CHRISTOPHER P. THORNTON AND THEODORE G. SCHURR

GENES, LANGUAGE, AND CULTURE: AN EXAMPLE FROM THE TARIM BASIN

Summary. *The Tarim Basin 'mummies' of western China continue to fascinate scholars and the general public alike due to their 'Caucasoid' features, well-preserved material culture, and putative 'European' origins. However, there have been some uncritical efforts to link these archaeological cultures to those of other ancient Eurasian groups (e.g. the Celts) by applying syllogistic reasoning to multi-disciplinary evidence. In an attempt to provide a more cautious synthesis of the prehistory of the Tarim Basin, this paper will briefly summarize the archaeological, physical, and linguistic evidence that has been used to model human settlement of this region. These data will then be related to recent molecular anthropology research on modern populations of Central Asia, focusing especially on the Uighur in relation to their neighbours. While the genetic history of the modern peoples of a particular region is not necessarily related to their prehistoric antecedents, it is argued that the Tarim Basin experienced a surprising cultural and biological continuity despite immigration from both east and west into Xinjiang Province. This conclusion has a number of possible political ramifications in the present day that must be addressed in future literature on the subject.*

INTRODUCTION

The relationship between genes, language, and culture[1] has been an issue of increasing importance in many fields of research due to the pre-eminence of biomolecular studies (particularly genetics) in both public and academic discourse at the start of the twenty-first century (Pollard 2001). However, there is no consensus among academics over how close this triadic relationship really is. Human geneticists often work under the assumption that biology, language and culture are very closely related, with notable examples of this approach being Sokal *et al*. (1993) and Cavalli-Sforza *et al*. (1994). Like many biological anthropologists, most linguists are aware of the problems in asserting direct relationships of this kind, but there is a visible minority that still utilizes genetic and/or archaeological data uncritically to support its comparative philological arguments (e.g. Parpola 1998). On the other side of the spectrum, many archaeologists today purport no relationship between genes, language, and culture in reaction

1 In this paper, we use the ambiguous term 'culture' to denote its material form; i.e. 'material culture'.

OXFORD JOURNAL OF ARCHAEOLOGY 23(1) 83–106 2004

to the 'pots equal people' mentality of the early twentieth century (e.g. Pluciennik 1996). While these archaeologists are undoubtedly correct in asserting that biology and culture should not be assumed to correlate one-to-one all of the time, in many instances they may be unwittingly throwing the baby out with the bathwater.

To demonstrate how biological research (i.e. genetics and osteometrics) can inform archaeological and linguistic interpretation, this paper seeks to summarize recent data on the Bronze and Iron Ages of Xinjiang Province, western China, in relation to molecular studies on the modern-day inhabitants of this area: the Uighur. Although there is no *a priori* reason why the genetic make-up of the modern peoples of this region should be related to their prehistoric antecedents, it will be demonstrated below that the population history of the Tarim Basin is marked by a surprising cultural and biological continuity, despite immigration from both east and west. This conclusion has a number of possible political ramifications in the present day that cannot be fully elucidated here, but an attempt will be made to examine critically the terminology used by scholars in addressing the prehistory of Xinjiang Province.

ARCHAEOLOGY

The archaeology of the Tarim Basin was catapulted into the media limelight in the mid-1990s with the popularization of the Bronze and Iron Age 'mummies' (actually naturally-desiccated corpses) found by Victor Mair in various regional museums (e.g. Hadingham 1994 or Mair 1995a). Since then, the wonderfully-preserved Tarim Basin mummies have become the focus of much popular and academic attention, due to their 'Caucasoid', 'Western', and 'European' physical appearance (i.e. fair hair, long nose, elongated skulls, high cranial vaults, etc.) and their surprising antiquity (*c*.2000–200 BC) (Mair 1995b; Barber 1999; Mallory and Mair 2000). Of course, this is not the first time that the mummies of Xinjiang have captured the people's interest. In the late nineteenth and early twentieth centuries, legendary European explorers such as Nikolay Przhevalsky, Sven Hedin, and Aurel Stein traversed the region and sent home riveting accounts of their travels to an eager audience (see Mallory and Mair 2000, 62–3). With the expulsion of Western explorers in the mid-twentieth century, the archaeology of the Tarim Basin became something of a *terra incognita* and the once-famous mummies fell into relative obscurity. For this reason, Victor Mair should be commended for reawakening international interest in the history of such an important area.

Although of incalculable importance, many of these famous mummies were not excavated with an attention towards the assemblage of goods or the burial types. As a result, we must focus on other finds for an archaeological synthesis. While a full review of the archaeological literature on the Xinjiang Province is unnecessary in light of the excellent treatises presented by Debaine-Francfort (1988), Mu (1992), Chen and Hiebert (1995), Wang (1996, 2001) and in the collection of papers edited by Mair (1998a), a brief summary of a few Bronze and Iron Age material cultures is essential to understanding many of the more controversial claims made about the history of the Tarim Basin.

The earliest date of human occupation in eastern Central Asia is unknown. However, there is some evidence of Mesolithic hunter-gatherers being present in the Pamir and Tian Shan mountains to the west and north by *c*.9000–8000 BC (Mallory and Mair 2000, 135). Similarly, the presence of Neolithic settlements or burials has never been documented, although surface scatters of microliths may be indicative of pre-Bronze Age occupation (Debaine-Francfort 1988). The first attested sites of the Tarim Basin are settled, agricultural communities in the

Bronze Age of the early second millennium BC (Mei 2000, 8). It is unclear whether this absence of evidence is truly evidence of absence, but the second millennium development of oasis-based agriculture is nearly contemporary with (but a few centuries later than) the rise of oasis-based agricultural communities in Bactria and Margiana of southern Central Asia (the Oxus civilization or 'BMAC;' see Chen and Hiebert 1995, 287) (Fig. 1). While there is no definitive evidence of contact between southern and eastern Central Asia until the late second millennium BC, Chen and Hiebert (1995) suggest that the *idea* of oasis-based agriculture may have spread from the BMAC region with the incursion of nomadic groups (in particular, the 'Andronovo')[2] into the Tarim Basin in the early second millennium BC (cf. Barber 1999).

Opponents of the 'Bactrian' model for Bronze Age origins in Xinjiang argue instead that the initial immigrants were part of the Afanasievo culture from the Altai–Minusinsk regions north of the Tarim Basin (e.g. Kuzmina 1998; Mallory and Mair 2000). This 'steppe' model rests upon the assumption by Russian scholars and others (e.g. Anthony 1998, 2001) that the Afanasievo culture derives ultimately from the Yamnaya culture of the Pontic Steppe region, due to similarities in burial types, grave goods, and wheeled vehicles (but see Frachetti 2002). This relationship is then extended into Xinjiang, where proponents of the 'steppe' hypothesis suggest parallels between the Afanasievo culture and the Gumugou (Qäwrighul) culture (*c*.2000–1500 BC) of the eastern Tarim Basin (Kuzmina 1998, 69–70). Regardless of whether the 'Bactrian' or 'steppe' hypotheses are correct, the general consensus is that Neolithic subsistence almost certainly did not come to Xinjiang from the east because the domestic plants and animals being utilized (sheep/goat, wheat, and barley) are distinctly western in origin (Mallory and Mair 2000, 135).

The first attested archaeological finds (including the earliest 'mummies') in the Tarim Basin date to the Bronze Age period, and are associated mainly with cultures to the south-east (the Gumugou) and north-east (the Yanbulake and Xintala) (Mallory and Mair 2000, 136–45). In addition, there is evidence of Bronze Age occupation along the southern rim of the Tian Shan mountains (the Haladun or Karadong culture), but few excavations have been reported from this region and, thus, will not be discussed here (see Chen and Hiebert 1995, 267–8). The Gumugou is probably the oldest of these cultures and notable for burials that are very similar to those of the Afanasievo culture (Kuzmina 1998, 69–70). However, there are three components that differentiate these burials from pure Afanasievo types: the lack of horses; the presence of Bactrian camels; and the lack of pottery (Chen and Hiebert 1995, 253). While the first two may provide evidence for the 'Bactrian' hypothesis, the lack of pottery is, for the time being, unexplained. It is worth noting, however, that similar burial traditions (i.e. using wooden vessels, instead of ceramics) continue in this region through the Iron Age and into the Han period (Chen and Hiebert 1995, 257), suggesting the long-term continuity of a local style that is distinct to eastern Central Asia.

Like the Gumugou culture, the Bronze Age Xintala (Yengidala) culture (*c*.1700–1300 BC) shows greatest similarities with the Steppe region, although more so with the Andronovo than with the Afanasievo (Mallory and Mair 2000, 145). This is particularly true for the metal artefacts, which display a level of technological skill previously unknown in the Tarim Basin, but which are very similar in form to Andronovo metal objects (Kuzmina 1998, 74; Mei 2000, 72). Although evidence of such long-distance interaction is not firmly attested, the Xintala

2 See Lambert-Karlovsky (2002, 64–5) for a discussion of the problems inherent to the concept of 'Andronovo' super-culture.

