New Scientist

WEEKLY 26 November 2022

CAN WE SAVE THE COLORADO RIVER?

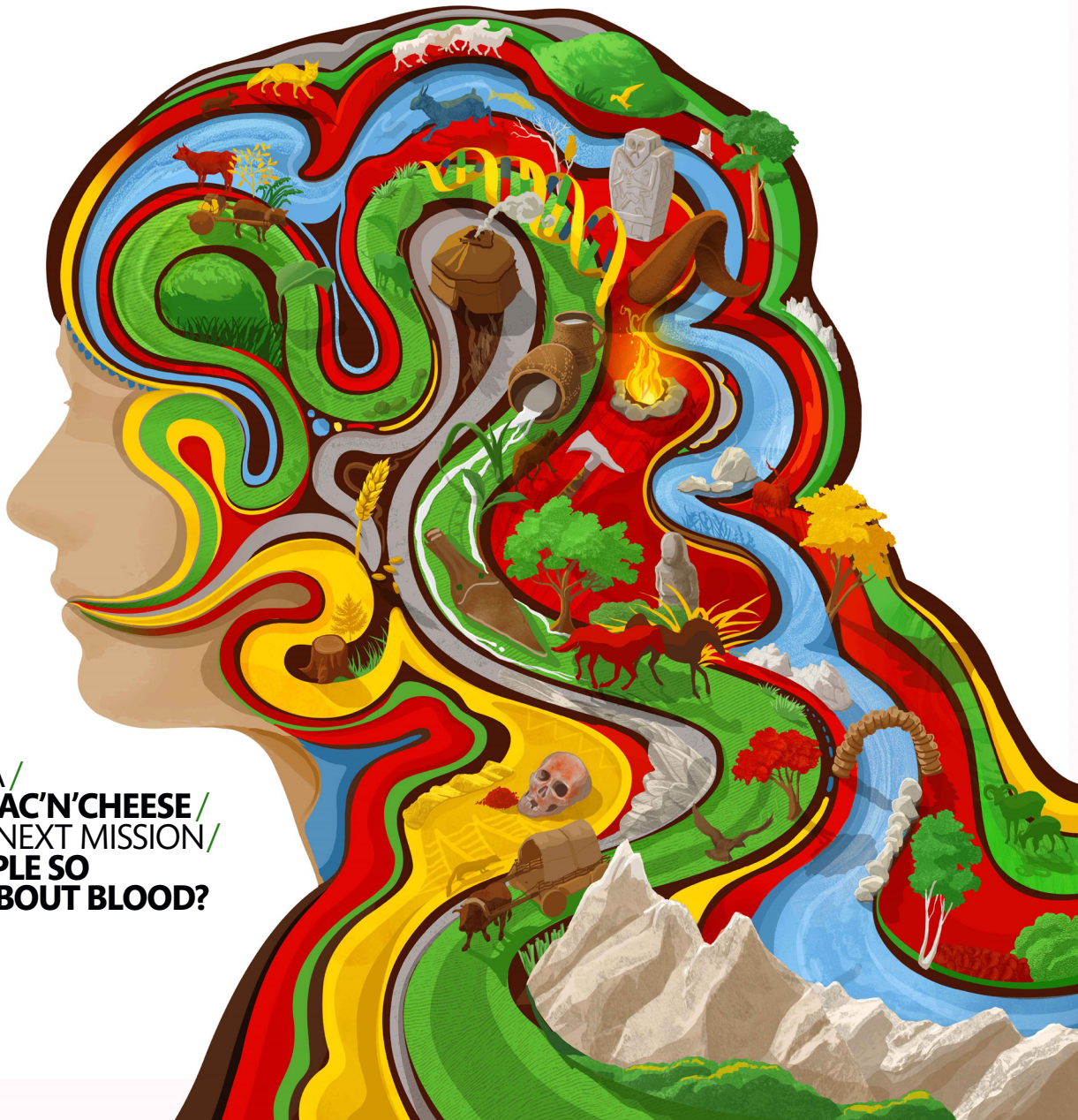
AI SEARCHES FOR NEW LAWS OF PHYSICS

FIRST DRUG TO DELAY TYPE 1 DIABETES

NASA FLIES AROUND THE MOON

MOTHER TONGUE

The hunt for the lost ancestral language of Europe



PLUS
ORCHID MANIA /
HACK YOUR MAC'N'CHEESE /
JAMES BOND'S NEXT MISSION /
WHY ARE PEOPLE SO SQUEAMISH ABOUT BLOOD?

