

HEALTH IN THE HIMALAYAS AND THE HIMALAYAN HOMELANDS

GEORGE VAN DRIEM
UNIVERSITY OF BERNE

ADAPTATIONS TO COLD AND HIGH ALTITUDE

In September 2012, the eminent Estonian population geneticist Toomas Kivisild convened and hosted a gathering at the University of Cambridge entitled *High Altitude and Cold: Adaptation to Extremes*. This joyous event was billed as ‘a conference to discuss the latest scientific results on human physiological and genetic adaptation to high altitudes and cold environments’. Memorable speakers included Hugh Montgomery, Andrew J. Murray, Miroslava Derenko, Cynthia Beal, Maanasa Raghavan, Dennis O’Rourke, Gianpiero Cavalleri, Anna di Rienzo and Bill Amos. In addition to population geneticists and physiologists, Toomas Kivisild invited to Cambridge a single linguist, who happened to find genetics symposia more enlightening and often more full of ebullience than linguistic gatherings.

Just how damaging life at high altitudes can be for people who are genetically ill-equipped to cope with cold high-altitude environments was demonstrated by Andrew J. Murray of Cambridge, who reported on an expedition called Extreme Everest. This scientific expedition aimed to study the metabolic response of heart and muscle tissue to high-altitude hypoxia. It appears that a failing heart is somehow unable to oxidise fatty acids whilst in a state of dilated cardiomyopathy. This inability has to do with hypoxia inducible factors (HIF) and metabolic regulation, with hypoxia inducible factor one-alpha (HIF-1 α) accumulating in hypoxic tissues. Fortunately, in healthy individuals, heart impairment caused by high altitude can repair itself after, say, about six months. Most interesting is how this impairment plays out at the biochemical level within the cell.

A loss of 20% to 25% of mitochondria was observed in muscle and skeletal tissue after an Everest summit. This reactive stress appears to come from reactive oxygen species, popularly known as free radicals,

and not as a result of the oxygen partial pressure (pO_2) in the tissues themselves. In this context, Dr. Murray exhorted us to protect our mitochondria. Biopsies of muscle tissue at sea level and at high altitude had shown a fall in mitochondrial density. This drop was not dramatic in the early phase, but, as one ascends higher, the levels of the mitochondrial uncoupling protein UCP3 increase dramatically, evidently in an attempt to protect the mitochondria by decreased uncoupling. Very dramatic biopsy changes were observed in subjects after summitting Everest, mainly on the cell membrane. Since it is still not understood why one population of mitochondria is depleted more than another, Dr. Murray assured us that a second Extreme Everest expedition was of the essence. In view of the fascinating findings, one could not begrudge these scientists the funding for a second Himalayan expedition.

Meanwhile, back in England, another group of scientists were indulging in their love for controlled experiments. One physiologist at the conference presented the results of an experiment in which he had put Tibetans and Englishmen into a large decompression chamber and set them to work on a treadmill. The subjects were hooked up to all sorts of medical measuring and monitoring equipment and taken up to artificially induced high altitudes. An interesting characteristic of the Tibetans in the experiment was that they had all been born and bred in England and had at no point in their lives been exposed to high-altitude environments. In fact, they were Britons of Tibetan extraction. In the experiment, therefore, their bodies had to rely on their inborn capacity to cope with alpine conditions.

Soon most test subjects of impeccable Brittanian pedigree were seen to be huffing and puffing and exhibiting physiological responses that, over any length of time, would have been incompatible with life. To the dismay of the experimenters, these test subjects had to be removed from the decompression chamber before their session could run its full course. By contrast, the British subjects of pure Tibetan extraction invariably experienced little sign of discomfort and even exhibited a demeanour of nonchalance throughout. One might even say that, at the histological level, the bodies of the subjects of Tibetan ancestry exuded an air of nonchalance, for their innate responses to high-altitude conditions are of such biochemical sophistication that their physiological indicators were observed to remain in the cool, calm and collected mode throughout the experiment.

This experiment vividly illustrated the inbuilt nature of the Tibetan capacity to cope with what Johann Grueber in 1661 described as ‘viato-

res vix respirare ob aëris subtilitatem queant', that is, travellers, but not the Tibetans, could hardly breathe in the thin air. The Austrian Jesuit recorded this observation after having traversed the Everest massif with his Walloon confrère Albert d'Orville on their way from Peking to Agra. Grueber also reported that Mount Everest was known locally to the Tibetans as ལྷང་གུར་ 'Langgur 'bull's tent' and claimed by them to be the 'mons omnium altissimus', the highest of all mountains.

It is obvious that the local Tibetan name 'Langgur was the source for the bookish Tibetan name ཇོ་མོ་མི་གཡོ་སྐྱང་བཟང་མ་ J'omomi'yo'langzangma 'Immovable Hardy Mistress Taurus', rather than the other way around. The literary Tibetan name is familiar to many Western travellers in its shortened form ཇོ་མོ་སྐྱང་མ་ J'omo'langma, which often ends up transmogrified in various guises on tourist posters and in popular travel guides. It stands to reason that the very people who had lived for so long in the high Himalayas would already have known which mountain was the world's highest long before Rādhānāth Sikdhar and Michael Hennessy of the Great Trigonometrical Survey of India first measured it to be so in 1852.

Yet just how long have the Tibetans been living in the Himalayas as anyone's guess until population geneticists and palaeontologists began to find answers to this age-old question. Genetic evidence indicates that the ancestors of the Tibetans have lived on the high plateau and in the Himalayan highlands for well over thirty millennia, whereas the Andeans have resided at high altitude for only just a dozen millennia. Consequently, Tibetan highlanders differ in their biological characteristics from the Nuñoan Peruvians. Under similarly hypoxic conditions, Tibetans have lower haemoglobin (Hb), more nitrogen oxides (NO) in the blood, and less oxygen partial pressure (pO_2) in their tissues. In their quest for the evolutionary basis of human adaptation to cold and high altitude, population geneticists have identified the now famous genes EPAS1 and EGLN1, and they optimistically announce that they will soon identify other loci.¹

Whereas both Andeans and Tibetans exhibit the high altitude adaptation gene EGLN1, the gene EPAS1 turns out to be somewhat of a Himalayan specialty and very much a product of natural selection. By contrast, Andeans appear to have relied heavily on developmental plasticity to realise short-term responses to living at high altitude, capitalising upon genetic propensities which we all share. For example, highlanders in the Andes are shorter in height than their closest relatives

at lower altitudes. Like Europeans, Hà Chinese and most other people, Andeans show a haemoglobin response which varies for displacement to high altitude, whereas Tibetans notably do not exhibit this response. At high altitudes, Tibetans have less haemoglobin.

Whilst Andeans respond by producing more haemoglobin to capture more oxygen from the air, the Tibetan EPAS1 gene codes for haemoglobin production in a way that represents the finely honed outcome of long natural selection. The Tibetan genome exhibits high-frequency single nucleotide polymorphisms (SNPs) near the EPAS1 gene associated with lower haemoglobin concentrations. Lower haemoglobin levels at high altitudes prevent free radical accumulation. Instead, Tibetans have enhanced nitrogen oxide levels in the blood, allowing more blood to circulate. Tibetans also breathe more deeply and more frequently. A long gradual evolution under the selective pressures of the Tibetan plateau and the Himalayan highlands is responsible for the physiological sophistication of the Tibetan genetic adaptations to cold and high altitude. The genetic architecture of the Tibetan adaptation to climate and altitude is unique.

A low haemoglobin response is not unique, however. In Ethiopia, the Amhara, for example, also exhibit a low haemoglobin response phenotype at high altitude, whereas the neighbouring Oromo have a very pronounced high altitude response. Yet the Amharic phenotype has a genetically distinct molecular basis, and the biochemical pathways which Tibetans exploit are not the same as those found in the Amhara. Although it has been understood for quite some years now that hypoxia inducible nuclear factors (HIF) bind to an enhancer element located 3' to the human erythropoietin gene, a great many disparate factors are known to affect the abundance and transcription of hypoxia inducible factors. Oxygen homeostasis pathways can be activated without hypoxia, for example, by exposure to extreme cold, although stress caused by temperature or poor nutrition remains inherently hard to quantify.

The Tibetan adaptation is of particular interest because this gene and the associated molecular markers reflect the antiquity of human habitation in the Himalayan region. More generally, it is observed that numerous alleles under strong selective pressure show dramatic clines the further one gets from Africa. At the same time, neutral variability declines with distance out of Africa, as reflected in the heterozygosity of alleles and cranial measurements, except where this appears to be under selective pressure. Natural selection has evidently favoured in-

creased variability in genes coding for resistance to pathogens which are under directed selection, and such markers increase in diversity as one moves out of Africa.

