6

TIBETO-BURMAN VS INDO-CHINESE

Implications for population geneticists, archaeologists and prehistorians

George van Driem

The Tibeto-Burman language family

Inklings of a TB language family first appeared in the eighteenth century, when western scholars observed that Tibetan was genetically related to Burmese. However, the precise contours of the TB language family were first defined in Paris in 1823 by the German scholar Julius Heinrich von Klaproth, the same man who first coined the term 'Indogermanisch'. In his *Asia Polyglotta*, Klaproth (1823a,b) defined TB as the language family which comprised Burmese, Tibetan and Chinese and all languages which could be demonstrated to be genetically related to these three. He explicitly excluded Thai (i.e. Daic) as well as Vietnamese and Mon (i.e. AA) because the comparison of lexical roots in the core vocabulary indicated that these languages were representatives of other distinct language phyla.

Julius Heinrich Klaproth was born on 11 October 1783 in Berlin and died on 28 August 1835 in Paris. As a young man he travelled to China in the years 1805–06 and again in 1806–07. He was widely read and mastered a good number of oriental tongues. He edited the *Asiatisches Magazin* in Weimar, became a foreign associate of the Société Asiatique after its founding in 1821 in Paris. He was the first to observe that the root for 'birch', a phytonym which Sanskrit shares with other Indo-European languages, was important to an understanding of the population prehistory of the subcontinent:

> Il est digne de remarque que le bouleau s'appelle en sanscrit भूर्ज्ज *bhourtchtcha*, et que ce mot dérive de la même racine que l'allemand *birke*, l'anglais *birch* et le russe, берeза (bereza), tandis que les noms des autres arbres de l'Inde ne se retrouvent pas dans les langues indo-germaniques de l'Europe. La raison en est, vraisemblablement, que les nations indo-germaniques venaient du nord, quand elles entrèrent dans l'Inde, où elles apportèrent la langue qui a servi de base au sanscrit, et qui a repoussé de la presqu'ile, les idiomes de la même origine que

> le malabar et le télinga, que ces nations, dis-je, ne trouvèrent pas dans leur nouvelle patrie les arbres qu'elles avaient connu dans l'ancienne, à l'exception du bouleau, qui croît sur le versant méridional de l'*Himâlaya*.
>
> (Klaproth 1830: 112–13)

This idea which was later seized upon by the Swiss linguist Adolphe Pictet, who coined the term 'linguistic palaeontology' in his 1859 study *Les origines indo-européennes ou les aryas primitifs: Essai de paléontologie linguistique*.

As far as I have been able to trace, Klaproth (1823: 380) was also the first to state clearly that the Formosan languages were members of the An family, genetically related to Malay and Malagasy. Klaproth carefully scrutinised the lexical and grammatical data available at the time, and, following the precedents set by Nicolaes Witsen (1692) and Phillip von Strahlenberg (1730), he was the first to be able to present an informed and comprehensive polyphyletic view of Asian languages and language families. In order to reconcile this view with his religious beliefs, Klaproth (1823: 43) devised a table of correspondence between Hindu and Biblical chronology, dating '*die große Ausbreitung des Indo-Germanischen Völkerstammes*' to a prehistoric period '*vielleicht schon vor der Noah'ischen Fluth*'. He identified and distinguished 23 main Asian linguistic stocks, which he knew did not yet represent an exhaustive inventory. Yet he argued for a smaller number of phyla because he recognised the genetic affinity between certain of these stocks and the distinct nature of others (Klaproth 1823a,b, 1831).

Klaproth was also the first to identify a family of languages comprising Chinese, the Burmese language of 'Awa', the language of the 'Tübeter' and related tongues, but specifically excluding languages such as Siamese, the Vietnamese language of Annam, the 'Moan' language of the 'Peguer' and so forth. Later German proponents of the TB theory had precocious intuitions about Chinese historical grammar. Scholars such as Carl Richard Lepsius (1861) and Wilhelm Grube (1881) mooted reflexes of TB historical morphology in Chinese. Lepsius even recognised that the tones of Chinese had arisen from the loss of older syllable-final segments and the loss of distinctions between older syllable-initial segments. Figure 6.1 shows a schematic view of Klaproth's model.

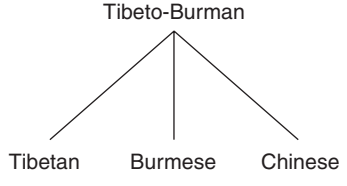


Figure 6.1 One of the language phyla identified by Klaproth (1823) in his polyphyletic view of Asian linguistic stocks.

Yet Klaproth's view of a polyglot Asian continent as the home to many distinct language phyla was not universally well-received. In January 1825, in a letter to Baron Paul Schilling von Canstadt, for instance, August Wilhelm von Schlegel described his distaste for the polyphyletic view of Asia presented by Klaproth (Körner 1930, I: 631), whereas Schlegel evidently found John Leyden's undifferentiated 'Indo-Chinese' view of Asian languages to be more palatable (1832: 21). To scholars in Europe, the two most important language families were what was known in the nineteenth century variously as Indo-European, Indo-Germanic or Aryan, and the Semitic family, later known as Hamito-Semitic and most recently as Afroasiatic. It did not come naturally to everyone to view the many distinct linguistic stocks of Asia as language families on an equal footing with Indo-European and Afroasiatic.

Personalities also played a role, and even the even-keeled Wilhelm von Humboldt made reference to the 'Ätzigkeit' of the brilliant Klaproth (Walravens 1999a). Moreover, between 1826 and 1829, the Société Asiatique in Paris was split apart by the feuding between the group comprising Klaproth, Abel Rémusat, Eugène Burnouf and Julius von Mohl and the 'fleuristes' or 'philologues-poètes', led by the acrimonious Silvestre de Sacy. The lines of animosities drawn in this conflict emanated far beyond Paris. Indeed, the professional perceptions of many a scholar of oriental languages were shaped by the constellation of likes and dislikes which existed between the linguists of the day as much as they were by substantive arguments, and arguably this is to some extent still the case in TB linguistics today. However, in the nineteenth-century personality conflicts also had the effect of exacerbating unstated but deeply rooted Eurocentric preconceptions.

The Indo-Chinese or Sino-Tibetan view

One sally against Klaproth's polyphyletic view of Asian languages was Friedrich Max Müller's Turanian theory, a putative language family encompassing each and every language of the Old World other than the 'Semitic' or Afroasiatic and 'Arian' or Indo-European languages (van Driem 2001). The Turanian view was highly influential in the British Isles and throughout the British Empire and continued to influence scholars after Müller's death in 1900, even though he had himself abandoned the theory in his lifetime.

Another more enduring challenge to the differentiated view of Asian linguistic stocks was originally named 'Indo-Chinese'. Indo-Chinese has a more chequered history than Turanian and still continues to lead a life of its own under the guise of 'Sino-Tibetan'. This view of languages originated with the Scottish physician and poet John Leyden. Leyden's work on 'Indo-Persic' lacked the profundity and erudition of the great Sanskrit scholar Henry Thomas Colebrooke (1765–1837), but his work on 'Indo-Chinese' was published in *Asiatick Researches* in 1808. Leyden's 'Indo-Chinese' encompassed Mon, which he called 'the Moan or language of Pegu', Balinese, Malay, Burmese, 'the Tai or Siamese' and 'the Law, or language of Laos', and Vietnamese or 'the Anam language of Cochin Chinese'. These 'Indo-Chinese' languages of the Asian continent shared a more immediate

genetic affinity with Chinese in Leyden's conception, but Indo-Chinese also explicitly included 'the inhabitants of the Eastern isles who are not immediately [*sic*] derived from the Chinese nations' (1806b: 1). In fact, Indo-Chinese encompassed all the languages spoken by 'the inhabitants of the regions which lie between India and China, and the greater part of the islanders in the eastern sea', which although 'dissimilar', according to Leyden, 'exhibit the same mixed origin' (1806a: 1).

After Leyden's death, the Indo-Chinese idea began to lead a life of its own. In 1837, the American missionary and linguist Nathan Brown used the term 'Indo-Chinese' to designate all the languages of eastern Eurasia. The fact that Brown's Indo-Chinese even included Korean and Japanese illustrates the appeal and dogged longevity of undifferentiated views in the face of more informed opinions. Later versions of Indo-Chinese excluded Japanese and Korean, and the AA languages were recognised as constituting a separate language family by the American Baptist missionary Francis Mason in 1854, when he saw evidence for a specific genetic relationship between the Mon-Khmer language Mon and the Munda language Kol. This newly recognised language family was known as Mon-Khmer-Kolarian for over half a century until Wilhelm Schmidt renamed it AA in 1906. After AA had been removed from Indo-Chinese, German scholars such as Emile Forchhammer (1882) and Ernst Kuhn (1889) continued to refer to what was left of the pseudophylum by the name 'indochinesisch', and in general the same practice was generally observed in the Anglo-Saxon literature. However, a few British scholars, for example, Sir Richard Temple (1903) and George Whitehead (1925), used the term 'Indo-Chinese' in precisely the opposite sense, to designate the AA or 'Mon-Khmer-Kolarian' language family which had been extracted from the expansive pseudophlyum.

After the removal of other phyla, Indo-Chinese had been whittled down to the original TB plus Daic (Figure 6.2 N.B. Daic has been excluded since the Second World War). However, in the confused Indo-Chinese conception, the putative language family consisted of a 'Tibeto-Burman' branch (i.e. the original TB minus Sinitic) and a 'Sino-Daic' branch, for example, August Conrady (1896), Franz Nikolaus Finck (1909). There was residual uncertainty about the genetic affinity of Vietnamese, particularly in the French scholarly community. André-Georges

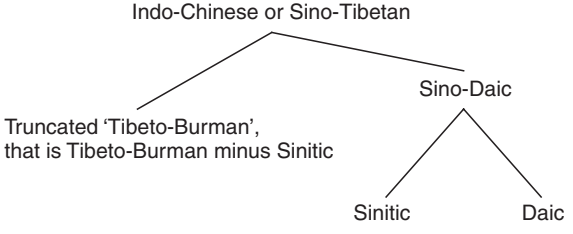


Figure 6.2 The Indo-Chinese or Sino-Tibetan theory.

