Man In India, 95 (4): 231–266 © Serials Publications

ANCIENT POPULATION MIGRATIONS IN NORTHEAST INDIA: A CLOSER LOOK AT THE ETHNOLINGUISTIC PREHISTORY

Manjil Hazarika

Four major linguistic families are found in Northeast India, i.e. Austroasiatic, Tibeto-Burman, Indo-European and Kradai. Previous studies on the reconstruction of the sequence of arrival of these people of different linguistic stocks have been based on simple philological and ethnographical considerations. The unfolding story revealed by population genetics and the linguistic palaeontology of the different language families turns out to be more complex than previously thought. The present paper aims at a fresh investigation of ethnolinguistic prehistory of Northeast India based on archaeological and relevant multidisciplinary data. The evidence presented in this paper is gathered from archaeobotanical, ecological, ethnographical, folkloristic, historical and human genetic sciences to inspire an interpretation of the available archaeological data for examining linguistic hypotheses of ancient population migrations and dispersals in this region.

Introduction

> 'When we are looking at the archaeology, we are looking at the past, but we are looking at just one version of the past, which is the material culture. Linguistics also gives a version of the past, and population genetics gives us another version of the past, and these three versions of prehistory can be correlated but they need not necessarily have anything to do with each other'
>
> George van Driem (2008: 101)

The use of independent evidence from different disciplines to reconstruct past population histories has proved to be of particular significance in recent years. Such independent evidence comprises archaeological, linguistic and genetic data. The archaeological record offers meaningful data on ancient material culture and the development of technology with a timeframe for the emergence of innovations. Historical linguistic data are useful for independent phylogenetic analysis of linguistic relationships which often complement archaeological data and provide clues about ancient migrations and possible events of admixture. The genetic data are extremely helpful to understand and interpret the biological relationships which obtain between modern people and the likely points of origin and expansion of their ancestors (Scheinfeldt *et al.* 2010: 8931). Notwithstanding the methodological differences between historical linguistics and archaeology, both disciplines aim to reconstruct the 'sequences of events, the one linguistic and the other material-cultural' (Spriggs and Blench

1998: 29). More importantly, connecting historical linguistic data with archaeology generates testable hypotheses (Blench 2004).

Our current knowledge of the population history in Northeast India is based on simplistic phylogenetic data. However, the hypotheses of historical linguistics should be tested through the archaeological record. In view of the usefulness of independent evidence to reconstruct population history of a particular area, this paper is an attempt to highlight the linguistic situation in Northeast India in terms of our understanding of the dispersals of ancient linguistic groups. Plausible migration histories of these groups have been addressed on the basis of linguistic, genetic, ethnographical, historical and folkloristic data. In this regard, one has to keep in mind caveats such as those voiced by George van Driem:

> 'Very often language seems to be less ambiguously correlated with the geographical distribution of genetic markers in the populations speaking the languages in question. So, can genes and languages generally be correlated and contrasted with each other in a more meaningful way than either can be with the fragments of material culture that happen to have resurfaced unscathed from the sands of time? On the one hand, the linguistic ancestors of a language community were not necessarily the same people as the biological ancestors of that community. At the same time, the genetic picture often shows sexual dimorphism in linguistic prehistory.' (van Driem 2011a: 24)

Linguistic situation in Northeast India

Northeast India is spread across over 262,000 km^2 and comprises the eight Indian states of Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Sikkim and Tripura. This vast region is known for its diverse landscapes and ecologies as well as cultural diversity. This region is an ethnic mosaic consisting of different tribal groups of various ethnic stocks, maintaining their traditional customs and practices, having self-sufficient economies, and thus creating a multicultural constellation of tribes and peoples. A great variety of languages are spoken by different linguistic communities in the region (Fig. 1). However, many of these languages are not known or are poorly understood and on the verge of extinction and often regarded as endangered languages (van Driem 2007a and 2007b). For a number of languages, only a small group of speakers continue to speak the language. Since these communities interact with neighbouring Assamese or Bengali speakers, they in most cases become fluent in these languages too. As most of these linguistic groups live close together, the possibility for the diffusion of linguistic features though contact situations over extended periods of time could have resulted in common linguistic features, even when these languages are not genetically related. Karbi a.k.a. Mikir is a Tibeto-Burman language, but Moral (1996: 44) has claimed the language shows evidence of having been exposed to the Austroasiatic language Khasi. Indeed, such an ancient contact situation between the two language communities finds possible archaeological corroboration in the form of ancient Khasi megaliths, the

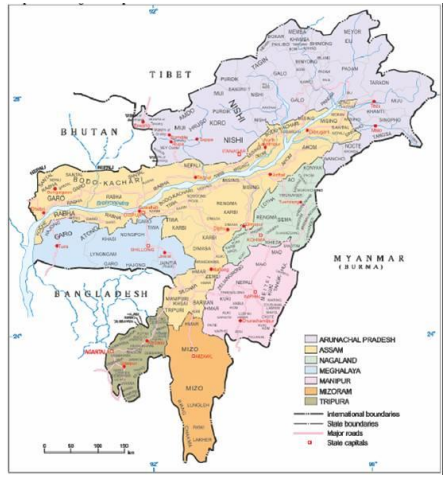

Figure 1: Ethnolinguistic map of Northeast India (after Blench 2014, originally published by Bishop's House, Guwahati)

distribution of which extend far beyond the current Khasi area well into the Mikir Hills a.k.a. the Karbi Anglong (van Driem 2001: 281).

In describing the linguistic importance of Northeast India, Post (2008: 5) asserts that it is 'without a doubt, and by any measure, the richest, most diverse, most linguistically significant area in the entire Asian continent, and is one of the top 3 or 4 most significant linguistic areas of the entire world'. Recently, Blench and Post (*in press*) on the basis of their work in Arunachal Pradesh show that urgent attention from linguists is required to document the lesser known languages spoken in the region considering their uniqueness and endangered status. The study also suggests that many of these languages or clusters could well be isolates, and that the Tibeto-Burman roots they demonstrate may well be borrowings.

Linguistic groups and theories of their dispersals

In terms of geography, Northeast India was always destined to play a crucial role in shaping the population prehistory of not just the Indian subcontinent but also East Asia and Southeast Asia (van Driem 2014a, 2014b). This region of India can rightly be equated with Northwest India, through which the country was linked with Western and Central Asia. Through these two corners of the subcontinent men, material culture and ideas have entered since prehistoric times and gave rise to the inestimable variety of races and cultures with which India is distinguished today (Chatterji 1970: 7-8). Mills (1928: 24) called Northeast India

'one of the great migration routes of mankind'. Medhi (2003: 322) refers to this region as the 'Great Indian Corridor' for the prehistoric and proto-historic movements of people from and to its neighbouring regions. The movement of people took place not only in the historical period but continues to the present as well.

The Ahom are a Kradai group which came to Northeast India from the kingdom of Pong in the upper Irrawaddy basin, a polity which straddled a part of upper Burma and the adjacent portion of the Chinese province Yúnnán. Around 1228, the Ahom were led across the Patkai range into the Brahmaputra valley under the leadership of Siukapha. The Ahom are ethnolinguistically related to the Shan, a prominent Tai group in Burma, and indeed the ethnonyms Ahom, Shan and Siam are all cognate. Historical phonology tells us that at the time that the Ahom polity was established in the Northeast, the name Ahom was still pronounced **asam*, and the present name of the Indian state, and ironically also of the Indo-Aryan language Assamese, derive from this native Kradai ethnonym (van Driem 2001: 329). The migration of this Kradai group into the Northeast led to the establishment of the Ahom kingdom that existed for almost six hundred years. Thanks to the chronicles known as *Buranjis* written during the Ahom period, the cultural, economic, social and linguistic history of the upper Brahmaputra valley from the 13th century AD is comparatively well known. Other Kradai groups of Northeast India, such as the Aiton, Khampti, Khamyang, Phake and Turung, are recorded as having come to Northeast India in the 17th and 18th centuries.

Before the advent of the Ahom, the inscriptional and numismatic evidence attests to the emergence of different principalities or smaller kingdoms in the valley since about the 5th century AD, particularly in the Dhansiri-Doiyang and Kapili-Jamuna valleys. The advent of Indo-Aryan colonists at this time is associated with the rise of the Kāmarūpa kingdom, which flourished from the 4th to the 13th century in what today is western Assam. The polity was characterised by an Indo-Aryan élite and a Tibeto-Burman populace speaking an early form of Bodo-Koch. Yet in reality the polity was alternatingly ruled by Indo-Aryan and indigenous but Hinduised Bodo-Koch dynasties, and the process of Hinduisation in the lower Brahmaputran valley may already have begun as early as the 1st century AD (van Driem 2001: 505-506). Barua (1933) suggested that some form of the Proto-Bodo-Koch language probably acted as a *lingua franca* throughout the Brahmaputra valley prior to the advent of the Ahom. More recently the same suggestion was put forward by DeLancey (2010: 29). The Indo-Aryan élite, however, already at this time had introduced the Prakrit which was to evolve into modern Assamese.

Sometimes speculations about the linguistic history appear to be sheer guesswork, as when Moral writes without explanation that the 'Tibeto-Burman tribes came through Burma and entered the hills and valleys of Assam in about

1000 BC. They gradually encroached upon the Austric settlers who have been settling here since 2000 to 2500 BC and forced most of them to take refuge in mountainous homes. That was how the Khasis thrived in their mountainous homes high on the hills of Meghalaya' (Moral 1996: 24, 52). In contrast to such guesswork, reasoned inferences have been put forward and careful correlations have been undertaken between different types of evidence in order to arrive at a reconstruction of the prehistory of the region with a transparent and adjustable argument structure, carefully weighing emerging archaeological, linguistic palaeontological, paleoethnobotanical, and ethnographical data. At the same time, some reconstructions of ethnolinguistic prehistory are guided by the theoretical frameworks of the scholars who propose them.

The agricultural and linguistic group dispersal hypothesis discussed on a global scale was proposed by Diamond and Bellwood (2003), who suggests an outward dispersal of farming populations, bearing their languages and culture from their original homeland (Fig. 2). The expansions of Austroasiatic and Tibeto-Burman language families and the Kradai to a certain extent from their agricultural homelands in China at different times and over different geographic ranges ostensibly determined the population history of Northeast India (Diamond and Bellwood 2003: 600). Crucial to their theory is that the founding dispersals of language families went hand in hand with the spread of agriculture along the Neolithic horizon, and Colin Renfrew is another proponent of this view. Before we evaluate this model in the context of the Northeast, let us first turn to different theories about the emergence and dispersal of the Austroasiatic and Tibeto-Burman language families.