Most recently, on the basis of the highly specific structure of the Tibetan haplotype and the fact that outside of Tibet the EPAS1 gene is exclusively found in Denisovan DNA, a team of geneticists have established that the celebrated hypoxia pathway gene is the result of introgression of a Denisovan gene into the ancestors of the Tibetans (Huerta-Sánchez *et al.* 2014). The Denisovans were a Palaeolithic people who lived in the Denisova cave in the Altai mountains of Siberia. Like the Neanderthals, this extinct variety of human is not really entirely extinct because these people met and interbred with the ancestors of many existing populations. A small percentage of DNA is shared between Denisovans and Australians and between Denisovans and a number of Asian populations. Today the EPAS1 gene is still exclusively shared between Denisovans and Tibetans. Interbreeding with Denisovans rendered the ancestors of the Tibetans specifically well-equipped to life in the Himalayan highlands and on the Tibetan plateau. In our understanding we must be ever wary of anachronisms in our thinking. When an ancestral highland Asian population interbred with the Denisovans, these people did not yet speak a language related to Tibetan, and ethnolinguistically they were, of course, not yet Tibetan in any modern sense of the word.

Not only can our susceptibility to altitude sickness or our inborn inurement to thin air and cold be used to garner information about our past. Microbial pathogens too can tell us how and where our ancestors moved about. Half of the people in the world today are infected with *Helicobacter pylori*, a pathogen which causes chronic gastritis and peptic ulcers and appears to be implicated in the development of duodenal and stomach cancers. Although first discovered only in 1982, this gram-negative proteobacterium has lived in human stomachs at least since the time that our ancestors emerged from Africa. Some *Helicobacter* enthusiasts regard this microbe as a benign commensal, and some have even hinted that its ubiquity and long coexistence with humans suggest that man's relationship with the bacterium might even be symbiotic. Yet any beneficial effects of this bug have yet to be demonstrated.

The phylogeny of the various strains of *Helicobacter pylori* reflect the peopling of the planet by anatomically modern humans such as ourselves. Compelling evidence has been adduced that human stomachs

were colonised by *Helicobacter pylori* long before we left Africa, and the oldest strains are found in the San populations of southern Africa. (Wirth *et al.* 2004, Linz *et al.* 2007). In the phylogeny of this pathogen, the first clade out of Africa to diverge is the clade *hp Sahul*, which traces the early littoral dispersal of humans out of Africa as far as Australia. The next most basal branch in the out-of-Africa tree is *hp Asia2*, which exists in two strains *hsp Ladakh* and *hsp India*, whereby *hsp Ladakh* is the older of the two. After the littoral dispersal of the first wave, the stomach pathogen traces the second major wave in the peopling of Asia as first disseminating through the welcoming mid hill habitats of the Himalayan corridor. Yet in terms of sampling, the Himalayan region remains one of the most understudied geographical areas in *Helicobacter* research (Moodley 2014).

Not all diseases have been with us as long as *Helicobacter pylori*. Leprosy is caused by the microbe *Mycobacterium leprae*, and initially it was thought that the different strains of this pathogen too would reflect the dispersals of anatomically modern humans out of Africa (Monot *et al.* 2005). However, in order to thrive and spread in a fecund manner, leprosy requires a host population that has attained some critical density, for leprosy is the least contagious of our communicable diseases. There is no palaeopathological evidence for leprosy, and this mycobacterium is unlikely to have thrived in our Pleistocene ancestors, who lived in small mobile bands of hunters and foragers. Instead leprosy appears to represent one of the first maladies associated with urban living. Its preference for dense sedentary populations of the human host population makes the phylogeny of the various strains of leprosy a good tracer for patterns of dispersal in the Holocene (Pinhasi *et al.* 2006).

Yet we need not resort to proxy studies such as pathogens to trace Holocene migrations. Both historical linguistics and human population genetics provide us with ample information about these more recent episodes of our past. We shall now define the Eastern Himalaya as a geographical area, and then investigate what human population genetics and linguistics tell us about our past.

THE EASTERN HIMALAYA AS A GEOGRAPHICAL AREA

Once anatomically man emerged from Africa, simple facts of geography dictated where he might go, at least to some extent. Molecular genetic findings tell us that our ancestors emerged in waves (Rasmussen *et al.* 2011). Between 75,000 and 62,000 years ago, the First Wave spilled out of Africa and followed a littoral route eastward across the Indian Subcontinent and Southeast Asia. The littoral dispersal should not be envisaged as bands of beachcombers, but as migrations through the vast and broad swathes of territory which encircle and abut on the southern coastline of the Asian continent.

There is genomic evidence for some secondary gene flow between New Guinea and Australia some 8,000 years ago at the time of the Sahul land bridge, but the actual colonisation of Australia took place some 50,000 years ago and involved a major sea crossing from Asia. How the Australians reached the great continent down under remains a mystery. Were the initial groups of settlers marooned after having been haplessly swept out to sea, or were the first Australians intrepid mariners?

What is clear is that the littoral route of the First Wave effectively culminated in a major seagoing adventure which defied simple facts of geography. What is equally evident is that the original Australians quickly abandoned their seagoing ways. The deep branching of maternal and paternal lineages of Australian populations vis-à-vis other populations around the Indian Ocean shows long-term isolation after initial settlement. The linguistic picture suggests that the maritime migration of the first Australians made landfall in the northwest of Australia. Seven-eighths of the continent is covered by Pama-Nyungan languages, whilst the northwestern region shows a diversity of language families that marks the area as the probable hearth for the peopling of the Australian continent. Meanwhile, strewn across the southern Asian littoral, ancient populations whose ancestors had emerged from Africa in the First Wave continued to flourish.

Then between 38,000 to 25,000 years ago, the Second Wave spilled out of Africa through the Levant. A small group branched off into Asia Minor and ventured further west into Europe, but the main body of the wave swept across South Asia and headed to East Asia. Other than those headed for Europe, these migrating peoples interbred with populations of the First Wave whom they met on their way. It is an inescapable consequence of geography that when our anatomically modern ancestors

emerged from Africa and took this inland route on their way to East Asia, Southeast Asia, Oceania, Siberia, the Americas and even Lapp-land, many of these ancestors must at one point have passed through the Eastern Himalaya and crossed the mighty Brahmaputra.

Geographically the Eastern Himalaya can be demarcated as a well defined area which served as a principal thoroughfare in the course of population prehistory. The Himalayas run over 3600 kilometres from the Hazārahjāt Highlands in the west to the Liángshān in the east. In the west, the Himalayas are punctuated by the Tirich Mir in the Hindu Kush at 7708m and by the K2 in the Karakoram at 8661m. In the east the Himalayas are punctuated by the ཀངས་དཀར་པོ་ Khàng Karpo in eastern Tibet at 6740m and the Hkakabo Razi in northern Burma at 5881m.

The Himalayas form no natural watershed, and many of the rivers are of greater antiquity than the mountains themselves. The Kālī Gaṇḍakī, which flows right through the Himalayan range and just beneath the Dhaulāgiri (8167m), has carved out the deepest river valley on the face of the planet. This invagination at the very heart of the Himalayan range, prominently visible to any airplane passenger flying across the Gangetic plain, bisects the Himalayas into two halves of roughly equal length.

The eastern half can be denominated as the Eastern Himalaya, beginning from the Dhaulāgiri and extending eastward across the Himalayas and sub-Himalayas, the Meghalaya, the lower Brahmaputran basin and its associated hills tracts, the eastern Tibetan Plateau and the Indo-Burmese borderlands into the Chinese provinces of Yúnnán and Sìchuān. The Eastern Himalaya constituted an area of pivotal importance to population prehistory and served as a cradle of ethnogenesis at different time depths in the past. New insights from linguistics, genetics and archaeobotany enable us tentatively to reconstruct the founding dispersals of a number of major language families in Asia and Oceania.



Figure 1: Geographical distribution of Trans-Himalayan languages



Figure 2: Geographical distribution of the major Trans-Himalayan subgroups. Each dot represents not just one language, but the putative historical geographical centre of each of 42 major linguistic subgroups.