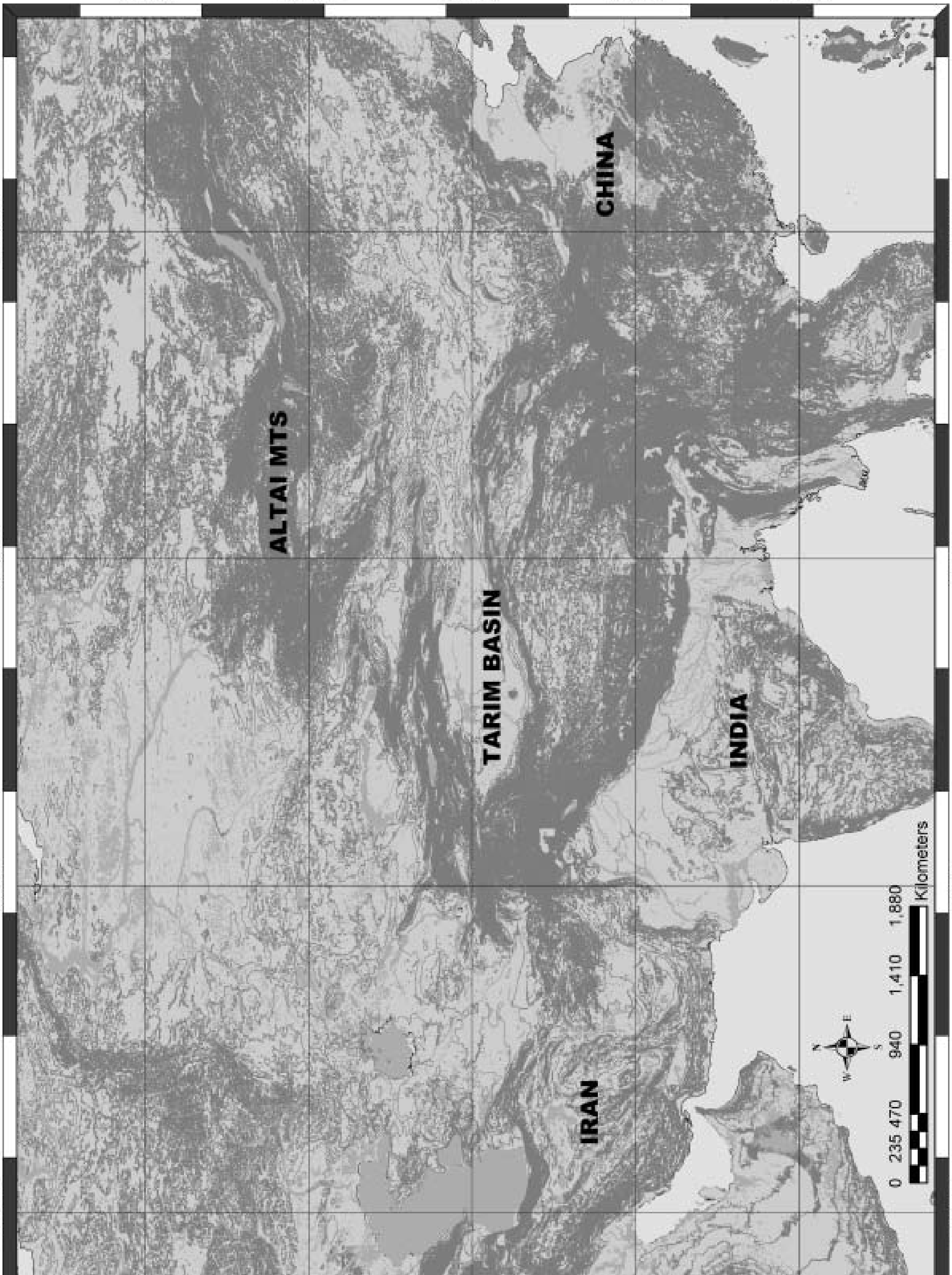

Figure 1A
A map of Eurasia showing the location of the Tarim Basin.

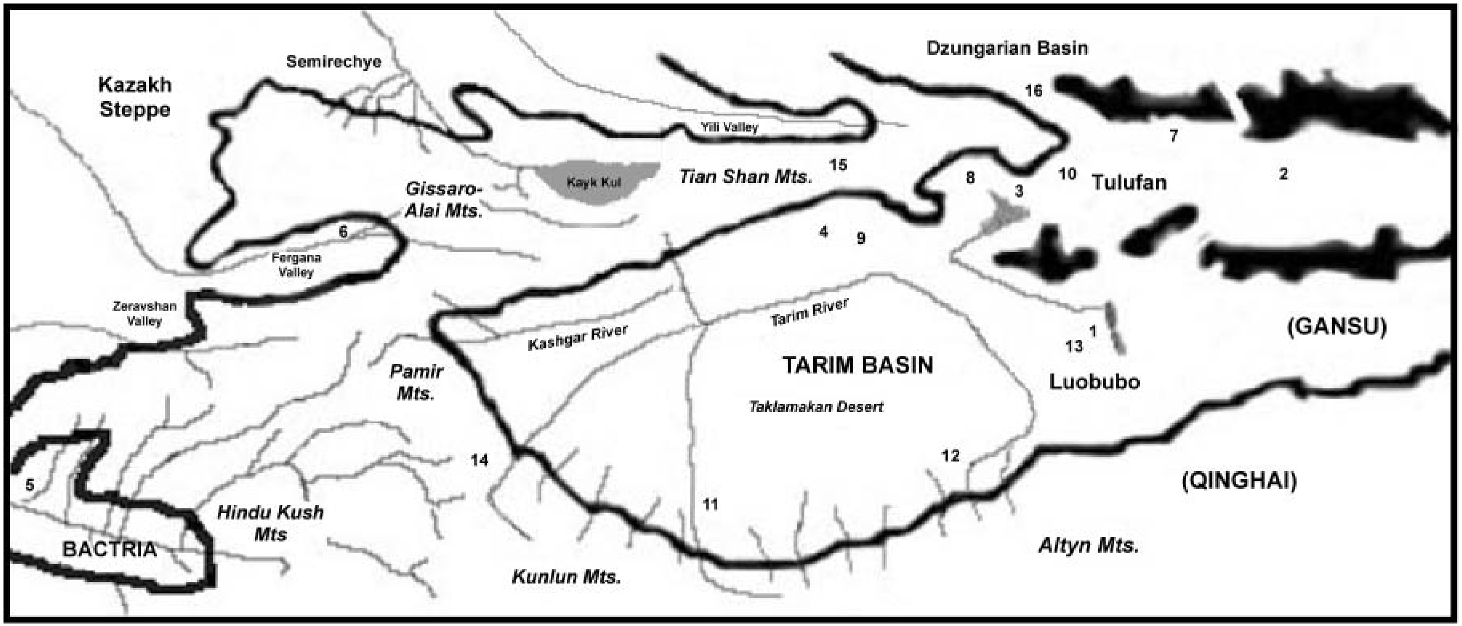

Figure 1B
Bronze and Iron Age sites mentioned in the text: 1. Gumugou; 2. Yanbulake; 3. Xintala; 4. Haladun; 5. Sapalli/Djarkutan; 6. Chust; 7. Sidaogou; 8. Chawuhugoukou; 9. Qunbake; 10. Alagou; 11. Shanpula; 12. Che'erchen; 13. Loulan; 14. Xiangbabai; 15. Zhaosu; 16. Ulumuqi (drawn by authors).

culture was well situated for contact with both the highlands of the Tian Shan and the eastern cultures of the Gansu Corridor (Chen and Hiebert 1995, 267; Debaine-Francfort 1995).

The other major Bronze Age culture, the Yanbulake (or Yanbulaq), demonstrates a burial tradition that, like Gumugou, continues into the Iron Age (*c*.1700–500 BC)[3] (Mallory and Mair 2000, 140–3). The pottery associated with the three contemporary types of burials is handmade with some stylistic similarities to the Siba and Xindian cultures of the Gansu Corridor and the Chust culture of the Fergana Valley in western Central Asia, although a few pieces are reminiscent of Andronovo or Afanasievo style (Chen and Hiebert 1995, 262; Li 2002). While the contemporary Xintala culture may provide evidence to support the 'steppe' hypothesis, the Yanbulake culture's possible architectural context, use of mud bricks, and burial style (type 1) demonstrate close parallels with the BMAC of southern Central Asia (Hiebert, pers. comm.). According to Chen and Hiebert (1995, 264), 'if there is any place in eastern Central Asia that may have had direct contact with the oases of the Oxus civilization, it is at Yanbulake.' It should be noted, however, that Chinese scholars (e.g. Li 2002) continue to maintain the lower chronology for this culture and see its fluorescence as directly related to influence from the eastern cultures of the Gansu Corridor.

The Late Bronze Age–Iron Age transition period (late second/early first millennium BC) was a time of significant socio-cultural change for the populations of the Tarim Basin (Mallory and Mair 2000, 147). First and foremost, this area experienced a shift from oasis-based agricultural subsistence to a horse-riding agro-pastoral way of life (Chen and Hiebert 1995, 285). Second, the metallurgical inventory expanded to include iron (probably brought from the Chust culture of western Central Asia), although copper-base alloys continued to be used and, for perhaps the first time, actively produced through the first millennium BC (Di Cosmo 1996, 91–4; Mei 2000). Third, the Bronze Age cultures of the north-east and northern rim became incorporated into four fairly distinct cultures (the Sidaogou, the Chawuhugoukou, and the Qunbake, with the poorly-understood Alagou culture situated between the first two), while the relatively unknown southern Tarim Basin began to fluoresce due to early Silk Road trade (e.g. the Shanpula and Che'erchen mummies) (Chen and Hiebert 1995, 270; Mallory and Mair 2000, 153–61; Mei and Shell 2002). Fourth, the mummies' remarkably preserved textiles underwent a dramatic shift from the modest, earth-toned wool and felt clothing of the Bronze Age to the complex colourful weaves of the Iron Age (Good 1998; Barber 1998, 1999). The famous plaid-design twill textiles that Barber (1998, 651) describes as 'proto-Celtic' due to their similarities to Hallstatt examples from Austria date from this period. Finally, in the first millennium BC, the highland regions to the north and west became connected by a series of comparable cemeteries (with Xiangbabai being the farthest south) often referred to as the Saka (Sai) culture (see Davis-Kimball *et al*. 1995; Chang *et al*. 2002), which had a great influence on the Iron Age and Han periods of Xinjiang (Chen and Hiebert 1995, 281–3; Mallory and Mair 2000, 162–3).

Although the first millennium was an era of great change in the Tarim Basin, the archaeological record suggests that there is also evidence of long-term cultural continuity within regional Bronze Age and Iron Age cultures (Chen and Hiebert 1995; cf. Mallory and Mair 2000, 152). It is only with the influx of the Han in the second century BC that the rise of the Silk Road significantly disrupted the cultural continuity of eastern Central Asia. Indeed, unless we

3 The dating of the Yanbulake cemetery is extremely contentious (see Mei 2000, 8). The majority of radiocarbon dates from the site cluster in the second millennium BC (*c*.1700–1100 BC), although the excavators assign the burials stylistically to the Bronze Age/Iron Age transition period (*c*.1300–500 BC).

take a fully migrationist position (*à la* Anthony 1990), there is no reason why these dramatic cultural shifts could not have had a mostly autochthonous origin while allowing for some outside impetus through trade contacts and small-scale immigration. As will be discussed in the next section, this argument may be supported by the biological evidence as well.

