Introduction

Hello, and welcome to the April 2012 issue of DNA Tribes® Digest. This month’s feature article explores genetic links in the Indus Valley, the location of the Harappan Civilization of the Bronze Age. In the past, Western academics have assumed the Harappan culture was replaced by Vedic invaders from the Eurasian steppe. However, this older model of a “Vedic invasion” is being reconsidered in the light of new archaeological, linguistic, and textual evidence.

In its place, a new alternative model is beginning to emerge that considers both Harappa and the Vedas as parts of an indigenous continuum of South Asian cultures that began in the Neolithic period. This article will discuss archaeological (Harappan) and cultural (Vedic and Buddhist) evidence, to provide a context for evidence of the Indus Valley’s genetic links with neighboring cultures of West Asia dating to the Neolithic period.

For a more in depth discussion, the book The Quest for the Origins of Vedic Culture by Edwin Bryant provides an up to date scholarly analysis of these topics.

Best regards,
Lucas Martin
DNA Tribes

DNA Tribes® is now on Facebook:
http://facebook.com/DNAtribes
Genetic Links in the Indus Valley

Historical Background

**Geographical Setting:** South Asian cultures have traditionally symbolized the world as a Jambudvipa (“island of jambu trees”), sometimes described as one of several lands that surround a central mountain (Mount Meru) like the petals of a lotus. Like the mythological “jambu island,” the Indian Subcontinent joins Asia at the Himalayas, which are surrounded by West Asia, Central Asia, the Eurasian Steppe, the Tarim Basin, and East Asia.

In this geographical location, South Asia has been secluded by mountains, but also connected to other parts of Asia by land and sea trade routes. Archaeologists have uncovered evidence here for a gradual development of agriculture beginning in Mehrgarh (between 7,000-3,300 BCE in present day Balochistan) and emerging as a large and sophisticated Bronze Age civilization known as the Indus Valley or Harappan civilization between 3,300-1,300 BCE (see Figure 1).

![Trade Links of the Indus Valley (Harappan) Civilization](image)

**Figure 1:** The Bronze Age Indus Valley Civilization (blue) and trade links (orange).

**Bronze Age Harappa:** Ancient Harappa flourished as a prosperous urban civilization in a large area of the northwestern Indian Subcontinent (see Figure 1), encompassing a larger area than any other Bronze Age civilization. Harappan society included large planned cities, standard weights and measures (similar to those used in Egypt), and a wide network of trade contacts that included the Persian Gulf, Central Asia, Mesopotamia, and possibly reaching as far as Egypt and Minoan Crete.1

In contrast to the absolute rulers of contemporary Egypt or Mesopotamia, the Harappan civilization was apparently more egalitarian, with no grand palaces or royal tombs discovered. Unlike the

---

1 For a discussion of archaeological evidence for Harappa’s trade contacts, see the video lecture “Meluhha: the Indus Civilization and its Contacts with Mesopotamia” by Dr. Jonathan Mark Kenoyer at [http://www.youtube.com/watch?v=8zcGLfLEbml](http://www.youtube.com/watch?v=8zcGLfLEbml).
imposing ziggurat temples of Mesopotamia, Harappan worship probably involved sacred groves (depicted in ancient seals) and washing ceremonies in locations such as the “Great Bath” of Mohenjo-daro. The Harappan economy emphasized village farming and cattle pastoralism, but also supported a rich tradition of fine trade crafts that were exported to Mesopotamia.

The Old Theory (Vedic Invasion): Despite this vivid archaeological record of Harappan life, the ancient Harappan language (or languages) is unknown, because the Indus Script remains undeciphered. When the ancient cities of Harappa were excavated, Western archaeologists generally assumed that the ancient Harappan culture had been replaced by invaders from the Eurasian Steppe and Central Asia at the end of the Harappan period around 1700-1300 BCE. The new Central Asian invaders were thought to be the composers of the *Rigveda* and other Vedic literature written in the Sanskrit language, ancestral to Hindi and other languages spoken throughout South Asia to the present day.