Figure 3: Thirty out of forty-two Trans-Himalayan subgroups lie south of the Himalayan divide, seven to the north and east, and five (Tshangla, Bodish, Nungish, Lolo-Burmese and Kachinic) straddle both flanks of the Himalayas



Figure 4: Geographical distribution of Hmong-Mien

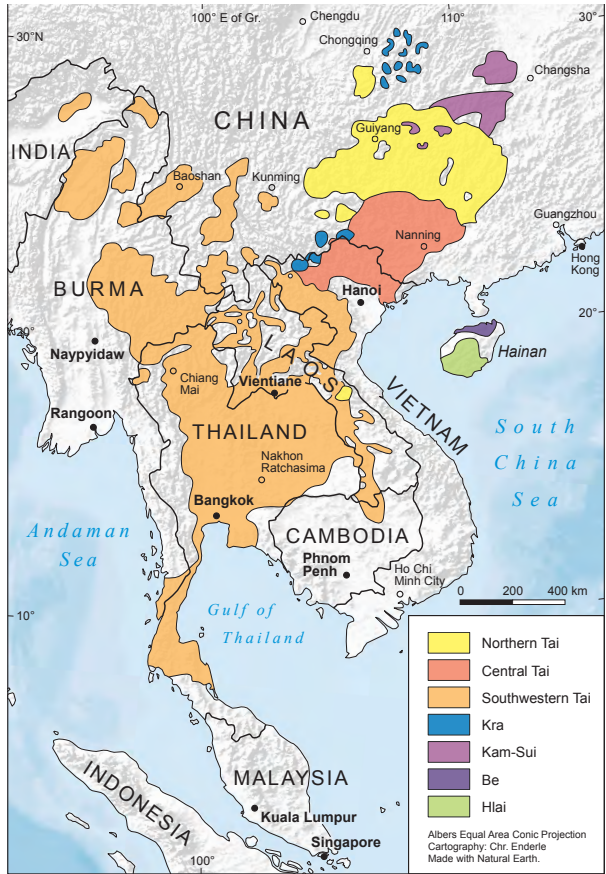


Figure 5: Geographical distribution of Kradaï

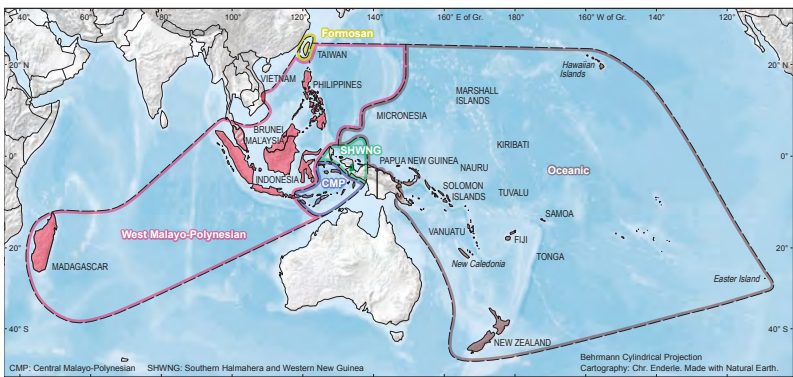


Figure 6: Geographical distribution of Austronesian



Figure 7: Geographical distribution of Austroasiatic

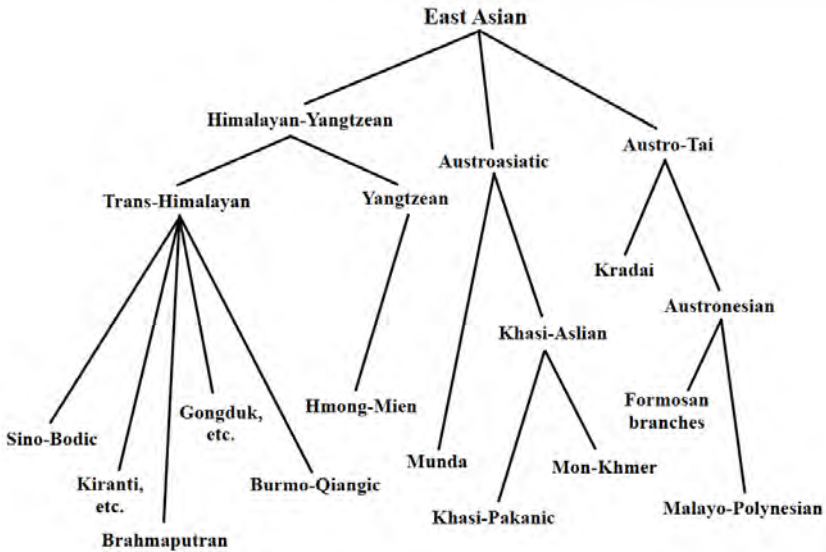


Figure 8: The 2012 Benares Recension: Revised East Asian phylogeny.

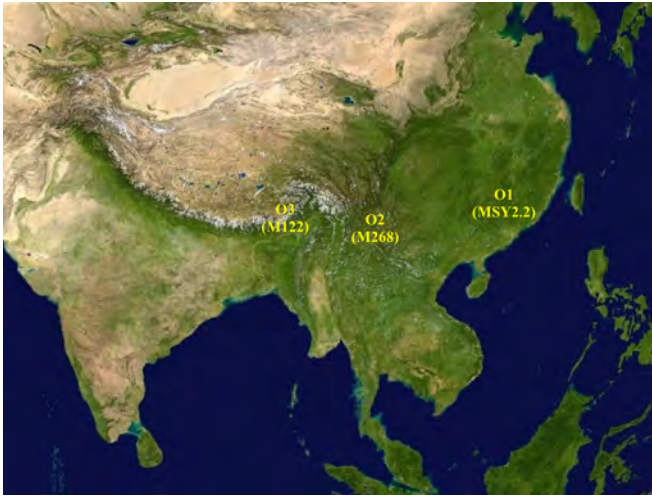


Figure 9: After the Last Glacial Maximum, the Y chromosomal haplogroup O (M175) split into the subclades O1 (M119), O2 (M268) and O3 (M122). Bearers of the O2 (M268) paternal lineage domesticated Asian rice.

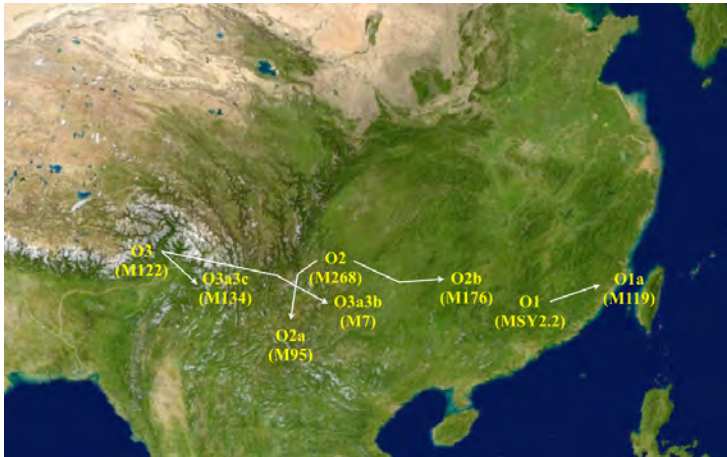


Figure 10: Paternal lineages branching into new subclades. Each event involved a linguistic bottleneck leading to language families that today are reconstructible as distinct linguistic phyla. The O1 (MSY2.2) lineage gave rise to the O1a (M119) subclade, which moved eastward to the Fújiàn hill tracts and across the strait to Formosa, which so became the *Urheimat* of the Austronesians. Bearers of O3a3b (M7) became the Proto-Hmong-Mien. In the Eastern Himalaya, the bearers of haplogroup O3a3c (M134) expanded and became the Trans-Himalayans. Haplogroup O2a (M95) is the Proto-Austroasiatic paternal lineage. The para-Austroasiatic fraternal clade O2b (M176) spread eastward, sowing seed along the way.

LANGUAGE AND GENES

There is a long lineage of scholars from Julius von Klaproth and Friedrich Max Müller who, since the early 19th century, have stressed that language and biological ancestry are two different things. There have been others too, like Sir William Jones, who from time immemorial have confounded language and race. Generally people throughout history have been inclined to speak the language spoken by their parents, but the language which we happen to speak today may very well not be the language of our parents. Since genes are invariably inherited by offspring from their biological parents, a probabilistic correlation may therefore exist between language and genes in human populations, though this need not necessarily be so.

The past encompassed vast stretches of time. There are many slices of the past, and a chronologically layered view of ethnolinguistic prehistory is therefore essential. When Neanderthals and anatomically modern humans coexisted in the Levant between 80,000 and 50,000 years ago, interaction between the two varieties of hominids entailed a modest Neanderthal paternal contribution of between 1% to 4% to the ancestry of mankind outside of Africa (Serre *et al.* 2004, Sankararam *et al.* 2012). When our ancestors subsequently set off in different directions to Europe and deeper into Asia and beyond, the people who moved east had a slightly higher proportion of Neanderthal ancestry than those who followed the Neanderthals into Europe and possibly contributed to their ultimate extinction there.

In terms of distinct chronological layers of prehistory, the telltale trace of Neanderthal DNA in Eastern African populations such as the Maasai represents a vestige of much later prehistorical contacts between peoples of the East African littoral and Asian populations across the Indian Ocean (Walls *et al.* 2013). Similarly, the linguistically reconstructible past likewise relates to more recent slices of prehistory. Not only is the time depth accessible to historical linguistics shallower than the time depth accessible to human genetics, but the spread of language families also happens to be a more recent phenomenon than the spread of our anatomically modern ancestors outside of Africa. Language families, linguistic phyla and the recognition of language isolates represent the maximal time depth accessible to historical linguistics because the relatedness of languages belonging to a recognised language family or linguistic phylum represents the limit of what linguists can empirical-

ly demonstrate. This epistemological barrier represents the linguistic event horizon.

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

Moreover, we must also take into account the potential skewing effects of natural selection, gene surfing, recurrent bottlenecks during range expansion and the sexually asymmetrical introgression of resident genes into incursive populations. Factors such as ancient population structure and possible ancient Y chromosomal introgression could also affect inferences and interpretations based on any single Y chromosomal locus when attempting to reconstruct migrations and elucidate the geographical origins of populations (Mendez *et al.* 2013, van Driem 2012b). Even with all these caveats in place, we must be especially aware of all provisos and qualifications included in our inferences and working hypotheses when attempting to understand East Asian ethnolinguistic phylogeography.

