Z, which are shared between Koreans and Japanese and virtually absent in Ryūkyūans and in the Ainu, testify to later migrations through the Korean peninsula to Japan, probably during the Yayoi agricultural intrusion 2,300 years ago. The presence of the mitochondrial lineage Y1 amongst the Ainu testifies to the migration of Siberian populations to the Japanese archipelago from the north [Kivisild et al. 2002; Tanaka et al. 2004].

3. Bottlenecks, refugia and the subcontinent as a prehistoric thoroughfare

The many paternal lineages from haplogroup F to T, which share the innovation M89, contain subclades whose time of divergence begins to come closer to the linguistically reconstructible past. Yet most of these molecular events lie well beyond the linguistic event horizon: The Y chromosomal F (M89) clade gave rise to the subclades H (M69), IJ (M429) and K (M9). The ancestral clade F* still appears to occur more often in the Indian subcontinent than elsewhere, suggesting that South Asia played a pivotal role as a prehistoric thoroughfare. The paternal lineage H (M69) is native to the Indian subcontinent and occurs in high frequencies in both low caste populations of the subcontinent as well as in the gypsies or Rroma [Rai et al. 2012]. The paternal clade IJ (M429) gave rise to the Y chromosomal haplogroups I (M170) and J (M304) (Figure 5).

The paternal clade K (M9) gave rise to the subclades L (M20), M (M106), NO (M214), P (M45), S (M230) and T (M272). The presence of the ancestral clade K* in a higher frequency in India than elsewhere again accentuates the role of the Subcontinent as an ancient staging area (Figure 6). The subclade P gave rise to the Y chromosomal haplogroups Q (M242) and R (M207), whereas the subclade NO (M214) gave rise to the haplogroups N (M231) and O (M175).

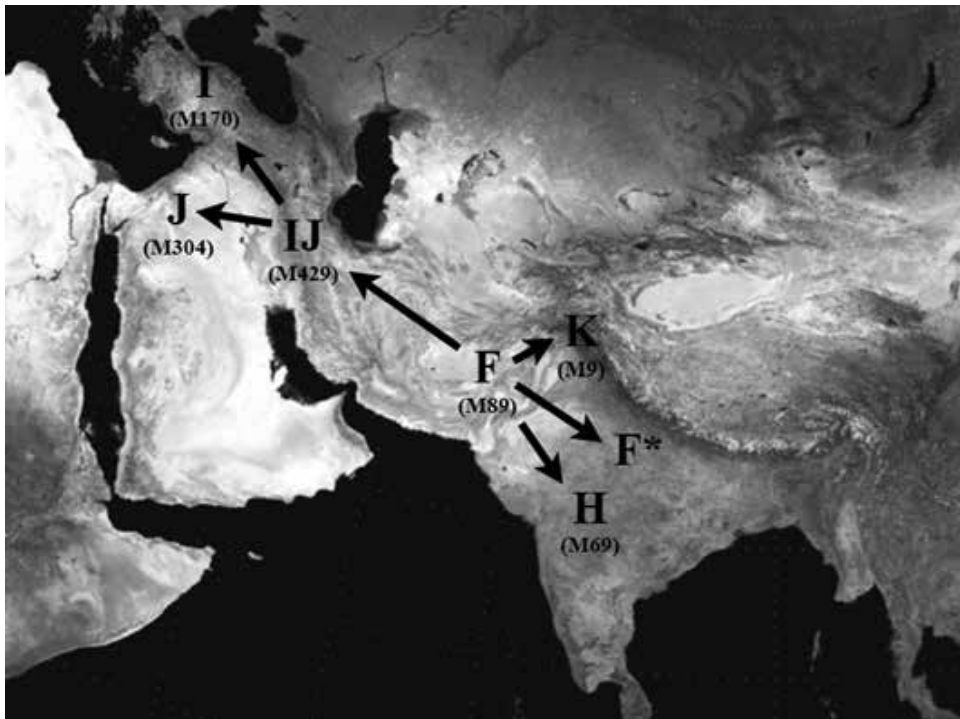


Figure 5. The split-up of paternal lineage F into the haplogroups H (M69), I (M170), J (M304) and K (M9)

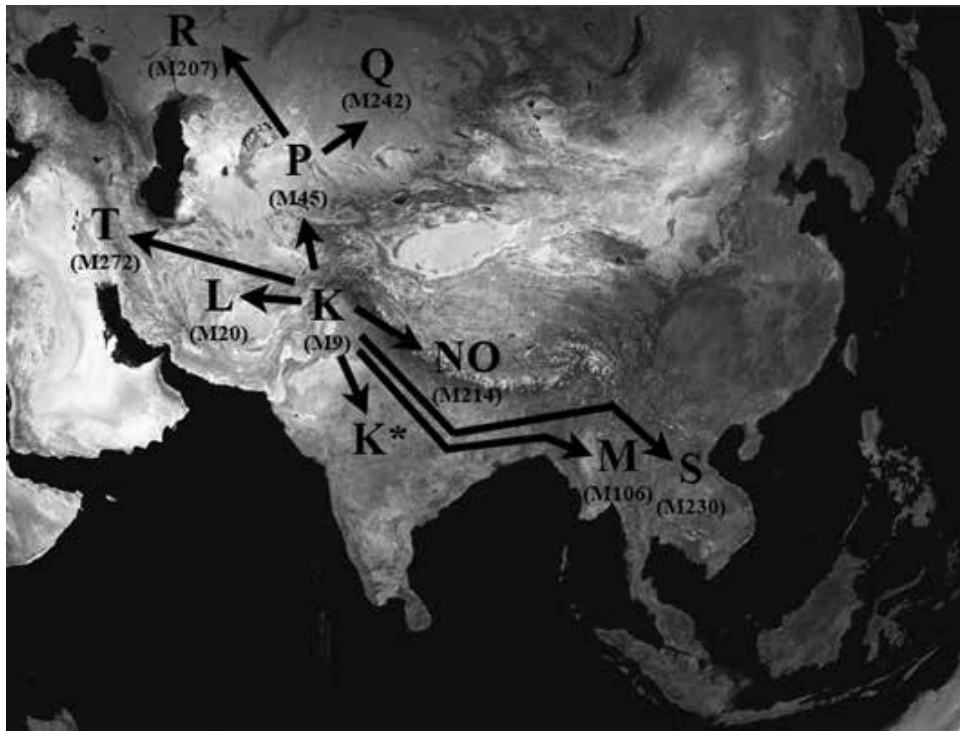


Figure 6. The split-up of paternal lineage K into the haplogroups L (M20), M (M106), NO (M214), P (M45), Q (M242), R (M207), S (M230) and T (M272)

Whilst the linguistic event horizon is an epistemological boundary beyond which historical linguistics by the comparative method is not equipped to venture, the discrete linguistic phyla recognised by linguists are themselves likely to represent the result of bottlenecks in linguistic prehistory. Not only were Palaeolithic human populations small in size, but the effective population size of any new paternal clade must have been smaller yet. Related and relevant to the phenomenon of bottlenecks, whether linguistic or genetic, is the phenomenon of Ice Age refugia. Perhaps the earliest attestation of the term ‘refugium’ used in the sense of an isolated non-glacial habitat during the Ice Age was by Canadian palynologist [Heusser 1955].

Recent phylogeographic studies [Stewart and Stringer 2012; Parducci et al. 2012] suggest that southern Tibet and the southeastern Himalayas could have harboured refuge areas for various organisms during the Last Glacial Maximum. Recent genetic studies on endemic species show that the prevalence of private haplotypes restricted to single populations could not all have evolved locally in just 14,000 years. Rather, the observed genetic diversity appears to reflect the fragmentation of once more widespread haplotypes before their isolation in refugia on the Tibetan plateau and in the Himalayas during the Last Glacial Maximum. Studies of private haplotypes of scattered juniper groves [Opgenorth et al. 2010] and endemic edaphous

beetles [Schmidt et al. 2011] as proxies for assessing moderately lower summer temperatures during the Last Glacial Maximum in southern Tibet are corroborated by studies of endemic flowering plants of the alpine steppe [Miehe et al. 2009, 2012], the Tibetan plateau pika [Ci et al. 2009], yaks [Qi et al. 2008] and monkshood [Wang et al. 2009]. Barley genotyping suggests that the Tibetan plateau may even represent the site of an independent alpine domestication of barley, involving the selective breeding for endurance to cold and drought and distinct from the domestication process which took place in the Fertile Crescent [Dai et al. 2012].

The geological and palaeontological evidence suggests that the Tibetan plateau might not yet have been entirely in the rain shadow at the time that anatomically modern humans spread across Asia, whilst the Kathmandu valley was a large palaeolake amidst verdant mountains 32,000 years ago until as recently as 15,000 years ago. The Himalayan region may very well have harboured suitable habitats for our hunter-gatherer ancestors. It has long been proposed that populations adapted to high altitude environments may not have suffered the diseases and parasites endemic to the jungles of the balmy plains [McNeill 1976]. Epidemiology may to a large extent have determined which ancient migrations were able to leave traces in today's genome and which did not. The ecological barrier between the highlands of the eastern Himalayan region and the lowlands of the Gangetic and Brahmaputran plains must have played a role in shaping population prehistory and thus the human environment.

Today's limited palaeontological survey data already clearly indicate that populations of hunter-gatherers were present on the Tibetan plateau in Palaeolithic times even though the Palaeolithic of this region is still virtually unknown [Madsen et al. 2006; Brantingham et al. 2007]. Genetic adaptations to the cold and to high altitude in populations of the Tibetan plateau are of a physiological sophistication suggestive of a long gradual evolution [Zhao et al. 2009; Yi et al. 2010; Xu et al. 2011; Wang et al. 2011; Peng et al. 2011; Qi et al. 2013]. At the same time, the ancient paternal lineage D, which is widespread in Tibet and throughout the central and eastern Himalayas attests to an ancient wave of peopling which passed through the Himalayan corridor [Qian et al. 2000]. These findings have been construed as support for the existence of cryptic refugia at high elevations during the Last Glacial Maximum.

4. East Asian and the linguistic event horizon

The East Asian linguistic hypothesis was proposed by Stanley Starosta in Périgueux in 2001, a year before he died of congestive heart failure in Hawai'i. Starosta conceived East Asian as an ancient linguistic phylum

encompassing Kradai, Austronesian, Tibeto-Burman, Hmong-Mien and Austroasiatic. Starosta was not the first to conceive of an East Asian superfamily. Once a polyphyletic view of numerous distinct Asian language families had been propounded by [Julius von Klaproth 1823], scholars began to advance proposals that might link some of these linguistic phyla together in the form of larger genetic constructs.

Gustave Schlegel in [Schlegel 1901, 1902] agreed with Klaproth in assessing Kradai to be unrelated to Sinitic, merely replete with Sinitic loans, and argued instead that Kradai was related to Austronesian. Schlegel's old theory was taken up by [Benedict 1942, 1976, 1990] under the guise of 'Austro-Thai', though this putative genetic link always constituted an ingredient in grander proposals such as Austric or 'Japanese/Austro-Tai'. Weera Ostapirat in [Ostapirat 2005, 2013] was the first to present methodologically sound and cogent historical comparative evidence that Kradai 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 Kradai language communities remained behind on the mainland. Much later, the Formosan exodus led to the spread of the Malayo-Polynesian branch throughout the Philippines, the Malay peninsula, the Indonesian Archipelago, Madagascar and Oceania.

Transgressing the linguistic event horizon, [Conrady 1916, 1922] and [Wulff 1934, 1942] each proposed a superfamily consisting of Austroasiatic, Austronesian, Kradai and Tibeto-Burman. Other than the neglect of Hmong-Mien, the mega-Austric superfamily envisaged by Conrady and Wulff already comprised all the constituents of Starosta's East Asian. [Benedict 1942], [Blust 1996] and [Peiros 1998] proposed an Austric superfamily comprising Austroasiatic, Austronesian, Kradai and possibly Hmong-Mien¹.

¹ Sino-Austronesian, staunchly and solely defended by [Sagart 1993, 2005], stands in stark contrast to these superfamilies. This highly restricted superfamily unites Sinitic and Austronesian and, more recently, 'Sino-Tibetan' and Austronesian into a single phylum. [Blust 2009: 707] writes: 'The Sino-Austronesian hypothesis is the product of an *idée fixe*'. Sagart's phylogeny of Formosan languages is rejected by [Blust 2009], [Winter 2010] and [Teng and Ross 2010]. Sagart misanalysed the Puyuma data in an attempt to assail the Nuclear Austronesian hypothesis proposed by [Ross 2009]. Nuclear Austronesian comprises all Austronesian languages other than Puyuma, Rukai and Tsou, the latter each representing primary branches of Austronesian. [Winter 2010] argues that the empirical basis for Sagart's hierarchical grouping of Formosan languages is flimsy and leads to an overly simplistic model of prehistoric migrations on Formosa. As an idea,

Starosta's evidence is meagre, yet primarily morphological in nature. Languages change so fast that traces of a genetic relationship between two languages are either obliterated or obscured beyond recognition after about a dozen millennia, give or take a few thousand years. Regular phonological correspondences and a common morphological system are the most compelling types of evidence for a genetic relationship between languages. When the time depth of a linguistic phylum is very great, morphological correspondences may be the only remaining vestige that evinces a genetic relationship.

Despite a scarcity of lexical correspondences between Itelmen a.k.a. Kamchadal and Chukchi-Koryak, the inclusion of Itelmen within a Luoravetlan a.k.a. Chukotko-Kamchatkan family is based on formal and semantic similarities in a small number of flecational morphemes. Yet such morphological evidence would necessarily be lacking altogether if the proto-language of a given family just happened to have been typologically analytical, like Mandarin is today. Most historical linguists modestly resign themselves to a maximal time depth beyond which the comparative method is unable to distinguish between correspondence and coincidence and therefore unable to establish a genetic relationship between languages. Yet a few linguists strive to gaze beyond the linguistic event horizon.

[Nichols 1986, 1992, 1995, 1998] attempts to salvage empirical evidence for deep genetic relationships between linguistic phyla and to detect the vestiges of ancient substrate influence exerted by one language phylum on another in the form of typological diagnostics. She argues that languages which are distant on a morphosyntactic spectrum ranging from wholly head-marking to wholly dependent-marking are unlikely to be genetically related. Nichols' diagnostics embolden her to speculate about ancient linguistic spread zones and bottlenecks. However, many of the world's languages exhibit both head-marking and dependent-marking morphology and are thus situated somewhere in the middle of the spectrum.

Sino-Austronesian has evolved since Sagart first explained his theory to me at the University of Hawai'i in August 1989. After abandoning his pre-1994 position that Sinitic was unrelated to Tibeto-Burman, Sagart was compelled to adhere to the outdated and empirically unsupported Sino-Tibetan family tree model, which he requires as an ingredient for his model. Sagart identifies the Middle Yǎngsháo culture in the 5th and 4th millennia BC as the Sino-Austronesian homeland based on 'regular correspondences' in the four words for pig, rice, net and millet, one of which is not reflected in Sinitic. The kindest assessment of Sagart's theory in print [van Driem 2005] evaluates his evidence as failing to meet the conventional standards of proof. Sino-Austronesian does not merit serious consideration as a hypothesis about linguistic phylogeny. Sagart's arguments for his Sino-Austronesian hypothesis have now been totally demolished by [Blust 2014].