Austroasiatic language family

Austroasiatic includes well-known languages and language groups such as Khasi, Munda, Nicobarese, Vietnamese and Khmer. This widespread language phylum in South and Southeast Asia comprises well over two hundred languages, yet over 90% of all Austroasiatics speak just one language Vietnamese. Whilst most languages are spoken by tiny language communities, only Khasi, Khmer and Vietnamese are large languages which have expanded successfully in historical times. Austroasiatic used to be conventionally divided into the two major branches Mon-Khmer and Munda. The world's leading authority on Austroasiatic linguistics, Gérard Diffloth (2005), challenged this traditional bifurcation of the family, stressing the time depth of the split between the Khasi-Khmuic and Mon-Khmer branches of the language family. Diffloth proposed a trifurcation at the deepest linguistically reconstructible level between Munda, Khasi-Khmuic and Mon-Khmer. Later, Diffloth (2009) reverted to a revised bipartite model of the language family, but with the two main branches

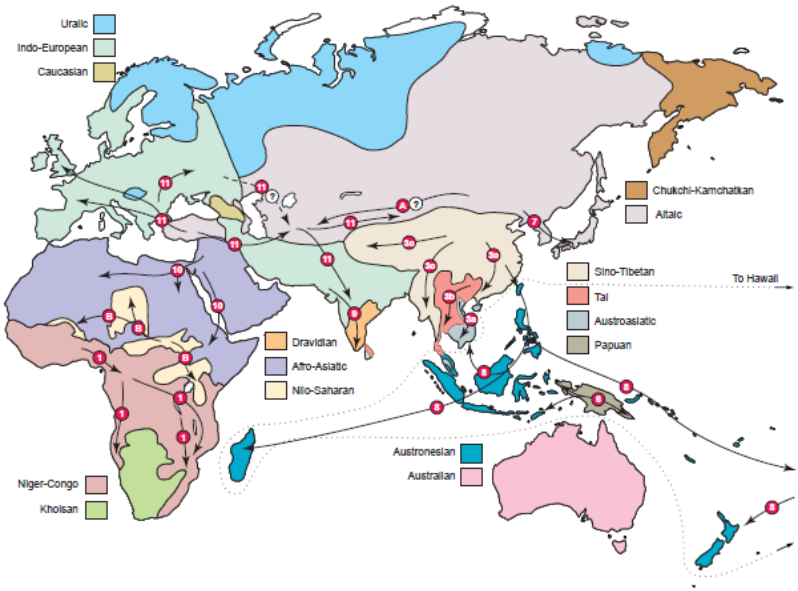

Figure 2: Language families of the Old World and their suggested expansions, 1 (Bantu), 3a to 3c (Austroasiatic, Kradai, and Tibeto-Burman respectively), 6 (Trans New Guinea), 7 (Japanese), 8 (Austronesian), 9 (Dravidian), 10 (Afro-Asiatic), 11 (Indo-European). Other possible examples mentioned only briefly: A (Turkic), B (Nilo-Saharan) (after Diamond and Bellwood 2003: 598)

of Austroasiatic now being Munda and Khasi-Aslian. Diffloth's current language family tree appears in print in an article by van Driem (2012c: 130, Fig. 19). The reproduced tree diagram deviates, however, in one minor respect from the model which Diffloth himself presented in 2009. The Munda languages are native to India, concentrated in and around Chota-Nagpur plateau. Munda is traditionally subdivided into North and South Munda, and this old view is reproduced in the diagram in van Driem (2012c: 130, Fig. 19). However, according to Diffloth's more meticulous 2005 phylogeny, reproduced here in Figure 4, Koraput is the first sub-branch to split off, the Kharia-Juang subgroup of languages and Lects is the second sub-branch to split off, and the remainder of Munda split into the two sub-branches Kherwarian and Korku at a younger time depth. Munda groups such as Juang, Gata, Bondo, Bodo Gadaba, Paranga and Saora occupy the Koraput and adjoining districts of Orissa, whilst the Kherwarian groups comprising Asur, Birhor, Ho, Korwa, Santhal, Turi and Munda are spread across

Jharkhand (Ranchi, Gumla, Lohardaga and Singhbhum districts), Orissa (Mayurbhanj, Keonjhar and Sundergarh districts), Madhya Pradesh (Raigarh and Jashpur districts) and West Bengal (Birbhum, Nadia and Bakura districts). A section of Korku inhabits the northeastern border areas of Maharastra.

Other than this detail, the tree diagram in van Driem (2012c: 130, Fig. 19) represents the state-of-the-art in Austroasiatic linguistics (Fig. 3). The large Khasi-Aslian trunk of the family, coordinate with Munda, splits into a Khasi-Pakanic and a Mon-Khmer branch. Khasian is the first sub-branch to split off of Khasi-Pakanic, and Khasian languages are spoken in the eastern Meghalaya and the Jaintia Hills. In addition to the dialects of Khasi, the Khasian subgroup comprises the languages Synteng, Lyngngam and Amwi a.k.a. War, which are clearly distinct but related languages. The second sub-branch to split off of Khasi-Pakanic is Khmuic, leaving behind a Pakano-Palaungic branch, which splits into the subgroups Pakanic and Palaungic at a younger time depth. Khmuic languages are found in northern Laos and northern Thailand, including the many dialects of Khmu, the Mal-Phrai languages and Mlabri. The Palaungic branch, formerly called Palaung-Wa, extends over northern Thailand and Laos, eastern Burma and southwestern Yúnnán. Eastern Palaungic contains several Palaung languages, the Riang dialects and Danau, whereas Western Palaungic contains the three language subgroups Waic, Angkuic and Lametic. The most differentiated of these is the Waic group, which includes Bulang, the many Lawa dialects and the Wa languages, totalling over half a million speakers. The Angkuic group includes several very small and nearly unknown languages such as Angku, U, Hu, Mok, Man Met and Kiorr (Diffloth and Zide 1992). The Pakanic languages are a fragmentary little known group in southern China.

The Mon-Khmer branch splits into a Khmero-Vietic and a Nico-Monic branch. Khmero-Vietic in turn splits into the Vieto-Katuic sub-branch, comprising the subgroups Vietic and Katuic, and the Khmero-Bahnaric sub-branch, comprising the subgroups Khmeric and Bahnaric. The Nico-Monic branch splits into Nicobarese languages, spoken on the Nicobars in the Andaman Sea, and the Asli-Monic sub-branch, comprising the subgroups Monic and Aslian. The Pearic subgroup occupies an indeterminate position within the Mon-Khmer branch. Pearic languages have undergone much contact influence from Khmer, and Khmer appears to have expanded largely at the expense of Pearic.

Robert von Heine-Geldern proposed in 1932 on archaeological and ethnolinguistic grounds that the Austroasiatic groups of the Indian subcontinent originated from an ancestral homeland in Southeast Asia. He also expressed the view, which at that time already reflected a widespread consensus amongst ethnographers, anthropologists and linguists, that the Austroasiatic presence in the north of the Indian subcontinent antedated the Dravidians and much later Indo-Europeans (van Driem 2001: 408, 416). His archaeological arguments

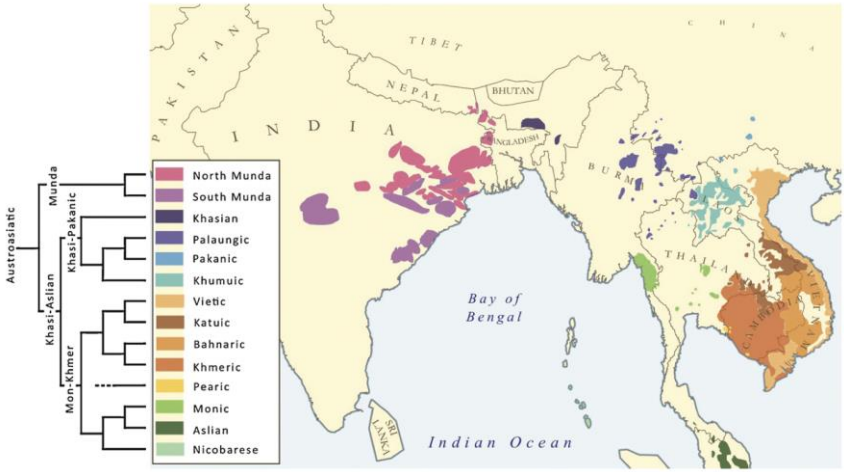

Figure 3: Geographical distribution of Austroasiatic subgroups, showing Diffloth's 2009 Austroasiatic phylogeny (after van Driem 2012c: 130, Fig. 19)

involving the *Schulterbeilkultur* have since Colani's work *mutatis mutandis* basically been applied to the Hoabinhian. Blench (2008: 163) adheres to this old theory, envisaging Austroasiatic as having spread from the Mekong valley westwards across a number of other river valleys. The geographical distribution and isolation of different Austroasiatic groups over a large area suggests to Blench that the Tibeto-Burman language groups subsequently spread southward. Peiros and Shnirelman (1997) argued that the reconstructed Austroasiatic lexicon does not contain any words associated with the sea coast. Moreover, linguistic palaeontology, according to Peiros and Schnirelman offer no clear indication of a tropical environment. Consequently, they propose an Austroasiatic homeland in mainland eastern Eurasia, where a non-tropical climate prevailed. They envisage a sub-tropical mountainous homeland along the Middle Yangtze.

Inferences made about the material culture based on the lexicon reconstructible for the proto-language is called Linguistic Palaeontology, a term coined by Swiss scholar Adolphe Pictet (1859). Linguistic paleontological study on the possible homeland of the Austroasiatic linguistic group by Gérard Diffloth (2005) demonstrates that speakers of the Austroasiatic proto-language were thoroughly familiar with rice agriculture, as evinced by the rich lexicon of rice agriculture terms reconstructible to the common ancient language. Even the names of the animal species in the reconstructed proto-Austroasiatic lexicon are

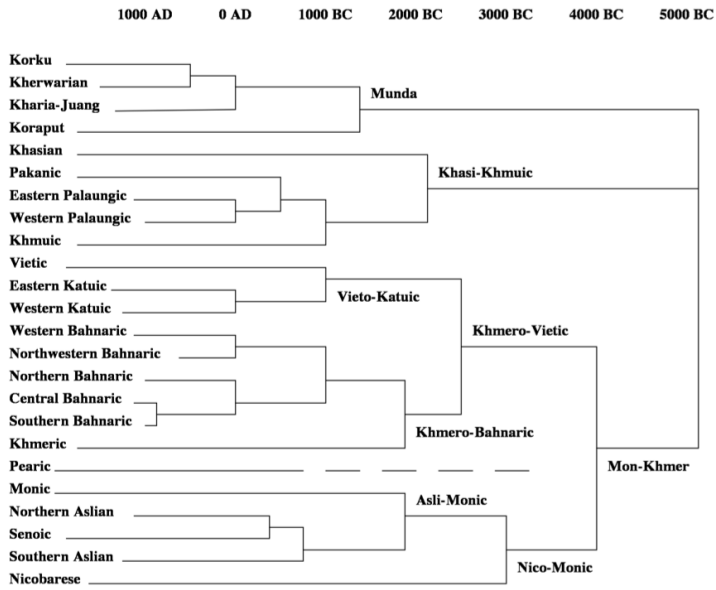

Figure 4: Tentative calibration of time depths for the various branches of the Austroasiatic language family by Gérard Diffloth (modified from Diffloth 2005 by van Driem 2011a: 16). As discussed in the text, Diffloth (2009) now holds the first bifurcation in the family to have been between Munda and Khasi-Aslian, the latter splitting into the branches Khasi-Khmuic and Mon-Khmer.

restricted to the humid tropics. The geographical distribution of the different branches of Austroasiatic also point towards a centre of the greatest historical diversity in the region encompassing the fertile flood plains of the Irrawaddy in Burma and the lower Brahmaputra in Assam and Bangladesh. Diffloth proposed a primary split between Munda and Khasi-Aslian in a location somewhere within the littoral arc of the Bay of Bengal.

According to Diffloth, the ability to reconstruct 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. To Diffloth this evidence suggested an area in northeastern India, the Indo-Burmese borderlands, Burma and Yúnnán which would indicate that the Austroasiatic homeland may have lain in the Northeast, or at least more towards Southeast Asia than to India proper.

At one time it appeared that the best archaeological correlate for the ancient rice cultivating culture might be the Hémǔdù archaeological assemblage (5000-

4500 BC) at the mouth of the Yangtze, which provided the best unambiguous evidence for a population for whom rice is the staple (van Driem 2007c). Yet the unfolding story revealed by rice genetics, the archaeology of rice and the linguistic palaeontology of the Austroasiatic language family turns out to be more complex. This tale is told by van Driem (2012a, 2012b) on the basis of the most recent findings of rice population genetics, reconstructible Austroasiatic fauna etyma provided by Diffloth and a critical assessment of the finds of paleobotanists, their reasoning and the significance of vast gaps in the archaeology of regions relevant to resolving the issue. A review of the population genetic literature undertaken by Kumar and Reddy (2003: 501) was unable to arrive at any consensus regarding provenance and migratory history of the Austroasiatic peoples during the peopling of India. Yet more recently, with the aid of new human population genetic research with better results, the story on rice has led to a new model of the provenance of Austroasiatic, distinguishing several distinct chronological layers (van Driem 2013), but we shall return to this model later. First, we shall turn to another Austroasiatic homeland hypothesis.

Paul Sidwell contends that, in spite of years of research, there is no general consensus on the relations between Austroasiatic branches, the age or diversity of the language family and an appropriate program for addressing these issues (Sidwell 2010: 117). On the basis of the geographical distribution of Austroasiatic speakers, Sidwell proposed that Austroasiatic language dispersed along an axis which ran roughly southeast to northwest along the middle course of the Mekong River as the greatest number of Austroasiatic branches are spoken along this axis. He calls this proposal the 'Austroasiatic Central Riverine Hypothesis' (Sidwell 2010: 118). This hypothesis is not at all inherently implausible, but, despite its novel label, Sidwell's hypothesis is basically Robert von Heine-Geldern's 1932 hypothesis, as reinterpreted by Blench (2008).