In attempting to understand such correlations, it will not suffice to keep just these caveats in mind, we should also beware of the insidious influence of obsolete ways of thinking. Myths are astonishingly robust. Myths continue to survive both in the popular imagination as well as in scientific discourse long after their obsolescence has been convincingly demonstrated. Categories such as the Mongoloid race or the Caucasoid race have no basis in biological fact, but they persist today. These racial categories originated in the attempt of Christoph Meiners (1813) to come to terms with phenotypical diversity of our species by setting up a classification of races based on what he imagined where the *uralte Stammvölker* or racial prototypes of mankind.

Similarly, in historical linguistics, the Sino-Tibetan family tree has long been outed as a false theory (van Driem 2014a). Yet this empirically unsupported model has continued to mislead a number of scholars even in recent years, and a subset of linguists who designate themselves as Sino-Tibetanists continue to adhere to this false phylogeny. Dispelling myths can be an arduous task because of the tenacity with which such narratives can take hold of the human mind. Modern adherents of Sino-Tibetan cannot be held accountable for the racist legacy of model, which is rooted in ‘scientific’ racism. Indeed, most linguists lack detailed knowledge of the history of their own field. Yet at one point students of historical linguistics come to shoulder responsibility for the models which they inherit from their mentors. Historical linguists must evaluate family trees and subgrouping hypotheses and be compelled by evidence and the lack thereof.

FATHER TONGUES

When studying the distribution of maternally inherited markers in the mitochondrial DNA and paternally inherited markers on the Y chromosome, population geneticists soon found that it was easier to find statistically relevant correlations between the language of a particular community and the paternally inherited markers prevalent in that community than between the language and the most salient maternally inherited markers found in that speech community. This Father Tongue correlation had already been described by a Swiss-Italian team (Poloni *et al.* 1997, 2000) well before the appearance of the seminal articles on Y chromosomal phylogeography by Underhill *et al.* (2000, 2001). Subsequent work has further refined the resolution of the Y chromosomal haplogroup tree, e.g. Karafet *et al.* (2008).

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 nucleo-

tide polymorphisms in whole genome studies are making headway (Li *et al.* 2008), but it is still too early to tell to what extent correlations of autosomal markers with language phyla will be identified that are as salient as the currently observed Father Tongue correlations.

There are a number of reasons why we might expect this outcome. Initial human colonisation of any part of the planet must have involved both sexes in order for a population of progeny to establish itself. Once a population is in place, however, subsequent migrations could have been heavily gender-biased. Subsequently, male intruders could impose their language whilst availing themselves of the womenfolk already in place. By contrast, correlations between maternal lineages and linguistic phylogeography discerned to date have been underwhelming. The Father Tongue hypothesis suggests that linguistic dispersals were, at least in most parts of the world, posterior to initial human colonisation and that many linguistic dispersals were predominantly later male-biased intrusions. Such patterns are observed worldwide.

The correlation of Niger-Congo languages with Y chromosomal haplogroups is a striking example (Wood *et al.* 2005). Likewise, the martial and male-biased historical spread of Hà 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.* 2006b).

A major Y-chromosomal lineage in Mesolithic Europe, and perhaps even the predominant paternal lineage in Palaeolithic Europe, was haplogroup I. The Basque community, whose language harks back to a pre-Indo-European epoch, are distinct from their immediate neighbours in Iberia and France in tracing as much as a third of their ancestry back to ancient Western European hunter-gatherer populations, as represented by the Mesolithic male skeleton recovered from the Loschbour rock shelter in Heffingen in Luxembourg. By contrast, local maxima of haplogroup I in Scandinavia and the western Balkans might reflect more recent expansions of this lineage (Lazaridis *et al.* 2014).

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 (van Driem 2007, 2012b), but the unfolding story of Y chromosomal R lineages will no doubt turn out to be complex. In an epoch anterior to the expansion of Indo-European from the Pontic Caspian, an older pre-Indo-European or even pre-Nostratic homeland could have lain in the Western Himalaya, as suggested by the presence of the ancestral clade R* in Indian populations (Underhill *et al.* 2010, 2014). This hypothesis exists in two versions. In this context, the presence of F* and K* in Indian populations represents additional molecular evidence for the pivotal position of South and Southeast Asia in population prehistory after the emergence of our ancestors from Africa (van Driem 2014b, Karafet *et al.* 2014). In this context, a fine-mesh study of populations inhabiting the Western Himalaya remains a desideratum.

Similarly, it has been proposed that the Y chromosomal lineage L, which shows a great diversity of subclades on the Iranian highland, can be identified as the possible marker of a patrilingual dispersal of Elamo-Dravidian emanating from a region which included the Bactria and Margiana of later prehistory (van Driem 2012b), and that one of these Y chromosomal L subclades will appear to be correlated with the patrilingual spread of Dravidian languages from the Indus Valley into southern India (van Driem 2014b).

I have also proposed that haplogroup Q, an offspring clade of Y chromosomal haplogroup P, could at a certain time depth represent a marker for the Greater Yenisseian linguistic phylum (van Driem 2008, 2014b). Obviously this is not to equate a marker such as Y-chromosomal haplogroup Q with some Greater Yenisseian 'ethnolinguistic stock'. Such simplifications would merely impede a detailed understanding of the phylogeny and chronology of the many Q subclades as well as the likely complexity of the spread of the Greater Yenisseian linguistic phylum. Populations forming local exceptions to the Father Tongue correlation, such as the Hungarians and the Balti, have been discussed elsewhere (van Driem 2012b, 2014b). At a certain time depth a Y chromosomal haplogroup marker may be a tracer for a paternally mediated linguistic spread.

In the following sections, the paternal ancestry of language communities speaking tongues of the East Asian linguistic phylum will be interpreted to sketch a plausible scenario for the founding dispersals of the constituent language families of the East Asian

phylum. In so doing, the paternal haplogroup O2a (M95) will be identified as a tracer for the predominantly paternal spread of Austroasiatic, haplogroup O3a3c (M134) with Trans-Himalayan, haplogroup O3a3b (M7) lineage with Hmong-Mien and O1 (MSY2.2) with Austro-Tai (van Driem 2007, 2012b, 2014b).² Against the background of the East Asian linguistic theory, linguistic ancestry will be seen to correlate well with paternal ancestry. Y chromosomal phylogeography captures formative episodes in the ethnolinguistic prehistory of Asia and the world. In the near future, whole-genome sequencing and genome-wide typing and selection scans may lead to an overhaul of Y chromosomal haplogroup nomenclature. Yet at the time depth reflected by Y chromosomal phylogeography, the interface of molecular genetics and historical linguistics tells a tale of the peopling of the world from a Holocene ethnolinguistic perspective.

THE TRANS-HIMALAYAN LANGUAGE FAMILY AND THE EAST ASIAN LINGUISTIC PHYLUM

The second most populous language family on the planet is Trans-Himalayan, also known as Tibeto-Burman. Most speakers of Trans-Himalayan languages today live to the north and east of the Himalayas (Figure 1), but most of the over 300 different languages and three fourths of the major Trans-Himalayan subgroups are located to the south of the Himalayan divide (Figure 2). The Trans-Himalayan language family was first recognised by Julius von Klaproth in 1823, who identified the family as consisting of Tibetan, Chinese, Burmese and related languages. This language family was called Tibeto-Burman by scholars in the British Isles, e.g. Hodgson (1857), Cust (1878), Forbes (1878), Houghton (1896).³

Following in the footsteps of scholars such as Witsen (1692) and Reilandus (1706, 1707, 1708), Julius von Klaproth challenged conventional wisdom in 1823 by proposing a polyphyletic view of Asian language families, and he distinguished the contours of many of the known Asian language families. The five major language families recognised today as forming part of the East Asian linguistic phylum are Trans-Himalayan, Hmong-Mien, Kra-dai, Austronesian and Austroasiatic (Figures 1, 4, 5, 6, 7).

Once Klaproth's polyphetic view had been in place for nearly a century, scholars began to discern possible long-distance relationships between the recognised language families. We might say that for linguistic taxonomy Klaproth's centrifugal step was gradually followed by a series of centripetal steps. Gustave Schlegel (1901, 1902) agreed with Klaproth in assessing 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, 1975, 1976, 1990) under the guise of 'Austro-Thai', though this putative genetic link constituted just an ingredient in his grand and poorly supported 'Japanese/Austro-Tai'.

Weera 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 language family 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.

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

In the first half of the 20th century, Conrady (1916, 1922) and Wulff (1934, 1942) each proposed a superfamily consisting of Austroasiatic, Austronesian, Kradai and Tibeto-Burman. Subsequently, Blust (1996) and Peiros (1998) proposed an Austric superfamily comprising Austroasiatic, Austronesian, Kradai and possibly Hmong-Mien. Then in

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

Starosta's posthumously published East Asian phylogeny was marred by editorial errors (Starosta 2005: 183), which were later corrected (van Driem 2005a: 322). A theory of linguistic relationship at this time depth lies at the frontier of what can be empirically demonstrated to the satisfaction of a methodologically rigorous historical linguist. This hypothesis will therefore remain an informed conjecture until solid historical linguistic evidence either further supports or overturns the model. At Benares in 2012, I presented the tweaked East Asian family tree depicted in Figure 8 (van Driem 2014b).

