

MOTHER TONGUE

JOURNAL OF THE ASSOCIATION FOR THE STUDY OF LANGUAGE IN PREHISTORY

ISSUE XI, 2006

CONFERENCE REPORT

Meetings held at Harvard University, October 21-22, 2006, entitled "*Asian Remnant Languages and the Year of the Australoid*". Joint sponsorship: Harvard University and ASLIP (Association for the Study of Language In Prehistory)

- 1 General Logic and Themes of the Conference.
By **Harold C. Fleming**, Gloucester, Massachusetts
- 23 Periods of Globalization over "the Southern Route" in Human Evolution (Africa, Southwest Asia, South Asia, Southeast Asia and Sahul and the Far East). A Meta-Review of Archaeology and Evidence for Symbolic Behavior
By **James B. Harrod**, Portland, Maine
- 85 Out of East Africa by 77,000 BP. A brief synopsis of remarks made by **Alison Brooks**, preliminary to a larger paper later.
- 91 Trombetti, Gatti and the Birth of the Indo-Pacific Concept
By **Jonathan Morris**, São Paulo, Brazil
- 124 New Light On Three South Asian Language Families
By **Franklin C. Southworth**, New York, New York
- 160 The prehistory of Tibeto-Burman and Austroasiatic in light of emergent population genetic studies.
By **George van Driem**, Leiden University (Netherlands)
- 212 Significance of Kusundas and Their Language in the Trans-Himalayan Region
By **B.K.Rana**, Harvard University
- 219 A Multilateral Look at Greater Austric.
By **John D. Bengtson**, Savage, Minnesota & Santa Fé Inst.
- 259 Outside relationships of Australian languages
By **Paul Black**, Darwin, Australia
- 262 The "Lost" Paper: A Belated Conference Postscript.
By **Paul Whitehouse**, London, England
- 275 Was There an Australian Substratum in Dravidian?
By **Vaclav Blazhek**, Masaryk University (Czech Republic)
- 295 Great Andamanese reconstruction underway; A condensed handout with tentative remarks on Papuan and Australian vis-à-vis external language families.
By **Timothy Usher**, San Francisco, CA & Santa Fé Inst.



T 6188

**The prehistory of Tibeto-Burman and Austroasiatic
in the light of emergent population genetic studies**

George van Driem

The Tibeto-Burman language family was first identified by Julius von Klaproth in 1823. The contours of the Austroasiatic language family were first recognised by Francis Mason in 1854. These two linguistic phyla represent keystones for our understanding of the ethnolinguistic prehistory of Asia. What light have recent population genetic studies begun to shed on the models of linguistic relationship? How can the pictures of our linguistic prehistory, our biological ancestry and the archaeological record be correlated to reconstruct the peopling of Asia? What type of questions can we ask of the three distinct data sets?

1. TIBETO-BURMAN

In 1823, Julius Heinrich von Klaproth presented a polyphyletic view of Asian linguistic stocks. He did not presume that the twenty-three distinct families which he had identified represented the definitive inventory. One of the linguistic phyla which he distinguished comprised Tibetan, Chinese and Burmese and all languages that could be demonstrated to be genetically related to these three. Klaproth explicitly excluded languages known today to be members of the Daic or Kra-Dai family, e.g. Thai, or members of the Austroasiatic family, e.g. Vietnamese and Mon (1823: 363-365).

Klaproth did not devise labels for each language phylum he identified. In 1852, John Logan became one of the first to use the term 'Tibeto-Burman' for the phylum identified by Klaproth encompassing Tibetan, Chinese and Burmese, to which Logan added Karen and numerous related languages. Charles Forbes noted that 'Tibeto-Burman' had become the accepted English term for this family (1878: 210). Robert Cust also treated 'Tibeto-Burman', including Karen, as a family distinct from the 'Tai' and 'Mon-Anam' families (1878). Bernard Houghton, who conducted research on languages of Burma, likewise recognised Chinese to be a member of Tibeto-Burman (1896: 28).

Klaproth's Tibeto-Burman outlasted other less well-informed models of language relationship, such as Japhetic, Atactic and Turanian. However, the empirically unsupported Indo-Chinese theory, renamed *sino-tibétain* in

1924, still has its adherents today.¹ The main tenet of the Sino-Tibetan model is that all non-Sinitic languages form a single unitary branch together denominated 'Tibeto-Burman'.² The truncated or pinioned 'Tibeto-Burman' of the Sino-Tibetanists must not be confused with Tibeto-Burman proper, which encompasses Sinitic as one of its subsidiary branches. Tibeto-Burman in its original sense is defined by Tibetan, Burmese and Chinese and furthermore comprises all demonstrably related languages. Diagram 2 illustrates the many new Tibeto-Burman languages and subgroups that have been recognised since 1823.

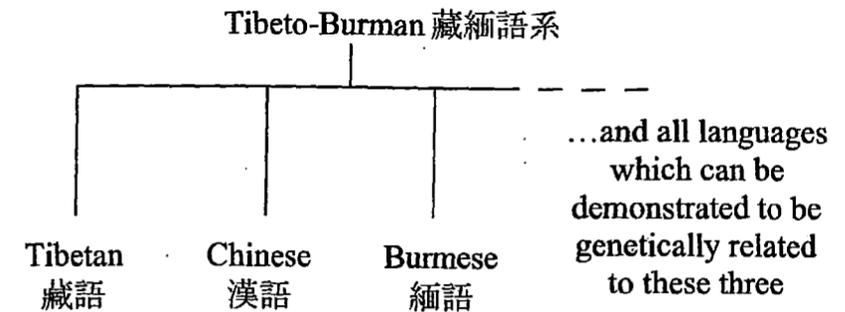


DIAGRAM 1: One of the language families identified by Julius Heinrich von Klaproth in his polyphyletic view of Asian linguistic stocks (1823). He explicitly excluded languages today known to be Kra-Dai or Daic (e.g. Thai, Lao, Shan) and known to be Austroasiatic (e.g. Mon, Vietnamese, Nicobarese, Khmer).

In 2001 in Cambridge, I introduced the metaphor of fallen leaves illustrated in Diagram 2 (van Driem 2001, 2002). The model attempts to identify the constituent branches of the family and draw the focus of attention back to the centre of Tibeto-Burman linguistic diversity, which lies in the eastern Himalayas and the Indo-Burmese borderlands. The patch of fallen leaves on the forest floor provides a more informative framework because all recog-

¹ The term *sino-tibétain* was coined as a new name for Indo-Chinese by Jean Przyluski (1924) and later introduced into English as 'Sino-Tibetan' by Przyluski and Luce (1931).

² The rise and fall of Sino-Tibetan and its racist underpinnings were discussed in the keynote address to the joint meeting of the 14th Annual Conference of the International Association of Chinese Linguistics and the 10th International Symposium on Chinese Languages and Linguistics at the Academia Sinica in Taipei on 27 May 2006.

nised subgroups are presented without a false or, at best, unsupported tree, such as Sino-Tibetan. The new metaphor still implies the existence of a tree, but we cannot lift our gaze from the forest floor to see the tree because we cannot look directly into the past. Instead, historical comparative work may enable us to see the shadows which the branches cast between the leaves on the forest floor.

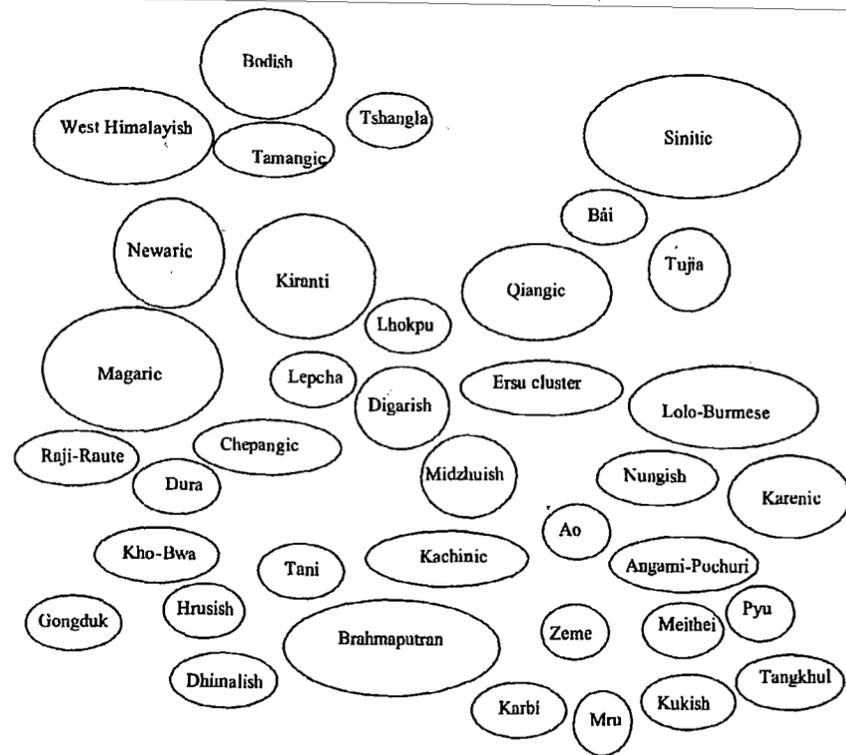


DIAGRAM 2: Tibeto-Burman subgroups identified since Julius von Klaproth. Brahmaputran may include Kachinic and Dhimalish. Other subgrouping proposals are discussed in the handbook (van Driem 2001).

Whether a language family appears to be more rake-like or more tree-like is often a function of the state of the art in historical comparative linguistics rather than a statement about linguistic phylogeny. With the inexorable progress of Indo-European studies, even the twelve branches of this most well-studied language family, once depicted in the pleasing shape

of August Schleicher's branching oak, have also gradually assumed a more rake-like appearance and so too come closer to the fallen leaves model.

The geographical distribution of the branches of the Tibeto-Burman language family reveals an intriguing pattern which raises questions and permits us to formulate hypotheses about the provenance of the linguistic ancestors of Tibeto-Burman language communities and the location of the Tibeto-Burman homeland. Future research will show the number of diamonds representing branches of the family to be more or less than shown in Diagram 3. Some groups may coalesce, and others may be split up. For example, the Dura language may one day be demonstrated to be a member of another known Tibeto-Burman subgroup, whereas 'Qiāngic', as currently conceived, could turn out not to be a valid clade at all but to consist of a number of independent clusters. In Diagram 2, the Ērsū cluster is another name for 'Southern Qiāngic', and may in fact consist of several subclusters. Qiāngic is 'Northern Qiāngic', which is currently supposed to include the rGyal-rongic group recognised by Jackson Sun (Sūn Tiānxīn) and Huáng Bùfán. In fact, the precise phylogenetic relationships between the diverse rGyal-rong languages, Ērgōng, Qiāng, Mi-ñag (Mùyǎ), Tangut, Ērsū, Lǔsū, Tosu (Duōxù), Nàmùyì, Shìxīng, Guìqióng, Choyo (Quèyù), Zhābà and Prinmi (Pǔmǐ) have yet to be demonstrated. In short, there is a lot of work left to be done in Sìchuān and Yúnnán provinces.

Just like British scholars of the nineteenth century, Jaxontov proposed a homeland in Sìchuān (1977). Subsequently, so did I (van Driem 1998). Peiros' classification based on the highest lexicostatistical diversity of primary taxa purportedly indicates 'a possible location of the homeland in the territories south of the Himalayas', whereas the location of Sinitic could be 'easily explained as the result of later migration' (1998: 217). In December 2004 at the 10th Himalayan Languages Symposium in Thimphu, I presented the argument of the internal linguistic diversity of the family for a Himalayan homeland for Tibeto-Burman. Questions of linguistic phylogeny are fundamentally resolved by historical linguistic comparison, but the location of the Tibeto-Burman homeland is not just a linguistic question.

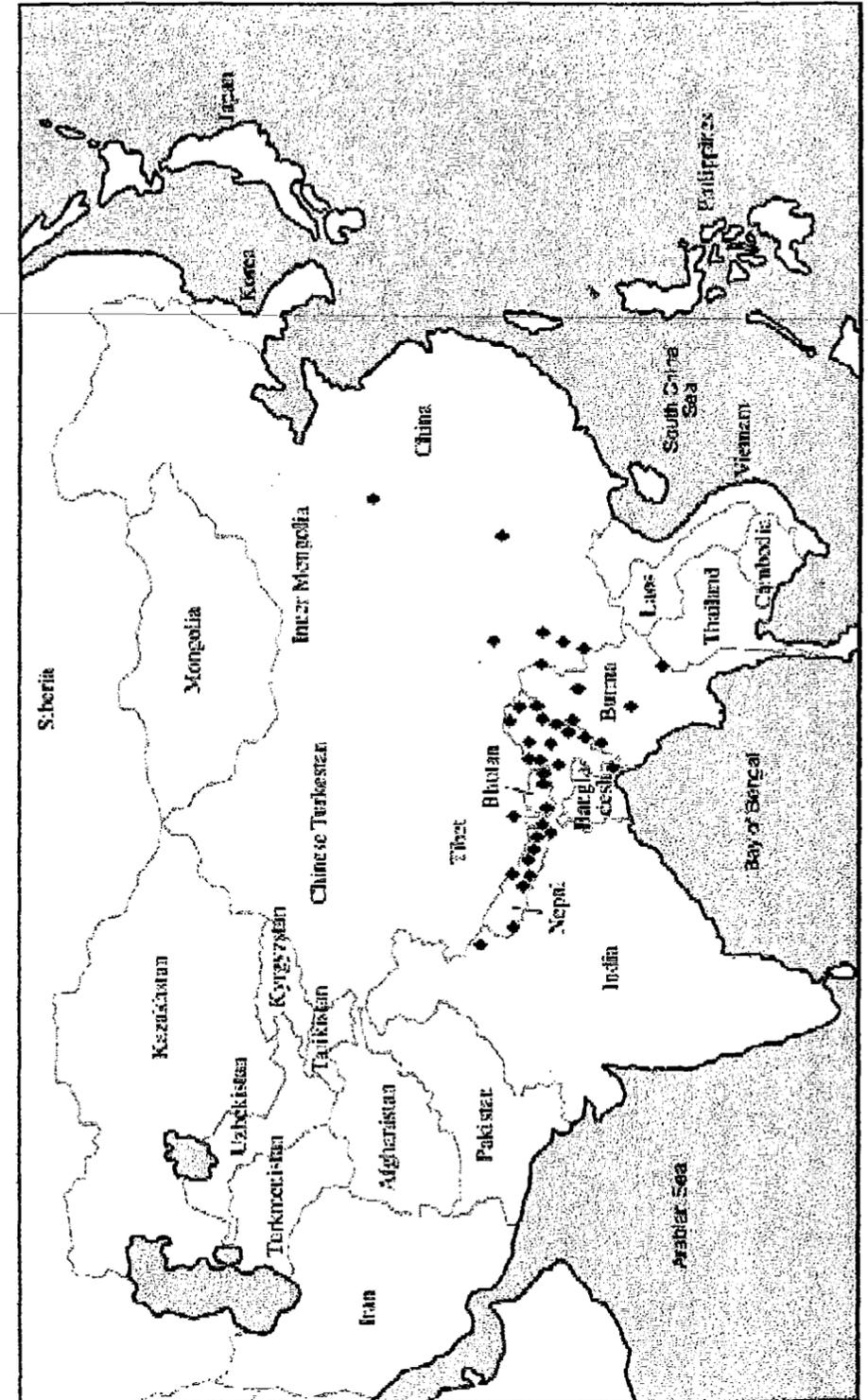
In addition to the comparative method, new mathematical models which aid lexicostatistical comparison may prove a useful tool. Elsewhere I have discussed the history of lexicostatistics since its invention by Rafinesque in 1831 (van Driem 2005). From the time of Dumont d'Urville (1834), the real advantages as well as the limitations of Rafinesque's method of lexicostatistics have become increasingly evident if the methodology is applied without the insights of historical linguistics. Hendrik Karel Jan Cowan (1959) was amongst the first to stress that practitioners of glottochronology

and lexicostatistics then appeared oblivious to the far greater probabilistic significance of structural correspondences between grammatical systems. A second flaw in the reasoning of glottochronology is that different languages are historically known to have changed at different rates. Finally, the validity of some of the mathematical models employed in glottochronology was also challenged, e.g. Bergsland and Vogt (1962), Chrétien (1962), Guy (1994).