Nonetheless, this proposition has been severely criticised by Peiros (2011), who argues that locating the original homeland of a proto-language and tracing possible migrations of speakers can be conducted based on the current geographical locations of genetically diverse languages. The current distribution pattern of Austroasiatic languages suggests a movement along river valleys of the Brahmaputra, Irrawaddy, Salween, Mekong and their tributaries, whereby the starting point of the original Austroasiatic migrations must have lain in present-day southern Sìchuān near the upper Yangtze. From this area, Peiros envisages the Proto-Austroasiatics moving into different parts of Southeast Asia. Stressing the present-day montane location of many Austroasiatic language communities, Peiros (2011) suggests that the homeland of the language group was located not at the bottom of a tropical valley, but on much higher ground.

Tibeto-Burman or Trans-Himalayan language family

The Tibeto-Burman or Trans-Himalayan language family includes over three hundred languages stretching from the Himalayas to East Asia and into Northeast India and Southeast Asia. The language family was first recognised by Julius von Klaproth in 1823, and in terms of numbers of speakers this language family is the second most populous on the planet. Like Austroasiatic and many other large language families, the distribution of speakers is lopsided and the result of recent developments in the historical period. Linguistically speaking, most Tibeto-Burmans belong to just a few language communities such as Cantonese, Burmese, Tibetan and Mandarin, whereas hundreds of Tibeto-Burman languages are spoken by anywhere between several hundred thousand and just a handful of speakers. The precise phylogeny of the language family is the object of research (van Driem 2001).

The Indo-Chinese or Sino-Tibetan model was inherited by a generation of American scholars after the Great Depression without any historical linguistic evidence ever having been presented for the family tree. Instead, the Chinese or Sinitic languages were treated as a distinct primary branch of this 'Sino-Tibetan' family because of racist typological arguments dating from the time of scientific racism which showed that Sinitic represented the lowest rung on the typological ladder of language evolution (van Driem 2003, 2005, 2007d, 2014a). A later generation of Sino-Tibetanists have tried to turn this argument around and attempted to propagate a Sinocentric view whereby the entire family ostensibly originated from the cradle of Chinese civilisation on the North China plain, e.g. LaPolla (2006). The writings of the so-called 'Sino-Tibetanists' cannot be understood unless one is aware that they use the term 'Tibeto-Burman' not in its original sense, to designate the family as a whole, but collectively to designate all non-Sinitic languages, which they believe to represent one single branch of the family coordinate with Sinitic. Since the 1990s, the Sino-Tibetanists have become increasingly embarrassed by the lack of evidence for their model and its origins in scientific racism, but old paradigms are often tenacious and some hold on to the obsolete label to save face.

The most state-of-the-art classification of the Tibeto-Burman languages is provided by George van Driem (2001, 2014a). Whilst most speakers of Tibeto-Burman languages live to the northeast of the Himalayas, most primary subgroups of the Trans-Himalayan linguistic phylum are located to the southwest of the Himalayan divide. Furthermore, the epicentre of diversity for the language family as a whole lies in Northeast India and the surrounding hill tracts (Fig. 5).

Unaware of the phylogenetic complexity of the ethnolinguistic groups in the Northeast, Matisoff (1991) attempted to put all of the subgroups of Northeast India with which he, as a Lolo-Burmanist, was least familiar into a single branch called 'Kamarupan' without presenting any historical linguistic evidence in the

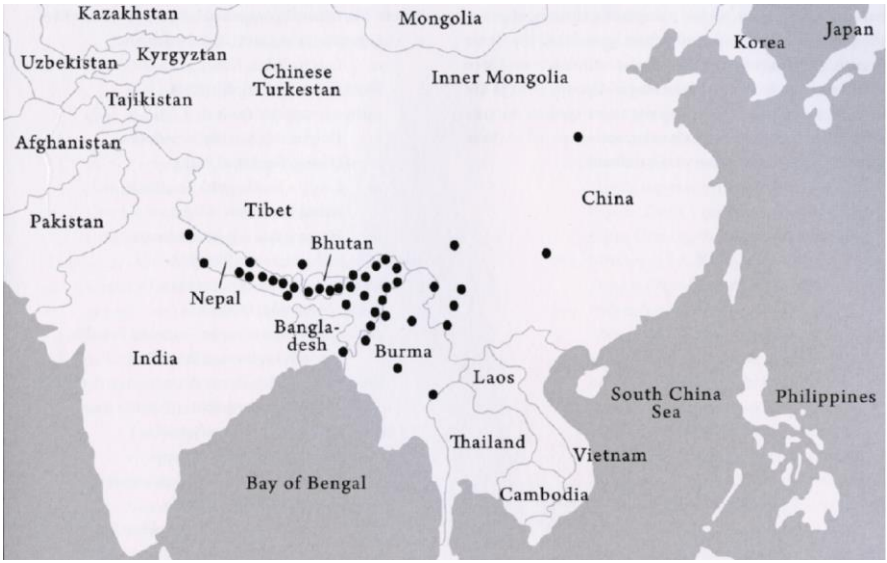

Figure 5: Asia showing the geographical distribution of the major branches of the Tibeto-Burman language family in which each diamond represents not a language, but a major subgroup. The centre of diversity lies in the eastern Himalayan region and Northeast India (after van Driem 2014a)

form of sound laws or shared innovations. This ungrounded catchall was criticised by Burling (1999: 169-171) and van Driem (1999: 50), and in his rejoinder Matisoff (1999) failed to adduce any argument in defence of the subgroup. Similar objections have been raised against the term 'Baric' which covers most of the languages of Northeast India (DeLancey 1991), but these criticism basically echoe the reservations of Robert Shafer himself, who explicitly used this label for a category which Shafer insisted did not represent a branch, subgroup or taxon, but a reservoir of subgroups whose precise phylogenetic relationships had yet to be worked out.

Synthesis of historical linguistics and prehistoric archaeological data

Above we have looked at the two major language families of greatest relevance to Northeast India, Austroasiatic and Tibeto-Burman, and we have mentioned language communities belonging to other linguistic phyla such as Kradai and Indo-European. In archaeology worldwide, the Neolithic period witnessed several important events of population prehistory including demographic change and expansion to new areas and further admixture among different groups. With the advent of agriculture and pastoralism, more complex societal issues emerged

including exchange of goods within and outside a territory. Population movements involved not only cultural exchange but also the expansion of language and genetic lineages to different areas. Although movements of people prior to Neolithic is also a matter of intense research, this post-Pleistocene period has especially attracted the attention of archaeologists, linguists and genetic scientists due to the availability of a larger dataset to examine the population dispersals independently in different parts of the world.

Application of an independent dataset to understand population history in Africa has provided meaningful insights, suggesting that the populations in close geographic proximity to each other as well as populations that speak linguistically similar languages are more likely to exchange genes. The geographical barrier also limits the gene flow as evident in the analysis of the northern African and sub-Saharan African populations (Scheinfeldt *et al.* 2010: 8931–8938). Another notable recent study of the application of historical linguistic and archaeological data involved the reconstruction of the Southern Jê languages of Brazil (de Souza 2011). A more complex multidisciplinary study based on archaeological, ecological, cultural, historical, social, linguistic and genetic data conducted in the Polynesia has provided vital insights into the human settlement and colonisation of the Pacific (Hurles *et al.* 2003).

The usefulness of historical linguistic data for solving and interpreting archaeological and historical problems has gained more attention in recent years which is evident from numerous publications (Bellwood 2005, Bellwood and Renfrew 2002, Blench and Spriggs 1997–9, Blench 2006, Enfield 2011, Forster and Renfrew 2006, Jin *et al.* 2001, Lamberg-Karlovsky 2002, McConvell and Evans 1997, Renfrew 1987, Renfrew *et al.* 2000, Sagart *et al.* 2005, Sanchez-Mazas *et al.* 2008, Southworth 2005a, van Driem 2001). There has been an attempt to find support or possible archaeological correlates for Pedersen's 1903 'Nostratic hypothesis', which argues that several of the world's language families are related in their origin, grammar and lexicon, and belong together in a larger unit of earlier origin. Nostratic includes Altaic, Afro-Asiatic, Indo-European, South Caucasian (Kartvelian), Uralic and Dravidian (Dolgopolsky 2008, Renfrew 2008).

The promising new trend of research integrating data from historical linguistics, prehistoric archaeology and molecular genetics for reconstructing the population prehistory of the world has made great strides forward. Barua, an Assamese scholar, was one of the first to attempt to correlate the prehistoric culture of the region with a particular ethnic and linguistic group. Barua was perhaps the first Indian scholar to follow Robert von Heine-Geldern in identifying the Neolithic culture of Assam with early Austroasiatics, and suggesting that the Khasis introduced the shouldered Neolithic hoe, terraced rice cultivation, megalithic burials, and matrilineal social system and later acquired the knowledge of iron smelting (Barua 1939: 6-18, 34-41). In this context, it may

be mentioned that in a recent publication, Prokop and Suliga (2013) report the stratigraphic evidence of iron smelting in the Khasi hills since 2040 ± 80 years BP (353 BC – 128 AD), which may be considered as the earliest undisputed evidence of this metal in Northeast India. Khasis are traditionally known for iron smelting and elaborate ways of erecting megalithic structures.

When exploring this interdisciplinary area, the potential veracity and therefore intrinsic interest of our reconstruction is limited by the reliability of the input from the various disciplines. For example, contact influence between languages is a well-documented phenomenon, whereby the grammar and lexicon of two language communities living in close proximity influence each other because of interaction. However, at this stage in the development of Tibeto-Burman historical linguistics, we must be wary as archaeologists of accepting any statements made by a linguist on the basis of phenomena such as contact influence *in the absence of historical comparative research* to support the statement. A typical example is LaPolla (2000), who envisages a mishmash of crossing lines of migration into the Himalayan region from China, explaining away similarities in some areas as the result of contact influence and providing no plausible account for the phylogenetic diversity observed in Northeast India (LaPolla 2001, 2006). In fact, LaPolla's 'reconstruction' manifestly derives from van Driem's 1998 study of possible Neolithic correlates of ancient Tibeto-Burman migrations, superficially recasting the data in a Sinocentric mould.

George van Driem has not recanted the scenario outlined in his 1998 study of possible Neolithic correlates with ancient Tibeto-Burman migrations, even though he no longer believes that these movements in the archaeological record necessarily date to the founding dispersal of the language family as a whole. In fact, details of van Driem's Neolithic scenario have been corroborated by archaeological research conducted since then. Rather, he now sees the episodes described in the study as corresponding to one slice or several slices of time in a long prehistory of peopling with a more complex chronological layering. As the story unfolds, the narrative has increased both in its complexity and in the empirical support for various aspects of the reconstruction (van Driem 2001, 2007c, 2007d, 2011a, 2012a, 2012b, 2012c, 2013, 2014b). Yet let us focus now on the archaeological aspects as they pertain to the Neolithic.

The archaeological record shows a connection between of Northeast Indian Neolithic and Neolithic assemblages in Sìchuan and Yúnnán than to those in Southeast Asia (van Driem 1998: 67-102). The Eastern Indian Neolithic wedges and tanged axes have clear parallels in upper Burma, Yúnnán and Sìchuan. Van Driem emphasises the correlations of Tibeto-Burman language dispersal with Neolithic expansion from South China. In his 1998 reconstruction, he still assumed that the proto-homeland of Tibeto-Burman language family lay in Yúnnán and Sìchuan, which he then argued was the present geographical centre of gravity of the language family as a whole. The first migration of the language

family out of this area was the Western Tibeto-Burman migration to the fluvial plains of the lower Brahmaputra and the surrounding hill tracts (van Driem 1998: 69). The Western Tibeto-Burman pioneers introduced the technologies of the Eastern Indian Neolithic and were probably the first farming communities of Northeast India. On the basis of these linguistic and archaeological correlations, the hypothesis emerged that the Neolithic culture or the early farming culture of Northeast India emanated from what today is southwestern China, and it was Tibeto-Burman groups of people who migrated from their original homeland and brought farming to Northeast India.