EAST ASIAN AND THE EASTERN HIMALAYAN HOMELAND

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

There is good reason to believe that the geographical locus of the ancestral haplogroup NO (M214) lay in the Eastern Himalaya. When the two paternal lineages N and O split up, the bearers of haplogroup N set out for East Asia just after the Last Glacial Maximum, braving ice and tundra, and, in a grand counterclockwise sweep, gradually migrated across northern Eurasia as far as west as Lapland (Rootsi *et al.* 2007, Derenko *et al.* 2007, Mirabal *et al.* 2009). I have identified this clade with the paternal spread of Fortescue's (1998, 2001) Uralo-Siberian linguistic phylum, comprising the Uralic, Yukagir, Eskimo-Aleut, Nivkh and Chukotko-Kamchatkan language families (van Driem 2014b). The

ancestral clade N* (M231) is still found in the highest frequency in northern Burma, Yúnnán and Sìchuān.

The fraternal clade O, which appears to be a marker for the linguistic ancestors of the hypothetical East Asian linguistic phylum, remained behind in the Eastern Himalaya. As temperature and humidity increased after the Last Glacial Maximum, the Y chromosomal haplogroup O (M175) split up into the subclades O1 (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 argument 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 9). The O2 (M268) clade split into O2a (M95) and O2b (M176), an event which took place just before the linguistic event horizon.

Asian rice, perhaps both *japonica* and *indica* rice, may have first been domesticated roughly in the area hypothetically imputed to O2 (M268), which would have included southern Yúnnán (van Driem 2011a, 2012a). 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 10). Despite its prevalence in Munda populations, the topology of haplogroup O2a does not support a South Asian origin for this paternal lineage (Kumar *et al.* 2007, Chaubey *et al.* 2010). 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). Whilst the bearers of the O2a

(M95) haplogroup became the *Stammväter* of the Austroasiatics, the other derivative paternal subclade O2b (M176) spread eastward, where they introduced rice agriculture to the areas south of the Yangtze. Though the bearers of the O2b (M176) haplogroup continued to sow seed as they continued to move ever further eastward, they left little or no linguistic traces, except maybe an Austroasiatic name for the Yangtze river, as proposed by Pulleyblank (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 major wave of peopling attested in the Japanese genome (Jin *et al.* 2009, Karafet *et al.* 2009b).

We can identify the O2b (M176) lineage with the Yayoi people, who introduced rice agriculture to Japan, perhaps as early as the second millennium BC, during the final phase of the Jōmon period (Tanaka *et al.* 2004, Hammer *et al.* 2006). The Yayoi appear also to have introduced other crops of continental inspiration to the Japanese archipelago such as millet, wheat and melons. The gracile Yayoi immigrants soon outnumbered the more robust and less populous Jōmon, who had been the first anatomically modern humans to populate Japan. The Y chromosomal haplogroup O2b and other O haplogroups in Japan are later arrivals but account for more than half of all Japanese paternal lineages, with their highest frequencies in Kyūshū.

A Father Tongue theory for Altaic which assumes no close affinity between Altaic and Uralo-Siberian entails that an antique C haplogroup, probably C3, represents an early trace of a paternally disseminated linguistic phylum at a great time depth. Factors such as changes in the ambient material world, social upheaval and cultural transformation, which are known to accelerate the pace of language change, have played an ever more salient role in human life since the Neolithic revolution.⁴ It is conceivable, therefore, that language may have changed at a slower tempo in Palaeolithic times. Scholars of the Altaic language family have reconstructed a very ancient linguistic relationship. No doubt, much of this old linguistic stratum was lost long ago. The remnants of this Father Tongue survive in Japan as Japanese and elsewhere in Asia as the other languages of the Altaic language family, i.e. Korean, Tungusic, Mongolic and Turkic.⁵ Another ancient Father Tongue, entirely distinct from Altaic and anciently introduced to Japan by the bearers of the Y chromosomal haplogroup D2 (M55), also survives today in the Japanese archipelago as Ainu.

At the dawn of the Holocene in the Eastern Himalaya, haplogroup O3 (M122) gave rise to the ancestral Trans-Himalayan paternal lineage O3a3c (M134) and the original Hmong-Mien paternal lineage O3a3b (M7). The bearers of the polymorphism O3a3c (M134) stayed behind in the area comprising Nepal, Sikkim, Bhutan, northeastern India, southeastern Tibet and northern Burma, whilst the bearers of the O3a3b (M7) lineage migrated eastward to settle in the areas south of the Yangtze. On their way, the early Hmong-Mien encountered the ancient Austroasiatics, from whom they adopted rice agriculture. The intimate interaction between ancient Austroasiatics and the early Hmong-Mien not only involved the sharing of knowledge about rice agriculture technology, but also left a genetic trace in the high frequencies of haplogroup O2a (M95) in today's Hmong-Mien and of haplogroup O3a3b (M7) in today's Austroasiatic populations.

On the basis of these Y chromosomal haplogroup frequencies, Cai *et al.* (2011: 8) observed that Austroasiatics and Hmong-Mien 'are closely related genetically' and ventured to speculate about 'a Mon-Khmer origin of Hmong-Mien populations'. It would be more precise to infer that the incidence of haplogroup O3a3b (M7) in Austroasiatic language communities of Southeast Asia indicates a significant Hmong-Mien paternal contribution to the early Austroasiatic populations whose descendants settled in Southeast Asia, whereas the incidence of haplogroup O3a3b (M7) in Austroasiatic communities of the Indian subcontinent is undetectably low. On the other hand, the incidence of Y chromosomal haplogroup O2a amongst the Hmong-Mien appears to indicate a slightly more modest Austroasiatic paternal contribution to Hmong-Mien populations than *vice versa*.

As the Hmong-Mien moved eastward, the bearers of haplogroup O2b (M176) likewise continued to move east. Even further east, the O1 (M119) paternal lineage gave rise to the O1a (M119) subclade, which moved from the Pearl River drainage eastward to the Mìn river drainage in the hill tracts of Fújiàn and across the strait to Formosa, which consequently became the *Urheimat* of the Austronesians (cf. Abdulla *et al.* 2009). Back west in the Eastern Himalaya, the bearers of Y chromosomal haplogroup O3a3c (M134) expanded further throughout Sìchuān and Yúnnán, north and northwest across the Tibetan plateau as well as further westward across the Himalayas and southward into the Indo-Burmese borderlands. In the southwest on the Brahmaputran plain, the early Tibeto-Burmans encountered Austroasiatics, who had preceded them.

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

The ancestral Trans-Himalayan paternal lineage O3a3c (M134) spread from the Eastern Himalaya in a northeasterly direction across East Asia to the North China plain. Subsequently, at a far shallower time depth, the Tibeto-Burman paternal lineage O3a3c (M134) spread from the Yellow River basin into what today is southern China, beginning with the Hàn expansion southward during the Qín dynasty in the third century BC. The ancestral Tibeto-Burman paternal lineage O3a3c (M134) is intrusively present in the Korean peninsula and beyond, although Uralo-Siberian populations such as the Evenki predominantly retain the paternal lineage N. The distribution map of major Trans-Himalayan linguistic subgroups shows the centre of linguistic phylogenetic diversity to be rooted squarely in the Eastern Himalaya, with outliers trailing off towards the loess plains of the Yellow River basin in the northeast. This geographical projection of Trans-Himalayan linguistic diversity appears to reflect the spread of the paternal O3a3c (M134) lineage putatively associated with this linguistic dispersal.

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

The Eastern Himalaya from the Dhaulāgiri to the Liángshān, and more particularly the region comprising Nepal, Sikkim, Bhutan, south-

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

NOTES

1 The EPAS1 gene encodes a protein named endothelial PAS domain protein 1 (EPAS1), alternatively known as hypoxia inducible factor 2-alpha (HIF-2 α). The egl-9 family hypoxia inducible factor 1 gene (EGLN1) encodes the enzyme hypoxia inducible factor prolyl hydroxylase 2 (HIF-PH2) or prolyl hydroxylase domain-containing protein 2 (PHD2)

2 Geneticists working in China might have been able to discern these correlations if they had had access to samples from the Himalayan region and if their interpretations had not been impaired by the faulty Sino-Tibetan framework. Su *et al.* (2000) related Y-chromosomal haplogroup M122 to “Sino-Tibetan”, and their understanding of this putative language family was based on an old edition of Grimes’ *Ethnologue*, which was replete with obsolete and misleading information and incorporated empirically unsupported family tree models. A shortage of information on language families and linguistic phylogeny likewise underlay the inability of Wen *et al.* (2004) and Shi *et al.* (2005) clearly to discern the correlations outlined here. The findings of Su *et al.* (2000) were discussed in my 2001 Périgieux paper, published as van Driem (2005a), and the findings of Wen *et al.* (2004) and Shi *et al.* (2005) were discussed in van Driem (2007).