<p style="text-align: center;">Kradai and Austronesian Gustave Schlegel (1901, 1902), Weera Ostapirat (2005)</p> <p style="text-align: center;">Austroasiatic and Austronesian Wilhelm Schmidt (1906), Lawrence Reid (1994, 2005)</p> <p style="text-align: center;">Austroasiatic, Austronesian, Kradai and Tibeto-Burman August Conrady (1916, 1922), Kurt Wulff (1934, 1942)</p> <p style="text-align: center;">Austroasiatic, Austronesian, Kradai and Hmong-Mien Paul Benedict (1942), Robert Blust (1996), Ilia Peiros (1998)</p> <p style="text-align: center;">Austroasiatic, Austronesian, Kradai, Tibeto-Burman and Hmong-Mien Stanley Starosta (2001)</p>

Figure 7. Some historical antecedents leading up to the East Asian hypothesis

In modern East Asia, many languages, such as Vietnamese, Thai, Cantonese, Mandarin, show little to no morphology and so do not occupy any place in this two-dimensional spectrum, but lie instead on a third typological axis, discerned by [Schlegel 1808], ranging from synthetic, i. e. agglutinative or flexional, to purely analytical and devoid of morphology. This widespread attrition of older morphology limits the scope for reconstructing the historical morphology of languages and linguistic subgroups which have come to be caught up in the maelstrom of what historically became the East Asian linguistic area.

[Fortescue 1998, 2011] advanced the theory of an ancient circumpolar Uralo-Siberian linguistic phylum comprising Uralic, Yukagir, Eskimo-Aleut, Nivkh and Chukotko-Kamchatkan, which he associates with Neolithic assemblages appearing across Siberia and circumpolar North America between 5000 and 3000 BC during the thermal maximum following the last Ice Age, several millennia after the land bridge across the Bering Strait had disappeared. Fortescue first pushes the comparative method as far as he can take it before introducing the notion of a ‘mesh’. Fortescue’s mesh not only denotes language families which derive from a single putative linguistic phylum at a time depth which lies beyond what is, strictly speaking, historically reconstructible, Fortescue’s mesh also represents a complex picture of language shift. Beyond the many known cases of language shift in the circumpolar region, Fortescue has sought typological traits which can be identified as the residue of language shifts in more distant prehistory. The mesh notion is invoked because the time depth of the relationship and the complexity of

language shifts which took place have left so little evidence intact that a genetic relationship is no longer demonstrable by conventional comparative means.

One of the predecessors of East Asian, [Schmidt 1906] Austric macrofamily, uniting Austroasiatic and Austronesian, was likewise based on morphological evidence drawn especially from Nicobarese. Today lexical evidence for Austric remains scarce [Diffloth 1994]. The arguments are still primarily morphological in nature, with Nicobarese still playing a star role. [Reid 1994] relates the Proto-Austroasiatic causative morphemes **<pa- ~ -ap->* and **<ka->* to the Proto-Austronesian causative prefixes **<pa->*, **<ka->* and **<paka->*, the Proto-Austroasiatic agentive marker **<ma- ~ -am->* with the Proto-Austronesian agentive **<mu- ~ -um->*, the Proto-Austroasiatic instrumental infixes **<-an->* and **<-in->* with the Malayo-Polynesian instrumental prefix **<paN->*, Proto-Austronesian instrumental morpheme **<ni- ~ -in->* and a Nancowry Nicobarese nominaliser suffix *<-a>* with a Proto-Austronesian ‘objective’ suffix **<-a>*.

The morphological evidence and its interpretation remain controversial due to the widespread nature and complexity and morphological processes involving infixation and discontinuous morphemes in Austroasiatic, including Nicobarese. By comparison with Austroasiatic, greater progress has been made in understanding the historical grammar of Austronesian, which may be both a function of the paucity of historical linguists working on Austroasiatic as well as the greater intractability and complexity of the linguistic problems confronting scholars of Austroasiatic historical grammar. Progress on the Austronesian side is in no small measure to the contributions of Reid himself, e. g. [Reid 2006, 2007, 2009a, 2009b, 2010]. Reid too envisaged an even larger macrofamily and contended 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].

The evidence adduced by Starosta for East Asian, though meagre, is morphological in nature. The ancient morphological processes shared by the families of this phylum 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*. By contrast, the structure of his family tree comprising Kradai, Austronesian, Tibeto-Burman¹, Hmong-Mien and Austroasiatic was based on

¹ Starosta accepted the Sino-Bodic hypothesis and rejected the Sino-Tibetan model. Due to an editorial error, the label ‘Sino-Tibetan’ appears in the posthumous version of

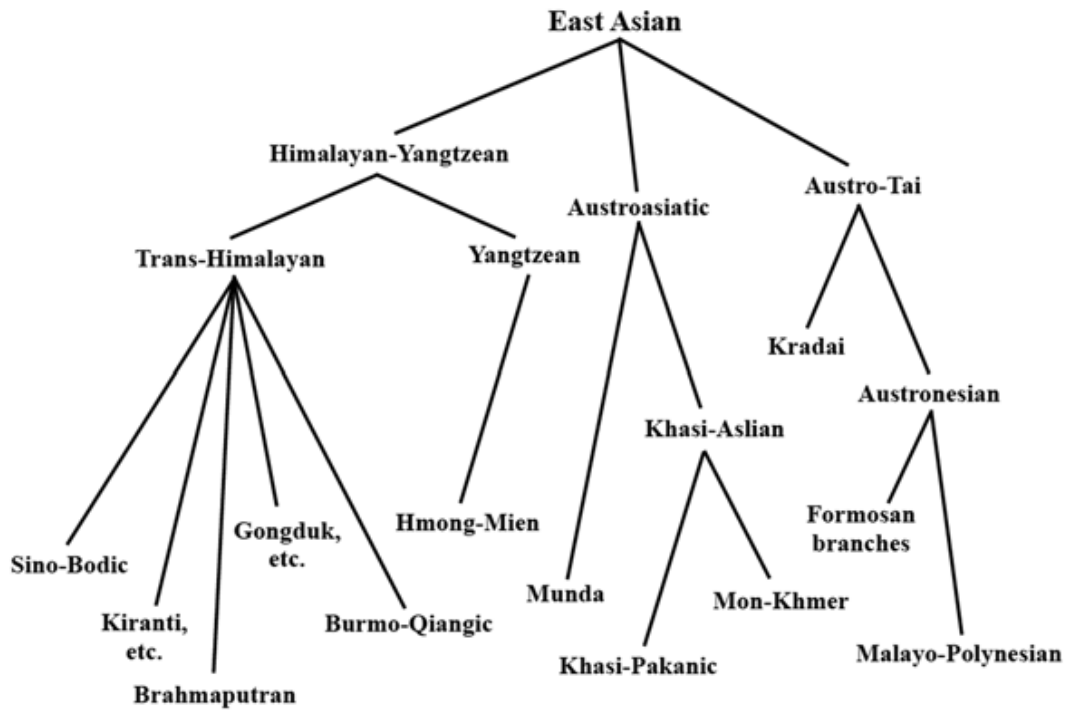


Figure 8. The 2012 Benares Recension: A revised East Asian phylogeny

sheer impressionism. Starosta wrote that the ‘potential utility’ of his hypothesis lay ‘in helping to focus scholars’ efforts on particular specific questions, resulting in the replacement of parts of this hypothesis with better supported arguments’ [Starosta 2005: 194]. At the 18th Himalayan Languages Symposium, held at Benares Hindu University in 2012, I presented the revised East Asian family tree depicted in Figure 8. The revised phylogeny is based on historical linguistic intuitions and other types of information about population prehistory.

5. Trans-Himalayan and other East Asian families

Trans-Himalayan is the world’s second most populous language family. Most speakers of Trans-Himalayan languages today live north of the Himalayas (Figure 9), but most of the over 300 different languages and three fourths

Starosta’s East Asian phylogeny [Starosta 2005: 183]. Starosta’s East Asian phylogeny, as presented at Périgueux in 2001, is reproduced correctly in [van Driem 2005: 322], rectifying the editorial misrepresentation. It remains a matter of conjecture whether the subordinate extra-Formosan status of Kradai in Starosta’s diagram might not also be a posthumous editorial enhancement, since this idea has chiefly been championed by one of the editors since Périgueux.



Figure 9. Geographical distribution of Trans-Himalayan languages¹

of the major Trans-Himalayan subgroups are located south of the Himalayan divide (Figure 11). The Trans-Himalayan linguistic phylum was first recognised by Julius von 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]. Yet confusingly, adherents of the Indo-Chinese tradition use the term ‘Tibeto-Burman’ in the sense of non-Sinitic, a putative taxon within the Indo-Chinese tree for which Sino-Tibetanists have perennially failed to adduce evidence². The Sino-Tibetanists’ tree was assailed

¹ The maps in Figures 9, 12, 13, 14, 15 and 25, drawn by Christiane Enderle, are reproduced here from [van Driem 2015] with gracious permission of George Mieke and Colin Pendry, editors of the *Flora of Nepal*.

² In 1807, the Scots amateur John Leyden proposed his exuberant but poorly informed Indo-Chinese theory to George Barlow, Governor General of India at Fort William, in which he claimed *a priori* that all the languages in Asia and Oceania shared a ‘common mixed origin’. Leyden died at the age of 35 after making a short but dazzling career in



Figure 10. Fallen Leaves. Thirty out of forty-two 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

by scholars who proposed other models, e. g. Sino-Burman [Ramstedt 1957], Sino-Himalayan [Bodman 1976, 1980] and Sino-Kiranti [Starostin 1994]. An historical account of scholarly thinking on Trans-Himalayan languages and their genetic relationships has been provided elsewhere [van Driem 2013].

The neutral geographical name Trans-Himalayan obviates the terminological confusion which arises from some scholars using the term ‘Tibeto-Burman’ in its original sense to denote the language family as a whole, whilst the Sino-Tibetanists use the term ‘Tibeto-Burman’ to denote a non-existent taxon in their empirically unsupported family tree. More Trans-Himalayan languages are known today than were recognised in Klaproth’s day. Today 42 subgroups can be identified, as shown in the updated Fallen Leaves model

the British colonial administration in Asia during the Napoleonic wars, but his hypothesis outlived him. In 1924, Indo-Chinese was renamed Sino-Tibetan by Jean Przyluski. By Przyluski’s time, Austroasiatic had been removed, but Sino-Tibetan then still comprised Sino-Daic, Hmong-Mien and a truncated version of ‘Tibeto-Burman’. This typologically and, in part, racially inspired theory played a large role in American academics after the Great Depression and in China after the Cultural Revolution [van Driem 2013].



Figure 11. Geographical distribution of the major Trans-Himalayan subgroups. Each dot represents not just one language, but the putative historical geographical centre of each of 42 major linguistic subgroups

in Figure 10. Geographical distribution provides one key to unravelling ethnolinguistic phylogeography. The geographical centre of gravity on the basis of the distribution of major Trans-Himalayan subgroups is more indicative of the location of the linguistic homeland for the phylum than the modern distribution of speech communities¹.

Vast swathes of what today is China are covered by Sinitic languages. Yet Manchuria was only sinicised after the Second World War, and much of southern China was only sinicised during the Qín dynasty beginning in the 3rd century BC. The global spread of English must be viewed in light of the fact that even an ancestral form of the language was not spoken on the British Isles until the fifth century AD. Viewed in terms of the distribution of other Trans-Himalayan subgroups, Sinitic represents a northeastward expansion toward the Yellow River basin. Tǔjiā and Bái most probably represent ethnolinguistic vestiges of the same ancient expansion. The internal phylogeny of Sinitic itself may also reflect this route of migration. The Càijiā 蔡家 language is spoken in the northwestern corner of Guìzhōu province [Bó 2004]. [Zhèngzhāng 2010] considers Càijiā to be a member of the same subgroup as Bái, whereas Sagart believes that both Càijiā and the Wǎxiāng 瓦鄉 dialect of western Húnán could

¹ The idea that the Tibeto-Burman homeland lay in the sub-Himalayan region was possibly first expressed by [Peiros 1998: 217].

represent the first sub-branches of the Sinitic subgroup to have split off from Proto-Sinitic, even before the splitting off of the Mǐn dialects [de Sousa 2012].

The idea of an Altaic substrate influence on Sinitic, and on Mandarin in particular, has been in the air for quite some time, e. g. [Poppe 1965; Ch'en 1976]. In terms of its salient typological features, [Hashimoto 1976a, 1976b, 1980, 1986] argued that Sinitic could be thought of as an ancient Tibeto-Burman language which underwent pidginisation in the mouths of an Altaic population in the northeast. Hashimoto adopted [Ballard 1979] metaphor, which depicts Sinitic as a mosaic of structural features representing a typological halfway house between Southeast and Northeast Asia. Some have responded critically, e. g. [Li 1995], whereas others have received the idea favourably, e. g. [Norman 1982; Wadley 1996]¹. In terms of gross syntactic element order, Sinitic is not unique. Mru in the Chittagong and Karen in the Tenasserim are likewise not verb-final. Did Sinitic undergo an actual process of creolisation, or was the language just subject to successive phases of extreme contact influence over the millennia as a *lingua franca* shifting its centre periodically from one capital to another? [Starostin 2008], posthumous argued that the lexical items shared between Proto-Altaic and Proto-Sinitic, but not with other Tibeto-Burman languages, indicated a not always friendly ancient contact situation which arose only after Sinitic had split off and migrated northeast to the lower Yellow River basin. Hashimoto's and Ballard's theory of Sinitic having arisen from a Tibeto-Burman creole has recently also been taken up by [Comrie 2008] and [DeLancey 2011].

The ancient eastward expansion of Tibeto-Burman to the North China Plain is likely also to have brought the Trans-Himalayan group ancestral to Sinitic into contact with ancient Proto-Austronesians on the eastern seaboard². If ancient Hmong-Mien were once a predominant group in what today is southern China, then future linguistic research could evaluate the hypothesis of ancient language contact between Hmong-Mien and Kra-dai in the vast hybrid zone south of the Yangtze, where these two linguistically distinct sets of populations interacted. In a later epoch, a new episode of language contact arose between Tibeto-Burman and the ancient Hmong-Mien and Kra-dai when Sinitic expanded southward. [Ballard 1979, 1984] argued that the Wú

¹ Wadley disbelieves the existence of a Manchu-Mandarin pidgin during the Qīng dynasty (1644–1911), but he is receptive to Hashimoto's and Ballard's central thesis regarding profound substrate influence on the formation of Sinitic at a greater time depth.

² I have suggested that Sagart's purported 'Sino-Austronesian' correspondences, if not just representing chance resemblances, could be a residue of such an early contact between Austronesian and Sinitic in the Lóngshān interaction sphere in the fourth and third millennia BC [van Driem 1998, 2005].