Since then, van Driem's model of Tibeto-Burman prehistory has changed, based first on linguistic evidence and then on supporting human population genetic evidence. Finally he assails the gaps in the archaeological record and the neglect of the Neolithic by archaeologists across vast swathes of the eastern Eurasian heartland. First of all, on the basis of his revised 'Fallen Leaves model' (Fig. 6) of the Trans-Himalayan linguistic phylum, van Driem (2011c: 141) pointed out that the centre of the language family lay not in Sìchuan but in Northeast India. The geographical centre point of the language family may well lie in Sìchuan in terms of the distribution of modern language communities, but, in terms of the distribution of major subgroups or recognised taxa within the Tibeto-Burman language family, the centre of gravity decidedly lies in Northeast India. In this respect, van Driem's reconstruction is growing closer to Peiros (1998), who argued that the Tibeto-Burman homeland must have been located in the sub-Himalayas.

Figure 6: The updated version of the 'Fallen Leaves model' depicting all recognised subgroups of the Trans-Himalayan linguistic phylum (after van Driem 2001, 2014a)

In his new reconstruction, the provenance of the Northeast Indian Neolithic is called into question, and the question is left unanswered because basically archaeology has failed to address this question since the pioneering work of Ahmed Hasan Dani (1960) and Tarun Chandra Sharma (1966) in the 1960s. The later episodes in van Driem's 1998 reconstruction, however, may correspond to the linguistic correlates initially envisaged. The establishment of the Dadiwan, Cishan and Peiligang cultures in the North China plain may represent ancient Sinitic culture, as van Driem (1998) initially proposed, or, perhaps more plausibly, the lure which motivated Proto-Sinitic population groups to move northeast to this area (van Driem 2007d). At a much shallower time depth, the spread of the Majiayao Neolithic from Gansu across eastern Tibet, reflected by sites such as mKhar-ro, as far as the western Himalayas, reflected by sites such as Burzahom in Kashmir, may correspond to the spread of Sino-Bodic, a hypothetical subgroup encompassing subgroups such as Sinitic, Bodish, West Himalayish, Tamangic and perhaps other closely related Tibeto-Burman taxa. Another important feature of van Driem's most recent reconstructions of ethnolinguistic population prehistory is that the various chronological layers at distinct time depths are carefully distinguished. Moreover, the notion of the linguistic event horizon underscores that archaeologists, or at least palaeontologists, and population geneticists can delve further back into prehistory than historical linguists can by means of the comparative method. Large population movements for which the clearest linguistic evidence can be adduced are often very recent, such as the historically attested Hàn spread into southern China from the 3rd century BC or the spread of Tibetic languages across the Tibetan plateau, beginning in the first millennium AD.

In agreement with van Driem's newest reconstructions, Blench (2009, 2011) echoes the view that the earliest speakers of Tibeto-Burman languages were highly diverse foragers living in an arc between the slopes of the Himalayas and Assam and Arunachal Pradesh at around 10,000 years ago. Some may have spoken unknown languages which are now manifest themselves only as relict groups such as Kusunda. Blench (2009, 2011) proposes that these foragers probably began to practise vegiculture such as taro and plantains and arboriculture of sago particularly in Northeast India and management of animals like *mithun* by 6000 BP. Around 5000 BP early Trans-Himalayan groups spread eastwards to China and Sinitic is one of the many migratory groups. These populations encountered other diverse populations with varied cultural backgrounds and agricultural practices. Cereals like buckwheat, foxtail (*Setaria*) and broomcorn (*Panicum*) millets were brought under domestication in the montane areas on the fringes of the Himalayas. Hybridised type of rice is the result of admixture of Indian and Chinese origin. Rice entered the lowland vegicultural zones rather later which pushed the taro into a lingering state (Blench 2009, 2011: 132-133). Yet Blench's reconstruction corresponds only to

one episode in the more complex chronologically layered reconstruction developed by van Driem (2014a). This entire ethnolinguistic reconstruction will not be recapitulated here. Rather, the most relevant portion thereof with respect to the Tibeto-Burman or Trans-Himalayan linguistic phylum is that the origins of the language family as a whole lay in or very near Northeast India at its deepest level. It is more than probable that many migrations involving Tibeto-Burmans in this area, however, belonged to more recent epochs of Holocene prehistory.

Now, let us turn to the most well-informed reconstruction of ethnolinguistic prehistory as regards the Austroasiatic language family. In view of the linguistic paleontological data and epicentre of phylogenetic diversity of Austroasiatic language communities, the geographical centre of gravity of the family may be proposed to have lain in the area around the northern coast of the Bay of Bengal covering the eastern extremity of South Asia and much of the southern littoral of Southeast Asia. On linguistic grounds alone, the original homeland of the Austroasiatic could have lain 'on either side of the Ganges and Brahmaputra delta' (van Driem 2011a: 16-17, 2011b: 361-362, 2012a: 191). However, recent linguistic palaeontological research pinpoints two groups, the ancient Austroasiatics and the ancestral Hmong-Mien, as the most likely candidates for the first cultivators of rice (van Driem 2012a: 193-4). The relationship of both Austroasiatic and Hmong-Mien to rice agriculture and their complex human population genetic relationship to each other, combined with recent advances and new insights into rice genetics, allow us to infer that rice agriculture was an early Austroasiatic technology (van Driem 2012b: 338). Furthermore, van Driem (2012a: 197) suggests that the ancient Austroasiatics may have favoured *Oryza nivara*, whereas the ancient Hmong-Mien may have favoured *Oryza rufipogon.* Both language families robustly reflect rice agriculture terminology (van Driem 2011a: 23, 2012c: 118). The ethnobotanical and rice genetics data also show a long history of rice cultivation in this part of India (for details see Hazarika 2005, 2006a, 2011c, 2014). Finally, and crucially, the origins of Austroasiatic according to the most well-informed interdisciplinary view to date must have lain in or very near Northeast India.

In other words, the Austroasiatic and the Tibeto-Burman language families are not only the most crucial languages for our understanding of the ethnolinguistic prehistory of Northeast India, but the yet largely unexplored archaeology of this ecologically and topographically complex region, which we have identified as a major corridor in the peopling of Asia (Hazarika, 2006b, 2011a, 2011b, 2012, 2013, 2014, 2016), is also crucial to our understanding of the prehistory of both Tibeto-Burman and Austroasiatic. Compared to the archaeological investigations carried out in the Ganges and Yangtze river valleys, northeastern India and particularly the Brahmaputra valley and the Indo-Burmese borderlands have not been explored in a scientific manner to the present day. In future, special attention must be devoted to the recovery of

palaeobotanical and palaeo-environmental information. One likely cause for the lack of evidence is the constant fluvial activities in the Brahmaputra flood plain which may very well have buried the archaeological signatures of the early rice cultivation or washed them out into the Bay of Bengal (van Driem 2011a: 23). However, as most of the prehistoric sites of Northeast India are located in the elevated areas rather than lowlands, and more attention should be given to exploring the foothills regions of the Himalayas and all of the hill tracts of the region in search of potential archaeological sites yielding fruitful data.

Emerging genetic data and its implications

There have been attempts to address the issue of the origin, antiquity and migration of the Northeast Indian tribes by population geneticists. One challenge to the advocates of the Farming Language Dispersal hypothesis (e.g. Diamond and Bellwood 2003), formulated long ago, was that 'the population wave of advance accompanying the spread of early farming should be reflected … in the genetic compositions of the resulting population' (Ammerman and Cavalli-Sforza 1971: 687). In fact, one component of the human population genetic story of our prehistory seems to match well with linguistic dispersals, but does not provide support for the Farming Language Dispersal theory propagated by Renfrew and Bellwood. All of the human population genetic data, however, inasmuch as they have a bearing on Asian ethnolinguistic prehistory underscore the crucial importance of Northeast India and therefore the urgency of a systematic programme of archaeological research throughout the region.

The population genetic story is a complex one, and in this section I shall attempt to recapitulate the argument in a nutshell. Readers who wish a more detailed account may refer to van Driem (2014b), where the present reconstruction is outlined, but opposing views are also addressed. At this juncture, it is useful to point out another difference between the disciplines other than that prehistory is only accessible to historical linguistics at a shallower time depth than it is to archaeology and population genetics. The state-of-the-art arguably advances or, at least, changes more rapidly in human population genetics than it does in either historical linguistics or archaeology. Early genetic studies on Indian populations (e.g. Mukherjee *et al.* 2001, Cordaux *et al.* 2003, Dutta *et al.* 2003), though the findings are still of great utility, now already appear outdated. On the basis of similar results, one group of geneticists at this time inferred that the ancestors of the Chinese migrated to the North China plain from the Himalayas (Chu *et al.* 1998), whilst another group argued that the ancestors of the Tibeto-Burman populations in the Himalayas migrated southwest from the North China plain (Su *et al.* 2000). Often genetic studies merely corroborate what we already know and expect on the basis of linguistics. For example, an autosomal microsatellite study showed that Bodish populations cluster together as against Mizo-Kuki-Chin language communities (Krithika *et*

al. 2007, 2008). The Shompen and other Nicobarese groups are shown to be genetically close to closely related Mon-Khmer groups on the Southeast Asian mainland (Trivedi *et al.* 2006), as opposed to the ethnolinguistically and phenotypically entirely distinct Andamanese, who live on the nearby Andaman Islands. By the same token, population genetic studies often corroborate what archaeologists such as myself have long inferred simply on the basis of facts of geography, namely that Northeast India served as a major corridor for the peopling of eastern Eurasia (Basu *et al.* 2003, Sahoo *et al.* 2006, Kumar *et al.* 2006, Reddy *et al.* 2007). Similarly, as against the hypothesis that the Himalayas acted as a barrier for human movements in the past (Cordaux *et al.* 2004), recent studies show that the Himalayas and especially the Terai acted as a pivotal passageway allowing multiple population interactions in different times (Fornarino *et al.* 2009). Archaeological record also suggests cultural interaction across Himalaya and its borderland since late Quaternary period (Hazarika 2011b, 2012, 2014, 2016).

In the past decade, the molecular clock based on calculated coalescence times used by population geneticists, though not yet as accurate as the dates of archaeologists using calibrated radiocarbon or accelerator mass spectrometry datings, have been improving. Speculations are no longer based just on haplotype frequency gradients. Rooted topologies are accorded due significance, and a higher resolution of molecular polymorphisms has now been attained. The upshot of the newest findings of population geneticists with respect to language families is that there is seldom any or only minor correlation with mitochondrial lineages. Instead, mitochondrial lineages, which reflect maternal ancestry, appear largely to reflect the oldest wave of peopling out of Africa in many areas, although subsequent movements are of course also in evidence. This tendency accounts for the fact that inferences on the basis of mitochondrial DNA have tended to emphasise early layers of population prehistory. For expansion, Hill *et al.* (2006) attribute the expansion of Hoabinhian from southern China and Vietnam into the Malay Peninsula with the arrival of the R9b and N9a mtDNA haplogroups. Soares *et al.* (2008) maintain that a close relationship obtains between the geographical extent of post-Last Glacial Maximum flake-blade industries and the mitochondrial haplogroup E lineages in Southeast Asia. A study on similar lines was conducted in Japan by Tanaka *et al.* (2004), and Peng *et al.* (2011) inferred on the basis of mitochondrial data that southern China and Southeast Asia served as the source of some post-Last Glacial Maximum dispersal. Such weak correlations are not highly informative. Moreover, ancient populations during the Last Glacial Maximum between 26.5 to 19 ka BP generally involved admixture rather than replacement, just as in the case of modern population movements (Chandrasekar *et al.* 2009).

By contrast, Y-chromosomal polymorphisms, which reflect paternal ancestry, correlate astonishingly well with the distribution of major language

families in Asia, although they are by no means congruent. This correlation has been termed the Father Tongue hypothesis. Van Driem is the first to stress that linguistic affinity and biological ancestry are two fundamentally distinct albeit probabilistically correlated entities, and that a molecular polymorphism cannot be construed as being identical with a marker for a particular ethnolinguistic affinity. The list of caveats which he adduces is longer, and the transparency of the argument structure enables inferences to be evaluated and reassessed in the light of emergent population genetic data and also revised in light of new insights from archaeology and historical linguistics as well. In a highly simplified version, therefore, a series of publications has identified the Y chromosomal haplogroup O2a (M95) with the Austroasiatics, the paternal lineage O3a3c (M134) with the Tibeto-Burmans, the paternal haplogroup O3a3b (M7) with Hmong-Mien, and the Y chromosomal lineage O1a (M119) with the spread of Austronesian (van Driem 2007c, 2007d, 2011a, 2011b, 2012a, 2012b, 2014b).