Haudricourt settled the question once and for all in 1954, and Vietnamese has been universally recognised as AA ever since.

Indo-Chinese was renamed 'sino-tibétain' by Jean Przyluski in 1924, and the name entered the English language in 1931 as 'Sino-Tibetan' when Przyluski and the British scholar Gordon Hannington Luce wrote an etymological note on the 'Sino-Tibetan' root for the numeral 'hundred'. A defining feature of the Indo-Chinese or ST theory, very much at variance with Klaproth's original TB theory, was that Chinese was not seen as a part of TB, whilst Daic was seen as the closest relative of Chinese. In the United States, Alfred Kroeber and Robert Shafer adopted the new term 'Sino-Tibetan' for Indo-Chinese. Chinese scholars similarly adopted the term *Hàn-Zàng* 'Sino-Tibetan', the contours of which are still the same as that of Conrady's 'Indo-Chinese' and Przyluski's antiquated 'Sino-Tibetan'.

Robert Shafer soon realised that Daic did not belong to the Indo-Chinese or ST family and in 1938 'prepared a list of words showing the lack of precise phonetic and semantic correspondence' between Daic and other Indo-Chinese languages. Armed with this list, Shafer travelled to France before the outbreak of the Second World War 'to convince Maspero that Daic was not Sino-Tibetan' (1955: 97–8). Instead, Henri Maspero managed to convince Shafer to retain Daic within ST.

When Paul Benedict moved to Berkeley in 1938 to join Kroeber's ST Philology project, he likewise exchanged the name Indo-Chinese for 'Sino-Tibetan'. Over a century after Klaproth had already identified Daic as a linguistic stock distinct from TB (inc. Chinese), Benedict too in 1942 ousted Daic from 'Sino-Tibetan', but he remained more resolute about this measure than Shafer. The removal of Sinitic from the 'Sino-Daic' branch of 'Sino-Tibetan' resulted in a tree model characterised by the retention of the heuristic artifact that Chinese was a separate trunk of the language family. In fact, this was the sole remaining feature which defined ST as a putative language family and distinguished it from the TB theory. For a brief spate in the 1970s, ST even consisted of a Chinese branch and a Tibeto-Karen construct, which in turn was divided into a Karen branch and an even more mutilated 'Tibeto-Burman' (Benedict 1972, 1976).

The tacit but always untested assumption of Sino-Tibetanists has been that all 'Tibeto-Burman' languages share unitary developments not found in Chinese and Karen. Great significance has been ascribed to superficial criteria such as word order. Though Karen was later put back into truncated 'Tibeto-Burman', adherents of ST have continued to assume the existence of as yet undemonstrated common innovations shared by all TB languages other than Sinitic.

Tibeto-Burman outlives Sino-Tibetan

In the 1990s, the time was ripe for the Indo-Chinese or ST paradigm to be replaced by the original TB theory of Klaproth. Three developments converged to yield insights heralding a return to the TB language family, that is, (1) a better

understanding of OC, (2) improved insights into the genetic position of Sinitic and an appreciation of its TB character, and (3) the exhaustive identification of all the TB subgroups.

The first development involved the production of better reconstructions of OC. Major advances in the historical phonology of Chinese were accompanied by new insights into Chinese historical morphology. New insights on the genetic position of Chinese vindicated Klaproth's and Lepsius' views. By the 1990s, the TB character of Sinitic had been amply demonstrated. In the history of the field no uniquely shared innovations have ever been adduced which could define truncated 'Tibeto-Burman' as a separate coherent taxon that would exclude Chinese and be coordinate with Proto-Sinitic. The new face of OC was of a language with a decidedly TB countenance and more closely allied with certain groups like Bodic and Kiranti. In fact, OC is less remote from the mainstream TB point of view than, say, Gongduk or Toto. A second development is that isoglosses possibly representing lexical innovations as well as uniquely shared morphological innovations in Brahmaputran appear to indicate that a more primary bifurcation in the language family is between subgroups such as Brahmaputran and the rest of the TB family whilst other lexical and grammatical features show that Sinitic is a member of a sub-branch, that I proposed, named Sino-Bodic.

The third development which has heralded a return to the original TB theory is the exhaustive charting of TB subgroups. Only recently have all the languages and language groups of the TB language family been identified with the discovery in Bhutan in the 1990s of the last hitherto unreported TB languages, namely Black Mountain and Gongduk. In addition to the identification of all basic subgroups, new members of already recognised subgroups have been discovered and rediscovered in Tibet, southwestern China, northeastern India and Nepal. In 1999, in an enclave around the shores of lake Ba-gsum or Brag-gsum in northern Koq-po rGya-mdaḥ in Tibet, Nicolas Tournadre identified the language Bag-skad [bɛ̠kɛʔ], spoken by an estimated 3,000 speakers and previously erroneously classified as a Tibetan dialect. Tournadre reports that this tongue is related to Dzala and other east Bodish languages of Bhutan. Similarly, Barām or 'Bhráhmú', a TB language reported by Hodgson in the mid-nineteenth century, but thought since to have gone extinct, was rediscovered in Gorkhā district in Central Nepal in the 1990s.

The basic outline of the TB family is shown in Figure 6.3. The model does not have the shape of a family tree, but this is not to claim that there is no Stammbaum. Not only is the branching pattern of the tree not within view, the constituent language subgroups of the family have only finally exhaustively been identified within the past decade. At present, we do not know the higher-order branching, but we have every reason to believe that these branches are there.

AQ: Map 6.1 not cited in text

This more candid but at the same time more comprehensive view of the language family confronts scholars with the immediate need to search for and identify the evidence which could support empirically defensible higher-order subgroups within TB, analogous to Italo-Celtic and Balto-Slavic in the

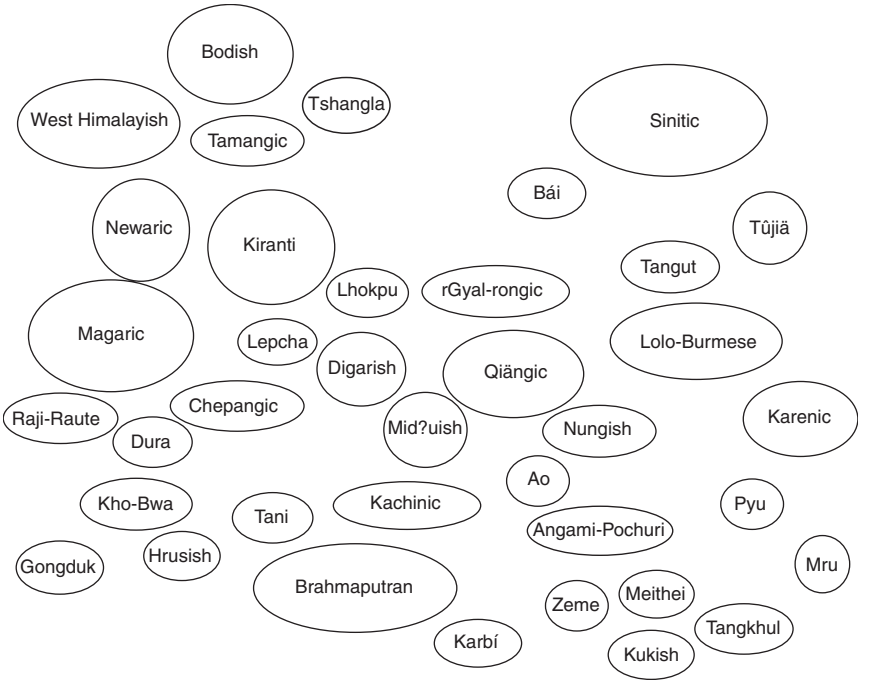


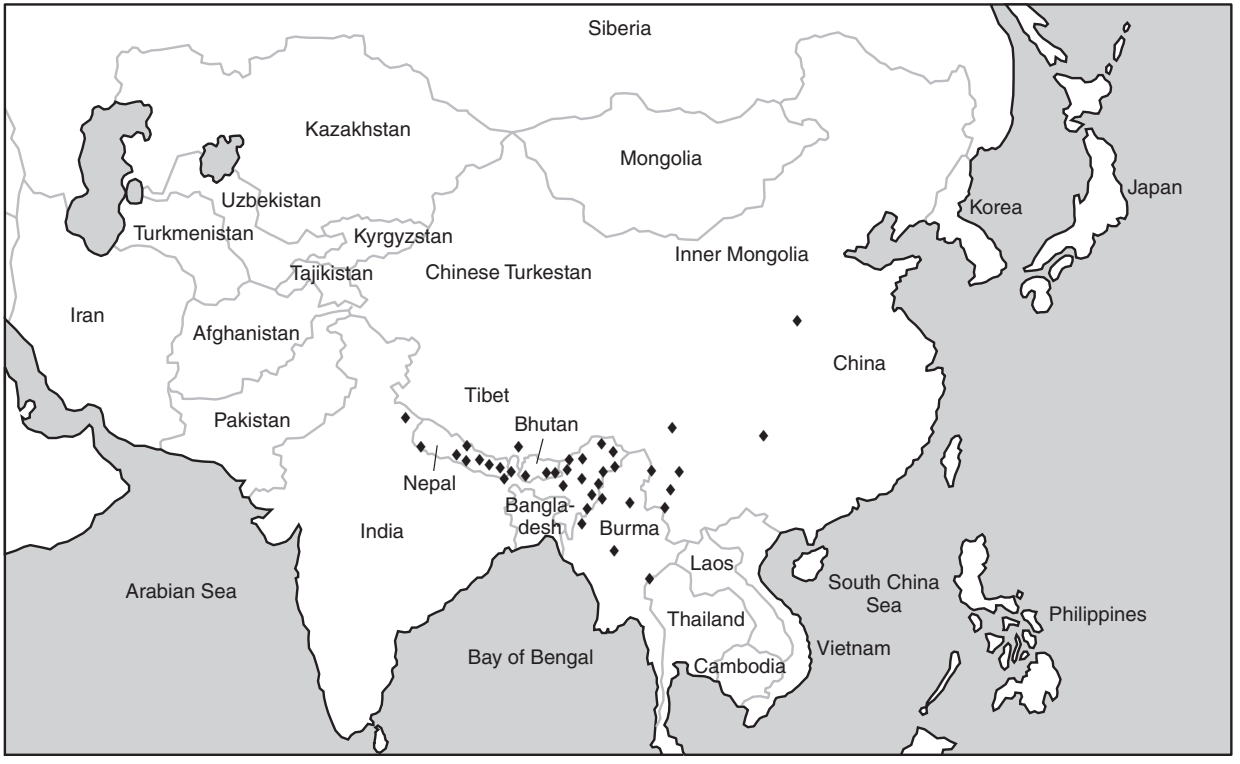
Figure 6.3 Informed but agnostic picture of Tibeto-Burman subgroups.