More recently, however, mathematical models used in glottochronology have undergone refinement, e.g. Gray and Atkinson (2003). Russell Gray is making every attempt to accommodate the criticisms of comparative linguists and so increasingly to incorporate historical linguistic insights into his mathematical model. Such models appear to work fine for Austronesian, a language family in which cognacy judgements are relatively non-controversial. However, the model can give false and misleading results when based on cognacy judgements for language families where such judgements are difficult and more controversial, e.g. Tibeto-Burman. In other cases, the putative phylogenetic construct is purely hypothetical and the cognacy judgements remain speculative, e.g. Sagart's Sino-Austronesian, Starostin's Sino-Caucasian (cf. van Driem 2005). Whenever many of the supposed cognates are not in fact cognate or the putative phylogenetic construct does not correspond to any reality that ever existed in the past, then the numbers churned out by the mathematical model will be meaningless, however good they may look.

As long as the caveats regarding lexicostatistical models are kept in mind, then there need not be much harm in using these potentially useful tools. The lexicostatistical attempt by Dèng and Wáng (2003) to arrive at a tree of some of the Tibeto-Burman languages spoken in China is a good beginning. Such studies will in time hopefully be extended to cover the Tibeto-Burman language family as a whole, most branches of which are represented exclusively outside of China.

DIAGRAM 3 next page: The geographical distribution of the major branches of the Tibeto-Burman. Each diamond represents not a language, but a major subgroup.



At the same time, attempts have been made by various scholars to go beyond the maximum time depth usually considered accessible to practitioners of historical linguistic comparison. For example, Dunn *et al.* (2005) have attempted to use typological features to go beyond the time depth of 8,000 years, give or take two millennia, reconstructible by conventional historical linguistics. They arrive at a tree for the hitherto unrelatable Papuan isolates of island Melanesia, which suggests to them a late Pleistocene dispersal, now visible only as vestigial structural similarities between the languages and no longer in the form of any reconstructible vocabulary or morphology. In a similar vein, Johanna Nichols (1992, 1998) has invoked her notion of a diffusion or spread zone to the Eurasian heartland, but Michael Fortescue (1998) has shown that such notions can only be meaningfully implemented when the comparative method has first taken us as far back as it can take us.

In the case of Tibeto-Burman, it would be premature to use typological comparison of this sort to attain benthic time depths. By the same token, expediency would appear to be the principal motivation behind a rush to use mathematical tools for lexicostatistical comparison at a time when most Tibeto-Burman languages have yet to be documented in adequate detail and historical linguistic comparison has yet to be carried out to anything approaching a satisfactory degree of refinement.³

2. AUSTROASIATIC

The contours of the Austroasiatic language family were identified by the American Baptist missionary Francis Mason (1854, 1860), when he realised that Munda languages of India, such as Kol and Ho, belonged to the same language family as Mon or 'Talaing', spoken in and around Pegu in Burma. Julius von Klaproth had previously recognised that there existed a family of languages encompassing Mon, Vietnamese (then more commonly known as 'Annamitic'), Khmer and Nicobarese. The family subsequently became known as Mon-Annam or Mon-Khmer, but, after Mason's addition of the Munda languages, the name 'Mon-Khmer-Kolarian' came into circulation for the phylum as a whole. Subsequently this unwieldy name was replaced by 'Austroasiatic', a coinage of the Austrian priest Wilhelm Schmidt (1904,

³ The Trans-Himalayan Database Programme <www.iias.nl/himalaya/> serves both traditional historical linguistic comparison and aims to collaborate with the lexicostatistical programme developed by Russell Gray and his associates.

1906). Schmidt's new label stuck even though his ideas about the family were decidedly fuzzier than those of some of his predecessors.

Austroasiatic is a sorely neglected field of linguistics that has been kept alive by a very few passionate and knowledgeable scholars.⁴ Paradoxically, the level of scholarship in Austroasiatic linguistics is such that this family, unlike Tibeto-Burman, at least has a tentative family tree. Currently, the most informed and authoritative Austroasiatic *Stammbaum* is the language family tree presented by Diffloth (2001, 2005), reproduced in modified form in Diagram 4. In contrast to earlier family trees, Diffloth's Austroasiatic family splits up into three major nodes, i.e. Munda, Khasi-Khmuic and a new 'Mon-Khmer'. In this new tripartite division, Munda is still one of the primary branches of Austroasiatic, representing the native heart of the Indian subcontinent. The Khasi-Khmuic branch represents what might be thought of as 'Inland Austroasiatic', and a more precisely delineated Mon-Khmer represents 'Littoral Austroasiatic'.

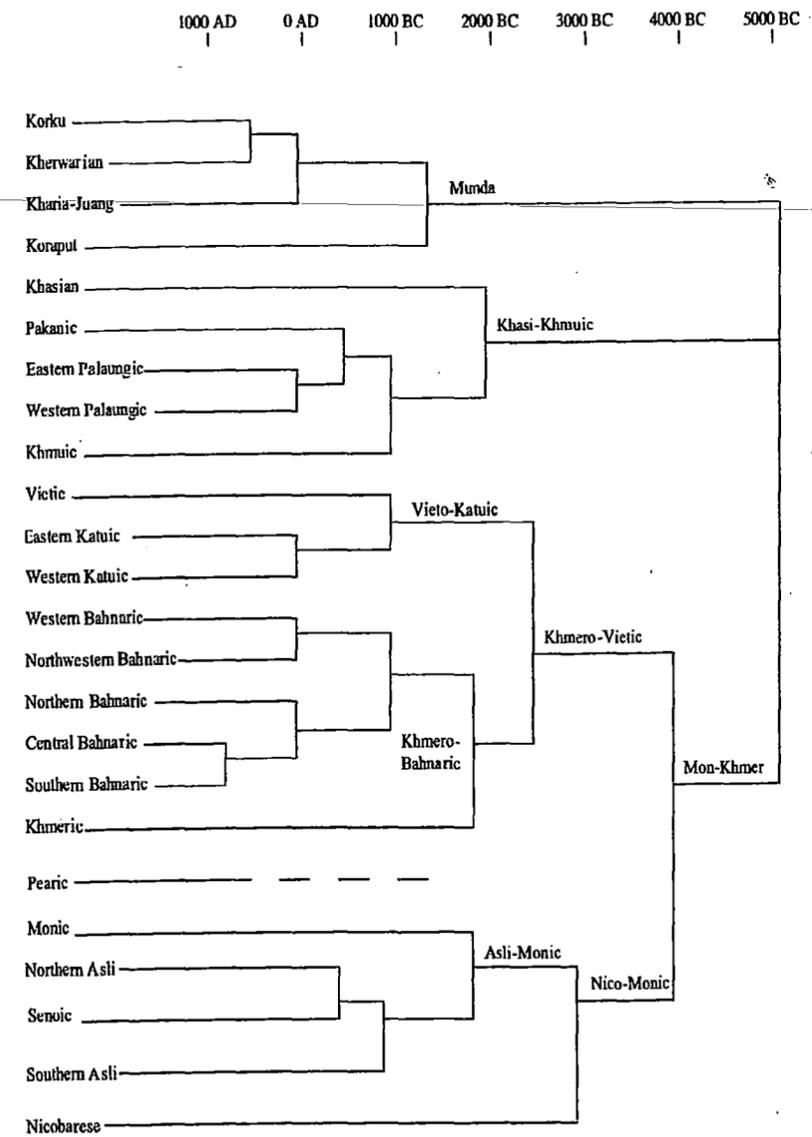
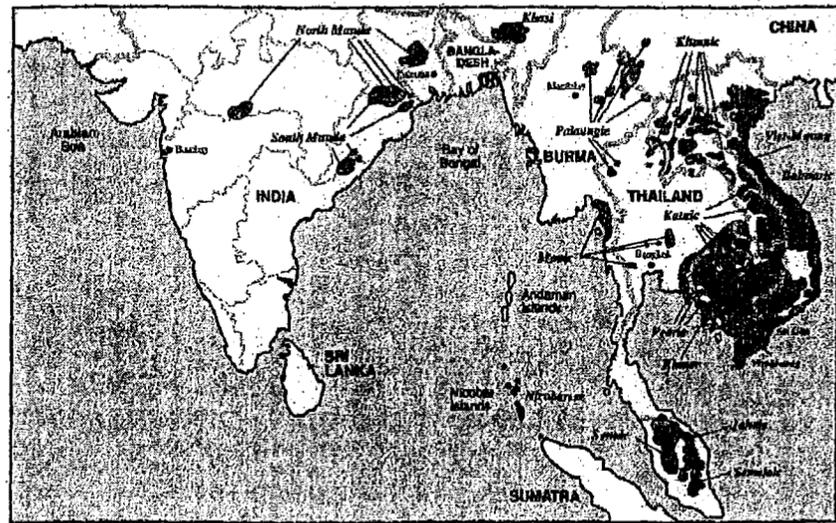
The new Mon-Khmer comprises Khmero-Vietic and Nico-Monic. Each of the two sub-branches of Mon-Khmer is further subdivided, with Nico-Monic consisting of Asli-Monic and Nicobarese, and Khmero-Vietic breaking up into Vieto-Katuic and Khmero-Bahnaric. Conspicuously, Diffloth had initially left out Pearic on purpose because its genetic affinity was still, as he put it, *en chantier*, but it is at least safe to say that its greatest genetic affinity is not with the Munda or Khasi-Khmuic branches, but with Mon-Khmer. Many more phylogenetic insights are contained in Diffloth's burgeoning, highly detailed but as yet unpublished Austroasiatic comparative database.

DIAGRAM 4 next page: Austroasiatic with Gérard Diffloth's tentative calibration of time depths for the various branches of the language family (modified from Diffloth 2001, 2005). The precise phylogenetic propinquity of Pearic, after Khmeric loan layers have been stripped off, remains uncertain except that Diffloth observes that Pearic is Mon-Khmer and not 'une espèce de vieux khmèr', as some scholars once maintained. This diagram arranges in a tree-shaped phylogeny the fourteen recognised branches of Austroasiatic, i.e. North Munda, South Munda, Khasian, Pakanic, Palaungic, Khmuic, Vietic, Katuic, Bahnaric, Khmeric, Pearic, Monic, Aslian and Nicobarese.

⁴ The International Conference on Austroasiatic Linguistics, a forum which convened only twice in the 1970s, has recently been resurrected. See <www.iias.nl/icaal/>.

Any reconstruction of Austroasiatic population prehistory must start out from the present and historically attested geographical distribution of Austroasiatic subgroups. Diagram 5 shows the geographical distribution of Austroasiatic subgroups with the exception of the recently documented enclaves of Pakanic in southern China. We can all look forward to Gérard Diffloth's new detailed Austroasiatic map which is currently in production. When we compare the new phylogenetic model for Austroasiatic with the geographical distribution of Austroasiatic subgroups, a number of hypotheses concerning the possible location of an Austroasiatic homeland suggest themselves. In fact, in the past the most diverse homeland sites have been proposed for Austroasiatic, and most of these are discussed in my handbook (van Driem 2001: 289-332).

DIAGRAM 5: Geographical distribution of Austroasiatic subgroups (van Driem 2001: 267). Recently documented Pakanic enclaves in southern China are not yet shown.



On the basis of linguistic palaeontology Diffloth has argued that the reconstructibility at the Proto-Austroasiatic level of words for tree monitor, ant eater, buffalo, mountain goat, bear cat, elephant, peacock, rhinoceros and bamboo rat as well as the rich reconstructible rice cultivation vocabulary imply that the Austroasiatic homeland was located in the tropics. The Hémūdù culture at the mouth of the Yangtze (5000-4500 BC) provides the best unambiguous evidence for a population for whom rice is the staple. The oldest direct evidence for domesticated rice, however, dates from 6500 BC and is from the Bāshídàng and Péngtóushān sites belonging to the Péngtóushān culture (7500-6100 BC) on the middle Yangtze in what today is Húnán and from the Jiǎhú culture (6000-7000 BC) on the Huái river further north in what today is Hénán.

Since the archaeological sites reflecting the oldest known rice cultivators are located along the middle Yangtze, Diffloth logically raised the palaeoclimatological question whether the faunal landscape which existed in this area at the putative time depth of Proto-Austroasiatic would be compatible with the environment suggested by linguistic palaeontology. Clearly, by the faunal criterion large tracts of the Indian Subcontinent and Southeast Asia also remain homeland candidates.

Any successful correlative study of the historical linguistic picture and the population genetics of the the modern language communities will have to provide an account for the manifest somatological or phenotypical difference between Munda speakers on one hand and speakers of Khasi-Khmuic and Mon-Khmer languages on the other, as well as comparably great differences between Aslian negrito groups and the linguistically closely related Nicobarese. The meaningfulness of any conjectures that we base on such correlative studies depends on the reliability of the linguistic reconstructions and language family tree as well as on the degree of resolution, refinement and thoroughness of sampling of our genetic assays.

Operating on the assumption that frequency gradients of Y haplogroups, mtDNA polymorphisms or autosomal haplotypes may correlate precisely or partially with the distribution of Austroasiatic language communities, we may still wonder whether such gradients necessarily reflect the people who introduced and disseminated any putative proto-language.

3. AUSTROASIATIC: LANGUAGES, GENES AND ARCHAEOLOGY

The linguistic ancestors of a language community were not necessarily the same people as the biological ancestors of that community. We invariably get all of our DNA from our biological parents, but only in most cases is our native language also that of our parents. So, notwithstanding the probabilistic correlation between languages and genes, the discrepancies between the two versions of prehistory can tell us at least as much about what went on in the past as the grand correlations.

The genetic picture also shows a certain sexual dimorphism in linguistic prehistory. In Baltistan, located in what today is northern Pakistan, the local Tibetan dialects are the most conservative of all Tibetan languages, preserving consonant clusters retained in Classical Tibetan orthography but wholly lost in most other Tibetan dialects. Yet the Balti abandoned the Tibetan script after they were converted to Islam in the fifteenth century, although native activists have in recent years begun reintroducing the Tibetan script, e.g. on shop signs, to the displeasure of central government authorities. Paradoxically, the old consonant clusters ceased to be pronounced as such in most areas throughout Tibet where the conservative indigenous orthography representing these phonological segments remained in use. Genetic studies of the Balti populations show intrusive Y haplogroups from the Near East, whereas the mitochondrial DNA of the Baltis is predominantly Tibetan mtDNA (Poloni *et al.* 1997, 2000, Zerjal *et al.* 1997, Quintana-Murci *et al.* 2001, Qamar *et al.* 2002). So, the religion of the Balti appears to be a paternal heritage, whilst the languages that they speak are literally mother tongues.

Genetic studies have suggested that the distribution of Indo-Aryan language communities in northern India patterns well with intrusive Y haplogroup frequencies emanating from the northwest, reflecting what many linguists and archaeologists had long thought about Indian prehistory. The picture of an Aryan invasion emerging from the R̥gveda, in the words of Mortimer Wheeler, 'constantly assumes the form of an onslaught upon the walled cities of the aborigines', i.e. the *purās*, and the Aryan god Indra is a *purandara* 'destroyer of aboriginal forts', who shattered ninety such strongholds (1966, 1968). Many scholars have connected this destruction of aboriginal fortresses and the conquest of subjugated Dasyus recounted in the Aryan hymns to the extinguishing of the Indus Valley civilisation. At any rate, the activities depicted were a predominantly male occupation. Genetic studies have suggested that the Y haplogroups L, R1a and R2 spread from the northwest along with Indo-Aryan language across northern India and to

Ceylon, whereas mitochondrial lineages prevalent in India are overwhelmingly indigenous to the Subcontinent (Kivisild *et al.* 1999a, 1999b, Wells *et al.* 2001, Cordaux *et al.* 2003, Kivisild *et al.* 2003, Baig *et al.* 2004, Cordaux *et al.* 2004a, Metspalu *et al.* 2004, Quintana-Murci *et al.* 2004, Thangaraj *et al.* 2005). At the same time, the spread of Indo-Aryan languages unambiguously attests to an ancient linguistic intrusion into the Subcontinent from the northwest.

So, were Vedic and Avestan introduced as father tongues? At the Indo-Pacific Prehistory Association conference in Taipei in 2002, I gave the name 'Father Tongue hypothesis' to the correlations observed between language spread and the geographical distribution of Y haplogroup frequencies by Poloni *et al.* (1997, 2000). One of the areas where this hypothesis appeared to hold was the linguistic intrusion of Indo-European into the Indian subcontinent from the northwest. A recent study by Sahoo *et al.* (2006), however, attempts to challenge this Y chromosome picture which has emerged from several previous population genetic studies.