PHYSICAL ANTHROPOLOGY

The use of bone measurements to induce classifications (i.e. 'osteometry') is fairly standard practice in zoology and other fields of macrobiology including hominid evolution. However, when these methods are applied to human (i.e. *Homo sapiens sapiens*) samples, the scientist becomes embroiled in the impassioned debate over the concept of biological 'races' (see Cartmill 1999). Many archaeologists reject outright the study of craniometry (i.e. skull measurements) as a valid means of determining ancient population movements because of the unclear relationship between environment, diet, phenotypes and genotypes (e.g. Lamberg-Karlovsky 2002, 64), while others simply tread cautiously when discussing such data (e.g. Renfrew 1992, 459–60). Western physical anthropologists are similarly divided on the issue of whether craniometry can serve as a direct indicator of geographically-based origins, although non-Western scholars have few qualms about assigning cultures to people based on cranial and dental morphological characteristics (see Zubrov 1996 for a Russian perspective).

In this section, we will review the results of craniometric work on Tarim Basin burials from the Bronze and Iron Ages, performed by a Chinese and an American physical anthropologist, using their terminologies. Although the causes that influence variation in cranial shape are not fully understood, the data being discussed here are assumed to be methodologically sound. Therefore, they will be used tentatively to discern biological differences between and within members of the ancient cultures of eastern Central Asia.

The pre-eminent expert on the craniometry of the prehistoric Tarim Basin is Kangxin Han of the Institute of Archaeology in Beijing (Mallory and Mair 2000, 236–43). After analysing over 300 skulls from Bronze and Iron Age burials, Han (1998) identified four major cranial shapes, which he labelled (following former Soviet classifications) Proto-European, Pamir-Fergana, Eastern Mediterranean, and Mongoloid. The first three of his types are subsumed under the category of 'Caucasoid' by Hemphill and Mallory (in press), which Han (1998, 559) defines as having an 'elongated, narrow, and high cranial vault, with relatively low and wide facial dimensions and strongly projecting nasal bones'. A 'Mongoloid' cranium, on the other hand, is defined as demonstrating an elongated and broad cranial vault, with relatively flat, high, and wide facial features and a low nasal projection (Han 1998, 562).

While such racial classifications are of questionable value in and of themselves, they are being viewed here as heuristic categories that help to show phenotypic clinal variation in Eurasian populations from east to west and north to south. Hence, from this perspective, these skeletal differences may reveal the relative genetic influences of eastern and western Eurasian peoples on ancient populations of the Tarim Basin as well as the approximate timing and nature of such interaction when placed in archaeological context.

The only Bronze Age crania examined by Han come from the type-sites of Gumugou and Yanbulake, the latter of which also includes burials from the Iron Age. All of the Gumugou crania are classified as 'Proto-European' (Han 1998, 559–60). However, Han has previously suggested a direct association between Gumugou and the Afanasievo and, for later burials, with the Andronovo (Chen and Hiebert 1995, 253 and 257). In contrast, the majority of Yanbulake crania are considered 'Mongoloid' (21 of 29), with the rest being 'Caucasoid' (Han 1998, 561; Chen and Hiebert 1995, 263). Of great interest is the assertion by Mallory and Mair (2000, 141) that 'Type I' burials at Yanbulake (i.e. shaft-pit graves, usually multiple interments placed in mud-brick tombs on platforms with abundant painted pottery) are exclusively of 'Mongoloid' individuals, while 'Type II' (i.e. shaft-pit graves with no platform, no mud-brick, single interments, and little painted pottery) and 'Type III' (i.e. mud-brick tombs at near-surface level, again with little painted pottery) burials were of 'Caucasoid' cranial type. Since neither Chen and Hiebert (1995) nor Han (1998) mention this association between burial types and cranial types, the original source of this information is unknown.

The Iron Age crania examined by Han (1998) come from the sites of Alagou (Alwighul) and Chawuhugoukou (Charwighul) in the north-east Tarim Basin (*c*.1000–400 BC), Zhaosu (Mongghul Kürä) and Xiangbaobao (Shanbabay) of the highland Saka culture (*c*.900–200 BC), and Shanpula (Sampul) and Loulan (Krorän) (*c*.200 BC–AD 200) in the southern Tarim Basin. The crania from Alagou represent the most diverse assemblage analysed by Han (1998, 560–1), with examples from all four of Han's racial types being present, including some showing intermediate features between 'Mongoloid' and 'Caucasoid' types. However, the 77 individuals from the proximal and synchronic site of Chawuhugoukou are characterized as entirely 'Caucasoid' (Han 1998, 564). To complicate the picture further, 17 contemporary examples from the more easterly site of Sidaogao were characterized as exclusively 'Mongoloid' (Chen and Hiebert 1995, 274), although these were not analysed by Han and, therefore, are not necessarily comparable.

The Saka (and/or Wusun) site of Zhaosu lies in the Tian Shan mountains west of Chawuhugoukou and dates from the later Iron Age period. The crania from this site are predominantly 'Caucasoid', although two females are classified as 'Mongoloid' (Han 1998, 563–4) but Han's criteria for sexing individuals is not made explicit. In contrast, the single cranium examined from the Saka cemetery of Xiangbabai is of 'Eastern Mediterranean' type, which connects it to the slightly later Silk Road site of Shanpula, where all 56 crania examined by Han (1998, 563) were also classified as 'Eastern Mediterranean' type. At the similar Silk Road site of Loulan, five male crania were also classified as 'Eastern Mediterranean' while a single female was labelled 'Mongoloid' (Han 1998, 562). In general, the trend from the Bronze Age to the early Han period is marked by a decrease in 'Mongoloid' and Han's 'Proto-European' cranial types, and an increase of 'Pamir-Fergana' and especially 'Eastern Mediterranean' types, which most closely resemble cranial types of the southern Pamirs and southern Central Asia (Mallory and Mair 2000, 243). Han (1998, 568) concludes that this immigration of 'Eastern Mediterranean' types is related to the origins of the Silk Road in the latter part of the first millennium BC.

Recent research by Brian Hemphill of California State University-Bakersfield has presented some intriguing new data on the origins of Xinjiang cultures (Hemphill and Mallory in press) by comparing crania from the Tarim Basin with roughly contemporary examples from the Steppe, southern Central Asia, and the Middle East (Hemphill 1999). The aim of this study was to determine whether either of the two origin hypotheses described above ('Bactrian' vs. 'steppe') could be supported with craniometric analysis. Unlike Han, who classified crania based on ethno-racial categories (e.g. 'Nordic', 'Alpine', etc.; Mallory and Mair 2000, 238), Hemphill used various statistical methods (including weighted pair average linkage analysis ('WPGMA'), neighbour-joining cluster analysis ('NJCA'), and multidimensional scaling analysis ('MDS')) to compare Tarim Basin samples from Gumugou, Alagou, and Loulan with possible origin populations, as well as with a modern population from southern China.

The general conclusion reached by Hemphill and Mallory (in press) was that neither the 'Bactrian' nor the 'steppe' hypotheses were fully supported by the statistical analysis of the craniometric data (see also Mallory and Mair 2000, 244). This was especially true for the 'steppe' hypothesis. Indeed, the Afanasievo and Andronovo samples of the Altai and Minusinsk regions remained phenetically distinct from any of the Tarim Basin crania. Although the Alagou and Loulan samples clustered consistently with samples from sites in southern Central Asia that were a millennium or two older (Sapelli depe and Djarkutan), only the crania from Gumugou followed suit in the NJCA. In all other analyses, Gumugou was a phenetic outlier that appeared to cluster with Indus Valley crania, although Hemphill (pers. comm.) does not consider this association to be statistically valid. While Hemphill and Mallory (in press) discount the 'Bactrian' hypothesis based on the isolation of the Gumugou samples, the relationship between the Margiana crania and the Alagou and Loulan individuals may reflect the increased cultural contact between southern Central Asia and the Tarim Basin in the Late Bronze Age/Early Iron Age transition period discussed above.

There is another conclusion that can tentatively be inferred from the data presented in Hemphill and Mallory (in press). Most cranial samples from the same discrete geographic region remain fairly similar relative to crania from other regions despite the passing of time and changing cultural patterns. This inference suggests at least two possible explanations: either the populations remained biologically stable despite increased intercultural interactions and changing lifestyles; or, the cranial features analysed by Hemphill have been shaped by the local environment, thereby nullifying any immigrant craniometric signature. As mentioned above, the evidence for long-term cultural continuity suggested by Chen and Hiebert (1995) may support the former hypothesis, although the latter also has scientific credibility.[4] As will be discussed in the next section, the idea of a biologically and culturally stable population lends some credence to the oft-discussed 'Indo-European' hypothesis for the origins of the Tocharians in the Tarim Basin.

HISTORICAL LINGUISTICS

It may seem preposterous that most studies into the history of the Xinjiang Province try to determine what languages the mummies spoke and how they tie into the great 'Indo-European' debate. However, this association is not as arbitrary as it may sound. In the nineteenth century, European explorers returned from Xinjiang with Buddhist manuscripts dating to the sixth-eighth centuries AD that had been written in languages and/or dialects of a previously unknown linguistic family: Tocharian (Mallory 1989, 56). The most surprising aspect of this new-found language family, which historical linguists can only reconstruct back to around 500–400 BC (Lubotsky 1998, 380), was that it is an extinct branch of the Indo-European family. Although considered by most scholars to be related to western Indo-European languages such as Celtic and Italian (e.g. Hamp 1998; Ringe *et al*. 1998), others have suggested that Tocharian may be closer to the Indo-Iranian languages of the Central Asian Steppe (e.g. Renfrew 1998, 209).

Despite our ability to read all three Tocharian languages,[5] we still have no idea who spoke them and when they were used (Mallory and Mair 2000, 333–4). Many Chinese scholars

4 This is an assumption based on the work of Boas (1911). However, see Sparks and Jantz (2002) for a recent rebuttal.

5 These are Tocharian A, B, and C. While the first two are the closest, they are still less similar to each other than Italian is to Romanian, although most scholars consider A to have been much older than B (Mallory and Mair 2000, 274–80).

(e.g. Lin 1998) associate Tocharian-speakers with the Ta-Yuezhi, a nomadic residents of the Tarim Basin mentioned in Chinese texts of the Han period. As told to the Han emperor by the great explorer Zhang Qian, the Yuezhi were defeated by the Hsiung-nu nomads (early Huns?) in the second century BC, at which point they moved westward out of Xinjiang to the land of 'Daxia', which they conquered (Lin 1998, 476; but see Mallory and Mair 2000, 333–4). The association between the Yuezhi and the Tocharian-speakers is then drawn from Strabo, who wrote of four nomadic tribes including the 'Tocharoi' who defeated one of the Greek rulers of Bactria ('Daxia'?) in the second century BC (Renfrew 1987, 65).

