JOURNAL OF THE ASIATIC SOCIETY

VOLUME LX No. 4 2018

THE ASIATIC SOCIETY
1 PARK STREET ❐ KOLKATA

ISSN 0368-3308

Edited and published by
Dr. Satyabrata Chakrabarti
General Secretary
The Asiatic Society
1 Park Street
Kolkata 700 016

Published in February 2019

Printed at
Desktop Printers
3A, Garstin Place, 4th Floor
Kolkata 700 001

Price : ₹ 400 (Complete vol. of four nos.)

CONTENTS

ARTICLES

The East Asian Linguistic Phylum : A Reconstruction Based on Language and Genes
George van Driem 1

Situating Buddhism in Mithila Region : Presence or Absence ?
Nisha Thakur 39

Another Inscribed Image Dated in the Reign of Vigrahapāla III
Rajat Sanyal 63

A Scottish Watchmaker — Educationist and Bengal Renaissance
Saptarshi Mallick 79

GLEANINGS FROM THE PAST

Notes on Charaka Sanhitá
Dr. Mahendra Lal Sircar 97

Review on Dr. Mahendra Lal Sircar's studies on Āyurveda
Anjalika Mukhopadhyay 101

BOOK REVIEW

Coin Hoards of the Bengal Sultans 1205-1576 AD from West Bengal, Bihar, Jharkhand, Assam and Bangladesh by Sutapa Sinha
Danish Moin 107

THE EAST ASIAN LINGUISTIC PHYLUM : A RECONSTRUCTION BASED ON LANGUAGE AND GENES

GEORGE VAN DRIEM

1. Trans-Himalayan

Mandarin, Cantonese, Hakka, Xiâng, Hokkien, Teochew, Pínghuṁ, Gṁn, Jṅn, Wú and a number of other languages and dialects together comprise the Sinitic branch of the Trans-Himalayan language family. These languages all collectively descend from a prehistorical Sinitic language, the earliest reconstructible form of which was called Archaic Chinese by Bernard Karlgren and is currently referred to in the anglophone literature as Old Chinese. Today, Sinitic linguistic diversity is under threat by the advance of Mandarin as a standard language throughout China because Mandarin is gradually taking over domains of language use that were originally conducted primarily in the local Sinitic languages. China is a demographic giant, and, within China, the speakers of Mandarin and other Sinitic languages together comprise the overwhelming majority of the population. As a consequence, the Trans-Himalayan family, to which Chinese belongs,is the second most populous family of languages in the world in terms of numbers of speakers.

This language family was first identified by Julius von Klaproth in 1823, who defined the family as consisting of Tibetan, Chinese, Burmese and all demonstrably related languages. This language family was originally called Tibeto-Burman in the British Isles, e.g. Hodgson (1857), Cust (1878), Forbes (1878), Houghton (1896). Most speakers of Trans-Himalayan languages live to the north and east of the Himalayas (Figure 1), whereas most of the over three hundred different languages and three fourths of the major Trans-Himalayan linguistic subgroups are located to the south of the Himalayan divide (Figure 2). Since 2004, the name Trans-Himalayan has increasingly come into use

because over time the original term "Tibeto-Burman" had come to be used in two opposing senses: both in its original meaning, but also in an historically and linguistically incorrect sense. The reason behind the neutral geographical name Trans-Himalayan is that the language family is spread across the Himalayas, both to the south and to the north of this greatest land barrier on our planet.

The use of the term "Tibeto-Burman"in two contradictory senses originated when a rival phylogenetic model named "Indo-Chinese" gained in popularity. This other theory of linguistic relationship differed from von Klaproth's well-informed Tibeto-Burman language family in that the Indo-Chinese construct contained all the languages of Asia and Oceania as far as Japan, Polynesia and Papua New Guinea. The theory was dreamt up by a Scotsman named John Caspar Leyden, who made a meteoric career as a British civil servant in Asia during the Napoleonic wars but then died at the age of 35 soon after he reached Java. The idea that all Asian and Oceanic languages shared some "common mixed origin" appealed to British colonial authorities, who were persuaded that they would be better able to rule over Asian peoples if a programme of linguistic research to understand Indo-Chinese language could be effectuated.

Over time, the hypothetical Indo-Chinese language family was whittled down in size, but the model also came to be tinged with racist overtones. Chinese has been a written language for millennia, during which the language served as the vehicle for an advanced civilisation. However, the eccentric and isolated position to which Chinese was relegated within the Indo-Chinese family tree was not based on any appreciation of the sophistication of Chinese culture, but on a racist appraisal of the Chinese language and people that arose amongst a particular breed of Western scholar at the time of the Opium Wars. The rebranding of the Indo-Chinese theory as "Sino-Tibetan" in 1924 helped to disguise these earlier racist underpinnings, but the model continued to treat Sinitic as a phylogenetic oddity.

Linguistically, the Sino-Tibetan model consisted of two branches, one of which was "Sino-Daic", and so represented a false family tree. When the Kradai languages were removed from "Sino-Tibetan", the reduced tree still represented a false phylogeny by sleight of relegating

all non-Sinitic languages to a single subgroup, which Sino-Tibetanists misleadingly labelled "Tibeto-Burman". No linguist has yet adduced any historical linguistic evidence that could unite this supposedly subordinate taxon into a single branch within the family tree. The Indo-Chinese or Sino-Tibetan model was assailed by scholars who proposed other models, e.g. Sino-Burman (Ramstedt 1957), Sino-Himalayan (Bodman 1973, 1980) and Sino-Kiranti (Starostin 1994).

Matisoff, who inherited the epistemologically flawed tree model from his mentor Benedict in the 1960s and championed this unsupported phylogeny for half a century, has at times retreated from face-saving denial and publicly recanted Sino-Tibetan on three occasions. This episode in linguistic history has already been recounted in greater detail elsewhere (van Driem 2014a, 2018). Yet today the model continues to mislead students and some scholars. Dispelling myths is an arduous task because of the tenacity with which such narratives take hold of the human mind. The default model was already presented by von Klaproth in 1823 in the form of his original Tibeto-Burman linguistic family, augmented by all the constituent linguistic subgroups which have since been recognised, some of which still remain to be properly validated (Figure 3).

Most of the speakers of Trans-Himalayan languages live within China, but most of the individual Trans-Himalayan languages are spoken exclusively outside of China. On the basis of evidence adduced by Schorer (2016), the Dura language of central Nepal has now been subsumed into the Magaric group, reducing the number of subgroups in the family from an erstwhile 42 to 41 branches. A total of 28 out of the 41Trans-Himalayan subgroups are found either exclusively or predominantly south of the Himalayan divide within the Indian subcontinent, viz. Tamangic, Newaric, Kiranti, Lepcha, Digarish, Lhokpu, Midźuish, Chepangic, Magaric, Tani, Siangic, Raji-Raute, Kho-Bwa, Ao, Zeme, Angami-Pochuri, Karbi, Mru, Brahmaputran, Gongduk, Hrusish, Black Mountain, Dhimalish, Tangkhul, Meithei, Pyu, Karenic and Kukish.

Sinitic languages are not the only Trans-Himalayan languages spoken inside China. In addition to Sinitic, six other Trans-Himalayan linguistic subgroups are found to the north and east of the Himalayas,

viz. Bái, Tǔjiā, rGyalrongic, Qiāngic, Ěkrsūish and Nàic.Sixbranches of the Trans-Himalayan family are represented by language communities distributed both on the north side and on the south flank of the Himalayas, viz. Tshangla, Bodish, Nungish, Lolo-Burmese, West Himalayish and Kachinic. Linguistic diversity south of the Himalayas may be even greater than shown in Figure 2. Ghale and Kaike together probably constitute a distinct linguistic subgroup, separate from Tamangic. Similarly, the diagram shown in Figure 3 subsumes the Northern Naga languages within Kachinic, a conjecture which still remains to be validated.

2. Austroasiatic

In 1823, von Klaproth recognised that Mon did not belong in the same phylum as Thai,Malay, Burmese, Chinese and Japanese, but that the Mon language, with respect to the other languages then documented, was in a phylum by itself. Francis Mason, born in York in 1799, immigrated across the Atlantic to Massachusetts in 1818, became a preacher in 1827 and was then sent to Burma by the American Baptist Missionary Union, where he worked on Pwo Karen and Sgaw Karen. In his 1854 article on the Mon language and then more explicitly in 1860, Mason identified a language family comprising both the Munda languages of India, such as Kol and Ho, and the "Talaing" language of Burma, i.e. Mon. Sir Arthur Purves Phayre (1873), the first Commissioner of British Burma from 1862 to 1867, popularised Mason's theory. This language family first went by the names of Mon-Annam or Mon-Khmer-Kolarian. In 1904, the Austrian priest Wilhelm Schmidt renamed the family Austroasiatic (Figure 4).

For much of the 20th century, scholars held that Austroasiatic consisted of a Munda and a Mon-Khmer branch (e.g. Cœdès 1948, Pinnow 1959, 1963). Austroasiatic languages of the Pakanic branch are spoken in Yúnnán province and northern Vietnam, and the attempt to determine the precise phylogenetic position of these languages within the family soon called into the question the structure of the Austroasiatic family tree as a whole (Benedict 1990). On the basis of chronologically ordered sound laws in his Austroasiatic database, Diffloth (2005) presented a revised phylogeny with the family's deepest

division lying between Munda in the west and Khasi-Aslian in the east. The Khasi-Aslian branch in turn split into Khasi-Pakanic and Mon-Khmer. The updated Austroasiatic family tree, showing the correct phylogenetic position of the Pearic branch for the first time, was presented by Diffloth at Agay in 2012 (Figure 5). The internal phylogeny of the Munda branch is the subject of ongoing research, but Heinz-Jūrgen Pinnow (1959, 1963) already divided Munda into South and North Munda, and his phylogeny was adopted by Zide and Zide (1976), who contributed a slight elaboration of the internal classification of South Munda.

3. Austro-Tai

Austronesian languages are spoken on Táiwā n and throughout insular Southeast Asia and beyond. The contours of the Austronesian language family first came into view when Frederick de Houtman, who spoke Malay after 26 months in captivity on Sumatra, reported that Malagasy was related to Malay. He had gleaned this insight from his cabin boy from Madagascar *die alreede goedt Duyts sprac: alsoo hy vvel vier Iaeren met onse Duytsche Natie ghevaren hadde* "who already spoke good Dutch, as he had sailed the seas with our Dutch nation for four years" (de Houtman 1603: v).