Their study is a major leap forward, but the sampling is still coarse, and the survey neglects to systematically distinguish between Turks, Kurds and other language communities in the Near East and between Indo-Iranian and Turkic language communities in Central Asia. A fine-mesh and more ethnolinguistically informed sampling remains a realisable goal. More crucially, the reasoning in Sahoo *et al.* (2006), edited by Colin Renfrew, omits to take note that Central Asia saw major incursions of Altaic populations in historical times. An ethnolinguistically low-resolution survey of present Central Asia Y chromosomal genography cannot be presumed to reflect the genography of the region during, say, the Bronze Age Andronovo culture and the Bactria Margiana archaeological complex.

In fact, the probable replacement of Y chromosomal lineages during the Altaicisation of Central Asia is consonant with the observation made by Sahoo *et al.* (2006) that the Y haplogroups E, I, G, J* and R1*, which have a combined frequency of 53% in Turks of Asia Minor and 24% in Central Asia, are virtually absent in India, except for a trickling of R1*. Also absent in India are haplogroups C3, D, N and O, which are 'specific to Central Asia', where they have a combined frequency of 36%. Likewise, the complete absence in India of the derived C3 lineages, which account for over 95% of the C haplogroup variation in Central Asia, 'cannot be ascribed to a recent admixture from the north' (op.cit. 845). At the same time, the J2 haplogroup, which appears to emanate from the Arabian Peninsula and, unlike haplogroups N and R1a, attains no high frequency in Ceylon, 'indicates an unambiguous recent external contribution, from West Asia

rather than Central Asia' (op.cit. 87), and indeed this gradient probably reflects the historically attested male-borne eastward spread of Islam.

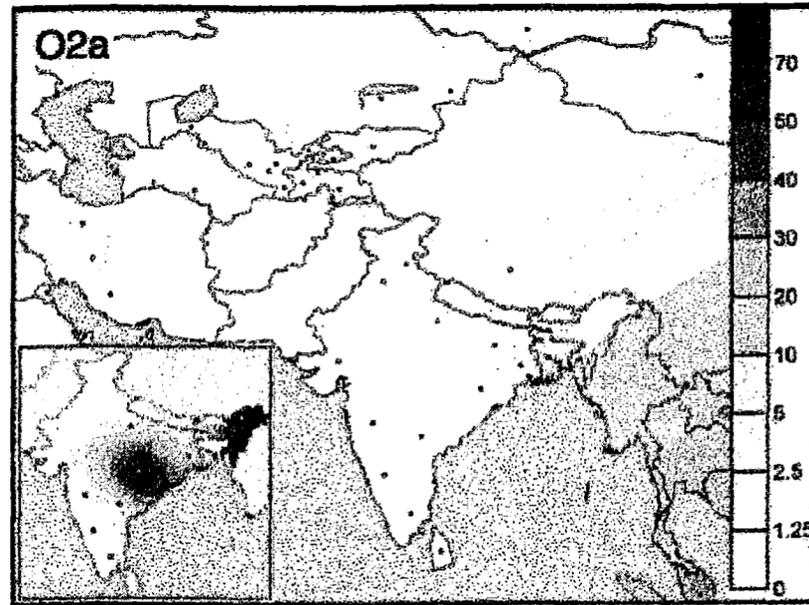
The population genetic work is ongoing, and some preliminary findings are prone to being interpreted prematurely in terms of their potential significance for population prehistory. Just two years ago, an article by Langstieh *et al.* (2004) created a stir amongst scholars of Khasi because the study addressed the provenance of the Garos and Khasis of the Meghalaya. This valuable contribution raised more questions than it answered. The researchers claimed that the population of the Meghalaya is homogeneous, whereas the Garos and Khasi tribes are linguistically unrelated. Judging from their median joining network, the Garos would appear to be an ethnic subset of the Khasis, something which suggests that the Garos are more homogeneous as a group than the Khasi tribes. Moreover, the purported homogeneity of the populations of the Meghalaya was based on comparison with the Chinese and North American Indians! Obviously it would be more meaningful to conduct fine-mesh genetic comparison of the Khasi with Pakanic, Palaungic and Khmuic language communities, who are their closest linguistic relatives. By the same token, fine-mesh studies should be undertaken to compare the Garos with the Bodos, Rabhas and Dimasas, who are their closest linguistic relatives, as well as other linguistically less related population groups of northeastern India.

Microsatellites or short tandem repeats (STR) are highly polymorphous, but the short tandem repeats chosen by Langstieh *et al.* were not necessarily the optimal choice as genetic markers for gauging differences between closely related populations. In all of the Himalayan groups which we have been testing — and our sampling represents a highly varied and heterogeneous collection of peoples and language communities — we do not always see that much variation in the short tandem repeats as these researchers have found in the Meghalaya (Kraaijenbrink *et al.* 2006a, 2006b, Parkin *et al.* 2006a, 2006b). So, the peoples of the Meghalaya show up as a highly heterogeneous population, but the researchers cannot yet know this for sure, for they have not been able to compare their findings with data on other relevant groups. Further studies will have to corroborate the impression that the Meghalaya may be an area where the antiquity and genetic heterogeneity of the populations is relatively great.

In addition to the studies already mentioned, relevant population genetic studies have begun to chart the autosomal lineages, the mitochondrial or maternal lineages and the Y chromosome haplogroups representing the paternal lineages of Austroasiatic language communities and neighbouring population groups, e.g., Ashma *et al.* (2002), Banerjee *et al.* (2005a, 2005b),

Debnath and Chaudhuri (2005), Ding *et al.* (2000), Kashyap *et al.* (2002a, 2002b), Krithika *et al.* (2005), Kumar *et al.* (2004), Maity *et al.* (2003), Sahoo *et al.* (2002), Shi *et al.* (2005, 2006), Singh *et al.* (2006), Sù *et al.* (1999, 2000), Thomas *et al.* (2004), Watkins *et al.* (2005).

DIAGRAM 6: Spatial frequency distribution of Y haplogroup O2a for caste populations [big map] and tribal populations [inset map], reproduced here from Sahoo *et al.* (2006: 846).



Much progress has been made in Y chromosome phylogeny since the seminal contribution by Underhill *et al.* (2001). A number of research teams have mooted a possible link between the distribution of Austroasiatic language communities and the M95 mutation, i.e. Y chromosomal haplogroup O2a, e.g. Sù *et al.* (2000), Kayser *et al.* (2003), Kivisild *et al.* (2003), Cordaux *et al.* (2004b). Frequency gradients for Y haplogroup O2a are mapped for the Indian Subcontinent and Southeast Asia by Sahoo *et al.* (2006) and are shown here in Diagram 6. At the same time, maternal lineages of Munda groups appear to be old and indigenous to the Subcontinent, as indeed can

be said of many Indian mitochondrial lineages (Kivisild *et al.* 1999a, 1999b, 2003).

Chaubey (2006) has ascertained that R7a is a salient mitochondrial haplogroup characterising Munda language communities in Jharkhand, Chattisgarh and Bihar and that the related mtDNA haplogroup R7b features saliently in maternal lineages of Dravidian tribal populations in the same geographical range, where they generally inhabit more southern portions of this area than the Munda language communities. The R7 lineage is hardly found outside of this geographical area, is exceedingly rare in caste populations and has not been found in Tibeto-Burman populations. Does R7 then represent the maternal signature of an ancient indigenous South Asian population to whom ancient bearers of the Y-chromosomal haplogroup O2a introduced Austroasiatic language from the northeast whilst other ancient males introduced Dravidian language from the west? What seems to be clear from the mitochondrial picture at any rate is that the Munda maternal lineage derives from early human settlers of the Subcontinent, whilst one of the predominant Y chromosome haplogroups in Austroasiatic language communities in India argues for a Southeast Asian paternal homeland for Austroasiatic.

Sahoo *et al.* (2006: 847) rightly caution 'against simplistic interpretations of either linguistic or genetic correlations'. By the same token, some of the formulations in Sahoo *et al.* (2006) provide grounds for cautioning against the use of a single explanatory model in our interpretation of the genetic, archaeological and linguistic data. Portions of the article reflect a *Hineininterpretieren* of the Farming-Language Dispersal theory into the genetic findings. This slant in no way diminishes the value of the correlation of the Y chromosomal haplogroup O2a with the geographical distribution of Austroasiatic language communities proposed by various research teams, viz. Sù *et al.* (2000), Kayser *et al.* (2003), Kivisild *et al.* (2003), Cordaux *et al.* (2004b), Sahoo *et al.* (2006). Yet this single-model interpretation of genetic findings raises a more general issue which is of central relevance to the ways that we think about the prehistory of language families such as Austroasiatic.

It is tempting to assume that genes, languages and archaeological horizons have always tended to move in tandem with the incremental spread of Neolithic agriculture and to convince ourselves that this model generates the most parsimonious explanations. In fact, realities on the ground were often more complex. This complexity is not only suggested by the dissonance between the different pictures of prehistory reconstructible through the three disciplines, but more so by the multi-layered nature of the distinct pictures which emerge from linguistics, population genetics and archaeology. For

example, Ossetian, an East Iranian language is spoken in an area which lies decidedly to the west of most West Iranian language communities, attesting to the ancient migration of the Alans and Sarmatians to the north central Caucasus.

The geographical distribution of gene frequencies not only reveals distinct migrations, sometimes in opposing directions at different time depths, but detailed future studies may also enable us to ascertain the relative chronology of the distinct layers of genetic diffusion at different times across the same areas. Archaeology defines specific cultural assemblages with definable horizons and identifiable colonial exponents. The farming-language dispersal model necessarily works in the case of Austronesian, where the geographical spread of the language family has to a major extent resulted from the colonisation of previously uninhabited insular environments emanating from Formosa, or perhaps from Hémūdù via Formosa.⁵ Yet we must question whether the latter theory has the same explanatory power to account for the spread of language families under the circumstances which prevailed on the land masses where most of prehistory unfolded. For an archaeologist contemplating language families, the urge is inevitably irresistible to associate the geographical spread of technologically advanced Neolithic civilisations into more backward areas with the spread of peoples and language families.

More fundamentally, the premisses of the farming-language spread theory ought to be questioned. The surplus generated by an agricultural economy and the stratified social and command structure enabled by a Neolithic lifestyle are held to have driven demographic spread into many areas. This argument is plausible, but this argument is not the crux of the farming-language dispersal theory. Crucial to the model is the tenet that the incremental spread of the Neolithic as such is associated with 'the foundation dispersals' of language families. This theory therefore presumes that the ancient spread of language families unfolded in the same direction as the demographic spread driven by Neolithic agriculture.

⁵ The Hémūdù culture at the mouth of the Yangtze (5000-4500 BC) provides the best unambiguous evidence for a population for whom rice is the staple. The oldest direct evidence for domesticated rice, however, dates from 6500 BC and is from the Bāshídàng and Péngtóushān sites belonging to the Péngtóushān culture (7500-6100 BC) on the middle Yangtze in what today is Húnán and from the Jǐāhú culture (6000-7000 BC) on the Huái river further north in what today is Hénán. Cultivated rice has been recovered from 南關里 Nánguānlǐ in southeastern Táiwān dating from ca. 3000 BC (Tsang 2004).

The very opposite may be what actually happened in many cases. Across the Fertile Crescent, agriculture was adopted by ethnolinguistically unrelated populations, and agriculture spread effortlessly across ethnolinguistic boundaries without disrupting them in any significant way. Sumerian pictographic script, developed ca. 3200 BC, appeared millennia after the invention of agriculture. Sumerian, Elamite, Akkadian,⁶ Hurrian, Hattic and other contemporaneous agricultural civilisations were in all likelihood not the first cultivators of the region. Yet even these antique agricultural language communities have left no surviving linguistic descendants. The earliest recorded and reconstructible history of the Near East bears witness to the permeability of linguistic boundaries for the dissemination of agriculture and crops.

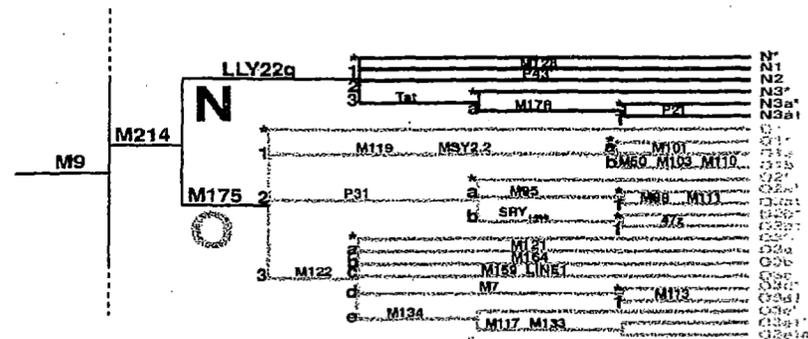
The Bronze Age of Asia Minor and Mesopotamia is characterised by a long period of incursive population movements into, rather than out of Anatolia and the Fertile Crescent, lured by the relative affluence of urban centres supported by agricultural surplus. Gutæans, Amorites, Kassites and other peoples were drawn in by the promise of the good life. Most linguistic reconstructions presume that Indo-European groups such as the Hittites and Mitanni likewise came to settle in Asia Minor and the Fertile Crescent from elsewhere. Toponymical evidence and details about the cults of certain deities have been used to argue that even the Sumerians originally migrated from an earlier northern homeland to lower Mesopotamia. Were the motivations of migrating peoples in agricultural and pre-agricultural societies and the complexity of their movements genuinely different and more monolithic at the Neolithic horizon than at later times in prehistory?

Tidings of technologically advanced urban societies may in the course of prehistory have provided ample motivation for migration, with enticing prospects of plunder and material advancement. We must consider such alternatives especially in those cases where the linguistic picture suggests a radically different view of prehistory than does the spread of material culture as reflected in the known archaeological record. The introduction of Proto-Sinitic, a branch of Tibeto-Burman, into the Yellow River basin is a case in point. This theory, which I shall call the Centripetal Migration model, is diametrically opposed to the centrifugal Farming-Language Dispersal theory. The Centripetal Migration model may also apply to portions of Austroasiatic prehistory.

⁶ Today Afroasiatic languages are spoken throughout this area, but none are descended directly from the extinct branch which Akkadian represents.

More crucially, an essential trait of the Centripetal Migration theory is that this model assumes that migrations in prehistory could have unfolded both in centrifugal and centripetal directions with respect to centres of technologically advanced and later urban civilisations. The motives for migrations were no doubt diverse, and no single model, such as the Farming-Language Dispersal theory, can account for all demographic developments and linguistic intrusions, even across the Neolithic horizon. Even the chief proponents of the Farming-Language Dispersal theory do not entertain the idea that all languages were spread by early farmers, e.g. Bellwood (2005). At the same time, we must also not lose sight of the fact that vast tracts of the Himalayas, Burma, northeastern India and neighbouring southwestern China remain archaeologically under-explored or unexplored.

DIAGRAM 7: The portion of the Y chromosome phylogenetic tree relevant to the Father Tongue hypothesis with regard to Austroasiatic, provided by Mark Jobling and Emma Parkin.



In conclusion, groundbreaking research in population genetics has begun to suggest that the geographical distribution of Austroasiatic may be connected to a well-defined Y chromosomal haplogroup. The Father Tongue hypothesis may also apply to Austroasiatic, either wholly or in part, on the basis of the population genetic studies completed to date. The veracity of the Father Tongue hypothesis is the inherent underlying assumption when geneticists propose that a particular Y haplogroup, say O2a, corresponds to the geographical spread of a particular language family, such as Austro-

asiatic. Diagram 7 illustrates the portion of the Y chromosomal phylogeny thought to be relevant to Austroasiatic.

However, Austroasiatic is an old language family, and we would expect the population history of this family to be at least as complex as that of Tibeto-Burman, if not more so. Careful correlation of linguistic and population genetic findings may enable us to reconstruct early language contact situations and ancient cases of language shift and linguistic intrusions that might account for the phenotypical difference between Munda speakers and Khasi-Khmuic and Mon-Khmer language communities as well as between the Aslian negrito populations, Aslian non-negrito populations and the Nicobarese. The Father Tongue hypothesis may not apply in all cases for the biological ancestry of all Austroasiatic language communities, just as language spreading solely via the paternal line cannot account for the linguistic identity of all Tibeto-Burman populations, e.g. maternal Balti vs. paternal Hân.