This “invasion theory” remains the traditional academic model for Indian prehistory, in part because it explains the similarity of Sanskrit to ancient Greek and Latin. However, archaeologists have not found clear evidence for a culture from the Eurasian Steppe or Central Asia that influenced South Asia in the relevant period. The best effort to address this lack of archaeological evidence is currently the “Kulturkugel” model, in which invaders spread a new Indo-European language without noticeably impacting the material culture of South Asia. Similarly, linguistic evidence for any pre-Vedic “substrate” language of the Indus Valley is somewhat limited.

A New Model (Vedic Harappans): To address this lack of evidence for Vedic invaders from Central Asia, some scholars are beginning to explore evidence for greater antiquity of the Vedic culture (dating to the Harappan period) and a South Asian geographical setting for the *Rigveda* and other texts.

Proposed evidence for the antiquity of Vedic cultures has included astronomical references in Vedic texts that date to 2,500 BCE and possibly older based on changed star positions (due to the precession of the equinoxes). Geographical terms in the *Rigveda* suggest a South Asian setting, including areas near the Indus Valley and as far east as the lower Ganges. Similarly, river names in the Punjab suggest the local antiquity of Sanskrit speaking cultures in northern India.²

Most importantly, the *Rigveda* itself does not mention any migration to northern India. In contrast, related Zoroastrian texts from Central and West Asia do mention a migration from an earlier homeland (possibly near the Hindu Kush Mountains). Early evidence from outside of India also includes West Asian Mitannian and Kassite cultures (contemporary with the Harappan Civilization), which used *Rigvedic* like deity names and the peacock (a South Asian animal) as an artistic motif. Taken together, this suggests the possibility that Vedic cultures were indigenous South Asians (possibly one of several Harappan cultures), appearing in West Asia through the trade links known to archaeologists.

The Language Puzzle and Evidence for Early Migrations: A new model of “Vedic Harappans” would however, create a new puzzle: if there was no Vedic invasion, how did Indo-European languages find their way to both Europe and South Asia? Archaeological evidence supports two major expansions into South Asia: (1) a Neolithic expansion (possibly from West Asia) between 6,000-4,500 BCE; and (2) an Iron Age expansion (possibly from Central Asia) between 800-200 BCE.

Neolithic (Pre-Harappan) Expansion: One possibility is that food producing cultures of West Asia brought Indo-European languages to South Asia during the Neolithic expansion (6,000-4,500 BCE).³

² Although the language(s) of the Indus script are unknown, S.R. Rao has proposed a Sanskrit decipherment based on a comparison to early Canaanite alphabets (related to the Phoenician, Hebrew, and later Greek scripts).
³ Linguists have proposed early Indo-European “Euphratic” language linked with Southern Mesopotamia, preserved as a substratum of ancient Sumerian and Akkadian. Since Sumerian dates to at least 3,000 BCE, Euphratic would be the oldest Indo-European language attested anywhere in the world.
This early date for the languages ancestral to Sanskrit would not contradict the Neolithic date for the Proto-Indo-European language that has been proposed by some linguists.4

**Iron Age (Shakya) Expansion:** The second expansion dating to 800-200 BCE has been associated with Shakya or Saka (Scythian related) cultures from Central Asia that influenced early Buddhist culture in India. For instance, the Sanskrit scholar Michael Witzel has suggested Central Asian links for some Shakya customs, such as the use of burial mounds (stupas) and Zoroastrian concepts in Buddhist literature.

First emerging in Sākyamuni’s native kingdom of Lumbini (in present day Nepal), Buddhism eventually spread outward from the Indian Subcontinent and flourished in the mercantile Silk Road oasis settlements of Central Asia. In the context of a “Vedic Harappans” model, these Shakyas might have been peripheral Harappan or Vedic influenced cultures from Central Asia that returned to the core Vedic location of India during the Iron Age.