Malagasy is now known to be a memberof the Maanyan subgroup of the Barito river area in southern Borneo (Dahl 1951), and both linguistic and archaeological evidence suggested that the colonisation of Madagascar by the Malagasy took place between 400 and 700 AD (Dahl 1951, Dewar 1996, Adelaar1996), but the earliest radiocarbon dates for human settlements in Madagascar only date from the beginning of the 8th century, after Indian influence had begun to make itself felt in insular Southeast Asia. A population genetic study has now lent support to this late historical date for the first human settlement of Madagascar or even for a later initial settlement than previously thought (Pierron et al. 2017).

Jacob le Maire travelled throughout the Indo-Pacific for the Dutch East India Company during his circumnavigation of the earth in the years 1615 and 1616. He observed that the languages of Polynesia were related to Malay and the languages of the Indonesian archipelago. On

the basis of the observations and materials provided by de Houtman and le Maire, Adriaan van Reeland (1708: 55-139) established the genetic affinity between Malagasy, Malay and the Polynesian languages and identified Malayo-Polynesian as a language family.

In 1624, on the southwestern coast of Formosa, the Dutch established Fort Zelandia on a sandy islet that bore the name *Taioan* in the now extinct local Formosan language Siraya. This Siraya toponym was recorded in Dutch sources variously as *Taioan, Teyouvan, Teyoan, Tayouan, Taiyouhan* and *Taiyouan*. Shifting sands soon joined the islet of *Taioan* to the Formosan coast, forming an isthmus. On the basis of the catechisms and linguistic materials produced by the missionary Daniel Gravius on Táiwān between 1647 to 1651, von Klaproth (1822) included the Formosan languages in the Malayo-Polynesian family, a year before he published his influential *Asia Polyglotta*. In 1904, the Austrian priest Wilhelm Schmidt renamed the language family Austronesian, and Blust (2009) furnished what is currently the authoritative phylogenetic model of the Austronesian family tree (Figure 6).

Languages of the Kradai family are also spoken in China, where the distinctness of these languages from Chinese has always been appreciated and where the family is today called *Zhuàng-Dòng* in Mandarin. On the basis of root etyma, von Klaproth recognised that Thai was not related to Sinitic, once the obvious Chinese loan words had been sifted out of the lexicon. Kradai previously went by Shafer's name Daic or Benedict's coinage Tai-Kadai, but in Thai and in English the term กระได Kradai has established itself as the name for the language family (Figure 7).

Schlegel (1901, 1902) first proposed a genetic relationship between Siamese and Austronesian without, however, presenting any evidence. Benedict (1975) coined the term "Austro-Thai" for this relationship, but for Benedict (1942, 1975, 1990) this putative genetic link constituted just one ingredient in his grander and poorly supported proposals. Weera Ostapirat (2005, 2013) became the first to present sound comparative linguistic evidence that Kradai and Austronesian represent coordinate branches of a language family which he named Austro-Tai.

4. Hmong-Mien

Hmong-Mien is traditionally called *Miáo-Yáo* in Mandarin and has long been recognised as a distinct family of languages in China. André Georges Haudricourt (1954) wrote the first historical phonology of Hmong-Mien. Kun Chang (1972, 1976) wrote a reconstruction of the Hmong-Mien tonal system and of the Proto-Hmong inventory of initials, representing the culmination of many years of work reflected in his earlier publications. A reconstruction of Hmong-Mien was written by Purnell (1970), followed by the reconstructions of Wáng Fǔshì (1994) and Wáng Fǔshì and Máo Zōngwǔ (1995), which have all been superseded by Ratliff's (2010) Hmong-Mien reconstruction. In terms of its restricted internal diversity, the Hmong-Mien language family looked to Stanley Starosta like a vestigial branch of what once may have been a greater linguistic phylum, which Starosta (2005) called "Yangtzean" in recognition of the historical distribution of Hmong-Mien peoples along and the Yangtze and south of the Yangtze (Figure 8).

5. Northeastern Asia

Not all languages of East Asia belong to the phylum denominated as "East Asian" by Starosta. Altaic languages are spoken in the northern parts of East Asia, and their distinctness from Chinese has always been understood. Nicolaes Witsen (1692) and Phillip von Strahlenberg (1730) first mooted the contours of the Altaic language family. Engelbert Kæmpfer stressed that Japanese was unrelated to either Chinese or Ainu and proposed that Japanese was genetically related to Turkic (1729: 63-65). Julius von Klaproth (1823) identified Mongolic, Tungusic and Turkic as branches of a coherent linguistic phylum. However, he assigned Japanese and Korean each to their own linguistic phylum. Philipp von Siebold (1832a: 238-244) became the first to add Japanese to this Altaic language family, and he soon asserted that the Korean and Japanese languages within this family stemmed from the same shoot (1832b, 1, VII: 10). More recently, Robbeets (2010, 2014) coined the term "Trans-Eurasian" for the language family traditionally known as Altaic and reserves the term "Altaic" for the higher-order subgroup comprising just Turkic, Mongolic and Tungusic.

Languages of the Chukotko-Kamchakthan and Uralic families and a number of Palaeosiberian languages spoken in northeastern Asia likewise do not constitute part of Starosta's East Asian.

6. East Asian

Just as Ostapirat united the Kradai and Austronesian families into an Austro-Tai linguistic phylum, others have sought to unite recognised language families into larger linguistic phyla.Wilhelm Schmidt (1906) and Lawrence Reid (1994, 2005) sought to unite Austroasiatic and Austronesian into an Austric macrofamily, but the historical evidence does not yet hold up to scrutiny (cf. Diffloth 1994). Similarly, August Conrady (1916, 1922) and later Kurt Wulff (1934, 1942) proposed uniting Austroasiatic, Austronesian, Kradai and Tibeto-Burman, whereas Paul Benedict (1942), Robert Blust (1996) and Ilia Peiros (1998) proposed a linguistic phylum comprising Austroasiatic, Austronesian, Kradai and possibly Hmong-Mien.

Quite unlike Ostapirat's Austro-Tai, no methodologically rigorous study has been conducted adducing solid evidence for any of these proposals. Finally, Stanley Starosta (2005) proposed uniting the Austroasiatic, Austronesian, Kradai, Tibeto-Burman and Hmong-Mien language families into an East Asian linguistic phylum. The shared morphological vestiges adduced by Starosta in support of his East Asian linguistic phylum comprised the agentive prefix *<m->, the patient suffix *<-n>, what he called the instrumental prefix <s-> and what he termed the perfective prefix *<n->.

Other than Ostapirat, all these scholars have ventured beyond the epistemological constraints of what I call the "linguistic event horizon", representing the maximal time depth accessible through methodologically sound linguistic reconstruction and the boundary beyond which any reconstructions are at one point reduced to sheer speculation. Therefore, a discussion of the merits of the evidence advanced by Starosta for the East Asian linguistic phylum strikes me as being of little utility, since I consider the phylum to lie at the linguistic event horizon and therefore doubt whether this issue can ever be conclusively resolved on the basis of firmly reconstructible linguistic evidence. Rather, Starosta proposed 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" (2005: 194). A tweaked version of Starosta's East Asian family tree is the 2012 in Benares recension(van Driem 2014b), shown in Figure 9.

7. The Father Tongue correlation

We ought not to lose sight of the fact that a proto-language can be reconstructed solely on the basis of linguistic evidence and that the linguistic ancestors of any modern language community were not necessarily the same people as the community's biological forebears. Although these points have long been reiterated from the time of Julius von Klaproth (1823) and Max Müller (1872), these lessons are often lost on some audiences. By the same token, each of us has countless ancestors via numerous lineages. There is no such thing as a pure race. In fact, in molecular genetic terms there is no such thing as race (Cavalli-Sforza, Menozzi and Piazza 1994, van Driem 2017b).

We are all members of one large human family. Moreover, even when languages and genes happen to exhibit a correlation, such a marker relationship should not be confused with identity. The correlation of a particular chromosomal marker with the distribution of a certain language family must not be simplistically equated with populations speaking languages of a particular linguistic phylum. Rather, molecular markers on the Y chromosome serve as proxies or tracers for the movements of paternal ancestors.

When studying the distribution of maternally inherited markers in the mitochondrial DNA and paternally inherited markers on the Y chromosome, a Swiss-Italian team of population geneticists soon found that it was easier to find statistically relevant correlations between the language of a particular community and the paternally inherited markers prevalent in that community than between the language and the most salient maternally inherited markers found in that speech community. This Father Tongue correlation was first described by Poloni et al. (1997, 2000).

On the basis of this finding, it was inferred that paternally inherited polymorphisms may serve as markers for linguistic dispersals in the past, and that a correlation of Y chromosomal markers with language

may point towards male-biased linguistic intrusions. The Father Tongue correlation is ubiquitous but not universal. Its preponderance allows us to deduce that a mother teaching her children their father's tongue must have been a prevalent and recurrent pattern in linguistic prehistory.

There are a number of reasons why we might expect this outcome. Palaeolithic populations were small, and the effective founder population sizes of the major modern paternal subclades must have been quite small, whilst new populations arise from the small surviving subsets that have passed through bottlenecks. In fact, molecular evidence indicates that the Y chromosome underwent a global bottleneck towards the end of the last ice age, when certain paternal clades started eradicating or out-competing other clades (Karmin et al. 2015). The founding dispersals of many major language families appear to be related to the robust spread and reproductive success of the bearers of a subset of Y chromosomal haplogroups that survived such bottlenecks.

As a consequence, the global phylogeography of Y chromosomal haplogroups is shallower in terms of time depth than the worldwide mitochondrial landscape. The initial human colonisation of any virgin part of the planet must have involved both sexes in order for a population of progeny to establish itself. Once a population is in place, however, subsequent migrations could have been heavily gender-biased. Subsequently, male intruders could impose their language whilst availing themselves of the womenfolk already in place. In this regard, population geneticist Toomas Kivisild (2014) has wryly characterised warfare as a sex-specific pathology linked to the Y chromosome.

Whereas the landscape of paternal lineages often appears to correlate with language at the comparatively shallower time depth of the linguistically reconstructible past, correlations between maternal lineages and linguistic phylogeography discerned to date have been underwhelming.The Father Tongue hypothesis suggests that linguistic dispersals were, at least in most parts of the world, posterior to initial human colonisation and that many linguistic dispersals were predominantly later male-biased intrusions. Such patterns are observed worldwide.

The resolution of the Y chromosomal tree is constantly being enhanced. Haplogroup labels are updated to reflect our improved understanding of phylogeny. Mutations numbers tend to remain unchanged, provided that the markers in question prove to be reliable in defining haplogroups. Conventional haplogroup labels of the Y Chromosome Consortium are still widely in use, but have been replaced here with the newer labels of the International Society of Genetic Genealogy, reflecting refinements incorporated up to the 12th of May 2017.