Notes
The extended version of the Brahmaputran hypothesis includes Kachinic, but for the sake of argument this diagram depicts the short variant of Brahmaputran, namely excluding Kachinic. Kachinic comprises the Sak languages and the Jinghpaw dialects. Likewise, Tangut is separately depicted, although Tangut is likely to be part of Qiāngic. Digarish is northern Mishmi, and Midźuish is southern Mishmi, that is the Kaman cluster. Bái is listed as a distinct group, whereas it may form a constituent of Sinitic, albeit one heavily influenced by Lolo-Burmese. Tǔjiā is a Tibeto-Burman language of indeterminate phylogenetic propinquity spoken in a few villages in northwestern Húnán. The Sino-Bodic hypothesis encompasses at least the groups called Sinitic, Kiranti, Bodish, West Himalayish, rGyal-rongic, Tamangic, Tshangla and Lhokpu and possibly Lepcha. Other hypotheses, such as the inclusion of Chepang and perhaps Dura and Raji-Raute within Magaric, are discussed in van Driem (2001)[1]

Indo-European language family. The burden of proof now lies squarely on the shoulders of Sino-Tibetanists who propagate truncated 'Tibeto-Burman' as a valid taxon to adduce evidence for this construct.

Implications for interpreting prehistory

The Neolithic Revolution and the spread of agriculture are widely thought to have been important factors in the dispersal of ancient populations and the spread of language families. However, the Fertile Crescent itself attests to the fact that agriculture was adopted by ethnolinguistically unrelated populations and that



Map 6.1 In this clutch of 39 diamonds, each diamond represents not a language, but the historical geographical centre of a primary taxon or subgroup of languages of the Tibeto-Burman family. In order to present a fair picture of the internal diversity of the Brahmaputran branch, the Dhimalish, Bodo-Koch and Konyak subgroups have each been represented by a diamond. Likewise, two separate diamonds indicate Kiranti and Newaric, the two constituent subgroups within the hypothetical and internally highly diverse Mahakiranti branch. The extinct Tangut language, however, is treated as a member of Qiāngic.

agriculture spread effortlessly across ethnolinguistic boundaries without affecting them in any significant way. Sumerian, Elamite, Akkadian, Hurrian, Hattic and other languages of early agricultural civilisations which have left no surviving linguistic descendants bear witness to the permeability of linguistic boundaries for the dissemination of agriculture. The Neolithic and Bronze Age of Asia Minor and Mesopotamia9 is characterised by a very long period of incursive population movements into, rather than out of Anatolia and the Fertile Crescent, driven or lured, it seems, by the relative affluence of urban centres supported by agricultural surplus.

Those who secondarily adopt a technique, tradition or cultural institution often improve upon it and excel in its exploitation beyond the attainments of its original innovators. In Dutch this is known as *de wet van de remmende voorsprong*, that is, the 'law' that the very group which has managed to get ahead of other groups by virtue of an innovation is also more prone to get bogged down at a later stage by shortcomings inherent to the prototypical version of the technology which originally gave them the edge over other groups. Meanwhile, other groups who did not have to invest the resources and effort to develop and implement the technology in the first place forge ahead by introducing a more refined and streamlined version of the innovation and are unhampered by having to replace or revamp an obsolete infrastructure. O'Connor (1995) and Blench (Chapter 2, this volume) have argued that irrigated rice agriculture in the Southeast Asian lowlands does not correlate with a spread at the language family level, but with spreads at a lower phylogenetic level.

By contrast, perhaps what the incursive Indo-Europeans did may have been nothing other than land-theft. Nevertheless, the spread of specific, well-defined Neolithic cultural assemblages remains a powerful tool in the reconstruction of ancient population movements and, more particularly, in the possible early dispersal of language families. The hypothesis that an agricultural dispersal may reflect the ancient spread of a language community underlies my reconstruction of the spread of the Sino-Bodic branch of TB (van Driem 1998, 1999, 2001, 2002). Yet the incentive for migration into affluent regions with an agricultural surplus is a factor to be reckoned with in TB prehistory too. The distribution of primary branches of TB suggests that it may be that the urban affluence of pre-TB agricultural populations was what drew the linguistic ancestors of early Sinitic civilisation to the Yellow River and North China Plain in the first place, just as Gutaeans, Kassites, Amorites and Indo-Europeans were drawn to the Fertile Crescent and Anatolia. Benedict once proposed that the Shāng may not have been Sinitic at all and that the Zhōu, who came from the West, may have been the bearers of the Proto-Sinitic language to the Yellow River basin, where they adopted the Shāng ideograms devised by a pre-TB population (1972: 197). In fact, the prosperous agricultural civilisation on the North China Plain may have lured the linguistic forebears of Sinitic, or perhaps Sino-Bodic, long before the Shāng period.

Quite often the archaeological record may not directly reflect such linguistic intrusions. Instead, rather than reflecting the spread of language families, archaeology

shows the regional discrepancies in technical advancement which may have motivated foreign linguistic intrusions. In particular, this may apply in the case of the early displacement of Sinitic outside of the TB core area as well as in the case of the advent of Indo-European groups to the Near East, such as the Hittites in Anatolia and the Mitanni in the Jazīrah. Not only did agriculture spread across linguistic boundaries from the very outset, the direction of linguistic intrusions in many episodes of prehistory may have been diametrically opposed to the direction of the spread of agriculture.

My reconstruction is based on a family tree model of TB, which presumes a clustering of groups and suggests a relative chronology. Yet, the model is not purely a Stammbaum as such. The problem with the TB family tree models proposed to date is that uniquely shared innovations are scarce, and higher-level subgroups are often defined by what later turn out to be shared retentions. The family tree in Figure 6.4 is not just a geographically inspired schema, for it incorporates subgroups which were discerned by Shafer and are still recognised on the basis of phonological and morphological criteria and lexical isoglosses. The model also incorporates Sino-Bodic, a higher-level subgrouping hypothesis involving Sinitic and those languages within TB which appear to be more immediately related to Sinitic than either are to, for example, Brahmaputran, Karbí and other genetically remote groups.

Although Sino-Bodic is associated with me (van Driem 1995, 1997), earlier versions of the Sino-Bodic hypothesis had previously suggested themselves to Walter Simon (1929), Robert Shafer (1955, 1966, 1967, 1968, 1974) and Nicholas Bodman (1980), on the basis of uniquely shared lexical items. In addition to the limited set of lexical isoglosses, I have described morphological features that

