Introduction

Hello, and welcome to the November 2012 issue of DNA Tribes® Digest. The ancestral links between world populations uncovered based on new DNA evidence often cross perceived ethnic and cultural boundaries. In particular, autosomal STR and SNP based analysis of world populations reveals genetic continuity between Europe and the Near East, as well as more distant areas of Central Asia, the Indian Subcontinent, and West Siberia.

A linguistic trace of ancient links between these regions is used around the world today: the Indo-European languages. These include not only European languages (such as Spanish, English, and Portuguese), but also over 200 Asian languages (such as Hindi, Kurdish, and Tajik). Although details vary, most theories suggest a shared origin for Indo-European languages somewhere near the border of Europe and West Asia (possibly Anatolia, the Balkan Peninsula, or the Black Sea) and sometime since the Neolithic period (when farming and pastoralist cultures expanded from the Fertile Crescent).

For clues to how the Indo-European languages might have spread, this month’s article explores genetic evidence for two routes of West Eurasian expansions into Asia: (1) Southern expansions through West Asia reaching the Indian Subcontinent (possibly related to the larger satem group of Indo-European languages in Asia); and (2) Northern expansions through the Eurasian Steppe into Siberia and the Tarim Basin (possibly related to the extinct Tocharian languages).

Best regards,
Lucas Martin
DNA Tribes
Genetic Evidence for Two Indo-European Expansion Routes in Asia (STR and SNP)

Historical Background

The Indo-European languages are today spoken in many parts of West Asia, the Indian Subcontinent, and Central Asia (see Figure 1). Most of the Indo-European languages in Asia are Indo-Iranian languages classified as part of the satem group (shaded cyan in Figure 1) that also includes the Baltic and Slavic languages of Eastern Europe. These include present day Hindi, Kurdish, and Tajik, as well as ancient Avestan and Saka and the less known Kalash and Ossetic languages.

However, early texts discovered in the Tarim Basin of Western China¹ attest two separate languages once spoken even further from Europe: the Tocharian languages (Kuchean and Agnean), both classified as centum languages. Generally, centum languages (shaded magenta in Figure 1) are found in Western Europe (such as the Romance, Germanic, and Celtic languages) and near the Aegean Sea and Anatolia (such as Greek and ancient Hittite). This leaves us with a puzzle: what processes carried the Indo-Iranian satem languages into large areas of Asia, but deposited the unrelated Tocharian centum languages in Western China?

Figure 1: Map of Indo-European languages in Asia (prior to the modern period). Centum languages are shaded in magenta; Satem languages are shaded in cyan. Languages that are unclassified in terms of the centum-satem isogloss are not shaded. Modern languages are in white text. Ancient languages and substrates (traces of a language preserved in the vocabulary of an unrelated language) are in yellow italics.

¹ For more information, see The Tarim Mummies by J. P. Mallory and V. H. Mair (general information); and The Mummies of Urumchi by Elizabeth Barber (analysis of tartan and twill textiles). For more in depth analysis, see The Bronze Age & Early Iron Age Peoples of Eastern Central Asia (vol. I and II) edited by Victor H. Mair.
Most scholars agree that the Indo-European languages originated sometime since the Neolithic period and somewhere near Anatolia, the Balkan Peninsula, or the Black Sea. These hypothetical Proto-Indo-Europeans then expanded outwards in search of new fields to farm, new pastures for their livestock, and/or new ores to mine, eventually differentiating into various sub-cultures and spreading their languages to distant parts of Europe and Asia.

The exact dates and mechanisms of this spread remain theoretical, primarily because the early Indo-European cultures did not leave written records. Instead, the earliest traces of Indo-European languages appear as loan words recorded by literate societies of the Fertile Crescent (see Table 1): these include early “Euphratic” vocabulary preserved in Sumerian and Akkadian records, as well as Indic technical and religious vocabulary used in Hittite, Kassite, Hurrian, and Canaanite societies. These texts demonstrate that Indo-European languages were spoken near the Fertile Crescent during the Bronze Age and probably before 3,000 BCE (possibly transmitted by migratory specialists or early merchants with trade links to Central Asia or the Harappan Civilization).