**Migration Summary:** In summary, the archaeological record generally supports the indigenous development of South Asian civilization since the Neolithic period (see Table 1). The Harappan Civilization emerged during a long period during which no major expansions into South Asia are attested. Shortly thereafter, the Vedic texts (which do not describe a migration from elsewhere) were recorded. This suggests that the Harappan Civilization and Sanskrit Vedic texts were indigenous expressions of related South Asian cultures.

<table>
<thead>
<tr>
<th>Period</th>
<th>Possible Origin</th>
<th>Possible Cultural Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>6,000 - 4,500 BCE</td>
<td>West Asia</td>
<td>Expansion of Neolithic food producing cultures.</td>
</tr>
<tr>
<td>3,300 - 1,300 BCE</td>
<td>Indigenous</td>
<td>Harappan Civilization. Indus script (undeciphered).</td>
</tr>
<tr>
<td>1,500 - 400 BCE</td>
<td>Indigenous</td>
<td>Vedic texts written in the Sanskrit language. Oral tradition of Vedic chant (Srauta) continues to present day.</td>
</tr>
<tr>
<td>800 - 200 BCE</td>
<td>Central Asia</td>
<td>Expansion of Shakya (Scythian related) cultures.</td>
</tr>
</tbody>
</table>

Table 1: Periods of the South Asian expansions attested in the archaeological records, together with the periods of the Harappan Civilization and composition of the Vedic texts.

---

4 See [http://language.psy.auckland.ac.nz/files/gray_and_atkinson2003/grayatkinson2003.pdf](http://language.psy.auckland.ac.nz/files/gray_and_atkinson2003/grayatkinson2003.pdf). However, a Neolithic date for Proto-Indo-European is not accepted by most linguists, who still generally support an origin for the Indo-European languages in the Chalcolithic (Copper Age) period.
STR Analysis of the North India Region

Genetic contributions to the North India region were identified based on autosomal STR data.\(^5\) Results are summarized in Table 2 and illustrated in Figure 2.

![Diagram of genetic contributions to the North India region](image)

**Figure 2**: Genetic contributions to the North India region (STR). For more about the world regions in DNA Tribes® STR based 15, 21 and 27 Marker Kit tests, see [http://dnatribes.com/populations.html](http://dnatribes.com/populations.html).

<table>
<thead>
<tr>
<th>World Region</th>
<th>Estimated Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eastern India</td>
<td>35.8%</td>
</tr>
<tr>
<td>Mesopotamian</td>
<td>27.4%</td>
</tr>
<tr>
<td>South India</td>
<td>22.0%</td>
</tr>
<tr>
<td>[Probably Caspian] N.W. European</td>
<td>7.8%</td>
</tr>
<tr>
<td>Altaian</td>
<td>3.4%</td>
</tr>
<tr>
<td>Other</td>
<td>3.6%</td>
</tr>
</tbody>
</table>

**Table 2**: Genetic contributions to the North India region (STR).

**Discussion**: Results in Table 2 indicate genetic links with several regions neighboring North India. Two of the largest components identified were from other regions within the Indian Subcontinent, including the Eastern India (35.8%) and South India (22.0%) regions. This might reflect genetic links within South

---

\(^5\) For more information about DNA Tribes® STR based 15, 21 and 27 Marker Kit tests, see [http://dnatribes.com/index.html](http://dnatribes.com/index.html).
Asia, including indigenous populations living in North India prior to the Neolithic period, as well as ongoing contacts within South Asia during the Vedic period.

Results also indicate genetic links with the Mesopotamian region (27.4%). This might express contacts dating to the expansion of Neolithic food producing cultures from West Asia between 6,000 and 4,500 BCE, possibly including the migrations that brought early Sanskrit related languages to South Asia. This might also suggest later Harappan links with Mesopotamia, attested in archaeological traces of trade networks and suggested by historical evidence of Vedic cultural concepts adopted by Mitannian (Hurrian) and Kassite cultures of West Asia.