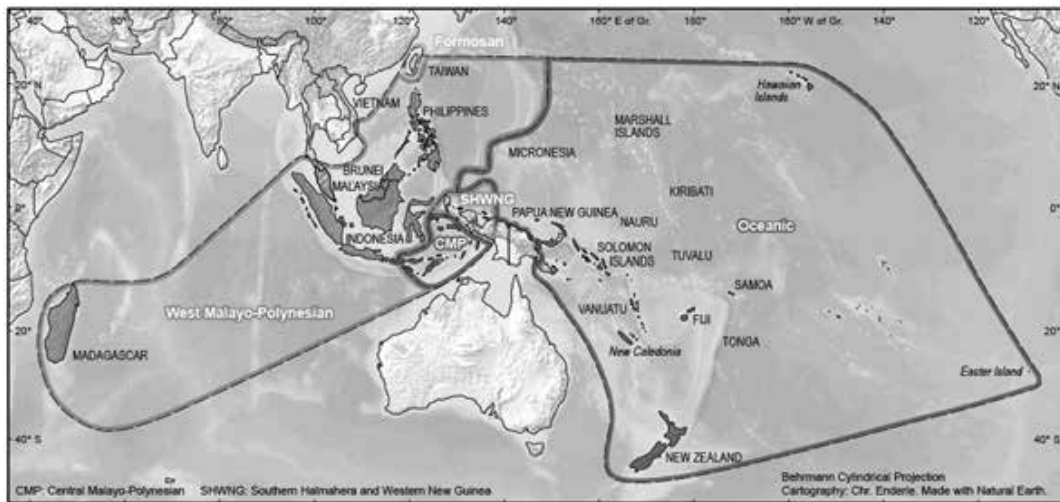


Figure 12. Geographical distribution of Austronesian

dialects arose from the expansion of Sinitic onto a Hmong-Mien substrate, and [Hagège and Haudricourt 1978: 163] proposed that Hakka likewise arose on a Hmong-Mien substrate and that Cantonese arose from the adoption of Sinitic by Kra-dai-speaking populations.

The Austronesian family is the geographically most widespread language phylum on the face of the planet, if we disregard the spread of Indo-European in the aftermath of European colonial expansion since Henry the Navigator. Linguistic, population genetic and archaeological evidence support the theory that the linguistic homeland of modern Austronesian language communities (Figure 12) lay on Formosa¹. The maritime expansion of Austronesian took place in spurts. The First Long Pause is a lull in the archaeological record, which suggests that there was little movement between 4300 and 3300 BC, before the agricultural colonisation of Taiwan began with the Dàpènkēng Neolithic. The colonisation of the Philippines was initiated ca. 2200 BC on the Batanes Islands and in northern Luzon [Bellwood and Dizon 2005]. The Second Long Pause, from 1200 to 200 BC, is a lull between the Lapita colonisation of western Polynesia and the colonisation of eastern Polynesia.

As in the case of Trans-Himalayan and Austronesian, an assessment of the geographical distribution of other East Asian linguistic phyla such as Hmong-Mien (Figure 13) and Kra-dai (Figure 14) must likewise be informed by an historical linguistic understanding of these language families. The expansion

¹ [Lemaréchal 2010] gingerly challenges the currently prevalent theory that Formosa and the Formosan languages represent the cradle of Austronesian phylum. [Peiros 2008] assails the Formosan homeland theory, claiming that Formosan subgroups collectively represent a single branch of Austronesian, and that the Austronesian *Urheimat* lay on the East Asian mainland, whence only one branch settled on Formosa.



Figure 13. Geographical distribution of Hmong-Mien



Figure 14. Geographical distribution of Kradai

of the Southwestern Tai languages into mainland Southeast Asia from the areas south of the Yangtze is an historically comparatively recent phenomenon. The expansion of Hmong-Mien groups into Southeast Asia constitutes an even more recent process. In the case of Austroasiatic (Figure 15), the spread of Vietnamese southward along the mainland Southeast Asian littoral from Tonkin, where its closest linguistic relatives Mùòng, Maleng, Chut, Arem, Aheu, Hung, Thô and Nguôn are still spoken, is also an historically comparatively recent phenomenon. Similarly, Khmer is believed to have spread at the expense of Pearic languages across the area that today is Cambodia during the Angkorian period and perhaps earlier. The deepest division within the Austroasiatic family lies between the Munda languages of India and the Khasi-Aslian languages of Southeast Asia [Diffloth 2009]. Within Khasi-Aslian, Mon-Khmer is coordinate with Khasi-Pakanic. [Diffloth 2012] has presented evidence that Pearic is a sister clade of Khmuic within Khasi-Pakanic, and is not ‘une espèce de vieux khmèr’, as has sometimes been thought (Figure 16). Pearic and Khmuic exclusively share the etymon **kloŋ* ‘cooked rice’ and other features. Pearic represents the older population of most of what today is Cambodia.

In the following sections, a new reconstruction of the prehistory of the East Asian phylum will be outlined on the basis of population genetic findings and other data and inferences about the past. Beforehand, for the



Figure 15. Geographical distribution of Austroasiatic

sake of reference and in tribute to Starosta, his reconstruction will be briefly recapitulated, though it differs from the view which will be presented here. Starosta envisaged the Proto-East-Asian ‘linkage’ or dialect continuum as

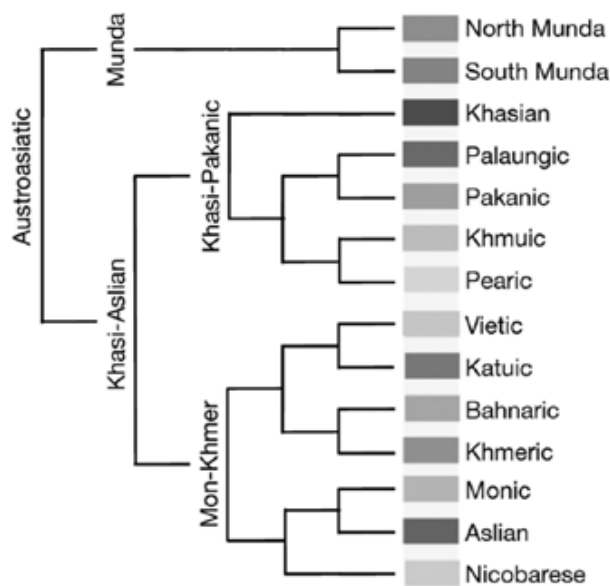


Figure 16. The phylogeny of Austroasiatic linguistic phylum [Diffloth 2009, 2012]

having lain in the region laced by the Hân, the Wèi and the central portion of the Yellow River in the period from 6500 to 6000 BC. He identified the Péilígǎng and Císhān Neolithic assemblages with East Asian, and he envisaged the linguistic ancestors of the Austronesians as the first group to have split off. He saw the Proto-Austronesians spreading to the coast and then down the eastern seaboard to establish the Hémǔdù and Dàwènkǒu Neolithic cultures of 5000 BC, ultimately to cross over to Formosa.

One migration from Formosa gave rise to the Malayo-Polynesian expansion into insular Southeast Asia, Oceania and parts of peninsular Southeast Asia, whereas another migration led back to the South China mainland, where it gave rise to Kra-dai or Daic¹. On the North China Plain, a second group, the ‘Yangtzeans’, split off and moved south and settle along the Yangtze, where they shifted from millet to rice agriculture. The Yangtzeans in turn later split up into the first Austroasiatic language communities, reflected in the Kūnmíng Neolithic of 4000 BC, and the Hmong-Mien, who appear in recorded history in what today is Húběi and northern Húnán as the Chǔ polity (770–223 BC) which challenged the Eastern Zhōu. Finally, somewhere in the central Yellow River basin, a third descendant group of East Asian remained. This third family was Tibeto-Burman. In Starosta’s conception, Tibeto-Burman split into Sino-Bodic, which he associated with the Yǎngsháo Neolithic of 5800 BC, and Himalayo-Burman, which he associated with the Dàdìwān Neolithic in Gānsù 6500 BC.

6. The population genetics of East Asian language families

The two paternal lineages N and O may have split up in the greater eastern Himalayan region. The highest frequency of the ancestral N* (M231) is still found in northern Burma, Yúnnán and Sìchūān, whilst the fraternal clade O appears to be a marker for the linguistic ancestors of the hypothetical East Asian linguistic phylum, comprising Kra-dai, Austronesian, Tibeto-Burman, Hmong-Mien and Austroasiatic. The evidence for refugia in southeastern Tibet suggests a possible putative point of origin for the expansion of the paternal lineage O. Whilst it remains a matter of speculation at this point whether or not the Tibetan plateau could have harboured refugia hospitable to human habitation during the Last Glacial Maximum, the entire southeastern and eastern declivity of High Asia furnishes numerous possible points of origin for paternal haplogroup O. Populations bearing the Y chromosomal O haplogroups colonised southeastern Eurasia, probably beginning from a locus in the eastern Himalayas. [Xue et al. 2006] speculate that the population expansion involved feeding on tubers as the climate warmed, but the early domestication of crops no doubt played a role for some subset of these ancient populations.

In some anthropological circles, it has recently become fashionable to refer to this eastern portion of highland Asia, comprising northeastern India, northern Burma, Yúnnán and Sìchūān, as ‘Zomia’. This term was coined by

¹ See footnote on p. 479.

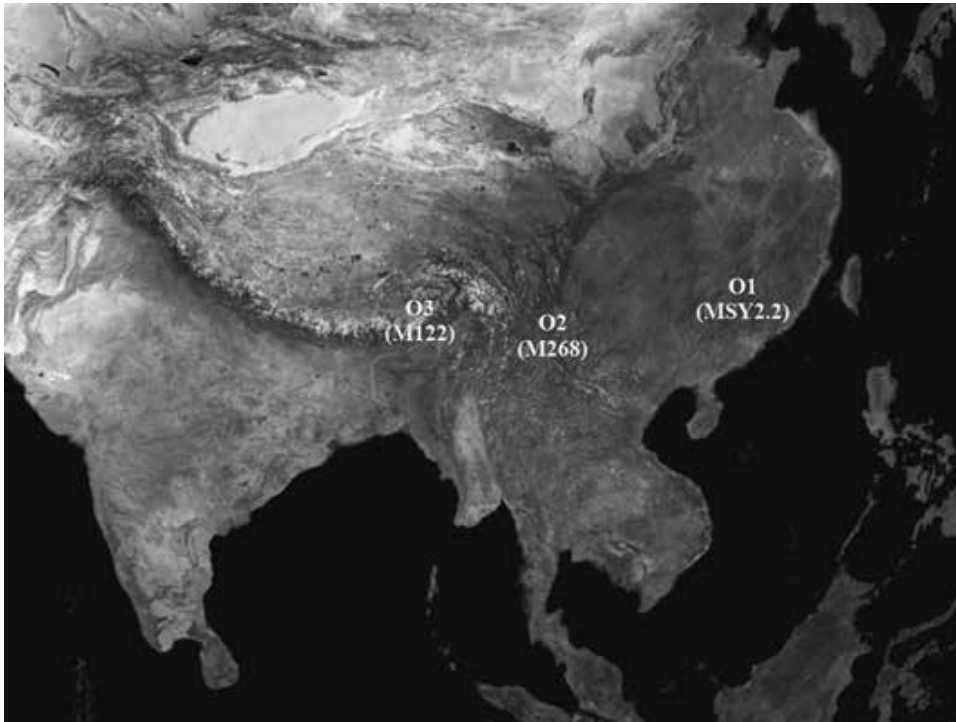


Figure 17. *After the Last Glacial Maximum, the Y chromosomal haplogroup O (M175) split into the subclades O1 (M119), O2 (M268) and O3 (M122). Bearers of the O2 (M268) paternal lineage domesticated Asian rice*

[Schendel 2002] in an obscure paper, one of the aims of which was to criticise the traditional areal disciplines in the humanities against the background of the organisational reshuffling within Dutch universities that was going on at the time, as perennially dictated from The Hague. The term Zomia was popularised by [Scott 2009], who attributes the ethnolinguistic diversity of the eastern Himalayan highlands to a tendency of ethnic minorities to head for the hills to escape the central authority of nation states. Certainly, communities have at times in the course of history fled to escape being subdued and enslaved by powerful polities, but historical linguistics and population genetics show us that the ethnolinguistic diversity of the eastern Himalayan region is of hoary antiquity, whilst the emergence of nation states is comparatively recent¹. In stark contrast to the political anthropologists' idyll of

¹ Though of little utility for understanding Asia's ethnolinguistic diversity, Scott's romantic fantasy about Zomia can be understood in terms of his palpable distaste for what [Weber 1919] called the 'Gewaltsmonopol des Staates', a concept already implicit in the writings of [Hobbes 1651]. [Toynbee 1976] described the earliest governments as arising when parasitic non-agricultural bands of brigands with superior weaponry or at least greater ruthlessness stumbled upon the idea of extorting a tax from sedentary agricultural populations in exchange for 'protection'. [Bodin 1573] pointed out that the principle of 'might makes right' essentially constitutes the difference between

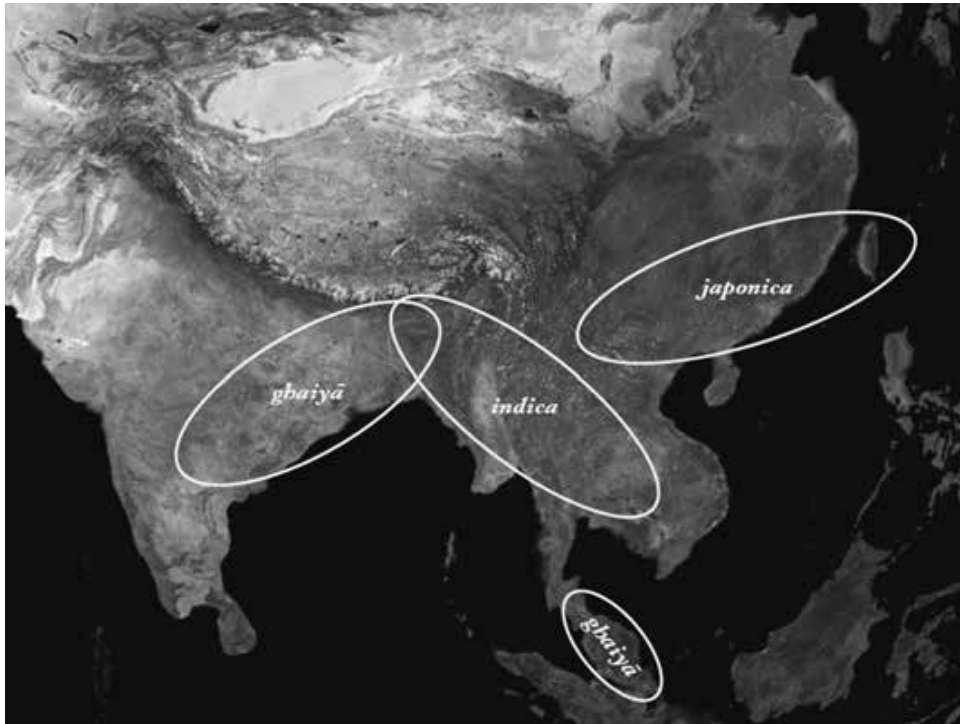


Figure 18. Geographical ranges for the possible domestication of ghaiyā or upland rice, wet indica rice and the japonica cultivar, based on the geographical distribution of genetic markers in the wild precursor *Oryza rufipogon* based on [Londo et al. 2007]

Zomia, the vast, topographically complex and ecologically diverse region of the eastern Himalayan highlands is likely to have been one of the primordial and pivotal cradles of ethnogenesis.