<table>
<thead>
<tr>
<th>Language</th>
<th>Earliest Record</th>
<th>Period</th>
<th>Location</th>
<th>Early Texts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euphratic (substrate in Sumerian and Akkadian)</td>
<td>pre-3000 BCE</td>
<td>Bronze Age</td>
<td>Mesopotamia</td>
<td>Vocabulary in Sumerian and Akkadian languages.</td>
</tr>
<tr>
<td>Hittite</td>
<td>c. 1600 - 1200 BCE</td>
<td>Bronze Age</td>
<td>Anatolia</td>
<td>Anitta text.</td>
</tr>
<tr>
<td>Greek (Mycenaeae)</td>
<td>c. 1450 BCE</td>
<td>Bronze Age</td>
<td>Aegean</td>
<td>Linear B Tablets.</td>
</tr>
<tr>
<td>Indic (substrate in Hittite, Kassite, Hurrian, and Canaanite)</td>
<td>c. 1400 - 1300 BCE</td>
<td>Bronze Age</td>
<td>East Mediterranean; Mesopotamia</td>
<td>Kikkuli Text; Kassite-Akkadian vocabulary; Mitanni treaties; Amarna letters.</td>
</tr>
<tr>
<td>Celtic, Illyrian, Iranian, Italic, Phrygian, Venetic</td>
<td>c. 1000 - 500 BCE</td>
<td>Iron Age</td>
<td>(various)</td>
<td>(various)</td>
</tr>
<tr>
<td>Armenian, Germanic, Lusitanian</td>
<td>c. 500 - 1 BCE</td>
<td>Iron Age</td>
<td>(various)</td>
<td>(various)</td>
</tr>
<tr>
<td>Slavic</td>
<td>c. 500 CE</td>
<td>Medieval</td>
<td>Eastern Europe</td>
<td>(various)</td>
</tr>
<tr>
<td>Albanian, Baltic</td>
<td>c. 1500 - 2000 CE</td>
<td>Medieval</td>
<td>Eastern Europe</td>
<td>(various)</td>
</tr>
</tbody>
</table>

Table 1: List of early texts attesting Indo-European languages. Note that languages might be older than these discovered texts, but the dates and geographical locations of those unrecorded Indo-European languages are unknown. For more information, see The Oxford Introduction to Proto-Indo-European and the Proto-Indo-European World by J. P. Mallory and D. Q. Adams, p. 14.

Unfortunately, the archaeological record of early Indo-European texts documents the spread of writing from the Fertile Crescent better than it provides evidence of when and where the first Indo-European languages were spoken. Most of Europe and Asia outside the Fertile Crescent remain a blank during this early period.

However, some models taking into account linguistic variation between these languages have identified a possible Proto-Indo-European origin near Anatolia during the Neolithic period.² This suggests that Indo-European languages in Asia might have accompanied west-to-east demographic expansions of food producing farming and/or pastoralist (animal herding) populations from West Asia.

² For a recent analysis, see “Mapping the origin of Indo-European” at http://language.cs.auckland.ac.nz/.
To explore genetic evidence for these early expansions into Asia, this article will use both autosomal STR and SNP analysis to identify the West Eurasian regional components of the Indo-European speaking regions of Asia. To evaluate whether more than one expansion route separately transmitted satem and centum languages into Asia, the genetic analysis will include the Altaian (STR) and West Siberian (SNP) regions in which the unusual Tocharian languages were once spoken.

**STR Analysis of West Eurasian Components in Asia**

To identify possible geographical links associated with the spread of Indo-European languages in Asia, West Eurasian (European and East Mediterranean) genetic components to the Mesopotamian, North India, Eastern India, and Altaian regions were identified using autosomal STR data (excluding local admixture between these listed regions). Results are summarized in Table 2 and illustrated in Figure 2.