No3414 £6.95 CAN$9.99

9 770262 407374 47

Mother tongue

Almost half the global population speaks a language that belongs to the same vast linguistic family tree – and we are closer than ever to working out where it first took root, finds **Andrea Valentino**

ANDY SMITH

MOTHER. There can scarcely be a more emotive word in the English language. We can imagine children howling it as they wake from nightmares, and centenarians whispering it on their death beds. A 2004 survey proclaimed it the most beautiful word in English, and artists have evoked it in countless poems and plays. Yet even though it can conjure home and hearth in a scant two syllables, mother is perhaps most remarkable for its deluge of cousins. From Dutch (*moeder*) to Czech (*matka*) to Bengali (*ma*), dozens of languages have words that share a common root with mother, tying English to a cobweb of tongues that straddles almost every continent.

Human societies can't exist without language, and no language family has shaped our world as much as Indo-European. It boasts well over 3 billion speakers, or an estimated 46 per cent of everyone on Earth. From the moment this language family was recognised, scholars have been searching for the answer to a weighty question. Who spoke the Indo-European mother tongue – dubbed Proto-Indo-European (PIE) – that splintered into the hundreds of daughters we hear today?

The quest has thrilled and frustrated experts for centuries, with the evidence sometimes pointing in opposing directions. Yet the field is far from deadlocked. With the power of DNA at their heels, geneticists are making new claims about PIE, a language that may predate civilisation. Meanwhile, linguistic studies now suggest we can trace the roots of Indo-European languages even further back than PIE, to the world that existed shortly after farming took hold in south-west Asia. Not that any of this is straightforward – or without controversy.

The story of PIE begins in the 18th century, when William Jones made a startling discovery. A judge working for the British administration in Calcutta (now Kolkata), Jones had always been a passionate linguist. After he moved to India, he became enthralled by Sanskrit, the ancient Indian language still used in modern Hinduism. Jones noticed similarities between Sanskrit and other languages – ones spoken very far from Calcutta. As he put it in a lecture in February 1786, no linguist could examine Greek, Latin and Sanskrit together "without believing them to have sprung" from some common ancestor.

Jones wasn't the first to make this sort of connection, but his lecture would fire a revolution. Soon, scholars were convinced that these languages were indeed related and belonged to what was, in 1813, termed the "Indo-European" language family. By the end of the 19th century, linguists had even begun to reconstruct PIE. They did so by analysing words that share a similar pronunciation and definition across many Indo-European languages – so-called cognates, including mother/*moeder*/*ma* – and by studying the sound changes known to occur as languages evolve. Without texts to help them – the earliest Indo-Europeans didn't have writing – this involved speculation. But in PIE words like *sénos* (old) and *móri* (sea), we may be hearing the distant cries of "senior" and "maritime".

Linguists also advanced a range of hypotheses about where the PIE speakers came from. Among the more popular was the idea that the first Indo-Europeans heralded from the westernmost region of the Eurasian steppe, north of the Black Sea, around 6000 years ago. If, after all, speakers eventually lived everywhere from Europe's Atlantic coast to the Himalayas, it made sense that they started somewhere in the middle. From there, the descendants of PIE speakers were thought to have headed variously south into the Anatolian peninsula,

north and west into Europe and east into southern Asia.

In the 20th century, however, the story became more complicated. Maps once promoted by linguists showed Indo-European populations moving along gracefully arcing lines, but archaeology spoke to a muddled set of relationships between groups judging by, for instance, the geographical distribution of pottery associated with particular ancient cultures. "Every time there's a new paper, it produces as many new problems as solutions," says archaeologist James Mallory at Queen's University Belfast, UK.