Although Sahoo *et al.* (2006) clearly favour a Southeast Asian homeland for Austroasiatic, their findings cannot yet conclusively establish that Southeast Asia is the point of origin for the O2a haplogroup. The exciting hypothesis that the O2a haplogroup may correlate with linguistic spread of Austroasiatic also remains to be demonstrated in convincing detail. A fine-mesh genetic sampling of all Austroasiatic populations — not just the most populous, national majority or prestige groups — will be required and the topology of the haplogroups in question will have to be determined in order to ascertain which precise area could be the probable point of origin of polymorphic genomic markers which could be correlated with the linguistic spread of Austroasiatic. Furthermore, the detailed geography of the entire O branch of Y chromosomal haplogroups has yet to be reconstructed at a satisfactorily high resolution. I call upon all interested parties to join forces and help us in this endeavour.

4. TIBETO-BURMAN: LINGUISTIC AND BIOLOGICAL ANCESTORS

What do genetic studies tell us about the spread of Tibeto-Burman? Pioneering work in the 1990s found the genetic distance between Mandarin speakers in the north and Tibetans to be far less than between southern Hân Chinese and Mandarin speakers (Cavalli-Sforza, Menozzi and Piazza 1994: 225), even though southern Hân populations such as the Cantonese and Min speak Sinitic languages. The genetic discrepancy between southern Hân and northern Hân then already appeared to corroborate what we knew about the

history of China, particularly with respect to Hàn linguistic and cultural expansion.

The Qin launched a brutal campaign to subdue the 'one Hundred 越 Yue' tribes of southern China in 221 BC, but resistance by indigenous population groups persisted fiercely, and Qin control over these areas was lost after the death of the first Qin emperor in 210 BC. The Hàn dynasties were able to consolidate Qin territorial gains and even expand further. In the south, the newly consolidated Sinitic state underwent territorial expansion into the eastern half of Yúnnán overthrowing the 滇 Diān kingdom in 109 BC, then subduing the region of 嶺南 Lǐngnán in 111-112 BC, an area comprising modern Guǎngxī and Guǎngdōng provinces, Hǎinán island and what today is northern Vietnam. Mountainous Fújiàn only became sinified much later, during the period of the Three Kingdoms in the aftermath of the Wú state's invasion of the southeast ca. 260 AD.

More recently, a population genetic study of 23 Hàn populations (Wen *et al.* 2004a) has further corroborated the picture which linguists and historians had of a martial and therefore male-biased Hàn expansion southward during the sinification of what today is southern China. Southern and northern Hàn populations were found to share roughly the same mean frequency of around 54% for the Y chromosomal haplogroups O3-M122 and O3e-M134, both characterised by the M122-C mutation. On the other hand, southern Hàn were found to have a higher frequency than northern Hàn, viz. 19% vs. 5%, for the mutation M119-C, characterising Y chromosomal haplogroups O1* and O1b, and the mutation M95-T, typifying haplogroups O2a* and O2a1. These haplogroups are known to be frequent in Daic, Austroasiatic and Hmong-Mien populations south of the Yangtze.

Moreover, southern Hàn were found to have an average frequency of 4% for the haplogroups O1b-M110, O2a1-M88 and O3d-M7, likewise frequent in pre-Sinitic populations south of the Yangtze, whereas these haplogroups were not found in northern Hàn. By contrast, the maternal lineages of southern Hàn showed an overall frequency of 36% for the mitochondrial haplotypes A, C, D, G, M8a, Y and Z, typically widespread in northern East Asia, as opposed to an overall frequency of 55% in northern Hàn. Mitochondrial lineages predominant in Daic, Austroasiatic and Hmong-Mien populations south of the Yangtze, i.e. haplotypes B, F, R9a, R9b and N9a, were found in a frequency of 55% in southern Hàn as opposed to 33% in northern Hàn.

In short, the southern Hàn paternal lineage shows preponderant northern Hàn penetration alongside a faint pre-Sinitic signature. Males from the north were the primary contributor to the paternal gene pool of southern Hàn populations, whereas the mitochondrial DNA of southern Hàn populations con-

tains roughly equal contributions from pre-Sinitic and Hàn maternal ancestors. The Father Tongue hypothesis appears to apply for Sinitic in the form of the Hàn demic expansion, at least on the basis of population genetic studies completed to date. Although there must be numerous contrary cases such as the Tibetan mother tongues of Baltistan, as a general principle the Father tongue hypothesis may at many times and in many places in prehistory have been an important mechanism in language shift.

The dynamics of a process whereby mothers passed on the language of their spouses to their offspring has major implications for our understanding of language change. If the language shift giving rise to the Sinitic languages and perhaps also the eastward spread of Indo-Aryan speech across northern India took place in this way, then could languages in some cases in the long course of prehistory have begun as languages belonging to another phylum until they reached the stage currently attained by Michif?

In origin at least, Michif is genetically an Algonquian language that was spoken by women who relexified the language with the French spoken by their husbands to such an extent that the genetic affinity has nearly been obscured (Bakker 1992, 1994, van Driem 2001: 169-173). If the process of relexification were to continue beyond the stage attained by Michif, then a language could conceivably change its genetic affinity even though the dynamics of the process would introduce a discontinuity with its past. Can such a process ever be reconstructed linguistically? A recent study of Chinese dialects indicates that the diversification of Sinitic languages did not proceed in a tree-like fashion (ben Hamed and Wāng 2006).

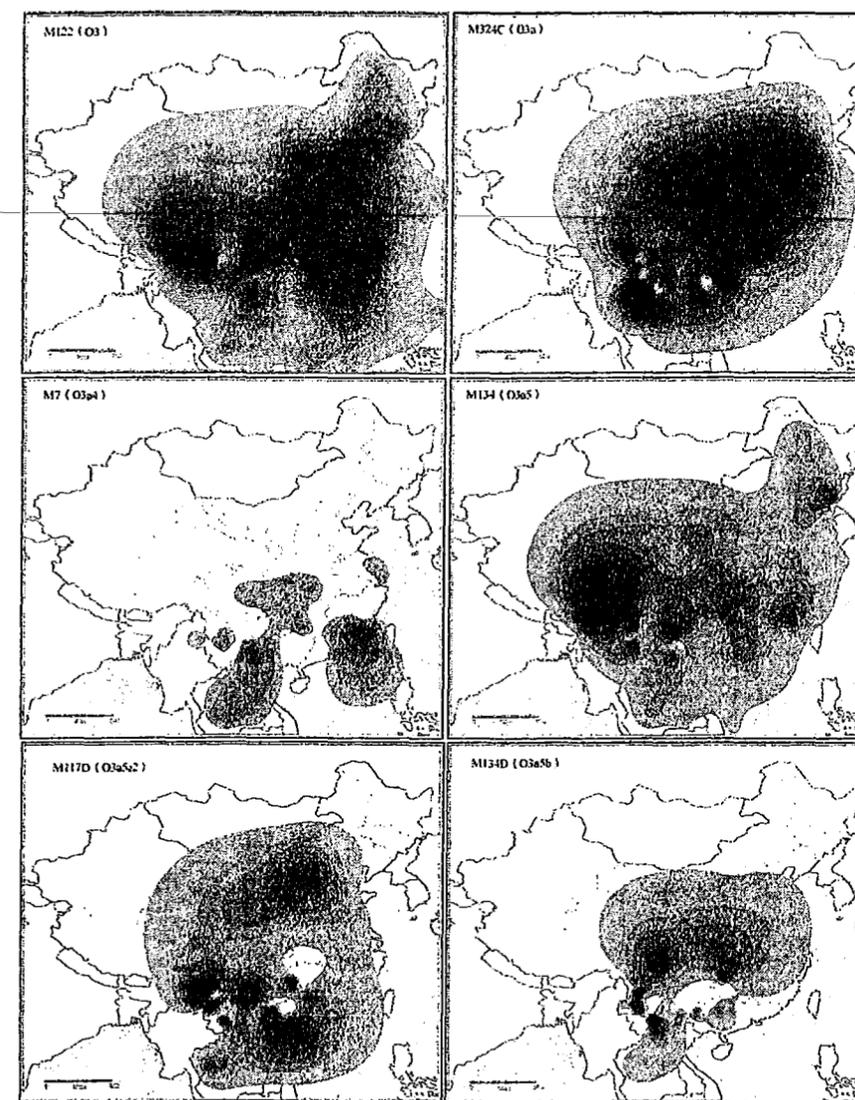
At a deeper time depth, what can we say about the origin of the Sinitic branch as such? Genes do not tell us which linguistic intrusions took place in prehistory. For this linguistic geography is a better indicator. Population genetics tells us about the spread of genotypes, whether this is caused by circumstances of origin, migration or natural selection. Geneticists have looked for markers which identify Hungarians as a Uralic language community and failed to come up with much. Even the Y chromosomal haplogroup N-TatC (N43), which is found at a high frequency throughout all Uralic language communities, does not seem to be prevalent in Hungary. Outside of Uralic speech communities, the haplogroup is also found at a high frequency amongst the Yakut, Even and Tuva.

Rather, Hungarians look genetically quite a lot like a Western Slavic language community, and there is very little trace at the moment of a Uralic genetic signature (Tambets *et al.* 2001). Perhaps the early Magyars who penetrated Pannonia introduced a Uralic language but not much else. Perhaps Uralic Y chromosome lineages died out in Hungary for some reason.

Whatever the case may be, the Hungarian language constitutes incontrovertible linguistic evidence that the Magyars did come to Pannonia. The historically attested Magyar linguistic intrusion may be genetically invisible, but the Hungarian language is linguistically palpably manifest. Given the extremely low population numbers which characterised prehistoric human demography, it stands to reason that no colossal throng of people was needed to effectuate a linguistic incursion.

By the same token, let us keep in mind that the linguistic ancestors of the Chinese were Tibeto-Burman, but there is no *a priori* reason for assuming that the biological ancestors of the Hân Chinese derived predominantly from ancient Tibeto-Burman speech communities. Earlier studies have been interpreted to indicate movements in all directions. However, work by our own team on the Y chromosome indicates that the linguistic ancestors of the Hân Chinese and at least some portion of Hân biological ancestry in the paternal line were the same people. Moreover, genetic studies do not reveal a simple picture of our past, but a multi-layered pattern of movements in different directions at different time depths, and sometimes these migrations are characterised by a certain sexual dimorphism or gender bias, whereby women quite often get left at home.

DIAGRAM 8 next page: Contour maps showing the geographical distribution of Y haplotype frequency, reproduced here from Shi (2005: 414), based on assays which were not yet able to include most Tibeto-Burman language communities. Haplogroup labels in their diagrams deviate from the conventional 2003 and 2005 nomenclatures of the Y Chromosome Consortium. What is called 'M7 (O3a4)' here is haplogroup M7 (O3d) in the 2003 Y Chromosome Consortium nomenclature or M7 (O3c) in the as of yet insufficiently verified 2005 Y Chromosome Consortium tree. What in this diagram is labelled as haplogroup 'M134 (O3a5)' is haplogroup M134 (O3e) in the nomenclature of the 2003 Y chromosome tree with its smaller and more reliably documented set of single nucleotide polymorphisms, or M134 (O3d) in 2005 Y Chromosome Consortium nomenclature. Both 'M117D (O3a5a2)' and 'M134D (O3a5b)' are subgroups of haplogroup M134 (O3e) or, in 2005 nomenclature, of haplogroup M134 (O3d). What in this diagram is labelled as 'M122 (O3)' is indeed M122 (O3), and the mutation 'M324 (O3a)' is purported to be a single nucleotide polymorphism (SNP) that characterises all lineages derived from M122 (O3), i.e. O3a through O3e (or O3a through O3d), except O3*.



The reduced polymorphism of northern populations of East Asia, which represent a subset of the haplotypes found in southern populations, was taken to reflect the peopling of the north after the Ice Age (Sù *et al.* 1999), whereas the extremely high frequency of H8, a haplotype derived from M122C, was seen as reflecting a genetic bottleneck effect that occurred during an ancient southwesterly migration about 10,000 years ago, suggesting a demic diffusion at the outset of the Neolithic (Sù *et al.* 2000, Ding *et al.* 2000, Shi *et al.* 2005). Another study suggested that Hân Chinese did not originate in the Yellow River basin but had more recently migrated to this area from southwestern China (Chǔ *et al.* 1998). Comparison of haplogroup frequencies exhibited by Tibetans vs. Tǔjiā, Bái and Lolo-Burmese groups showed all Tibeto-Burman groups to have a high frequency of the Y-chromosomal haplogroups O3e and O3*, with the average hovering approximately around 40%. The findings were interpreted as supporting a slightly male-biased infiltration from the Bodish area in Amdo into Yúnnán and Húnán about two and a half millennia ago, though 'the less drastic bias between male and female lineages' suggested that these putative southward migrations 'likely occurred with the involvement of both sexes rather than as conquests involving expedition forces primarily consisting of male soldiers' (Wen *et al.* 2004b).

These pioneering genetic studies are highly insightful, but they are limited by the fact that most Tibeto-Burman language communities and even most branches of the language family are exclusively represented outside of China. The picture of the Tibeto-Burman past has been rendered far more complete by findings of our research team, which has conducted the most extensive sampling of Tibeto-Burman populations in the Himalayan region (Kraaijenbrink *et al.* 2006a, 2006b, Parkin *et al.* 2006a, 2006b, Tyler-Smith *et al.* 2006).⁷ Tibeto-Burman language communities in northeastern India

⁷ Our findings are not contradicted by other recent genetic studies on populations of East and Central Asia (Hu *et al.* 2005a, Hu *et al.* 2005b, Hu *et al.* 2005c, Hu *et al.* 2005d, Liang *et al.* 2005, Lin *et al.* 2005, Xue *et al.* 2005, Yang *et al.* 2005, Zhu *et al.* 2005, Chen *et al.* 2006, Kang and Li 2006, Shi *et al.* 2006, Zhang *et al.* 2006) or relevant portions of the Indian subcontinent (Ashma and Kashyap 2002, Sahoo and Kashyap 2002, Kashyap *et al.* 2002a, Kashyap *et al.* 2002b, Maity *et al.* 2003, Kumar *et al.* 2004, Langstich *et al.* 2004, Thomas *et al.* 2004, Banerjee *et al.* 2005a, Banerjee *et al.* 2005b, Debnath and Chaudhuri 2005, Krithika *et al.* 2005, Watkins *et al.* 2005, Singh *et al.* 2006). DNA testing on probably Tibeto-Burman mummies in the Kālī Gaṇḍakī valley has yet to test for the relevant markers (Alt *et al.* 2003).

and southwestern China have yet to be sampled in as rigorous and fine-mesh a fashion as has been done in Nepal and Bhutan.

If we were to assume the veracity of the Father Tongue hypothesis for Tibeto-Burman in general with the exception of cases such as Baltistan, then our team has identified a Y-chromosomal haplogroup that may be specifically correlated to the early spread of Tibeto-Burman language communities. Frequency gradient maps for the relevant haplogroup would be somewhat indicative of a possible location for the Tibeto-Burman homeland. More instructive, however, would be the identification of a precise geographical locus, if one can be said to exist, for the root of the topology of the relevant haplogroup. Our results will be published in due course in an appropriate population genetics journal, and I am not at liberty to detail the findings here. However, suffice it to say that one highly plausible interpretation of these findings would be commensurate with one of the scenarios outlined in the following section of this paper.

Far away to the south, in the Brahmaputran basin and the Indo-Burmese borderlands, however, some of the spread of Tibeto-Burman may have been at the expense of indigenous Austroasiatic populations who were assimilated linguistically. The Y haplogroup O2a is represented at a frequency of 77% in Austroasiatic groups in India and 47% in Tibeto-Burman groups of northeastern India (Sahoo *et al.* 2006). This patterning could suggest that Tibeto-Burman paternal lineages may have partially replaced indigenous Austroasiatic lineages in the northeast of the Indian Subcontinent and that Austroasiatic populations preceded the Tibeto-Burmans in this area, as linguists and ethnographers have speculated for over a century and a half.

