

News & views

Linguistics

Tracking the origin of Transeurasian languages

Peter Bellwood

A triangulation of linguistic, archaeological and genetic data suggests that the Transeurasian language family originated in a population of grain farmers in China around 9,000 years ago, and that agriculture underpinned its spread. **See p.616**

The Transeurasian language family is one of the most widely distributed language families of the modern world, and one school of thought says that it encompasses the Turkic, Mongolic, Tungusic, Koreanic and Japonic language groups (Tungusic languages are spoken in northeastern China and Siberia). On page 616, Robbeets *et al.*¹ present their use of an innovative approach to resolve the long-standing debate about the origins and spread of this language family.

Nowadays, about 80 Transeurasian languages are spoken in a region that extends across more than 8,000 kilometres of the Asian continent (Fig. 1). The speakers of these languages span territories from Istanbul in the west to Tokyo and Seoul in the east, and that extend north to a Russian republic called Sakha, which reaches the Arctic coastline of Siberia and is home to the reindeer-herding and Turkic-language-speaking Yakut people (also known as the Sakha). A distribution on this scale is remarkable, and is on a par with the dispersals of other extensive language families such as Indo-European and Sino-Tibetan. Historical examples of speakers of Transeurasian languages include the Mongols, who were ruled by Genghis Khan and Kublai Khan; the Ottoman Turks, who captured Constantinople (now Istanbul) in 1453; and the Manchu (a Tungusic people), who ruled China during the Qing Dynasty (1644–1912).

Transeurasian language speakers clearly travelled far and wide in ancient times, so we might ask what prompted these dispersals in the first place. Did the early Transeurasians live mainly as roving pastoralists who tended cattle, sheep and horses, like many historical Turkic, Tungusic and Mongolic peoples in central Asia? Or did they live as settled farmers relying on cereals, legumes and pig husbandry,

more like the early Koreans and Japanese, and, indeed, the neighbouring Chinese (who speak Sino-Tibetan languages that are unrelated to the Transeurasian language family)?

Robbeets and colleagues, a large group of multidisciplinary and multinational authors, did not focus on the Transeurasian languages alone, but also considered the ancient

genomes of some of the potential speakers of these languages, and assessed archaeological traces that these populations left behind. The authors sought to reveal where the ancestral Transeurasian speakers originated, to understand what kind of lifestyle they had, and to determine how and when these languages spread. Robbeets *et al.* did this by ‘triangulating’ relevant results from the three disciplines of linguistics, archaeology and genetics.

First, the authors examined the internal history and family tree of the Transeurasian language family. The concept of a language family assumes that the member languages have inherited a body of cognate (that is, non-borrowed) vocabulary from a common ancestral proto-language. To investigate the existence of an ancestral Proto-Transeurasian language, and to reconstruct the meanings of words for key cultural concepts within it, Robbeets and colleagues used a statistical approach to analyse the most comprehensive set of cognate Transeurasian basic and agropastoral vocabulary terms available. The results of this linguistic analysis led the authors to locate the Transeurasian linguistic



Figure 1 | The origin and spread of Transeurasian languages. Speakers of the Transeurasian language family spanned a wide geographical region, as indicated by the regions depicted, which are associated with current and ancient speakers of the listed language groups of this family. Robbeets *et al.*¹ present an analysis of linguistic, archaeological and genetic data that suggests that this language family originated in the Liao River Valley in China approximately 9,000 years ago, and then spread initially towards Korea and far-eastern Russia by 7,000 years ago. These data indicate that the earliest population to speak a Transeurasian language was a group of people who lived in the Liao River Valley, cultivated millet grain and farmed pigs. This community had a genetic heritage associated with ancient people from around the Amur River Valley (on the border between Russia and China). The research also suggests that this population of early Transeurasian-language speakers was distinct from the people who first spoke languages of the Sino-Tibetan language family (which includes Chinese languages) – individuals living near the Yellow River in China. (Image based on Fig. 1 of ref. 1.)

homeland in the Liao Valley and adjacent lowlands of northeastern China (Fig. 1) about 9,000 years ago, among a group of farmers who were among the first in the world to grow a type of grain called millet, and who also reared pigs and had knowledge of weaving².

The authors then compared the linguistically reconstructed lifestyle of the early Trans-eurasian speakers with the archaeological record of northeastern Asia, examining a database of 255 Neolithic, Bronze Age and Iron Age archaeological sites that date between 8,400 and 1,700 years ago. This analysis supports a Neolithic millet- and pig-farming community as corresponding to the primary ancestry of Transeurasian languages, located in the Liao Valley from around 8,500 years ago and onwards (as supported by radiocarbon dating of archaeological sites), in circumstances that enabled strong population growth. This viewpoint is consistent with an independent archaeological appraisal, published this year, of the chronology of early cultivation of broomcorn millet (*Panicum miliaceum*), indicating that this cultivation was under way by at least 4500 BC along the eastern branch of the Liao River³.

Finally, Robbeets *et al.* compared ancient genomes from many skeletal samples to identify the potential genetic ancestry of the Neolithic population of northeastern China. This evidence suggests that these early farmers of the Liao Valley had an Amur genomic ancestry (named after the Amur River, located on the border between China and Russia), a genetic heritage that was once widespread throughout northeastern China and the most easterly parts of Russia. This observation brings the genetic results of Robbeets *et al.* in line with the results of another analysis of ancient DNA from northeastern Asia, published earlier this year, that established an Amur genetic ancestry for modern speakers of Mongolic and Tungusic languages⁴. However, ancient DNA from the oldest Liao Valley farmers, who lived around 8,500 years ago, is still lacking, and this absence renders elusive any certainty about their origins.