**Figure 2**: West Eurasian (European and East Mediterranean) STR components in the Mesopotamian, North India, Eastern India, and Altaian regions. Percentages from regions outside of West Eurasia are listed as “Other.”

**Discussion**: Results in Table 2 indicate multiple West Eurasian (European and East Mediterranean) genetic components in the studied Mesopotamian, North India, Eastern India, and Altaian regions. Indo-European languages have been spoken in all of the studied regions prior to the modern period (see Figure 1). However, no single western genetic component was shared between all studied regions (see Table 2).

Instead, the western genetic components in Asia are expressed according to two groups (“West Asian” and “Eurasian Steppe”), each suggesting a separate route of early population expansions or migrations. First, a “West Asian” group of genetic components from populations near the Mediterranean

---

3 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).
Sea (highlighted in orange and yellow in Table 2) characterizes the Mesopotamian, North India, and Eastern India regions.

Primary among these “West Asian” components was an Aegean component (related to populations living near the southern Balkan Peninsula). This Aegean component is largest for the Mesopotamian region (39.1%), but is also found in the neighboring North India (17.5%) and Eastern India (2.8%) regions. Accompanying this Aegean component is a secondary Italian component found in the Mesopotamian (2.0%) and North India (8.4%) regions.

The geographical distribution of this Aegean component suggests a pattern of expansions into West Asia from near the southern Balkan Peninsula (possibly via Anatolia), eventually reaching as far as the North India and Eastern India regions (where Indo-Iranian languages are spoken). This geographical pattern would be generally consistent with an origin of Indo-European languages among early Neolithic (farming and/or pastoralist) populations living near Anatolia or the Balkan Peninsula proposed by some linguists (see Historical Background).

However, an Aegean genetic component was not identified for the Altaian region that includes populations of Central Asia and Western Siberia, where the unusual Tocharian languages are attested by early texts discovered in the Tarim Basin. Instead of the “West Asian” genetic links with populations near the Mediterranean Sea, Altaian populations are characterized by a different group of “Eurasian Steppe” genetic links related to Northwest Europe (highlighted blue in Table 2). These included Celtic (12.6%) and Norse (10.8%) components, which are not identified for the Mesopotamian or Eastern India regions and only expressed as a smaller 4.6% Celtic component in North India.

This suggests that the unusual (Northwest European and Mediterranean like) Tocharian languages might have reached the Tarim Basin through a process that was separate from the general spread of satem Indo-European languages (including Indo-Iranian) into West Asia. More detailed analysis suggesting possible archaeological and linguistic periods associated with these separate language expansions is in the Conclusion section of this article.

An additional Urals component (related to indigenous Uralic speaking populations near the Ural Mountains) is also identified for several studied regions. This is highest in the Altaian region (11.9%), but also found in the Mesopotamian (4.6%) and North India (4.2%) regions. This might express genetic traces of pre-Indo-European expansions of Uralic speaking populations during the Mesolithic period, as well as later Indo-European population movements between Eastern Europe, West Siberia, Central Asia, and the Indian Subcontinent.4

<table>
<thead>
<tr>
<th>STR Region</th>
<th>Levantine</th>
<th>Aegean</th>
<th>Italian</th>
<th>Portuguese</th>
<th>Basque</th>
<th>Norse</th>
<th>Celtic</th>
<th>Scythian</th>
<th>Urals</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mesopotamian</td>
<td>29.7%</td>
<td>39.1%</td>
<td>2.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>4.6%</td>
<td>4.6%</td>
<td>24.6%</td>
</tr>
<tr>
<td>North India</td>
<td>0.0%</td>
<td>17.5%</td>
<td>8.4%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>4.6%</td>
<td>0.0%</td>
<td>4.2%</td>
<td>65.3%</td>
</tr>
<tr>
<td>Eastern India</td>
<td>0.0%</td>
<td>2.8%</td>
<td>0.0%</td>
<td>1.8%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>2.4%</td>
<td>0.0%</td>
<td>93.0%</td>
</tr>
<tr>
<td>Altaian</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>2.7%</td>
<td>10.8%</td>
<td>12.6%</td>
<td>0.0%</td>
<td>11.9%</td>
<td>62.0%</td>
</tr>
</tbody>
</table>