The identification of the Y chromosomal haplogroup O2a (M95) with Austroasiatic was borne out in the genetic study by Chaubey *et al.* (2010), correcting inaccuracies or imprecision in earlier publications e.g. Kumar and Reddy (2003), Reddy *et al.* (2007), Kumar *et al.* (2007). Ancient Austroasiatics must have brought their language into the Indian subcontinent from Southeast Asia, perhaps during the Neolithic period. The paternal lineage of the Munda correlates well with the language, but the maternal lineages are older and indigenous to the Subcontinent. For the purposes of the population genetic study, mainland Southeast Asia begins at the Indo-Burmese borderlands also includes adjacent portions southwestern and southern China, and the exact locus of origin of Austroasiatic within this vast region has not been precisely identified. Previous studies had already shown a complex interaction and genetic exchange between ancient Tibeto-Burmans and ancient Austroasiatics in this region (Cordaux *et al.* 2004, Sahoo *et al.* 2006, Kashyap *et al.* 2006), whereby as much as 47% of the Tibeto-Burman populations in Northeast India also received the paternal lineage which we have hypothetically identified with Austroasiatic, which would indicate that the Tibeto-Burman paternal lineages may have partially replaced by incursive Austroasiatic lineages arriving in Northeast India (van Driem 2007d: 237).

The Y chromosomal haplogroup O3e (M134), associated with the Tibeto-Burman may be tied to one or several refuge areas in the eastern Himalayan region during the last Ice Age (van Driem 2011a: 27-29, 2014b). The identification of this paternal lineage with the Tibeto-Burman language family is corroborated by other studies, e.g. Wang *et al.* (2012). In his reconstruction of the dispersal of Tibeto-Burman from some locus in or near Northeast India and the eastern Himalayan region, there were ancient groups that moved away which led to linguistic descendants such as Bai, Tujia, Qiangic, Ersuic and Sinitic.

There were also groups, whose ancestors might have left the Himalayan region, or at least remained north of the great Himalayan divide, only to return in a later epoch and colonise portions of the southern flanks, such as Bodish, West Himalayish and Tamangic. Yet the majority of Tibeto-Burman subgroups and, indeed, the greater part of linguistic diversity within the Trans-Himalayan phylum appears to have long been a feature of the eastern Himalayan region. Ethnolinguistic diversity has been preserved better in the mountains and hill tracts. By contrast, the Brahmaputran plains and portions of the Arunachal hills may have been more prone to the effects of migrations and regional creoles.

The archaeology of Northeast India is poorly studied. Yet we know that Northeast India acted as a corridor for several multidirectional dispersal events in the late Pleistocene and early Holocene (Hazarika 2012, 2013, 2014). This fact is particularly visible in the dispersal of Hoabinhian traits from Southeast Asia to as far as the northwestern sub-Himalayas through Northeast Indian corridor. These events are reflected in the archaeological record as much as in the molecular genetic evidence. The painstaking research carried out by the late Gudrun Corvinus (2007) in Nepal has provided a cultural sequence right from the Lower Palaeolithic until the Neolithic period. However, it is not clear if the later cultures such as Patu culture in the eastern Sivaliks and the Brakhuti culture in the western Terai can be linked with a particular linguistic group. In Northeast India, the archaeological record of only a few areas like the Garo Hills provided good stratigraphic evidence of late Pleistocene and early Holocene culture. The discovery of Hoabinhian or Hoabinhian-like industry in the Siwaliks, Nepal, the Garo Hills, Manipur, Arunachal Pradesh and Bhutan suggests a population movement across the area (Hazarika 2012, 2013, 2014). A systematic investigation of the archaeological record of this region is urgently needed to shed more light on human movements from Africa across Asia to Australia, Oceania and the Americas. Northeast India served at a strategic thoroughfare, not just once but repeatedly during many distinct episodes of ethnolinguistic prehistory, and it is high time that good modern archaeology with meticulous stratigraphies and sound dating techniques explore this vast key region.

Memories of origin and migration: data from folklore

After a detailed discussion of the historical linguistic and genetic data, let us now turn to some of the folkloristic data prevalent among the Tibeto-Burman and Austroasiatic groups of the region. The folkloristic data is generously recorded in the proceedings of the *North East India History Association* (NEIHA), the foremost academic platform for the historians of Northeast India. The NEIHA volumes II (1981) and IV (1983) have particularly emphasised on the importance of these data in reconstructing history of the lesser known ethnic groups of the region.

There is no general consensus about the origin of the Bodo, one of the branches of the Tibeto-Burman subgroup Brahmaputran (van Driem 2001). This ethnic group forms an important and large section of the population of the Brahmaputra valley. They are considered to be one of the earliest settlers of the valley. Traditional sources hazarded conjectures that even northwestern China or Tibet might have been their original homeland (Endle 1911: 3). Bodo folklore suggests the course of at least one ancestral migration, generally inferred by linguists and ethnographers: A section of the Bodo moved towards the west along the foothills of the Himalayas up to the river Mech between India and Nepal, who later became known as Mech (Sanyal 1973: 2). An example of Bodo folk tradition describes themselves: "Of all the mountains, highest and whitest is the Father (probably referring to the snow-clad Himalayas), of all the rivers, longest and biggest is the Mother (probably referring to the Brahmaputra), we are *Korosa Aris*, First born sea race and our line is continuous" (Mosahary 1983: 46-48). The Bodo language used here appeared to Mosahary (1983: 47-48) to represent a mixture of three different languages, Bodo, Kokborok and Dimasa. The reference to the Himalayas as the father and the Brahmaputra as the mother of all rivers evinces familiarity with the geography of Northeast India. The use of the term *Korosa Aris* 'first born people' or 'first settlers' may suggest that the Bodo see themselves as the first inhabitants of the Brahmaputra valley. Barua (1933: 1-2) has suggested that the term translated as 'sea race' suggests familiarity with the Bay of Bengal and the fact that Bodo-Koch peoples inhabited the low-lying areas of Sylhet, Myemensing and all of eastern Bengal before the advent of the Indo-Aryans. Alternatively, the term translated as 'sea race' could refer even to the Brahmaputra river, for this vast and expansive river is spoken about in this way in, for example, the *Lauhitya Sagara*, (Vasu 1922: 19-20). Lauhitya or Luhit or Luit is another name of Brahmaputra and *Sagara* means sea in Assamese.

Similarly, a possible ancient migration route is suggested in a popular Garo song (Rongmuthu 1997: 222 cited in Sangma 1983: 71) which portends to commemorate that the Garos once occupied eastern and central Tibet (Playfair 1909 (reprint): 9). Another Garo legend narrates the ancient Garo way of life in their original homeland (Marak 1982: 22 cited in Sangma 1983: 70) which tells a story that it was the time where the women and the men still used tree barks or hand-made curved woods and finished cut bamboo culms for covering their private parts and similarly it was the time when the people extracted fire for their use, making sparks out of the dried bamboos and crystal stones. The legends depict a prehistoric way of life. Marak (1982 cited in Sangma 1983: 71-73) maintains that the ancient Garo, or perhaps the ancient Bodo-Koch peoples, branched off into several groups led by different chieftains whose names were Jappa Jalimpa, Sukpa Bonggipa, Auk Raja, Asilik Gitel and Raja Sirampa and migrated up the courses of the rivers Tursa and Tista to Tibet, up to the source of

the river Brahmaputra and the source of Chindwin, Salwin and Irrawady rivers. These ancestral Bodo-Koch people ostensibly entered a place recorded as *A'sng jimjim Chiga Dare Gongdingding*, which Marak identifies with present-day Bhutan, and *Nokcholbari*, purportedly Kalimpong. They carried agricultural implements such as *Janggil ma'rori*, *Ki'me matjanggi*, *Kawa*, *Silcha*, *Gaanti*, *Susuak* and *Gitchi bangje* and yak tails as ornaments. The ancient Bodo-Koch then moved towards the *A'song Patari Chiga Su'unchi*, which Marak identifies with Koch Bihar, and subsequently to the *Rangamati*, where a Garo song describes the life led by their ancestors (Rongmuthu 1997: 224 cited in Sangma 1983: 71): The song described their homeland as a place with a granary of agricultural crops and a storehouse of wealth and property, fruit garden of most precious stones, sanctuary of wild elephants, vast cattle farm and place which was full of rich vegetation (Sangma 1983: 71). Other Garo legends indicate that their ancestors once inhabited all over the Brahmaputran plain (Marak 2004a, 2004b). If there is any truth to these oral traditions, then at the time depth to which they refer the various Bodo-Koch language communities would not yet have been differentiated into groups such as the Garo, who subsequently settled in the Garo Hills. So we must construe these Garo and Bodo as pertaining to the ancestral Bodo-Koch people if there is any veracity to them at all.

The Mizos claim *Chinlung* as their original homeland, but this place has not been identified. Lalrimawia (1981: 26-28), however, has proposed to relate the name to toponyms in China which on romanised maps seem to have similar sounding names. Similarly, Bhattacharjee (1983: 79) fancifully traces the Hmars to an ancient Chinese city state called 'Singlang'. The names of many agricultural implements bear the prefix *kawi-* among the Mizo which Lalthangliana (1977: 9-11) suggests must have been borrowed from the Burmese. Lalrimawia (1981: 28-31) tells us that Khampat in Burma is believed to be the oldest Mizo town, where an earthen rampart is still visible. Mizo lore recalls that famine compelled their ancestors to migrate the Indo-Burmese hill tracts a.k.a. the Chin hills in the 14th century AD. Some early settlements from this period still existence, e.g. Seipui, Suaipui Saihmun, Bochung, all toponyms which are actually also clan names. In the 16th century, the Mizo arrived at present-day Mizoram after crossing the Tiau River. This last wave of migrants from the Chin Hills is also called Lushai. In fact, such dispersal may have been the inevitable result of a lifestyle in which *jhum* cultivation played a central role.

Trade links between Northeast India and neighbouring regions

There is virtually no evidence of any Chalcolithic and Bronze or Iron Age in the Northeast Indian context. The reason behind the absence of subsequent cultural developmental stages from Neolithic to the emergence of early states or kingdom has yet to be addressed. The ancient period of Assam from 4th – 5th century AD witnesses the emergence of several political and cultural centres (Boruah 2007).

Excavations at the sites of Ambari (7th century AD) (Dhavalikar 1973) and the Dhansiri-Doyang valley (5th century AD) (Dutta 2000-2001, Sharma 2007) reveal rich archaeological record with art and architectural pieces, ancient settlements, religious sites and pottery. Sharma (2007) addressed the questions of site formation process, context and cultural links in the Dhansiri-Doyang valley and asserted that the earliest state formation in the region was not necessarily due to the inflow of Indo-Europeans into Assam, but as a result of intermittent trade between India and what today is China, in which Tibeto-Burman speaking communities played a crucial role. The architectural remains of the site of Deo-Parbat near Numaligarh of Golaghat show resemblances with Southeast Asian architecture. Yet the early contacts between India and East and Southeast Asia remain to be elucidated in coherent way.

Noted historian Ramesh Chandra Majumdar (1990: 635) mentions three routes connecting India and China in the ancient period, i.e. the route running across the Hindu Kush to Bactria and then through Central Asia to China, the route from eastern India through upper Burma to what today is southwestern China, the sea route along the Southeast Asian coast. He neglects to mention the many trade routes through the Himalayas themselves. The Kali Gandaki route through Mustang and the route through the Kathmandu valley, for example, are generally held to date at least from Neolithic times. These routes contributed to regular commercial interaction between these two large regions even prior to the Common Era. The Chinese envoy Chang-k'ien observed bamboo and textiles of southwestern China during his visit to Bactria (*circa* 127 BC) being sold in the local market. These objects were brought to eastern India through Yúnnán and Burma and carried to northern India and Afghanistan (Majumdar 1990: 645). The Chinese pilgrim I-Tsing (635–713 AD) who travelled in India during the period 671-695 AD recorded that a group of Buddhist priests travelled from China to India through Burma about five hundred years before his time. Travellers could reach lower Burma through the Arakan and upper Burma through passes in the Patkai range or Manipur hills. Through this route Chinese wares came from Yúnnán and Sìchuan to northern India in the 2nd century BC (Majumdar 1990: 646, 649, 652). One interesting correlations of cultural interaction from either side of the Himalaya may be made on the basis of identical gold masks discovered at the burial site of Malari in Uttarakhand in India and Quta cemetery of Ngari in Tibet. These burial sites show similarities with the burial sites of Nepal particularly Mustang and western Tibet (Bhatt 2011). In this context, it would be pertinent to mention about the discovery of the gold mask at the burial site of Sekta in Manipur by A.K. Sharma (1994). These sites are tentatively dated to early centuries of the Common Era.

