*J Biosci* (2019)44:58
DOI: 10.1007/s12038-019-9881-7

Indian Academy of Sciences

# Early 'Aryans' and their neighbors outside and inside India

M Witzel

*Department of South Asian Studies, Harvard University, Cambridge, MA 02138, USA*

*(Email, witzel@fas.harvard.edu)*

Data from archaeology, linguistics, population genetics, and from early Vedic texts, which deal with religion, mythology and rituals, have to be assembled and closely compared in order to gain a comprehensive picture of the early 'Aryans'. Such interdisciplinary dialogue is necessary in order to establish areas of overlap of data. This paper attempts to indicate a western Central Asian origin of the Indo-Aryan speakers, in the steppe belt near the Urals, from where they moved, via the Inner Asian Mountain belt and Bactria, into India. Their gradual migration entailed acculturation with previous populations, their languages and cultures.

**Keywords.** Aryans; Indo-Aryans, migration; Ural steppe belt; Inner Asian Mountain belt; interdisciplinary comparison; Veda; ancient DNA

Over the past year, newspapers and interviews have bombarded us with headlines such as 'The new Indus–Saraswati excavations show no Aryan[1] invasion' or the like. Apart from the usual journalistic hype, such articles evidenced a remarkable lack of balanced, multidisciplinary stance, in some cases even by the archaeologists, the excavators themselves.

However, any serious understanding of the early history of human settlement in the Indian subcontinent has to be multidisciplinary, and evolve in constant dialogue with a number of relevant disciplines.

It is absolutely necessary that we can actually understand how far from certainty are some of the conclusions of some of our colleagues in the sciences – or not. All the various results from the humanities and the sciences have to agree with each other. *If* there is an outlier in data – that sticks out like a proverbial sore thumb – something must be wrong with the data and must be explained within reason.

In general, recent advances in a number of scientific fields and in the humanities have greatly enlarged and enhanced our view of the oldest and archaic India.

In the mid- to the late-nineteenth century, before Indian (pre)history was rediscovered, such as with Asoka's inscriptions, our view of archaic and ancient history was restricted to some ancient monuments and South Indian menhir graves, to some linguistic data that linked India with Central Asia and Europe (Proto-Indo-European) and to its related mythology; and later on, to some anthropology, focusing, as it was the trend then, on body shapes, etc. All of this has resulted in a scenario of an earlier Dravidian occupation in India, overlaid by an Indo-Aryan. Ethnicity ('race') was then often confused with linguistic affiliation, although the early Vedic pioneer Max Müller protested against such simplifications already in the 1870s.

Currently, our view of archaeology, understanding of ancient texts, of developments in comparative linguistics (Indo-European, Dravidian, Munda, etc.), and most recently, of human genetics – especially that based on ancient DNA – have enormously widened our horizons.

I will pick up and comment on some such developments, especially those in various scientific fields when their results mutually overlap. As mentioned, a truly multidisciplinary approach is needed.

We have practiced this at my University since 1999 in our yearly roundtables that involved textual scholars, archeologists, indologists, and later on also geneticists, aiming at understanding each other's research problems, at a view of how colleagues have reached their conclusions. For a long time, a textual scholar did just take over the results and certain aspects, say, of genetics. However, we did not know *how* such scholars had come to their conclusions.

## 1. Important sources of early history

1. Archaeology has set a certain framework for early history, especially from the beginning of agriculture; though unlike Egypt, in South Asia we cannot rely on an accumulation of data from architecture, art, texts, and religion, some of which are largely missing for early periods.

Further, *relative* dating had originally been fairly vague, relying on the layering of pottery styles, but there have been major recent advances in $C_{14}$ dating of organic materials

**http://www.ias.ac.in/jbiosci**
Published online: 05 July 2019

enhanced by dendrochronology, thermo-luminescence, the study of ice cores, etc., that allow for much more precise *absolute* dates.

For example, the new data relating to climate change by Gayatri Kathayat *et al*., last year, that were retrieved from the ice cores in Himalayan caves have helped to correlate climate change with the dates of major changes in civilizations, such as the Harappan civilization, the early Vedic period, etc. Importantly, they blend well with the textual data.