Hypotheses on how an Indo-European language came to be spoken by the Yuezhi are even more speculative. One of the more extreme versions is Henning's (1978) assertion that the Guti ('Yuezhi'?) tribes, who spoke proto-Tocharian, fled to eastern Central Asia after being expelled from Babylonia in the late third millennium BC (Mallory and Mair 2000, 282; Lamberg-Karlovsky 2002, 74). Haskins (in Opie 1995) took this argument a step further when he proposed that the Yuezhi/Guti were the 'Massagetai' of Herodotus ('Massa-getai' means 'Great Getai (Guti?)' just as 'Ta-Yuezhi' means 'Great Yuezhi'), a people that he assigned to the Altai region of the eastern Eurasian Steppe.

Besides this problematic use of linguistic and historical data, there are two immediate problems with the Yuezhi–Guti–Tocharian analogy. First, the languages translated from the documents discovered in the Tarim Basin were labelled 'Tocharian' arbitrarily by western scholars wishing to associate them historically with the people mentioned in Greek and Indic texts as the Tocharoi/Tukhara (Renfrew 1987, 65; see also Mallory and Mair 2000, 333). As a result, they merely created their own evidence. Second, there is nothing in the texts or anything found on the mummies themselves to suggest an association between these important archaeological finds and this extinct linguistic family. In fact, 'Yuezhi', 'Guti', and 'Tocharoi' are more likely to be collective ethnonyms placed upon non-literate 'barbarian' tribes by their literate neighbours in the way of 'Celts', 'Scythians', and American 'Indians'. Despite the inherent problems with the name, we can be fairly confident through self-referential passages that speakers of the 'Tocharian B' language lived in the northern Tarim Basin some time before the invasion of the Turkic-speaking Uighur tribes in the eighth century AD (Mallory and Mair 2000, 280).

Regardless of what these extinct Indo-European languages are labelled, it is still possible to assess the extent of linguistic exchange between Tocharian-speakers and other linguistic groups, which may reflect the amount and direction of cultural interaction. For example, the passing of loan words from Tocharian to Old Chinese (e.g. the word for 'honey', which occurred some time in the first millennium BC), is extremely limited (Lubotsky 1998). This fact seems to reflect the archaeological and biological evidence discussed above, which suggests only limited influence from the east into the Tarim Basin. On the other hand, the rise of the Silk Road is reflected in the increase of words borrowed by Tocharian from Indo-Iranian languages such as Avestan, Sanskrit, Gandhari, Khotanese Saka, and Sogdian (Mallory and Mair 2000, 254–5). Although these patterns are intriguing, historical linguistic studies of eastern Central Asia languages do little to further our understanding of the Bronze and Iron Age cultures of Xinjiang for the simple reason that they were prehistoric, i.e. non-literate. However, they leave us with an important question: how did a language with ties to European, Anatolian, and Indo-Iranian languages end up in the most remote region of the Asian continent?

MOLECULAR ANTHROPOLOGY

The method purported by numerous scholars as the answer to all problems 'Indo-European' is genetics, despite the fact that this question is inherently a linguistic one. Attempts to assign direct relationships between biology and language have a history far earlier than the work of Cavalli-Sforza *et al*. (1994), including Griffith Taylor's (1921) correlation between the cephalic index of a cranium and a particular Indo-European language (see also Mallory and Mair 2000, 233). Citing the work of Sokal (1988), Mair (1998b) purports that the true origins of Tarim Basin cultures can *only* be established by analysing DNA extracted from the mummies. In turn, these data will explain how Tocharian came to be spoken in Xinjiang.

However, this sort of syllogistic argument relating genes, language, and culture does not necessarily work. While language or culture *can* serve as genetic barriers, thereby leading to biologically and socio-culturally differentiated populations, these boundaries are not defined based on scientific 'facts' but on *perceived* socio-cultural and/or biological differences such as 'ethnicity' or 'race' (see Mirza and Dungworth 1995; Macbeth 1997). Indeed, there are many famous examples of significant linguistic transitions marked by little cultural or genetic distinction, e.g. in Hungary (Semino *et al*. 2000) and Turkey (Comas *et al*. 1996), while there are other examples of linguistic affinities between groups who have significant genetic and/or cultural differences, such as Sardinians and Italians (Francalacci 1998, 544).

Despite its limitations, genetics does provide the best indication of population history, especially when using non-recombining, sex-linked genetic systems such as mitochondrial DNA (mtDNA) and Y-chromosome DNA (NRY) from ancient samples (i.e. 'archaeogenetics'). Until recently, ancient DNA data had been published from only one Tarim Basin mummy due to political obstacles during sample collection. This individual was found to belong to mtDNA haplogroup H and to have an HVR1 sequence that is identical with that of the 'Cambridge Reference Sequence' (CRS)[6] (Francalacci 1998). As it turns out, haplogroup H is the most frequent mtDNA lineage in most modern European populations, appearing at ~40 per cent in these groups, and the CRS is the most common mtDNA HVR1 sequence within this haplogroup (Comas *et al*. 1996; Richards *et al*. 2000; Di Benedetto *et al*. 2001). Mallory and Mair (2000, 247) highlight this fact to support their claim that the Tarim Basin mummies are 'Europeans'. However, it should be noted that haplogroup H is also found among populations of the Near East, Central Asia, and South Asia (Comas *et al*. 1998; Kivisild *et al*. 1999; Richards *et al*. 2000), and, in fact, has a number of diagnostic polymorphisms that are not usually targeted in archaeogenetic research (Finnilä *et al*. 2001; Herrnstadt *et al*. 2002). Therefore, the high frequency of this mtDNA lineage in modern European populations does not necessitate a European origin or ancestry for the Tarim Basin mummy analysed by Francalacci (1998).

A more extensive archaeogenetic study[7] has recently been completed by Dr. Yinqiu Cui (2002) at Jilin University in China. Dr. Cui analysed mtDNA extracted from 11 Bronze Age mummies from Luobunoer (i.e. Lopnur, *c*.3800 BP) and 20 Iron Age mummies from Tulufan (i.e. Turfan, *c*.2500 BP). Her comparison of these ancient samples with modern mtDNA from

6 Haplogroup H was also prevalent in the 15 DNA samples extracted from Bronze Age burials in the Kargaly region of the Ural Mountains (Kulikov and Poltaraus 2002).

7 This important dissertation was kindly provided to us by Victor Mair.

Central Asian, East Asian, and European populations provides some intriguing conclusions. First, NJCA based on genetic distances ('D_A') between the Luobunoer and other samples showed that the earliest mummies of the southern Tarim Basin grouped closely with the modern Sardinian and Basque samples without evidence for any mtDNA contribution from the east (Cui 2002, 85–6). Second, the mismatch distribution of ancient Luobunoer sequences was bimodal, suggesting the Luobunoer 'population' had undergone two periods of rapid demographic expansion followed by rapid decline. In contrast, the Iron Age samples (Turfan) displayed a more gradual expansion and greater pairwise differences characteristic of modern Central Asian populations (ibid., 82). In addition, the later Iron Age samples from the northern Tarim Basin were extremely similar to the modern Uighur and Kazak mtDNA samples analysed by Comas *et al*. (1998). These data lend further support to the idea of long-term biological continuity in Central Asian communities.

Although these results could be of importance for understanding the history of the Tarim Basin, there is still much that remains unresolved. For example, Cui's (2002) samples were taken from only two small areas of a large, culturally diverse region, and are not necessarily indicative of the wider pattern of genetic variation in this area. In addition, her aDNA sample set is quite small relative to those typically analysed for modern populations, thereby making it difficult to compare the two kinds of data. A further problem is that some of her sequences exhibit combinations of mutations (motifs) that are known to define West Eurasian or East Eurasian haplotypes (e.g. C, T, or U), while others have motifs that are novel and atypical for these same haplogroups. These atypical sequences often occur because of damaged nucleotides in the DNA sample or due to the misalignment of sequence data. Despite these technical limitations, it is fair to say that Cui's (2002) work gives us additional evidence with which to evaluate models of human settlement in the Tarim Basin.

Obviously, more ancient DNA samples from Tarim Basin individuals need to be extracted and analysed before any conclusive arguments can be posited on the genetic demography of the prehistoric cultures of this region. Until such data become available, we must use a more indirect method for reconstructing the phylogenetic history of eastern Central Asia: the molecular genetic analysis of modern populations. An assertion of this nature generally meets with immediate resistance from historians, who point out that the modern people of the Tarim Basin (the Uighurs) were one of a number of Altaic-speaking tribes who arrived in the early first millennium AD and rose to dominance *c*.AD 800 (Gladney 1998, 817; Mallory and Mair 2000, 100–1). This date corresponds with the end of the use of Tocharian as a written language (Mallory and Mair 2000, 280–1). Such a coincidence may explain Mair's (1998b) emphasis on analysing ancient DNA. He may believe that the Turkic tribes 'erased' any earlier genetic signals, just as they 'erased' the Tocharian language.

There are two arguments to support the use of genetic data from modern Uighurs to research the prehistoric past of the Xinjiang Province. First, as argued above for Hungary and Turkey, the replacement of a language does not necessitate a dramatic shift in the regional gene pool, especially when the language was introduced by a nomadic group. As seen with Altaic-speakers of central Siberia (Pakendorf *et al*. 2003), the nomads' contribution to the local gene pool may be rendered insignificant by the populous sedentary societies that they conquered. Second, as demonstrated by archaeological, craniometric, and ancient genetic data, the history of the Tarim Basin involved a great deal of cultural and biological continuity. While the rise of the Silk Road and the nomadic incursions of the early first millennium AD undoubtedly had an impact upon the gene pool, these historical events may be detectable in the genetic composition of modern Uighurs. Thus, although this method cannot determine which groups spoke Indo-European languages, it can delineate possible suspects.