5. TIBETO-BURMAN: LINGUISTIC ANCESTORS AND MATERIAL CULTURE

Linguistic palaeontology has begun to suggest that the early speakers of Tibeto-Burman languages, or a subset thereof, were already agriculturalists as well as hunters. On the other hand, the Limbu, Lohorung, Dumi and other Kiranti groups in the eastern Himalayas retain lore whereby their ancestors once only practised hunting and gathering and then one day became cultivators. The transition to a sedentary agricultural lifestyle no doubt occurred in the hoary past, yet the memory of this episode is kept alive as if it were a recent historical event. Could the Kiranti ancestors have been farmers who were forced by circumstances at some point to revert to a hunter-gatherer existence, only for their descendants in some later period to return to sedentary agriculturalism? The antiquity of oral traditions is difficult to ascertain, yet millets must have played a key role in Tibeto-Burman culture for a long

time, as attested by reflexes for *Setaria italica* in languages as far flung as Old Chinese 稷 *btsik* in the Yellow River basin and Lhokpu *cá'kto* 'foxtail millet' in modern southwestern Bhutan.⁸

Both foxtail millet *Setaria italica* and broomcorn millet *Panicum miliaceum* and were staples in what today is northern China, where they are first found to occur in the Péiligǎng culture (6200-5000 BC). No archaeological sequence provides evidence for their prior domestication, and neither northern China nor Korea have yet yielded any archaeological data on subsistence for the period between 10000 and 6500 BC (Crawford 2006: 80-81, 91), even though by far 'most archaeological fieldwork has taken place in the eastern half of China' (Underhill and Habu 2006). Domesticated foxtail millet derives from green foxtail millet, i.e. *Setaria italica*, subsp. *viridis*. Broomcorn millet is known to grow throughout Eurasia as a weed, and the wild form has been denominated subspecies *rudérale*. The early Neolithic in northern China is therefore in effect defined by the appearance of ceramic communities, although the appearance of ceramic communities in Korea and Japan are conventionally not interpreted as representing agricultural communities (Underhill and Habu 2006).

For Kiranti groups of eastern Nepal no sacred ritual can be performed without millet beer and distilled millet spirits. This applies particularly to ceremonies to commemorate and revere the ancestors, at which millet beer and millet brandy are indispensable. In Nepal, *Setaria* and *Panicum* have in many areas been replaced with finger millet *Eleusine coracana*, a crop ultimately of African provenance. Yet in parts of Nepal as well as in Bhutan, *Panicum*, *Setaria* and other millets are still widely cultivated, though these crops are on the decline due to our headlong global rush towards 'improved' monocultures. Amongst the Gongduk community in Bhutan, for example, broomcorn and foxtail millet are prized as the staples sacred to the tribal ancestors.

In the Himalayas, groups which have undergone either strong Aryan religious and Hindu cultural influence or the influence of Buddhism emanating from the Tibetan plateau preserve less faithfully the Tibeto-Burman cultural

⁸ The Lhokpu are an inbred and genetically highly distinct group within the Himalayan region as a whole (Kraaijenbrink *et al.* 2006a, Parkin *et al.* 2006a). The impact of matrilocality and cross-cousin endogamy is clearly discernible in the genetic signature of this language community. Many of the ancient Tibeto-Burman groups may have been matrilineal, matrilocal societies with uxori-local marriage such as the modern Lhokpu and Gongduk of Bhutan.

heritage retained amongst groups like the Kiranti in eastern Nepal and the Gongduk of eastern Bhutan. Wherever the older stratum of shamanism and sacrifice has been retained, the role of millet beer and millet brandy takes centre stage. I cannot help but look with Kiranti eyes at the plethora of elaborate bronze ritual vessels for beer and distilled spirits which appear in the Shāng and Zhōu period. These diverse ornate liquor vessels have been labelled by archaeologists variously as 觚 *gū* 'beaker', 壺 *hú* 'liquor vessel', 尊 *zūn* 'liquor vessel', 觥 *gōng* 'goblet', 觥 *gōng* 'animal-shaped liquor receptacle', 杯 *bēi* 'beer bowl', 卣 *yǒu* 'spiced millet liquor vessel', 罍 *léi* 'vessel for libations in honour of the ancestors', 盃 *huì* 'vessel for mixing liquor', 彝 *yí* 'large liquor container', 罍 *léi* 'liquor receptacle', 罍 *léi* 'liquor receptacle, modified from the *léi*', 角 *jué* 'decanter' and 爵 *jué* 'decanter'. These receptacles were used for storing, blending, serving beer and spirits brewed from the millets *Setaria* and *Panicum*, sacred to the ancestors of the Kiranti, the Gongduk and the Chinese.

So, were broomcorn millet and foxtail millet first cultivated in what today is northern China, where evidence of their domestication appears as early as 6200 BC, or were they first domesticated somewhere in the expanse of territory between Shānxī and the eastern Himalayas, where these crops are still cultivated by indigenous Tibeto-Burman peoples today? Are the Lhokpu descendants of early agricultural colonists from the Yellow River basin who forged their way across the Tibetan plateau, over the towering Himalayas and down its southern flanks into the dense malarious jungles on the western duars in search of arable land? Or did the linguistic ancestors of the ancient Chinese migrate up from the jungles of the Brahmaputran plain across the white peaks of the Himalayas to make a long trek to what is now the North China plain in search of fertile fluvial plains far away?

Before we cast our inquiry in such a mould, we must ask what the first domestication of crops can tell us about the spread of language families. The Neolithic spans a vast stretch of time, and this long period was no doubt not characterised by demographic stasis. The linguistically reconstructible past has a shallower time depth than the prehistory of human habitation in the region. Ancient humans inhabited at least the foothills of the Himalayas,⁹

⁹ The Himalayas are the most prominent barrier along the Movius line, beyond which *Homo erectus* populations colonising eastern Asia either lost or abandoned their advanced Acheulian stone knapping technologies. Based on an archaeological survey in Dāng Deukhūrī in western Nepal and in the foothills of the eastern Terai along the Rāto Kholā, Corvinius concluded that palaeolithic 'sites are rare and that hand-axe makers were not frequent

and anatomically modern humans later inhabited even the Tibetan plateau in palaeolithic times (Zhang *et al.* 2003, Madsen *et al.* 2006). Yet the palaeolithic of the Himalayan region, including the Tibetan plateau, remains largely unexplored and unknown. As for the early Neolithic inhabitants, the very first cultivators may not have left any linguistic descendants at all. This point was made clear for the Bronze Age Near East in the previous section.

There is no reason to think that prehistoric events did not transpire in a parallel fashion in the Yellow River basin. A reasoned correlation of the archaeological record with the reconstructible linguistic past and the complex picture emerging from population genetic studies may help us reconstruct some of what actually happened. Archaeology, comparative linguistics and population genetics give us three different versions of prehistory, and in the handbook (van Driem 2001), I argued for keeping these three different versions of prehistory distinct. In a similar vein, Karafet *et al.* (2001) argue for a 'multilayered, multidirectional and multidisciplinary framework' and insist that 'more realistic models for the underlying processes leading to the modern population structure of East Asia will have to accommodate more complex multidirectional biological and — especially — cultural influences than earlier explanatory paradigms' (2001: 626).

Cultural traits, crops and the names for crops could have come along with a community of speakers but are also known to diffuse back and forth across language boundaries or to be adopted by newcomers to an area from an older resident population. So this view varies fundamentally from a programme that seeks to see genes and languages spreading monolithically in tandem with Neolithic agriculture as attested in the archaeological record. What archaeology tells us is the prehistory of material culture, which may often be a reflection not of population movements but of socioeconomic discrepancies which drove ancient peoples to migrate towards the centres of affluence which lured them with the promise of a better life. The distribution of major Tibeto-Burman subgroups mapped in Diagram 3 suggests the tracks of a northeasterly migration from the Tibeto-Burman core area in the fertile hills and river valleys of Sichuān and the eastern Himalayas to the

occupants of the valleys' (1996: 48), corroborating Pandey's earlier survey (1987). Gaillard reported that Acheulian industries are rare throughout the Śivālik hills north of the Gangetic basin (1996). Yet Acheulian bifacials have been found at Chinese sites as far flung as Lán-tián, Bósè, Zhōukōudiàn and Dīngcūn, which underscores that the Movius line was a permeable barrier. It is possible that stone tools were largely replaced by tools of bamboo, rattan or some other more perishable material to the east of the Movius line.

loess plains of northern China by an ancient group that was linguistically ancestral to the Chinese.

Different scenarios have been proposed to account for the modern geographical distribution of Tibeto-Burman language communities. Here I shall discuss three such possible versions of prehistory, which may be numbered Scenario 1, 2 and 3. Scenario 2 exists in several versions, which we may call Scenario 2a, 2b and 2c. Over the past years, I have argued the case for Scenarios 1, 2b and 2c. Scenario 2a was first implied by Paul Benedict, and Scenario 3 is a model of population prehistory proposed for Tibeto-Burman by Peter Bellwood.

Scenario 1 envisages Proto-Tibeto-Burman originating in what today is Sichuān province, whence early Tibeto-Burmans spread to the southwest onto the Brahmaputran plain, introducing themselves and the Eastern Indian Neolithic culture to resident Austroasiatic populations. Another group, which we might call Proto-Sino-Bodic, moved to the northeast seeding the Péilǐgǎng (6500-5800 BC), Císhān (6000-5600 BC) and Dàdìwān (6500 to 5200 BC) Neolithic cultures along the Yellow River. Other groups remained in Sichuān and spread across the fertile hills of Yúnnán province in the south. The Mǎjiāyáo Neolithic (3900-1700 BC) succeeded the Dàdìwān culture in eastern Gānsù and adjacent parts of Qīnghǎi and Níngxià. Sinitic remained in the east and can be associated with the Yǎngsháo culture (5500-2700 BC), which succeeded the Péilǐgǎng and Císhān cultures on the North China plain, whereas the expansion of Bodic into the Himalayas is associated with the sudden appearance of colonial exponents of the Mǎjiāyáo Neolithic in eastern Tibet at mKhar-ro and in Kashmir at Burzahom at the same time that the core area in Gānsù shrank during a period of climate change between the Mǎjiāyáo phase (2700-2300 BC) and the Bànsān phase (2200-1900 BC) of the Mǎjiāyáo sequence. This, in a nutshell, is the scenario which I outlined in several previous publications (van Driem 1998, 2001, 2002).¹⁰

¹⁰ In their archaeological discussion of the Sichuān homeland hypotheses, Aldenderfer and Zhāng 'agree with van Driem that Sichuan is a likely source for a Neolithic package' which gave rise to cultures on the Yellow River (2004: 39). Yet Aldenderfer and Zhāng (2004: 37) appear to think that I do not include the mKhar-ro site near Chab-mdo or any other Tibetan archaeological sites in my model. The Tibetan archaeological site mKhar-ro or mKhar-chu, which I discuss at length (van Driem 2001: 430-431), is sinicised in the Chinese archaeological literature with characters that are correctly romanised as Kǎruò, and which Aldenderfer and Zhāng incorrectly transcribe as 'Karou'. Sites should be named properly in accordance with archaeological convention. Their misunderstanding again provides

Scenario 2, discussed as an alternative proposal in the same publications, plays out at an utterly different time depth. This alternative view does not see the ancient Tibeto-Burmans as the people who seeded the early Neolithic cultures seen at the Péilígǎng, Císhān and Dàdìwān sites. Rather, the ancient Tibeto-Burmans emerged from the linguistic core area, drawn by the riches of the affluent Yellow River basin and introduced themselves and their language only in the late Neolithic or Bronze Age. The point of departure in this scenario again is Sìchuān.

The version of this model which we shall call Scenario 2a develops a proposal first put forward by Benedict that the Shāng may not yet have been Sinitic at all. Instead, the Zhōu, who came from the west, were the bearers of the Proto-Sinitic language who 'became fused with, or perhaps immersed in' the pre-Tibeto-Burman language spoken by the Shāng (1972: 197). My own variations on this theme are Scenario 2b, which envisages that the prosperous agricultural civilisation in the Yellow River basin may have lured the linguistic forebears of Sinitic, or perhaps even Sino-Bodic, before the Shāng period, and alternatively Scenario 2c, whereby Tibeto-Burman could have been introduced or re-introduced to the Yellow River basin more than once in the course of prehistory. Each version of Scenario 2 presumed that tidings of the technologically advanced societies already in place throughout the Yellow River basin would have provided ample motivation for the move, with enticing prospects of plunder, riches and material advancement.

There are possible archaeological correlates for the Bronze Age linguistic intrusion proposed by Scenario 2. As compared to eastern China, the vast southwestern region has not received nearly as much attention from archaeologists. Fortunately, some progress has been made since Zhāng Guāngzhí (1977, 1986) lamented the lack of fieldwork in Sìchuān. At the same time,

the context for my assertion that: 'Numerous artificial problems in Tibetan toponymy and cartography currently result from the practice of listing only the sinified version of Tibetan place names in Hànyǔ Pīnyīn romanisation without providing the actual place names' (loc. cit.). Incorrect Hànyǔ Pīnyīn transcriptions merely exacerbate the problem. Aldenderfer and Zhāng identify mKhar-ro or Kāruò as a colonial exponent of the Mǎjiāyáo Neolithic in Gānsu, but their cursory familiarity with the literature leads them to think that they are the first to do so. In fact, a good number of Chinese archaeologists (e.g. Xīzàng etc. 1979, Ān 1992) had already identified mKhar-ro or Kāruò as a colonial exponent of the Mǎjiāyáo Neolithic, and my model followed this consensus. Aldenderfer and Zhāng do not differentiate between language spread by demic diffusion and language intrusion by colonial migration, and they inexplicably attempt to interpret 'Karou' as the result of demic diffusion from Sìchuān.

the grand scale on which the earth is being ripped apart in many parts of Sìchuān, including even the Mínjiāng river valley, for highway networks, dams and large industrial projects may already have obliterated a great deal of potential archaeological sites, especially along rivers and at many of the best sites for ancient human habitation. When archaeological fieldwork is conducted in the region, excavations unearth spectacular new sites such as the major but previously unknown Bronze Age civilisation at Sānxīngdū, which only fully came to light in 1986, and the discovery in recent years of the earlier Neolithic civilisation along the precipitous upper reaches of the Mínjiāng river, a tributary of the Yangtze. In 茂縣 Màoxiàn county, the 波西 Bōxī (4000 BC), 營盤山 Yíngpánshān (3500-3000 BC) and 沙烏都 Shāwūdū cultures (2500 BC), situated on the largest pieces of fertile flat land along the Mínjiāng river, on the way from Chéngdū to gZi-rtsa-sde-dgu,¹¹ have been identified as possible antecedents of the Sānxīngdū culture, located 40 km northeast of Chéngdū.

Sānxīngdū has been associated with the ancient 蜀 Shǔ polity. In terms of chronology, the earliest period of habitation, Period I, is the 寶墩 Bǎodūn phase, which lasted from 2800 to 2000 BC and is contemporary with the Shāwūdū culture upstream in the Mínjiāng river valley. The spectacular Bronze Age culture at Sānxīngdū is represented by Periods II and III, which together lasted from 2000 to 1200 BC. The apogee of the Sānxīngdū culture is therefore contemporaneous with and somewhat precedes the Shāng period (1700-1100 BC) at Ānyáng. The later 巴蜀 Bāshǔ period (1200-800 BC) at Sānxīngdū is contemporary with the Western Zhōu (1100-771 BC) centred at Hào near Xī'ān. Dragons and physiognomic motifs on some of the bronze 鏡 nǎo 'ritual bells' and 盤 pán 'basins, dishes' of the late Shāng period at Ānyáng are in fact stylistically reminiscent of earlier Sānxīngdū iconography in Sìchuān.

The striking imagery of the Sānxīngdū culture has led archaeologists to speculate that the society was theocratic in nature, with sacrifice playing a central role. Brewing beer and distilling alcohol were evidently of pivotal cultural importance. In addition to elaborate bronze cooking vessels, musical instruments and a variety of water containers, the Sānxīngdū people also had a variegated repertoire of ritual vessels for beer and distilled spirits just like those of the Shāng and Zhōu further east. Archaeological speculation about the ritual importance of alcohol and of blood sacrifice to haunting

¹¹ gZi-rtsa-sde-dgu [zətsazdergə] is the local place name. The official Tibetan name is gYu-tsha-sde-gu [ɣ jəts'azdergə], and the Mandarin name is Jiūzhàngōu.

goblin-like deities is reminiscent of the ritual importance of alcohol in many Tibeto-Burman cultures and of the blood sacrifice practised by the Kiranti and other Himalayan groups, which in olden days even entailed human sacrifice, as recorded in the Dumi grammar.