Chakrabarti and Lahiri (1986) discuss the Assam-Burma route to China in the early days by considering the historical and archaeological sources. There may have been trade routes linking the Mauryan capital Pataliputra with southern

China through the Brahmaputra valley and the Bhamo area of North Burma (Chakrabarti 2006). Gupta (2006: 90-107) speculates about the Indian Ocean trade network extending overland to Yúnnán through the Assam valley at the beginning of the Common Era, and supports his conjectures with archaeological data from eastern India, Northeast India and southern China. Salles (2004) also argues for overland trade via Northeast India and Burma. In this context Gupta (2006) underscores the importance of sites like Ambari in Guwahati and Sekta in Manipur. Sarma (2006) discusses archaeological data from the Dhansiri-Doyang and Kopili-Jamuna valleys supporting such an overland trade network, emphasising that these two river valleys linked the sites at Sekta and Ambari. Glass beads in the burials at Sekta indicate trade links with mainland India as well as Southeast Asia (Singh 1997: 29). Hsüan-tsang (c. 602–664), the well known Chinese pilgrim to India who reached Kāmarūpa in 638 AD records the profuse use of glass beads in Assam (Kanungo 2006). The Nagas traditionally wear ornaments of marine origin like conch shells and cowries, Indo-Pacific glass beads traded by sea from the southeast Indian coast, and carnelian beads from western India (Kanungo 2006: 155) suggesting a long distance exchange network although the Nagas reside in isolation from the seas.

In later historical times, the trade connecting Tibet with Assam and Bengal through Bhutan involved dyes, *endi* or *eri* cloth, cocoons, areca nuts, tea, tobacco from India and wool, salt and musk from Tibet. Pilgrims from Tibet visited the shrine at Hajo near Guwahati (Rahul 1970: 10). The Mönpa or Memba of the Mechuka valley were middlemen in the trade between Tibet and tribes of Arunachal Pradesh. The Memba themselves procured commodities like dyes, hides, cane and chillies from the Tagin and the Romo and exchanged these for salt and wool in Tibet. The Memba exported rice, corn, ginger, chilly, bamboos, hides, dyes, silk and butter and imported wool, woollen clothes, salt, tea, weapons, tools, copper and brass vessels, sweets, walnuts, peaches, dried cheese, dry meat, religious manuscripts, painted scrolls, images, gold and silver ornaments, precious stone (Billorey 1981: 18-22).

The possible role played by Buddhism for population movements from India to Southeast Asia and vice versa has yet to be investigated. A growing body of archaeological data on Buddhism in Northeast India leads us to suggest that this region may have played a role for the spread of Buddhism in the early days. Surya Pahar in the district of Goalpara, Assam is an interesting site which authenticates the coexistence of Shaivaist, Buddhist and Jain traditions. Situated at the banks of Brahmaputra the site on excavation reveals a sanctum sanctorum besides numerous icons. The site is also marked with rock-engraved figures of Buddha, Shivalingas and numerous icons of gods, goddesses and anthropomorphs in the hilly terrain. Chakrabarti (2006) cites the Sunga bowl with incurved rim as a diagnostic pottery type of ca. 200 BC as a surface find at Tezpur. Daparbatiya in Tezpur and Madan Kamdev near Guwahati are two other

noteworthy sites with ruins of ancient temples and different figures of gods and goddesses, human figurines, animals and decorative designs and motifs. The architectural style in the Brahmaputra valley was influenced by the Gupta architectural style (Sarma 1988). The ruins of the Bhismaknagar and Rukmininagar now in Arunachal Pradesh represent the earliest evidence of contact of the people of this region with the mainland India. The archaeological record of the pre-Ahom period in Brahmaputra valley has been discussed in great details by Choudhury (1985).

Concluding Remarks

From the above discussion it is clear that several independent bodies of evidence all pointing towards several distinct waves of migration into or through Northeast India. A historical linguistic reconstruction of the dispersal of ancient linguistic groups as well as associated genetic and ethnographic data suggests movements of people through Northeast India since the late Pleistocene or early Holocene period, which continued even into later periods.

The present paper is a systematic attempt to address the prehistory of Northeast India by combining multidisciplinary data based on archaeological, linguistic, genetic, and folkloristic information. This study has put forward a strong case for a multidisciplinary approach to archaeological research in areas such as Northeast India, where the archaeological record is extremely fragmentary. The paper empirically has demonstrated the contributions of the Austroasiatic and Tibeto-Burman linguistic communities in the making of the prehistoric scenario of Northeast India. Prehistoric movements of these linguistic groups in different directions throughout Northeast India are in evidence. However, the possible time frame for the migration and the waves of peopling is not entirely clear. Archaeological data support the introduction of new cultural traits rather than a local indigenous development of material culture from pre-Neolithic to the Neolithic in the region, thus suggesting dispersals involving newer populations. The folklore traditions prevalent among resident groups suggest a possible East Asian origin in the form of various strands of oral lore passed down from generation to generation.

The linguist Franklin C. Southworth, who has contributed a volume *Linguistic Archaeology of South Asia* (2005a), argues for collaborative work between archaeologists and linguists: “if what linguists say makes sense to archaeologists … then the door is open for conversations about the ways in which the two disciplines can serve to support, supplement, and question each other’s conclusions. If linguists can produce rigorous reconstructions, which provide close matches to archaeological findings, then prehistorians will have more reason to trust linguistic reconstructions of more intangible things, such as social structure and ideology. Such a dialogue may well lead to further refinements in methods of reconstruction, which will produce even better

matches with the archaeological record. On the linguistic side, the first step in that direction must be to present a clear picture of what can and cannot be done: while linguists may be confident in our ability to reconstruct the *forms* of ancient words, we must convey clearly the inherent problems involved in reconstructing the *meanings* of those forms" (2005b: 21). However, one must also keep in mind the admonition of Lamberg-Karlovsky: "Linguists too often assign languages to archaeological cultures, while archaeologists are often too quick to assign their sherds a language … Linguists cannot associate an archaeological culture with words, syntax and grammar, and archaeologists cannot make their sherds utter words. We need a third arbiter, which may or may not offer some degree of resolution to the relationships between archaeological culture and language … our genes … In the context of a renewed fashion of relating archaeology, culture and language, it is well to remember that neither sherds nor genes are destined to speak specific languages, nor does a given language require a specific ceramic type or genetic structure" (2002: 75).

The absence of absolute dates from the sites of Northeast India creates a big hazard in our attempts to augment our understanding the cultural sequence of the region to the fullest. More archaeological excavations could enable us to obtain sufficient dating material for both the historical as well as the prehistoric period. Excavations may also provide us with sufficient material to examine and analyse the transitional phases and also investigate and understand the probable coexistence of these two distinct phases. Today the first attempts have already been made to address the issue of the origin, antiquity and migration of the Northeast Indian tribes by population geneticists. All of the human population genetic data, however, inasmuch as they have a bearing on Asian ethnolinguistic prehistory underscore the crucial importance of Northeast India and therefore the urgency of a systematic programme of archaeological research throughout the region.

Address for communication

Manjil Hazarika, Department of Archaeology, Cotton College State University, Guwahati, India and Himalayan Languages Project, Institut für Sprachwissenschaft, Universität Bern, Switzerland, email: hazarikam@gmail.com.

Acknowledgements

The present work is based on my doctoral research at the Bern University, Switzerland under the guidance of Prof. Dr. George van Driem. I am grateful to him for his valuable advice, guidance, care and constant encouragement at every time. I am also thankful to my co-supervisor Prof. Dr. Miroslav Novák for his kind suggestions in my work. I would like to express my gratitude to Indian Council of Historical Research for granting me with a Foreign Travel Grant for attending the multidisciplinary conference at Bern. Finally, I am thankful to the organisers Prof. Dr. George van Driem and Dr. Maxwell P. Phillips for inviting me to present a paper at this conference.

References

Ammerman, A. and L.L. Cavalli-Sforza. 1971. 'Measuring the rate of spread of early farming in Europe', *Man* (New Series), 6: 673-688.

Barua, K.L. 1933. *Early History of Kamarupa.* Guwahati: Lawyers Book Stall.

Barua, K.L. 1939. 'Prehistoric culture of Assam', *Journal of the Assam Research Society*, 7(2): 35-41.

Basu, A., N. Mukherjee, S. Roy, S. Sengupta, S. Banerjee, M. Chakraborty, B. Dey, M. Roy, B. Roy, N.P. Bhattacharyya, S. Roychoudhury and P.P. Majumder. 2003. 'Ethnic India: a genomic view with special reference to peopling and structure', *Genome Research*, 13: 2277-2290.

Bellwood, P. 2005. *First Farmers: The Origins of Agricultural Societies*. Oxford: Blackwell Publishing.

Bellwood, P. and C. Renfrew (eds.). 2002. *Examining the Farming/Language Dispersal Hypothesis.* Cambridge: McDonald Institute for Archaeological Research.

Bhatt, R.C., 2011. Malari Cave Burials, Funerary Gold Mask and Evolved Ceramic Tradition in Uttrakhand Himalaya, India. Paper presented at the International Conference on the Prehistory of the Tibetan Plateau at the Center for Tibetan Studies of Sichuan University, August 21-24, 2011.

Bhattacharjee, R. 1983. 'The Hmar tribe – their society and some of its aspects', *Proceedings of North East India History Association*, 79.

Billorey, R.K. 1981. 'Oral history in Northeast India', *Proceedings of the North East India History Association*, II: 14-22.

Blench, R. 2004. 'Archaeology and language: methods and issues', pp. 52-74 in J. Bintliff, ed., *A Companion to Archaeology*. Oxford: Basil Blackwell.

Blench, R. 2006. *Archaeology, Language and the African Past*. Lanham: Alta Mira Press.

Blench, R. 2008. 'Re-evaluating the linguistic prehistory of South Asia', pp. 159-178 in T. Osada and A. Uesugi, eds., *Occasional Paper 3: Linguistics, Archaeology and the Human Past*. Kyoto: Indus Project, Research Institute for Humanity and Nature.

Blench, R. 2009. 'If Agriculture cannot be reconstructed for Proto-Sino-Tibetan what are the Consequences?' *Paper Presented at the 42nd Conference on Sino-Tibetan Language and Linguistics, Chiang Mai*, November 2-4, 2009. www.rogerblench.info

Blench, R. 2011. 'The role of agriculture in the evolution of mainland Southeast Asian language phyla', pp. 125-152 in N.J. Enfield, ed., *Dynamics of Human Diversity*. Canberra: Pacific Linguistics.

Blench, R. 2014. 'The contribution of linguistics to understanding the foraging/farming transition in NE India', pp. 99-109 in T. Jamir and M. Hazarika, eds., *51 Years after Daojali Hading: Emerging perspectives in the Archaeology of Northeast India - Essays in Honour of Tarun Chandra Sharma*. New Delhi: Research India Press.

Blench, R. and M. Spriggs (eds.). 1997-9. *Archaeology and Language*, 4 vols. London: Routledge.

Blench, R.M. and M. Post. (*in press*). 'Rethinking Sino-Tibetan phylogeny from the perspective of Northeast Indian languages' in N. Hill, ed. *Selected papers from the 16th Himalayan Languages Symposium, September 2010*. New York: Mouton de Gruyter.

Boruah, N. 2007. *Early Assam, State Formation, Political Centres, Cultural Zones*. Guwahati: Spectrum Publications.

Burling, R. 1999. 'On "Kamarupan", *Linguistics of the Tibeto-Burman Area*, 22(2): 169-171.

Chakrabarti, D.K. 2006. *The Oxford Companion to Indian Archaeology: the Archaeological Foundation of Ancient India (Stone Age to AD 13th Century).* Oxford: Oxford University Press.

Chakrabarti, D.K. and N. Lahiri. 1986. 'The Assam-Burma route to China', *Man and Environment*, 10: 123-134.

Chandrasekar, A., S. Kumar, J. Sreenath, B.N. Sarkar, B.P. Urade, S. Mallick, S.S. Bandopadhyay, P. Barua, S.S. Barik, D. Basu, U. Kiran, P. Gangopadhyay, R. Sahani, B. Venkata, R. Prasad, S. Gangopadhyay, G.R. Lakshmi, R.R. Ravuri, K. Padmaja, P.N. Venugopal, M.B. Sharma and V.R. Rao. 2009. 'Updating phylogeny of mitochondrial DNA macrohaplogroup M in India: dispersal of modern human in South Asian corridor', *PLoS ONE*, 4 (10): e7447.