A similar effort to reconstruct the history of the Tarim Basin can be made with classic genetic markers (e.g. blood polymorphisms), which provide considerable information about biparentally inherited genetic variation (Renfrew 1992, 460). In an important paper, Tongmao Zhao (1998) summarized the past two decades of research by Chinese scholars into immunoglobulin (antibody) allotypes and blood group haplotypes of Xinjiang populations. This work has demonstrated that the Uighur possess genetic markers specific both to western Eurasian populations (e.g. immunoglobulin haplotype $Gm^{f;b}$) and to eastern Eurasian populations (e.g. the 'Diego' blood group antigen Di^{a}). These results have been confirmed recently by Iwasaki *et al.* (2000), who demonstrated a similar mix between eastern and western Eurasian ABO blood group allele frequencies, and by Xiao *et al.* (2002) using eight Alu insertion polymorphisms.

Although analyses of classical genetic markers can produce a useful summary of a population's biological ancestry, they cannot provide the precision needed for reconstructing its migratory history. For this purpose, mtDNA and NRY provide a high-resolution method for assessing female and male ancestry, respectively, based on high frequencies of certain diagnostic haplogroups in specific geographic regions or cultural groups. While there are various terminologies in the literature for these haplogroups (e.g. Zerjal *et al.* 2002, 469), in this paper, we will use the mtDNA haplogroup nomenclature of Richards *et al.* (2000) and the NRY haplogroup nomenclature of Underhill *et al.* (2000).[8]

The earliest high-resolution genetic analyses involving Uighur populations stemmed from research into the mtDNA diversity of Anatolian populations, where scientists continue to argue over the significance of genetic input from the Turkish incursion of the eleventh century AD (see Comas *et al.* 1996 and Di Benedetto *et al.* 2001 for contrasting opinions). Because the Uighurs and the Turks probably originated in the same general area of Central Asia (given their linguistic similarities), it was assumed that they share a common genetic ancestry. The results presented in Comas *et al.* (1998) support this hypothesis to some extent by demonstrating that the Turkic-speaking peoples of Central Asia (i.e. Uighurs, Kazakhs, and Kirgiz) cluster (via

8 A revised nomenclature has recently been proposed for NRY haplotypes (Y-Chromosome Consortium 2001). Although somewhat cumbersome, this is being used in lieu of previous naming systems. Using the revised nomenclature, we obtain the following haplogroup designations:

SNP Marker Designation	Underhill *et al.* (2000, 2001)	Y-Chromosome Consortium
M1	III + IV	DE*
M1 + M96	III	E
M1 + M174	IV	D
M45	VII	P*
M17	within IX	R1a1
M45	VII	P*
M89	VI	F*
M119	within VII	O1*
M122	within VII	O3*
M130	V	C
M173	IX	R1
M175	VII	O

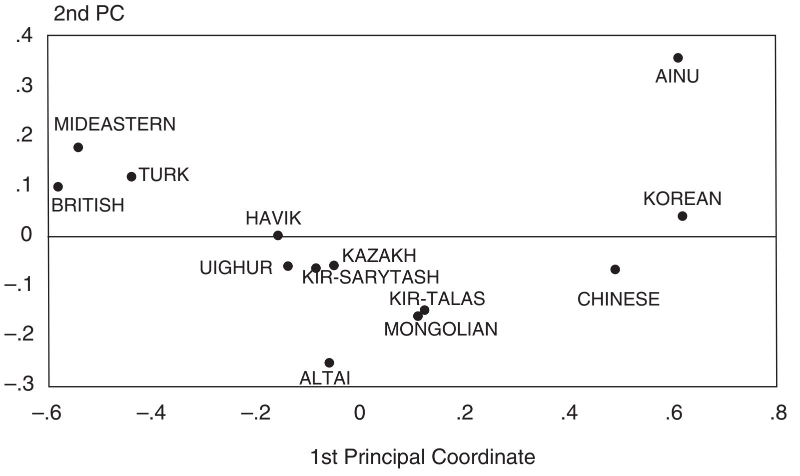

Figure 2
Principal-coordinate analysis of a distance matrix generated from mtDNA sequences among Asian and European populations. The first principal coordinate explains 51.9% of the variation, and the second explains 21.1%, for a combined percentage of 73.0% (modified from Comas *et al*., Am. J. Hum. Genet. 63, 1824–1838. © 1998 by The American Society of Human Genetics. All rights reserved).

MDS and NJCA) closer to Anatolian populations than to Chinese or other East Asian populations (see also Owens and King 1999). However, they also noted that 54.5 per cent of Uighur mtDNA sequences are characteristic of East Asian populations, whereas only 34.5 per cent were distinctly western (Comas *et al*. 1998, 1832) (Fig. 2).

This latter conclusion is supported by the mtDNA work of Yao *et al*. (2000), who noted a shorter genetic distance between Uighurs and Southeast/East Asians than between the former and modern Turks. Indeed, in a comparison of mtDNA sequences from southern European, Anatolian, and Turkic-speaking Central Asian populations (Di Benedetto *et al*. 2001, 152), Central Asians demonstrated a relatively high frequency of characteristically East Asian (i.e. A, C, D, and E) and Southeast Asian (i.e. B, F, and M) haplogroups, and exhibited characteristically European haplogroups (e.g. H, J, T, T1, and U) at low frequencies (see also Schurr and Wallace 2002). At the same time, while showing the differential contribution of eastern and western Eurasian populations, this data set is somewhat imprecise in assigning the geographic origins of these haplogroups. In this regard, it would be very informative to identify the specific subclusters within the haplogroups that are present in the Uighurs and Central Asian populations, as they can be geographically provenienced at a much finer scale than just the haplogroups themselves (e.g. haplogroup U in Kivisild *et al*. 1999).

In an attempt to clarify some of the questions arising out of these mtDNA studies, a number of researchers have begun to accumulate significant amounts of data on the NRY haplotype composition of Central Asian populations. For example, Perez-Lezaun *et al*. (1999), who analysed eight STR polymorphisms in the same samples used by Comas *et al*. (1998), demonstrated less genetic similarity between the Uighurs and both Kazaks and Kirgiz than was observed with mtDNA. However, when similar analyses targeting a greater number of NRY polymorphic sites (both biallelic markers and microsatellites) were performed, the Uighur consistently clustered (via MDS and NJCA) with the Uzbeks and other Turkic-speaking peoples

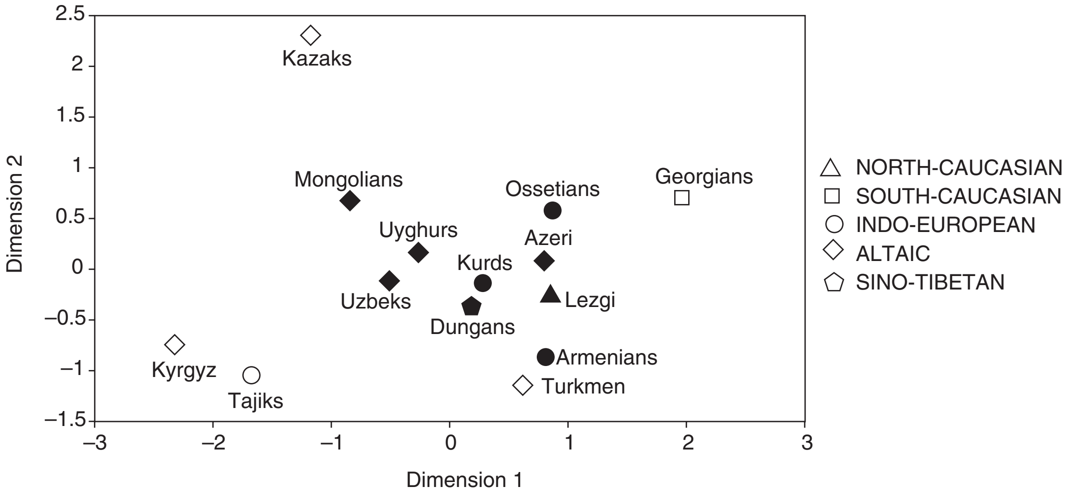

Figure 3
Multidimensional scaling analysis of population pairwise values of FST based on NRY microsatellite haplotypes. Symbol shapes indicate language affiliation, blackened symbols represent high-diversity populations, and unblackened symbols represent low-diversity populations (from Zerjal *et al*., Am. J. Hum. Genet. 71, 466–482. © 2002 by The American Society of Human Genetics. All rights reserved).

of Central Asia (Wells *et al*. 2001; Karafet *et al*. 2001; Zerjal *et al*. 2002). Furthermore, the Kazaks were found to be genetically distinct from the neighbouring Uighurs, although the former had a relatively low intra-popular diversity, suggesting a past bottleneck or founder effect (Zerjal *et al*. 2002, 473) (Fig. 3).

Several NRY haplogroups appear at significant frequencies in the Uighur (Wells *et al*. 2001; Zerjal *et al*. 2002). These include M130, which is prevalent in Central Asia, but absent in Europe and the Middle East; M89, which is most prevalent in the Middle East and sometimes considered diagnostic of the Neolithic spread throughout Eurasia; M9, which is most prevalent in Southeast Asia, but common to Eurasia; and M17, which is common to Eurasia, and sometimes considered diagnostic for the spread of horse domestication from the Pontic Steppe. Two haplogroups conspicuously absent from the Uighur population are M1 or 'YAP+' (and its derivatives M96 and M174), which is present at low frequencies in populations across the entire Eurasian supercontinent, and M173, which is common in western Eurasia but almost entirely absent from East and South Asian populations.

The data from both Wells *et al*. (2001) and Zerjal *et al*. (2002) are consistent on two points concerning the NRY haplotype composition of the Uighur population. First, the genetic contribution of Han Chinese males appears to be fairly limited, based on the absence of M119 and M122 haplotypes, which are common in Han populations (Su *et al*. 1999). These data contrast with the mtDNA data discussed above, which show more genetic input from the east than from the west. Such a pattern suggests the presence of a socio-cultural barrier preventing unions between East Asian men and Tarim Basin women but not the reverse. This may, in turn, explain the prevalence of purportedly female 'Mongoloid' crania in the sample analysed by Han (1998). Second, the Uighur are closer to some Middle Eastern populations (e.g. the Kurds) than they are to many other Turkic-speaking populations of Central Asia, a result that Zerjal *et al*.

(2002, 478) attributes to gene flow along the Silk Road.[9] However, this association may be earlier than the Silk Road, and could instead suggest a Middle Eastern (e.g. 'Bactrian') origin for some of the prehistoric inhabitants of the Tarim Basin. Attempts to assign genetic subclusters to the larger haplogroups, such as U2i ('Indian') vs. U2e ('European') (Quintana-Murci *et al*. 2002), may help to clarify this issue.