Table 2: West Eurasian (European and East Mediterranean) STR components in the Mesopotamian, North India, Eastern India, and Altaian regions. Percentages from regions outside of West Eurasia5 are listed as “Other.” The “West Asian” group of components (most typical of the Mesopotamian, North India, and Eastern India regions) is highlighted in orange; the “Eurasian Steppe” group of components (most characteristic of the Altaian region) is highlighted in blue.

---

4 For more discussion and genetic analysis, see http://dnatribes.com/dnatribes-digest-2012-10-01.pdf
5 Percentages listed under “Other” included East Asian, Native American, Sub-Saharan African, and South Indian genetic components of the studied Mesopotamian, North India, Eastern India, and Altaian regions.
SNP Analysis of West Eurasian Components in Asia

West Eurasian (European and East Mediterranean) genetic components in the Persian, Balochi, Indus Valley, North India, and West Siberian regions were identified using autosomal SNP data (excluding local admixture for these listed regions). Results are summarized in Table 3 and illustrated in Figure 3.

Discussion: Results in Table 3 indicate multiple West Eurasian (European and East Mediterranean) genetic components of the studied Persian, Balochi, Indus Valley, North India, and West Siberian regions. Indo-European languages have been spoken in all of the studied regions prior to the modern period (see Figure 1 in the Historical Background).

For all studied regions, the largest West Eurasian genetic component is North Caucasus. This is highest in the Balochi region (42.1%), which was the site of early agricultural communities of Mehrgarh and later was absorbed in the Bronze Age Harappan Civilization. North Caucasus components are also found at similar levels in the Persian (38.2%) and Indus Valley (37.4%) regions. Substantial North Caucasus components are also identified for the North India (21.3%) and West Siberian (15.1%) regions.

6 For more information about DNA Tribes® SNP analysis, see http://dnatribes.com/snp.html.
7 Percentages listed under “Other” included Arabian, Arctic, East Asian, East Siberian, Egyptian, Mesoamerican, South India, Tibetan, and West African genetic components of the studied regions.
8 The “North India” STR region differs from the “North India” SNP region: the SNP region is specific to northern parts of India and related Brahmin populations throughout India; in contrast, the more inclusive STR region encompasses populations of the “Balochi”, “Indus Valley,” and “North India” SNP regions.
Another West Eurasian genetic component found in all studied regions is East Mediterranean. This is highest in the Persian (30.4%) and Balochi (22.0%) regions, both including populations located near the urban centers and trade routes of ancient Mesopotamia.

Taken together, the North Caucasus and East Mediterranean components are found in all studied regions in which Indo-European languages were spoken prior to the modern period. This suggests population expansions from one or both of these regions (possibly associated with the spread of farming and pastoralism) might have carried the Indo-European languages eastward into Asia.

The East Mediterranean component is substantially smaller in the Indus Valley and North India regions. However, it is possible the Indo-European languages first spread through population expansions to an initial area in West Asia (such as the Persian and/or Balochi regions), and later came into use in neighboring areas (such as the Indus Valley and North India regions) through cultural diffusion and/or secondary local migrations within Asia.

A smaller Baltic-Urals component is found for the West Siberian (11.3%), Indus Valley (8.8%), and North India (3.4%) regions, but is not found in the Persian and Balochi regions. This suggests the Baltic-Urals component might not have been involved in the primary Indo-European expansions in Asia.

Instead, the Baltic-Urals component might express genetic traces of Uralic language expansions in Central Asia, West Siberia, and Northern Europe. These Uralic expansions probably took place during the Mesolithic period and pre-date the later Indo-European expansions. However, this Baltic-Urals component might also have been further transmitted by Indo-European speaking populations active in Central Asia and Siberia during later periods (possibly including Bronze Age Andronovo and Iron Age Saka cultures).