3 Historically, the empirically well supported Tibeto-Burman or Trans-Himalayan model of language relationship has had to compete with a number of empirically unsupported “Sino” family trees, which were essentially founded upon methodologically flawed comparisons, bewilderment about the historical grammar of Chinese and, more generally, inadequate knowledge of the Trans-Himalayan languages: Sino-Tibetan (Przyluski 1924), Sino-Yeniseian (Schmidt 1926), Sino-Caucasian (Bouda 1950), Sino-Burman (Ramstedt 1957), Sino-Indo-European (Pulleyblank 1966), Sino-Himalayan (Bodman 1973), Sino-Austronesian (Sagart 1993), Sino-Kiranti (Starostin 1994), Sino-Mayan (Jones 1995) and Sino-Uralic (Gão 2008). This episode of linguistic history has been recounted elsewhere (van Driem 2014a).

4 Presumably other factors propelling language change, however, would have been just as operative in Palaeolithic times as they are today, e.g. economy of articulation, emphasis for clarity, politeness and avoidance, taboo, misunderstanding, analogy, back formation, prestige and pretentiousness, group identity signalling as well as changes by structural necessity whereby one change or set of changes inexorably impels another.

5 Martine Robbeets (2010) coined the term Trans-Eurasian for the language family traditionally known as Altaic. She reserves the term Altaic for a putative higher-order subgroup within Trans-Eurasian comprising Turkic, Mongolic and Tungusic, whereas Koreanic and Japonic in her model are treated as branches that have a phylogenetic status coordinate with ‘Altaic’ *sensu* Robbeets within Trans-Eurasian.

BIBLIOGRAPHY

- Abdulla, Mahmood Ameen, and Ikhlaq 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, Jin Li, Jongsun Jung, Daorong 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 Kulawonganuchai, 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 Palittapongpim, 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. Tabada, Adrian Tan, Katsushi Tokunaga, Sissades Tongsimma, 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.
- Bellwood, Peter, Geoffrey Chambers, Malcolm Ross 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, 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: Human Relations Area Files.
- Benedict, Paul King. 1976. 'Sino-Tibetan: another look', *Journal of the American Oriental Society*, 96 (2): 167-197.
- Benedict, Paul King. 1990. *Japanese/Austro-Tai*. Ann Arbor: Karoma Publishers.
- 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.
- 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.
- Bouda, Karl. 1950. 'Die Sprache der Buruscho', *Eusko-Jakintza, Revue d'Études Basques*, IV: 37-50.
- Cai Xiaoyun, Zhendong Qin, Bo Wen, Shuhua Xu, Yi Wang, Yan Lu, Lanhai Wei, Chuanchao Wang, Shilin Li, Xingqiu Huang, Hui Li and the Genographic Consortium. 2011. 'Human migration through bottlenecks from Southeast Asia into East Asia

- during Last Glacial Maximum revealed by Y chromosomes', *Public Library of Science*, 6 (8): e24282.
- Chaubey, Gyaneshwer, Mait Metspalu, Ying Choi, Reedik Mägi, Irene Gallego Romero, Siiri Rootsi, Pedro Soares, Mannis van Oven, Doron M. Behar, Siiri Rootsi, Georgi Hudjashov, Chandana Basu Mallick, Monika Karmin, Mari Nelis, Jüri Parik, Alla Goverdhana Reddy, Ene Metspalu, George van Driem, Yali Xue, Chris Tyler-Smith, Kumarasamy Thangaraj, Lalji Singh, Maito Remm, Martin B. Richards, Marta Mirazon Lahr, Manfred Kayser, Richard Villems and Toomas Kivisild. 2010. 'Population genetic structure in Indian Austroasiatic speakers: The role of landscape barriers and sex-specific admixture', *Molecular Biology and Evolution*, 28 (2): 1013-1024.
- Conrady, August. 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, 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.
- 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.
- van Driem, George. 2005a. '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. 2005b. '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, George. 2007. 'Austroasiatic phylogeny and the Austroasiatic homeland in light of recent population genetic studies', *Mon-Khmer Studies*, 37: 1-14.
- van Driem, George. 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, George. 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, George. 2011b. 'The Trans-Himalayan phylum and its implications for population prehistory', *Communication on Contemporary Anthropology*, 5: 135-142.
- van Driem, George. 2012a. 'The ethnolinguistic identity of the domesticators of Asian rice', *Comptes Rendus Palevol*, 11 (2): 117-132.
- van Driem, George. 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, 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.
- Dulik, Matthew C., Sergey I. Zhadanov, Ludmila P. Osipova, Ayken Askapuli, Lydia Gau, Omer Gokcumen, Samara Rubinstein and Theodore G. Schurr. 2012. 'Mitochondrial DNA and Y chromosome variation provides evidence for a recent common ancestry between native Americans and indigenous Altaians', *American Journal of Human Genetics*, 90 (2): 229-246.
- Forbes, Charles James F.S. 1878. 'On Tibeto-Burman languages', *Journal of the Royal Asiatic Society of Great Britain and Ireland*, X: 210-227.
- 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.
- Gāo Jīngyī. 2008. *Hányǔ yú Běi'ou Yǔyán: Hányǔ yú Wūlā'er Yǔyán jí Yìn'ou Yǔyán Tóngyuán Tànjiū* ['Chinese and Nordic languages: Investigating correspondences between Chinese and Indo-European and Uralic']. Peking: Zhōngguó Shèhuì Kēxué Chūbǎnshè.
- Hammer, Michael F., Tatiana M. Karafet, Hwayong Park, Keiichi Omoto, Shinji Haruhara, Mark Stoneking and Satoshi Horai. 2006. 'Dual origins of the Japanese: Common ground for hunter-gather and farmer Y chromosomes', *Journal of Human Genetics*, 51: 47-58.
- Hodgson, Brian Houghton. 1857. 'Comparative Vocabulary of the Languages of the broken Tribes of Népal', *Journal of the Asiatic Society of Bengal*, XXVI: 317-371.
- Houghton, Bernard. 1896. 'Outlines of Tibeto-Burman linguistic palæontology', *Journal of the Royal Asiatic Society*, 1896: 23-55.
- Huerta-Sánchez, Emilia, and Xin Jin, Asan, Zhuoma Bianba, Benjamin M. Peter, Nicolas Vinckenbosch, Yu Liang, Xin Yi, Mingze He, Mehmet Somel, Peixiang Ni, Bo Wang, Xiaohua Ou, Huasang, Jiangbai Luosang, Zha Xi Ping Cuo, Kui Li, Guoyi Gao, Ye Yin, Wei Wang, Xiuqing Zhang, Xun Xu, Huanming Yang, Yingrui Li, Jian Wang, Jun Wang and Rasmus Nielsen. 2014. 'Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA', *Nature*, doi:10.1038/nature13408
- Hurles, Matthew E., Bryan C. Sykes, Mark A. Jobling and Peter Forster. 2005. 'The dual origin of the Malagasy in Island Southeast Asia and East Africa: Evidence from maternal and paternal lineages', *American Journal of Human Genetics*, 76: 894-901.
- Jin, 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, Timothy A., Lih-Chun Hong, Maude E. Phipps, Mark Stoneking, Mahmood Ameen, Juli Edo, HUGO Pan-Asian SNP Consortium and Naruya Satou. 2013. 'Evolutionary history of continental Southeast Asian: 'Early train' hypothesis based on genetic analysis of mitochondrial and autosomal DNA data', *Molecular Biology and Evolution*, 29 (11): 3513-3527.
- Jones, Robert Reese. 1995. 'Two-Hundred CALMSEA ['Culturally Appropriate Lexico-statistical Model for Southeast Asia'] Comparisons in Sino-Tibeto-Mayan', paper