DISCUSSION

In this paper, the archaeological, biological, and linguistic evidence pertaining to the history of the Tarim Basin cultures of the Bronze and Iron Ages has been summarized and compared with similar information from modern populations of the region. The archaeological evidence suggests the occurrence of three discernible prehistoric events. First, the initial settling of the oases by peoples from the west in the early second millennium BC. Second, the transition to semi-nomadic pastoralism, horse-riding, and the manufacture of colourful textiles (including plaid-twill) and iron in the late second/early first millennium BC. Third, the arrival of the Han Empire from the east and the beginnings of the Silk Road in the late first millennium BC. Despite these major cultural transformations, archaeologists support both the synchronic cultural heterogeneity of the Tarim Basin (i.e. it should not be treated as one society) and its diachronic cultural homogeneity (i.e. it should not be perceived as a series of popular replacements).

The physical anthropological evidence put forward by Han suggests the initial colonization from both the west (at Gumugou) and from the east (at Yanbulake), although the former is probably earlier. In the Iron Age, Han (1998) observes a decline in 'Mongoloid' and the initial 'Western' type, and an influx of new 'Western' types from the highlands of the west and south. Hemphill, on the other hand, believes that the initial colonizers of the Tarim Basin were from an as-of-yet undefined origin (perhaps more western Eurasian than eastern Eurasian). The Iron Age crania that he examined cluster exclusively with southern Central Asian samples, mirroring the directionality of the archaeological changes of the first millennium BC. More importantly, Hemphill's work demonstrates a high degree of biological continuity within discrete geographic locations over time, for reasons that remain open to debate.

The historical linguistic evidence, which formulates the questions most are attempting to answer, is unfortunately the most tenuous. It is probably safe to assume that Tocharian was a written language used in the northern Tarim Basin from at least the mid-first millennium BC until the late first millennium AD. After this time, it appears to have been replaced by Turkic and Indo-Iranian languages. Inter-linguistic comparison demonstrates little interaction between Old Chinese and Tocharian, although increasing word exchange with Indo-Iranian languages is documented during the time of the Silk Road. Furthermore, Tocharian itself shows little relation to its geographic neighbours. Indeed, its closest linguistic relatives are probably the Celtic, Italic, and Anatolian languages.

The new data of Cui (2002) from the Tarim Basin mummies, in combination with the philological arguments concerning Tocharian's linguistic relatives, tempt us to assign the Tocharian arrival to one of the two suggested 'European' immigrations. However, it must also be stated that there is no substantive argument for why Tocharian could not have arrived in the Early Iron Age from southern Central Asia, or even in some unremarkable time period not

9 This may also explain the presence of identical haplogroups (although at different frequencies) observed in both Uighurs and South Asian populations (Wells *et al*. 2001, 10246).

distinguished by significant cultural or biological changes. Until such evidence is brought forth, the eastern frontier of the Indo-European question will remain forever unanswered and, perhaps, unanswerable.

Genetic analysis of the modern population of the Tarim Basin merely confirms what one would expect from a Central Asian population: an intermediate position between East and West (i.e. 'Central *Eur*asian'). However, the sex-linked genetic evidence reveals an interesting contrast between the mtDNA, which shows greater input from the east, and NRY, which shows greater input from the west. The latter of these genetic markers has been the most researched in recent years, and, while the results are far from conclusive, they do provide possibilities for the origins and development of Xinjiang populations. The fact that such an assessment runs counter to the recent data presented in Cui (2002) may simply demonstrate the problems associated with her small sample sizes and technical analysis of DNA sequence data. On the other hand, Cui's work seems to suggest a fairly distinct demographic break between the Bronze Age and Iron Age populations that she studied, which, in turn, may imply a genetic bottleneck followed by significant genetic influx in the Early Iron Age period. Only further studies of mtDNA and NRY variation in both ancient and modern samples and the contextualization of these data with archaeological, biological, and linguistic evidence can help to delineate the biocultural history of this complex and fascinating region.

EPILOGUE

The role of a synthetic paper such as this one is not to make conclusions where none are possible, but to summarize the available data in order to provide new directions for future studies. In a field such as molecular anthropology, the rapid pace of research and publication makes this task all the more necessary. Geneticists are often misinformed on the ambiguities of archaeological and linguistic interpretation, and begin inadvertently to produce data for data's sake. Similarly, archaeologists and historical linguists often lack even a basic understanding of genetics, both its strengths and its limitations, and, therefore, look for answers in the wrong places. In this paper, we have attempted to bridge these conceptual gaps by demonstrating how genetic studies of modern populations can inform and be informed by anthropological, and linguistic research.

Besides the specific suggestions made above for how to improve our understanding of the history of the Tarim Basin (i.e. scientific excavation of actual settlements, identification of diagnostic mutations within haplogroups, etc.), there are three general areas of research that still require considerable scrutiny before an informed synthesis can be achieved. First, and perhaps most obviously, more DNA data are needed from C^{14}-dated, scientifically excavated, and culturally provenienced human remains (whether 'mummified' or not) across a wide geographical and chronological expanse. Although synchronic comparisons of ancient DNA from Tarim Basin and other Eurasian populations would undoubtedly be informative, a diachronic approach targeting the relationship between ancient DNA and modern DNA diversity within a distinct geographical region could elucidate larger questions about the levels of biological continuity, mixing, and/or replacement in relation to cultural and/or linguistic stasis or change. An example of this approach can be seen with Holocene populations of the Cis-Baikal region (Mooder *et al*. 2003; Schurr 2003).

Similarly, ancient DNA research may provide the key to understanding the patterns observed by physical anthropologists in ancient crania. In this regard, we recognize that patterns of craniometric variation may reflect the influences of cultural, biological, and/or environmental factors, and do not represent evidence supporting racial typologies of human groups. At the same time, we have based our discussion of Han and Hemphill's craniofacial data on the assumption that there are differences between eastern and western Eurasian crania, and that these data reflect the regional divergence of local gene pools in those general areas over many millennia. Thus, by extracting ancient DNA from the specimens examined by Han and Hemphill, we would be able to begin scientifically exploring the conception of geographically defined biological populations held by many physical anthropologists, forensic scientists, and medical practitioners (Cartmill 1999, 652). Furthermore, a comparison between ancient DNA data and cranial markers within small, distinct populations (e.g. a cemetery) could reveal the genealogical relationships among the ancient inhabitants.

A third area of research that we have highlighted in this paper is molecular anthropology, a field that serves as both a producer and a consumer of interpretations of the past. The relationship between archaeological and/or historical linguistic data and modern genetics is undeniably indirect, thus requiring a great deal of caution. However, the fact remains that there *is* a connection between the present-day biological diversity of a region and the cultures and languages of its past, and it is our role as anthropologists to determine the extent of this association. To accomplish this goal, a greater emphasis is needed on teaching multi-disciplinary approaches to the past (e.g. Day 2001). This will allow the next generation of scholars to traverse scholastic lines more easily and to formulate a more holistic understanding of the Tarim Basin and beyond.

In addition to championing a multi-disciplinary approach to the prehistory of eastern Central Asia, we have re-examined the claims of Indo-European origins of Tarim Basin peoples. Our comparative work has generally shown that the narrative espoused by Mallory and Mair (2000), Barber (1999), Anthony (1995) and others, i.e. that the famous mummies are the progeny of 'proto-Celtic' 'Europeans' from the Pontic Steppe who migrated thousands of kilometres across two vast mountain ranges and the entire Eurasian Steppe just to settle on the outskirts of one of the most inhospitable deserts in the world, is not the only interpretation for these data. As Stefan Zimmer (1998, 178) writes, 'Indo-European Cultural Studies have to be kept under strict methodological control . . . in order to avoid gross misunderstandings.' This is especially true when introducing molecular studies to the equation, as it requires a 'self-critical awareness of the world both of and beyond genetics' (Pluciennik 1996, 14), particularly in a politically sensitive area such as Xinjiang Province (cf. Light 1999).

In general, a greater reflexivity among scholars of the Tarim Basin is needed, particularly in relation to the actual vocabulary being used to describe archaeological cultures and human remains. It is our opinion that ill-defined words such as 'European' or 'Celtic' are problematic in academic discussions of populations in eastern Central Asia/western China who were, and still are, constantly renegotiating their cultural identity (see Chvyr' 1994; Gladney 1996; Rudelson 1997). A more critical self-awareness of the actual lexicon that we are using in our research will help us to avoid 'gross misunderstandings' such as those evinced in politically biased works such as Deavin (1997). For the modern-day inhabitants of this region, the relationship between genes, language, and culture in the past, as well as in the present, is of far greater consequence than perhaps we are willing to admit.

Acknowledgements

A number of scholars have contributed useful comments and, at times, important criticisms that have helped us to come to terms with this complex but fascinating subject. We wish to thank Michael Frachetti, Brian Hemphill, Fred Hiebert, Karl Lamberg-Karlovsky, Nathan Light, Victor Mair, Jianjun Mei, and Colin Renfrew for all their help. Needless to say, the opinions expressed in this paper should not be taken to reflect the views of anyone but the authors themselves.

Department of Anthropology
University of Pennsylvania
325 University Museum
3260 South Street
Philadelphia, PA 19104-6398
cpt2@sas.upenn.edu
tgschurr@sas.upenn.edu

REFERENCES

ANTHONY, D.W. 1990: Migration in archaeology: the baby and the bathwater. *American Anthropologist* 92(4), 895–914.

ANTHONY, D.W. 1995: Horse, wagon, & chariot: Indo-European languages and archaeology. *Antiquity* 69, 554–65.

ANTHONY, D.W. 1998: The opening of the Eurasian steppe at 2000 BCE. In Mair, V.H. (ed.), 1998a, 94–113.

ANTHONY, D.W. 2001: Tracking the Tarim Mummies. *Archaeology* 54(2), 76–84.

BARBER, E.W. 1998: Bronze Age cloth and clothing of the Tarim Basin: the Loulan (Loulan) and Qumul (Hami) evidence. In Mair, V.H. (ed.), 1998a, 647–55.

BARBER, E.W. 1999: *The Mummies of Ürümchi* (New York).

BOAS, F. 1911: *The Mind of Primitive Man* (New York).

CARTMILL, M. 1999: The status of the race concept in physical anthropology. *American Anthropologist* 100(3), 651–60.

CAVALLI-SFORZA, L.L. 2000: *Genes, Peoples, and Languages* (New York).