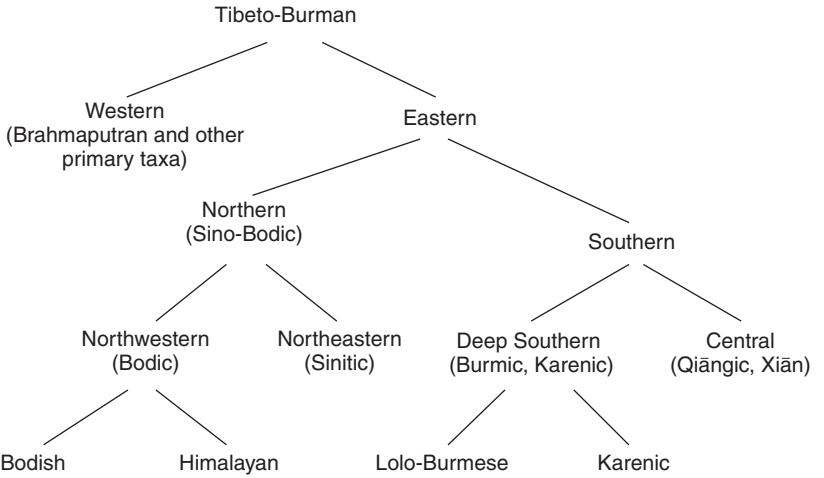
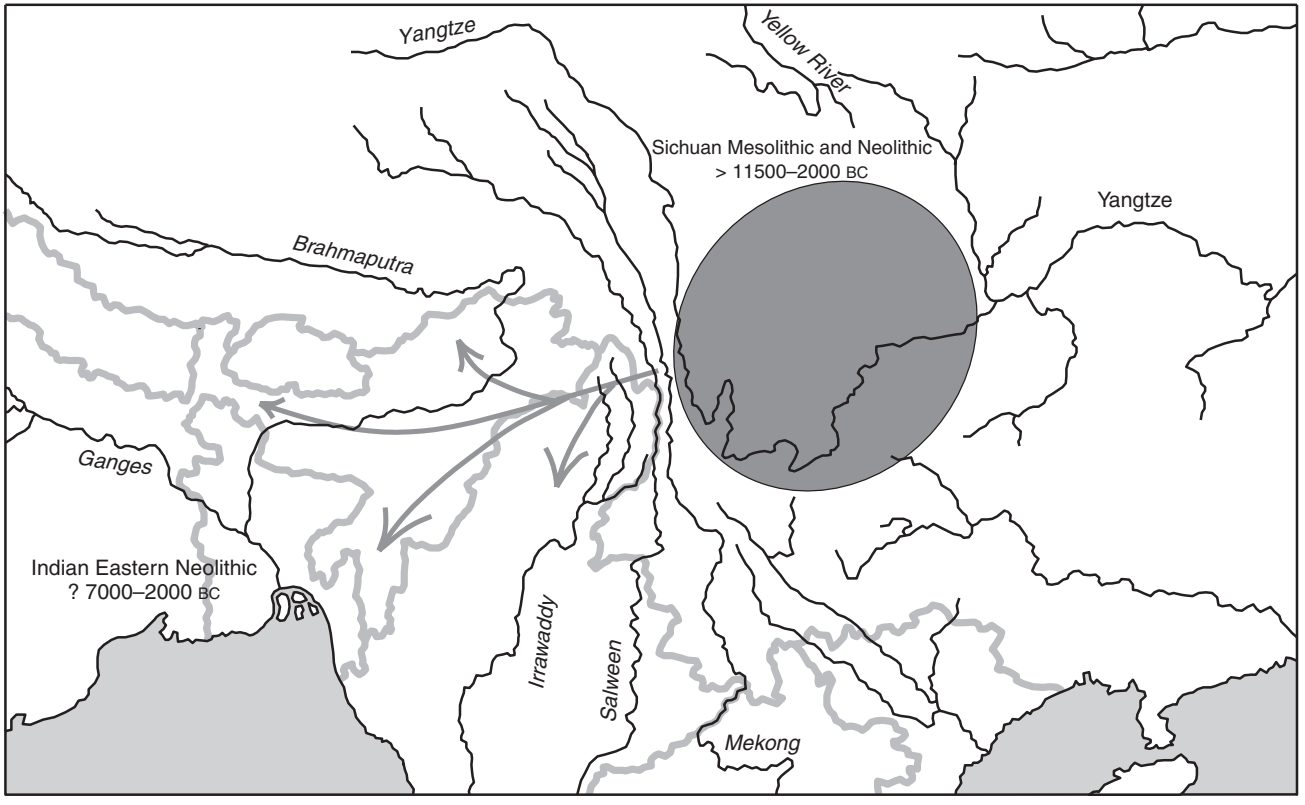


Figure 6.4 Linguistically inspired archaeological interpretation of the geographical dispersal of Tibeto-Burman groups.

appear to bolster the identification of Sino-Bodic as a subgroup (van Driem 1997). By contrast, the constellation of subgroups which I collectively name Western TB represents a number of primary branches which I assume had split off at an early stage and settled in northeastern India, originating from a TB proto-homeland which I locate in Sìchuān, as British scholars in the nineteenth century had already proposed, even though they did not have access to modern-day linguistic, archaeological and genetic evidence. Here I shall briefly outline the model again and adduce additional supporting arguments from recent research on haplotypes on the Y chromosome. I shall also point out linguistic and archaeological weaknesses in the model, which leave room for an alternative version of the reconstructed linguistic dispersal.

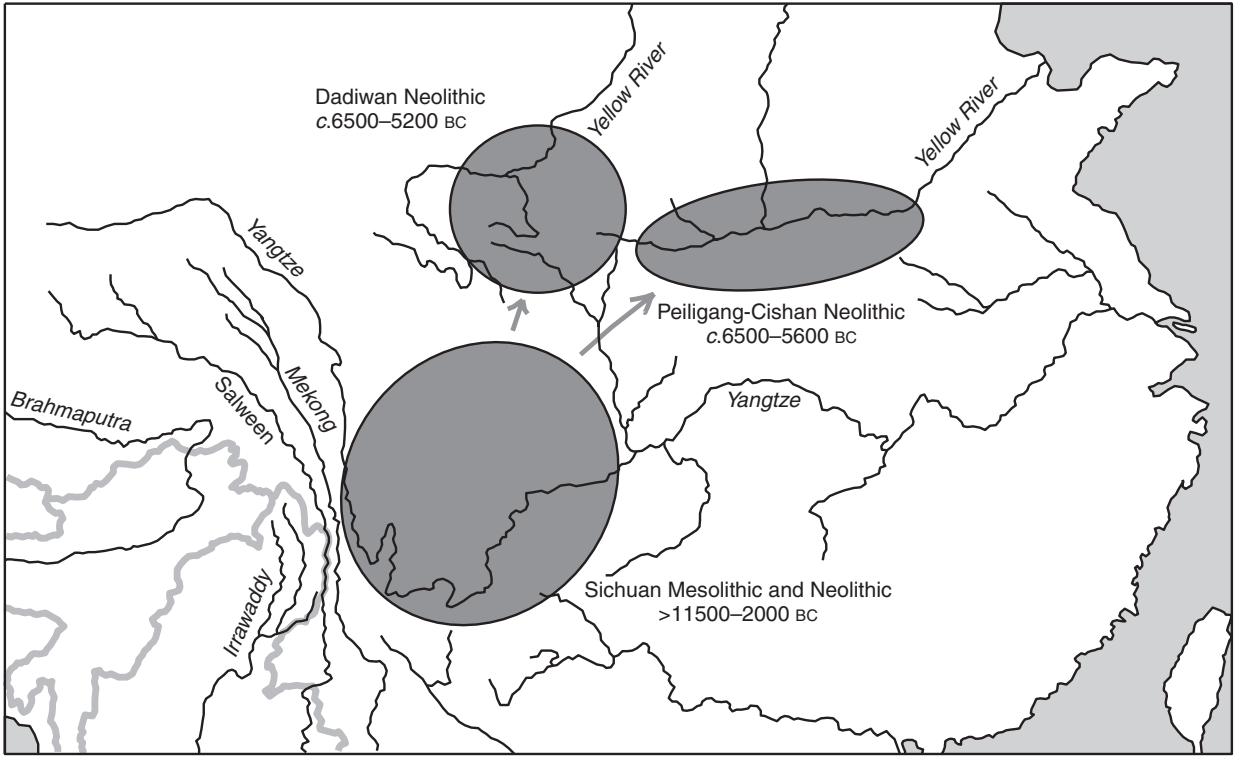
Though primarily linguistically inspired, my theory represents an interpretation of the archaeological record in light of TB subgrouping hypotheses and the geographical distribution of modern and historically attested communities. The theory depicted schematically in Figure 6.4 is illustrated in Maps 6.2–6.5. The differences between Figure 6.3 and Figure 6.4 illustrate the linguistic and the archaeological view between which some correlation is sought. Western TB in particular is not just a linguistic hypothesis, but an archaeological theory about the population history of the TB area informed by linguistic insights about the primary nature of subgroups in the Himalayas and northeastern India. From a phylogenetic perspective, Western TB is analogous to the Formosan language groups within An. Like Formosan, Western TB is not a single taxon, but a collection of primary taxa within the family. Rather, it is the remaining branch, Eastern TB, which may constitute a possible genetic unit, just as MP is a single primary branch within An. It is therefore more fitting to speak of an Eastern than of a Western TB hypothesis, if there is such a thing as the latter. Brahmaputran is just one of the many taxa collectively referred to as Western TB. The short variant of Brahmaputran consists of the Dhimalish, Bodo-Koch and Konyak subgroups, and the extended version of the Brahmaputran hypothesis includes Kachinic, that is, the Sak languages and the Jinghpaw dialects. Some other Western TB taxa in the northeast of the Subcontinent include the Kho-Bwa cluster, Hrusish, Midźuish, Nungish, Digarish, Tani, Karbí, Ao, Angami-Pochuri, Zeme, Tangkhul and Gongduk.

The various ways of reconstructing prehistory, that is, archaeology, linguistics and genetics, measure three independent quantities which are merely probabilistically correlated and which, moreover, may divide into taxa which may correspond to quite different time depths. Discrepancies between the chromosomal and the linguistic pictures of the past indicate that, in some cases, a larger incursive population may have adopted a language of a smaller population already resident in the area which they had settled, such as the case of Bulgarian, whereas some languages borne by ruling élites have been adopted by a larger dominated resident population, such as the case of Hungarian. The racial heterogeneity of TB populations in northeastern India, particularly the phenotypic difference between Brahmaputran language communities and other TB groups in the northeast, has been noted ever since the earliest British accounts of the area.