However, a Northwest European (5.5%) component is found in the West Siberian region that includes populations near the Tarim Basin, where the unusual Tocharian languages were spoken. This Northwest European component is not found in the other studied regions of Asia (except for a trace 0.6% percentage in the Persian region). This suggests the Northwest European component was not involved in general Indo-European expansions in Asia, and instead reached West Siberia Asia through a separate Eurasian Steppe migration (such as the early Afanasevo expansion to Siberia).

### Table 3: West Eurasian (European and East Mediterranean) SNP components in the Persian, Balochi, Indus Valley, North India, and West Siberian regions. Percentages from outside of West Eurasia are listed as “Other.”

<table>
<thead>
<tr>
<th>SNP Region</th>
<th>North Caucasus</th>
<th>East Mediterranean</th>
<th>Baltic Urals</th>
<th>Northwest European</th>
<th>Iberian</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persian</td>
<td>38.2%</td>
<td>30.4%</td>
<td>0.0%</td>
<td>0.6%</td>
<td>3.2%</td>
<td>27.5%</td>
</tr>
<tr>
<td>Balochi</td>
<td>42.1%</td>
<td>22.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>35.9%</td>
</tr>
<tr>
<td>Indus Valley</td>
<td>37.4%</td>
<td>8.4%</td>
<td>8.8%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>45.4%</td>
</tr>
<tr>
<td>North India</td>
<td>21.3%</td>
<td>3.0%</td>
<td>3.4%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>72.3%</td>
</tr>
<tr>
<td>West Siberian</td>
<td>15.1%</td>
<td>1.4%</td>
<td>11.3%</td>
<td>5.5%</td>
<td>0.0%</td>
<td>66.8%</td>
</tr>
</tbody>
</table>

10 Archaeological evidence suggests cultural interactions between the Indus Valley civilization, the Central Asian BMAC, and the Sintashta-Petrovka culture of the Eurasian Steppe. These interactions helped generate the Andronovo horizon, which might have provided a context in which Indo-Iranian, Uralic, Tocharian, and European IE languages interacted (possibly influencing the development of satem languages related to Proto-Balto-Slavic in Eastern Europe). See E. E. Kuzmina, *The Prehistory of the Silk Road*.  
11 Percentages listed under “Other” included Arabian, Arctic, East Asian, East Siberian, Egyptian, Mesoamerican, South India, Tibetan, and West African genetic components of the studied regions.
Conclusion: Two Routes of Indo-European Expansions in Asia

In summary, both STR and SNP analyses identified two groups of West Eurasian genetic components possibly related to Indo-European (IE) language expansions in Asia (listed in Table 4): (1) a West Asian IE related group of components (highlighted orange); and (2) a Eurasian Steppe IE group of related components (highlighted blue). Both STR and SNP data also identified possible genetic traces of earlier Uralic (pre-IE) language expansions (highlighted in purple). The map in Figure 5 illustrates early archaeological cultures that might have transmitted Indo-European languages via West Asia and the Eurasian Steppe.

<table>
<thead>
<tr>
<th>Population Data</th>
<th>Possible Route</th>
<th>Affected Regions</th>
<th>Genetic Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autosomal STR</td>
<td>West Asian IE</td>
<td>Mesopotamian, North India, Eastern India</td>
<td>Aegean</td>
</tr>
<tr>
<td>Autosomal SNP</td>
<td>West Asian IE</td>
<td>Persian, Balochi, Indus Valley, North India</td>
<td>North Caucasus, East Mediterranean</td>
</tr>
<tr>
<td>Autosomal STR</td>
<td>Eurasian Steppe IE</td>
<td>Altaian</td>
<td>Celtic, Norse</td>
</tr>
<tr>
<td>Autosomal SNP</td>
<td>Eurasian Steppe IE</td>
<td>West Siberian</td>
<td>North Caucasus, N.W. European</td>
</tr>
<tr>
<td>Autosomal STR</td>
<td>Uralic (partly pre-IE)</td>
<td>Altaian</td>
<td>Urals</td>
</tr>
<tr>
<td>Autosomal SNP</td>
<td>Uralic (partly pre-IE)</td>
<td>West Siberian, Indus Valley, North India</td>
<td>Baltic-Urals</td>
</tr>
</tbody>
</table>