Sānxīngdū has yielded numerous ornate bronze ornaments and tools, but far more daunting is the vast arsenal of well honed jade weaponry, such as 璋 zhāng 'axes' and 戈 gē 'dagger axes', adzes, blades, swords and spear points in addition to numerous jade chisels and other lithic tools. Rather than a peaceful demic diffusion, the expansion of Tibeto-Burman into Shānxī may very well have been a military affair. The martial campaigns which heralded the Tibeto-Burman linguistic intrusion into northeastern China may have left no more testimony in the archaeological record than did the successive invasions of Gutæans, Amorites and Kassites into the Fertile Crescent. Rather the principal, telling legacy of this intrusion today is the Sinitic branch of languages. As Máo's Red Army demonstrated in 1935-36, the rugged mountain strongholds of Sìchuān and the eastern perimeter of the Tibetan plateau are a strategic area from which to launch a military campaign into Shānxī. In their case, however, this area also served as a place of refuge for which many troops were ill prepared, with fatal consequences.

Even if Sinitic were only introduced to Shānxī as late as the Zhōu period, as Benedict proposed, then the turbulent maelstrom of cultural changes and military conflicts which have characterised Hàn expansion from the second century BC to the present day as well as the succession of distinct prestige vernaculars emanating from shifting capitals in the course of over two millennia are more than adequate to account for the aberrant appearance of modern Chinese dialects when compared to reconstructible Old Chinese. A language spoken in the thick of things incurs change more rapidly than languages sheltering in the undisturbed periphery. Once an ancient variety of Tibeto-Burman speech had been introduced into the political centre of what was to remain the most powerful polity in East Asia, the language would change more quickly than those varieties of Tibeto-Burman spoken in less easily traversable terrain.

Benedict's and my versions of Scenario 2 have different implications for the nature of the Shāng script. For Benedict, the Shāng spoke a pre-Sinitic language, whilst my versions, 2b and 2c, do not exclude the possibility that the Shāng script might already represent an early Sinitic language. In favour of Benedict's view, it can be pointed out that only half of the nearly five thousand Shāng period characters have been deciphered with certainty, and the extant corpus consists entirely of highly abbreviated divinatory fragments. No critical study has been undertaken to ascertain precisely what per-

centage of Shāng characters consist of a phonetic and a semantic component, and how the phonetic components in composite Shāng characters compare with those in Zhōu characters with the aim of testing the hypothesis of a possible language shift between the two periods, whereby the script was adopted by an early Sinitic population from a non-Sinitic one. Indeed, undertaking any such study of the Shāng character corpus with the aim of assessing this hypothesis dispassionately would already be flying in the face of orthodoxy.

Tangential to Benedict's hypothesis is the question of the origin of the Shāng script itself. Nativists such as Keightley oppose the idea of a foreign inspiration for the earliest Sinitic script because of the lack of similarity between Shāng oracle bone inscriptions and 'Sumerian, Egyptian or Hittite' writing (2006: 177). Few would take issue with the lack of similarity between cuneiform, Egyptian hieroglyphs and the Shāng characters that appeared in the second millennium BC. However, the old theory that the Shāng writing system was inspired from a foreign model does not look to cuneiform or hieroglyphic writing, but logically looks to the two earlier writing systems that were closest to the Yellow River basin both in time and in space, i.e. the Indus and Proto-Elamite pictographic scripts.

These two logosyllabic writing systems could have travelled eastward via the same, then already ancient trade route as did contemporary Bronze Age technologies. The Indus and Proto-Elamite scripts are not only structurally similar¹² to the early Shāng writing system but also similar in terms of individual graphemes, as I have illustrated previously (2001: 355-358). Is it mere coincidence that the Western Zhōu ideograms show great resemblance to contemporaneous Late Bactrian glyphs, whilst the earlier Shāng script more closely resembles its nearest precursors, Indus and Proto-Elamite writing? Or are such differences in style and parallel developments in style to be explained away merely as a function of the difference in medium involving the transition from scapulae and plastrons to bronzes in which shapes could be carefully fashioned in the malleable clay of the moulds?

Nativists look for precursors to the Shāng script in the decorative glyphs found on local ceramics, whilst ignoring likely Central Asian antecedents.

¹² Whilst modern Chinese writing is ideogrammatic in that it consists of characters or ideograms representing morphemes, the Shāng writing system is widely held to have been logographic, whereby each character represented a word. I shall not entertain the theory of Vandermeersch (1980) and Hansen (1993) that Shang writing was ideographic in the sense of representing ideas or things directly rather than representing language.

Followers of this line of inquiry should at least include the 巴蜀 Bāshǔ pictographs on Sānxīngduī pottery in their deliberations. Some have ventured to speculate on the erstwhile existence of texts of a more elaborate nature on perishable materials during the Shāng period, not one of which has survived. If such speculation is warranted, then how much more probable is it that specimens of Indus and Proto-Elamite writing on perishable materials could have made their way along the main eastbound trade artery to the Yellow River basin by the sixteenth century BC and inspired the writing system of the Shāng in the first place? Or could the idea of script have travelled via Sānxīngduī itself, where hoards of tusks, cowrie shells and other objects likewise attest to long-distance trade?

More fundamentally, the search for precursors of the Shāng script in the decorative motifs on pottery reminds us that semasiography, viz. communication by pictorial or symbolic representation, was already a finely developed art in the Upper Palaeolithic. Franco-Cantabrian glyphs which appeared between 60,000 and 40,000 years ago, some spectacular specimens of which are kept at the Museum of Natural History in Brussels, resemble symbolic writing systems far more than do the decorations on East Asian ceramics of the fourth and third millennium BC. Glottography, viz. visual representation of spoken language, is attested from 3200 BC in Sumer, and recent finds at Abydos by Günter Dreyer's team may now push back the date for glottography to 3400 BC and its earliest attestation west to Egypt. Subsequently a plethora of writing systems had evolved in West Asia and the eastern Mediterranean before the Shāng writing system appeared nineteen centuries later in the sixteenth century BC.

Yet even if we envisage the Shāng as speakers of some early form of Sinitic, then the linguistic ancestors of the Chinese would still very much have been relative latecomers, arriving millennia after cultivation had begun to be practised along the Yangtze and Yellow River basins. This is the key feature of Scenario 2. Recently, a study of human leukocyte antigen (HLA) diversity on the genomic region known as the major histocompatibility complex (MHC) purportedly found support for the old linguistic view 'that Altaic speakers in northern China have been switching to Chinese *en masse* in historical times' (Sanchez-Mazas *et al.* 2005: 290). At their current state of temporal resolution, these genetic findings are compatible with Benedict's version of Scenario 2, whereby Chinese arose in a process of language shift, with the Zhōu imposing the Proto-Sinitic language onto a Shāng population speaking some pre-Tibeto-Burman tongue, conceivably perhaps even some early form of Altaic. In fact, Benedict's suggestion about the origin of Sinitic gave expression to older widespread linguistic conjectures regarding the

linguistic prehistory of eastern Asia. Hashimoto's altaicisation hypothesis¹³ about Mandarin originating as a Manchu-Chinese pidgin (1986) can be seen as representing an even later stage in a long-term and intermittent process, the first stage of which was envisaged by Benedict.

An alternative view, which here I have called Scenario 2b, different from the hypothesis advanced by Benedict, envisages the Tibeto-Burman linguistic intrusion onto the North China plain as having first occurred either as early as the Lóngshān horizon, during the subsequent Èrlítóu period, or as late as the Shāng period. The Dàwènkǒu culture of Shāndōng and the Yǎngshóu cultural assemblage were superseded by the more advanced Late Neolithic Lóngshān culture in the middle of the third millennium BC. Population size increased in the Lóngshān period (2600-1900 BC), and jade and ceramic prestige objects proliferated, especially in Shāndōng and southern Shānxī. The walls surrounding many Lóngshān settlements indicate an increased concern with the protection of resources, although one rammed earth wall was also already found to surround a late Yǎngshóu period settlement at Xīshān near modern Zhèngzhōu (Liu and Chen 2006, Underhill and Habu 2006).

Settlement nucleation in the subsequent 二里頭 Èrlítóu period, which emerges ca. 1900 BC, ostensibly during the time of the mythical Xià dynasty, has been taken to indicate increasing craft specialisation and changing patterns of resource management. The burial practices and stratified urban architecture of the Èrlítóu period indicate the emergence of a more complex political infrastructure in what today is northern China. Perhaps the new Èrlítóu social order was established by the first incursive Tibeto-Burmans from the southwest, whereas the walls surrounding Lóngshān settlements were the fortifications which had been intended to repel them, but in vain.

Whichever scenario happens to be our favourite, what would appear to be incontestable is that the Hàn and Tibetan expansions are both historical and relatively recent, and could possibly have effaced and assimilated many Tibeto-Burman and allophylian groups in their paths. Yet even Hàn linguistic and cultural expansion appears not to have been so imperious as to have entirely swept away the Tǔjiǎ, Bái and diverse Lolo-Burmese, Qiāngic, Hmong-Mien and Daic language communities which remain scattered

¹³ Hashimoto wondered whether the typology of Mandarin could be explained as the result of the altaicisation of Chinese or the sinicisation of an Altaic languages, which would have involved either the 'Altaic replacement of Chinese syntax or the Chinese replacement of Altaic lexicon and morphology' (1986: 95).

throughout central and southern China. Neither did the Tibetan expansion annihilate all of the Zhangzhung literary legacy or attestations of other languages still preserved in the Dūnhuáng documents. These residual islets of retention and the now mute testimonies represent vestiges of the older situation.

The expansions of these two branches of Tibeto-Burman, viz. Bodish and Sinitic, are seen precisely in the areas of high mobility where we would expect them to have occurred, i.e. martial expansion across the vast rolling treeless and sparsely populated high alpine plateau of Tibet in the case of Bodish or across the more easily traversable East, where mountainous areas such as Fújiàn were colonised only belatedly by the Hàn. Another important feature of the model underlying each version of Scenario 2 is the temporal dimension. Although Sīchuān is treated as the point of departure for the establishment of early Sinitic in Shǎnxī or perhaps an early Sinitic intrusion even further east into Shāndōng, the ultimate homeland of Tibeto-Burman, as suggested by the diversity observed between the distinct branches of the Tibeto-Burman family, would be expected to have lain far closer to the eastern Himalayas.

Scenario 2c envisages that Tibeto-Burman could have been introduced or re-introduced to the Yellow River basin more than once in the course of prehistory. The Sinitic heartland within the eastern half of what today is China was not politically unified before the Qín dynasty in 221 BC. Rather, monarchs from the house of Zhōu ruled over a constellation of distinct polities in the Yellow River and Yangtze basins during the first millennium BC. It is conceivable that the Shāng, the Zhōu and the Qín¹⁴ could all have spoken different early forms of Tibeto-Burman that influenced each other and ultimately led to the emergence of a Sinitic creole subsequently regularised by the Hàn. Many structural features which Sinitic languages share with young creoles are itemised by Aronoff, Meier and Sandler (2005). It is conceivable that an early migration of ancient Tibeto-Burman speakers to

¹⁴ In the fourth century BC, the Qín were described as 夷 Yí 'barbarians', and later sources such as the 史記 *Shiji* 'Historical Records' written around the beginning of the first century BC described the Qín as similar to the 戎 Róng or 狄 Dī, who strove to emulate Zhōu ritual and tradition. The ethnolinguistic composition of the Qín state must have been complex, and statues distinguished between 故秦民 gù Qín mín, the native Qín population, and 客 kè and 邦客 bāng kè, foreign and subject populations (Shelach and Pines 2006: 205, 217, 220).

gart represent the residue of an early Sino-Austronesian contact situation, then even this could imply that the Proto-Tibeto-Burmans, or some major subset thereof, lived as far east as Shāndōng in the Lóngshān period.

On the other hand, it does not seem that the correspondences necessarily represent anything but a collection of coincidental resemblances, with the exception of a tantalising correspondence first identified as a loan word into Tibetan by Hendrik Kern (1889: 5), viz. Austronesian *beRas 'husked rice' vs. Tibetan *hbras* 'rice'. Sagart has added the Old Chinese cognate 糲 *bma-rat-s*, and pointed out a second rice term Austronesian *Sumay 'rice as food' corresponding to Old Chinese 米 **mij?* 'grain of cereal' and Garo *may* 'paddy'. Kern believed that this loan correspondence pointed to the source whence the ancestors of the Tibetans had first acquired familiarity with rice.

If the veracity of either the Sino-Tibetan or the Sino-Austronesian hypothesis can ever be convincingly demonstrated, then this would compel us to decide in favour of Scenario 3. Yet at present the linguistic evidence for either hypothesis is not compelling. Whilst the lack of conclusive linguistic evidence does not support Scenario 3, neither does it invalidate Bellwood's model. Another line of reasoning which might sustain Bellwood's homeland hypothesis would be to argue that the current distribution of Tibeto-Burman groups could be accounted for if the Himalayas had for millennia served as a refuge area for people fleeing from more belligerent groups raiding, pillaging and waging war across more traversable terrain. The question formulated in the opening paragraph of this section alludes to this possibility. In other words, the present distribution of Tibeto-Burman linguistic diversity could arguably be a function of refuge areas and the traversability of terrain. Populations with cults possibly demanding horrific sacrifice, such as those suggested to some minds by Sānxīngduī iconography alongside the more concrete evidence found at sites such as Anyáng, could have been amongst the repellent influences driving other Tibeto-Burman groups into ever more remote and sheltered alpine recesses.

Finally, Bellwood's Tibeto-Burman homeland in Scenario 3 extends all across Shǎnxī and abuts against Sīchuān, the homeland of Scenario 1. The disparity, therefore, is greater between Scenarios 2 and 3 than between Scenarios 1 and 3. The merit of Scenario 2 is that linguistic prehistory is reconstructed on the basis of the linguistic diversity situation, whereas the archaeological record is treated as testimony of the prehistory of material culture. One interpretation of the emerging population genetic data could support a version of Scenario 2, but this too may change as more data are analysed and interpreted and this multi-faceted story continues to unfold in ways perhaps unforeseen.

the Himalayan region was followed by a later migration of a Tibeto-Burman group back to the Yellow River basin.

Let us turn to a third view, which I shall call Scenario 3. Bellwood places the homeland of Tibeto-Burman, which he refers to as 'Sino-Tibetan', in an elongated region stretching along the lower course of the Yellow River in the northeast deep into Shānxī in the southwest. From this oblong territory he envisages the language family spreading into Gānsù and southwest into the Himalayas. The idea of agricultural dispersals in the Neolithic is an enthralling model. Such an interpretation of the archaeological record is an obvious one and was already pioneered by scholars such as Robert von Heine-Geldern. Yet the farming-language dispersal theory advocated principally by Bellwood and Renfrew differs essentially from associating the rapid spread of a specific and well-defined cultural assemblage such as, for instance, the Mǎjiāyáo Neolithic in Gānsù and its recognisable colonial exponents in eastern Tibet and Kashmir with the putative spread of ancestral Sino-Bodic groups across the Tibetan Plateau towards certain parts of the Himalayas.

Instead, the farming-language dispersal theory envisages genes and language spreading in tandem with the incremental spread of Neolithic agriculture. Weaknesses in this theory have been discussed in Section 3, where it was argued that the very opposite actually have happened in many instances in the course of prehistory. We must consider such an alternative especially in those cases where the linguistic picture suggests a radically different view of prehistory than does the spread of material culture as reflected in the known archaeological record. What can be said in favour of Scenario 3? The Sino-Tibetan hypothesis that there exist some shared innovations that unify all non-Sinitic languages within a truncated 'Tibeto-Burman' group remains empirically unsupported. Yet the hypothesis remains an intrinsically interesting one, especially from a Sinocentric perspective, and it cannot be excluded that some linguistic evidence for it may be found one day. Clearly an empirically unsupported linguistic hypothesis cannot buttress the case for Scenario 3.