Otherwise, regarding absolute dates, there is no direct connection with extra-Indian dates – except for the Indus trade with Sumeria.

2. Linguistics supplies elaborate data as preserved in texts and in *reconstructed* languages. They are interlinked to some extent, with dated near-eastern texts. Linguistics, in spite of some recent partisan sniping, is a natural science as far as the development of sounds is concerned; they can be measured by movements in the mouth and by the MHz frequency of sounds. Good examples are the many loanwords in Vedic Sanskrit of substrate languages (see below) and the early form of Vedic attested in the Mitanni Kingdom of northern Mesopotamia around 1500 BCE.

3. The early texts themselves, usually in Vedic Sanskrit (while the so-called Indus inscriptions cannot be read – they may not even be recorded language but symbols such as the Mesopotamian Kudurru) of the ensuing periods (1300 sqq) present a special, unique situation: the Vedic texts, notably the earliest (Ṛgveda) are virtual tape recording as text, not even a syllable or tonal accent, were allowed to be changed in transmission (until today!). In addition, the ensuing Vedic texts were codified after composition: thus we have a clearly levelled, relative, though not an absolutely, dated set of texts down to the lifetime of the Buddha (revised date ca. 460–380 BCE). Its beginning can be linked with written and dated near eastern (Mitanni) texts.

This allows, although the texts are religious in content, to establish a clear line of development of not just religious and ritual data but also – usually mentioned as asides – of those relating to society, ethnicities and even polities. For example, the spread of the Ṛgvedic set up, in its Kuru version, can be followed up to the borders of Bengal and the Vidhyas in the South.

Importantly, no evidence of the Mature Indus civilization with cities (2600–1900 BCE) is seen in the Vedic texts: there are no cities, no international trade, no seals, none of the major Harappan religious features. The Vedic texts do not correlate with or overlap at all with the Harappan civilization, one of the several reasons to date the Vedas much later than 1900 BCE.

4. Paleontology and genetics, especially ancient DNA, add important facts to this emerging picture. While paleontology has been used and misused in the late nineteenth and early twentieth centuries, often with sweeping generalized results,[2] population DNA by now has been used for some 20 years, originally based on modern Indian data. This has delivered some general scenarios of ancient North Indian and ancient South Indian populations.

However, there was a 'revolution' over the last 10 years or so, once ancient DNA could be extracted and studied. The increasing volume of aDNA data has revolutionized our view of the prehistory of Africa, Europe, and the Americas, and also for the South Asian subcontinent. This is in need of some elaboration. Regarding the maternal side, we have the typical South Asian DNA, which was brought by the Out of Africa movement: it is still prominent as haplogroups M and N (and R). The male side is more complicated. The most interesting haplogroup on the Y chromosome, in the current context, is R1a1, a typical Central Asian one and Western Central Asian one.

More recently, we have gained access to ancient genetic evidence (aDNA), which as in Europe and Africa, allows for a much more detailed genetic history. Such aDNA is found inside bones, especially in the Petrous bones of the inner ear, which often is well preserved.

Now, the prepublication paper (2018)[3] has established the first ancient DNA from South Asia, found in the Swat Valley – which is mentioned, by the way, in the Ṛgveda as a 'good place'. For the first time, we can see who was around, at 1250 BCE in that area; that is, as expected by linguists and textual scholars, not South Asian but predominantly steppe aDNA, from 1250 BCE.

For more than a year, there have been press reports and rumors about another paper resulting from the V Shinde excavation of the Harappan city of Rakhigarhi in Haryana; these reports say that they could sequence just one skeleton. Other skeletons, also at other graveyards in Haryana, etc., were of no use for aDNA due to contamination. In the present case, the aDNA reportedly does not show the Central Asian haplogroup R1a1: of course not, as a Harappan specimen is just too early.