Long before the linguistically reconstructible past, at a time that lay well beyond the linguistic event horizon, the paternal haplogroup K (M9) was centred in the area between South Asia and Southeast Asia, where the ancestral *K appears to have been situated. This clade spawned many successful paternal lineages, some of which moved into insular Southeast Asia, i.e. the haplogroups S(M69) and M(M304), whereas other clades moved back westward into South Asia and beyond, viz. the haplogroups Q(M242), R (M201), T(M89) and L(M429) (Karafet et al. 2015). The geographical locus of yet another descendant subclade lay in the Eastern Himalaya, i.e. the ancestral haplogroup NO (M214).

Millennia after the two paternal lineages N and O had split up, the bearers of haplogroup N set out for East Asia just after the last glacial maximum, braving ice and tundra, and — in a grand counterclockwise sweep — migrated across northern Eurasia as far as west as Lappland, whilst the ancestral form *N appears to have been situated in northern Burma, Yúnnán and Sìchuān (Rootsi et al. 2007, Derenko et al. 2007, Mirabal et al. 2009, Ilumäe *et al.* 2016).

In previous publications (van Driem 2014b, 2016, 2017b), I identified the clade N (M231) with the paternal spread of Fortescue's Uralo-Siberian linguistic phylum. Fortescue (1998, 2011, 2017) adduced evidence for an ancient circumpolar Uralo-Siberian linguistic phylum comprising Uralic, Yukagir, Eskimo-Aleut, Nivkh and Chukotko-Kamchatkan, an ancient linguistic relationship at the very horizon of the linguistically reconstructible past that is rendered more discernible when the known cases of language shift in the circumpolar region are carefully taken into account.

8. The East Asian dispersal

The molecular marker M175 defining the paternal clade O is overwhelmingly shared by the linguistic ancestors of what Starosta called the East Asian linguistic phylum, as observed in the cited studies, where it was proposed that the four major East Asian language families, i.e. Austroasiatic, Trans-Himalayan, Hmong-Mien and Austro-Tai, had resulted from prehistoric bottlenecks correlated with specific paternal lineages. The non-random correlation of the subclades of this particular Y chromosomal haplogroup with the four recognised language families enables us to infer the following sequence of events.

Millennia before the end of the last glacial maximum, the paternal lineage O (M175) split into the subclades O2 (M122) and O1 (F265, M1354), as shown in Figure 10. The two subclades can be putatively assigned to two geographical loci, with the haplogroup O1 (F265, M1354) moving eastward into East Asia south of the Yangtze, whilst bearers of the O2 (M122) haplogroup settled in the general region of the Eastern Himalaya. Subsequently, over the course of time, as temperature and humidity increased after the last glacial maximum, haplogroup O split further into the paternal lineages that serve as tracers for the spread of Trans-Himalayan, Hmong-Mien, Austroasiatic and Austro-Tai.

The O1 (F265, M1354) lineage south of the Yangtze split into the subclades O1b(M268) and O1a (M119), with the latter moving eastward to the Fújiàn hill tracts and across the strait to settle on Formosa, which so became the *Urheimat* of the Austronesians (cf. Abdulla et al. 2009). The founding dispersal of the Austro-Tai language familycan be traced through a correlation of the current geographical range of Austro-Tai languages with the chronology and spread of the molecular proxies defining the paternal haplogroups O1b (M268) and O1a (M119).

Subsequently, the paternal subclade O1b(M268) gave rise to the filial subclades O1b2(M176) and O1b1a1a(M95). The bearers of haplogroup O1b1a1a (M95)became the progenitors of the Austroasiatics (van Driem 2007). The Austroasiatics spread throughout the Salween drainage and thence to southern Yúnnán, northern Thailand and western Laos. In time, the Austroasiatics would spread as far as the Mekong delta, the Malay peninsula and the Nicobars,

Figure 1 : Geographical distribution of Trans-Himalayan languages (This map and Figures 3 through 7 are reproduced from van Driem 2015, with the gracious permission of Colin Pendry and Georg Miehe).

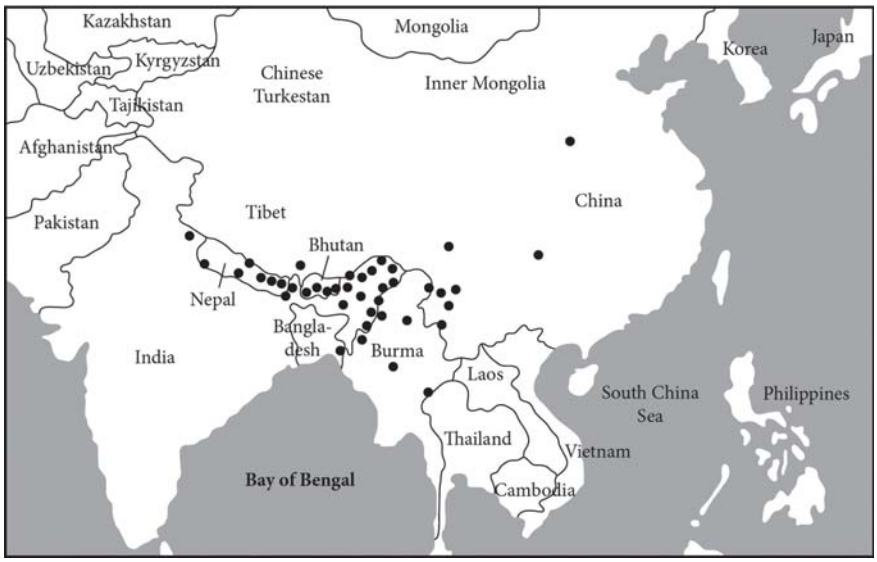

Figure 2 : Geographical distribution of Trans-Himalayan subgroups. Each dot represents not just one language, but the historical geographical centres of 41 major linguistic subgroups, each comprising anywhere between one language to several dozen closely related languages. Out of 41 linguistic subgroups, 28 lie to the south of the Himalayan divide, seven to the north and east, and six subgroups straddle both flanks of the Himalayas.

Figure 3 : The 41 major Trans-Himalayan linguistic subgroups or "fallen leaves" of the Trans-Himalayan language family. At variance with previous depictions, the diagram shows just 41 subgroups instead of 42. Schorer (2016) has shown Dura to be Magaric, thereby reducing the number of subgroups by one. Linguistic diversity south of the Himalayas is likely to be greater than shown here. Ghale and Kaike probably make up a linguistic subgroup distinct from Tamangic. Here the Northern Naga languages have been expediently subsumed within Kachinic, but this conjecture still remains to be validated. The number and arrangement of subgroups in this heuristic diagram will necessarily change over time as advances in historical linguistics are made and the structure of the family tree is discovered.

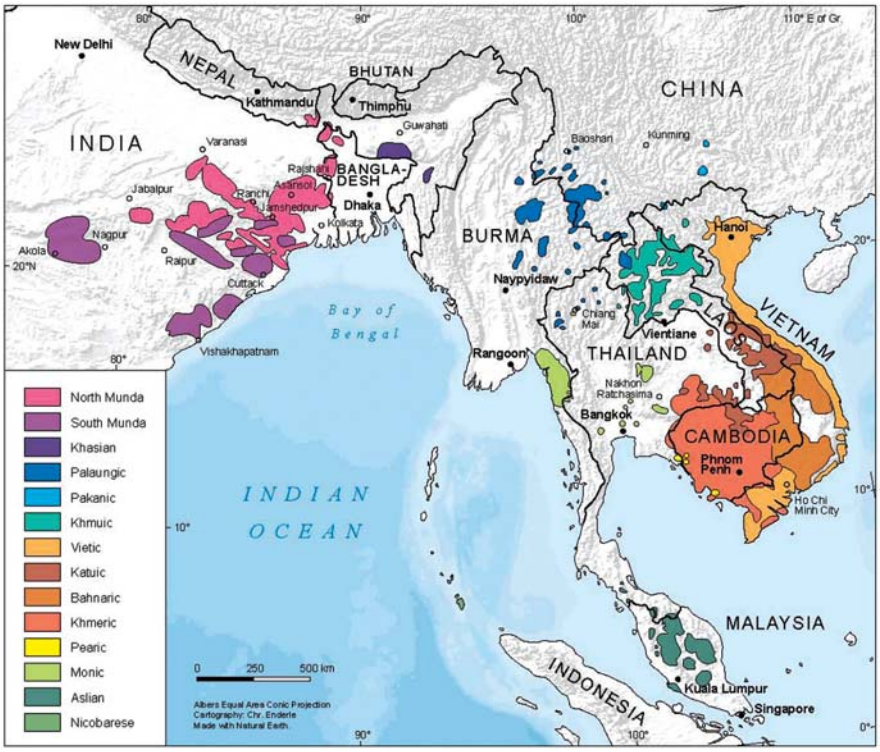

Figure 4 : Geographical distribution of Austroasiatic languages

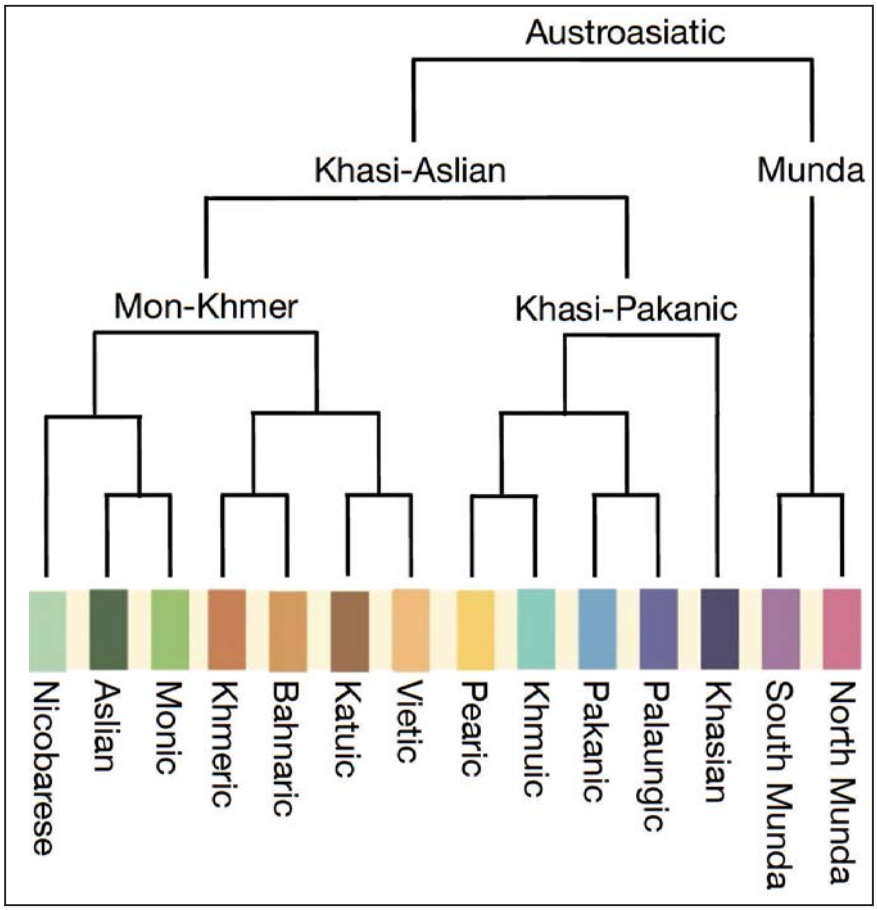

Figure 5 : The family tree of Austroasiatic(Diffloth 2012). Unlike the Khasi-Aslian branch, the internal phylogeny of the Munda branch has not been established.