As temperature and humidity increased after the Last Glacial Maximum, the Y chromosomal haplogroup O (M175) split up into the subclades O1 (M119), 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 ar-

governments and bands of robbers and pirates: ‘Nous auons dit en premier lieu droit gouvernement, pour la difference qu’il y a entre les Republicques, & les troupes de voleurs & pirates... La souueraineté est la puissance absoluë & perpetuelle... la puissance de donner loy à tous, & à chacun en particulier... Sous ce mefme puissance de donner, & casser la loy, font compris tous les autres droicts, & marques de souueraineté: de forte qu’à parler proprement on peut dire qu’il n’y a que ceste seule marque de souueraineté: attendu que tous les autres droicts font compris en cestui-là comme decerner la guerre, ou faire la paix...’ (1576: 1, 125, 197, 199). [Orwell 1949] predicted that governmental transgressions against the rights of the individual and infringements upon privacy and personal freedom would ultimately get worse. In the epilogue to his two-tome history on several millennia of transfers and cultural exchanges across the Indian Ocean, [Beaujard 2012] poses the timely question as to whether the current episode of globalisation under American hegemony can possibly culminate in a benign outcome.

gument and without any claim to geographical precision. Whereas the haplogroup O1 (M119) 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 17). 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 18). 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 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 19). 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 2010; Chaubey et al. 2010]. 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 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) here¹. 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

¹ [Ferlus 1996] proposed that one of the prominent Proto-Austroasiatic etyma for rice originally denoted taro, but this argument is refuted by [Diffloth 2011], who shows that the two cultigens are reconstructible to separate roots which have been consistently distinguished throughout linguistically reconstructible Austroasiatic prehistory. The domestication of taro [Rao et al. 2010] is as important to understanding Austroasiatic prehistory as rice. Despite attempts by [Bradley 2012; Blench's 2009] claim still holds true that no rice agricultural terminology can be confidently reconstructed for Tibeto-Burman. As has long widely been presumed, the ancient Tibeto-Burmans probably first cultivated not rice, but foxtail millet *Setaria italica* and broomcorn millet *Panicum mileaceum*.

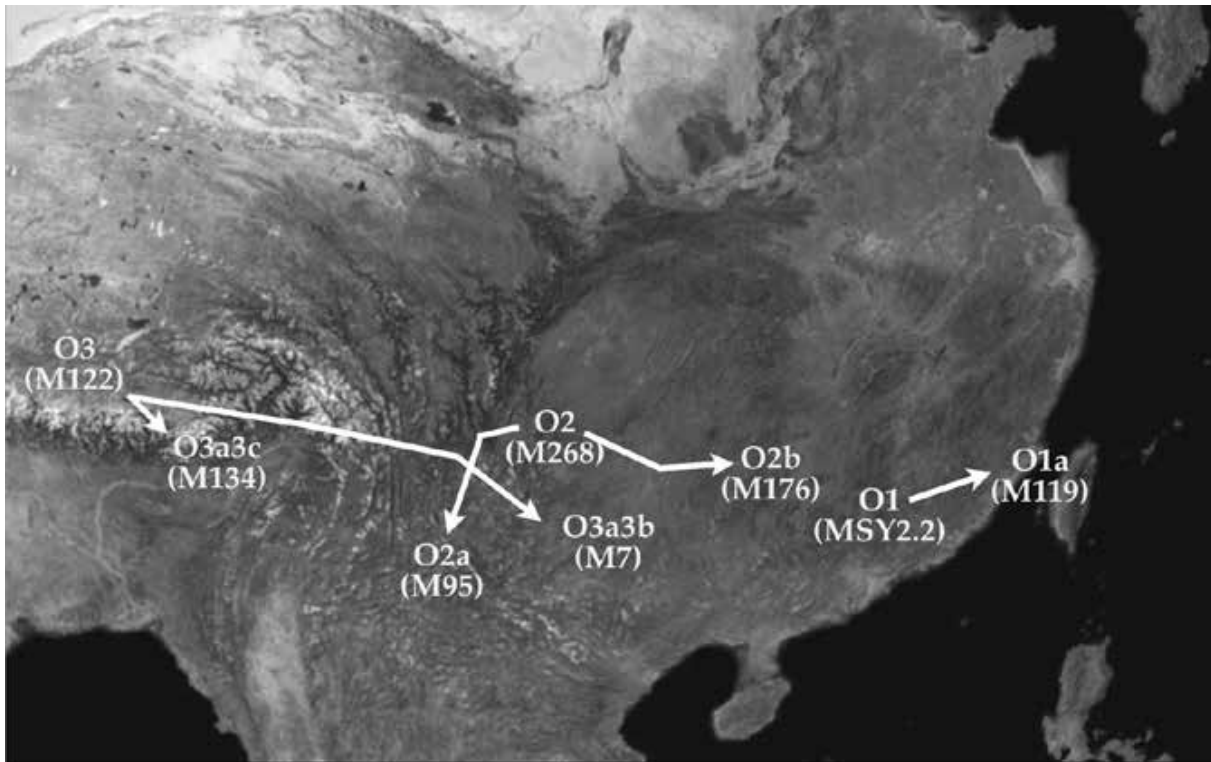


Figure 19. Paternal lineages branching into new subclades. Each event involved a linguistic bottleneck leading to language families that today are reconstructible as distinct linguistic phyla. The O1 (MSY2.2) lineage in the the Pearl River drainage gave rise to the O1a (M119) subclade, which moved eastward 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, who migrated eastward to areas south of the Yangtze. On their way, they adopted rice agriculture from the ancient Austroasiatics. In the eastern spurs of the Himalayas, 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 and leaving behind an old toponym for the Yangtze, later borrowed by Old Chinese as 江 *kʰroŋ.

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 1983], reflected as the toponym borrowed by Old Chinese as 江 *kʰroŋ (modern Mandarin: jiāng). This para-Austroasiatic paternal lineage moved as far as the Korean peninsula and represents the second wave of peopling attested in the Japanese genome [Jin et al. 2009; Karafet et al. 2009b]. Whereas the maternal lineage represented by mitochondrial subclade D4 points toward an immediate provenance on the East Asian mainland, the Y chromosomal haplogroup tells this more specific story.

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. Along the way followed by their ancestors northward toward the Korean peninsula, the earliest attested domestic millet dates from before 6000 BC at 興隆溝 *Xīnglōnggōu* near 赤峰 *Chífēng*, where a Neolithic culture without sickles has been described [Zhào 2005]. 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 presence of Y chromosomal haplogroup O2b and other O haplogroups in Japan is more recent, but accounts for more than half of all Japanese paternal lineages, with their highest frequencies in Kyūshū. The Y chromosomal haplogroup N is present only as a marginal paternal clade in Japan. If we assume that the paternal lineage N represents a marker for Altaic, then this vestige may be the tenuous genetic trace of the population who once bore an Altaic language to the Japanese archipelago which eventually evolved into modern Japanese.

At the dawn of the Holocene in the southeastern Himalayas and the eastern declivity of the Tibetan plateau, haplogroup O3 (M122) gave rise to the ancestral Trans-Himalayan or Tibeto-Burman 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 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 South-

east 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 (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 and across the strait to Formosa, which consequently became the *Urheimat* of the Austronesians, cf. [Abdulla et al. 2009]. Back west in the eastern Himalayas, the bearers of Y chromosomal haplogroup O3a3c (M134) expanded eastward into Sìchuān and Yúnnán, 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.

If we assume a linguistic dispersal in which languages were spread by populations in which a particular paternal lineage was dominant, as outlined in the scenario above, 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 or Tibeto-Burman paternal lineage O3a3c (M134) spread from the eastern Himalayas 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 of course predominantly bear 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 Himalayas, 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.

Correlations between linguistic, archaeology and genetics must inform a chronologically layered view of ethnolinguistic prehistory [Bellwood et al. 2011; van Driem 2011b]. Any attempt to span great stretches of time must distinguish numerous chronological layers. The post-glacial movements to the north and east that gave rise to the Trans-Himalayan, Austroasiatic, Hmong-Mien and Austronesian linguistic phyla took place long before the northeasterly spread of ancient Tibeto-Burmans to the putative early locus of Sino-Bodic. A distinct process was the incremental spread of diverse ancient Trans-Himalayan groups throughout the Himalayas, where linguistic and genetic evidence indicates the presence of pre-Tibeto-Burman populations.

More recent than this was the southward spread of Sino-Bodic that brought Sino-Bodic groups into contact with the ancient Hmong-Mien, early Austroasiatics, Austronesians and with other Tibeto-Burman groups. More recent yet was the Bodic spread across the Tibetan plateau spilling over into the Himalayas, as evinced by the distribution of Bodish, East Bodish, Tamangic, West Himalayish and several other groups. The spread of Tibeto-Burman groups from Yúnnán into Southeast Asia, e. g. Karen, Pyu and later Lolo-Burmese, constitutes such a recent episode that parts of these movements are historically attested. Also historically documented is the Hàn spread, well reflected in linguistics and genetics, which assimilated both other Tibeto-Burman communities as well as non-Tibeto-Burman groups. The historically documented spread of Tibetic (i. e. Bodish) across the Tibetan plateau is even more recent.

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].

At a shallower time depth, ancient mitochondrial DNA recovered in northeastern Thailand at the Bronze Age site Noen U-loke, dating from 1500

BC to 500 AD, and the Iron Age site Ban Lum-Khao, dating from 1200 to 400 BC, shows the closest affinity to the Chao-Bon, a modern Austroasiatic population who speak the Monic language Nyahkur. The Chao-Bon are directly descendant from the Mon of the Dvārāvātī kingdom, which from the 7th to the 13th century extended from the Andaman Sea near present-day Mergui in the southern Tenasserim in the west to the alluvial plains of what today is central Thailand in the east. By contrast, the majority Thai of modern Thailand show the greatest autosomal affinity with Kra-dai language communities of southern China and with the Southern Chinese Hân population, whilst the mitochondrial lineages of the Khmer show greater affinity with today's national majority Thai than do other Austroasiatic groups [Lertrit et al. 2008; Abdulla et al. 2009].

At the same time, the Y chromosomal haplogroup frequencies of Hân Píngguà dialect speakers in Guǎngxī province in a dendrogram clustering are shown to be intermediate between Hmong-Mien, Austroasiatic and Kra-dai language communities on one hand and Austronesian populations on the other [Wen et al. 2004; Gan et al. 2008]. These findings dovetail with the late historical expansion of the Thai from southern China into Southeast Asia in the first part of the second millennium as well as with their subsequent subjugation by the Khmer.

7. Other Asian language families of the subcontinent and models of migration

Human genetic studies suggest that the paternal lineages N and O may have split in East Asia between 30,000 and 20,000 years ago (Figure 20). The spatial dynamics of the two haplogroups diverged greatly after the split, and the ancient Asian populations which bore the Y chromosomal haplogroups N and O are calculated to have undergone an effective expansion between 18,000 and 12,000 years ago [Xue et al. 2006]. Ancient bearers of the N haplogroup moved north from East Asia after the Last Glacial Maximum and, in a grand counterclockwise sweep, migrated across northern Eurasia as far as west as Lapland [Rootsi et al. 2007; Derenko et al. 2007; Mirabal et al. 2009] (Figure 21). The Y chromosomal haplogroup N appears to be a marker for the linguistic ancestors of Fortescue's Uralo-Siberian linguistic phylum, comprising Uralic, Yukagir, Eskimo-Aleut, Nivkh and Chukotko-Kamchatkan.

The absence of haplogroup N in the Americas and its prevalence throughout Siberia allow us to infer that the spread of the paternal lineage N northward must have taken place after the paternal founder lineages had already

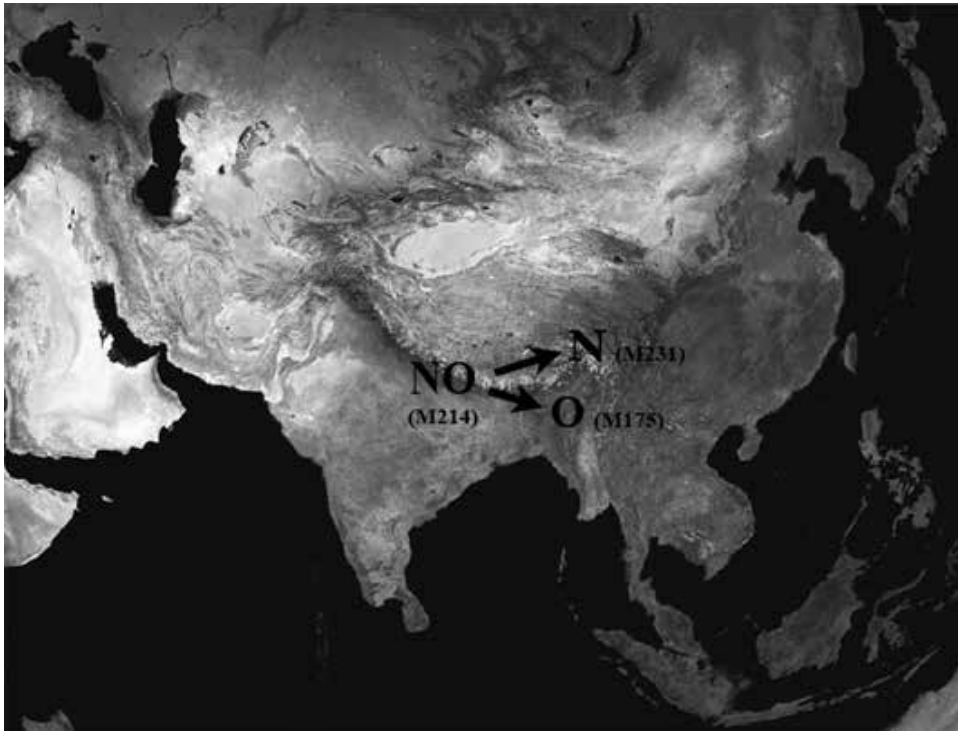


Figure 20. *The split-up of paternal lineage NO (M214) into the haplogroups N (M231) and O (M175)*

established themselves in the Americas [Rootsi et al. 2007; Fu et al. 2013]. We may therefore imagine that the Greater Yenisseian paternal haplogroup Q must have expanded across Siberia and by way of Beringia colonised the Americas, where it became the predominant paternal lineage, before the Hyperborean intrusion of Y chromosomal N lineages replaced it in the sparsely populated north. The N (M231) lineages differentiated into N*, N1 (M128), N2 (P43) and N3 (Tat). The most prevalent haplogroup N3 (Tat) is widespread throughout the Uralo-Siberian area, spreading as far west as Scandinavia. The ancestral N* (M231) is still found in the highest frequency in the area encompassing northern Burma, Yúnnán and Sìchuān, with N1 (M128) particularly frequent in the Altai region and to a lesser extent in Manchuria, and N2 (P43) showing an especially high frequency on both the Yamal and Tamyр peninsulas in northern Siberia¹.

Previously I proposed that haplogroup Q, an offspring clade of Y chromosomal haplogroup P, could be a marker for the Greater Yenisseian lin-

¹ Evidently inspired by story of the Y chromosomal haplogroup N, [Gāo 2008, 2012] has sought to find a link between Sinitic and Uralic. This hypothesis suffers from some of the same difficulties as Sino-Austronesian. Could reconstructible linguistic vestiges have possibly ever survived in Sinitic and Uralic from a time beyond Uralo-Siberian and East Asian, or could linguistic contact influence of such hoary antiquity ever be plausibly demonstrated?