Genetic clues

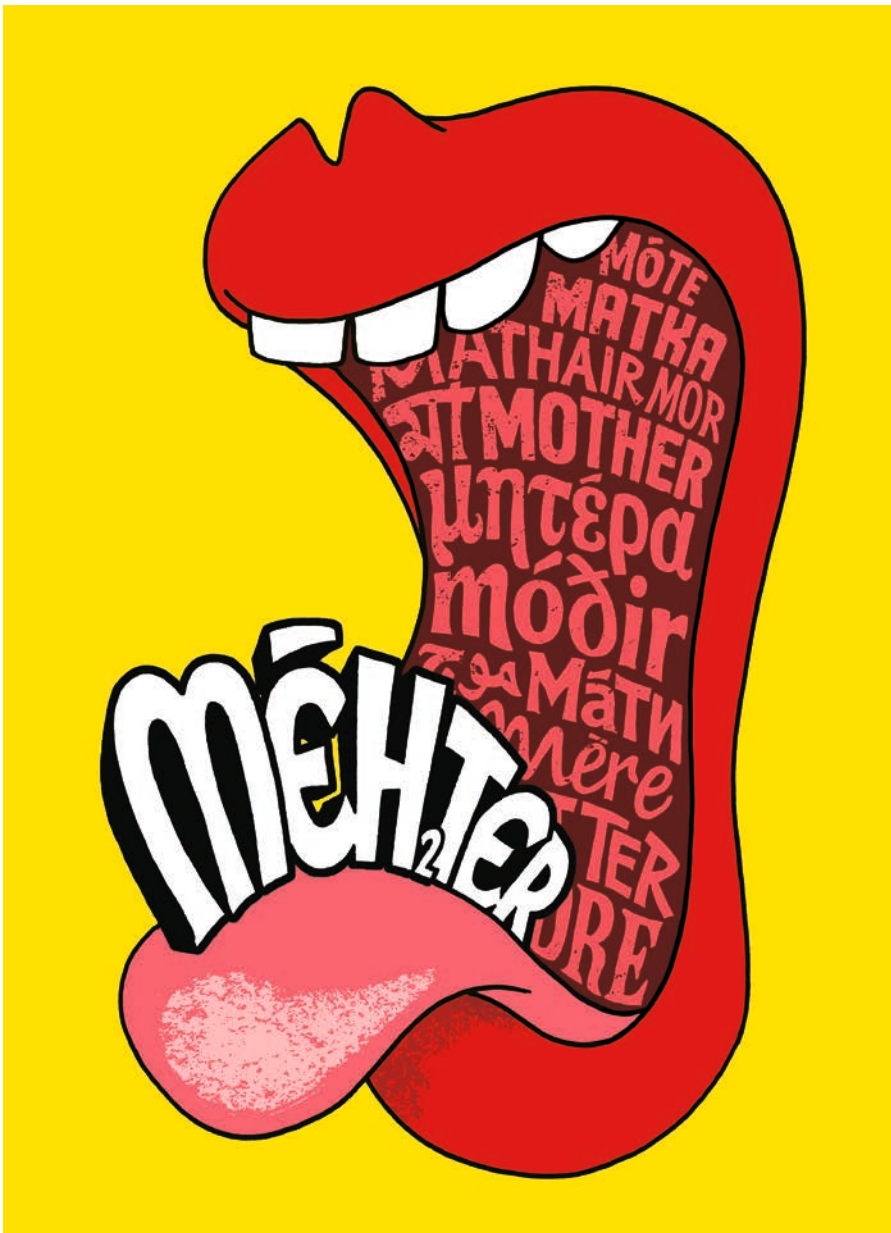
More recently, research into the PIE speakers has been transformed to an even greater extent by genetics. Iosif Lazaridis, a geneticist at Harvard University, recalls a time as recently as 2014 when his colleagues had fragments of ancient DNA from a mere 10 individuals across the whole of Europe. Now, there are hundreds of samples to investigate, taken from ancient graves as far apart as Croatia and Iran.

These developments are already offering up clues for understanding the spread of Indo-European languages. In 2015, for example, Lazaridis and other researchers studied the genetic data of dozens of ancient Europeans who lived between 3000 and 8000 years ago. Their work hinted that the spread of Indo-European languages was at least partly due to a massive migration from the Eurasian steppe.

Some geneticists speculated that this migration was more akin to an invasion, and in recent years, the migrants in question – known to history as the Yamnaya – have gained a fearsome reputation. Portrayed as skilled herders, possibly armed, potentially travelling on wheeled carts, it has been suggested they left a trail of destruction across the continent.