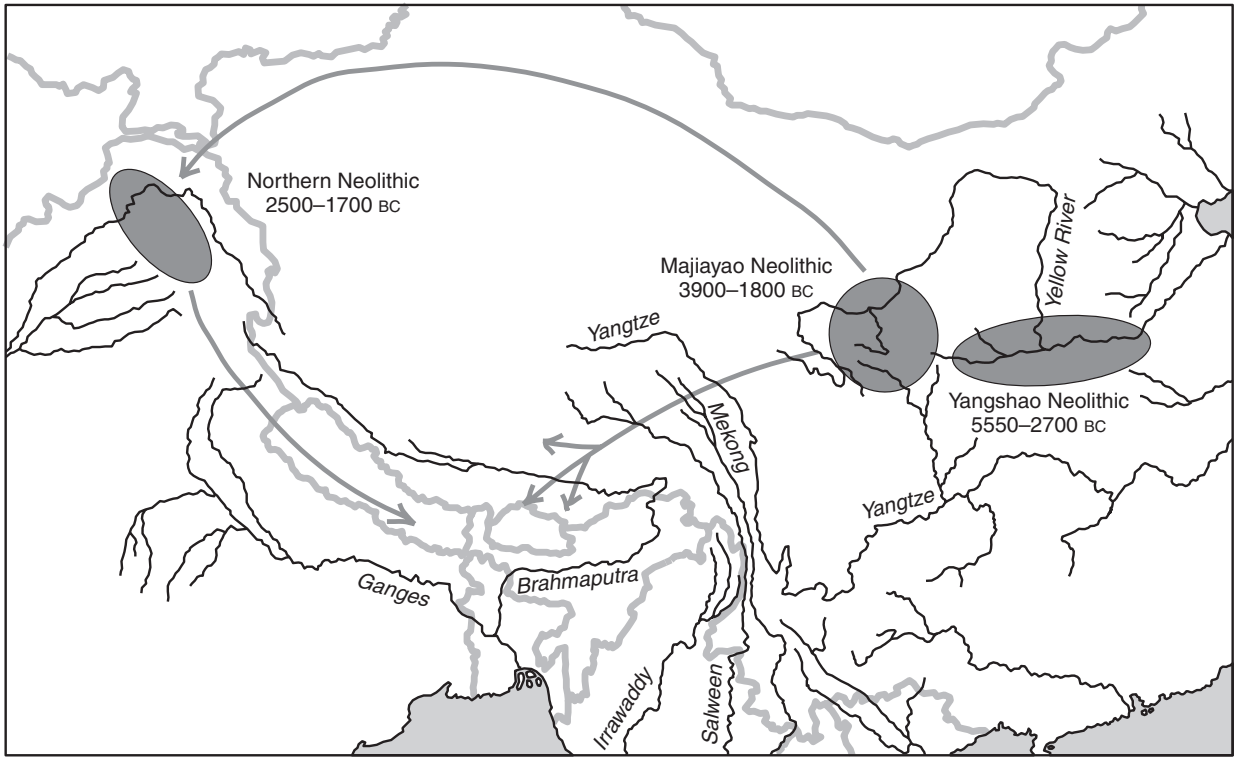


: Please
ck '?'
bol

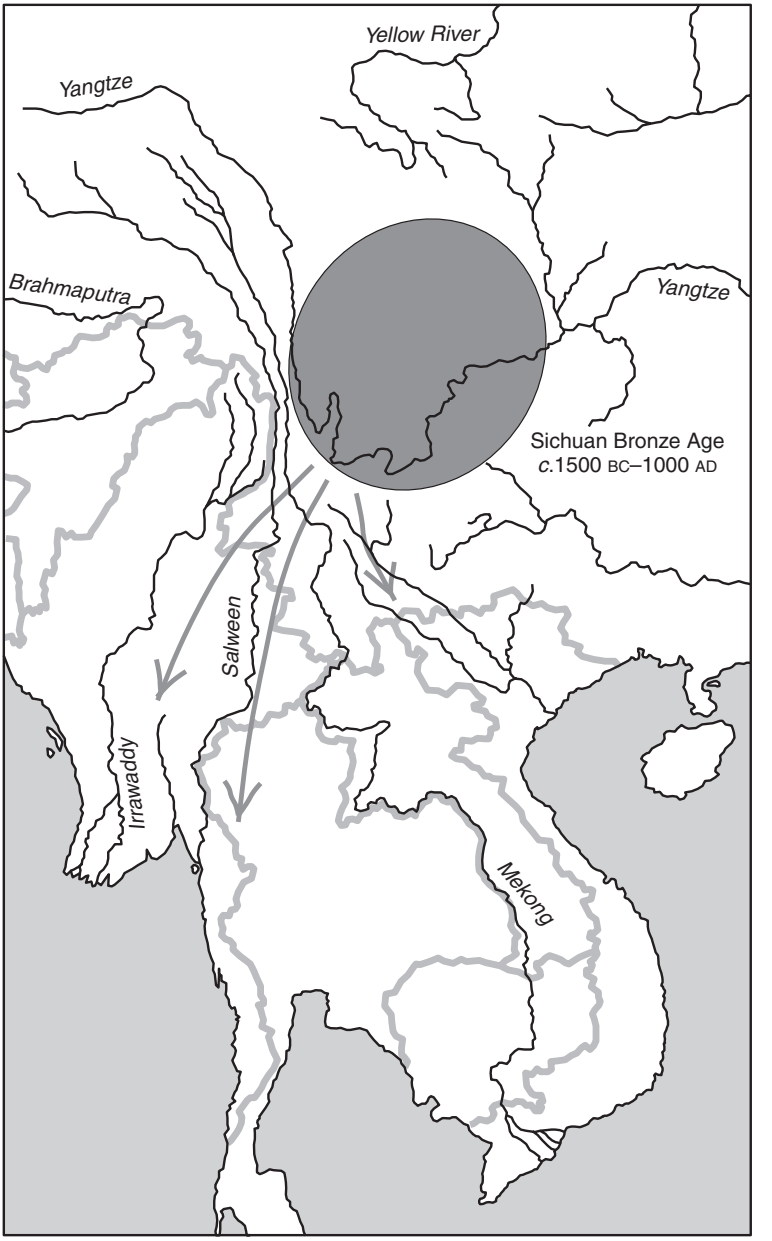
Map 6.2 Lower Brahmaputra basin and surrounding hill tracts colonised by Western Tibeto-Burmans bearing the technologies from Sìchuān which were to become known as the Indian Eastern Neolithic, an *Auswanderung* possibly set in motion before the seventh millennium BC.



Map 6.3 The establishment of the early Neolithic Péilígǎng-Císhān and Dàdìwān civilisations in the Yellow River basin by Northern Tibeto-Burmans before the beginning of the sixth millennium BC.



Map 6.4 One offshoot of the late Neolithic Mǎjiāyáo cultural complex migrates South through northern Sìchuān and eastern Tibet into Sikkim, whereas another offshoot migrates to the southwest across the Himalayas to establish the northern Neolithic civilisation in Kashmir. Northwestern Tibeto-Burmans peopled the Himalayas, both from the northeast, colonising Sikkim and Nepal, and from the west, colonising the western Himalayas and the Tibetan plateau.



Map 6.5 The exodus of deep southern Tibeto-Burmans into peninsular Southeast Asia had begun by the first millennium BC, and the process seems never to have completely come to a halt, as Lolo-Burmese groups have continued to trickle into Thailand from Yúnnán in recent history.

Modern genetic studies occasionally corroborate old theories of population history which were exclusively inspired by, and based on, language and old-fashioned racial somatology. For example, Basu *et al*. (2001) recently studied haplotype frequencies of (CTG)n repeat and three other biallelic markers in and around the myotonic dystrophy locus in 13 ethnolinguistically and geographically diverse populations of India. Their findings support the traditional ethnographic conception that certain tribal groups such as the AA Lodhas and Santhal represent 'the most ancient inhabitants' of the Subcontinent and may be identified as the 'descendants of modern humans who arrived in India on one of the early waves of "out-of-Africa" migration' (2001: 316, 317). Likewise, in keeping with the conceptions of traditional Indian ethnography, their findings suggest that tribal populations have 'remained relatively more isolated than the caste populations', and 'the boundaries of caste populations, especially those of middle and lower ranks, have been more fluid than those of tribal populations' (2001: 316).

Until recently the state of the art was such that the interpretation of the chromosomal picture using classical markers sometimes only provided a limited glimpse of events in prehistory in the absence of supporting archaeological or linguistic evidence. But a spectrum of markers is now available which ranges from slowly evolving biallelic markers to rapidly evolving minisatellites. Binary haplotypes with very low mutation rates represent unique event polymorphisms which occurred at large intervals in human evolution. These are known as 'unique mutation events', abbreviated UME, and include the single-base pair substitutions described by Underhill *et al*. (1997, 2001). By contrast, some rapidly evolving loci on the Y chromosome, such as the minisatellite locus *MSY1* studied by Jobling *et al*. (1998), exhibit a mutation rate of between 2 per cent to 11 per cent per generation. Intermediate between these two extremes are markers which evolve with moderate rapidity, such as Y chromosome microsatellite loci known as short tandem repeats, abbreviated STR, which Kayser *et al*. (2001) have shown to be a powerful tool in reconstructing population history. Though still problematic in some respects, the findings of studies on these different types of polymorphisms allow statistical analyses which may be of some utility in evaluating competing models of the peopling of Eurasia reconstructed on the basis of linguistic and archaeological evidence.

Any model of TB prehistory will have to account for the racial affinities of some Western TB groups, for example, Toto, Raji, Raute, Dhimal and some other Brahmaputran groups. The intriguing racial variation of TB and non-TB groups in the Subcontinent, already evident to earlier generations of ethnographers, is being charted in greater detail by current population genetic studies, such as those currently being conducted by Peter de Knijff and myself in the Himalayan region. Both the collection of genetic samples and the interpretation of the results must be conducted in an ethnolinguistically informed way.

In this context, two apparently conflicting sets of findings have recently been obtained by teams of geneticists looking at TB populations in China and the Greater Himalayan region. Yet, the discrepancy between these findings may be

more apparent than real, and may very well correspond to different realities situated at different time depths. The hypothesis of a TB homeland in Sìchuān has recently found unexpected corroboration in the findings of the Chinese Human Genome Diversity Project, whose ethnolinguistically informed assays of population groups in China have shown that genetically East Asian populations can be derived from Southeast Asian populations and that, therefore, populations ancestral to the Chinese may not have originated in the Yellow River basin but could have migrated to this area in a northeasterly direction from southwestern China (Chu *et al*. 1998). This information was still unavailable when I first proposed that the Tibeto-Burman homeland lay in Sìchuān on linguistic grounds.

Another team of geneticists has found a strong genetic affinity amongst population groups of the TB language family in the form the prevalence of a T to C mutation at Y chromosome locus *M122*, whereas the extremely high frequency of H8, a haplotype derived from *M122C*, reflects the results of a genetic bottleneck effect that occurred during an ancient southwesterly migration (Su *et al*. 2000). The latter group of geneticists attempted to relate the geographical distribution of TB populations with a migration from the middle Yellow River basin about 10,000 BP, and to conjecture that the earliest Neolithic cultures of this area might have been associated with the putative TB homeland. However, there are two flaws in this interpretation.