CAVALLI-SFORZA, L.L., MENOZZI, P. and PIAZZA, A. 1994: *The History and Geography of Human Genes* (Princeton, NJ).

CHANG, C., TOURTELLOTTE, P., BAIPAKOV, K.M. and GRIGORIEV, F.P. 2002: *The Evolution of Steppe Communities from the Bronze Age through Medieval Periods in Southeastern Kazakstan (Zhetysu)* (Sweet Briar–Almaty).

CHEN, K.T. and HIEBERT, F.T. 1995: The late prehistory of Xinjiang in relation to its neighbors. *Journal of World Prehistory* 9(2), 243–300.

CHVYR', L.A. 1994 (1995–1996): Notes on the ethnic self-awareness of the Uighur. *Anthropology and Archeology of Eurasia* 34(3), 48–67.

COMAS, D., CALAFELL, F., MATEU, E., PEREZ-LEZAUN, A. and BERTRANPETIT, J. 1996: Geographic variation in human mitochondrial DNA control region sequence: the population history of Turkey and its relationship to the European populations. *Molecular Biology and Evolution* 13(8), 1067–77.

COMAS, D., CALAFELL, F., MATEU, E., PEREZ-LEZAUN, A., BOSCH, E., MARTINEZ-ARIAS, R., CLARIMON, J., FACCHINI, F., FIORI, G., LUISELLI, D., PETTENER, D. and BERTRANPETIT, J. 1998: Trading genes along the Silk Road: mtDNA sequences and the origin of Central Asian populations. *American Journal of Human Genetics* 63, 1824–38.

CUI, Y. 2002: *Ancient DNA Analysis of Human Population in Xinjiang, China* (Ph.D. Dissertation, Jilin University).

DAVIS-KIMBALL, J., BASHILOV, V.A. and YABLONSKY, L.T. (eds.) 1995: *Nomads of the Eurasian steppes in the early Iron Age* (Berkeley, CA).

DAY, J.V. 2001: *Indo-European Origins: The Anthropological Evidence* (Washington DC, The Institute for the Study of Man).

DEAVIN, M. 1997: Aryans: culture bearers to China – New evidence of ancient European migration to the Orient. *National Vanguard Magazine* 117.

DEBAINE-FRANCFORT, C. 1988: Archeologie du Xinjiang des origines aux Han: premiere partie. *Paleorient* 14(1), 5–29.

DEBAINE-FRANCFORT, C. 1995: *Du Néolithique à L'âge du Bronze en Chine du Nord-ouest: la culture de Qijia et ses connexions* (Paris, Mémoires de la Mission archéologique française en Asie centrale, t.6, Editions Recherche sur les civilizations).

DI BENEDETTO, G., ERGUVEN, A., STENICO, M., CASTRI, L., BERTORELLE, G., TOGAN, I. and BARBUJANI, G. 2001: DNA diversity and population admixture in Anatolia. *American Journal of Physical Anthropology* 115, 144–56.

DI COSMO, N. 1996: Ancient Xinjiang between Central Asia and China: The nomadic factor. *Anthropology and Archeology of Eurasia* 34(4), 87–101.

FINNILÄ, S., LEHTONEN, M.S. and MAJAMAA, K. 2001: Phylogenetic Network for European mtDNA. *American Journal of Human Genetics* 68, 1475–84.

FRACHETTI, M. 2002: Bronze Age exploitation and political dynamics of the Eastern Eurasian Steppe zone. In Boyle, K., Renfrew, C. and Levine, M. (eds.), *Ancient Interactions: East and West in Eurasia* (Cambridge, McDonald Institute for Archaeological Research), 161–70.

FRANCALACCI, P. 1998: DNA analysis on ancient desiccated corpses from Xinjiang (China): further results. In Mair, V.H. (ed.), 1998a, 537–47.

GLADNEY, D.C. 1996: Relational alterity: constructing Dungan (Hui), Uygur, and Kazakh identities across China, Central Asia, and Turkey. *History and Anthropology* 9(4), 445–77.

GLADNEY, D.C. 1998: Ethnogenesis and ethnic identity in China: considering the Uyghurs and Kazaks. In Mair, V.H. (ed.), 1998a, 812–34.

GOOD, I. 1998: Bronze Age cloth and clothing of the Tarim Basin: the Chärchän evidence. In Mair, V.H. (ed.), 1998a, 656–70.

HADINGHAM, E. 1994: The mummies of Xinjiang. *Discover* 15(4), 68–77.

HAMP, E.P. 1998: Whose were the Tocharians? Linguistic subgrouping and diagnostic idiosyncrasy. In Mair, V.H. (ed.), 1998a, 307–46.

HAN, K. 1998: The physical anthropology of the ancient populations of the Tarim Basin and surrounding areas. In Mair, V.H. (ed.), 1998a, 558–72.

HEMPHILL, B.E. 1999: Biological affinities and adaptations of Bronze Age Bactrians: IV. A craniometric investigation of Bactrian origins. *American Journal of Physical Anthropology* 108, 173–92.

HEMPHILL, B.E. and MALLORY, J.P. in press: The Bronze Age settlement of Xinjiang: A craniometric investigation (accepted to *American Journal of Physical Anthropology*).

HENNING, W.B. 1978: The first Indo-Europeans in history. In Ulmen, G.L. (ed.), *Society and History: Essays in Honor of Karl August Wittfogel* (The Hague), 215–32.

HERRNSTADT, C., ELSON, J.L., FAHY, E., PRESTON, G., TURNBULL, D.M., ANDERSON, C., GHOSH, S.S., OLEFSKY, J.M., BEAL, M.F., DAVIS, R.E. and HOWELL, N. 2002: Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *American Journal of Human Genetics* 70(5), 1152–71.

IWASAKI, M., KOBAYASHI, K., SUZUKI, H., ANAN, K., OHNO, S., GENG, Z., LI, G. and INOKO, H. 2000: Polymorphism of the ABO blood group genes in Han, Kazak and Uyghur populations in the Silk Route of northwestern China. *Tissue Antigens* 56, 136–42.

KARAFET, T., XU, L., DU, R., WANG, W., FENG, S., WELLS, R.S., REDD, A.J., ZEGURA, S.L. and HAMMER, M.F. 2001: Paternal population history of East Asia: sources, patterns, and microevolutionary processes. *American Journal of Human Genetics* 69, 615–28.

KIVISILD, T., BAMSHAD, M.J., KALDMA, K., METSPALU, M., METSPALU, E., REIDLA, M., LAOS, S., PARIK, J., WATKINS, W.S., DIXON, M.E., PAPIHA, S.S., MASTANA, S.S., MIR, M.R., FERAK, V. and VILLEMS, R. 1999: Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. *Current Biology* 9(22), 1331–4.

KULIKOV, E.E. and POLTARAUS, A.B. 2002: Molecular analysis of mt-DNA of Bronze Age samples from Karghaly region. Paper presented at the *Earliest Stages of Mining and Metallurgy in Northern Part of Eurasia: Kargaly Complex* conference in Kargaly–Orenburg (Moscow, Archaeological Institute of the Russian Academy of Science).

KUZMINA, E.E. 1998: Cultural connections of the Tarim Basin people and pastoralists of the Asian steppes in the Bronze Age. In Mair, V.H. (ed.), 1998a, 63–93.

LAMBERG-KARLOVSKY, C.C. 2002: Archaeology and language: the Indo-Iranians. *Current Anthropology* 43(1), 63–88.

LI, S. 2002: The interaction between Northwest China and Central Asia during the second millennium B.C.: an archaeological perspective. In Boyle, K., Renfrew, C. and Levine, M. (eds.), *Ancient Interactions: East and West in Eurasia* (Cambridge, McDonald Institute for Archaeological Research), 171–82.

LIGHT, N. 1999: Hidden discourses of race: imagining Europeans in China. Based on 'A fly crushed between stallions: Uyghur history on the margins of empire' presented at the Association of Asian Studies annual meeting in Boston, March 1999 (URL: http://www.utoledo.edu/~nlight/uyghhst.htm).

LIN, M. 1998: Qilian and Kunlun – the earliest Tokharian loan-words in Ancient Chinese. In Mair, V.H. (ed.), 1998a, 476–82.

LUBOTSKY, A. 1998: Tocharian loan words in Old Chinese: chariots, chariot gear, and town building. In Mair, V.H. (ed.), 1998a, 379–90.

MACBETH, H. 1997: What is an ethnic group? A biological perspective. In Clarke, A. and Parsons, E. (eds.), *Culture, Kinship and Genes* (New York), 54–66.

MAIR, V.H. 1995a: Mummies of the Tarim Basin. *Archaeology* 48(2), 28–35.

MAIR, V.H. 1995b: Prehistoric Caucasoid corpses of the Tarim Basin. *Journal of Indo-European Studies* 23(3–4), 281–308.

MAIR, V.H. (ed.) 1995c: *Journal of Indo-European Studies* 23(3–4).

MAIR, V.H. (ed.) 1998a: *The Bronze Age and Early Iron Age Peoples of Eastern Central Asia* (Washington D.C., Institute for the Study of Man).

MAIR, V.H. 1998b: Priorities. In Mair, V.H. (ed.), 1998a, 4–44.

MALLORY, J.P. 1989: *In Search of Indo-Europeans* (London).

MALLORY, J.P. and MAIR, V.H. 2000: *The Tarim Mummies* (London).

MEI, J. 2000: *Copper and Bronze Metallurgy in Late Prehistoric Xinjiang* (Oxford, BAR Int. Ser. 865).

MEI, J. and SHELL, C. 2002: The Iron Age cultures in Xinjiang and their Steppe connections. In Boyle, K., Renfrew, C. and Levine, M. (eds.), *Ancient Interactions: East and West in Eurasia* (Cambridge, McDonald Institute for Archaeological Research), 213–34.

MIRZA, M.N. and DUNGWORTH, D.B. 1995: The potential misuse of genetic analyses and the social construction of 'race' and 'ethnicity'. *Oxford Journal of Archaeology* 14(3), 345–54.

MOODER, K.A., SCHURR, T.G. and BAMFORTH, F. 2003: Mitochondrial DNA and archeology: The genetic characterisation of prehistoric Siberian hunter-gatherers. In McKenzie, H. and Weber, A. (eds.), *Middle Holocene Hunter-Gatherers of the Baikal Boreal Forest: Current Problems and Future Directions* (Edmonton, Canadian Circumpolar Institute).