Unfortunately, as mentioned, this result was turned into a political one. Immediately, some newspapers broadcast: 'no Central Asians' in the Indus Valley; and: 'no Aryan invasion'. Obviously, what we expect is South Asian DNA. That is also what the excavators and their geneticist say: they add that there was some Iranian influence, which is not surprising, given the introduction of 'Iranian' agriculture. Even the name for 'wheat' (*godhūma*) is a Sanskritized near eastern word (see below).

In the sequel, I will not deal with these branches of sciences separately, rather I will draw their results together, and describe them in historical fashion. All of this is an ongoing project: archaeological finds (aDNA) genetic research, even textual and linguistic data, are constantly in a flux; we get closer to the solution, but we are not there yet.

## 2. First anatomically modern humans, 50 kya

After the exodus from Africa around 65,000 years ago (65 kya), there are few traces of early anatomically modern humans (AMH) in India: there are some stray finds of AMH tools, as some contend, already before the 74 kya explosion

of the Toba volcano. AMH seem to have been present in India before and after that, although the Toba volcano ashes covered large parts of SE Asia and India.

Other human remnants are under the waters of the Indian Ocean, as the sea level around 60–50 kya was about 100 m lower than today. There is some evidence for AMH with rock paintings in the Narmada corridor, but they are much younger than 50 kya. By about 30 kya we have some occasional early data from Chitral[4] and Sri Lanka.

Finally, modern DNA provides some indications of the early, ex-Africa settlement: the female mtDNA haplogroups M and N, derived from the African $L_3$, are still very prominent in India.

During this long period, many individual tribes and languages will have developed: clear remnants are the Andamanese pygmies and their (almost extinct) languages: Onge, Jarawa, and Sentinelese, as well an early Australian substrate layer in Tamil.[5] The ex-Africa migrants, who had reached Australia already by 50 kya, necessarily must have traveled through India. The intermittent SE Asian languages have not even been touched by such investigations, and as for India, very little has been done in this regard.

Thus, the question of the 'original' Indian language(s) has to remain open until we have better reconstructions of the various current language families (Dravidian, Munda, etc.), combined with the currently visible (as well as the still to be investigated) substrates: what lies below the IA languages Sindhi and Punjabi? Further, a widespread northern Indian 'Language X' seems to have existed in the Gangetic Plains[6] and well beyond (see below).

For this early period, we obviously do not have any texts (unless reconstructed for mythology),[7] but just some rock paintings. However, they all are difficult to date if their creators have used natural mineral colors that do not contain animal fat, etc. (which can be dated by the $C_{14}$ dating method).

Eventually, future finds of aDNA may help, though they have not been recovered so far, due to bad preservation in the Indian Monsoon climate (see below on Swat and Rakhigarhi). The best chance for aDNA recovery therefore lies in the western hills, like Swat or Akra, not in the Indus or UP plains, or in South India. We have to wait and see.

Incidentally, paleontological data are generally too vague.[8] Also, body shape shifts fairly quickly according to climate and nutrition (as can easily be observed in medieval Europe or post-World War II Japan).

## 3. Agriculture

The introduction of agriculture some 8000 years ago brings about major changes. This takes place from the near-east and Iran, around 6000 BCE, and is first attested in Mehrgarh, SE Baluchistan.[9] It followed up on a pastoral culture with caprids, this too based on a clear near eastern impact. Sheep and goats have been domesticated first in the eastern Mesopotamian mountains. The impact of agriculture likewise proceeded via Iran: actually, it took some 2000 years before wheat was acculturized to the Indian climate.[10] Even the Sanskrit word for wheat is of near-eastern/Iranian descent: Sanskrit *godhūma* (with folk etymology 'cow smoke'!), via Old Iranian *gantuma*, comes from near eastern *xand*, etc.[11] (However, the Zebu is a local Indian domesticate.)