First of all, the study by Su *et al*. (2000) sampled only six populations from the pivotal, ethnolinguistically most heterogeneous TB heartland in northeastern India. The samples from this area were limited to a 'Kachari' individual, a Rabha, a Naga, an Adi, a Nishi and an Apatani. Their study left most key TB population groups untouched. Conjectures were advanced about prehistoric migrations to the Himalayas, but, other than the three sample populations from Arunachal Pradesh, no Himalayan populations were tested. Fifteen samples, constituting half of the test material, were obtained from individuals representing Hàn Chinese populations settled in various provinces of China. The remaining samples were from several TB populations resident in China, that is, Nakhi, Bái, Yí, Jīnuò, Jinghpaw, Yúnnán Lahu and Tǔjiā. Finally, there were two Tibetan samples, one from Lhasa and one from Yúnnán, and a single Karen sample from Southeast Asia. The assay was therefore limited and did not sample most of the key TB language communities in the Himalayas about whose ancestors inferences were made. The second problem is that the interpretative framework was based on the phylogenetic model presented by Matisoff (1991), in which an Indo-Chinese or 'Proto-Sino-Tibetan' Ursprache at its deepest time depth is presumed to have split east–west into 'Proto-Chinese' and 'Proto-Tibeto-Burman'. Problems with this model have been discussed earlier.

At a far greater time depth, ethnolinguistically informed assays of the population of eastern Asia on the basis of 30 microsatellites made by Chu *et al*. (1998) have shown that the ethnolinguistic composition of China is reflected in the genetic complexity, and that the peopling of eastern Asia probably occurred in

a northward movement from Southeast Asia. These results have been corroborated in a study of 19 biallelic loci on the Y chromosome, which demonstrated that northern populations in eastern Asia only represent a subset of the haplotypes found in southern populations, which show greater polymorphism on the whole than northern populations (Su *et al*. 1999).

Craniometric and skeletal evidence is still routinely used by archaeologists and palaeontologists to reconstruct population history. For example, Brown (1998) and Demeter (2000) argue for major morphological changes in population in the far East between various phases of the post-Pleistocene or between the Mesolithic and Neolithic periods. Hopefully, it will be possible in future to make such findings square with the new insights of genomic studies. Particularly in view of the phenotypic variation sometimes observed within single populations, it will hopefully be undertaken to extract DNA from such crania for study. Recent work by Ding *et al*. (2000) has also shown that northern and southern haplotype clusters blend across a cline without any abrupt change, so that no clear genetic support has yet been identified that might corroborate linguistic theories connecting Chinese to Caucasian, for example, the Sino-Caucasian theory advocated by Starostin, or connecting Chinese genetically with Indo-European, as Pulleyblank does. Yet all these investigations have merely scratched the surface of a vast terrain which lies to be charted and have begun to make possible an integrated vision of the genetic, linguistic, historical, archaeological and anthropological data.

Three arguments support the identification of Sìchuān as the TB homeland. The first is the centre of gravity argument based on the present and historically attested geographical distribution of TB language communities. Sìchuān encompasses the area where the upper courses of the Brahmaputra, Salween, Mekong and Yangtze run parallel to each other within a corridor just 500 km in breadth. The second argument is that archaeologists identify the Indian Eastern Neolithic, associated with the indigenous TB populations of northeastern India and the Indo-Burmese borderlands, as a Neolithic cultural complex which originated in Sìchuān and spread into Assam and the surrounding hill tracts of Arunachal Pradesh, the Meghalaya, Tripura, the Mizoram, Manipur, Nagaland and Chittagong before the third millennium BC (Dani 1960; Sharma 1967, 1981, 1989; Thapar 1985; Wheeler 1959).

Archaeologists have estimated the Indian Eastern Neolithic to date from between 10,000 and 5,000 BC (Sharma 1989; Thapar 1985). If these estimates are taken at face value, it would mean that northeastern India had shouldered adzes at least three millennia before they appeared in Southeast Asia. Whilst some archaeologists may give younger estimates for the Indian Eastern Neolithic, a solid stratigraphy and calibrated radiocarbon dates are still unavailable for this major South Asian cultural assemblage. The Indian Eastern Neolithic appears intrusively in the northeast of the subcontinent and represents a tradition wholly distinct from the other Neolithic assemblages attested in India. Assuming that the Indian Eastern Neolithic was borne to the Subcontinent by ancient TBs, then if the younger estimates for this cultural assemblage can be substantiated by solid

dating, the linguistic fracturing of subgroups would have to have occurred earlier in Sìchuān before the migrations, as I had suggested previously (1998, 2001).

The third argument for a TB homeland in Sìchuān is that archaeologists have argued that southwestern China would be a potentially promising place to look for the precursors of the Neolithic civilisations which later took root in the Yellow River Valley (Chang 1965, 1977, 1986, 1992; Chêng 1957). The Dàdìwān culture in Gānsù and Shǎnxī, and the contiguous and contemporaneous Péilígǎng-Císhān assemblage along the middle course of the Yellow River share common patterns of habitation and burial, and employed common technologies, such as hand-formed tripod pottery with short firing times, highly worked chipped stone tools and non-perforated semi-polished stone axes. The Dàdìwān and Péilígǎng-Císhān assemblages, despite several points of divergence, were closely related cultural complexes, and the people behind these civilisations shared the same preference for settlements on plains along the river or on high terraces at confluences. Whereas the Sìchuān Neolithic represented the continuation of local Mesolithic cultural traditions, the first Neolithic agriculturalists of the Dàdìwān and Péilígǎng-Císhān cultures may tentatively be identified with innovators who migrated from Sìchuān to the fertile loess plains of the Yellow River basin. The technological gap between the earlier local microlithic cultures and the highly advanced Neolithic civilisations which subsequently come into flower in the Yellow River basin remains striking. Yet a weakness in this third argument lies in the archaeological state of the art. Just as it is difficult to argue for a possible precursor in Sìchuān in face of a lack of compelling archaeological evidence, neither can the inadequate state of the art in Neolithic archaeology in southwestern China serve as an argument for the absence of such a precursor.

Moreover, agricultural dispersals and linguistic intrusions may be distinct issues altogether. The concentration within a contiguous geographical region of all major high-order TB subgroups other than Tǔjiā and Sinitic constitutes a linguistic argument for an early TB linguistic intrusion into the area that today is northern China. If the Dàdìwān culture in Gānsù and Shǎnxī, and the contiguous Péilígǎng-Císhān assemblage along the middle course of the Yellow River are indeed primary Neolithic civilisations, then the eccentric location of Sinitic and Tǔjiā may even trace the route of the early migration out of TB homeland to the affluent and more technologically advanced agricultural societies in the Yellow River basin. In other words, since the linguistic evidence puts the TB heartland in southwestern China and northeastern India, an archaeological precursor in Sìchuān for the Dàdìwān and Péilígǎng-Císhān cultures would fit the hypothesis that the displacement of Sinitic to northern China was the result of an early TB archaeological dispersal. The absence of any such precursor in Sìchuān would fit a theory of early migration from the northern end of the ancient TB dialect continuum to the affluent areas of pre-TB agricultural civilisations along the Yellow River.

I collectively refer to the ancient TB populations, who either bore with them from Sìchuān to the loess plateau the technologies of polished stone tools and

cord-marked pottery or were enticed to the loess plateau by the affluence of the technologically more advanced agricultural civilisations there, as 'Northern Tibeto-Burmans'. I identify these Northern TBs as the likely linguistic ancestors of the Sino-Bodic groups. Subsequent technological developments were both innovated and introduced comparatively rapidly in the North, whereas relatively egalitarian small-scale agricultural societies persisted in southwestern China until the Bronze Age. This hypothesis places the split between Northern and Southern TB in the seventh millennium BC, just before the dawn of the Dàdìwān and Péilígǎng-Císhān civilisations.

I identify the spread of Bodic groups from Gānsù with the dispersal of the Mǎjiāyáo and Yǎngsháo Neolithic cultures and the cultivars broomcorn millet (*Panicum miliaceum*) and foxtail millet (*Setaria italica*), first domesticated on the North China Plain, into the Himalayan region in the third millennium BC. Sino-Bodic would have split up into Sinitic and Bodic before this date. This dispersal proceeded along two routes. The Mǎjiāyáo Neolithic culture spread westward along the main ancient Inner Asian trade route across the Himalayas to establish the genetically related Northern or Kashmir Neolithic in Kashmir and Swāt. At the same time, the Mǎjiāyáo cultural assemblage spread southward from Gānsù through eastern Tibet into southeastern Tibet, Bhutan and Sikkim to establish the Neolithic cultures of Chab-mdo and northern Sikkim, both of which have been identified as colonial exponents of the Mǎjiāyáo Neolithic. Moreover, these colonial exponents make their appearance in Kashmir, eastern Tibet and Sikkim in the second half of the third millennium BC, so that the final phase of these movements coincides precisely with the Banshan phase of the Mǎjiāyáo cultural assemblage, which covers the period between 2,200 and 1,900 BC and is characterised by a marked geographical contraction of the original Mǎjiāyáo core territory.

My reconstruction of TB dispersals, presented in greater detail elsewhere (van Driem 1998, 1999, 2001), is outlined here in Maps 6.2 to 6.5. On the whole, this reconstruction still fits the known facts well. Yet the weaknesses in this model must be recognised. First of all, Sìchuān and southwestern China in general remains archaeologically inadequately researched, despite the significance of the area's prehistory. A second problem is that the linguistic state of the art gives us no real relative chronology for the splitting off of the main taxa of the language family, as shown in Figure 6.3. None the less, the sheer number of major language groups in the Himalayan region and the northeast of the Subcontinent provides a good idea of where and when it would be most fruitful to look for likely archaeological correlates for the dispersal of ancient TB populations. The lopsided geographical distribution of most major TB groups in the Himalayas and northeastern India, the likely linguistic affinity of Sinitic with Bodic, and the possible affinity of 'Deep Southern' with 'Central' Tibeto-Burman groups have inspired the tree schema outlined in Figure 6.4.