- presented at the XXVIII International Conference on Sino-Tibetan Languages and Linguistics, University of Virginia at Charlottesville, October 1995.
- Karafet, Tatiana M., J.S. Lansing, Alan J. Redd and Joseph C. Watkins. 2005. 'Balinese Y-chromosome perspective on the peopling of Indonesia: Genetic contributions from pre-Neolithic hunter-gatherers, Austronesian farmers and Indian traders', *Human Biology*, 77 (1): 93-114.
- Karafet, Tatiana M., Fernando L. Mendez, Monica B. Meilerman, Peter A. Underhill, Stephen L. Zegura and Michael F. Hammer. 2008. 'New binary polymorphisms reshape and increase resolution of the human Y-chromosomal haplogroup tree', *Genome Research*, 18: 830-838.
- Karafet, Tatiana M., Stephen L. Zegura and Michael F. Hammer. 2009. '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.
- von Klaproth, Julius Heinrich. 1823. *Asia Polyglotta*. Paris: A. Schubart.
- Kumar, Vikrant, Banrida T. Langstieh, Silpak Biswas, J. Pradeep Babu, T. Nageswara Rao, Kumarasamy Thangaraj, Alla G. Reddy, Lalji Singh and B. Mohan Reddy. 2006. 'Asian and non-Asian origins of Mon-Khmer and Mundari speaking Austroasiatic populations of India', *American Journal of Human Biology*, 18: 461-469.
- Kumar, Vikrant, Arimanda N.S. Reddy, Jagedeesh P. Babu, Tiprisetti N. Rao, Banrida T. Langstieh, Kumarasamy Thangaraj, Alla G. Reddy, Lalji Singh and Battini M. Reddy. 2007. 'Y-chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations', *BioMed Central Evolutionary Biology*, 7: 47, doi:10.1186/1471-2148-7-47.
- Lazaridis, Iosif, and Nick Patterson, Alissa Mittnik, Gabriel Renaud, Swapan Mallick, Karola Kirisanow, Peter H. Sudmant, Joshua G. Schraiber, Sergi Castellano, Mark Lipson, Bonnie Berger, Christos Economou, Ruth Bollongino, Qiaomei Fu, Kirsten I. Bos, Susanne Nordenfelt, Heng Li, Cesare de Filippo, Kay Prüfer, Susanna Sawyer, Cosimo Posth, Wolfgang Haak, Fredrik Hallgren, Elin Founer, Nadin Rohland, Dominique Delsate, Michael Francken, Jean-Michel Guinet, Joachim Wahl, George Ayodo, Hamza A. Babiker, Graciela Bailliet, Elena Balanovska, Oleg Balanovsky, Ramiro Barrantes, Gabriel Bedoya, Haim Ben-Ami, Judit Bene, Fouad Berrada, Claudia M. Bravi, Francesca Brisighelli, George B.J. Busby, Francesco Cali, Michail Churnosov, David E.C. Cole, Daniel Corach, Larissa Damba, George van Driem, Stanislav Dryomov, Jean-Michel Dugoujon, Sardana A. Fedorova, Irene Gallego Romero, Marina Gubina, Michael Hammer, Brenna Henn, Tor Hervig, Ugur Hodoglugil, Aashish R. Jha, Sena Karachanak-Yankova, Rita Khusainova, Elza Khusnutdinova, Rick Kittles, Toomas Kivisild, William Klitz, Vaidutis Kučinskis, Alena Kushniarevich, Leila Laredj, Sergei Litvinov, Theologos Loukidis, Robert W. Mahley, Béla Melegh, Ene Metspalu, Julio Molina, Joanna Mountain, Klemetti Näkkäljärvi, Desislava Nesheva, Thomas Nyambo, Ludmila Osipova, Jüri Parik, Fedor Platanov, Olga Posukh, Valentino Romano, Francisco Rothhammer, Igor Rudan, Ruslan Ruizbakiev, Hovhannes Sahakyan, Antti Sajantila, Antonio Salas, Elena B. Starikovskaya, Ayele Tarekegn, Draga Toncheva, Shahlo Turdikulova, Ingrida Uktveryte, Olga Utevska, René Vazquez, Mercedes Villena, Mikhail Voevoda, Cheryl Winkler, Levon Yempiskoposyan, Pierre Zalloua, Tatijana Zemunik, Alan Cooper, Cristian Capelli, Mark G. Thomas, Andres Ruiz-Linares, Sarah A. Tishkoff, Lalji Singh, Kumarasamy Thangaraj, Richard Villems, David Comas, Rem Sukernik, Mait Metspalu, Matthias Meyer, Evan E. Eichler, Joachim Burger, Montgomery Slatkin, Svante Pääbo, Janet Kel-