Chatterjii, S.K. 1970. *The Place of Assam in the History and Civilization of India*. Guwahati: Gauhati University.

Chaubey, G., M. Metspalu, Y. Choi, R. Mägi, I. Gallego Romero, S. Rootsi, P. Soares, M. van Oven, D.M. Behar, S. Rootsi, G. Hudjashov, C.B. Mallick, M. Karmin, M. Nelis, J. Parik, A. Goverdhana Reddy, E. Metspalu, G. van Driem, Y. Xue, C. Tyler-Smith, K. Thangaraj, L. Singh, M. Remm, M. B. Richards, M. Mirazon Lahr, M. Kayser, R. Villems and T. Kivisild. 2010. 'Population genetic structure in Indian Austroasiatic speakers: the role of landscape barriers and sex-specific admixture', *Molecular Biology and Evolution*, 28 (2): 1013-1024.

Chu, J., 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: 11763-68.

Cordaux, R., G. Weiss, N. Saha and M. Stoneking. 2004. 'The northeast Indian passageway: a barrier or corridor for human migrations?', *Molecular Biology and Evolution*, 21: 1525-33.

Cordaux, R., N. Saha, G.R. Bentley, R. Aunger, S.M. Sirajuddin and M. Stoneking. 2003. 'Mitochondrial DNA analysis reveals diverse histories of tribal populations from India', *European Journal of Human Genetics*, 11: 253–264.

Corvinus, G. 2007. *Prehistoric Cultures in Nepal: From the Early Palaeolithic to the Neolithic and the Quaternary Geology of the Dang-Deokhuri Dun Valleys*, eds. Gisela Freund and Ludwig Reisch, 2 vols. Wiesbaden: Harrassowitz Verlag.

Dani, A.H. 1960. *Prehistory and Protohistory of Eastern India*. Calcutta: Firma L. Mukhopadhyay.

de Souza, J.G. 2011. 'Linguistics, archaeology, and the histories of language spread: the case of the Southern Jê languages, Brazil', *Cadernos de Etnolingüística*, 3 (2): 1-16.

DeLancey, S.C. 1991. 'Sino-Tibetan languages', pp. 445-449 in W. Bright, ed. *International Encyclopedia of Linguistics.* New York: Oxford University Press.

DeLancey, S.C. 2010. 'Towards a history of verb agreement in Tibeto-Burman', *Himalayan Linguistics*, 9(1): 1-38.

Dhavalikar, M.K. 1973. 'Archaeology of Gauhati', *Bulletin of the Deccan College Post-Graduate and Research Institute*, 31-32: 137-49.

Diamond, J. and P. Bellwood. 2003. 'Farmers and their languages: the first expansions', *Science*, 300 (25 April): 597-603.

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

Diffloth, G. 2009. 'More on Dvaravati Old Mon', *Paper presented at the Fourth International Conference on Austroasiatic Linguistics*, Mahidol University at Salaya, 29 October 2009.

Diffloth, G. and N. Zide. 1992. 'Austroasiatic languages', pp. 137-142 in W. Bright, ed., *International Encyclopedia of Linguistics* 1. New York: Oxford University Press.

Dolgopolsky, A. 2008. *Nostratic Dictionary*. Cambridge: McDonald Institute for Archaeological Research.

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

van Driem, G. 1999. 'A new theory on the origin of Chinese', *Bulletin of the Indo-Pacific Prehistory Association, Bulletin* 18 (2): 43-58.

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

van Driem, G. 2003. 'Tibeto-Burman vs. Sino-Tibetan', pp. 101-119 in B. Bauer and G. Pinault, eds., *Language in Time and Space: A Festschrift for Werner Winter on the Occasion of his 80th Birthday*. Berlin: Mouton de Gruyter.

van Driem, G. 2005. 'Sino-Austronesian vs. Sino-Caucasian, Sino-Bodic vs. Sino-Tibetan, and Tibeto-Burman as default theory', pp. 285-338 in Y.P. Yadava, G. Bhattarai, R.R. Lohani, B. Prasain and K. Parajuli, eds., *Contemporary Issues in Nepalese Linguistics*. Kathmandu: Linguistic Society of Nepal.

van Driem, G. 2007a. 'South Asia and the Middle East', pp. 283-347 in C. Mosely, ed. *Encyclopaedia of the World's Endangered Languages*. London: Routledge.

van Driem, G. 2007b. 'Endangered languages of South Asia', pp. 303-341 in M. Brenzinger, ed., *Handbook of Endangered Languages*. Berlin: Mouton de Gruyter.

van Driem, G. 2007c. 'Austroasiatic phylogeny and the Austroasiatic homeland in light of recent population genetic studies', *Mon-Khmer Studies*, 37: 1-14.

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

van Driem, G. 2008. 'Linguistic population prehistory of the greater Himalayan region: interpretations of emergent genetic data', pp. 98-112 in V.G. Tomachev and A.M. Zholdasbekova, eds., *Proceedings of the International Conference 'First Greet Migrations of Peoples' under the auspices of UNESCO (UNESCO Headquarters, Paris, 19 June 2008)*. Paris: The Permanent Delegation of the Republic of Kazakhstan to UNESCO.

van Driem, G. 2011a. 'Lost in the sands of time somewhere north of the Bay of Bengal', pp. 13-38 in M. Turin and B. Zeisler, eds., *Himalayan Languages and Linguistics: Studies in Phonology, Semantics, Morphology and Syntax*. Leiden: Brill.

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

van Driem, G. 2011c. 'The Trans-Himalayan phylum and its implications for population prehistory', *Communication on Contemporary Anthropology*, 5: 135-142.

van Driem, G. 2012a. 'Glimpses of the ethnolinguistic prehistory of northeastern India', pp. 187-211 in T. Huber and S. Blackburn, eds., *Origins and Migrations in the Extended Eastern Himalayas*. Leiden: Brill.

van Driem, G. 2012b. 'Etyma, shouldered adzes and molecular variants', pp. 335-361 in A. Ender, A. Leemann and B. Wälchli, eds., *Methods in Contemporary Linguistics.* Berlin: Mouton de Gruyter.

van Driem, G. 2012c. 'The ethnolinguistic identity of the domesticators of Asian rice', *Comptes Rendus Palevol*, 11 (2): 117-132.

van Driem, G. 2013. 'East Asian ethnolinguistic phylogeography', *Bulletin of Chinese Linguistics*, 7 (1): 1-55.

van Driem, G. 2014a. 'Trans-Himalayan', pp. 11-40 in N. Hill and T. Owen-Smith, eds., *Trans-Himalayan Linguistics.* Berlin: Mouton de Gruyter.

van Driem, G. 2014b. 'A prehistoric thoroughfare between the Ganges and the Himalayas', pp. 60-98 in T. Jamir and M. Hazarika, eds., *50 Years after Daojali-Hading: Emerging Perspectives in the Archaeology of Northeast India*. New Delhi: Research India Press.

Dutta, H.N. 2000-2001. 'Archaeological importance of Doiyang-Dhansiri valley of Assam', *Puratattva*, 31: 92-97.

Dutta, R., V.K. Kashyap, R. Trivedi and T.S. Vasulu. 2003. 'Genomic diversity in North-East, India: genetic relationship among five ethnic populations from Manipur based on VNTR polymorphism', *International Journal of Human Genetics*, 3(1): 21-27.

Endle, Rev. S. 1911. *The Kacharis.* London: Macmillan and Co. Ltd.

Enfield, N.J. (ed). 2011. *Dynamics of Human Diversity: The Case of Mainland Southeast Asia*. Canberra: Pacific Linguistics.

Fornarino, S., M. Pala, V. Battaglia, R. Maranta, A. Achilli, G. Modiano, A. Torroni, O. Semino and S.A. Santachiara-Benerecetti. 2009. 'Mitochondrial and Y-chromosome diversity of the Tharus (Nepal): a reservoir of genetic variation', *BMC Evolutionary Biology*, 9 (154), doi: 10.1186/1471-2148-9-154.

Forster, P. and C. Renfrew (eds.). 2006. *Phylogenetic Methods and the Prehistory of Languages*. Cambridge: McDonald Institute for Archaeological Research.

Gupta, S. 2006. 'Indo-Chinese crossroads: the Assam-Yunnan axis in the Indian Ocean Interaction Sphere (1st-5th century CE)', *Journal of Indian Ocean Archaeology*, 3: 90-107.

Hazarika, M. 2005. *Neolithic culture of Northeast India with Special Reference to the Origins of Pottery and Agriculture*. Unpublished MA dissertation, Deccan College Post Graduate and Research Institute, Pune.

Hazarika, M. 2006a. 'Understanding the process of plants and animal domestication in Northeast India: a hypothetical approach', *Asian Agri-History*, 10 (3): 203-212.

Hazarika, M. 2006b. 'Neolithic culture of Northeast India: a recent perspective on the origins of pottery and agriculture', *Ancient Asia*, 1: 25-43.

Hazarika, M. 2011a. 'Archaeological record of human 'movements' in Northeast India', pp. 288-305 in T. Baruah, ed., *People of Contemporary Northeast India*. Guwahati: Pratisruti Publishers.

Hazarika, M. 2011b. 'Late quaternary archaeological record of Northeast India and its bearing on the prehistory of Himalaya and adjoining regions', *Paper presented at the International Conference on the Prehistory of Tibetan Plateau*, Sichuan University, Chengdu, China during 21-24 August, 2011.

Hazarika, M. 2011c. 'Origin and development of rice agriculture: perspectives from Northeast India', *Paper presented at the International Symposium on Rice and Language across Asia: Crops, Movement, and Social Change*, Cornell University, Ithaca, New York, USA during 22 – 25 September, 2011.

Hazarika, M. 2012. 'Lithic industries with Palaeolithic elements in Northeast India', *Quaternary International*, 269: 48-58.

Hazarika, M. 2013. 'Prehistoric cultural affinities between Southeast Asia, East Asia and Northeast India: an exploration', pp. 16-25 in M.J. Klokke and V. Degroot, eds., *Unearthing Southeast Asia's Past, Selected Papers from the 12th International Conference of the European Association of South East Asian Archaeologists*, vol 1. Singapore: National University of Singapore Press.

Hazarika, M. 2014. *Austroasiatic and Tibeto-Burman Linguistic Dispersals in Northeast India: An Archaeological Perspective*. Unpublished PhD dissertation, Bern University, Switzerland.

Hazarika, M. 2016. 'Tracing Post-Pleistocene Human Movements and Cultural Connections of the Eastern Himalayan Region with the Tibetan Plateau', *Journal of Archaeological Research in Asia*, 5: 44-53.

Hill, C., P. Soares, M. Mormina, V. Macaulay, W. Meehan, J. Blackburn, D. Clarke, J.M. Raja, P. Ismail, D. Bulbeck, S. Oppenheimer and M. Richards. 2006. 'Phylogeography

and Ethnogenesis of aboriginal Southeast Asians', *Molecular Biology and Evolution*, 23 (12): 2480-2491.

Hurles, M.E., E. Matisoo-Smith, R.D. Gray and D. Penny. 2003. 'Untangling Oceanic settlement: the edge of the knowable', *Trends in Ecology and Evolution*, 18 (10): 531-540.

Jin, L., M. Seielstad and C. Xiao (eds). 2001. *Genetic, Linguistic and Archaeological Perspectives on Human Diversity in Southeast Asia.* River Edge, NJ: World Scientific Publishing Co.,

Kanungo, A.K. 2006. 'Indian Ocean and the Naga ornaments', *Bulletin of Indo-pacific Prehistory Association*, 26: 154-162.

Kashyap, V.K., S. Guha, T. Sitalaximi, G. Hima Bindu, Seyed E. Hasnain and R Trivedi. 2006. 'Genetic structure of Indian populations based on fifteen autosomal microsatellite loci', *BMC Genetics*, 7:28, doi: 10.1186/1471-2156-7-28.

Krithika, S., S. Maji and T.S. Vasulu. 2007. 'Geographic contiguity, patterns of gene flow and genetic affinity among the tribes of Arunachal Pradesh, India', *International Journal of Human Genetics*, 7(3): 267-276.