Figure 21. *The counterclockwise spread of the paternal lineage N (M231), based on [Rootsi et al. 2007]*

guistic phylum [van Driem 2008: Figure 22]. A migration route from Central Asia across Beringia to the Americas is suggested by the phylogeny of the Y chromosomal haplogroup Q (M242). The ancestral paternal lineage of the Americans is still found in high frequency in the Yeniseians, and the movement across Siberia to Beringia must have taken place before the Y chromosomal haplogroup N moved north. The Y chromosomal haplogroup Q (M242) established itself as the predominant paternal lineage of the Americas, most probably in a single migration process, with the bottleneck of geography conditioning the resultant genetic bottleneck. The Q1a3a1 (M3) mutation specific to the Americas only arose after the initial colonisation of the New World and, as populations moved south, this new lineage overwhelmingly came to replace the ancestral lineages Q1a (MEH2) and Q1b (L275), which are still most prevalent in Central Asia and northeastern Siberia. The coalescence times for the haplogroups in question is compatible with conventional archaeological wisdom that the Americas were uninhabited before the Clovis culture, and that men bearing the paternal lineage Q (M242) were also the bearers of the first Clovis projectile points into the heart of North America [Zegura et al. 2003; Seielstad et al. 2003; Schnurr 2004; Rootsi et al. 2007; Karafet et al. 2009a; Malyarchuk et al. 2011; Reguiero et al. 2013].

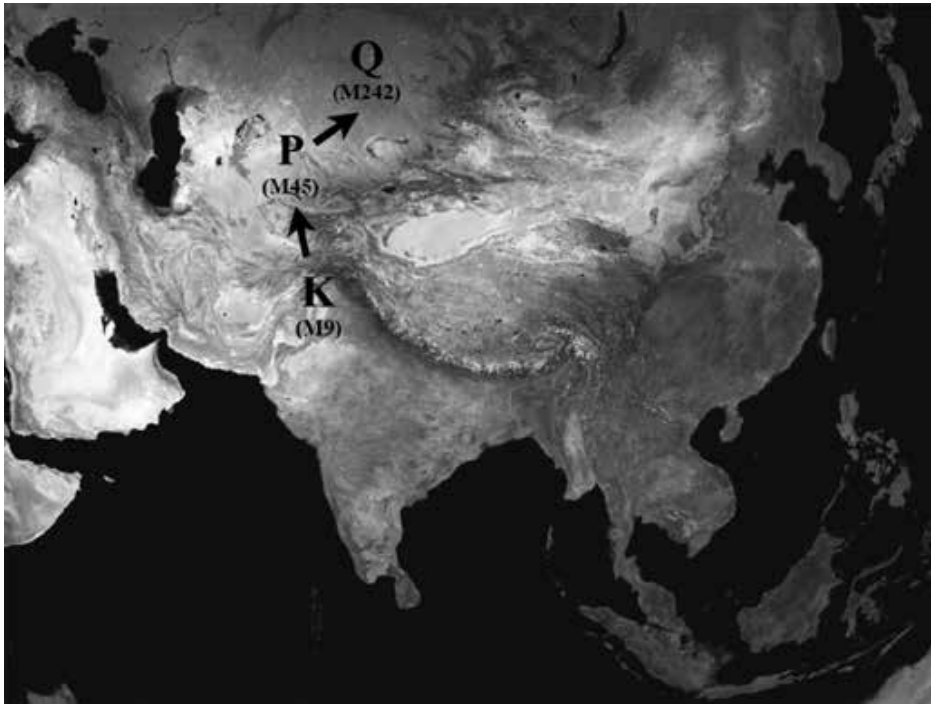


Figure 22. *The Y chromosomal haplogroup Q is a possible patrilineal marker for the spread of the Greater Yenisseeian linguistic phylum [van Driem 2008]*

Following [Toporov 1969, 1971] I adduced some additional linguistic evidence for the hypothesis that Burushaski might be distantly related to Yenisseeian [van Driem 2001], but because of a reinterpretation of the significance of the Karasuk cultural assemblage I later rejected the archaeologically inspired name Karasuk for the putative language family in favour of Greater Yenisseeian [van Driem 2008].¹ The Burushaski may have retained a Greater Yenisseeian language but largely lost the presumed paternal lineage Q (M242). In fact, the ethnolinguistic history of the Burushaski may be a rather complex tale and is no doubt intimately tied up with the Indo-Iranian groups which surround them.

At the time of the release of the new Kusunda material by [Watters 2006], possible correspondences suggested by the Kusunda realis suffix <-ən ~ -n>, the Burushaski plural agent-subject suffix <-an> (except for genderless third person) and the Ket plural subject-agent suffix <-(V)n ~ -(V)ŋ>, by the Kusunda plural suffix <-da> and the Ket distributive prefix <d-> and by a number of other typological parallels between Kusunda and Greater Yenisseeian

¹I retain the conventional English spelling Yenisseeian, which follows the original German and Dutch sources, where the doubling of the *s* ensured a voiceless pronunciation, although a newer Russian-inspired spelling with a single *s* has recently come into use among some English speaking linguists.

verbal agreement morphology raised the question in my mind whether Kusunda might be the remnant of the same ancient Greater Yenisseian migration into the Himalayas.

Subsequently, evidence was adduced to argue that the Yenisseian languages are genetically connected to the Athabaskan-Eyak-Tlingit a.k.a. Na-Dené languages of North America [Vajda 2010]. Vajda called his hypothesis ‘Dene-Yeniseic’. I introduced the term ‘Dene-Yenisseian’ to refer to the putative linguistic phylum comprising Athabaskan-Eyak-Tlingit, Yenisseian, Kusunda and Burushaski, with no explicit hypothesis at present for the structure of the phylogenetic tree connecting these four branches [van Driem 2014]. In this context, the Y chromosomal lineages of the Kusunda tell us an illuminating story of immediate relevance to the Greater Yenisseian hypothesis and to Himalayan ethnolinguistic prehistory, but these data have yet to be published in the population genetic literature.

In a methodologically rigorous appraisal, [Gerber 2013] challenges the idea that a linguistic relationship can even ever be reconstructed at the putative time depth assumed for such a distant genetic relationship. In other words, entities such as Greater Yenisseian or Dene-Yenisseian may very well lie beyond the linguistic event horizon. At the same time, Gerber’s critical and detailed discussion of possible correspondences provides ample food for thought and numerous leads for further research. Recently, [Gerber forthcoming] has adduced additional data and analyses, which justify him in replacing my name ‘Dene-Yenisseian’ with his own coinage Dene-Kusunda for the putative linguistic phylum.

South of the great Himalayan divide, the paternal haplogroup H appears to be associated with the indigenous populations of the Indian subcontinent, and this paternal lineage remains preponderant in today’s tribal population, scheduled castes and the gypsies [Rai et al. 2012]. The Indian origins of the the gypsies or Roma was already argued in 1783 by Grellmann on the basis of their language, culture and physical appearance. The veracity of this thesis was linguistically demonstrated by August Friedrich Pott in 1844, who narrowed the provenance down to North India. This linguistic proof was popularised in Prosper Mérimée’s 1846 novella *Carmen*. Subsequent linguistic analysis yielded specific inferences about the Roma route of migration.

[Grierson 1922] propagated the idea that the ‘Gipsy languages’ were of ‘Dardic origin’, but [Turner 1926] demonstrated that the Romani languages were not Dardic, but belonged to the same central Indo-Aryan subgroup as Hindi. The presence of Burushaski loans in Romani [Berger 1959], the lack of Arabic loans and the presence of Dardic, Georgian, Ossetian, Armenian and mediaeval Greek loans [Hancock 1995] indicated that the Roma migrated

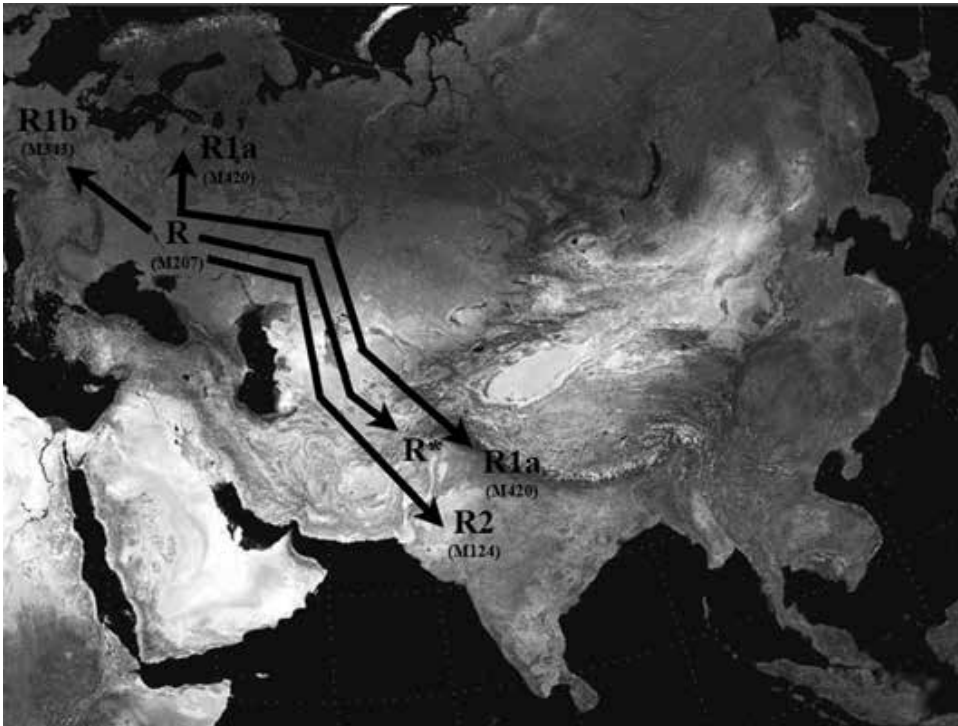


Figure 23. The dissemination of paternal lineage R (M207) and its major subclades

to Europe by a northerly route, beginning around Gilgit in the northernmost Hindu Kush, thence along the southern Caspian littoral, the southern flank of the Caucasus, the southern shoreline of the Black Sea, across the Bosphorus, and subsequently spreading across Europe from the 13th century onwards.

A legacy of the migration is that some Roma refer to themselves Sinti, an adjectival form derived from Sindh, the name of the Indus river. The Roma in England were known as gypsies because it was thought that they were 'gyptians' from Egypt. German *Zigeuner*, French *tzigane* and names in several other European languages derive from a designation for a Manichaean sect that practised sorcery and soothsaying in the last centuries of the Byzantine Empire. Roma populations are also found in eastern Europe, the Balkan, Scandinavia and throughout the Near East.

The name by which Roma designate themselves is *Rroma* (singular *Rrom*), whereby the double *rr* in Romani orthography represents an uvular 'r' [R] as opposed to an apical 'r' [r]. The autonym *Rroma* is held to be cognate with डोम *Doma*, a collective term for the ancient aboriginal populations of the Indian subcontinent who were subjugated and often enslaved by the early Aryans. Many *Doma* remained outcasts or tribals, whereas some were assimilated into the lower strata of the caste system by the Aryans as they colonised the Gangetic Plain from the west to east [Trail 1828; van Driem

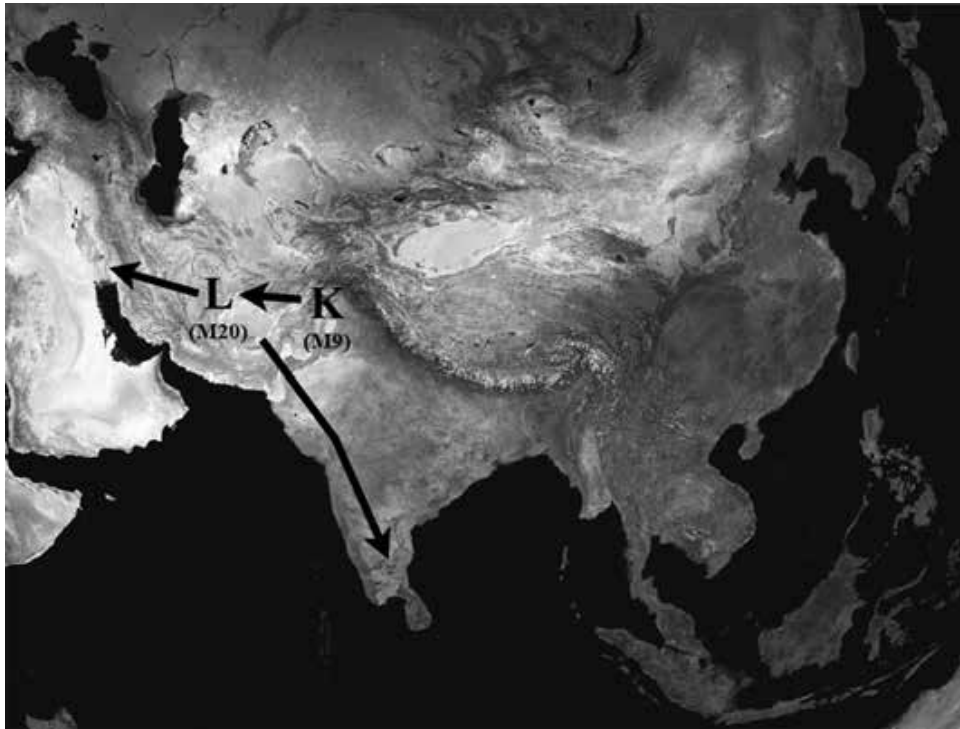


Figure 24. *Y chromosomal lineage L shows a great diversity of subclades on the Iranian highland and is the possible marker of a patrilingual dispersal of Elamo-Dravidian emanating from the Bactria and Margiana region*

2001]. [Morgenstierne 1973] argued that the endangered Ḍomākī language spoken by several hundred Ḍoma in Gilgit and Yasin, belonging to the minstrel and blacksmith castes, represents an ethnolinguistic remnant of the early Rroma migration through what today is northern Pakistan.

From 2001, several population genetic studies corroborated the Indian ancestry of the Rroma, comprising mtDNA, Y-chromosomal and autosomal studies. Most recently, however, [Rai et al. 2012] for the first time provided clear population genetic evidence on the basis of the H1a1a paternal signature for the identification of the paternal ancestry of the the Rroma specifically with the Ḍoma of northwestern India. The Ḍoma are a subset of the Śūdra caste, representing an indigeous population of the Indian subcontinent who were incorporated into the lowest tier of the *caturavarna* caste system. This finding corroborates inferences made by linguists and ethnographers in the past and sheds light on the possible sociolinguistic nature of the migration which brought an Indo-Aryan speaking population as far west as the Irish Sea and beyond.

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, 2012a: Figure 23], but the unfolding story of Y chromosomal R lineages will no doubt turn out to



Figure 25. Geographical distribution of Dravidian

be complex [Underhill et al. 2010]. 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 2012a: Figure 24]. One of these Y chromosomal L subclades appears to be correlated with the patrilingual spread of Dravidian languages from the Indus Valley into south India (Figure 25).

The Beluch show haplogroup L at a greater frequency than any other group in Pakistan, far more so than the Brahui [Qamar et al. 2002; Haber et al. 2012]. This paradoxical finding can, however, permit us to make inferences about the historical sociolinguistic situation in the aftermath of the Indo-Aryan incursion into the territory of the Indus civilisation and the Himalayas. The complex intercourse and extensive bilingualism which have historically characterised the Brahui-Baluchi commensality has been documented and described by [Bray 1909, 1934; Emeneau 1962a, 1962b] and [Elfenbein 1982, 1983, 1987].