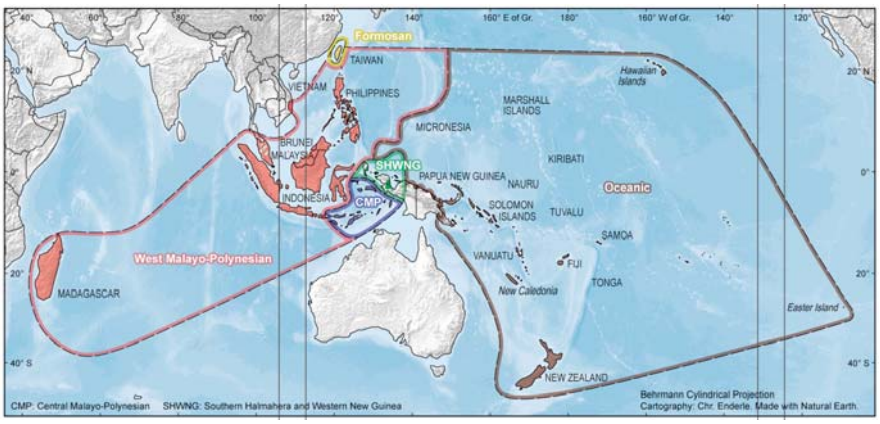

Figure 6 : Geographical distribution of Austronesian. The abbreviations CMP and SHWNG stand for the linguistic subgroups Central Malayo-Polynesian and South Halmahera West New Guinea respectively. Formosan on Táiwān comprises at least nine primary branches of the language family, whilst all other groups shown represent the later geographical dispersal of the single-branch Malayo-Polynesian, which branched into West Malayo-Polynesian and Central East Malayo-Polynesian. Subsequently, Central East split into Central and East. Finally, East Malayo-Polynesian branched into the subgroups South Halmahera West New Guinea and Oceanic.

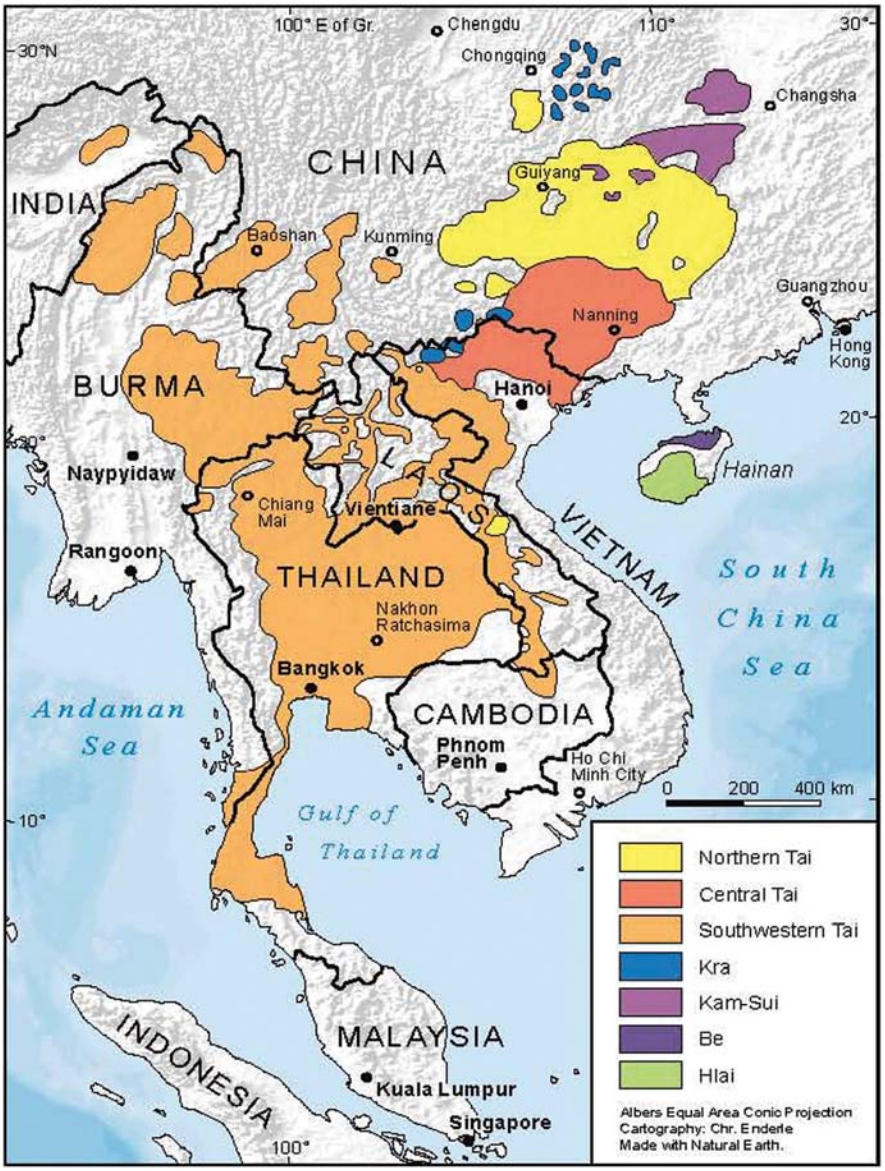

Figure 7 : Geographical distribution of Kradai languages. The spread of Southwestern Tai languages into mainland Southeast Asia first took place in historical times. Austronesian languages of the Monic group were still spoken in the Dvāravatī kingdom that flourished in what today is central Thailand until the 13th century.

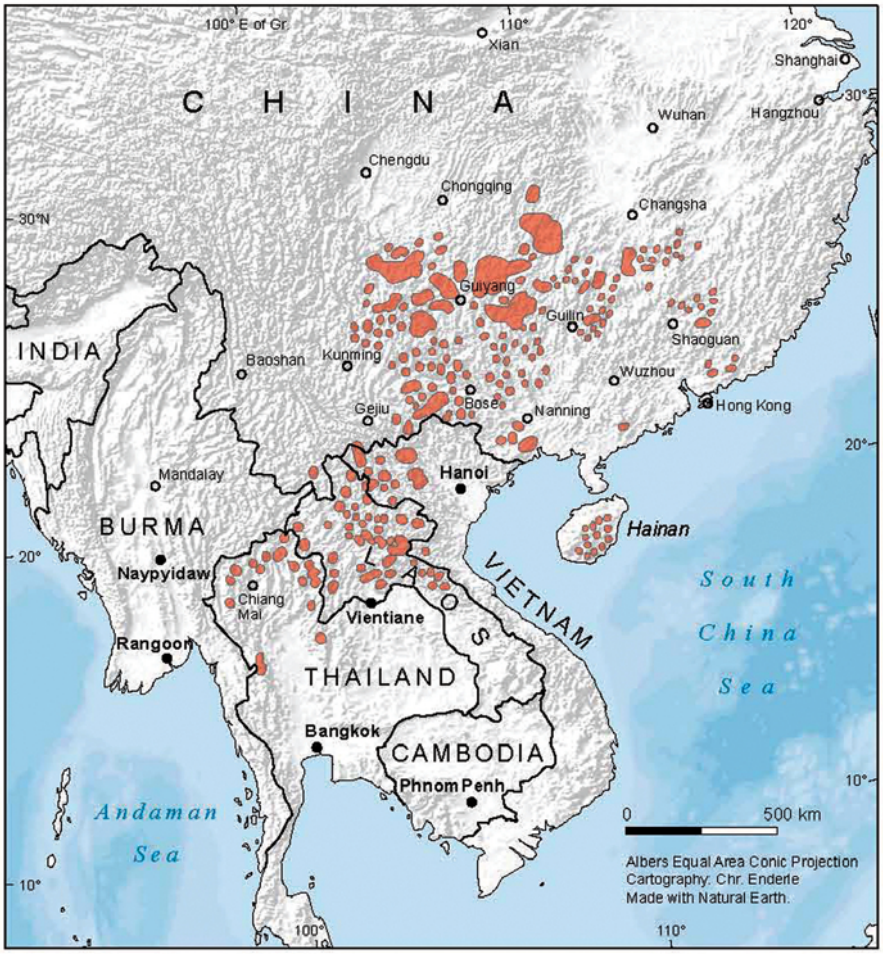

Figure 8 : Geographical distribution of Hmong-Mien languages. The Hmong-Mien migrated into mainland Southeast Asia relatively recently. Historically, Hmong-Mien peoples live along and south of the Yangtze.

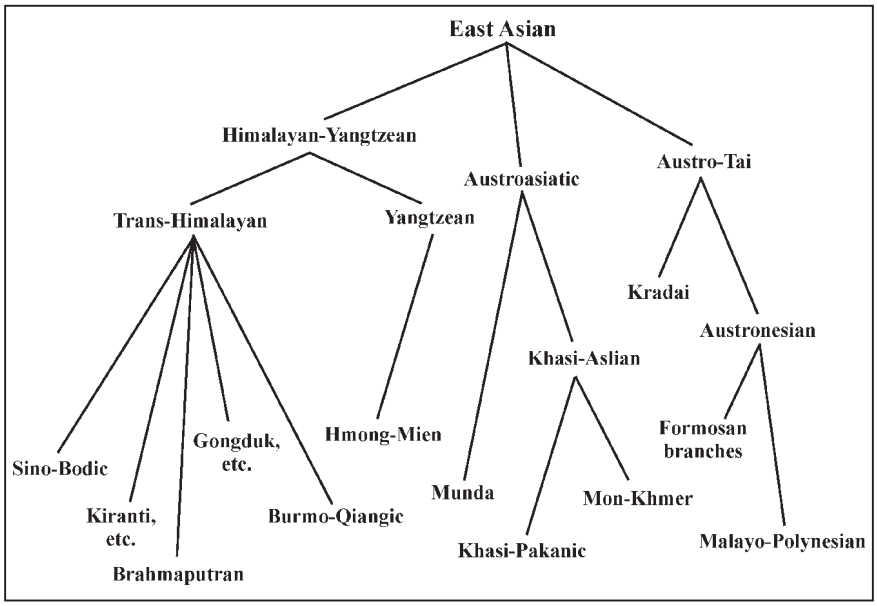

Figure 9 : The 2012 Benares recension of Stanley Starosta's East Asian linguistic phylum presented at Périgueux in 2001 (Starosta 2005, van Driem 2014b).