The near-eastern influence is supported by some modern DNA (L, J2, etc., with western origins), and now even by some aDNA from the western rims of the Indus area (Shahr-i Sokhta on the E. Iran/Afghan border).[12]

This heavy 'Iranian' influence on the Indus area stands in remarkable contrast with the development of *indigenous* agriculture in the eastern Gangetic plains such as mung beans and legumes, and later on (2500–2000 BCE) also rice hybrids (along with the indigenous water buffalo), whose names are neither IA, nor Dravidian, nor Munda. Instead, they stem from a lost northern Indian substrate, called 'Language X' by Masica.[13] Incidentally it seems to have been much more widespread than that appears in Masica's original study of the Hindi substrate. Such words are found all the way to Maharashtra and western Punjab; this needs a more detailed study. Unfortunately, there is no enthusiasm in India to carry out such substrate studies: following the traditional medieval Sanskrit: Prakrit: Deśī classification, all non-IA words are clubbed together as '*deśī*'.

Similarly, the Deccan and South India developed their own local brand of agriculture, focusing on millet (much later also on rice). There are many varieties of local and 'foreign' millet that have moved in both directions, including an early variety from east Africa.[14] This type of agriculture existed next to extensive cattle herding, which has left huge ash mounds and menhir-like graves. The language of these people before (reconstructed) Dravidians appeared is uncertain. There must have been several substrates, such as those found in Tamil or in the Nilgiris.[15]

A word must be added about some claims for early Indian agriculture (such as a few grains at Lahuradeva, UP), or domesticated rice in eastern UP (wrong identification of domestic rice instead of spikelets, etc.). In fact, it has been well known for some 15 years that Indica rice is a hybrid of *Oryza japonica* (stemming from southern China) and local wild rice (*Oryza nivāra*).[16] Instead, domesticated rice emerged in India only by 2500–2000 BCE. (Similar spurious claims are made for early smelted iron in eastern UP, usually by the same archaeologists.) Such 'outliers' must not be taken on faith but must always be taken with so many grains of salt.

The three major Indian agricultural areas do not readily overlap. Likely, they also developed separate languages and cultures. For example, Masica's 'Language X' is quite different from the Ṛgvedic substrate;[17] for the presumably Dravidian-speaking Deccan; see Southworth[18] (2006).

As always, language is a good yardstick. Its substrates often reveal multiple layers, such with Nahali (Central

India): below the current Indo-Aryan speech we find, subsequently, layers of Dravidian, Munda, and some 25% of 'original' Indian languages. Such language change can be recent: for example, the Mushahar 'mice eaters' in UP, like the Nahali, speak Indo-Aryan now but were Munda speaking in the recent past. Genetic study might reveal their relationship with other Munda-speaking tribes in Central or East India. Likewise, the Tharu in the sub-Himalayan belt of India and Nepal speak Indo-Aryan now but are culturally different from their immediate neighbors, and have acquired near-immunity against malaria. I have detected (but not published) a substratum in their current language.

When it comes to attested language change, the current Newar population of the Kathmandu Valley is well represented with place and personal names in inscriptions since ca. 400 CE (probably indicating a Kirāta substrate); brief, actual texts start in 990 CE. These two Tibeto-Burmese languages definitely indicate the ancestral speech of the modern Newars, although a number of migrations from the north and south have occurred since then, as is now actually seen in genetic studies.

The Kirātas are mentioned already in the Atharvaveda (in the western Himalayas), then in some inscriptions and medieval texts of the Kathmandu Valley, and now survive in eastern Nepal (as Rai and Limbu tribes).

Taking a larger view on remnant and substrate languages we have: Burushaski in northernmost Pakistan, a Kashmiri substrate, a northern Indus substrate (found in the Ṛgveda), the Tharu, the highly endangered Kusunda (C. Nepal), the Rajbanshi (Nepal/Bengal border), the Mushahar, the Gangetic (agricultural) substrate (non-Dravidian, non-Munda), the central Indian Nahali, that of several Nilgiri populations (Toda, Irula etc.), the Australian substrate in Tamil,[19] the remnants of the Vedda language in Sri Lanka, and the endangered languages of the Andaman populations. To access the prehistory of the Harappan civilization, we badly need a study of the Sindhi and Punjabi substrate, as well as that of the Bhili substrate, but that is not in sight.[20] An incipient dictionary of Indian substrate words can be found online.[21]