It is certainly true that these migrants completely changed the linguistic landscape: Basque is now the only European language that predates the coming of the Yamnaya. But as Lazaridis and his colleagues have continued their genetic investigations, they have developed a more nuanced picture of what this massive migration really looked like.

Earlier this year, Lazaridis co-published three major new studies on the ancient people of an area he and his colleagues called the Southern Arc, a swathe of territory spanning from the Balkans to the Middle East. The research suggests that relations between local ›



farmers and the incoming herders may not have been so violent, at least not everywhere.

Studying DNA from a high-status burial in Greece, dating to about 1450 BC and complete with a stunning gold-hilted sword, the geneticists found that the individual, dubbed the Griffin warrior, had no steppe ancestry. This seems to discount what Lazaridis calls a "social caste hierarchy" – newcomers at the top, natives at the bottom – especially since humbler Greek graves did show signs of steppe ancestry. For Guus Kroonen, a linguist at Leiden University in the Netherlands, it is possible that Europe's native farmers adopted Yamnaya pastoralism voluntarily, keen to emulate a new and exciting way of life.

Yet, as Lazaridis is learning, not all his ideas have found such easy acceptance – particularly those concerning the earliest history of Indo-European. His latest genetics research finds shortcomings in the suggestion that the PIE speakers lived on the steppe to the north of the Black Sea. Instead, Lazaridis contends that they originated to the east and south of this region.

Again, he stakes his claim on DNA. If the steppe thesis were correct, and the Indo-Europeans originated in the north before some spread south onto the Anatolian peninsula, ancient burials on the north and south sides of the Black Sea should show signs of a shared history. But his team found no steppe ancestry in Anatolian burials, suggesting that people from both places heralded from somewhere else. As Lazaridis explains, that "somewhere" is probably the Caucasus, to the east of the Black Sea. "Anatolia doesn't have this steppe ancestry," he says, but ancestry from the Caucasus is seen in ancient burials both on the steppe and on the Anatolian peninsula.

If Lazaridis is correct, the story of the PIE speakers will need to undergo a significant rewriting. Put simply, it implies that PIE as many have imagined it – a language emerging on the steppe roughly 6000 years ago – is only the second stage in a linguistic journey that may have started centuries earlier in the Caucasus, possibly in what is now Armenia. This means that nomenclature, if nothing else, must change. One straightforward way to think about it could be to imagine traditional PIE as merely "PIE 2", with "PIE 1" serving as the proud parent to both PIE 2 and to the now-extinct Indo-European languages that were spoken on the Anatolian peninsula millennia ago.

But speak to the linguists and they are far from convinced by Lazaridis's idea. Some feel the geneticists came up with this new scenario by ignoring the contributions linguists have made to the PIE story. Alwin Kloekhorst, a linguist at Leiden University, notes something missing from one of the Southern Arc studies in which Lazaridis and his colleagues set out their new scenario. "There were no linguists involved in the whole paper," he says. "I think that's a bit annoying, to be honest."

In particular, linguists have reconstructed PIE words associated with pastoralism and herding, the sorts of activities you would expect from nomads on the steppe. But this way of life wasn't common in the Caucasus. There, people were crop farmers, from around

"The linguistic journey may have started in the Caucasus, possibly in what is now Armenia"

New Scientist audio
You can now listen to many articles – look for the headphones icon in our app **newscientist.com/app**

THE DEEP ROOTS OF LANGUAGE

With linguistic detective work, we can reconstruct some of the vocabulary of Proto-Indo-European (PIE), a language spoken roughly 6000 years ago. Now, some linguists have set their sights on an older hypothetical language called Proto-Indo-Uralic, spoken perhaps 9000 years ago (see main story). Can we rewind the language clock even further?

The answer is both yes and no. If the goal is to reconstruct ancient vocabulary, PIE is probably about as far back as we can go, says Guus Kroonen at Leiden University, the Netherlands. Even PIE would have been difficult to reconstruct if not for the fact that linguists gleaned important clues from ancient texts – some more than 3000 years old – written in Indo-European languages including Greek and Sanskrit, he says.

But if the goal is merely to identify ancient languages rather than reconstruct them, we might go back further, says Alwin Kloekhorst, also at Leiden University. For instance, some scholars think the Indo-European and Uralic language families might be distantly related to Mongolian and Turkic, which, if correct, "would imply that a Proto-Indo-Uralo-Mongolo-Turkic language may have been spoken before 12,000 BC", he says.