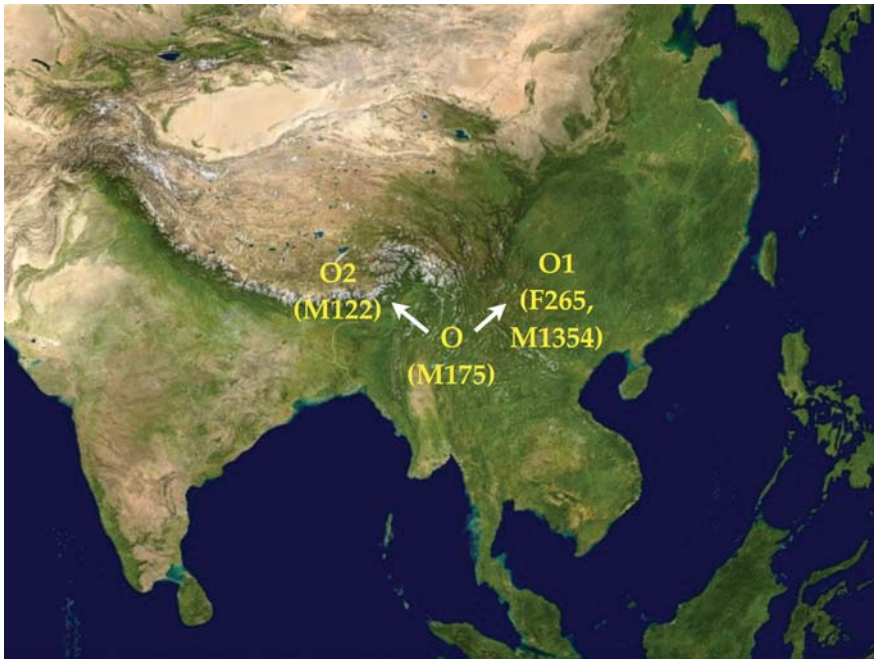

Figure 10 : After the last glacial maximum, the Y chromosomal haplogroup O (M175) split into the subclades O1 (F265, M1354) and O2 (M122).

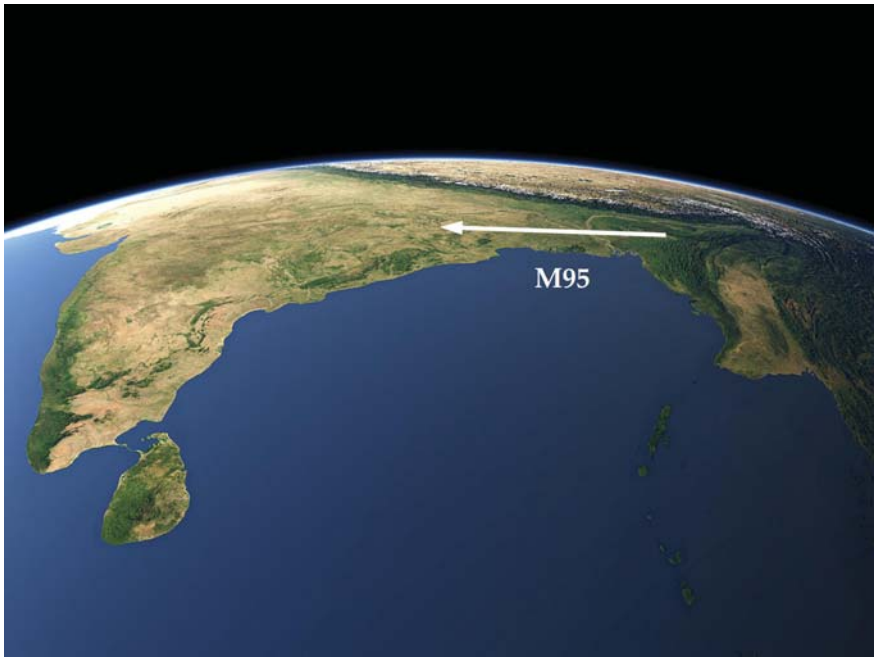

Figure 11 : A male-biased linguistic intrusion introduced both Austroasiatic language and a paternal lineage, haplogroup O1b1a1a (M95), into the indigenous population of the Choṭā Nāgpur.

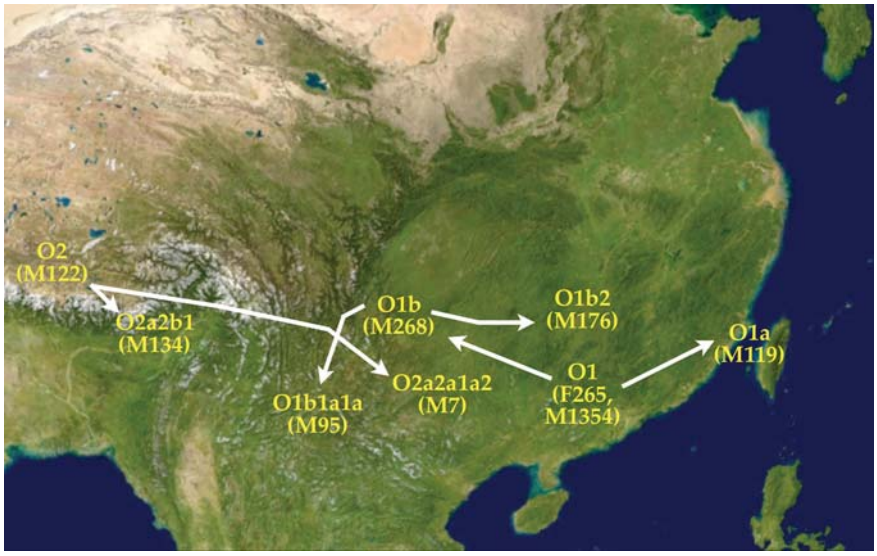

Figure 12 : At a more recent time depth, paternal lineages branched into new subclades, and each event involved a linguistic bottleneck leading to language families that today are reconstructible as distinct linguistic phyla. The O1 (F265, M1354) lineage gave rise to the O1a (M119) and O1b (M268) subclades. The former moved eastward to the Fújiàn hill tracts and across the strait to Formosa, which so became the *Urheimat* of the Austronesians. Bearers of the paternal lineage O1b (M268) domesticated Asian rice and spawned the paternal subclades O1b1a1a (M95) and O1b2 (M176). Haplogroup O1b1a1a (M95) is the Proto-Austroasiatic paternal lineage, whereas the para-Austroasiatic fraternal clade O1b2 (M176) spread eastward, sowing seed along the way. The haplogroup O2 (M122) gave rise to the paternal subclades O2a2b1 (M134) and O2a2a1a2 (M7). The spread of the molecular marker O2a2b1 (M134) from the Eastern Himalaya serves as a tracer for the dissemination of people speaking languages of the Trans-Himalayan family, whereas the paternal lineage O2a2a1a2 (M7) serves as a tracer for the spread of people speaking languages of the Hmong-Mien family.

and their paternal lineage would also spread deep into insular Southeast Asia. However, the prominent paternal lineage O1b2 (M176), which I have previously referred to as "para-Austroasiatic", does not appear to be correlated with any extant linguistic group today.

Secondarily, bands of male Austroasiatics introduced their language and their paternal lineage, O1b1a1a (M95), to the indigenous peoples of the Choṭā Nāgpur. Chaubey et al. (2010) showed that the Munda branch of Austroasiatic had arisen as the result of a sexually biased linguistic intrusion into the Indian subcontinent from the region to the north of the Bay of Bengal (Figure 11). As a consequence of the comparatively younger date and the nearly absolute gender asymmetry of this linguistic intrusion, it appears that the deepest division within the Khasi-Aslian trunk of Austroasiatic, i.e. the split between Khasi-Pakanic and Mon-Khmer, might be more indicative of the geographical location of the Austroasiatic homeland than the split between Munda and Khasi-Aslian. If we accept this line of reasoning, then the point of dispersal for Khasi-Aslian would appear to have lain in the area between South Asia proper and mainland Southeast Asia proper.

The spread of haplogroup O1 (F265, M1354) reflects the paternal founding dispersals of both Austro-Tai and Austroasiatic as well as the geographical spread of a para-Austroasiatic paternal subclade that evidently left no modern linguistic descendants. Our data from the Himalayan region and the data from populations elsewhere in Asia indicate that the geographical range and the chronology of spread of haplogroup O2a2b1 (M134) traces the founding dispersal of the Trans-Himalayan language family, whereas the paternal lineage O2a2a1a2 (M7) serves as a molecular proxy for the founding and spread of Hmong-Mien.

About twelve thousand years ago, at the dawn of the Holocene, in the southeastern Himalayas and eastern slopes of the Tibetan Plateau, haplogroup O2 (M122) gave rise to the ancestral Trans-Himalayan paternal lineage O2a2b1 (M134) and the "Yangtzean" or Hmong-Mien paternal lineage O2a2a1a2 (M7), as shown in Figure 12.It is a reasonable conjecture that the bearers of the polymorphism O2a2b1 (M134) at first remained in the Eastern Himalaya, which today also continues to represent the centre of phylogenetic diversity of the Trans-Himalayan

language family based on the geographical distribution of primary linguistic subgroups. After the bearers of the O2a2a1a2 (M7) lineage migrated eastward to settle in the areas south of the Yangtze, they were followed by early Trans-Himalayan language communities that spread from northeastern India into southeastern Tibet and northern Burma.

On their way, the early Hmong-Mien encountered the ancient Austroasiatics, from whom they adopted rice agriculture. Three principal populations of cultivated rice *Oryza sativa* are distinguished, comprising the families of cultivars known as *ahu*, *indica* and *japonica* rice. Earlier population genetic research on the genome of wild and cultivated varieties of rice supported the hypothesis that Asian rice was domesticated twice (Kovach *et al.* 2007, Sweeney and McCouch 2007, Kovach *et al.* 2009), but molecular evidence adduced in subsequent work demonstrated that the three principal cultivars represent three independent domestication processes, and that the introgression of domesticated traits occurred not just unidirectionally from *japonica* into *ahu* and *indica* rice, but multidirectionally from *ahu* and *indica* into *japonica* as well (Londo *et al.* 2006, McNally*et al.* 2009, Civáň*et al.* 2015).