Table 4: Summary of West Eurasian genetic components in Asia identified by autosomal STR and SNP analysis

West Asian IE (Southern Route): Both STR and SNP data identified genetic components that might have accompanied Indo-European language expansions through West Asia (possibly from near Anatolia). These included Aegean (STR) and North Caucasus and East Mediterranean (SNP) genetic components (orange in Table 4).

These “West Asian” components were shared by most studied regions and might express genetic traces of early agricultural expansions (possibly through Anatolia) that transmitted ancestral Indo-Iranian languages to South Asia. The exact period in which these expansions took place is unknown; however, potential archaeological periods include Mehrgarh III (Togau Phase) (4,300 – 3,800 BCE)12; Mature Harappan (2,600 BCE)13; or the Bactria–Margiana Archaeological Complex (BMAC 2,300 BCE).

Linguistic models have typically suggested a relatively recent transmission of Indo-European languages to South Asia via the Eurasian Steppe (for instance, proposing that the Proto-Indo-Iranian language originated in the Andronovo culture). More recent research has questioned this northern steppe invasion model,14 but has not outlined an accepted new model of when and how Indo-European languages spread to West Asia, Central Asia, and the Indian Subcontinent. Based on the geographical distribution of these genetic components, this article’s genetic analysis suggests instead a southerly West

12 An influx of new collective burial customs, ceramic styles, and copper technology appeared during the Mehrgarh III (Togau) period (probably from West Asia). See G. Possehl, The Indus Civilization pp. 34-35.
13 The Mature Harappan involved extensive founding of new settlements in South Asia and introduced new and different technologies, including: bronze, baked bricks, mastery of sailing, faience, hydrological technology, stamp seals, and writing. See G. Possehl, The Indus Civilization pp. 51-52. The R-37 population was distinguished from other Harappans by its West Asian characteristics (similar to Mehrgarh III). Ibid., p. 175.
Asian expansion of Indo-European languages in Asia related to early agricultural populations expanding from near Anatolia.

**Eurasian Steppe IE (Northern Route):** Both STR and SNP data identified genetic components that might have accompanied separate Indo-European language expansions through the Eurasian Steppe. These included Celtic and Norse (STR) and North Caucasus and Northwest European (SNP) genetic components (highlighted in blue in Table 4). These components were not found in most studied regions (except for a small percentage of a Celtic like component in the North India STR region), which suggests a separate path of Indo-European migrations to Siberia.

This apparently separate group of West Eurasian genetic components might correspond to similarly separate archaeological cultures and languages in Siberia and the Tarim Basin: the Afanasevo archaeological culture and the Tocharian languages (see Figure 5). Afanasevo (approximately 3,500-2,500 BCE) was an early Copper Age culture that separated from early Yamna cultures, around the time when agricultural settlements near the Balkan Peninsula were dispersing due to climate change.\(^\text{15}\)

In the process of expanding eastwards, Afanasevo was influenced by cultures living near the North Caucasus (possibly Hurro-Urartian or Alarodian speaking populations) and possibly transmitted the Tocharian languages to the Tarim Basin prior to the (separate) expansion of the Andronovo culture and Iranian (Saka) languages.\(^\text{16}\) Afanasevo links with Europe and the North Caucasus might in part correspond to the Northwest European and North Caucasus SNP components in West Siberia.