Other linguists have identified even older languages, such as Proto-Afro-Asiatic. This may have been spoken 18,000 years ago and gave rise to hundreds of languages now spoken in parts of Africa and western Asia.

the 6th millennium BC, several thousand years before the PIE speakers emerged.

If there really was a PIE 1 spoken in the Caucasus, you would expect it to include crop terminology – and for at least some of those terms to have been carried north and incorporated into PIE 2.

Kroonen says this didn't happen. Though it boasted a single term for cereal, he says the oldest form of Indo-European lacked words for peas, lentils and other legumes. In fact, even when the Yamnaya did, much later, migrate further west into Europe and begin farming in earnest, some linguists believe they borrowed local words to describe their new crops. As Kroonen says, *fagiolo* (Italian) and *fasóli* (Greek), both meaning bean, may ultimately have been derived from some ancient and now extinct European tongue unrelated to the Indo-European languages. "It seems that the [language] package is very much incomplete," says Kroonen.

Finnish connection

Lazaridis accepts these criticisms gracefully. He describes Kroonen's critique as "very interesting" and stresses that genetics can't hope to solve every mystery alone. All the same, he is eager to hold the line against some linguistic incursions. When it comes to the absence of farming terms in early Indo-European, for instance, he says the people who moved from the Caucasus to the steppe – the PIE 1 speakers – may not have been farmers at all. As Lazaridis explains, agriculturalists shared the region with hunter-gatherers. If the PIE 1 speakers were hunter-gatherers, they would have had little need for legume terminology. Another possibility, says Lazaridis, is that PIE 1 speakers were farmers who had farming words, which they abandoned when they moved to the steppe and stopped farming.

With so much back and forth, progress can feel slow. But while linguists and geneticists currently have their disagreements, there is surely room for optimism. If Lazaridis is happy to reconcile his work with the linguistic evidence, linguists like Kroonen seem equally willing to meet him and other geneticists halfway. Kroonen says one of the "great strengths of the genetic revolution" has been the ability of scientists to look for population movements postulated by linguists.

Mallory, who wrote his opus on the Indo-Europeans some three decades ago and can perhaps look at the field with more distance, seems sanguine too. "With any field or any science that enters into a debate, it will initially be fairly simplistic interpretations," he says. "But it gets more and more contentious as you get more data." Not that this is necessarily a problem, he adds, given that the evidence of past migrations we can now glean from ancient DNA is far more sophisticated than anything we could work out from archaeological remains or linguistic tinkering alone.

Yet even as the field argues, some linguists are turning their gaze to an even remoter past. No language appears from nowhere and even PIE must have had precursors in the depths of time (see "The deep roots of language", left). For Kloekhorst, that could be what he and some other scholars call Proto-Indo-Uralic (PIU). This idea seeks to dovetail the Indo-European family of languages with the Uralic family, a set spoken by about 25 million people that includes Hungarian and Finnish. Giving his super-family a potential homeland near the Ural mountains in Russia, Kloekhorst says he has detected certain similarities in Indo-European and Uralic grammar and vocabulary. Examples include the apparent resemblance between the English "me" and the Finnish *minä*, and English "water" and Finnish *vesi*. He concedes that this shared inheritance is "very small", but equally stresses that the inheritance involves basic, everyday words that are the most conservative parts of any language, and so the least likely to have been borrowed later.

Kloekhorst's approach turns the linguistic clock back bewilderingly far: if it existed, PIU would have been spoken around 9000 years ago. It leads to a scenario in which PIU split and gave rise on the one hand to the common ancestor of the Uralic languages and on the other hand to PIE 1. From there, the Uralic languages would have continued to develop in the north of Eurasia, while PIE 1 speakers moved south and west to the Caucasus.

Kroonen thinks most linguists would say the PIU idea hasn't yet been confirmed. Any similarities between the Indo-European and Uralic families could just be coincidence, he says. Nevertheless, the idea is gaining ground: 2019 saw the publication of a major book on PIU. It is a serious hypothesis, says Kroonen.

Scholars have already spent more than 200 years searching for the mother of the languages spoken by billions. They may just be getting started. ▮

Andrea Valentino is a journalist based in New York