Genomic insights into the origin of prehistoric populations in East Asia
22 February 2021

East Asia today harbors more than a fifth of the world's population and some of the most deeply branching modern human lineages outside of Africa. The archaeological record attests that this region was also one of the earliest centers of animal and plant domestication. However, its genetic diversity and deep population history remain poorly understood relative to many other parts of the world. In a new study, a team of international researchers analyzes genome-wide data for 166 ancient individuals spanning 8,000 years and 46 present-day groups, and provides insights into the formation of East Asian populations.

Researchers have long debated whether the peopling of East Asia by modern humans occurred mainly via a coastal or interior route. The answer is probably both. "Indigenous Andaman islanders of the Bay of Bengal, Indigenous Tibetans, ancient Taiwanese, and ancient and modern Japanese all derive ancestry from a deep shared lineage that split from other East Asian lineages more than 40,000 years ago," says David Reich, co-senior author of the study, who is a Professor of Genetics and Human Evolutionary Biology at Harvard University and an Investigator of the Howard Hughes Medical Institute. "The simplest way to explain this is if some of the earliest modern humans in East Asia spread along a coastal route linking southeast Asia, coastal China, and the Japanese Archipelago." In contrast, the 40,000 year old Tianyuan individual—along with present-day and early Holocene people from Mongolia and the Amur River Basin—derive almost all their ancestry from a different early splitting lineage. This lineage plausibly expanded by an interior route and contributed in the highest proportion to ancient Mongolians, people of the Amur River Basin, and early farmers from the West Liao and Yellow River regions.

How East Asian populations got to be how they are today

Today there is a fairly high degree of genetic homogeneity in East Asia, reflecting mixtures of deeply splitting lineages that occurred especially in the last ten thousand years. To connect the population structure of present-day East Asians to ancient lineages, co-first author Chuan-Chao Wang, Director of the Institute of Anthropology at Xiamen University, led a collection of genome-scale data from 383 people from 46 diverse populations from East Asia. The study team also generated genome-scale ancient DNA data from 166 individuals from mainland China, Taiwan, Mongolia, Japan and Russia, ranging in age up to about 8,000 years ago. Wang led the statistical analysis of these data, along with previously published data from other research groups, to understand how the ancient and modern people were connected with each other. Say Wang: "With high resolution modern and ancient data, we can connect the dots between ancient and modern East Asians."

Testing the extent to which present-day language distributions reflect ancient farming expansions

Increases of population densities in and around centers of domestication created demographic
expansions that in principle could have spread people and languages. The authors tested several hypotheses about whether particular farming expansions were connected to language spreads.

The first is the 'Transeurasian hypothesis': the theory that languages of the Mongolic, Turkic, Tungusic, Koreanic, and Japonic families may descend from a proto-language associated with the expansion of early millet farmers around the West Liao River in northeast China. The study confirms a shared genetic link between present-day people speaking Transeurasian languages: they all have large proportions of ancestry from the deeply splitting Tianyuan-related lineage. However, this link dates to before farming expansions, as Mongolic, Turkic, and Tungusic speakers do not have the characteristic mixture of ancient lineages that is a signature of West Liao River farmers. However, West Liao River farmers do appear to be an excellent candidate for the primary source population for present-day Korean and Japanese people.

A second hypothesis concerns the origins of Sino-Tibetan languages. The authors show that most of the ancestry of both Tibetans and Han Chinese comes from a population closely related to Yellow River farmers who lived around 3,000 BCE. "The findings show that Neolithic farmers cultivating foxtail millet in the Upper and Middle Yellow River Basin expanded southwest toward the Tibetan Plateau to spread present-day Tibeto-Burman languages, and east and south towards the Central Plains and eastern coast to spread Sinitic languages including the linguistic ancestor of Han Chinese," says co-first author Huqin Zhang, a professor at Xi'an Jiaotong University, China.

A third hypothesis concerns the spread of languages and rice farming. Previous ancient DNA analysis correlates the spread of rice farming across Southeast and South Asia with Austroasiatic, Tai-Kadai, and Austronesian languages. Through modeling of diverse Indigenous groups from southern China, this study strengthens the connections to these diverse language groups.

How West and East Eurasian peoples met and mixed

Mongolia falls near the eastern extreme of the Eurasian Steppe, and archaeological evidence shows that throughout the last ten thousand years the region was a conduit for cultural exchanges between East and West Eurasia. For example, the Afanasievo culture, an eastward extension of the Yamnaya steppe pastoralist culture, impacted Mongolia around the turn of the third millennium BCE, although in contrast to Europe, where Yamnaya ancestry became permanently established after it first arrived, a near-complete displacement of the ancestry brought by these migrants occurred in Mongolia following the decline of the Afanasievo culture. Strikingly, ancestry from the initial Yamnaya expansion survived in northwest China at least into the Iron Age. "This increases the plausibility of the theory that the Yamnaya expansion was a vector of spread, not only of the ancestors of all spoken Indo-European languages, but also of the early-branching Tocharian languages attested from written records in Iron Age western China," says co-senior author Johannes Krause, professor and director at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany.

"Our analysis also documents subsequent ancestry impacts from the west on Mongolia," says co-senior author Ron Pinhasi of the University of Vienna. "One of these brought Yamnaya ancestry back to Mongolia more than a millennium after it first appeared, this time mixed with about a third ancestry from European farmers." Pinhasi adds: "The other impact from the west began about 4,500 years ago, bringing ancestry along the Inner Asian Mountain Corridor related to the people of Iran and Central Asia. However, this ancestry made only transient early contributions to populations in Mongolia and western China, and does not appear again in our ancient DNA time for two millennia."