If Robbeets and colleagues are correct, the implication must be that the early Trans-eurasians and Sino-Tibetans were in the first instance derived, respectively, from two genetically distinct and geographically separate Amur and Yellow river source populations, even though substantial genetic mixing between them occurred from later Neolithic times onwards. The search for the DNA of the first Neolithic farmers in East Asia remains a matter of urgency⁵.

The historical reconstruction favoured by Robbeets *et al.* aligns the origins of the Transeurasian language family with a widely supported hypothesis about farming and language dispersal that I have long favoured^{6–8}. This model offers a way to explain the initial

expansions of many of the world's key language families, including Sino-Tibetan⁹, as being related to the substantial population growth that occurred during the behavioural shifts from a hunter-gatherer lifestyle to food production. It is true that many later Trans-eurasian-language-speaking populations who moved into more-arid regions of central Asia, especially those of Turkic, Mongolic and Tungusic speakers, were primarily herders rather than cultivators. However, their dispersals were aided by extensive cultural interaction after 2300 BC with populations in western Eurasia that prompted the introduction of domesticated cattle, sheep and horses, and the adoption of dairy production¹⁰.

Robbeets and colleagues' work provides an important contribution from at least two perspectives. It provides computational linguistic support for the existence of a unified Transeurasian language family, a concept that has divided linguists, some of whom have long regarded Turkic, Mongolic and Tungusic as being members of an Altaic language family separate from Japonic and Koreanic languages¹¹. The authors' analysis also puts the focus on the initial development of agriculture in northeastern Asia as

being the underlying driver of the expansion of Transeurasian languages, rather than the later development of specialized pastoralism in central Asia. The study by Robbeets *et al.* stands as a welcome contribution to the debate over the origins of Transeurasian languages.

Peter Bellwood is at the School of Archaeology and Anthropology, The Australian National University, Canberra, Australian Capital Territory 2601, Australia. e-mail: peter.bellwood@anu.edu.au

1. Robbeets, M. *et al.* *Nature* **599**, 616–621 (2021).
2. Nelson, S., Zhushchikhovskaya, I., Li, T., Hudson, M. & Robbeets, M. *Evol. Hum. Sci.* **2**, e5 (2020).
3. Wang, J. *et al.* *Archaeol. Anthropol. Sci.* **13**, 156 (2021).
4. Wang, C.-C. *et al.* *Nature* **591**, 413–419 (2021).
5. Bellwood, P. *Science* **361**, 31–32 (2018).
6. Bellwood, P. *First Farmers* (Blackwell, 2005).
7. Bellwood, P. & Renfrew, C. (eds) *Examining the Farming/Language Dispersal Hypothesis* (McDonald Inst., 2002).
8. Bellwood, P. *Evol. Hum. Sci.* **2**, e37 (2020).
9. Zhang, M., Yan, S., Pan, W. & Jin, L. *Nature* **569**, 112–115 (2019).
10. Honeychurch, W. *et al.* *Archaeol. Res. Asia* **26**, 100264 (2021).
11. Vovin, A. in *The Global Prehistory of Human Migration* (ed. Bellwood, P.) 197–203 (Wiley, 2015).

The author declares no competing interests. This article was published online on 10 November 2021.

Neurodegeneration

Principles of Parkinson's disease disputed by model

Zak Doric & Ken Nakamura

Mice in which organelles called mitochondria are disrupted in vulnerable neuronal cells provide a new model of Parkinson's disease. The pattern of neurodegeneration challenges long-held ideas about the disease's motor symptoms. **See p.650**

Mitochondria – the cellular organelles crucial for energy production – have long been thought to be disrupted in the neurodegenerative disorder Parkinson's disease. Mitochondrial function relies on the activity of various enzyme complexes, and dysfunction of complex I in particular has been implicated in Parkinson's disease. González-Rodríguez *et al.*¹ report on page 650 that they genetically engineered mice to lack the gene encoding a complex I subunit called NDUFS2 in neurons that produce the neurotransmitter molecule dopamine. One of the brain regions that has dopamine neurons, the substantia nigra, contains the neurons that die first in individuals with Parkinson's disease. The mice lacking NDUFS2 show neurodegeneration, providing a model of Parkinson's disease (named the MCI-Park model) based on mitochondrial dysfunction.

Moreover, the authors' findings challenge the long-held idea that loss of dopamine release in the target structure of substantia nigra dopamine neurons, a brain region called the striatum, is sufficient to elicit the movement deficits associated with Parkinson's disease.

Mutations in genes encoding mitochondrial-associated proteins, including PINK1, parkin, DJ-1 and CHCHD2, have been shown to cause Parkinson's disease in humans^{2,3}, demonstrating that mitochondrial dysfunction alone can cause forms of this disorder that are driven by individual mutations. Sporadic Parkinson's disease (cases for which there is no clear single genetic cause) might also stem from mitochondrial impairment. Individuals with this form of the disease have lower than normal levels of mitochondrial complex I in dopamine neurons of the substantia nigra^{4–6}.