In previous studies (van Driem 2011, 2012, 2017c), evidence from linguistic palaeontology, rice genetics and human population genetics enabled the ethnolinguistic identification of the domesticators of Asian rice with the early Yangtzeans, the ancient Austroasiatics and the bearers of para-Austroasiatic haplogroup O1b2 (M176). The latter para-Austroasiatic paternal lineage advanced as far as the Korean peninsula and also represents a major wave of immigration recorded in the Japanese genome, representing the probable lineage of the Yayoi people, who introduced rice agriculture to Japan, as early as the second millennium BC, during the final phase of the Jōmon period.

The intimate interaction between ancient Austroasiatics and the ancestral Hmong-Mien not only involved the sharing of knowledge about rice agriculture, but also left a genetic trace in the high frequencies of haplogroup O1b1a1a (M95) in today's Hmong-Mien language communities and of haplogroup O2a2a1a2 (M7) in today's Austroasiatic populations. In an earlier study on Y chromosomal

haplogroups, Cai et al. (2011: 8) once ventured to speculate about "a Mon-Khmer origin of Hmong-Mien populations". Instead, the incidence of haplogroup O2a2a1a2 (M7) in Austroasiatic language communities of Southeast Asia evinces a significant Hmong-Mien paternal contribution to early Austroasiatic populations, whereas the incidence of haplogroup O1b1a1a (M95) amongst the Hmong-Mien indicates a comparatively lower Austroasiatic paternal contribution to Hmong-Mien populations. As the Hmong-Mien moved eastward, the bearers of para-Austroasiatic paternal lineage O1b2 (M176) likewise continued to move east.

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

Finally, the ancestral Trans-Himalayan paternal lineage O2a2b1 (M134) spread further northeast to the North China plain. The complex history of Sinitic populations featured successive constellations of dynastic empires governed from geographically ever shifting capitals, whereby subjugated and neighbouring populations as well as immigrants were absorbed. Not surprisingly therefore, Hàn Chinese populations tend to represent an amalgam of East Asian paternal lineages. Yet even in modern Hàn Chinese populations, the molecular marker associated with the spread of a Trans-Himalayan father tongue from the eastern Himalayan region, i.e. haplogroup O2a2b1 (M134), taken together with its subclade O2a2b1a1 (M117), occurs in a much higher frequency than any other O haplogroup subclade, and approximately twice as frequently as the next most

frequent fraternal subclade O2a1c (002611) (Yan et al. 2011, Wang et al. 2013, Yao et al. 2017).

On the grand time scale of the present narrative, China as a cultural and linguistic entity represents only a rather recent phenomenon, and Kwang-chih Chang (1986: 242) cautioned us against anachronisms that arise from applying the label "Chinese" to archaeological cultural assemblages or to peoples of the distant past. Likewise, in his valuable pioneering interdisciplinary studies, Bill Wang (1998) stressed the distinctness of the"three windows on the past"afforded by archaeology, linguistics and genetics.

Only at a much shallower time depth did the Trans-Himalayan paternal lineage O2a2b1 (M134) spread in tandem with early Sinitic speaking populations southward from the Yellow River basin into southern China during the Qín dynasty in the third century BC (Mountain et al. 1992, Wen et al. 2004). The martial and male-biased historical spread of Hm̀n Chinese during the gradual cultural sinification of the region south of the Yangtze involved both the spread of language and the introduction of paternal lineages and is historically documented in the Chinese chronicles. The paternal lineage O2a2b1 (M134) is also intrusively present in the Korean peninsula and beyond.

Not only do the geographical distribution of Trans-Himalayan linguistic subgroups and the chronology and spread of Y chromosomal haplogroups take us back to an eastern Himalayan homeland and the northeastern part of the Indian subcontinent, basic facts of geography in light of our ultimate African origins prompt us to direct our gaze back to the East Himalaya. The sub-Himalayan hill tracts offered a rich corridor replete with lush habitats for early human populations at a time that our ancestors still practised hunting and foraging as their principal subsistence strategies. The presence of Y chromosomal haplogroup D and other molecular markers in the Himalayas likewise attest to numerous older waves of peopling that passed through this region.

References

Abdulla, Mahmood Ameen, and Ikhlak Ahmed, Anunchai Assawamakin, Jong Bhak, Samir K. Brahmachari, Gayvelline C. Calacal, Amit Chaurasia, Chien-Hsiun Chen, Jieming Chen, Yuan-Tsong Chen, Jiayou Chu, Eva Maria C.

Cutiongco-de la Paz, Maria Corazon A. de Ungria, Frederick C. Delfin, Juli Edo, Suthat Fuchareon, Ho Ghang, Takashi Gojobori, Junsong Han, Sheng-Feng Ho, Boon Peng Hoh, Wei Huang, Hidetoshi Inoko, Pankaj Jha, Timothy A. Jinam, Li Jin, Jongsun Jung, Daoroong Kangwanpong, Jatupol Kampuansai, Giulia C. Kennedy, Preeti Khurana, Hyung-Lae Kim, Kwangjoong Kim, Sangsoo Kim, Woo-Yeon Kim, Kuchan Kimm, Ryosuke Kimura, Tomohiro Koike, Supasak Kulawonganunchai, Vikrant Kumar, Poh San Lai, Jong-Young Lee, Sunghoon Lee, Edison T. Liu, Partha P. Majumder, Kiran Kumar Mandapati, Sangkot Marzuki, Wayne Mitchell, Mitali Mukerji, Kenji Naritomi, Chumpol Ngamphiw, Norio Niikawa, Nao Nishida, Bermseok Oh, Sangho Oh, Jun Ohashi, Akira Oka, Rick Ong, Carmencita D. Padilla, Prasit Palittapongarnpim, Henry B. Perdigon, Maude Elvira Phipps, Eileen Png, Yoshiyuki Sakaki, Jazelyn M. Salvador, Yuliana Sandraling, Vinod Scaria, Mark Seielstad, Mohd Ros Sidek, Amit Sinha, Metawee Srikummool, Herawati Sudoyo, Sumio Sugano, Helena Suryadi, Yoshiyuki Suzuki, Kristina A. Tabbada, Adrian Tan, Katsushi Tokunaga, Sissades Tongsima, Lilian P. Villamor, Eric Wang, Ying Wang, Haifeng Wang, Jer-Yuarn Wu, Huasheng Xiao, Shuhua Xu, Jin Ok Yang, Yin Yao Shugart, Hyang-Sook Yoo, Wentao Yuan, Guoping Zhao, Bin Alwi Zilfalil and the Indian Genome Variation Consortium. 2009. Mapping human genetic diversity in Asia, *Science,* 326: 1541-1545.

Adelaar, Karl Alexander ("Sander"). 1996. Malagasy culture-history: Some linguistic evidence,pp. 487-500 in Julian Reade, ed., *The Indian Ocean in Antiquity.* London: Kegan Paul International.

Benedict, Paul King. 1942. Thai, Kadai, and Indonesia: A new alignment in southeastern Asia, *American Anthropologist,* 44: 576-601.

Benedict, Paul King. 1975. *Austro-Thai: Language and Culture, with a Glossary of Roots.* New Haven, Connecticut: Human Relations Area Files Press.

Benedict, Paul King. 1990. How to tell Lai: An exercise in classification, *Linguistics of the Tibeto-Burman Area*, 13 (2): 1-26.

Blust, Robert. 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, Robert. 2009. *The Austronesian Languages.* Canberra: Pacific Linguistics.

Bodman, Nicholas Cleaveland. 1973. Some Chinese reflexes of Sino-Tibetan s-clusters, *Journal of Chinese Linguistics*, 1 (3): 383–396.

Bodman, Nicholas Cleaveland. 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.

Cai Xiaoyun, Zhendong Qin, Bo Wen, Shuhua Xu, Yi Wang, Yan Lu, Lanhai Wei, Chuanchao Wang, Shilin Li, Xingqiu Huang, Huiń Li and the Genographic Consortium. 2011. Human migration through bottlenecks from Southeast

Asia into East Asia during Last Glacial Maxiumum revealed by Y chromosomes, *Public Library of Science,* 6 (8): e24282.

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

Chang Kun (Zhāng Kūn 張琨).1972. The reconstruction of Proto-Miao-Yao tones, *Bulletin of the Institute of History and Philology*, 44:541-628.

Chang Kun (Zhāng Kūn 張琨).1976. Proto-Miao initials, *Bulletin of the Institute of History and Philology*, 47:155-218.

Chang Kwang-Chih (張光直 Zhāng Guāngzhí). 1986. *The Archaeology of Ancient China* (4th edition).New Haven, Connecticut: Yale University Press.

Chaubey, Gyaneshwer, and Mait Metspalu, Ying Choi, Reedik Māgi, Irene Gallego Romero, Siiri Rootsi, Pedro Soares, Mannis van Oven, Doron M. Behar, Siiri Rootsi, Georgi Hudjashov, Chandana Basu Mallick, Monika Karmin, Mari Nelis, Jūri Parik, Alla Goverdhana Reddy, Ene Metspalu, George van Driem, Yali Xue, Chris Tyler-Smith, Kumarasamy Thangaraj, Lalji Singh, Maido Remm, Martin B. Richards, Marta Mirazon Lahr, Manfred Kayser, Richard Villems and Toomas Kivisild.2010. Population genetic structure in Indian Austroasiatic speakers: The role of landscape barriers and sex-specific admixture, *Molecular Biology and Evolution,* 28 (2): 1013-1024.

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

Cœdès, Georges. 1948. Les langues de l'Indochine, *Conférences de l'Institut de Linguistique de l'Université de Paris*, VIII: 63-81.

Conrady, August. 1916. Eine merkwürdige Beziehung zwischen den austrischen und den indochinesischen Sprachen, pp. 475-504in *Aufsātze zur Kultur- und Sprachgeschichte vornehmlich des Orients: Ernst Kuhn zum 70. Geburtstage am 7.Februar 1916 gewidmet von Freunden und Schulern*. Munchen: Verlag von M. & H. Marcus.

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

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

Dahl, Otto Christian. 1951. *Malgache et maanyan: une comparaison linquistique* (Avhandlinger utgitt av Instituttet, 3). Oslo: Egede Instituttet.

Derenko, Miroslava, Boris Malyarchuk, Galina Denisova, Marcin Wozniak, Tomasz Grzybowski, Irina Dambueva and Ilia Zakharov. 2007. Y-chromosome haplogroup N dispersal from south Siberia to Europe, *Journal of Human Genetics*, 52: 763-770.

Dewar, Robert E. 1996. The archaeology of the early settlement of Madagascar, pp. 471-486 in Julian Reade, ed., *The Indian Ocean in Antiquity.* London: Kegan Paul International.