MU, S. 1992: Comments on the painted pottery in Xinjiang. In Jarrige, C. (ed.), *South Asian Archaeology 1989* (Madison, Wisconsin), 201–12.

OPIE, J. 1995: Xinjiang remains and 'the Tocharian Problem'. *Journal of Indo-European Studies* 23(3–4), 431–8.

OWENS, K. and KING, M.-C. 1999: Genomic views of human history. *Science* 286, 451–3.

PAKENDORF, B., WIEBE, V., TARSKAIA, L.A., SPITSYN, V.A., SOODYALL, H., RODEWALD, A. and STONEKING, M. 2003: Mitochondrial DNA evidence for admixed origins of Central Siberian populations. *American Journal of Physical Anthropology* 120, 211–24.

PARPOLA, A. 1998: Aryan languages, archaeological cultures, and Sinkiang: where did Proto-Iranian come into being and how did it spread? In Mair, V.H. (ed.), 1998a, 114–47.

PEREZ-LEZAUN, A., CALAFELL, F., COMAS, D., MATEU, E., BOSCH, E., MARTINEZ-ARIAS, R., CLARIMON, J., FIORI, G., LUISELLI, D., FACCHINI, F., PETTENER, D. and BERTRANPETIT, J. 1999: Sex-specific migration patterns in Central Asian populations, revealed by analysis of Y-chromosome short tandem repeats and mtDNA. *American Journal of Human Genetics* 65, 208–19.

PLUCIENNIK, M. 1996: Genetics, archaeology and the wider world. *Antiquity* 70, 13–14.

POLLARD, A.M. 2001: Overview – archaeological science in the biomolecular century. In Brothwell, D.R. and Pollard, A.M. (eds.), *Handbook of Archaeological Sciences* (Chichester), 295–9.

QUINTANA-MURCI, L., WELLS, S., CHAIX, R., SAYAR, H., OTTOLENGHI, C., AYUB, Q., TYLER-SMITH, C., MEHDI, S.Q., BANDELT, H.J. and MCELREAVEY, K. 2002: *Mitochondrial DNA phylogeography and population history of West, South and Central Asia* (HGM2002 Poster Abstracts, Poster 573: http://hgm2002.hgu.mrc.ac.uk/Abstracts/Publish/WorkshopPosters/WorkshopPoster11/hgm0573.htm).

RENFREW, C. 1987: *Archaeology and Language* (Cambridge).

RENFREW, C. 1992: Archaeology, genetics and linguistic diversity. *Man* 27(3), 445–78.

RENFREW, C. 1998: The Tarim Basin, Tocharian, and Indo-European Origins: A view from the West. In Mair, V.H. (ed.), 1998a, 202–14.

RICHARDS, M., MACAULAY, V., HICKEY, E., VEGA, E., SYKES, B., GUIDA, V., RENGO, C., SELLITTO, D., CRUCIANI, F., KIVISILD, T., VILLEMS, R., THOMAS, M.G., RYCHKOV, S., RYCHKOV, O., RYCHKOV, Y., GOLGE, M., DIMITROV, D., HILL, E., BRADLEY, D., ROMANO, V., CALI, F., VONA, G., DEMAINE, A., PAPIHA, S.S., TRIANTAPHYLLIDIS, C., STEFANESCU, G., HATINA, J., BELLEDI, M., DI RIENZO, A., NOVELLETTO, A., OPPENHEIM, A., NORBY, S., AL-ZAHERI, N., SANTACHIARA-BENERECETTI, A.S., SCOZZARI, R., TORRONI, A. and BANDELT, H.J. 2000: Tracing European founder lineages in the Near Eastern mtDNA pool. *American Journal of Human Genetics* 67, 1251–76.

RINGE, D., WARNOW, T., TAYLOR, A. and MICHAILOV, A. 1998: Computational cladistics and the position of Tocharian. In Mair, V.H. (ed.), 1998a, 391–414.

RUDELSON, J.J. 1997: *Oasis Identities* (New York).

SCHURR, T.G. 2003: Molecular genetic diversity of Siberian populations: Implications for ancient DNA studies of archeological populations from the Cis-Baikal region. In McKenzie, H. and Weber, A. (eds.), *Middle Holocene Hunter-Gatherers of the Baikal Boreal Forest: Current Problems and Future Directions* (Edmonton, Canadian Circumpolar Institute).

SCHURR, T.G. and WALLACE, D.C. 2002: Mitochondrial DNA diversity in Southeast Asian populations. *Human Biology* 74(3), 431–52.

SCHURR, T.G., MAGGI, W., GROVEMAN, M.A., FOWLER, K. and WALLACE, D.C. Submitted: Mitochondrial DNA variation and the genetic origins of the Kalasha of northern Pakistan. *American Journal of Physical Anthropology.*

SEMINO, O., PASSARINO, G., QUINTANA-MURCI, L., LIU, A., BERES, J., CZIEZEL, A. and SANTACHIARA-BENERECETTI, A.S. 2000: MtDNA and Y chromosome polymorphisms in Hungary: inferences from the Paleolithic, Neolithic and Uralic influences on the modern Hungarian gene pool. *European Journal of Human Genetics* 8, 339–46.

SOKAL, R.R. 1988: Genetic, geographic, and linguistic distances in Europe. *Proceedings of the National Academy of Sciences USA* 85, 1722–6.

SOKAL, R.R., JACQUEZ, G.M., ODEN, N.L., DI GIOVANNI, D., FALSETTI, A.B., MCGEE, E. and THOMSON, B.A. 1993: Genetic relationships of European populations reflect their ethnohistorical affinities. *American Journal of Physical Anthropology* 91, 55–70.

SPARKS, C.S. and JANTZ, R.L. 2002: A Reassessment of Human Cranial Plasticity: Boas Revisited. *Proceedings of the National Academy of Sciences USA* 99(23), 14636–9.

SU, B., JIN, L., UNDERHILL, P., MARTINSON, J., SAHA, N., MCGARVEY, S.T., SHRIVER, M.D., CHU, J., OEFNER, P., CHAKRABORTY, R. and DEKA, R. 2000: Polynesian origins: Insights from the Y chromosome. *Proceedings of the National Academy of Sciences USA* 97(15), 8225–8.

SU, B., XIAO, J., UNDERHILL, P., DEKA, R., ZHANG, W., AKEY, J., HUANG, W., SHEN, D., LU, D., LUO, J., CHU, J., TAN, J., SHEN, P., DAVIS, R., CAVALLI-SFORZA, L., CHAKRABORTY, R., XIONG, M., DU, R., OEFNER, P., CHEN, Z. and JIN, L. 1999: Y chromosome evidence for a northward migration of modern humans into eastern Asia during the last Ice Age. *American Journal of Human Genetics* 65, 1718–24.

TAYLOR, G. 1921: The evolution and distribution of race, culture, and language. *Geographical Review* 11, 54–119.

UNDERHILL, P.A., SHEN, P., LIN, A.A., JIN, L., PASSARINO, G., YANG, W.H., KAUFFMAN, E., BONNE-TAMIR, B., BERTRANPETIT, J., FRANCALACCI, P., IBRAHIM, M., JENKINS, T., KIDD, J.R., MEHDI, S.Q., SEIELSTAD, M.T., WELLS, R.S., PIAZZA, A., DAVIS, R.W., FELDMAN, M.W., CAVALLI-SFORZA, L.L. and OEFNER, P.J. 2000: Y chromosome sequence variation and the history of human populations. *Nature Genetics* 26, 358–61.

UNDERHILL, P.A., PASSARINO, G., LIN, A.A., SHEN, P., MIRAZÓN LAHR, M., FOLEY, R.A., OEFNER, P.J. and CAVALLI-SFORZA, L.L. 2001: The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann. Hum. Genet.* 65, 43–62.

WANG, B. 1996: A preliminary analysis of the archaeological cultures of the Bronze Age in the region of Xinjiang. *Anthropology and Archeology of Eurasia* 34(4), 67–86.

WANG, B. 2001: *Xinjiang gu shi: Gu dai Xinjiang ju min ji qi wen hua (The ancient corpses of Xinjiang: the peoples of ancient Xinjiang and their culture)* (Urumqi).

WELLS, R.S., YULDASHEVA, N., RUZIBAKIEV, R., UNDERHILL, P.A., EVSEEVA, I., BLUE-SMITH, J., JIN, L., SU, B., PITCHAPPAN, R., SHANMUGALAKSHMI, S., BALAKRISHNAN, K., READ, M., PEARSON, N.M., ZERJAL, T., WEBSTER, M.T., ZHOLOSHVILI, I., JAMARJASHVILI, E., GAMBAROV, S., NIKBIN, B., DOSTIEV, A., AKNAZAROV, O., ZALLOUA, P., TSOY, I., KITAEV, M., MIRRAKHIMOV, M., CHARIEV, A. and BODMER, W.F. 2001: The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proceedings of the National Academy of Sciences USA* 98(18), 10244–9.

XIAO, F.-X., YANG, J.-F., CASSIMAN, J.-J. and DECORTE, R. 2002: Diversity at eight polymorphic Alu insertion loci in Chinese populations shows evidence for European admixture in a ethnic minority population from northwest China. *Human Biology* 74(4), 555–68.

YAO, Y.-G., LU, X.-M., LUO, H.-R., LI, W.-H. and ZHANG, Y.-P. 2000: Gene admixture in the Silk Road region of China: Evidence from mtDNA and melanocortin 1 receptor polymorphism. *Genes and Genetic Systems* 75, 173–8.

ZERJAL, T., WELLS, R.S., YULDASHEVA, N., RUZIBAKIEV, R. and TYLER-SMITH, C. 2002: A genetic landscape reshaped by recent events: Y-chromosomal insights into Central Asia. *American Journal of Human Genetics* 71, 466–82.

ZHAO, T. 1998: The Uyghurs, a Mongoloid-Caucasoid mixed population: genetic evidence and estimates of Caucasian admixture in the peoples living in northwest China. In Mair, V.H. (ed.), 1998a, 548–57.

ZIMMER, S. 1998: Modern necromancy, or how to make mummies speak. *Journal of Indo-European Studies* 26(1–2), 163–80.

ZUBROV, A.A. 1996: Problema termina ⟨raca⟩ i racovykh classifikatsij v sovremennoj fizicheskoj antropologii (Problem of the term 'race' and racial classification in modern physical anthropology). *Ethnograficheskoe Obozrenie* 1996(1), 15–23.