We may infer that, in part, the Beluch represent the *in situ* descendants of the ancient Dravidians who peopled the Indus civilisation. The ancestors of the Beluch, however, opted to assimilate linguistically to the incursive

Indo-Europeans and retained the originally predominant Elamo-Dravidian haplogroup L at the highest frequency. By contrast, the Brahui retained the original Dravidian language of their ancestors and consequently also inherited the lower social status of the subjugated Indus population. Ironically, the lower status connected with the retention of a pre-Indo-Aryan ethnolinguistic identity rendered the Brahui language community more prone to male-biased genetic contributions from incursive Indo-European groups through hypergamy practised by succeeding generations of Brahui women.

Confronted with the overwhelming growing body of evidence demonstrating the prevalence of the Father Tongue correlation, [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]. Those collaborating with Renfrew continue to seek Y chromosomal correlates for the spread of the agricultural horizon in the Neolithic, even when the reasoning continues to be strained and the purported correlates are not manifestly evident, e. g. [Arun Kumar et al. 2012]. In order to buttress Renfrew’s widely doubted hypothesis of an Indo-European homeland in Asia Minor, Forster and Renfrew propose a correlation of Indo-European with the Y chromosomal haplogroup J2a.

In a similar vein, [Wolff 2010] attributes the spread of the language families Tibeto-Burman, Austronesian, Kra-dai, Hmong-Mien and Austroasiatic all to the spread of rice cultivation, a highly simplistic view criticised by [Blust 2011]. 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 peopling of Oceania by Austronesian populations, the colonisation of previously uninhabited lands invariably involved both sexes and the introduction of a linguistic 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.

The reasoning of [Bellwood and Renfrew 2002] is therefore flawed, and the Centripetal Migration model was proposed as an alternative to their centrifugal Farming Language Dispersal theory [van Driem 2007]. Bellwood and Renfrew argue that the surplus generated by an agricultural economy and the stratified social and command structure enabled by a Neolithic lifestyle drove demographic spread into many areas. They claim that the incremental spread of the Neolithic led to ‘the foundation dispersals’ of language families with the ancient spread of linguistic phyla unfolding in the same direction as the demographic spread driven by Neolithic agriculture. The very opposite may be what actually happened in many cases. Across the Fertile Crescent, agriculture was adopted by ethnolinguistically unrelated populations, and agriculture spread effortlessly across ethnolinguistic boundaries without disrupting them in any significant way. The phylogeography of barley DNA suggests that the spread of the Neolithic across Europe may have involved a complex history of interaction between foragers and farmers [Jones et al. 2012].

Sumerian pictographic script, developed ca. 3200 BC, appeared millennia after the invention of agriculture. Sumerian, Elamite, Akkadian,¹ Hurrian, Hattic and other contemporaneous agricultural civilisations were in all likelihood not the first cultivators of the region. Yet even these antique agricultural language communities have left no surviving linguistic descendants. The earliest recorded and reconstructible history of the Near East bears witness to the permeability of linguistic boundaries for the dissemination of agriculture and crops. The Bronze Age of Asia Minor and Mesopotamia is characterised by a long period of incursive population movements into, rather than out of Anatolia and the Fertile Crescent, lured by the relative affluence of urban centres supported by agricultural surplus. Gutæans, Amorites, Kassites and other peoples were drawn in by the promise of the good life. Most linguistic reconstructions presume that Indo-European groups such as the Hittites and Mitanni likewise came to settle in Asia Minor and the Fertile Crescent from elsewhere. Toponymical evidence and details about the cults of certain deities have been used to argue that even the Sumerians originally migrated from an earlier northern homeland to lower Mesopotamia.

Were the motivations of migrating peoples in agricultural and pre-agricultural societies genuinely different at the Neolithic horizon than at later times? Tidings of technologically advanced urban societies may in the course of prehistory have provided ample motivation for migration, with enticing prospects of plunder and material advancement. We must consider such al-

¹ Today Afroasiatic languages are spoken throughout this area, but none are descended directly from the extinct branch of the family represented by Akkadian.

ternatives especially in those cases where the linguistic picture suggests a radically different view of prehistory than does the spread of material culture as reflected in the known archaeological record. The introduction of Proto-Sinitic to the Yellow River basin may have been likewise inspired.

The Centripetal Migration model was named to contrast with the centrifugal reasoning of the Farming Language Dispersal proponents. In fact, ancient population movements may have unfolded both in centrifugal and centripetal directions with respect to centres of technologically advanced and later urban civilisations. The Centripetal Migration model acknowledges that the motives for migrations were diverse and that no model as simple as the Farming Language Dispersal theory could therefore account for all linguistic intrusions, even at the time of the Neolithic horizon.

With reference to Forster and Renfrew's wilful interpretation of the Y chromosomal haplogroup J2, I previously argued in the context of the Indian subcontinent that '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' and the ancient maritime trade across the Arabian Sea, 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 patrilingual dispersal of Elamo-Dravidian emanating from a region which encompassed the Bactria and Margiana of later prehistory.

Contrary to both the Pontic Caspian and Asia Minor homeland theories, the presence of the ancestral clade R* in Indian populations could be construed as evidence for the hypothesis of an ultimate Indian homeland for Indo-European. This hypothesis exists in two versions. The old myth of all Indo-European languages deriving from Sanskrit originated with Sir William Jones' garbled understanding of the Scythian linguistic theory. Consequently, this idea is, naturally enough, not countenanced in the scholarly literature today.¹ Another more sophisticated version of the hypothesis, however is that

¹ In 1647, the Scythian linguistic phylum outlined by Marcus van Boxhorn encompassed Latin, Greek, Germanic, Baltic, Slavic, Celtic, Indo-Iranian, including Sanskrit. In 1647, the Scythian language family did not yet contain Albanian, which Rasmus Rask first suggested was Indo-European at the beginning of the 19th century. Albanian was only demonstrated to be Indo-European in 1835 by Joseph Ritter von Xylander. In 1647, Scythian

of a pre-Indo-European or even pre-Nostratic homeland in India. The presence of F* and K* in Indian populations represents additional molecular evidence for the even more daring hypothesis that the Indian subcontinent may have been the ultimate primordial fatherland of most of linguistic and genetic phyla outside of Africa.

References

- Abdulla, Mahmood Ameen* et al. and the Indian Genome Variation Consortium. 2009. ‘Mapping human genetic diversity in Asia’, *Science*, 326: 1541–1545.
- Arunkumar, G., Soria-Hernanz, D. F., Kavitha, V. J.* et al. 2012. ‘Population differentiation of Southern Indian male lineages correlates with agricultural expansions predating the caste system’, *Public Library of Science PLoS One*, 7 (11): e50269.
- Ballard, W. L.* 1979. ‘Chinese — A bastard at the Sino-Tibetan family reunion’, paper presented at the 12th International Conference on Sino-Tibetan Languages and Linguistics, Paris.
- Ballard, W. L.* 1984. ‘The Mother Soup: A South China recipe for tonometamorphogenesis’, *Computational Analyses of Asian and African Languages*, 22: 43–70.
- Beaujard, P.* 2012. *Les mondes de l’Océan Indien — Tome I: De la formation de l’état au premier système-monde afro-eurasien (4e millénaire av. J.-C. — 6e siècle*

likewise did not yet include Hittite, Luvian and Palaic because the clay tablets on which these extinct languages were recorded in cuneiform script had not yet been discovered, and were later recognised as Indo-European by Bedřich Hrozný only in 1915. Manuscripts written in Tocharian languages were not discovered until the beginning of the 20th century. The Scythian language family was renamed Indo-Germanic by Malte-Brun in 1810, and this name has largely been replaced by the synonymous term Indo-European, which was first used by the English polymath Thomas Young in 1813 in a book review of *Mithridates* by Johann Christoph Adelung. Portions of this story are told in detail elsewhere [van Driem 2001: 1039–1051, 2005: 285–291]. By contrast, William Jones believed that most of the languages ‘from the China Seas to Persia’, including Latin and Greek, all derived from Sanskrit. Accordingly he called this language family the ‘Indian branch’. Jones’ two other language families were the ‘Tartarian’ and ‘Arabian branches’. His three branches derived from Noah’s three sons, whereas languages not belonging to these three branches were considered by Jones to be ‘antediluvian’ vestiges, i. e. remnants from before the Biblical Flood. In Jones’ conception, Sanskrit was ancestral to Latin, Chinese, Ancient Egyptian, Japanese, the languages of Ethiopia, Peruvian, the Celtic languages, Mexican, Greek and Phoenician, whose speakers all ‘had a common source with the Hindus’. Jones’ study of Hindi led him to believe that Hindi was unrelated to Sanskrit. Instead, Hindi was of ‘Tartarian or Chaldean origin’ [Jones 1786, 1792, 1793]. Nonetheless, the absurd myth ascribing the discovery of the Indo-European language family to Jones is astonishingly robust.

apr. J.-C.), Tome II: L'Océan Indien au cœur des globalisations de l'ancien monde (7e-15e siècles). Paris: Armand Colin.

Bellwood, P., Colin R. (eds.). 2002. Examining the Farming/Language Dispersal Hypothesis. Cambridge: McDonald Institute for Archaeological Research.

Bellwood, P., Eusebio, D. 2005. 'The Batanes archaeological project and the 'Out of Taiwan' hypothesis for Austronesian dispersal', *Journal of Austronesian Studies*, 1 (1): 1–33.

Bellwood, P., Geoffrey Ch., Malcolm, R. and Hsiao-chun Hung. 2011. 'Are "cultures" inherited: Multidisciplinary perspectives on the origins and migrations of Austronesian-speaking peoples prior to 1000 bc', pp. 321–354 in Benjamin W. Roberts and Marc van der Linden, eds., *Investigating Archaeological Cultures: Material Culture, Variability and Transmission*. Dordrecht: Springer.

Benedict, P. K. 1942. 'Thai, Kadai, and Indonesia: A new alignment in southeastern 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.

Bielmeier, R. forthcoming. *Comparative Dictionary of Tibetan Dialects* (three volumes). Berlin: Mouton de Gruyter.

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

Blust, R. 1996. 'Beyond the Austronesian homeland: The Austric hypothesis and its implications for archaeology', pp. 117–160 in Ward H. Goodenough, ed., *Prehistoric Settlement of the Pacific*. Philadelphia: American Philosophical Society.

Blust, R. 2009. *Austronesian Languages*. Canberra: Pacific Linguistics.

Blust, R. 2011. 'Dempwolff reinvented: A review of Wolff (2010)', *Oceanic Linguistics*, 50 (2): 560–579.

Blust, R. 2014. 'Some recent proposals concerning the classification of the Austronesian languages', *Oceanic Linguistics*, 53 (2): 300–391.

Bó Wéncé. 2004. 'Càijiāhuà gàikuàng', *Mínzú Yǔwén*, 2004 (2): 68–81.

Bodin, I. 1576. *Les six livres de la Republique*. Paris: Chez Jacques du Puys, Librairie Iuré à la Samaritaine.

Bodman, Nicholas Cleaveland. 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', pp. 34–199 in Frans van Coetsem and Linda R. Waugh, eds., *Contributions to Historical Linguistics: Issues and Materials*. Leiden: E.J. Brill.

Bradley, D. 2012. 'Proto-Tibeto-Burman grain crops', *Rice*, 4 (3–4): 134–141.

Brantingham, P. J., Gao Xing, Olden, J. W. et al. 2007. 'A short chronology for the peopling of the Tibeian plateau', *Developments in Quaternary Sciences*, 9: 129–150.

Bray, D. de Saumarez. 1909. *The Brahui Language, Part I: Introduction and Grammar.* Calcutta: Superintendent of Government Printing.

Bray, D. de Saumarez. 1934. *The Brahui Language, Part II: The Brahui Problem, and Part III: Etymological Dictionary* [reprinted 1978 at Quetta: Brahui Academy].

Cai, X., Qin Z., Wen, B. et al. 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.

Chaubey, G., Metspalu, M., Choi, Y. et al. 2011. ‘Population genetic structure in Indian Austroasiatic speakers: The role of landscape barriers and sex-specific admixture’, *Molecular Biology and Evolution*, 28 (2): 1013–1024.

Chaubey, G. 2010. *The Demographic History of India: A Perspective based on Genetic Evidence.* Tartu: Tartu University Press.

Ch'en Chieh-hsien. 1976. ‘The decline of the Manchu language in China during the Ch'ing period’, pp. 137–154 in *Altaica Collecta: Berichte und Vorträge der XII. Permanent International Altaistic Conference (3.-8. Juni 1974 in Bonn/Bad Honneg).* Wiesbaden: Otto Harrassowitz.

Ci, H.X., Lin G. H., Cai Z. Y. et al. 2009. ‘Population history of the plateau pika endemic to the Qinghai-Tibetan plateau based on mtDNA sequence data’, *Journal of Zoology*, 279: 396–403.

Comrie, B. 2008. ‘The areal typology of Chinese: Between North and Southeast Asia’, pp. 1–21 in Redouane Djamouri, Barbara Meisterernst and Rint Sybesma, eds. *Chinese Linguistics in Leipzig: Chinese Linguistics in Europe*, 2. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale.

Conrady, A. 1916. ‘Eine merkwürdige Beziehung zwischen den austrischen und den indochinesischen Sprachen’, pp. 475–504 in *Aufsätze zur Kultur- und Sprachgeschichte vornehmlich des Orients: Ernst Kuhn zum 70. Geburtstag am 7. Februar 1916 gewidmet von Freunden und Schülern.* München: Verlag von M. & H. Marcus.

Conrady, A. 1922. ‘Neue austrisch-indochinesische Parallelen’, pp. 23–66 in *Asia Major: Hirth Anniversary Volume.* London: Robsthan and Company.

Currat, M., Excoffier, L. 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.

Cust, R. N. 1878. *A Sketch of the Modern Languages of East India.* London: Trübner and Company.

Dai, F., Nevo, E., Wu, D. et al. 2012. ‘Tibet is one of the centers of domestication of cultivated barley’, *Proceedings of the National Academy of Sciences*, 109 (42): 16969–16973.

DeLancey, S. 2011. ‘On the origins of Sinitic’, pp. 51–64, Vol. 1 in Zhuo Jing-Schmidt, ed., *Proceedings of the 23rd North American Conference on Chinese Linguistics.* Eugene: University of Oregon.

Derenko, M., Malyarchuk, B., Denisova, G. et al. 2007. ‘Y-chromosome haplogroup N dispersal from south Siberia to Europe’, *Journal of Human Genetics*, 52: 763–770.

Diffloth, G. 1994. 'The lexical evidence for Austric, so far', *Oceanic Linguistics*, 33 (2): 309–321.

Diffloth, G. 2009. 'More on Dvaravati Mon', paper presented at the Fourth International Conference on Austroasiatic Linguistics (icaal 4) at Mahidol University, Salaya, 29 October 2009.

Diffloth, G. 2011. 'Austroasiatic word histories: boat, husked rice and taro', pp. 295–314 in Nick J. Enfield, ed., *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*. Canberra: Pacific Linguistics.

Diffloth, G. 2012. 'The four registers of Pearic', conférence plénière at the 22nd Meeting of the Southeast Asian Linguistics Society, Agay, 2 June 2012.

van Driem, G. 1998. 'Neolithic correlates of ancient Tibeto-Burman migrations', pp. 67–102 in Roger Blench and Matthew Spriggs, eds., *Archaeology and Language II*. London: Routledge.