Results also suggest some genetic similarity to populations of Northwest Europe (7.8%). This might reflect genetic traces of early Neolithic or Copper Age expansions related to the spread of Indo-European languages. These language expansions might have originated elsewhere, but left genetic traces in peripheral areas (such as Northwest Europe). One possibility is that these genetic connections in distant locations of Asia and Europe express ancestry from ancient populations that are no longer extant or not yet represented by available STR data (such as Caspian Sea and North Caucasus populations).

Finally, results indicated genetic links with the Altaian region (3.4%). This might express contacts with Siberia via Central Asia, perhaps including the expansions of Shakya cultures during the Iron Age.

---

6 Linguists have variously proposed Proto-Indo-European origins in several locations, such as: Anatolia, the Balkans, the Pontic Steppe, or the Transcaucasus (to name just a few). Most of these locations are near meeting points between West Asia and Europe, consistent with a role in mediating between early agricultural technologies from the Fertile Crescent and hunting-fishing lifeways in peripheral “wilderness” areas of West Eurasia.

7 Based on separate SNP data, some populations near the Caspian Sea and North Caucasus are (somewhat) more similar to Northwest Europeans compared to other West Asians. These include Dargins, Lezgins, Chechens, and Balkars. STR data for these populations are not yet available. However, SNP analysis that includes these populations identifies a Caucasus-Anatolian component in Indus populations (see next section of this article), instead of the Mesopotamian and Northwest European links identified based on STR data.
SNP Analysis of Indus Valley Populations

Regional admixture components in Indus Valley samples and several neighboring populations (excluding local Indus Valley admixture components) were identified based on autosomal SNP data. Results are summarized in Table 3 and illustrated in Figure 3.

Discussion: Results in Table 3 indicate two primary contributions to all studied populations: South India and Caucasus-Anatolian. Smaller Baltic-Urals, Arabian, and Mongolian components were also identified in some populations, primarily near the periphery of the Indian Subcontinent.

South India components were highest towards the interior of the Indian Subcontinent, for instance: Kurmi (82.0%); Gujarat (80.7%); and Kshatriya Uttar Pradesh (72.6%). South India components were found in populations speaking multiple languages throughout the Indus Valley region, including: Sindhi (52.7%); Brahui (33.3%); Burusho (47.9%); and Kalash (39.6%). Beyond the Indus Valley proper, smaller South India contributions were also found in Central Asian populations, such as Tajik (19.9%) and Turkmen (12.6%). These South Asian components indicate genetic continuity with more southern and eastern parts of the Indian Subcontinent that extends throughout the Indus Valley and beyond to parts

---

For more information about DNA Tribes® SNP analysis, see [http://www.dnatribes.com/snp.html](http://www.dnatribes.com/snp.html).

---

8 For more information about DNA Tribes® SNP analysis, see [http://www.dnatribes.com/snp.html](http://www.dnatribes.com/snp.html).
of Central Asia. This might express contacts facilitated by the extensive trade network of the Bronze Age Harappan Civilization, as well as other periods.

The second substantial genetic component identified for all studied populations was from the Caucasus-Anatolian region. Caucasus-Anatolian components were largest in westerly populations, such as: Balochi (61.9%) and Brahui (59.8%). Caucasus-Anatolian components were also high in Central Asian populations, such as Turkmen (72.5%) and Tajik (57.2%). This component generally was smaller in more easterly populations, such as Brahmin Uttar Pradesh (22.9%) and Kurmi (17.5%).

Given the archaeological record of India, this Caucasus-Anatolian component could reflect expansions from West Asia, such as during the spread of Neolithic food producing cultures into South Asia between 6,000 and 4,500 BCE. This West Asian expansion might have provided opportunities for the spread of Indo-European languages, perhaps related to the early Euphratic substrate preserved in early languages of Southern Mesopotamia.