## 4. The Harappan civilization

Against this background, highlighting the diversity that has always been a hallmark of Indian history, linguistic and otherwise, we can now take a look at the first major Indian civilization. It has often been called, for short, the Indus civilization (2600–1900 BCE),[22] but its technical, archaeological name is Harappan, after its first-find spot, Harappa in Punjab.[23] The currently propagated name 'Sindhu-Sarasvati' merely is a political maneuver, emphasizing the late Harappan settlements in the Ghagghar-Hakra (Sarasvatī) valley and underlining the supposedly Vedic nature of the Harappan civilization.[24]

The establishment of the Harappan civilization created a clear break with the preceding small village cultures. The new ones are: city planning, canalization, water management, development of certain crafts, international trade with Mesopotamia, and use of seals and tablets, the meaning of which remain unexplained.[25] Other typical enigma are: the lack of temples and palaces that are so typical for the contemporaneous near-east. Apparently, the Harappan civilization was not dominated by a royal/priestly class, although separate 'citadels' exist next to the general living quarters of the towns. Their exact function remains unclear.[26] Also, the existence or mutual relation of the handful of the more or less evenly distributed Harappan city 'states' is unclear: Harappa, Mohenjo-daro, Rakhigarhi, Kalibangan, Ganweriwala, Dholavira, etc.

A major question relates to the water management by this culture: it has recently been alleged that the major center of the Harappan civilization was situated in the area of the Ghagghar-Hakra river (the later Vedic Sarasvatī). However, it is clear now that this river was not perennial even during the heyday of the Harappan civilization:[27] it does not carry Himalayan isotopes that would point to a glacier-fed, perennial river. Instead, its wide *post*-Ice Age river bed was only partially and temporally filled by monsoon rains.[28] The rains were on the wane even during the Mature Harappan civilization.[29] Then, how could the Harappan civilization flourish? The answer lies in the many oxbow lakes of the major Punjab rivers[30] and in the evidence of the very common Harappan seals depicting 'water carriers' (two vessels balanced on a carrying pole). Clearly, given these ubiquitous representations, there must have been a pressing need for such individual irrigation.[31]

As far as Harappan religion and mythology are concerned, we have some 5000 'seals' and oblong tablets, the latter with obvious mythological themes on one of their sides. But, the nature of these myths is not evident at all. Clearly, it is not reflected in the succeeding Vedic texts[32] and even in later texts: who is the tiger goddess or the person in the tree? The so-called 'yogic' figure merely represents a normal Asian sitting posture and the deity(?) as such is found similar to the Celtic 'horned one' (Cernunnos) all over northern Eurasia, for example at Gundestrup in west Denmark. Śiva is not involved. Also, the ahistorical link of a man fighting a buffalo with the millennia-later Hindu deity Mahiṣāsuramardinī is a fantasy. Even the typical 'Indian' *namaste* gesture seen in some clay figures is found in distant Bactria (BMAC) and in the contemporaneous Jōmon period of Japan.[33]

All these items do *not* figure in the succeeding Vedic culture that only later on (ca. 1200–400 BCE) existed in the northern area of the Indus civilization of the greater Punjab. At best, one can point to the custom of putting *sindura* color in the hair parting line; and some items of 'low level' religion (Piśāca, Kimīdin, Cumuri, etc.) have survived the influx of the IAs. But there are no Indus 'high gods' or bull-fighting or tiger myths. This restriction extends to material culture: only some traditional features such as the width of the axle of carts, production of sea shell bangles and beads, traditional shapes of pottery have survived. Such survival,

however, is not surprising: it is common everywhere during the change to a new culture – unless there was a total replacement of the population – which definitely did not take place in the post-Harappan case (see below).

The above discussion was necessary as to clearly distinguish the Harappan civilization from what followed, the Vedic civilization. It existed in partially the same area (Greater Punjab), but was clearly later than the former.