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

Diffloth, Gérard. 2005. The contribution of linguistic palaeontology to the homeland of Austroasiatic, pp. 77-80 in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together the Archaeology, Linguistics and Genetics*.London: Routledge Curzon.

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

van Driem, George. 2005. Tibeto-Burman vs. Indo-Chinese: Implications for population geneticists, archaeologists and prehistorians, pp. 81-106 in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds. *The Peopling of East Asia: Putting Together the Archaeology, Linguistics and Genetics.* London: Routledge Curzon.

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

van Driem, George. 2011. 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, George. 2012. The ethnolinguistic identity of the domesticators of Asian rice, *Comptes Rendus Palīvol,* 11 (2): 117-132.

van Driem, George. 2014a. Trans-Himalayan, pp. 11-40 in Nathan Hill and Thomas Owen-Smith, eds., *Trans-Himalayan Linguistics*. Berlin: Mouton de Gruyter.

van Driem, George. 2014b. A prehistoric thoroughfare between the Ganges and the Himalayas, pp. 60-98 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.

van Driem, George. 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.

van Driem, George. 2016. The Eastern Himalayan corridor in prehistory, pp. 467-524, Vol. II in Elena Nikolaevna Kolpačkova, ed., *Проблемы китайского и общего языкознания — Problems in Chinese and General Linguistics*. St. Petersburg: Izdatel'stvo Studija NP-Print.

van Driem, George. 2017a. The epistemological event horizon in archaeology, pp. xv-xvii in Manjil Hazarika, *Prehistory and Archaeology of Northeast India.* New Delhi: Oxford University Press.

van Driem, George. 2017b.The Eastern Himalaya and the Mongoloid myth, pp. 12-41 in Sanjoy Hazarika and Reshmi Banerjee, eds., *Gender, Poverty and Livelihood in the Eastern Himalaya*. Abingdon: Routledge.

van Driem, George. 2017c.The domestications and the domesticators of Asian rice, pp. 183-214 in Martine Robbeets, ed., *Language Dispersal Beyond Farming.* Amsterdam: John Benjamins.

van Driem, George. 2018. Linguistic history and historical linguistics, *Linguistics of the Tibeto-Burman Area,* 41 (1): 106-127.

Forbes, Charles James F. S. 1878. On Tibeto-Burman languages,*Journal of the Royal Asiatic Society of Great Britain and Ireland,* 10: 210–227.

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

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

Fortescue, Michael. 2011. The relationship of Nivkh to Chuktko-Kamchatkan revisited, *Lingua,* 121: 1359-1376.

Fortescue, Michael. 2017. Correlating Palaeo-Siberian languages and populations: Recent advances in the Uralo-Siberian hypothesis, *Man In India*, 97 (1): 47-68.

Haudricourt, André Georges. 1954. Introduction à la phonologie historique des langues miao-yao, *Bulletin de l'École Francaise d'Extréme-Orient,* XLIV : 555-574.

Hodgson, Brian Houghton. 1857. Comparative Vocabulary of the Languages of the broken Tribesof Nīpál, *Journal of the Asiatic Society of Bengal*, 26 : 317–371.

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

de Houtman, Frederick. 1603. *Spraeck ende woord-boeck, Inde Maleysche ende Madagaskarsche Talen met vele Arabische ende Turcsche Woorden: Inhoudende twaelf tsamensprekinghen inde Maleysche ende drie inde Madagaskarsche Spraken met alderhande woorden endenamen, ghestelt naer ordre vanden A.B.C. alles int Nederduytsch verduytst.* Amsterdam:Jan Evertsz. Cloppenburch.

Huang, Xuehui, and Nori Kurata, Xinghua Wei, Zi-Xuan Wang, Ahong Wang, Qiang Zhao, Yan Zhao, Kunyan Liu,Hengyun Lu, Wenjun Li, Yunli Guo, Yiqi Lu, Congcong Zhou, Danlin Fan, Qijun Weng, Chuanrang Zhu, Tao Huang, Lei Zhang, Yongchun Wang, Lei Feng, Hiroyasu Furuumi, Takahiko Kubo, Toshie Miyabayashi, Xiaoping Yuan, Qun Xu,Guojun Dong, Qilin Zhan, Canyang Li, Asao Fujiyama, Atsushi Toyoda, Tingting Lu, Qi Feng, Qian Qian, Jiayang Li and Bin Han. 2012. A map of rice genome variation reveals the origin of cultivated rice, *Nature,* 490: 497-503.

Ilumäe, Anne-Mai, and Maere Reidla, Marina Chukhryaeva, Mari Järve, Helen Post, Monika Karmin, Lauri Saag, Anastasiya Agdzhoyan, Alena Kushniarevich, Sergey Litvinov, Natalya Ekomasova, Kristiina Tambets, Ene Metspalu, Rita Khusainova, Bayazit Yunusbayev, Elza K. Khusnutdinova, Ludmila P. Osipova, Sardana Fedorova, Olga Utevska, Sergey Koshel, Elena Balanovska, Doron M. Behar, Oleg Balanovsky, Toomas Kivisild, Peter A. Underhill, Richard Villems and Siiri Rootsi. 2016. Human Y chromosome haplogroup N:A non-trivial time-resolved phylogeography that cuts across language families, *American Journal of Human Genetics*, 99: 163-173.

Kæmpfer, Engelbert. 1729. *De Beschryving van Japan, behelsende een Verhaal van den Ouden en Tegenwoordigen Staat en Regeering van dat Ryk, van deszelfs Tempels, Paleysen, Kasteelenen andere Gebouwen; van deszelfs Metalen, Mineralen, Boomen, Planten, Dieren, Vogelenen en Visschen, van de Tydrekening, en Opvolging van de Geestelyke en Wereldlyke Keyzers, van de Oorspronkelyke Afstamming, Godsdiensten, Gewoonten en Handwerkselen der Inboorlingen, en van hunnen Koophandel met de Nederlanders en de Chineesen, beneven seene Beschryving van het Koningryk Siam*. Amsterdam: Balthasar Lakeman.

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

Karmin, Monika, and Lauri Saag, Mário Vicente, Melissa A. Wilson Sayres,Mari Järve, Ulvi Gerst Talas, Siiri Rootsi, Anne-Mai Ilumäe, Reedik Mägi,Mario Mitt, Luca Pagani, Tarmo Puurand, Zuzana Faltyskova, Florian Clemente, Alexia Cardona, Ene Metspalu, Hovhannes Sahakyan, Bayazit Yunusbayev, Georgi Hudjashov, Michael DeGiorgio, Eva-Liis Loogväli, Christina Eichstaedt, Mikk Eelmets, Gyaneshwer Chaubey, Kristiina Tambets, Sergei Litvinov, Maru Mormina, Yali Xue, Qasim Ayub, Grigor Zoraqi, Thorfinn Sand Korneliussen, Farida Akhatova, Joseph Lachance, Sarah Tishkoff, Kuvat Momynaliev, François-Xavier Ricaut, Pradiptajati Kusuma, Harilanto Razafindrazaka, Denis Pierron, Murray P. Cox, Gazi Nurun Nahar Sultana, Rane Willerslev, Craig Muller, Michael Westaway, David Lambert, Vedrana Skaro, Lejla Kovačević, Shahlo Turdikulova, Dilbar Dalimova, Rita Khusainova, Natalya Trofimova, Vita Akhmetova, Irina Khidiyatova, Daria V. Lichman, Jainagul Isakova, Elvira Pocheshkhova, Zhaxylyk Sabitov, Nikolay A. Barashkov, Pagbajabyn Nymadawa, Evelin Mihailov, Joseph Wee Tien Seng, Irina Evseeva, Andrea Bamberg Migliano, Syafiq Abdullah, George Andriadze, Dragan Primorac, Lubov Atramentova, Olga Utevska, Levon Yepiskoposyan, Damir Marjanović, Alena Kushniarevich, Doron M. Behar, Christian Gilissen, Lisenka Vissers, Joris A. Veltman, Elena Balanovska, Miroslava Derenko, Boris Malyarchuk, Andres Metspalu, Sardana Fedorova, Anders Eriksson, Andrea Manica, Fernando L. Mendez, Tatiana M. Karafet, Krishna R. Veeramah, Neil Bradman, Michael F. Hammer, Ludmila P. Osipova, Oleg Balanovsky, Elza K. Khusnutdinova, Knut Johnsen, Maido Remm, Mark G. Thomas, Chris Tyler-Smith, PeterA. Underhill, Eske Willerslev, Rasmus Nielsen, Mait Metspalu,Richard Villems and Toomas Kivisild. 2015. A recent bottleneck of Y chromosome diversity coincides with a global change in culture, *Genome Research*, 25 (4): 459-466.

Kivisild, Toomas. 2014. Mapping Chromosome diversity in Eurasia, paper presented at the conference *Migrations and Transfers in PrehistoryAsian and Oceanic Ethnolinguistic Phylogeography*, University of Bern, 28 to 30 July 2014.

von Klaproth, Julius Heinrich. 1822. Sur la langue des indigènes de l'Île de Formose, *Journal Asiatiqué*, I : 193-202.

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

Kovach, Michael J., Megan T. Sweeney and Susan McCouch. 2007. New insights into the history of rice domestication, *Trends in Genetics*, 23 (11): 578-587.

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

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

Mason, Francis. 1854. The Talaing language, *Journal of the American Oriental Society*, IV: 277-289.

Mason, Francis. 1860. *Burmah, Its people and Natural Productions, or Notes on the Nations, Fauna, Flora and Minerals of Tenasserim, Pegu and Burmah*. Rangoon: Thomas Stowe Ranney.

McNally,Kenneth L., and Kevin L. Child, Regina Bohnert, Rebecca M. Davidson, Keyan Zhao, Victor J. Ulat,Georg Zeller, Richard M. Clark, Douglas R. Hoen, Thomas E. Bureau, Renée Stokowski, Dennis G. Ballinger,Kelly A. Frazer, David R. Cox, Badri Padhukasahasram, Carlos D. Bustamante, Detlef Weigel, David J. Mackill,Richard M. Bruskiewicha, Gunnar Rätsch, C. Robin Buell, Hei Leung and Jan E. Leach. 2009. Genome-wide SNP variation reveals relationships among landraces and modern varieties of rice, *Proceedings of the National Academy of Sciences of the United States of America,* 106 (30): 12273-12278.

Mirabal, Sheyla, Maria Reguiero, Alicia M. Cadenas, Luigi Luca Cavalli-Sforza, PeterA. Underhill, Dmitry A. Verbenko, Svetlana A. Limborska and René J. Herrera. 2009. Y chromosome distribution within the geolinguistic landscape of northwestern Russia, *European Journal of Human Genetics*, 2009: 1-14.

