

Rare Ancient DNA Provides Window Into a 5,000-Year-Old South Asian Civilization

The Indus Valley Civilization flourished alongside Mesopotamia and Egypt, but the early society remains shrouded in mystery. July 1, 2021

[Smithsonian Magazine](#)

- *Brian Handwerk*



Mohenjo Daro, or "Mound of the Dead" is an ancient Indus Valley Civilization city that was abandoned in the 19th century BCE as the Indus Valley Civilization declined and was not rediscovered until the 1920s. Built around 2500 BCE, it was one of the largest settlements of the ancient Indus Valley civilization, and one of the world's earliest major urban settlements. Photo by SM Rafiq Photography / Getty Images.

During the last few millennia B.C., beginning roughly 5,000 years ago, great civilizations prospered across Eurasia and North Africa. The ancient societies of Mesopotamia and Sumer in the Middle East were among the first to introduce written history; the Old, Middle and New Kingdoms of Egypt established complex religious and social structures; and the Xia, Shang and Zhou dynasties ruled over ever advancing communities and technologies in China. But another, little understood civilization prevailed along the basins of the Indus River, stretching across much of modern Afghanistan and Pakistan and into the northwestern regions of India.

This [Indus Valley Civilization \(IVC\)](#), also called the [Harappan civilization](#) after an archaeological site in Pakistan, has remained veiled in mystery largely due to the fact that scholars have [yet to make sense of the Harappan language](#), comprised of fragmented symbols, drawings and other writings.

Archaeological evidence gives researchers some sense of the daily lives of the Harappan people, but scientists have struggled to piece together evidence from ancient DNA in the IVC due to the deterioration of genetic material in the hot and humid region—until now.

For the first time, scientists have sequenced the genome of a person from the Harappan or Indus Valley Civilization, which peaked in today's India-Pakistan border region around 2600 to 1900 B.C. A trace amount of DNA from a woman in a 4,500-year-old burial site, painstakingly recovered from ancient skeletal remains, gives researchers a window into one of the oldest civilizations in the world. The work, along with a comprehensive analysis of ancient DNA across the Eurasian continent, also raises new questions about the origins of agriculture in South Asia.

The ancient Harappan genome, sequenced and described in the journal [Cell](#), was compared to the DNA of modern South Asians, revealing that the people of the IVC were the primary ancestors of most living Indians. Both modern South Asian DNA and the Harappan genome have a telltale mixture of ancient Iranian DNA and a smattering of Southeast Asian hunter-gatherer lineages. "Ancestry like that in the IVC individuals is the primary ancestry source in South Asia today," co-author David Reich, a geneticist at Harvard Medical School, said in a statement. "This finding ties people in South Asia today directly to the Indus Valley Civilization."

The genome also holds some surprises. Genetic relationships to Steppe pastoralists, who ranged across the vast Eurasian grasslands from contemporary Eastern Europe to Mongolia, are ubiquitous among living South Asians as well as Europeans and other people across the continent. But Steppe pastoralist DNA is absent in the ancient Indus Valley individual, suggesting similarities between these nomadic herders and modern populations arose from migrations after the IVC's decline.

These findings influence theories about how and when Indo-European languages spread widely across the ancient world. And while shared ancestry between modern South Asians and early Iranian farmers has fueled ideas that agriculture arrived in the Indo-Pakistani region via migration from the Fertile Crescent of the Middle East, the ancient Harappan genes show little contribution from that lineage, suggesting that farming spread through an exchange of ideas rather than a mass migration, or perhaps even arose independently in South Asia.

"The archaeology and linguistic work that had been carried out for decades was really at the forefront of our process," says [Vagheesh Narasimhan](#), a Harvard University genomicist and co-author of the study. "These projects bring a new line of genetic evidence to the process, to try to show the impact that the movement of people may have had as part of these two great cultural transformations of agriculture and language."

The large, well-planned cities of the IVC included sewer and water systems, as well as long-distance trade networks that stretched as far as Mesopotamia. But despite its former glory, the civilization was unknown to modern researchers until 1921, when excavations at Harappa began to uncover an ancient city. The Harappans have remained something of a mystery ever since, leaving behind extensive urban ruins and a mysterious language of symbols and drawings, but few additional clues to their identity. What ultimately befell the Harappan civilization is also unclear, though [a changing climate has been posited](#) as part of its downfall.