An alternative proposal to a TB homeland in Sìchuān would be to identify the earliest Neolithic cultures along the Yellow River basin and on North China Plain with the TB homeland. However, if the TB homeland were to have lain in the

Yellow River basin, then we would be hard pressed to find a plausible archaeological correlate for the spread of Brahmaputran language communities, which once extended beyond Assam and the Meghalaya and formerly covered much of the area that is now Bangladesh and West Bengal. Furthermore, it must be kept in mind that the early Neolithic civilisation on the Yellow River is distinct from the cultural assemblages of the middle Yangtze basin, the succeeding stages of which ultimately spread as far afield as Oceania in the course of the millennia. Both the Yellow River and the middle Yangtze civilisations represent ancient agricultural societies nearly as old as those of the Fertile Crescent.

Clearly, the first and foremost desiderata are that the archaeology of Sìchuān and northeastern India be better understood, that a fine-grid and ethnolinguistically informed genome study of the greater Himalayan region be carried out, and that a new look be taken at subgroups within TB, whereby the same methodological rigour of sound laws and shared innovation is applied which has characterised Indo-European studies. My reconstruction of TB language dispersals will remain sensitive to revision and modification based on new data and new insights.

An intriguing theory involving a remote linguistic relationship with TB is the Sino-Austronesian theory proposed by Laurent Sagart (1994, 2001 and this volume) connecting TB with An. Because Sagart initially recognised possible Sino-An correspondences in Chinese material more than in TB, he was originally inclined to identify the Sino-Austronesian unity with the Lóngshān cultural horizon. However, there is an alternative way of viewing the Sino-Austronesian evidence and the archaeological record. The Lóngshān coastal interaction ensued upon a northward expansion of PAN or Austro-Tai culture from its ancient homeland in southern and southeastern China, and this northward expansion of early Ans would have brought them into contact with early Northern TBs. The ensuing contact situations between An and the Sino-Bodic branch of TB could have involved the ancient exchange of vocabulary between the two language families. The way to test this would be to determine whether items shared by An and TB are indeed limited to the Sino-Bodic branch of TB, including rice terms such as Malay beras and Tibetan ḥbras, a correspondence already pointed out by Hendrik Kern in 1889. The Lóngshān interaction sphere is an obvious candidate in terms of time and place for early contacts between ancient Ans and ancient Tibeto-Burmans, particularly the Dàwènkǒu Neolithic of Shāndōng with its well-established ties both with the other coastal cultures of the Lóngshān interaction sphere as well as with the ancient Northern TB Lóngshān Neolithic civilisation.

However, the archaeological record presents earlier possible correlates for contact between ancient Daic or Austro-Tai and ancient Northern TB culture. For one, impressions of rice contained within the walls of ceramic vessels from the sixth millennium BC indicate that the Yǎngsháo Neolithic maintained some degree of interaction with the probably Daic rice-cultivating civilisations south of the Qínlǐng mountains along the Yangtze. However, the first reported instance of recovery of actual rice remains in the Yellow River basin dates from the beginning of the second millennium BC, associated with the Lóngshān culture of Hénán,

though some rice impressions found on potsherds would appear to be of earlier date (Wú 1996). A much later candidate for an archaeological reflection of intense interaction between ancient Northern TBs on the Yellow River and ancient Daic peoples on the middle Yangtze, some time after the Lóngshān horizon, is the Qūjiālǐng and Shíjiāhé culture, which expanded from the middle Yangtze into peripheral regions rapidly and on a grand scale, even replacing the Yǎngsháo culture in southern and southeastern Hénán in the middle of the third millennium BC (Chang 1996).

Abbreviations

AA Austro-Asiatic
An Austronesian
MP Malayo-Polynesian
OC Old Chinese
ST Sino-Tibetan
TB Tibeto-Burman

Notes

1 Jackson Sun (Sūn Tiānxīn) of the Academia Sinica argues that Guìqióng, spoken in west-central Sìchuān (cf. van Driem 2001: 498), may represent a separate subgroup in its own right, whereas Sūn Hóngkāi of the Chinese Academy of Social Sciences suspects that Guìqióng is a Qiāngic language heavily influenced lexically and phonologically by its Lolo-Burmese neighbours. Conversely, Sūn Hóngkāi believes that Báimǎ, spoken in central northern Sìchuān, is a separate Tibeto-Burman subgroup which has previously been misidentified as a Tibetan dialect, whereas Jackson Sun believes it is a Tibetan dialect. Sūn and Sūn agree, however, that the solutions to the controversy will only come through the detailed analysis and documentation of both languages. Only linguistic field work leading to the detailed description of undocumented Tibeto-Burman languages will render possible the comparative work which will enable us to build a tree of genetic subgroup relationships.

Bibliography

Basu, P., Majumder, P., Roychoudhury, S. and Bhattacharya, N. (2001) 'Haplotype analysis of genomic polymorphisms in and around the myotonic dystrophy locus in diverse peoples of India', *Human Genetics* 108, 4: 310–17.

Benedict, P. (1942) 'Thai, Kadai, and Indonesian: a new alignment in southeastern Asia', *American Anthropologist* 44: 576–601.

——(1972) *Sino-Tibetan: A Conspectus*, Cambridge: Cambridge University Press.

——(1976) 'Sino-Tibetan: another look', *Journal of the American Oriental Society* 96, 2: 167–97.

Blench, R. (2001) 'From the mountains to the valleys: understanding ethnolinguistic geography in Southeast Asia', paper presented at the Colloque *Perspectives sur la Phylogénie des Langues d'Asie Orientales* at Périgueux, France, 30 August.

Bodman, N. (1980) 'Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship', in Frans van Coetsem and Linda R. Waugh (eds) *Contributions to Historical Linguistics: Issues and Materials*, 34–199, Leiden: E.J. Brill.

Brown, N. (1837) 'Comparison of Indo-Chinese languages', *Journal of the Asiatic Society of Bengal* VI, 2: 1023–38.

Brown, P. (1998) 'The first Mongoloids: another look at Upper Cave 101 and Minatogawa 1', *Acta Anthropologica Sinica* 17: 260–75.

Chang, K.C. (1965) 'Relative chronologies of China to the end of Chou', in Robert W. Ehrich (ed.) *Chronologies in Old World Archaeology*, 503–26, Chicago: Chicago University Press.

—— (1977) *The Archaeology of Ancient China* (3rd edn), New Haven, Connecticut: Yale University Press.

——(1986) *The Archaeology of Ancient China* (4th edn), New Haven, Connecticut: Yale University Press.

——(1989) 'Xin shíqi shídai de Taiwan haixia', *Kaogu* 6: 541–50.

——(1992) 'China' in Robert W. Ehrich (ed.) *Chronologies in Old World Archaeology* (3rd edn, 2 vols), Chicago: Chicago University Press.

Chêng Tê-K'un (1957) *Archaeological Studies in Szechwan*, Cambridge: Cambridge University Press.

Chu, J.Y. , Huang, W., Kuang, S.Q., Wang, J.M., Xu, J.J., Chu, Z.T., Yang, Z.Q., Lin, K.Q., Li, P., Wu, M., Geng, Z.C., Tan, C.C., Du, R.F. and Jin, L. (1998) 'Genetic relationship of populations in China', *Proceedings of the National Academy of Sciences of the United States of America* 95: 11763–8.

O'Connor, R. (1995) 'Agricultural change and ethnic succession in Southeast Asian states: a case for regional anthropology', *Journal of Asian Studies* 54, 4: 968–96.

Conrady, A. (1896) *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten*, Leipzig: Otto Harrassowitz.

Dani, A. (1960) *Prehistory and Protohistory of Eastern India, with a Detailed Account of the Neolithic Cultures in Mainland Southeast Asia*, Calcutta: Firma K.L. Mukhopadhyay.

Demeter, F. (2000) *Histoire du peuplement humain de l'Asie extrême-orientale depuis le pléistocène supérieur récent*, thèse de doctorat, Université de la Sorbonne à Paris, 19 décembre.

Ding, Y.C., Wooding, S., Harpending, H., Chi, H.C., Li, H.P., Fu, Y.X., Pang, J.F., Yao, Y.G., Xiang Yu, J.G., Moyzis, R. and Zhang, Y.P. (2000) 'Population structure and history in East Asia', *Proceedings of the National Academy of Sciences* 97, 25: 14003–6.

D'jakonov, I. (1968) *Predystorija Armjanskogo Naroda: Istorija Armjanskogo Nagor'ja s 1500 po 500g. do n.è., Xurrity, Luvijcy, Protoarmjane.* Erevan: Akademija Nauk Armjanskoj Sovetskoj Socialističeskoj Respubliki.

van Driem, George (1997) 'Sino-Bodic', *Bulletin of the School of Oriental and African Studies* 60, 3: 455–88.

——(1998) 'Neolithic correlates of ancient Tibeto-Burman migrations', in Roger Blench and Matthew Spriggs (eds) *Archaeology and Language II,* London: Routledge.

——(1999) 'A new theory on the origin of Chinese', in Peter Bellwood and Ian Lilley (eds) *Bulletin of the Indo-Pacific Prehistory Association Bulletin, 18* (*Indo-Pacific Prehistory: The Melaka Papers, Vol. 2*), 43–58, Canberra: Australian National University.

——(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.