<table>
<thead>
<tr>
<th>Population</th>
<th>Caucasus-Anatolian</th>
<th>South India</th>
<th>Baltic-Urals</th>
<th>Arabian</th>
<th>Mongolian</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arain Punjab Pakistan</td>
<td>43.6%</td>
<td>56.4%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Balochi Pakistan</td>
<td>61.9%</td>
<td>35.6%</td>
<td>-</td>
<td>1.2%</td>
<td>-</td>
<td>1.4%</td>
</tr>
<tr>
<td>Brahmin Uttar Pradesh India</td>
<td>22.9%</td>
<td>69.4%</td>
<td>4.5%</td>
<td>-</td>
<td>-</td>
<td>3.3%</td>
</tr>
<tr>
<td>Brahmin Uttaranchal India</td>
<td>15.4%</td>
<td>53.9%</td>
<td>-</td>
<td>-</td>
<td>13.5%</td>
<td>17.2%</td>
</tr>
<tr>
<td>Brahui Pakistan</td>
<td>59.8%</td>
<td>33.3%</td>
<td>-</td>
<td>4.6%</td>
<td>-</td>
<td>2.2%</td>
</tr>
<tr>
<td>Burusho Pakistan</td>
<td>40.6%</td>
<td>47.9%</td>
<td>3.6%</td>
<td>-</td>
<td>5.4%</td>
<td>2.6%</td>
</tr>
<tr>
<td>Gujarati India</td>
<td>19.3%</td>
<td>80.7%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kalash Pakistan</td>
<td>53.9%</td>
<td>39.6%</td>
<td>5.8%</td>
<td>-</td>
<td>-</td>
<td>0.7%</td>
</tr>
<tr>
<td>Kshatriya Uttar Pradesh India</td>
<td>22.3%</td>
<td>72.6%</td>
<td>1.8%</td>
<td>-</td>
<td>-</td>
<td>3.3%</td>
</tr>
<tr>
<td>Kurmi</td>
<td>17.5%</td>
<td>82.0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.6%</td>
</tr>
<tr>
<td>Makrani Pakistan</td>
<td>55.9%</td>
<td>30.7%</td>
<td>-</td>
<td>8.3%</td>
<td>-</td>
<td>5.1%</td>
</tr>
<tr>
<td>Meena India</td>
<td>28.6%</td>
<td>70.0%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.4%</td>
</tr>
<tr>
<td>Nepal</td>
<td>17.4%</td>
<td>53.8%</td>
<td>5.4%</td>
<td>-</td>
<td>6.1%</td>
<td>17.4%</td>
</tr>
<tr>
<td>Pashtun Pakistan</td>
<td>51.0%</td>
<td>44.6%</td>
<td>3.6%</td>
<td>-</td>
<td>-</td>
<td>0.8%</td>
</tr>
<tr>
<td>Sindh Pakistan</td>
<td>44.8%</td>
<td>52.7%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.5%</td>
</tr>
<tr>
<td>Tajik</td>
<td>57.2%</td>
<td>19.9%</td>
<td>10.5%</td>
<td>-</td>
<td>8.6%</td>
<td>3.7%</td>
</tr>
<tr>
<td>Turkmen</td>
<td>72.5%</td>
<td>12.6%</td>
<td>-</td>
<td>5.6%</td>
<td>-</td>
<td>9.3%</td>
</tr>
</tbody>
</table>

Table 3: Regional admixture in Indus Valley and neighboring populations (excluding Indus Valley admixture) based on autosomal SNP data. For SNP admixture components not excluding Indus Valley admixture, see http://dnatribes.com/dnatribes-snp-admixture-2012-03-12.pdf.

Several smaller genetic components were also identified. These included Baltic-Urals components in populations at the northern fringe of South Asia, for instance: Tajik (10.5%); Kalash (5.8%); and Nepal (5.4%). This might express links with Eurasian Steppe populations via Central Asia, such as possible Scythian related Shakya expansions between 800 and 200 BCE. Notably, these components were found in both western and eastern populations, suggesting the possibility of multiple migration routes (for instance, southward from the eastern Tarim Basin).
Similarly, Mongolian components were identified in several populations, such as: Brahmin Uttarakhand Pradesh (13.5%); Tajik (8.6%); and Nepal (6.1%). These components might express similar contacts with Asian Steppe populations via Central Asia, attested during the Iron Age as well as later periods (such as the Kushan and later Turkic expansions).