Krithika, S., S. Maji, T.S. Vasulu. 2008. 'A microsatellite guided insight into the genetic status of Adi, an isolated hunting-Gathering tribe of Northeast India', *PLoS ONE*, 3(7): e2549, doi: 10.1371/journal.pone.0002549.

Kumar, V. and B.M. Reddy. 2003. 'Status of Austro-Asiatic groups in the peopling of India: an exploratory study based on the available prehistoric, linguistic and biological evidences', *Journal of Bioscience*, 28: 507-522.

Kumar, V., A.N.S. Reddy, J.P. Babu, T.N. Rao, B.T. Langstieh, K. Thangaraj, A.G. Reddy, L. Singh and B.M. Reddy. 2007. 'Y chromosome evidence suggests a common paternal heritage of Austro-Asiatic populations', *BMC Evolutionary Biology*, 7: 47.

Kumar, V., B.T. Langsiteh, S. Biswas, J.P. Babu, T.N. Rao, K. Thangaraj, A.G. Reddy, L. Singh and B.M. Reddy. 2006. 'Asian and non-Asian origins of Mon-Khmer and Mundari-speaking Austro-Asiatic populations of India', *American Journal of Human Biology*, 18: 461-469.

Lalrimawia. 1981. 'Pre-historic and historic migrations of the Mizos', *Proceedings of the North East India History Association*, II: 23-33.

Lalthangliana, B. 1977. *History of the Mizo in Burma*. Aizawl.

Lamberg-Karlovsky, C.C. 2002. 'Archaeology and language - the Indo-Iranians', *Current Anthropology*, 43 (1): 63-88.

LaPolla, R.J. 2000. 'Subgrouping in Tibeto-Burman: can an individual-identifying standard be developed? How do we factor in the history of migrations and language contact?', *Paper presented at the 33rd International Conference on Sino-Tibetan Languages and Linguistics*, October 2-6, 2000, Bangkok, Thailand. To appear in B. Bickel, L.A. Grenoble, D.A. Peterson and A. Timberlake, eds., *What's where why? Language Typology and Historical Contingency*. Amsterdam and Philadelphia: John Benjamins.

LaPolla, R.J. 2001. 'The role of migration and language contact in the development of the Sino-Tibetan language family', pp. 225-254 in R.M.W. Dixon and A.Y. Aikhenvald, eds., *Areal Diffusion and Genetic Inheritance: Case Studies in Language Change.* New York & Oxford: Oxford University Press.

LaPolla, R.J. 2006. 'Sino-Tibetan languages', pp. 393-397 in K. Brown, ed., *Encyclopedia of language and linguistics, 2nd edition.* London: Elsevier.

Majumdar, R.C. 1990 (6th edition). 'Colonial and cultural expansion', pp. 635-659 in R.C. Majumdar, ed., *The Age of Imperial Unity.* Bombay: Bharatiya Vidya Bhavan.

Marak, C.R. 2004a. 'Introduction', pp. v-xvi in C.R. Marak and S. Som, eds., *Rivers and Culture – Focus on Garo Hills.* Bhopal: Indira Gandhi Rashtriya Manav Sangrahalaya.

Marak, C.R. 2004b. 'Some aspects of the A.Chik culture in the Brahmaputra valley, pp. 1-12 in C.R. Marak and S. Som, eds., *Rivers and Culture – Focus on Garo Hills.* Bhopal: Indira Gandhi Rashtriya Manav Sangrahalaya.

Marak, J. 1982 (second edition). *Garo History.* part 1.

Matisoff, J.A. 1991. 'Sino-Tibetan linguistics: present state and future prospects', *Annual Review of Anthropology*, 20: 469-504.

Matisoff, J.A. 1999. 'In defence of Kamarupan', *Linguistics of the Tibeto-Burman Area*, 22(2): 173-182.

McConvell, P. and N. Evans (eds.), 1997. *Archaeology and Linguistics, Aboriginal Australia in Global Perspective*. Melbourne: Oxford University Press.

Medhi, D.K. 2003. 'Potters and pottery of the Assam region', in J. Miksic, ed., *Earthenware of Southeast Asia*. Singapore: University Press.

Mills, J.P. 1928. *Assam Review,* March.

Moral, D. 1996. 'North-East India as a linguistic area', *Mon-Khmer Studies*, 27: 43-53.

Mosahary, R.N. 1983. 'The Bodos: their origin, migration and settlement in Assam', *Proceedings of North East India History Association,* 42-70.

Mukherjee N., A. Nebel, A. Oppenheim and P.P. Majumder. 2001. 'High-resolution analysis of Y-chromosomal polymorphisms reveals signatures of population movements from Central Asia and West Asia into India', *Journal of Genetics*, 80: 125-135.

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

Peiros, I. 2011. 'Some thoughts on the problem of the Austro-Asiatic homeland', *Journal of Language Relationship,* 6: 101–113.

Peiros, I. and V. Shnirelman. 1997. 'Rice in South-East Asia: regional interdisciplinary approach', pp. 379–389 in R. Blench and M. Spiggs, eds., *Archaeology and Language* 2. London and New York: Routledge.

Peng, M.S., M.G. Palanichamy, Y.G. Yao, B. Mitra, Y.T. Cheng, M. Zhao, J. Liu, H.W. Wang, H. Pan, W.Z. Wang, A.M. Zhang, W. Zhang, D. Wang, Y. Zou, Y. Yang, T.K. Chaudhuri, Q.P. Kong and Y.P. Zhang. 2011. 'Inland post-glacial dispersal in East Asia revealed by mitochondrial haplogroup M9a'b', *BMC Biology*, 9(2).

Pictet, A. 1859. *Les origines indo-europeennes ou les aryas primitifs: essai de paleontologie linguistique* (premiere partie, seconde partie). Joel Cherbuliez, Paris.

Playfair, A. 1909. *The Garos.* Guwahati and Calcutta: United Publishers.

Post, M.W. 2008. 'North East Indian linguistics: present state and future prospects'. *Paper presented at Conference on North East Indian linguistics*, Lakhimpur, 8 September 2008.

Prokop, P. and I. Suliga. 2013. 'Two thousand years of iron smelting in the Khasi Hills, Meghalaya, North East India', *Current Science*, 104 (6): 761-768.

Rahul, R. 1970. *The Himalayan Borderlands.* Delhi: Vikas Publications.

Reddy, B.M., B.T. Langstieh, V. Kumar, T. Nagaraja, A.N.S. Reddy, A. Meka, A.G. Reddy, K. Thangaraj and L. Singh. 2007. 'Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia', *PLoS ONE*, 11: e1141.

Renfrew, C. 1987. *Archaeology and Language, the Puzzle of Indo-European Origins*. London: Jonathan Cape.

Renfrew, C. 2008. 'Preface', pp. v-viii in A. Dolgopolsky, ed., *Nostratic Dictionary.* Cambridge: McDonald Institute for Archaeological Research.

Renfrew, C., A. McMahon and L. Trask (eds.), 2000. *Time Depth in Historical Linguistics*. Cambridge: McDonald Institute for Archaeological Research.

Rongmuthu, D.S. 1997. *Apasong Agana.* Tura.

Sagart, L., R. Blench and A. Sanchez-Mazas (eds). 2005. *The peopling of East Asia: putting together archaeology, linguistics and genetics*. London and New York: Routledge Curzon.

Sahoo, S., A. Singh, G. Himabindu, J. Banerjee, T. Sitalaximi, S. Gaikwad, R. Trivedi, P. Endicott, T. Kivisild, M. Metspalu, Richard Villems, and V. K. Kashyap. 2006. 'A prehistory of Indian Y chromosomes: evaluating demic diffusion scenarios', *Proceedings of National Academy of Science*, 103 (4): 843-848.

Salles, J-F. 2004. 'Archaeology and history of Bangladesh: recent perspective', pp. 185-218 in H.P. Ray and C.M. Sinopoli, eds., *Archaeology as History in Early South Asia*. New Delhi: ICHR and Aryan Books International.

Sanchez-Mazas, A., R. Blench, M. Ross, I. Peiros and M. Lin (eds). 2008. *Human migrations in continental East Asia and Taiwan - Matching archaeology, linguistics, and genetics.* London: Routledge.

Sangma, M.N. 1983. 'Traditions of Garo migration', *Proceedings of the North East India History Association*, IV: 71-78.

Sanyal, C.C. 1973. *The Meches and Totos of North Bengal*. Darjeeling: North Bengal University.

Sarma, P.C. 1988. *Architecture of Assam*. New Delhi: Agam Kala Prakashan.

Sarma, P.J. 2006. 'Roles played by two river valleys of Assam in the Indo-Chinese trade network', *Journal of Indian Ocean Archaeology*, 3: 108-116.

Sarma, P.J. 2007. *Archaeology of Early Settlements in Dhansiri-Doyang Valley, Assam: A Study in Regional Perspective*. Unpublished PhD dissertation, Deccan College Post-graduate and Research Institute, Pune.

Scheinfeldt, L.B. S. Soi and S.A. Tishkoff. 2010. 'Working toward a synthesis of archaeological, linguistic, and genetic data for inferring African population history', *Proceedings of the National Academy of Science*, 107: 8931–8938.

Sharma, A.K. 1994. *Manipur: The Glorious Past*. New Delhi: Aryan Books International.

Sharma, T.C. 1966. *Prehistoric Archaeology of Assam - A Study of the Neolithic Culture*. Unpublished PhD dissertation, University of London, London.

Sidwell, P. 2010. 'The Austroasiatic central riverine hypothesis', *Journal of Language Relationship*, 4: 117–134.

Singh, O.K. 1997. *Stone Age Archaeology of Manipur*. Imphal: Amusana Institute of Antiquarian Studies, Manipur.

Soares, P., J.A. Trejaut, J.H. Loo, C. Hill, M. Mormina, C.L. Lee, Y.M. Chen, G. Hudjashov, P. Forster, V. Macaulay, D. Bulbeck, S. Oppenheimer, M. Lin and M.B. Richards. 2008. 'Climate change and postglacial human dispersals in Southeast Asia', *Molecular Biology and Evolution*, 25 (6): 1209-1218.

Southworth, F.C. 2005a. *Linguistic Archaeology of South Asia*. London and New York: Routledge Curzon.

Southworth, F.C. 2005b. 'Proto-Dravidian agriculture', *Paper presented at the 7th ESCA Round Table Conference*, Kyoto, June 2005.

Spriggs, M. and R. Blench. 1998. 'Correlating archaeological and linguistic hypotheses', in R. Blench and M. Spriggs, eds., *Archaeology and Language II: Archaeological Data and Linguistic Hypotheses*. London: Routledge.

Su, B., C. Xiao, R. Deka, M. T. Seielstad, D. Kangwanpong, J. Xiao, D. Lu, P. Underhill, L. Cavalli-Sforza, R. Chakraborty and L. Jin. 2000. 'Y chromosome haplotypes reveal prehistoric migrations to the Himalayas', *Human Genetics*, 107: 582–590.

Tanaka, M., V.M. Cabrera, A.M. González, J.M. Larruga, T. Takeyasu, N. Fuku, L. Guo, R. Hirose, Y. Fujita, M. Kurata, K. Shinoda, K. Umetsu, Y. Yamada, Y. Oshida, Y. Sato, N. Hattori, Y. Mizuno, Y. Arai, N. Hirose, S. Ohta, O. Ogawa, Y. Tanaka, R. Kawamori, M. Shamoto-Nagai, W. Maruyama, H. Shimokata, R. Suzuki and H. Shimodaira. 2004. 'Mitochondrial genome variation in Eastern Asia and the peopling of Japan', *Genome Research*, 14: 1832-1850.

Trivedi, R., T. Sitalaximi, J, Banerjee, A. Singh, P.K. Sircar and V.K. Kashyap. 2006. 'Molecular insights into the origins of the Shompen, a declining population of the Nicobar archipelago', *Journal of Human Genetics*, 51: 217–226.

Vasu, N.N. 1922. *A Social History of Kamarupa,* vol. 1. Culcutta: Bisvakosh Press.

Wang, H., Y. Li, F. Sun, M. Zhao, B. Mitra, T.K. Chaudhuri, P. Regmi, S. Wu, Q. Kong and Y. Zhang. 2012. 'Revisiting the role of the Himalayas in peopling Nepal: insights from mitochondrial genomes', *Journal of Human Genetics*, 57: 228–234