For STR analysis, comparable North Caucasus population data are not available for a direct comparison to SNP results. Instead, West Eurasian genetic links in the Altaian region are expressed as Celtic and Norse sub-regional components (both sub-regions within Northwest Europe). These long distance genetic links do not necessarily suggest culturally “Celtic” or “Norse” migrations in Asia, but might express traces of shared ancestry from ancient populations (once living near the North Caucasus) preserved in geographically peripheral locations (the British Isles and Central Asia).\(^\text{17}\)

The Tocharian and (probably related) Yuezhi and Kushan cultures were eventually absorbed by Indo-Iranian and Turkic (Uyghur) populations.\(^\text{18}\) Nevertheless, their signature might persist as the North Caucasus and Northwest European genetic components still found in Central Asia and West Siberia.

**Uralic (partly pre-IE):** Both STR and SNP analysis identified Urals (STR) and Baltic-Urals (SNP) genetic components in Western Siberia. These components might in part express traces of Uralic expansions during the Mesolithic period, predating the Indo-European expansions described in this article. However, it is possible that Indo-European cultures near the Eurasian Steppe also helped transmit Urals and Baltic-Urals components between Eastern Europe, Central Asia, and the Indian Subcontinent.

---


\(^\text{16}\) Afanasevo material culture was distinct from both Andronovo and Karasuk and possibly involved different population movements. However, Tocharian speakers later adopted (Indic) Sanskrit personal names, and related Kushan cultures played a key role in the Silk Road transmission of Buddhism. Notably, this Tocharian-Kushan adoption of Indian cultural ideas paralleled the Kassite transmission of Indic deity names (such as Surias) to Mesopotamia. Comp. Kassite *Suqamuna*, Sanskrit *Sakyamuni*.

\(^\text{17}\) Notably, Celtic (British Isles) populations are associated with slightly higher levels of the North Caucasus SNP component than other populations of Northwest Europe (possibly expressing traces of Proto-Celtic migrations to Western Europe from near the Black Sea). See [http://dnatribes.com/dnatribes-digest-2012-08-01.pdf](http://dnatribes.com/dnatribes-digest-2012-08-01.pdf).

\(^\text{18}\) Although no specifically Tocharian link is known, medieval writers noted a non-Turkic non-Persian language “unlike any other human tongue” still spoken near Central Asia. See R. Frye, *Ibn Fadlan’s Journey to Russia* p.74.
Figure 5: Possible Indo-European expansions in Asia along two routes (West Asian and Eurasian Steppe).
Update: New STR Populations for November 2012

We are pleased to announce several new STR populations added to our database:

**New African populations:**
- Arabic speakers (Doukkala, Morocco) (71)
- Bantu (South Africa) (Zulu, South Sotho, and Tswana) (50)
- Berber (Khenifra, Morocco) (75)
- Herero (Namibia) (13)
- Khoe (Nama) (Namibia) (26)
- San (Namibia and South Africa) (138)

**New East Asian populations:**
- Hunan, China (586)
- Korea (197)
- Uyghur (Kashgar, Western China) (326)

**New European populations:**
- Austria (131)
- Austria (132)
- Belgium (206)
- Czech Republic (200)
- Denmark (199)
- Finland (171)
- Finland (210)
- France (208)
- Germany (331)
- Germany (331)
- Greece (208)
- Hungarian (Vojvodina, Serbia) (291)
- Hungary (223)
- Italy (151)
- Italy (152)
- Italy (464)
- Lodz, Poland (400)
- Montenegro (200)
- Norway (200)
- Poland (205)
- Republic of Ireland (304)
- Slovakia (246)
- Slovenia (207)
- Spain (284)
- Spain (284)
New South Asian populations:
- Maldives (301)

New Global Diaspora populations:
- Angolan (Lisbon) (480)
- Afro-Caribbean (UK) (309)
- Brazilians (Lisbon, Portugal) (207)
- Cabo Verde (213)
- Cape Coloured (Mixed) (South Africa) (97)
- Mestizo (Mexico) (1640)
- South Asian (UK) (373)

Updates: STR Updates to incorporate these new populations are available using the $24.99 “Update Your Analysis” option through our secure checkout system at http://dnatribes.com/order_addons.html.