AQ: please check page range

van Driem (2002) ‘Tibeto-Burman replaces Indo-Chinese in the 1990s: review of a decade of scholarship’, *Lingua* 111: 79–102.
Finck, F. (1909) *Die Sprachstämme des Erdkreises*, Leipzig: B.G. Teubner.
Forchhammer, E. (1882) ‘Indo-Chinese languages’, *Indian Antiquary* XI: 177–89.
Grube, W. (1881) *Die sprachgeschichtliche Stellung des Chinesischen*, Leipzig: T.O. Weigel.
Haudricourt, A.-G. (1953) ‘La place du vietnamien dans les langues Austro-asiatiques’, *Bulletin de la Société Linguistique de Paris* 49, 1: 122–8.
——(1954) ‘De l’origine des tons en viêtnamien’, *Journal Asiatique* 242: 68–82.
Jobling, M., Bouzekri, N. and Taylor, P.G. (1998) ‘Hypervariable digital DNA codes for human paternal lineages: MVR-PCR at the Y-specific minisatellite MSY1 (DYF155S1)’, *Human Molecular Genetics* 7: 643–53.
Kayser, M., Krawczak, M., Excoffier, L., Dieltjes, P., Corach, D., Pascali, V., Gehrig, C., Bernini, L., Jespersen, J., Bakker, E., Roewer, L. and de Knijff, P. (2001) ‘An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations’, *American Journal of Human Genetics* 68: 990–1018.
von Klaproth, J. (1823a) *Asia Polyglotta*, Paris: A. Schubart.
——(1823b) *Asia Polyglotta: Sprachatlas*, Paris: A. Schubart.
——(1830) ‘Réponse à quelques passages de la préface du roman chinois intitulé: *Hao khieou tchhouan*, traduit par M.J.F. Davis’, *Journal Asiatique* V: 97–122.
——(1831) *Asia Polyglotta* (2nd edn), Paris: Heideloff & Campe.
Körner, J. (ed.) (1930) *Briefe von und an August Wilhelm Schlegel* (2 vols), Zürich: Amalthea-Verlag.
de Knijff, P. (2000) ‘Messages through bottlenecks: on the combined use of slow and fast evolving polymorphic markers on the human Y chromosome’, *American Journal of Human Genetics* 67: 1055–61.
Kuhn, E. (1889) ‘Beiträge zur Sprachenkunde Hinterindiens’, *Sitzungsberichte der Königlichen Bayerischen Akademie der Wissenschaften* (München), Philosophisch-philologische Classe, II: 189–236.
Lepsius, C. (1861) ‘Über die Umschrift und Lautverhältnisse einiger hinterasiatischer Sprachen, namentlich der Chinesischen und der Tibetischen’, *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin,* aus dem Jahre 1860: 449–96.
Leyden, J. (*c.*1806a) *Plan for the Investigation of the Language, literature, History and Antiquities of the Indo-Chinese Nations*. [69-page manuscript held by the British Library, ADD. MSS 26,564; later published with changes as Leyden (1808)].
——(*c.*1806b) *Plan for the Investigation of the Language, Literature and History of the Indo Chinese Nations*. [11-page manuscript held by the British Library, ADD. MSS 26,565].
——(1808) ‘On the languages and literature of the Indo-Chinese nations’, *Asiatic Researches* X: 158–289.
Mason, F. (1854) ‘The Talaing language’, *Journal of the American Oriental Society* IV: 277–89.
Matisoff, J. (1991) ‘Sino-Tibetan linguistics: present state and future prospects’, *Annual Review of Anthropology* 20: 469–504.
Pictet, A. (1859) *Les origines indo-européennes ou les aryas primitifs: Essai de paléontologie linguistique (première partie)*, Paris: Joël Cherbuliez.
——(1863) *Les origines indo-européennes ou les aryas primitifs: Essai de paléontologie linguistique (seconde partie)*, Paris: Joël Cherbuliez.
Przyluski, J. (1924) ‘le sino-tibétain’, and ‘les langues austroasiatiques’, in Antoine Meillet and Marcel Cohen (eds) *Les Langues du Monde*, Paris, Librairie Ancienne Édouard Champion.

Przyluski, J., and Luce, G. (1931) 'The number "a hundred" in Sino-Tibetan', *Bulletin of the School of Oriental Studies* VI, 3: 667–8.

Sagart, L. (1990) 'Chinese and Austronesian are genetically related', Paper presented at the *23rd International Conference on Sino-Tibetan Languages and Linguistics* (5–7 October), University of Texas at Arlington.

——(1994) 'Proto-Austronesian and Old Chinese evidence for Sino-Austronesian', *Oceanic Linguistics* 33, 2: 271–308.

——(2001) 'Lexical evidence for Austronesian-Sino-Tibetan relatedness', Paper presented at the *Conference on Connections across the Southern Pacific*, Hong Kong City University, 18 January.

von Schlegel, A. (1832) *Réflexions sur l'étude des langues asiatiques adressées à Sir James Mackintosh, suivie d'une lettre à M. Horace Hayman Wilson*, Bonn: Weber.

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

——(1955) 'Classification of the Sino-Tibetan languages', *Word, Journal of the Linguistic Circle of New York* 11, 94–111.

——(1966) *Introduction to Sino-Tibetan, Part I*, Wiesbaden: Otto Harrassowitz.

——(1967) *Introduction to Sino-Tibetan, Part II*, Wiesbaden: Otto Harrassowitz.

——(1968) *Introduction to Sino-Tibetan, Part III*, Wiesbaden: Otto Harrassowitz.

——(1974) *Introduction to Sino-Tibetan, Part IV*, Wiesbaden: Otto Harrassowitz.

Sharma, T.C. (1967) 'A note on the neolithic pottery of Assam', *Man, Journal of the Royal Anthropological Institute (n.s.)* 2, 1: 126–8.

——(1981) 'The neolithic pattern of north-eastern India', in M.S. Nagaraja Rao (ed.) *Madhu: Recent Researches in Indian Archaeology and Art History*, Delhi: Agam Kala Prakashan.

——(1989) 'Neolithic: Eastern region', in A. Ghosh (ed.) *An Encyclopaedia of Indian Archaeology* (2 vols), Vol. 1, 58–60, New Delhi: Munshiram Manoharlal Publishers.

Simon, W. (1929) 'Tibetisch-chinesische Wortgleichungen, ein Versuch', *Mitteilungen des Seminars für Orientalische Sprachen an der Friedrich-Wilhelms-Universitāt zu Berlin* XXXII, 1: 157–228.

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

Su, B., Xiao, J.H., Underhill, P., Deka, R., Zhang, W.L., Akey, J., Huang, W., Shen, D., Lu, D.R., Luo, J.C, Chu, J.Y., Tan, J.Z., Shen, P.D., Davis, R., Cavalli-Sforza, L., Chakraborty, R., Xiong, M.M., Du, R.F., Oefner, P., Chen, Z. and Jin, L. (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: 1718–24.

Su, B., Xiao, C.J., Deka, R., Seielstad, M., Kangwanpong, D., Xiao, J.H., Lu, D.R., Underhill, P., Cavalli-Sforza, L., Chakraborty, R. and Jin, L. (2000) 'Y chromosome haplotypes reveal prehistorical migrations to the Himalayas', *Human Genetics* 107, 6: 582–90.

Temple, R. (1903) *Census of India, 1901,* Vol. III: *The Andaman and Nicobar Islands: Report on the Census*, Calcutta: Office of the Superintendent of Government Printing, India.

Thapar, B.K. (1985) *Recent Archaeological Discoveries in India*, Paris: United Nations Educational Scientific and Cultural Organization.

AQ: Please check the change of Part II to Part III in Schmidt 1968. Is it Okay?

Tsang, C.H. (2001) 'Recent discoveries of the Tap'enkeng culture in Taiwan: Implications for the problem of Austronesian origins', Paper presented at the Colloque *Perspectives sur la Phylogénie des Langues d'Asie Orientales* at Périgueux, France, 30 August.

Underhill, P., Jin, L., Lin, A., Mehdi, S.Q., Jenkins, T., Vollrath, D., Davis, R.W., Cavalli-Sforza, L. and Oefner, P. (1997) 'Detection of numerous Y chromosme biallelic polymorphisms by denaturing high-performance liquid chromatography', *Genome Research* 7: 996–1005.

Underhill, P., Passarino, G., Lin, A., Shen, P., Mirazón Lahr, M., Foley, R.A., Oefner, P. and Cavalli-Sforza, L. (2001) 'The phylogeny of Y chromosme binary haplotypes and the origins of modern human populations', *American Journal of Human Genetics* 65: 43–62.

Walravens, H. (1999a) *Julius Klaproth (1783–1835): Leben und Werk*, Wiesbaden: Otto Harassowitz.

——(1999b) *Julius Klaproth (1783–1835): Briefe und Dokumente*, Wiesbaden: Otto Harassowitz.

Wheeler, R. (1959) *Early India and Pakistan to Ashoka*, London: Thames and Hudson.

Whitehead, G. (1925) *Dictionary of the Car-Nicobarese Language*, Rangoon: American Baptist Mission Press.

Witsen, N. (1692) [2nd impression 1705] *Noord en Oost Tartarye, ofte Bondig Ontwerp van eenige dier Landen en Volken, welke voormaels bekent zijn geweest, beneffens verscheide tot noch toe onbekende, en meest nooit voorheen beschreven Tartersche en Nabuurige Gewesten, Landstreeken, Steden, Rivieren, en Plaetzen, in de Noorder en Oostelykste Gedeelten van Asia en Europa* (2 vols), Amsterdam: François Halma.

Wu, Y.L. (1996) 'Prehistoric rice agriculture in the Yellow River valley', in Ian C. Glover and Peter Bellwood (eds) *Indo-Pacific Prehistory: The Chiang-Mai Papers* (Proceedings of the 15th Congress of the Indo-Pacific Prehistory Association, Chiang Mai, Thailand, 5–12 January 1994), 223–4, Vol. 2, Canberra: Australian National University.

Zhang, C. (1996) 'The rise of urbanism in the middle and lower Yangtze river valley', in Peter Bellwood and Diane Tillotson (eds) *Indo-Pacific Prehistory: The Chiang-Mai Papers* (Proceedings of the 15th Congress of the Indo-Pacific Prehistory Association, Chiang Mai, Thailand, 5–12 January 1994), 63–8, Vol. 3, Canberra: Australian National University.