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

van Driem, G. 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.

van Driem, G. 2005. 'Sino-Austronesian vs. Sino-Caucasian, Sino-Bodic vs. Sino-Tibetan, and Tibeto-Burman as default theory', pp. 285–338 in Yogendra Prasada Yadava, Govinda Bhattarai, Ram Raj Lohani, Balaram Prasain and Krishna Parajuli, eds., *Contemporary Issues in Nepalese Linguistics*. Kathmandu: Linguistic Society of Nepal.

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', pp. 39–59, Vol.1 in Brigitte Huber, Marianne Volkart and Paul Widmer, eds., *Chomolangma, Demawend und Kasbek: Festschrift für Roland Bielmeier zu seinem 65. Geburtstag (2 vols.)*. Halle: International Institute for Tibetan and Buddhist Studies.

van Driem, G. 2011a. 'Rice and the Austroasiatic and Hmong-Mien homelands', pp. 361–389 in Nick J. Enfield, ed., *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*. Canberra: Pacific Linguistics.

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', pp. 335–361 in Andrea Ender, Adrian Leemann and Bernhard Wälchli, eds., *Methods in Contemporary Linguistics*. Berlin: Mouton de Gruyter.

van Driem, G. 2013. 'Trans-Himalayan', in Nathan Hill and Tom Owen-Smith, eds., *Trans-Himalayan Linguistics*. Berlin: Mouton de Gruyter.

van Driem, G. 2014. ‘A prehistoric thoroughfare between the Ganges and the Himalayas’, pp. 60–98 in Tiatoshi Jamir and Manjil Hazarika, eds., *Fifty Years after Daojali-Hading: Emerging Perspectives in the Archaeology of Northeast India*. New Delhi: Research India Press.

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

Dulik, M. C., Zhadanov, S. I., Osipova, L. P. et al. 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.

Elfenbein, J. 1982. ‘Notes on the Balochi-Brahui commensality’, *Transactions of the Philological Society*, 1982: 77–98.

Elfenbein, J. 1983. ‘The Brahui problem again’, *Indo-Iranian Journal*, 25: 103–132.

Elfenbein, J. 1987. ‘A Periplus of the “Brahui Problem” ’, *Studia Iranica*, 16: 215–233.

Emeneau, M. B. 1962a. *Dravidian Borrowings from Indo-Aryan* (University of California Publications in Linguistics, xxvi). Berkeley: University of California.

Emeneau, M. B. 1962b. *Brahui and Dravidian Comparative Grammar* (University of California Publications in Linguistics, xxvii). Berkeley: University of California Press.

Ferlus, M. 1996. ‘Du taro au riz en Asie du Sud-Est: petite histoire d’un glissement sémantique’, *Mon-Khmer Studies*, 25: 39–49.

Forbes, C. J. 1878. ‘On Tibeto-Burman languages’, *Journal of the Royal Asiatic Society of Great Britain and Ireland*, X: 210–227.

Forster, P., Renfrew, C. 2011. ‘Mother tongue and Y chromosomes’, *Science*, 333: 13901–1391.

Fortescue, M. 1998. *Language Relations across Bering Strait: Reappraising the Archaeological and Linguistic Evidence*. London: Cassell.

Fortescue, M. 2011. ‘The relationship of Nivkh to Chuktko-Kamchatkan revisited’, *Lingua*, 121: 1359–1376.

Fu Q., Meyer M., Gao X. et al. 2013. ‘DNA analysis of an early modern human from Tianyuan cave, China’, *Proceedings of the National Academy of Science*, 110 (6): 2223–2227.

Gan, R. J., Pan, S. L., Mustavich, L. F. et al. 2008. ‘Pingua population as an exception of Han Chinese’s coherent genetic structure’, *Journal of Human Genetics*, 53: 303–313.

Gāo Jīngyī. 2008. *Hányǔ yú Běi’ōu Yǔyán: Hányǔ yú Wūlā’ěr Yǔyán jí Yīn’ōu Yǔyán Tóngyuán Tànjiū* [‘Chinese and Nordic languages: Investigating correspondences between Chinese and Indo-European and Uralic’]. Peking: Zhōngguó Shèhuì Kēxué Chūbǎnshè.

Gāo Jīngyī. 2012. ‘Official colours of Chinese regimes: A panchronic philological study with historical accounts of China’, *Trames*, 16 (66/61), 3: 237–285.

Gerber, P. 2013. *A Critical Approach to Distant Genetic Relationships: A Comparative Examination of the Verbal Morphology of Kusunda, Burushaski, Yenisseian and Athabaskan-Eyak-Tlingit*. University of Bern Bachelor’s Thesis, 17 May 2013.

Gerber, P. forthcoming. ‘The Dene-Kusunda hypothesis: A critical account’.

Glynis, J., Huw, J., Charles, M. P., Jones, M. K. et al. 2012. ‘Phylogeographic analysis of barley DNA as evidence for the spread of Neolithic agriculture across Europe’, *Journal of Archaeological Science*, 39: 3230–3238.

Green, R. E., Krause, J., Briggs, A. W. et al. 2010. ‘A draft sequence of the Neandertal genome’, *Science*, 328 (5929): 710–722.

Haber, M., Platt, D. E., Ashrafian Bonab M. et al. 2012. ‘Afghanistan’s ethnic groups share a Y-chromosomal heritage structured by historical events’, *Public Library of Science One*, 7 (3): e34288.

Hagège, C., Haudricourt, A.-G. 1976. *La phonologie panchronique*. Presses Universitaires de France.

Hammer, M. F., Karafet, T. M., Park, H. et al. 2006. ‘Dual origins of the Japanese: Common ground for hunter-gather and farmer Y chromosomes’, *Journal of Human Genetics*, 51: 47–58.

Hashimoto Mantarō. 1976a. ‘Language diffusion on the Asian continent’, *Computational Analyses of Asian and African Languages*, 3: 49–64.

Hashimoto Mantarō. 1976b. ‘The agrarian and the pastoral diffusion of language’, pp. 1–14 in *Genetic Relationship, Diffusion and Typological Similarities of East and Southeast Asian Languages*. Tokyo: Japan Society for the Promotion of Sciences.

Hashimoto Mantarō. 1980. ‘Typogeography of phonotactics and suprasegmentals in the languages of the East Asian continent’, *Computational Analyses of Asian and African Languages*, 13: 153–164.

Hashimoto Mantarō. 1986. ‘The altaicization of Northern Chinese’, pp. 76–97 in *John McCoy and Timothy Light, eds., Contributions to Sino-Tibetan Studies*. Leiden: E. J. Brill.

Heusser, C. J. 1955. ‘Pollen profiles from the Queen Charlotte Islands, British Columbia’, *Canadian Journal of Botany*, 1955, 33 (5): 429–449.

Hobbes, T. 1651. *Leviathan, or the Matter, Forme, & Power of a Common-Wealth Ecclesiasticall and Civill*. London: Andrew Crooke at the Green Dragon in St. Paul’s Church-yard.

Hodgson, B. H. 1857. ‘Comparative Vocabulary of the Languages of the broken Tribes of Népal’, *Journal of the Asiatic Society of Bengal*, XXVI: 317–371.

Houghton, B. 1896. ‘Outlines of Tibeto-Burman linguistic palæontology’, *Journal of the Royal Asiatic Society*, 1896: 23–55.

Hudjashov, G., Kivisild, T., Underhill, P.A. et al. 2007. ‘Revealing the prehistoric settlement of Australia by Y chromosome and mtDNA analysis’, *Proceedings of the National Academy of Sciences*, 104 (21): 8726–8730.

- HUGO Pan-Asian SNP Consortium, Abdulla, M.A., Ahmed, I. et al.* 2009. 'Mapping human genetic diversity in Asia', *Science*, 326: 1541–1545.
- Hurles, M. E., Bryan C. S., Mark A. J., Forster, P.* 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.
- Huxley, T. H.* 1870 'On the geographical distribution of the chief modifications of mankind', *Journal of the Ethnological Society of London*, 2: 404–412.
- Jin, Han-Jun, Chris Tyler-Smith and Wook 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., Hong, L. C., Phipps, M. E. et al.* 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.
- Jones, W.* 1807a [1786]. 'The third anniversary discourse, on the Hindus, delivered 2 February, 1786', pp. 24–46, Vol. iii in John Shore, Lord Teignmouth. *The Works of Sir William Jones* (13 vols). London: John Stockdale.
- Jones, W.* 1807b [1792]. 'The ninth anniversary discourse, on the origin and families of nations, delivered 23 February, 1792', pp. 185–204, Vol. iii in John Shore, Lord Teignmouth. *The Works of Sir William Jones* (13 vols). London: John Stockdale.
- Jones, W.* 1807c [1793]. 'The tenth anniversary discourse, on Asiatick history, civil and natural, delivered 28 February, 1793', pp. 205–228, Vol. iii in John Shore, Lord Teignmouth. *The Works of Sir William Jones* (13 vols). London: John Stockdale.
- Karafet, T. M., Lansing, J.S., Redd, A. J., Watkins, J. C.* 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., Mendez F. L., Meilerman M.B. et al.* 2008. 'New binary polymorphisms reshape and increase resolution of the human Y-chromosomal haplogroup tree', *Genome Research*, 18: 830–838.
- Karafet, T. M., Zegura, S.L., Hammer M. F.* 2009a. 'The peopling of the Americas as viewed from the Y chromosome', pp. 127–136 in Peter N. Peregrine, Ilia Peiros and Murray Feldman, eds., *Ancient Human Migrations: A Multidisciplinary Approach*. Salt Lake City: University of Utah Press.
- Karafet, T. M., Zegura, S.L., Hammer M. F.* 2009b. 'Y-chromosome Japanese roots', pp. 137–148 in Peter N. Peregrine, Ilia Peiros and Murray Feldman, eds., *Ancient Human Migrations: A Multidisciplinary Approach*. Salt Lake City: University of Utah Press.
- Kivisild, T., Tolk, H.-V., Parik, J. et al.* 2002. 'The emerging limbs and twigs of the East Asian mtDNA tree', *Molecular Biology and Evolution*, 19 (10): 1737–1751.
- von Klaproth, J.H.* 1823. *Asia Polyglotta*. Paris: A. Schubart.

- Kumar, S., Reddy, R., Koneru, P., Urade, B.P.* et al. 2007. ‘Reconstructing Indian-Australian phylogenetic link’, *BioMed Central Evolutionary Biology*, 9: 173.
- Kumar, V., Langstieh, B. T., Pradeep, Babu J., Rao T. N.* et al. 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., Reddy, A. N. S., Babu, J. P.* et al. 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.
- Lemaréchal, A.* 2010. *Comparative Grammar and Typology: Essays on the Historical Grammar of the Austronesian Languages*. Leuven: Peeters.
- Lertrit, P., Poolsuwan, S., Thosorat, R.* et al. 2008. ‘Genetic history of Southeast Asian populations as revealed by ancient and modern human mitochondrial DNA analysis’, *American Journal of Physical Anthropology*, 137 (4): 425–440.
- Li, Charles N.* 1995. ‘Language contact in China: Is Mandarin derived from a pidgin?’, pp. 131–148 in Jacek Fisiak, ed., *Linguistic Change under Contact Conditions*. Berlin: Mouton de Gruyter.
- Li H., Wen B., Chen S. J.* et al. 2008 ‘Paternal genetic affinity between western Austronesians and Daic populations’, *BMC Evolutionary Biology*, 8: 146, doi:10.1186/1471–2148–8-146.
- Lì Jīn, Underhill, P. A., Vishal Doctor* et al. 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.
- Li, Jun Z., Devin M. Absher, Hua Tang* et al. 2008. ‘Worldwide human relationships inferred from genome-wide patterns of variation’, *Science*, 319: 1100–1104.
- Madsen, D. B., Ma Haizhou, Brantingham P. J., Gao Xing* et al. 2006. ‘The Late Upper Palaeolithic occupation of the northern Tibetan plateau margin’, *Journal of Archaeological Science*, 33: 1433–1444.
- Malyarchuk, B., Derenko M., Denisova G., Maksimov A.* et al. 2011. ‘Ancient links between Siberians and Native Americans revealed by subtyping the Y chromosome haplogroup Q1a’, *Journal of Human Genetics*, 56: 583–588.
- McNeill, W. H.* 1976. *Plagues and Peoples*. Garden City: Anchor Books.
- Mendez F. L., Krahn, T., Schrack, B.* et al. 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.
- Miehe, G., Miehe, S., Kaiser K.* et al. 2009. ‘How old is pastoralism in Tibet? An ecological approach to the making of a Tibetan landscape’, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 276: 130–147.
- Miehe, Georg, Bach K., Miehe S.* et al. 2012. ‘Alpine steppe plant communities of the Tibetan highlands’, *Applied Vegetation Science*, 14 (4): 547–560.
- Mirabal, S., Regueiro, M., Cadenas A. M.* et al. 2009. ‘Y chromosome distribution within the geo-linguistic landscape of northwestern Russia’, *European Journal of Human Genetics*, 2009: 1–14.

Nichols, J. 1986. 'Head-marking and dependent-marking grammar', *Language*, 62 (1): 56–199.

Nichols, J. 1992. *Linguistic Diversity in Time and Space*. Chicago: Chicago University Press.

Nichols, J. 1995. 'Diachronically stable structural features', pp. 337–356 in H. Andersen, ed., *Historical Linguistics 1993: Papers from the Eleventh International Congress of Historical Linguistics*. Amsterdam: John Benjamins.

Nichols, J. 1998. 'The Eurasian spread zone and the Indo-European spread', pp. 220–266 in Roger Marsh Blench and Matthew Spriggs, eds., *Archaeology and Language II: Archaeological Data and Linguistic Hypotheses*. London: Routledge.

Norman, J. 1982. 'Four notes on Chinese-Altaic linguistic contacts', *Tsing Hua Journal of Chinese Studies*, 1982 (1–2): 243–247.

Opgenoorth, L., Giovanni G. V., Mao Kangshan et al. 2010. 'Tree endurance on the Tibetan plateau marks the world's highest known tree line of the Last Glacial Maximum', *New Phytologist*, 185: 332–342.

Orwell, G. 1949. *Nineteen Eighty-Four*. London: Secker and Warburg.

Ostapirat, W. 2005. 'Kra-Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution', pp. 107–131 in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon.

Ostapirat, W. 2013. 'Austro-Tai revisited', 23rd Annual Meeting of the Southeast Asian Linguistic Society, Chulalongkorn University, 29 May 2013.

Parducci L., Jørgensen T., Tollefsrud M. M. et al. 2012. 'Glacial survival of boreal trees in northern Scandinavia', *Science*, 335: 1083–1086.

Peiros, I. 1998. *Comparative Linguistics in Southeast Asia*. Canberra; Pacific Linguistics.

Peiros, I. 2008. 'The Formosan language family', pp. 182–210 in Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin, eds., *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*. London and New York: Routledge.

Peng Y., Yang Z., Zhang H. et al. 2011. 'Genetic variations in Tibetan populations and high-altitude adaptation at the Himalayas', *Molecular Biology and Evolution*, 28 (2): 1075–1081.

Poloni, E. S., Semino, O., Passarino G. et al. 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, E. S., Ray N., Schneider S., Langaney A. 2000. 'Languages and genes: Modes of transmission observed through the analysis of male-specific and female-specific genes', pp. 185–186 in Jean-Louis Dessalles and Laleh Ghadakpour, eds., *Proceedings: Evolution of Language, 3rd International Conference 3–6 April 2000*. Paris: École Nationale Supérieure des Télécommunications.