Mountain, Joanna L., William Shì-Yuán Wáng, Du Ruofu, Yuan Yida and Luigi Luca Cavalli-Sforza. 1992. Congruence of genetic and linguistic evolution in China, *Journal of Chinese Linguistics*, 20 (2): 315-331.

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

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

Peiros, Ilia. 1998. *Comparative Linguistics in Southeast Asia.* Canberra: PacificLinguistics.

Phayre, Arthur Purves. 1873. On the history of Pegu, *Journal of the Asiatic Society of Bengal*, XLII(I): 23-57.

Pierron, Denis, and Margit Heiske, Harilanto Razafindrazaka, Ignace Rakoto, Nelly Rabetokotany, Bodo Ravololomanga, Lucien M.-A. Rakotozafy, Mireille Mialy Rakotomalala, Michel Razafiarivony, Bako Rasoarifetra, Miakabola Andriamampianina Raharijesy, Lolona Razafindralambo, Ramilisonina, Fulgence Fanony, Sendra Lejamble, Olivier Thomas, Ahmed Mohamed Abdallah, Christophe Rocher, Amal Arachiche, Laure Tonaso, Veronica Pereda-loth, Stéphanie Schiavinato, Nicolas Brucato, Francois-Xavier Ricaut, Pradiptajati Kusuma, Herawati Sudoyo, Shengyu Ni, Anne Boland, Jean-Francois Deleuze, Philippe Beaujard, Philippe Grange, Sander Adelaar, Mark Stoneking, Jean-Aimé Rakotoarisoa, Chantal Radimilahy and Thierry Letellier. 2017. Genomic landscape of human diversity across Madagascar, *Proceedings of the National Academy of Sciences* <www.pnas.org/cgi/doi/10.1073/pnas.1704906114>.

Pinnow, Heinz-Jürgen. 1959. *Versuch einer historischen Lautlehre der Kharia-Sprache.* Wiesbaden : Otto Harrassowitz.

Pinnow, Heinz-Jürgen. 1963. The position of the Munda languages within the Austroasiatic language family, pp. 140-152 in H.L. Shorto, ed., *Linguistic Comparison in South East Asia and the Pacific.* London : School of Oriental and African Studies.

Poloni, Estella Simone, Ornella Semino, Giuseppe Passarino, A.S. Santachiara-Benerecetti, I. Dupanloup, André Langaney and Laurent Excoffier. 1997. Human genetic affinities for Y chromosome P49a, f/*Taq*I haploptypes show strong correspondence with linguistics, *American Journal of Human Genetics,* 61: 1015- 1035 (cf. the erratum published in 1998 in the *American Journal of Human Genetics,* 62: 1267).

Poloni, Estella Simone, Nicolas Ray, Stefan Schneider and André Langaney. 2000. Languages and genes: Modes of transmission observed through the analysis of male-specific and female-specific genes, 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.

Purnell, Herbert C., Jr. 1970. *Toward a Reconstruction of Proto-Miao-Yao.*Ithaca, New York: Cornell University Ph. D. dissertation.

Ramstedt, Gustaf John. 1957 [posthumous]. *Einführung in die altaische Sprachwissenschaft:Lautlehre* (Suomalais-Ugrilaisen Seuran Toimituksia, Mémoires de la Société Finno-Ougrienne,104, 1). Helsinki: Suomalais-Ugrilaisen Seura.

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

Reddy, B. Mohan, B.T. Langstieh, Vikrant Kumar, T. Nagaraja, A.N.S. Reddy, M. Aruna, K. Thangaraj, A.G. Reddy and Lalji Singh. 2007. Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia, *Public Library of Science One*, 2 (11): e1141. doi:10.1371/journal.pone.0001141.

Reid, Lawrence A. 1994. Morphological evidence for Austric, *Oceanic Linguistics,* 33 (2): 323-344.

Reid, Lawrence 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.

Relandus, Hadrianus [i.e Adriaan van Reeland]. 1706, 1707, 1708. *Dissertationum Miscellanearum, Pars Prima, Pars Altera, Pars Tertia et Ultima*. Trajecti ad Rhenum (i.e. Utrecht): Gulielmus Broedelet (esp. *Pars Tertia et Ultima*, Chapter XI Dissertatio de linguis insularum quarundam orientalium).

Rootsi, Siiri, Lev A. Zhivotovsky, Marian Baldovič, Manfred Kayser, Ildus A. Kutuev,Rita Khusainova, Marina A. Bermisheva, Marina Gubina, Sardana A. Federova, Anne-Mai Ilumäe, Elza K. Khusnutdinova, Mikhail I. Voevoda, Ludmila P. Osipova, Mark Stoneking, Alice A. Lin, Vladimir Ferak, Jüri Parik, Toomas Kivisild, Peter A. Underhill and Richard Villems. 2007. A counter-clockwise northern route of the Y-chromosome haplogroup N from Southeast Asia towards Europe, *European Journal of Human Genetics,* 15: 204-211.

Sahoo, Sanghamitra, Anamika Singh, G. Himabindu, Jheeman Banerjee, T. Sitalaximi, Sonali Gaikwad, R. Trivedi, Phillip Endicott, Toomas Kivisild, Mait Metspalu, Richard Villems and V. K. Kashyap. 2006. A prehistory of Indian Y chromosomes: Evaluating demic diffusion scenarios, *Proceedings of the National Academy of Sciences of the United States of America*, 103 (4): 843–848.

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

Schmidt, Wilhelm. 1904. *Grundzüge einer Lautlehre der Khasi-Sprache in ihren Beziehungen zu derjenigen der Mon-Khmer-Sprachen, mit einem Anhang: Die Palaung-, Wa- und Riang-Sprachen der mittleren Salwin*. Munchen: Kaiserliche Akademie.

Schmidt, Wilhelm. 1906. Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens, *Archiv fur Anthropologie,* Neue Folge, v: 59-109.

Schorer, Nicolas Michael. 2016. *The Dura Language: Grammar and Phylogeny.* Leiden: Brill.

von Siebold, Philipp Franz Balthazar. 1832a. Verhandeling over de afkomst der Japanners, *Verhandelingen van het Bataviaasch Genootschap van Kunsten en Wetenschappen*, 13: 183-275.

von Siebold, Philipp Franz Balthazar. 1832b. *Nippon: Archiv zur Beschreibung von Japan* (four volumes). Leyden:ausgegeben unter dem Schutze Seiner Majestät des Königs der Niederlande bei dem Verfasser.

Starosta, Stanley. 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 [Editorial misrepresentations that had crept into Starosta's posthumously published East Asian phylogeny were rectified that same year (cf. van Driem 2005).].

Starostin, Sergej Anatol'evič. 1994. The reconstruction of Proto-Kiranti, paper presented at the 27ème Congrès International sur les Langues et la Linguistique Sino-Tibètaines, Centre International d'Études Pèdagogiques, Sèvres, 14 octobre 1994.

von Strahlenberg, Phillip Johann. 1730. *Das Nord- und Östliche Theil von Europea und Asia, in so weit das gantze Russische Reich mit Siberien und grossen Tatarey in sich begreiffet, in einer Historisch-Geographischen Beschreibung der alten und neueren Zeiten, und vielen andern unbekannten Nachrichten vorgestellet, nebst einer noch niemahls and Licht gegebenen* Tabula Polyglotta *von zwei und dreyβiglei Arten Tatarischer Völcker Sprachen und einem Kalmuckischen* Vocabulario. Stockholm: In Verlegung des *Autoris*.

Sweeney, Megan T., and Susan R. McCouch. 2007. The complex history of the domestication of rice, *Annals of Botany*, 100: 951-957.

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

Wáng Fǔshì (王辅世). 1994. *Miáo Yǔ Gǔyīn Gǔunǐ* ["Reconstruction of the sound system of proto-Miao"]. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Wáng Fǔshì(王辅世). and Máo Zongwǔ(毛宗武). 1995. *Miáoyáo Yǔ Gǔyīn Gòunǐ* ["Reconstruction of the sound system of proto-Miao-Yao"]. Peking: Zhōngguó Shéhuì Kēxué Chūbǎnshè.

Wáng, William Shì-Yuán (王士元). 1998. Three windows on the past, pp. 508-534, Vol. 1 in Victor Henry Mair, ed., *The Bronze Age and Early Iron Age Peoples of Eastern Central Asia*. Washington and Philadelphia: Institute for the Study of Man and the University of Pennsylvania Museum.

Wen Bo, Li Hui, Lu Daru, Song Xiufeng, Zhang Feng, He Yungang, Li Feng, Gao Yang, Mao Xianyun, Zhang Liang, Qian Ji, Tan Jingze, Jin Jianzhong, Huang

Wei, Ranjan Deka, Sù Bīng, Ranajit Chakroborty and Jīn Lì. 2004. Genetic evidence supports demic diffusion of Han culture, *Nature,* 431: 302-305.

Witsen, Nicolaes. 1692. *Noord en Oost Tartarye, ofte Bondigh Ontwerp van Eenige dier Landen, en Volken, zo als voormaels bekent zyn geweeft, Beneffens Verncheide tot noch toe onbekende, en meent noit voorheen befchreve Tarterfche en nabeurige gewe∫ten, lant∫treken, ∫teden, rivieren, en plaetzen, in de Noorder en Oosterlykste Gedeelten van Asia en Europa* (2 vols.). Amsterdam: François Halma.

Wulff, Kurt. 1934. *Chinesisch und Tai: Sprachvergleichende Untersuchungen* (Det KongeligeDanske Videnskabernes Selskab, Historisk-filologiske Meddelelser, xx, 3). Copenhagen : Levin & Munksgaard.

Wulff, Kurt. 1942 [posthumous]. *Über das Verhältnis des Malay-Polynesischen zum Indochinesischen*.Copenhagen: Munksgaard.

Yan, Shi, and Chuan-Chao Wang, Hui Li, Shi-Lin Li, Li Jin and the Genographic Consortium. 2011. An updated tree of Y-chromosome haplogroup O and revised phylogenetic positions of mutations P164 and PK4,*European Journal of Human Genetics*, 19(9):1013-1015.

Yao, Xiaotian and Senwei Tang, Beilei Bian, Xiaoli Wu, Gang Chen and Chuan-Chao Wang. 2017. Improved phylogenetic resolution for Y-chromosome haplogroup O2a1c-002611, *Scientific Reports*, 7: 1146.

Zide, Norman H., and Arlene R.K.Zide. 1976. 'A Proto-Munda cultural vocabulary: Evidence for early agriculture', pp. 295-334, Vol. II, Part II in Philip N. Jenner, Laurence C. Thompson and Stanley Starosta, eds.,*Austroasiatic Studies* (Oceanic Linguistics Special Publication13). Honolulu: University of Hawaii.