A second potential argument in favour of Scenario 3 might be sought in the Sino-Austronesian theory, the veracity of which, likewise, has yet to be demonstrated. I have already discussed in detail the evidence marshalled in support of the Sino-Austronesian theory elsewhere (van Driem 2005). If Sino-Austronesian were demonstrated to have existed as an ancient genetic unity comprising Tibeto-Burman and Austronesian, as Sagart believes, then this could imply that the Proto-Tibeto-Burmans were on the North China plain at the Lóngshān horizon. Yet, if the correspondences adduced by Sa-

6. BIBLIOGRAPHY

- Aldenderfer, Mark, and Zhāng Yínóng. 2004. The prehistory of the Tibetan Plateau to theseventh century A.D.: Perspectives from China and the West since 1950, *Journal of World Prehistory*, 18 (1): 1-55.
- Allard, Francis. 2006. Frontiers and boundaries: The Han empire from its southern periphery, pp. 233-254 in Miriam T. Stark, ed., *Archaeology of Asia*. Oxford: Blackwell. [actually appeared in early 2005]
- Alt, Kurt W. Joachim Burger, Angela Simons, Werner Schön, Gisela Grupe, Susanne Hummel, Birgit Grosskopf, Werner Vach, Carlos Buitrago Téllez, Christian-Herbert Fischer, Susan Möller-Wiering, Sukra S. Shrestha, Sandra L. Pichler and Angela von den Driesch. 2003. Climbing into the past — first Himalayan mummies discovered in Nepal, *Journal of Archaeological Science*, 30: 1529-1535.
- Ān Zhīmín. 1992. Neolithic communities in eastern parts of Central Asia, pp. 153-189 in Ahmad Hasan Dani and Vadim Mixajlovič Masson, eds., *History of Civilizations of Central Asia* (Vol. 1: The dawn of civilization: earliest times to 700 BC). Paris: United Nations Educational Scientific and Cultural Organization.
- Aronoff, Mark, Irit Meier and Wendy Sandler. 2005. The paradox of sign language morphology, *Language*, 81 (2): 301-344.
- Ashma, R., and V.K. Kashyap. 2002. Genetic polymorphism at 15 STR loci among three important subpopulations of Bihar, India, *Forensic Science International*, 130: 58-62.
- Baig, M.M., A.A. Khan and K.M. Kulkarni. 2004. Mitochondrial DNA diversity in tribal and caste groups of Maharashtra (India) and its implication on the their genetic origins, *Annals of Human Genetics*, 68: 453-460.
- Banerjee, Jheelam, R. Trivedi and V.K. Kashyap. 2005a. Y-STR haplotypes in autochthonous tribal population of the Chotanagpur Plateau, India, *Journal of Forensic Science*, 50 (3), paper JFS2004531.
- Banerjee, Jheelam, R. Trivedi and V.K. Kashyap. 2005b. Polymorphism at 15 short tandem repeat AmpF ℓ STR ® Identifier ™ loci in three aboriginal populations in India: An assessment in human identification, *Journal of Forensic Science*, 50 (5), paper JFS2005151.
- Bellwood, Peter. 2006. Asian farming diasporas? Agriculture, languages and genes in China and Southeast Asia, pp. 96-118 in Miriam T. Stark, ed., *Archaeology of Asia*. Oxford: Blackwell. [actually appeared in early 2005]
- Benedict, Paul King. 1972. *Sino-Tibetan: A Conspectus*. Cambridge: Cambridge University Press.

- Bergsland, Knut, and Hans Vogt. 1962. On the validity of glottochronology, *Current Anthropology*, 3: 115-153.
- 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 Kwang-Chih (i.e. Zhāng Guāngzhī). 1977. *The Archaeology of Ancient China* (3rd edition). New Haven, Connecticut: Yale University Press.
- Chang Kwang-Chih (i.e. Zhāng Guāngzhī). 1986. *The Archaeology of Ancient China* (4th edition, revised and enlarged). New Haven, Connecticut: Yale University Press.
- Chaubey, Gyaneshwer (i.e. Jñāneśvara Caube). 2006. 'Phylogeography of haplogroup R7', paper presented at the European Molecular Biology Organisation workshop on Human Evolution and Disease, Centre for Cellular and Molecular Biology at Hyderabad, 7 December 2006.
- Chen Teng, Lai Jianghua, Chen Yanjiong and Li Shengbin. 2006. Population genetics of nine STR loci from Baoan population in NW China, *Forensic Science International*, 157: 218-220.
- Chrétien, C. Douglas. 1962. The mathematical models of glottochronology, *Language*, 38: 11-37.
- Chú Jiāyòu and W. Huang, S.Q. Kuang, J.M. Wang, J.J. Xu, Z.T. Chu, Z.Q. Yang, K.Q. Lin, P. Li, M. Wu, Z.C. Geng, C.C. Tan, R.F. Du and Jin Li. 1998. Genetic relationship of populations in China, *Proceedings of the National Academy of Sciences of the United States of America*, 95: 11763-11768.
- Cordaux, Richard, Nilmani Saha, Gillian R. Bentley, Robert Aunger, S.M. Sirajuddin and Mark Stoneking. 2003. Mitochondrial DNA analysis reveals diverse histories of tribal populations from India, *European Journal of Human Genetics*, 11: 253-264.
- Cordaux, Richard, G. Weiss, Nilmani Saha and Mark Stoneking. 2004. The northeast Indian passageway: A barrier or corridor for human migrations?, *Molecular Biology and Evolution*, 21: 1525-1533.
- Corvinus, Gudrun. 1996. The prehistory of Nepal after ten years of research, pp. 43-55, Vol. 1 in Peter Bellwood, ed., *Indo-Pacific Prehistory: The Chiang-Mai Papers* (Proceedings of the 15th Congress of the Indo-Pacific Prehistory Association, Chiang Mai, Thailand, 5 to 12 January 1994). Canberra: Australian National University.
- Cowan, Hendrik Karel Jan. 1959. A note on statistical methods in comparative linguistics, *Lingua*, VIII (3): 233-246.

- Crawford, Gary W. 2006. East Asian plant domestication, pp. 77-95 in Miriam T. Stark, ed., *Archaeology of Asia*. Oxford: Blackwell. [actually appeared in early 2005]
- Cust, Robert N. 1878. *A Sketch of the Modern Languages of East India*. London: Trübner and Company.
- Debnath, M., and T.K. Chaudhuri. 2005. HLA-A and HLA-B distribution in the Toto — a vanishing sub-Himalayan tribe of India, *Tissue Antigens*, 67: 64-65.
- Dèng Xiǎohuá and Wáng Shiyuán. 2003. Zàngmiǎn yǔzú yǔyán de shùlǐ fēnlèi jíqí fēnxī, *Mínzú Yǔwén*, 2003 (4): 8-18.
- Diffloth, Gérard. 2001. 'Tentative calibration of time depths in Austroasiatic branches', paper presented at the Colloque «Perspectives sur la Phylogénie des Langues d'Asie Orientales» at Périgueux, France, 30 August 2001.
- 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.
- Ding Yuan-Chun, Stephen Wooding, Henry C. Harpending, Chi Han-Chang, Li Hai-Peng, Fu Yun-Xin, Pang Jun-Feng, Yao Yong-Gang, Xiang Yu Jing-Gong, Robert Moyzis and Zhang Ya-ping. 2000. Population structure and history in East Asia, *Proceedings of the National Academy of Sciences of the United States of America*, 97: 14003-14006.
- van Driem, George. 1998. Neolithic correlates of ancient Tibeto-Burman migrations, pp. 67-102 in Roger Blench and Matthew Spriggs, eds., *Archaeology and Language II*. London: Routledge.
- van Driem, George. 2001. *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region, containing an Introduction to the Symbiotic Theory of Language* (2 vols.). Leiden: Brill.
- van Driem, George. 2002. Tibeto-Burman phylogeny and prehistory: Languages, material culture and genes, pp. 233-249 in Peter Bellwood and Colin Renfrew, eds., *Examining the Farming/Language Dispersal Hypothesis*. Cambridge: McDonald Institute for Archaeological Research.
- van Driem, George. 2005. Sino-Austronesian vs. Sino-Caucasian, Sino-Bodic vs. Sino-Tibetan, and Tibeto-Burman as default theory, pp. 285-338 in Yogendra Prasada Yadava, Govinda Bhattarai, Ram Raj Lohani, Balaram Prasain and Krishna Parajuli, eds., *Contemporary*

- Issues in Nepalese Linguistics*. Kathmandu: Linguistic Society of Nepal.
- van Driem, George. 2006. The diversity of the Tibeto-Burman language family and the linguistic ancestry of Chinese, keynote address to the joint meeting of the 14th Annual Conference of the International Association of Chinese Linguistics and the 10th International Symposium on Chinese Languages and Linguistics at the Academia Sinica in Taipei, Saturday, 27 May 2006 [60-page manuscript submitted for publication].
- Du Hong, Liang Weibo, Lv Meili, Chen Guodi and Lin Zhang. 2005. Allele frequency distributions for 9 STR loci of Tibetan population in Chinese Tibet, *Journal of Forensic Science*, 50 (3), paper JFS20044523.
- Dumont d'Urville, Jules Sébastien César. 1830a, 1832a, 1833a, 1834a, 1835a, 1833b; 1832b, 1834b; 1832c, 1835b; 1834c. *Voyages de découvertes de l'Astrolabe exécuté par ordre du Roi pendant les années 1826-1827-1828-1829 sous le commandement de M. J. Dumont d'Urville: Zoologie* (4 vols., volumes II and III each in two parts), *Botanique* (2 vols.), *Faune Entomologique* (2 vols.), *Philologie* (1 vol.) Paris: J. Tastu [except for the volume *Philologie*, which was published by le Ministère de la Marine].
- Dunn, Michael, Angela Terrill, Ger Reesink, Robert A. Foley and Stephen C. Levinson. 2005. Structural phylogenetics and the reconstruction of ancient language history, *Science*, 309: 2072-2075.
- Forbes, Charles James F.S. 1878. On Tibeto-Burman languages, *Journal of the Royal Asiatic Society of Great Britain and Ireland*, New Series, X: 210-227.
- Fortescue, Michael. 1998. *Language Relations across Bering Strait: Reappraising the Archaeological and Linguistic Evidence*. London: Cassell.
- Gaillard, Claire. 1996. Processing sequences in the Indian lower Palaeolithic: Examples from the Acheulian and Soanian, pp. 57-67, Vol. 1 in Peter Bellwood, ed., *Indo-Pacific Prehistory: The Chiang-Mai Papers* (Proceedings of the 15th Congress of the Indo-Pacific Prehistory Association, Chiang Mai, Thailand, 5 to 12 January 1994). Canberra: Australian National University.
- Gray, Russell D., and Quentin D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin, *Nature*, 426: 435-439.
- Guy, Jacques B.M. 1994. On lexicostatistics and glottochronology, *Dhumbadji!*, 2 (1): 55-87.

- ben Hamed, Mahé, and Wāng Fēng. 2006. Stuck in the forest: Trees, networks and Chinese dialects, *Diachronica*, 23 (1): 29-60.
- Hansen, Chad. 1993. Chinese ideographs and Western ideas, *Journal of Asian Studies*, 52: 373-399.
- Hashimoto, Mantaro Joseph. 1986. The altaicization of Northern Chinese, pp. 76-97 in John McCoy and Timothy Light, eds., *Contributions to Sino-Tibetan Studies*. Leiden: E.J. Brill.
- Houghton, Bernard. 1896. Outlines of Tibeto-Burman linguistic palaeontology, *Journal of the Royal Asiatic Society*, 1896: 23-55.
- Hu, W.-H., J. Lu, Y.-P. Lei, B.-W. Chen, W.-R. Tang, Y.-N. Cun, Y.-L. Dong, S.-J. Tan. H.-J. Yu, C.-J. Xiao. 2005. HLA-DPB1 allelic frequency of the Lisu ethnic group in the Southwest China [sic] and evolutionary relationship of Lisu with other populations, *Tissue Antigens*, 65: 289-292.
- Hu, W.-H., J. Lu, Y.-P. Lei, B.-W. Chen, W.-R. Tang, Y.-N. Cun, Y.-L. Dong, S.-J. Tan. H.-J. Yu, C.-J. Xiao. 2005. Polymorphism of the DPB1 locus in Hani ethnic group of south-western China, *International Journal of Immunogenetics*, 32: 421-423.
- Hu, S.P., X.J. Yu, J.W. Liu and K.L. Cai. 2005. Analysis of STR polymorphisms in the Chao Shan population in South China, *Forensic Science International*, 147: 93-95.
- Hu, S.P., D.Q. Wu, X.H. Xu, J.W. Liu and B. Li. 2005. Genetic profile of 15 STR loci in the Min Nan population in Southeast China, *Forensic Science International*, 152: 263-265.
- Jaxontov, Sergej Evgen'evič. 1977. Jazyki vostočnoj i jugovostočnoj Azii v pervom tysjačiletii do našej èry, pp. 98-108 in Nikolaj Nikolaevič Čeboksarov, Mixail Vasil'evič Krjukov i Mixail Viktorovič Sofronov, eds., *Rannjaja ètničeskaja istorija i narody vostočnoj Azii*. Moscow: Nauka.
- Kang Longli and Li Shengbin. 2006. X-chromosome STR polymorphism of Luoba ethnic group living in Tibet (SW China), *Forensic Science International*, 156: 88-90.
- Karafet, Tatiana, Xu Liping, Du Ruofu, William Wang, Feng Shi, R. Spencer Wells, Alan J. Redd, Stephen L. Zegura and Michael F. Hammer. 2001. Paternal population history in East Asia: Sources, patterns and microevolutionary processes, *American Journal of Human Genetics*, 69: 615-628.
- Kashyap, V.K., N. Sarkar and R. Trivedi. 2002. Allele frequencies for STR loci of the PowerPlex™ 16 multiplex system in five endogamous populations of India, *Forensic Science International*, 126: 178-186.

- Kashyap, V.K., Saurav Guha and R. Trivedi. 2002. Concordance study on 15 STR loci in three major populations of the Himalayan state Sikkim, *Journal of Forensic Science*, 47 (5): paper JFS2001333 475.
- Keightley, David N. 2006. Marks and labels: Early writing in Neolithic and Shang China, pp. 177-201 in Miriam T. Stark, ed., *Archaeology of Asia*. Oxford: Blackwell. [actually appeared in early 2005]
- Kern, Hendrik. 1889. Het stamland der Maleisch-Polynesische volken, *Tijdschrift voor Nederlandsch-Indië*, XVIII (2): 1-9.
- Kivisild, Toomas, Katrin Kaldma, Mait Metspalu, Jüri Parik, Surinder S. Papiha and Richard Villems. 1999a. The place of the Indian mitochondrial DNA variants in the global network of maternal lineages and the peopling of the Old World, pp. 135-152 in Surinder S. Papiha, R. Deka and R. Chakraborty, eds., *Genomic Diversity: Applications in Human Population Genetics*. New York: Kluwer Academic, Plenum Publishers.
- Kivisild, Toomas, Michael J. Bamshad, Katrin Kaldma, Mait Metspalu, E. Metspalu, M. Reidla, S. Laos, Jüri Parik, W. Scott Watkins, Mary E. Dixon, Surinder S. Papiha, S.S. Mastana, M.R. Mir, V. Ferak and Richard Villems. 1999b. Deep common ancestry of Indian and Western-Eurasian mitochondrial DNA lineages, *Current Biology*, 9 (22): 1331-1334.
- Kivisild, Toomas, S. Rootsi, Mait Metspalu, S. Mastana, Katrin Kaldma, Jüri Parik, E. Metspalu, M. Adojaan, H.-V. Tolk, V. Stepanov, M. Gölge, E. Usanga, Surinder S. Papiha, C. Cinniöglu, R. King, Luca Luigi Cavalli-Sforza, Peter A. Underhill and Richard Villems. 2003. The genetic heritage of the earliest settlers persist both in Indian tribal and caste populations, *American Journal of Human Genetics*, 72: 313-332.
- von Klaproth, Julius Heinrich. 1823. *Asia Polyglotta*. Paris: A. Schubart.
- Kraaijenbrink, Thirsa, and George van Driem, Karma Tshering of Gaseló and Peter de Knijff. 2006a. Allele frequency distribution for 21 autosomal STR loci in Bhutan, *Forensic Science International*, [in press].
- Kraaijenbrink, Thirsa, and George van Driem, Jean Robert M.L. Opgenort, Nirmal Man Tuladhar and Peter de Knijff. 2006b. Allele frequency distribution for 21 autosomal STR loci in Nepal, *Forensic Science International*, [in press].
- Krithika, S., R. Trivedi, V.K. Kashyap and T.S. Vasulu. 2005. Genetic diversity at 15 microsatellite loci among the Adi Pasi population of

- Adi tribal cluster in Arunachal Pradesh, India, *Legal Medicine*, 7: 306-310.
- Kumar, Vikrant, Debashis Basu and B. Mohan Reddy. 2004. Genetic heterogeneity in northeastern India: Reflection of tribe-caste continuum in the genetic structure, *American Journal of Human Biology*, 16: 334-345.
- Langstieh B.T., B. Mohan Reddy, Kumarasamy Thangakraj, Vikrant Kumar and Lalji Singh. 2004. Genetic diversity and relationships among the tribes of Meghalaya compared to other Indian and Continental populations, *Human Biology*, 76 (4): 569-590.
- Liang Weibo, Lv Meili, Xu Yonchun, Liao Miao, Zhou Bin, Jia Yu and Lin Zhang. 2005. STR loci D19S400's allele frequency distribution in ten populations, *Journal of Forensic Science*, 50 (3), paper JFS20044524.
- Lin, Marie, Chu Chen-Chung, Richard E. Broadberry, Yu Lung-Chih, Loo Jun-Hun and Jean A. Trejaut. 2005. Genetic diversity of Taiwan's indigenous peoples, pp. 232-249 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.
- Liu Li and Chen Xingcan. 2006. Sociopolitical change from Neolithic to Bronze Age China, pp. 149-176 in Miriam T. Stark, ed., *Archaeology of Asia*. Oxford: Blackwell. [actually appeared in early 2005]
- Madsen, David B., Ma Haizhou, P. Jeffrey Brantingham, Gai Xing, David Rhode, Zhang Haiying and John W. Olsen. 2006. The late Upper Palaeolithic occupation of the northern Tibetan Plateau margin, *Journal of Archaeological Science*.
- Maity, B., S.C. Nunga and V.K. Kashyap. 2003. Genetic polymorphism revealed by 13 tetrameric and 2 pentameric STR loci in four Mongoloid tribal populations, *Forensic Science International*, 132: 216-222.
- Mason, Francis. 1854. The Talaiing 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.
- Matisoff, James Alan. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibeto-Burman Reconstruction*. Berkeley: University of California Press.
- Metspalu, Mait, Toomas Kivisild, Ene Metspalu, Jüri Parik, Georgi Hudjashov, Katrin Kadma, Piia Serk, Monika Karmin, Doron M. Behar, M. Thomas P. Gilbert, Phillip Endicott, Sarabjit Mastana, Surinder S.