Poppe, Nikolaus ('Nicholas Poppe'). 1965. Introduction to Altaic Linguistics. Wiesbaden: Otto Harrassowitz.

Pulleyblank, E. G. 1983. 'The Chinese and their neighbours in prehistoric and early historic times', pp. 411–466 in David N. Keightley, ed., *The Origins of Chinese Civilization*. Berkeley: University of California Press.

Qamar R., Ayub Q., Mohyuddin A. et al. 2002. 'Y-chromosomal DNA variation in Pakistan', *American Journal of Human Genetics*, 70: 1107–1124.

Qi Xuebin, Han Jianlin, Blench R. et al. 2008. 'Understanding yak pastoralism in Central Asian highlands', pp. 427–442 in Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin, eds., *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*. London and New York: Routledge.

Qi X., Cui C., Peng Y. et al. 2013. 'Genetic evidence of Palaeolithic colonisation and Neolithic expansion of modern humans on the Tibetan plateau', *Molecular Biology and Evolution*, doi: 10.1093/molbev/mst093

Qian, Yaping, Binzhi Qian, Bing Su et al. 2000. 'Multiple origins of Tibetan Y chromosomes', *Human Genetics*, 106: 453–454.

Quintana-Murci L., Krausz C., Zerjal T. et al. 2001. 'Y-chromosome lineages trace diffusion of people and languages in southwestern Asia', *American Journal of Human Genetics*, 68: 537–542.

Rai N., Chaubey G., Tamang R. et al. 2012. 'The phylogeography of Y-chromosome haplogroup H1a1a-M82 reveals the likely Indian origin of the European Romani populations', *Public Library of Science PLoS One*, 7 (11): e48477.

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.

Rao, V. R., Matthews, P. J., Eyzaguirre P. B., Hunter D. (eds.) 2010. *The Global Diversity of Taro: Ethnobotany and Conservation*. Rome: Biodiversity International.

Rasmussen, M., Guo X., Wang Y. et al. 2011. 'An aboriginal Australian genome reveals separate human dispersals into Asia', *Science*, 334 (6052): 94–98.

Reddy, B. M., Langstieh B. T., Kumar, V. et al. 2007. 'Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia', *Public Library of Science One*, 2 (11): e1141. doi:10.1371/journal.pone.0001141

Reguiero, M., Alvarez J., Rowold D., Herrera R. J. 2013. 'On the origins, rapid expansion and genetic diversity of native Americans from hunting-gatherers to agriculturalists', *American Journal of Physical Anthropology*, 150: 333–348.

Reich D., Green R. E., Kircher M. et al. 2010. 'Genetic history of an archaic hominin group from Denisova Cave in Siberia', *Nature*, 468: 1053–1060.

Reid, L. A. 1994. 'Morphological evidence for Austric', *Oceanic Linguistics*, 33 (2): 323–344.

Reid, L. A. 2005. 'The current status of Austric: a review and evaluation of the lexical and morphosyntactic evidence', pp. 132–160 in Laurent Sagart, Roger

Blench and Alicia Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon, London.

Reid, L. A. 2006. 'On the origin of Philippine vowel grades', *Oceanic Linguistics*, 45: 457–473.

Reid, L. A. 2007. 'Another look at the marking of plural personal noun constructions in Austronesian languages', *Oceanic Linguistics*, 46 (1): 232–252.

Reid, L. A. 2009a. 'The reconstruction of a dual pronoun to Proto-Malayo-Polynesian', pp. 461–477 in Bethwyn Evans, ed., *Discovering History through Language: Papers in Honour of Malcolm Ross*. Canberra: Pacific Linguistics.

Reid, L. A. 2009b. 'On the diachronic development of C1V1-reduplication in some Austronesian languages', *Morphology*, 19 (2): 231–261.

Reid, L. A. 2010. 'Palauan velar nasals and the diachronic development of Proto-Malayo-Polynesian noun phrases: A response to Blust', *Oceanic Linguistics*, 49 (2): 436–477.

Rootsi S., Zhivotovsky L. A., Baldovic M. et al. 2007. 'A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe', *European Journal of Human Genetics*, 15: 204–211.

Ross, M. D. 2009. 'Proto Austronesian verbal morphology: A reappraisal', pp. 295–326 in Alexander Adelaar and Andrew Pawley, eds., *Austronesian Historical Linguistics and Culture History: A Festschrift for Robert Blust*. Canberra: Pacific Linguistics.

Sagart, L. 1993. 'Chinese and Austronesian: evidence for a genetic relationship', *Journal of Chinese Linguistics*, 21 (1): 1–62.

Sagart, L. 2005. 'Sino-Tibetan-Austronesian: An updated and improved argument', pp. 161–176 in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon.

Sahoo, Sanghamitra, Anamika Singh, G. Himabindu et al. 2006. 'A prehistory of Indian Y chromosomes: Evaluating demic diffusion scenarios', *Proceedings of the National Academy of Sciences of the United States of America*, 103 (4): 843–848.

van Schendel, W. 2002. 'Geographies of knowing, geographies of ignorance: jumping scale in Southeast Asia', *Environment and Planning D: Society and Space*, 20: 647–668.

von Schlegel, F. 1808. *Ueber die Sprache und Weisheit der Indier: Ein Beitrag zur Begruendung der Alterthumskunde, nebst metrischen Uebersetzungen indischer Gedichte*. Heidelberg: Mohr und Zimmer.

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)*, II: 76–87.

Schlegel, G. 1902. *Siamese Studies (T'oung Pao, New Series II, Volume II, Supplement)*. Leiden.

Schmidt, J., Opgenoorth L., Martens J., Mieke G. 2011. 'Neo-endemic ground beetles and private tree haplotypes: Two independent proxies attest to a moderate

Last Glacial Maximum summer temperature depression of 3–4°C for the southern Tibetan plateau’, *Quaternary Science Reviews*, 30: 1918–1925.

Schmidt, W. 1906. ‘Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens’, *Archiv für Anthropologie, Neue Folge*, V: 59–109.

Schnurr, T. G. 2004. ‘The peopling of the New World: Perspectives from molecular anthropology’, *Annual Review of Anthropology*, 33: 551–583.

Scott, J. C. 2009. *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*. New Haven: Yale University Press.

Seielstad, M., Yuldasheva N., Singh N. et al. 2003. ‘A novel Y-chromosome variant puts an upper limit on the timing of first entry into the Americas’, *American Journal of Human Genetics*, 73: 700–705.

Sengupta, Sanghamitra, Zhivotovsky L.A., King R et al. 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.

Shah A.M., Tamang R., Moorjani P. et al. 2011. ‘Indian Siddis: African descendants with Indian admixture’, *American Journal of Human Genetics*, 89: 154–161.

Shi Hong, Hua Zhong, Yi Peng et al. 2008. ‘Y chromosome evidence of earliest modern human settlement in East Asia and multiple origins of Tibetan and Japanese populations’, *Biomed Biology*, 6 (45): doi: 10.1186/1741–7007–6-45.

de Sousa, Hilário. 2012. ‘The southern Sinitic languages at the edge of mainland Southeast Asia’, paper read for the Berner Zirkel für Linguistik, University of Bern, 21 November 2012.

Sprigg, R.K. 2002. *Balti-English English-Balti Dictionary*. London: Routledge Curzon.

Stanyon, R., Sazzini M., Luiselli D. 2009. ‘Timing the first human migration into eastern Asia’, *Journal of Biology*, 8: 18, doi:10.1186/jbiol115.

Starosta, S. 2005 (posthumous). Proto-East-Asian and the origin and dispersal of languages of East and Southeast Asia and the Pacific, pp. 182–197 in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon.

Starostin, S. A. 1994. ‘The reconstruction of Proto-Kiranti’, paper presented at the 27ème Congrès International sur les Langues et la Linguistique Sino-Tibétaines, Centre International d’Études Pédagogiques à Sèvres, 14 octobre 1994.

Starostin, S. A. 2008 (posthumous). ‘Comparing Japanese and Korean’, pp. 254–262 in Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros and Marie Lin, eds., *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*. London and New York: Routledge.

Stewart, J. R., Stringer, C. B. 2012. ‘Human evolution out of Africa: The role of refugia and climate change’, *Science*, 335: 1317–1321.

Tamang, R., Thangaraj K. 2012. ‘Genomic view on the peopling of India’, *Investigative Genetics*, 3: 20.

- Tanaka M., Cabrera V. M., González A. M. et al.* 2004. 'Mitochondrial genome variation in eastern Asia and the peopling of Japan', *Genome Research*, 10A: 1832–1850.
- Teng, S. F., Ross M.D.* 2010. 'Is Puyuma a primary branch of Austronesian? A reply to Sagart', *Oceanic Linguistics*, 49 (2): 543–558.
- Thangaraj, K., Chaubey G., Kivisild T. et al.* 2005. 'Reconstructing the origin of Andaman islanders', *Science*, 308 (5724): 996.
- Thangaraj, K., Sridhar V., Kivisild T. et al.* 2006a. '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.
- Thangaraj, K., Chaubey, G., Reddy, A.G. et al.* 2006b. 'Unique origin of Andaman islanders: Insight from autosomal loci', *Journal of Human Genetics*, 51: 800–804.
- Thangaraj, K., Sridhar V., Singh, V.K. et al.* 2006c. 'Response to comment on "Reconstructing the origin of Andaman islanders"', *Science*, 311: 470b.
- Thangaraj, K., Sridhar V., Kivisild T. et al., Ayyasamy Vanniarajan, Ismail Thanseem, Alla G. Reddy and Lalji Singh.* 2006d. 'In situ origin of deep rooting lineages of mitochondrial macrohaplogroup M in India', *BMC Genomics*, 7: 151, doi:10.1186/1471–2164–7–151.
- Thanseem, I., Thangaraj, K., Chaubey G. et al.* 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.
- Toporov, V. N.* 1969. 'K voprosu o tipologičeskoj blizosti enisejskix jazykov i burušaski', pp. 217–220 in *Proisxoždenie aborigenov Sibiri i ix jazykov (Materialy mežvuzovskoj konferencii 11–13 maja 1969 g.)*. Tomsk.
- Toporov, V. N.* 1971. 'Burushaski and Yenisean languages: Some parallels', *Travaux Linguistiques de Prague*, 4: 107–125.
- Toynbee, A.* 1976 [posthumous]. *Mankind and Mother Earth: A Narrative History of the World*. London: Oxford University Press.
- Underhill P. A., Shen P., Lin A. A. et al.* 2000. 'Y chromosome sequence variation and the history of human populations', *Nature Genetics*, 26: 358–361.
- Underhill, P. A., G. Passarino, A. A., Lin, P. Shen et al.* 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., Myres, N. M., Rootsi S. et al.* 2010. 'Separating the post-glacial coancestry of European and Asian Y chromosomes within haplogroup R1a', *European Journal of Human Genetics*, 18: 479–484.
- Vajda, E. J.* 2010. 'A Siberian link with Na-Dene languages', pp. 33–99 in James Kari and Ben A Potter, eds., *The Dene-Yeniseian Connection*. Fairbanks: University of Alaska.
- Wadley, S. A.* 1996. 'Altaic influences on Beijing dialect: The Manchu case', *Journal of the American Oriental Society*, 116 (1): 99–104.
- Wang B., Zhang Y.B., Zhang F. et al.* 2011. 'On the origin of Tibetans and their genetic basis in adapting high-altitude environments', *Public Library of Science PLoS One*, 6(2): e17002.

- Wang, L., Abbot R. J., Zheng W. et al.* 2009. 'History and evolution of alpine plants endemic to the Qinghai-Tibetan plateau: *Aconitum gymnantrum* (Ranunculaceae), *Molecular Ecology*, 18: 709–721.
- Watters, D. E.* 2006. 'Notes on Kusunda grammar: A language isolate of Nepal', *Himalayan Linguistics Archive*, 3: 1–182.
- Weber, M.* 1919. *Politik als Beruf*. München: Verlag von Duncker und Humblot.
- Wen B., Li H., Lu D. et al.* 2004. 'Genetic evidence supports demic diffusion of Han culture', *Nature*, 431: 302–305.
- Winter, B.* 2010. 'A note on the higher phylogeny of Austronesian', *Oceanic Linguistics*, 49 (1): 282–287.
- Wolff, J. U.* 2010. *Proto-Austronesian Phonology with Glossary* (2 volumes). Ithaca, New York: Cornell Southeast Asia Program Publications.
- Wood E. T., Stover D. A., Ehret C. et al.* 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.
- Xu S., Li S., Yang Y. et al.* 2011. 'A genome-wide search for signals of high-altitude adaptations in Tibetans', *Molecular Biology and Evolution*, 28 (2): 1003–1011.
- Xue Y., Zerjal T., Bao W. et al.* 2006. 'Male demography in East Asia: A north-south contrast in human population expansion times', *Genetics*, 172: 2431–2439.
- Yan Si, Chuan-Chao, Wang, Hui Li et al.* 2011. 'An updated tree of Y-chromosome haplogroup O and revised phylogenetic positions of mutations P164 and PK4', *European Journal of Human Genetics*, 19: 1013–1015.
- Yi X., Liang Y., Huerta-Sanchez E. et al.* 2010. 'Sequencing of 50 human exomes reveals adaptation to high altitude', *Science*, 329 (5987): 75–78.
- Zegura, S. L., Karafet T. M., Zhivotovsky L. A., Hammer M. F.* 2003. 'High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas', *Molecular Biology and Evolution*, 21 (1): 164–175.
- Zeisler, B.* 2005. 'On the position of Ladakhi and Balti in the Tibetan language family', pp. 41–64 in John Bray, ed., *Ladakhi Histories: Local and Regional Perspectives*. Leiden: Brill.
- Zeisler, B.* 2009. 'Reducing phonological complexity and grammatical opaqueness: Old Tibetan as a lingua franca and the development of the modern Tibetan varieties', pp. 75–95 in Enoch O. Aboh and Norval Smith, eds., *Complex Process in New Languages*. Amsterdam: John Benjamins.
- Zeisler, B.* 2016. 'Book review: Thomas Owen-Smith and Nathan Wayne Hill (eds.). *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area* (Trends in Linguistics 226) 2014, Berlin/Boston: de Gruyter, viii,

444 pp. ISBN 978–3–11–031074. €95,95 (print or e-book)', *Journal of South Asian Languages and Linguistics*, 3 (2): 233–249.

Zerjal T., Dashnyam B., Pandya A. et al. 1997. 'Genetic relationships of Asian and Northern Europeans, revealed by Y chromosomal DNA analysis', *American Journal of Human Genetics*, 62: 1174–1183.

Zhao M., Kong Q.P., Wang H.W. et al. 2009. 'Mitochondrial genome evidence reveals successful Late Palaeolithic settlement on the Tibetan plateau', *Proceedings of the National Academy of Sciences*, 106 (50): 21230–21235.

Zhào Zhìjūn. 2005. *Zhíwù káogǔxué jíqí xīn jìnzhǎn* ['Palaeobotany and its recent progress'], *Káogǔ*, 2005 (7): 42–49.

Zhèngzhāng Shàngfāng. 2010. 'Càijiāhuà Báiyǔ kuānxì jí cígēn bǐjiào', pp. 389–400 in Pān Wǔyún and Shěn Zhōngwěi, eds., *Yánjūzhī Lè*, *The Joy of Research*, II. Shanghai: Shanghai Educational Publishing House.