- so, David Reich and Johannes Krause. 2014. 'Ancient human genomes suggest three ancestral populations for present-day Europeans', *Nature*.
- Lǐ Huī, Bo Wen, Shu-Juo Chen, Bing Su, Patcharin Pramoongjago, Yangfan Liu, Shangling Pan, Zhendong Qin, Wenhong Liu, Xu Cheng, Ningning Yang, Xin Li, Dinhbinh Tran, Daru Lu, Mu-Tsu Hsu, Ranjan Deka, Sangkot Marzuki, Chia-Chen Tan and Jin Li. 2008. 'Paternal genetic affinity between western Austronesians and Daic populations', *BMC Evolutionary Biology*, 8: 146, doi:10.1186/1471-2148-8-146.
- Li, Jun Z., Devin M. Absher, Hua Tang, Audrey M. Southwick, Amanda M. Casto, Sohini Ramachandran, Howard M. Cann, Gregory S. Barsh, Marcus Feldman, Luigi Luca Cavalli-Sforza and Richard M. Myers. 2008. 'Worldwide human relationships inferred from genome-wide patterns of variation', *Science*, 319: 1100-1104.
- Linz, Bodo, François Balloux, Yoshan Moodley, Andrea Manica, Hua Liu, Philippe Roumagnac, Daniel Falush, Christiana Stamer, Franck Prugnolle, Schalk Willem van der Merwe, Yoshio Yamaoka, David Y. Graham, Emilio Perez-Trallero, Torkel Wadstrom, Sebastian Suerbaum and Mark Achtman. 2007. 'An African origin for the intimate association between humans and *Helicobacter pylori*', *Nature*, 445: 915-918.
- Meiners, Christoph. 1813 [posthumous]. *Untersuchungen über die Verschiedenheiten der Menschennaturen (die verschiedenen Menschenarten) in Asien und in den Südländern, in den Ostindischen und Südseeinseln, nebst einer historischen Vergleichung der vormahligen und gegenwärtigen Bewohner dieser Continente und Eylande*. Tübingen: in der J.G. Cotta'schen Buchhandlung.
- Mendez, Fernando L., Thomas Krahn, Bonnie Schrack, Astrid-Maria Krahn, Krishna R. Veeramah, August E. Woerner, Forka Leyepey Mathew Fomine, Neil Bradman, Mark G. Thomas, Tatiana M. Karafet and Michael F. Hammer. 2013. 'An African American paternal lineage adds an extremely ancient root to the human Y chromosome phylogenetic tree', *American Journal of Human Genetics*, 92: 454-459.
- Mirabal, Sheyla, Maria Reguiero, Alicia M. Cadenas, Luigi Luca Cavalli-Sforza, Peter A. Underhill, Dmitry A. Verbenko, Svetlana A. Limborska and René J. Herrera. 2009. 'Y chromosome distribution within the geo-linguistic landscape of north-western Russia', *European Journal of Human Genetics*, 2009: 1-14.
- Monot, Marc, Nadine Honoré, Thierry Garnier, Romulo Araoz, Jean-Yves Coppée, Céline Lacroix, Samba Sow, John S. Spencer, Richard W. Truman, Diana L. Williams, Robert Gelber, Marcos Virmond, Béatrice Flageul, Sang-Nae Cho, Baohong Ji, Alberto Paniz-Mondolfi, Jacinto Convit, Saroj Young, Paul E. Fine, Voahangy Rasolofo, Patrick J. Brennan and Stewart T. Cole. 2005. 'On the origin of leprosy', *Science*, 308 (5724): 1040-1042.
- Moodley, Yoshan. 2014. 'The population structure of *Helicobacter pylori* across Asia offers insights into prehistoric human migrations', paper presented at the conference *Migrations and Transfers in Prehistory: Asian and Oceanic Ethnolinguistic Phylogeography*, University of Bern, 29 July 2014.
- Müller, Friedrich Max. 1901 [posthumous]. *My Autobiography*. London: Longmans, Green and Company.
- 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; Pacific Linguistics.
- Pinhasi, Ron, Robert Foley and Helen D. Donaghue. 2006. 'Reconsidering the antiquity of leprosy', *Science*, 312: 846.
- Poloni, Estella Simone, Ornella Semino, Giuseppe Passarino, A.S. Santachiara-Benecetti, I. Dupanloup, André Langaney and Laurent Excoffier. 1997. 'Human genetic affinities for Y chromosome P49a.f/TaqI haplotypes show strong correspondence with linguistics', *American Journal of Human Genetics*, 61: 1015-1035 (cf. the erratum published in 1998 in the *American Journal of Human Genetics*, 62: 1267).
- Poloni, Estella Simone, Nicolas Ray, Stefan Schneider and André Langaney. 2000. 'Languages and genes: Modes of transmission observed through the analysis of male-specific and female-specific genes', 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.
- Przyluski, Jean. 1924. 'le sino-tibétain', pp. 361-384 in Antoine Meillet and Marcel Cohen, eds., *Les langues du monde*. Paris: Librairie Ancienne Édouard Champion.
- Pulleyblank, Edwin George. 1966. 'Chinese and Indo-Europeans', *Journal of the Royal Asiatic Society of Great Britain & Ireland (New Series)*, 98 (1): 9-39.
- Pulleyblank, Edwin George. 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.
- 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.
- Rasmussen, Morten, and Xiaosen Guo, Yong Wang, Kirk E. Lohmueller, Simon Rasmussen, Anders Albrechtsen, Line Skotte, Stinus Lindgreen, Mait Metspalu, Thibaut Jombart, Toomas Kivisild, Weiwei Zhai, Anders Eriksson, Andrea Manica, Ludovic Orlando, Francisco de la Vega, Silvano Tridico, Ene Metspalu, Kasper Nielsen, María C. Ávila-Arcos, J. Víctor Moreno-Mayar, Craig Muller, Joe Dortch, M. Thomas P. Gilbert, Ole Lund, Agata Wesolowska, Monika Karmin, Lucy A. Weinert, Bo Wang, Jun Li, Shuaishuai Tai, Fei Xiao, Tsunehiko Hanihara, George van Driem, Aashish R. Jha, François-Xavier Ricaut, Peter de Knijff, Andrea B. Migliano, Irene Gallego-Romero, Karsten Kristiansen, David Lambert, Søren Brunak, Peter Forster, Bernd Brinkmann, Olaf Nehlich, Michael Bunce, Michael Richards, Ramneek Gupta, Carlos Bustamante, Anders Krogh, Robert A. Foley, Marta Mirazón Lahr, François Balloux, Thomas Sicheritz-Pontén, Richard Villems, Rasmus Nielsen, Wang Jun, Eske Willerslev. 2011. 'An aboriginal Australian genome reveals separate human dispersals into Asia', *Science*, 334 (6052): 94-98.
- 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, London.
- Relandus, Hadrianus [Adriaan van Reeland]. 1706, 1707, 1708. *Dissertationum Miscellanearum, Pars Prima, Pars Altera, Pars Tertia et Ultima*. Trajecti ad Rhenum: Gulielmus Broedelet.
- Robbeets, Martine. 2010. 'Trans-Eurasian: Can verbal morphology end the controversy?', pp. 81-114 in Lars Johanson and Martine Robbeets, eds. *Trans-Eurasian Verbal Morphology in a Comparative Perspective: Genealogy, Contact, Chance* (Turcologica 78.) Wiesbaden: Harrassowitz.
- Rootsi, Siiri, Lev A. Zhivotovskiy, 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.
- Sagart, Laurent. 1993. 'Chinese and Austronesian: evidence for a genetic relationship', *Journal of Chinese Linguistics*, 21 (1): 1-62.
- 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. 1906. 'Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens', *Archiv für Anthropologie*, Neue Folge, V: 59-109.
- Schmidt, Wilhelm. 1926. *Die Sprachfamilien und Sprachenkreise der Erde* (2 vols., the second of which is an atlas). Heidelberg: Carl Winter.
- Shi Hong, Dong Yong-li, Wen Bo, Xiao Chun-Jie, Peter A. Underhill, Shen Pei-dong, Ranajit Chakraborty, Jin Li and Sù Bīng. 2005. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122. *American Journal of Human Genetics* 77: 408-419.
- 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.
- Starostin, Sergej Anatol'evič. 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.
- Su, Bing, Chunjie Xiao, Ranjan Deka, Mark T. Seielstad, Daoroong Kangwanpong, Junhua Xiao, Daru Lu, Peter Underhill, Luca Cavalli-Sforza, Ranajit Chakraborty and Li Jin. 2000. 'Y chromosome haplotypes reveal prehistorical migrations to the Himalayas', *Human Genetics*, 107 (6): 582-590.
- Tanaka, Masashi, Vicente M. Cabrera, Ana M. González, José M. Larruga, Takeshi Takeyasu, Noriyuki Fuku, Li-Jun Guo, Raita Hirose, Yasunori Fujita, Miyuki Kurata, Ken-ichi Shinoda, Kazuo Umetso, Yoshiji Yamada, Yoshiharu, Oshida, Yuzo Sato, Nobutaka Hattori, Yoshikuni Mizuno, Yasumichi Arai, Nobuyoshi Hirose, Shigeo Ohta, Osamu Ogawa, Yasushi Tanaka, Ryuzo Kawamori, Masayo Shamoto-Nagai, Wakako Maruyama, Hiroshi Shimokata, Ryota Suzuki and Hideo-toshi Shimodaira. 2004. 'Mitochondrial genome variation in eastern Asia and the peopling of Japan', *Genome Research*, 10A: 1832-1850.
- Thangaraj, Kumarasamy, V. Sridhar, Toomas Kivisild, Alla G. Reddy, Gyaneshwer Chaubey, Vijay Kumar Singh, S. Kaur, P. Agrawal, A. Rai, J. Gupta, C.B. Mallick, N. Kumar, T.P. Velavan, R. Suganthan, D. Udaykumar, R. Kumar, R. Mishra, A. Khan, , C. Annapurna and Lalji Singh. 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, Kumarasamy, Gyaneshwer Chaubey, Vijay Kumar Singh, Ayyasamy Vanniarajan, Ismail Thanseem, Alla G. Reddy and Lalji Singh. 2006b. '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, Ismail, Kumarasamy Thangaraj, Gyaneshwer Chaubey, Vijay Kumar Singh, Lakkakula V.K.S. Bhaskar, B. Mohan Reddy, Alla G. Reddy and Lalji Singh. 2006. 'Genetic affinities among the lower castes and tribal groups of India: Inference from Y chromosome and mitochondrial DNA', *BMC Genetocs*, 7: 42. doi:10.1186/1471-2156-7-42.
- Underhill, Peter A., Peidong Shen, Alice A. Lin, Jin Li, Giuseppe Passarino, Wei H. Yang, Erin Kauffman, Batsheva Bonn -Tamir, Jaume Bertranpetit, Paolo Francalacci, Muntaser Ibrahim, Trefor Jenkins, Judith R. Kidd, S. Qasim Mehdi, Mark T. Seielstad, R. Spencer Wells, Alberto Piazza, Ronald W. Davis, Marcus W. Feldman, Luigi Luca Cavalli-Sforza and Peter J. Oefner. 2000. 'Y chromosome sequence variation and the history of human populations', *Nature Genetics*, 26: 358-361.
- Underhill, Peter A., G. Passarino, A.A. Lin, P. Shen, Marta Miraz n-Lahr, Robert A. Foley, P.J. Oefner and Luigi Luca Cavalli-Sforza. 2001. 'The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations', *Annals of Human Genetics*, 65: 4-62.
- Underhill, Peter A. Natalie M. Myres, Siiri Rootsi, Mait Metspalu, Lev A. Zhivotovsky, Roy J. King, Alice A. Lin, Cheryl-Eliliane T. Chow, Ornella Semino, Vincenza Battaglia, Ildus Kutuev, Mari J rve, Gyaneshwer Chaubey, Qasim Ayub, Aisha Mohyuddin, S. Qaim Mehdi, Sanghamitra Sengupta, Evgeny I. Rogaev, Elza K. Khusnutdinova, Andrey Pshenichnov, Oleg Balanovsky, Elena Balanovska, Nina Jeran, Dubravka Havas Augustin, Marian Baldovi , Ren  J. Herrera, Kummarasamy Thangaraj, Vijay Singh, Lalji Singh, Partha Majumder, Pavao Rudan, Dragan Primorac, Richard Villems and Toomas Kivisild. 2010. 'Separating the post-glacial coancestry of European and Asian Y chromosomes within haplogroup R1a', *European Journal of Human Genetics*, 18: 479-484.
- 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, S  Bing, Ranajit Chakroborty and Jin Li. 2004. 'Genetic evidence supports demic diffusion of Han culture', *Nature*, 431: 302-305.
- Wirth, Thierry, Xiaoyan Wang, Bodo Linz, Richard P. Novick, J. Koji Lum, Martin Blaser, Giovanna Morelli, Daniel Falush and Mark Achtman. 2004. 'Distinguishing human ethnic groups by means of sequences from *Helicobacter pylori*: Lessons from Ladakh', *Proceedings of the National Academy of Science*, 101 (14): 4746-4751.
- Witsen, Nicolaes. 1692. *Noord en Oost Tartarye, ofte Bondig Ontwerp van eenige dier Landen en Volken, welke voormaels bekend zijn geweest, beneffens verscheide tot noch toe onbekende, en meest nooit voorheen beschreven Tartersche en Nabuurige Gewesten, Landstreeken, Steden, Rivieren, en Plaetzen, in de Noorder en Oostelykste Gedeelten van Asia en Europa* (2 volumes). Amsterdam: Fran ois Halma.
- Wood, Elizabeth T., Daryn A. Stover, Christopher Ehret, Giovanni Destro-Bisol, Gabriella Spedini, Howard McLeiod, Leslie Louie, Mike Bamshad, Beverly I. Strassmann, Himla Soodyall and Michael F. Hammer. 2005. 'Contrasting patterns of Y chromosome and mtDNA variation in Africa: Evidence for sex-biased demographic processes', *European Journal of Human Genetics*, 13: 867-876.
- Wulff, Kurt. 1934. *Chinesisch und Tai: Sprachvergleichende Untersuchungen*. Copenhagen: Levin & Munksgaard.
- Wulff, Kurt. 1942 [posthumous]. * ber das Verh ltnis des Malay-Polynesischen zum Indochinesischen*. Copenhagen: Munksgaard.