Scientists have a notoriously difficult time recovering ancient DNA in South Asia, where the subtropical climate typically makes genetic preservation impossible. It took a massive, time-consuming effort to produce the genome from remains found in the cemetery at Rakhigarhi, the Harappans' largest city, located in the modern Indian state of Haryana. Scientists collected powder from 61 skeletal samples, but just one contained a minute amount of ancient DNA. That sample was sequenced as much as possible, generating 100 different collections of DNA fragments, called libraries, each of which were too incomplete to yield their own analysis.

"We had to pool 100 libraries together and sort of hold our breath, but we were fortunate that that yielded enough DNA to then do high resolution population genetics analysis," Narasimhan says. "I think if anything, this paper is a technical success story," he adds, noting that the approach holds promise for sourcing DNA in other challenging locales.

A single sample is not representative of a widespread population that once included a million or more people, but a related [study](#) published today in *Science* lends some wider regional context. Several of the same authors, including Narasimhan and Reich, and dozens of international collaborators, authored the largest ancient DNA study published to date. Among the genetic sequences from 523 ancient humans are individuals from sites as far flung as the Eurasian Steppe, eastern Iran and Iron Age Swat Valley in modern Pakistan.

The team found that among many genetically similar individuals, a handful of outliers existed who had ancestry types completely different from those found around them.

Eleven such individuals found at sites in Iran and Turkmenistan were likely involved in interchange with the Harappan civilization. In fact, some of these outlier individuals were buried with artifacts culturally affiliated with South Asia, strengthening the case that they were connected to the IVC. “This made us hypothesize that these samples were migrants, possibly even first-generation migrants from South Asia,” Narasimhan says. The IVC genome from Rakhigarhi shows strong genetic similarities to the 11 genetic outliers in the large study of ancient humans, supporting the idea that these individuals ventured from the Harappan civilization to the Middle East. “Now we believe that these 12 samples, taken together, broadly represent the ancestry that was present in [South Asia] at that time.” The first evidence of agriculture comes from the Fertile Crescent, dating to as early as 9,500 B.C., and many archaeologists have long believed that the practice of growing crops was brought to South Asia from the Middle East by migrants. Earlier DNA studies seemed to bear out this idea, since South Asians today have significant Iranian ancestry.

“I really found their analysis to be very exciting, where they look at ancient DNA samples from different time scales in Iran and try to correlate how the Iranian ancestry in South Asians is related to those different groups,” says [Priya Moorjani](#), a population geneticist at UC Berkeley not involved in the *Cell* study of the IVC genome.

However, the new analysis shows that the first farmers of the Fertile Crescent appear to have contributed little, genetically, to South Asian populations. “Yet similar practices of farming are present in South Asia by about 8,000 B.C. or so,” says Moorjani, a co-author on the wider population study of South and Central Asia. “As we are getting more ancient DNA, we can start to build a more detailed picture of how farming spread across the world. We’re learning, as with everything else, that things are very complex.”

If farming did spread from the Fertile Crescent to modern India, it likely spread via the exchange of ideas and knowledge—a cultural transfer rather than a significant migration of western Iranian farmers themselves. Alternatively, farming could have arisen independently in South Asia, as agricultural practices started to sprout up in many places across Eurasia during this time.

Ancient IVC ancestry holds other mysteries as well. This civilization was the largest source population for modern South Asians, and for Iron Age South Asians as well, but it lacks the Steppe pastoralist lineages common in later eras. “Just like in Europe, where Steppe pastoralist ancestry doesn’t arrive until the Bronze Age, this is also the case in South Asia,” Narasimhan says. “So this evidence provides information about the timing of arrival of this ancestry type, and their movement parallels the linguistic phylogeny of Indo-European languages, which today are spoken in places as far away as Ireland to New Delhi.”

The authors suggest Indo-European languages may have reached South Asia via Central Asia and Eastern Europe during the first half of the 1000s B.C., a theory [evidenced by some genetic studies](#) and similarities between Indo-Iranian and Balto-Slavic languages.

Narasimhan hopes that more genetic data can help clear up this ancient puzzle—especially by exploring where DNA dovetails or differs with findings from other lines of evidence.

“We’re trying to look at when and how archaeological cultures are associated with a particular genetic ancestry, and whether there’s any linguistic connections,” he says. “To understand human history, you really need to integrate these three lines.”

Brian Handwerk is a freelance writer based in Amherst, New Hampshire.