Another smaller genetic component found in some populations was Arabian, for instance Makrani (8.3%) and Brahui (4.6%). This might express maritime links, such as the Meluhhan trade contacts attested during the Harappan Bronze Age.

**Conclusion**

Both STR and SNP based analyses indicated substantial genetic links between the Indus Valley and both the interior of the Indian Subcontinent and West Asia. Archaeological evidence for a population expansion (possibly from West Asia) between 6,000 and 4,500 BCE might relate to genetic links with the Mesopotamian region (STR) and Caucasus-Anatolian region (SNP). Expansions of food producing cultures during this period might have provided an opportunity for the Indo-European languages (ancestral to Vedic Sanskrit) to reach the Indian Subcontinent.

In contrast, genetic links with Siberian populations were smaller. These included relatively small Altaian (STR), Baltic-Urals (SNP), and Mongolian (SNP) genetic components. These genetic links might express later and less extensive population expansions from the Eurasian Steppe and Central Asia, such as possible Shakya migrations during the Iron Age.

In addition, results also suggested genetic expansions from India to Central Asia. This included South India components identified in Kalash, Tajik, and Turkmen populations near the periphery of the Indus Valley region. These genetic links might express population expansions from South Asia, such as during the period of the Bronze Age Harappan Civilization. Future research might explore South Asian genetic links in more distant locations (such as the BMAC and Urals), where evidence for Vedic influences in material culture have been suggested by archaeologists.

In summary, results are consistent with emerging alternative models of South Asian prehistory, in which the Vedic cultures were descended from indigenous Harappans already resident in South Asia. Rather than a putative “Indo-European invasion” from Central Asia in the late Bronze Age, results suggest the possibility of an earlier and more peaceful “Indo-European diffusion” of food producing cultures from West Asia during the Copper Age.
We are pleased to announce a new update for DNA Tribes® SNP analysis:

New SNP populations: Several new populations have been incorporated in DNA Tribes® SNP:

New African populations: Somalia

New European populations: Basque Spain, Galicia Spain, Germany and Netherlands, Poland and West Slavic (mixed), Scandinavia, Western Scotland and Ireland

New Middle Eastern populations: Arab (Doha, Qatar)

New Modern Diasporic populations: African (Doha, Qatar), Canary Islands, Persian and South Asian (Doha, Qatar)

New South Asian populations: Bengali India, Bhunjia India, Brahmin Tamil Nadu India, Brahmin Uttar Pradesh India, Brahmin Uttaranchal India, Chamar India, Chenchu India, Dharkar India
New South Asian populations (continued):

- Dhurwa India
- Dusadh India
- Gond India
- Hakkipikki India
- Kanjar India
- Kol India
- Kshatriya Uttar Pradesh India
- Kurmi India
- Kurumba India
- Lambadi India
- Mawasi India
- Kol India
- Kshatriya Uttar Pradesh India
- Kurmi India
- Kurumba India
- Lambadi India
- Mawasi India
- Kol India
- Kshatriya Uttar Pradesh India
- Kurmi India
- Kurumba India
- Lambadi India
- Mawasi India
- Kol India
- Kshatriya Uttar Pradesh India
- Kurmi India
- Kurumba India
- Lambadi India
- Mawasi India
- Kol India
- Kshatriya Uttar Pradesh India
- Kurmi India

Enhanced Population Comparison: DNA Tribes® SNP now features a new enhanced Member Similarity comparison of your genotype to world populations. The new match algorithm allows smaller samples to appear higher in your population rankings, for a smoother listing of populations.

New World Regions: Regional admixture analysis now includes the Iberian and Northwest European regions (both formerly part of the Atlantic European region), for a total of 21 world regions. For a complete list of regions, see page 3 of any sample SNP report at http://dnatribes.com/snp.html.


New Sample Reports: Updated DNA Tribes® SNP reports for several world populations are available at http://dnatribes.com/snp.html. New SNP analysis orders (Sale Price $49.99) and updates to your personal DNA Tribes® SNP report can also be ordered at this link.