- Papiha, Karl Skorecki, Antonio Torroni and Richard Villems. 2004. Most of the extant mtDNA boundaries in South and Southeast Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans, *BioMed Central, Genetics*, 2004 (5): 26 [25 pages].
- Nichols, Johanna. 1992. *Linguistic Diversity in Time and Space*. Chicago: Chicago University Press.
- Nichols, Johanna. 1998. The Eurasian spread zone and the Indo-European spread, pp. 220-266 in Roger Marsh Blench and Matthew Spriggs, eds., *Archaeology and Language II: Archaeological Data and Linguistic Hypotheses*. London: Routledge.
- Pandey, R.N. 1987. Palaeo-environment and prehistory in Nepal, *Contributions to Nepalese Studies*, 14 (2): 111-124.
- Parkin, Emma J., Thirsa Kraayenbrink, Karma Tshering of Gaselô, George van Driem, Peter de Knijff and Mark Jobling. 2006a. 26-locus Y-STR typing in a Bhutanese population sample, *Forensic Science International*, 161 (1): 1-7.
- Parkin, Emma J., Thirsa Kraayenbrink, Jean Robert M.L. Opgenort, George van Driem, Nirmal Man Tuladhar, Peter de Knijff and Mark Jobling. 2006b. High-resolution Y-STR haplotypes in a Nepalese population sample: Isolation and drift in the Himalayas [in press].
- Peiros, Ilija [i.e. Il'ja Iosifovič Pejros] 1998. *Comparative Linguistics in Southeast Asia*. Canberra: Pacific Linguistics.
- Poloni, Estella Simone, Ornella Semino, Giuseppe Passarino, A.S. Santachiara-Benerecetti, I. Dupanloup, André Langaney and Laurent Excoffier. 1997. Human genetic affinities for Y chromosome P49a,f/TaqI haplotypes show strong correspondence with linguistics, *American Journal of Human Genetics*, 61: 1015-1035 (cf. the erratum published in 1998 in *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.

- Przyluski, Jean, and Gordon Hannington Luce. 1931. The number 'a hundred' in Sino-Tibetan, *Bulletin of the School of Oriental Studies*, VI (3): 667-668.
- Qamar, Raheel, Qasim Ayub, Aisha Mohyuddin, Agnar Helgason, Kehkashan Mazhar, Atika Manoor, Tatiana Zerjal, Chris Tyler-Smith and S. Qasim Mehdi. 2002. Y-chromosomal variation in Pakistan, *American Journal of Human Genetics*.
- Quintana-Murci, Lluís, C. Krausz, Tatiana Zerjal, S. Hamid Sayar, Michael F. Hammer, S. Qasim Mehdi, Qasim Ayub, Raheel Qamar, Aisha Mohyuddin, U. Radhakrishna, Mark A. Jobling Chris Tyler-Smith and Ken McElreavey. 2001. Y-chromosome lineages trace diffusion of people and languages in southwestern Asia, *American Journal of Human Genetics*, 68: 537-542.
- Quintana-Murci, Lluís, Raphaëlle Chaix, R. Spencer Wells, Doron M. Behar, Hamid Sayar, Rosaria Scozzari, Chiara Rengo, Nadia al-Zahery, Ornella Semino, A. Silvana Santachiara-Benerecetti, Alfredo Coppa, Qasim Ayub, Aisha Mohyuddin, Chris Tyler-Smith, S. Qasim Mehdi, Antonio Torroni and Ken McElreavey. 2004. Where West meets East: The complex mtDNA landscape of the Southwest and Central Asian corridor, *American Journal of Human Genetics*, 74.
- Sahoo, Sanghamitra, and V.K. Kashyap. 2002. Genetic variation at 15 autosomal microsatellite loci in the three highly endogamous tribal populations of Orissa, India, *Forensic Science International*, 130: 189-193.
- 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*, 103 (4): 843-848.
- Sanchez-Mazas, Alicia, Estella S. Poloni, Guillaume Jacques and Laurent Sagart. 2005. HLA genetic diversity and linguistic variation in East Asia, pp. 273-296 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.
- 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*. München: Kaiserliche Akademie.

- Schmidt, Wilhelm. 1906. 'Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens', *Archiv für Anthropologie, Neue Folge*, V: 59-109.
- Shelach, Gideon, and Yuri Pines. 2006. Secondary state formation and the development of local identity: Change and continuity in the state of Qin (770-221 BC), pp. 202-230 in Miriam T. Stark, ed., *Archaeology of Asia*. Oxford: Blackwell. [actually appeared in early 2005]
- Shi Hong, Dong Yong-li, Wen Bo, Xiao Chun-Jie, Peter A. Underhill, Shen Pei-dong, Ranajit Chakraborty, Jin Li and Sù Bing. 2005. Y-chromosome evidence of southern origin of the East Asian-specific haplogroup O3-M122, *American Journal of Human Genetics*, 77: 408-419.
- Shi, L., S.B. Xu, J. Ohashi, H. Sun, J.K. Yu, X.Q. Huang, Y.F. Tao, L. Yu, S. Horai, J.Y. Chu and K. Tokunaga. 2006. HLA-A, HLA-B and HLA-DRB1 alleles and haplotypes in Naxi and Han populations in southwestern China (Yunnan province), *Tissue Antigens*, 67: 38-44.
- Singh, Anamika, R. Trivedi and V.K. Kashyap. 2006. Genetic polymorphism at 15 tetrameric short tandem repeat loci in four aboriginal tribal populations of Bengal, *Journal of Forensic Science*, 51 (1): 183-187.
- Sù Bing, Xiao Junhua, Peter Underhill, Ranjan Deka, Zhang Weiling, Joshua Akey, Huang Wei, Shen Di, Daru Lu, Luo Jingchun, Chu Jiayou, Tan Jiazhen, Shen Peidong, Ron Davis, Luca Cavalli-Sforza, Ranajit Chakraborty, Xiong Momiao, Du Ruofu, Peter Oefner, Chen Zhu and Jin Li. 1999. Y-chromosome evidence for a northward migration of modern humans into eastern Asia during the last Ice Age, *American Journal of Human Genetics*, 65: 1718-1724.
- Sù Bing, Chunjie Xiao, Ranjan Deka, Mark T. Seielstad, Daoroong Kangwanpong, Junhua Xiao, Daru Lu, Peter Underhill, Luca Cavalli-Sforza, Ranajit Chakraborty and Jin Li. 2000. Y chromosome haplotypes reveal prehistorical migrations to the Himalayas, *Human Genetics*, 107 (6): 582-590.
- Tambets, Kristiina, Siiri Rootsi, Toomas Kivisild and Richard Villems. 2001. The concepts of Richard Indreko about the origins of the Finno-Ugric speakers and the population genetics of the extant north-east European populations, *Trames, Journal of the Humanities and Social Sciences*, 5 (55/50), (1): 59-74.
- Thangaraj, Kumarasamy, Vempati Sridhar, Toomas Kivisild, Alla G. Reddy, Gyaneshwer Chaubey, Vijay Kumar Singh, Suminder Kaur, Pooja Agarawal, Amit Rai, Jalaj Gupta, Chandana Basu Mallick, Niraj Kumar, Thrimulaisamy P. Velavan, Rajanbabu Suganthan, Divi Udaykumar, Rashmi Kumar, Rachana Mishra, Arif Khan, Chitikineni Anna-

- purna and Lalji Singh. 2005. 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*, 116: 507-517.
- Thomas, R,m S.B. Nair and M. Banerjee. 2004. HLA-B and HLA-C alleles and haplotypes in the Dravidian tribal populations of southern India, *Tissue Antigens*, 64: 58-65.
- Tsang, Cheng-hwa. 2004. Recent discoveries of the Tapenkeng culture in Taiwan, pp. 63-74 in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*. London: Routledge Curzon.
- Tyler-Smith, Chris, Denise Rejane Carvalho-Silva, Thirsa Kraaijenbrink, Emma J. Parkin, Nirmal Man Tuladhar, Jean Robert Oppenort, Karma Tshering of Gaselô, George van Driem, Guido Barbujani, Mark A. Jobling and Peter de Knijff. 2006. 'Genetic, linguistic and geographical boundaries in the Himalayan region], paper presented at the European Molecular Biology Organisation workshop on Human Evolution and Disease, Centre for Cellular and Molecular Biology at Hyderabad, 7 December 2006.
- Underhill, Anne P., and Junk Habu. 2006. Early communities in East Asia: Economic and sociopolitical organisation at the local and regional levels, pp. 121-148 in Miriam T. Stark, ed., *Archaeology of Asia*. Oxford: Blackwell. [actually appeared in early 2005]
- Underhill, Peter A., Giuseppe Passarino, A.A. Lin, P. Shen, M. Mirazón-Lahr, Robert A. Foley, P.J. Oefner and Luca Luigi Cavalli-Sforza. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations, *Annals of Human Genetics*, 65: 43-62.
- Vandermeersch, Léon. 1980. *Wangdao ou la voie royale: Recherches sur l'esprit des institutions de la Chine archaïque, Tome II: Structures politiques, les rites*. Paris: École française d'Extrême-Orient.
- Watkins, W.S., B.V.R. Prasad, J.M. Naidu, B.B. Rao, B.A. Bhanu, B. Ramachandran, P.K. Das, P.B. Gai, P.C. Reddy, P.G. Reddy, M. Sethuraman, M.J. Bamshad and L.B. Jorde. 2005. Diversity and divergence among tribal populations of India, *Annals of Human Genetics*, 69: 680-692.
- Wells, R. Spencer, Nadira Yuldasheva, Ruslan Ruzibakiev, Peter A. Underhill, Irina Evseeva, Jason Blue-Smith, Jin Li, Sù Bing, Ramasamy Pitchappan, Sadagopal Shanmugalakshmi Karuppiiah Balakrishnan, Mark Read, Nathaniel M. Pearson, Tatiana Zerjal, Matthew T. Webster, Irakli Zholoshvili, Elena Jamarjashvili, Spartak Gambarov, Beh-

- rouz Nikbin, Ashur Distiev, Oganazar Aknazarov, Pierre Zalloua, Igor Tsoy, Mikhail Kitaev, Mirsaid Mirrakhimov, Ashir Chariev and Walter F. Bodmer. 2001. The Eurasian heartland: A continental perspective on Y-chromosome diversity, *Proceedings of the National Academy of Sciences*, 98: 10244-10249.
- Wen Bo, Li Hui, Lu Daru, Song Xiufeng, Zhang Feng, He Yungang, Li Feng, Gao Yang, Mao Xianyun, Zhang Liang, Qian Ji, Tan Jingze, Jin Jianzhong, Huang Wei, Ranjan Deka, Sù Bīng, Ranajit Chakraborty and Jin Li. 2004. Genetic evidence supports demic diffusion of Han culture, *Nature*, 431: 302-305.
- Wen Bo, Xie Xuanhua, Gao Song, Li Hui, Shi Hong, Song Xiufeng, Qian Tingzhi, Xiao Chunjie, Jin Jianzhong, Sù Bīng, Lu Daru, Ranajit Chakraborty and Jin Li. 2004. Analyses of genetic structure of Tibeto-Burman populations reveals sex-biased admixture in southern Tibeto-Burmans, *American Journal of Human Genetics*, 74 (5): 856-865.
- Wheeler, Robert Eric Mortimer. 1966. *Civilizations of the Indus Valley and Beyond*. London: Thames and Hudson.
- Wheeler, Robert Eric Mortimer. 1968. *The Indus Civilization* (Supplementary Volume to the Cambridge History of India) (3rd edition). Cambridge: Cambridge University Press.
- Xīzàng Zìzhìqū Wénwù Guǎnlǐ Wěiyuán Huì ['Managing Committee for Cultural Affairs of the Autonomous Region of Tibet']. 1979. Xīzàng Chāngdū Kǎruò yízhǐ shǐjué jiǎn bào ['Brief report of an initial excavations of the site at mKhar-ro, Chab-mdo, Tibet'], *Wénwù*, 1979 (9): 22-28.
- Xue Yali, Tatjana Zerjal, Bao Weidong, Zhu Suling, Si-Keun Lim, Shu Qunfang, Xu Jiujin, Du Ruofu, Fu Songbin, Li Pu, Yang Huanming and Chris Tyler-Smith. 2005. Recent spread of a Y-chromosomal lineage in northern China and Mongolia, *American Journal of Human Genetics*, 77: 1112-1116.
- Yang Yajun, Xie Xiaodeng, Wang Xunling, Liu Yanhong, Zhou Ruixia, Zhang Weihua, Xu Jiujin and An Lizhe. 2005. Allele frequency distribution at seven Y-chromosomal STR loci among five ethnic groups from northwest China, *Journal of Forensic Science*, 50 (3), paper JFS2004469.
- Zerjal, Tatjana, B. Dashnyam, A. Pandya, Manfred Kayser, Lutz Roewer, F.R. Santos, W. Schiefenhover, N. Fretwell, M.A. Jobling, S. Harihara, K. Shimuzu, D. Semjidmaa, Antti Sajantila, P. Salo, M.H. Crawford, E.K. Ginter, Oleg V. Efgrafov and Chris Tyler-Smith. 1997. Genetic

- relationships of Asian and Northern Europeans, revealed by Y chromosomal DNA analysis, *American Journal of Human Genetics*, 62: 1174-1183.
- Zhang, D.D., S.H. Li, Y.Q. He and B.S. Li. 2003. Human settlement of the last glaciation on the Tibetan plateau, *Current Science*, 84 (5): 701-704.
- Zhang, X.L., and C.Y. Zhang, Y. Yu, F. Chen, P. Li and S.B. Fu. 2006. CCR2 allele polymorphisms in 15 Chinese ethnic populations, *International Journal of Immunogenetics*, 33: 45-48.
- Zhu Bofeng, Wang Zhenyuan, Wu Qiangju, Maimait. Eli Sedike, Zhu Jun, Huang Ping, Xu Yongcheng and Liu Yao. 2005. Genetic analysis of 15 STR loci of Chinese Uigur ethnic population, *Journal of Forensic Science*, 50 (5): 1-2.