## 5. End of the Harappan civiliation

The dissolution and demise of the great Harappan civilization has been much debated, but it has recently become increasingly clear that the impact of the climate shift around 2000 BCE also affected the Middle East, Oman, Iran, etc.: that means, already the Mature Harappan civilization, which survived a bit longer, until 1900 BCE, was due to good water management.[34]

Even then, the climate deterioration had led to a massive movement of Indus populations eastwards into the Haryana/Delhi area (and a similar situation occurred from the southern Harappan civilization in Sindh towards western Gujarat). At the same time, the eastward move by the agricultural populations left many wide open spaces in the Greater Punjab, which pastoralists could now exploit. This is exactly what seems to have happened at the end of the Mature Harappan civilization: the Indo-Aryans, coming down from the Afghan mountains (and beyond), occupied Gandhāra and the Greater Punjab, though not without resistance by some of the remaining Indus people.[35] They were constantly looking for 'wide open spaces' (*urugavyūti*) with good access to water for their cattle.

This process has not yet been properly understood: it certainly was *not* an 'invasion' with killing/replacement of the Indus people, as it is frequently mischaracterized by some politically motivated writers.

In fact, there were some substantial remnants of Indus people (along river banks!) that did interact with the IAs. This is clearly seen in the many loanwords taken over even into the IAs' sacred text, the Ṛgveda.[36] Incidentally, almost none of them have a Dravidian etymology[37] and hardly a Munda or Burushaski one.[38]

The archaeology of this period is little known,[39] but it may still deliver surprises: the recently discovered items at Sinauli (just east of Delhi) include carts (not 'chariots'),[40] and copper inlaid coffins and swords, that clearly belong to the late Bronze Age (thus before 1000 BCE). This find may point to the survival of an extra-Harappan organized society. However, we need better dating (going beyond the Late Bronze Age label).

It is still unclear what kind of cultures precede/overlapped here[41] with the Harappan civilization and the newly arrived Vedic culture. Again, they were probably different from those of lower Gangetic area with its Language X, which does reflect a different one than the Ṛgvedic substrate. We can, however, deduce some of the languages spoken east of the Punjab from data in the late Ṛgveda (book 10) and from those in the succeeding Yajurveda Saṃhitās.[42] They are easily identified in these texts (since ca. 1000 BCE). However, the extent of their contributions is not easily seen in archaeology.[43] Again, the eventual discovery of ancient DNA will help. So far it is available only for Swat; the Rakhigarhi data of the Harappan period are non-conclusive (see above).

## 6. Origins of the Indo-Aryans

The ultimate origins of the population speaking early IA (pre-Ṛgvedic) is to be located near the Ural mountains. The main pastoral people of the area (Arkhaim, Sintashta) were in close contact with the early Uralic speakers in the Russian woodland (*taiga*) belt: many dozens of early IA loanwords into the various Uralic languages attest to this, sometimes borrowed even with the nominative *–s*, as in *pakas* 'Bhaga'; they cover IA mythology (*asura*, etc.), people (*orya* 'Aryan slave'!), as well as many items of material culture.[44]

Importantly, even some aspects of typical IA social culture have turned up in the archaeology of the Ural area, such as the Vrātya young men association (*Männerbund*) that celebrated a 'dog killing' (*śvagn-in*) ritual in winter. Remnants of it have recently been discovered at Krasno-Samarskoye, just west of the Urals.[45] In the same area, some 40 years ago, a grave has been found with a headless body; instead a horse head had been substituted, just as in the Ṛgvedic Dadhyañc myth.[46]

We cannot follow up all the stages of the migration southwards of these early IA speakers; let it just be mentioned that there was a major impact on IA by the (unknown) language of the Bactria-Margiana Archaeological Complex (BMAC). It included even major Ṛgvedic gods such as Indra, Śarva (Rudra), Gandharva, etc.[47]

However, a major milestone is the appearance of early IA speakers in the near east, on the northern rims of Mesopotamia. They have left many loanwords in a non-related Caucasus language, Mitanni Hurrite, from ca. 1600 BCE onwards. The Mitanni kingdom covered northern Iraq and Syria; IA speakers will have reached it from its eastern border in the Zagros mountains.

The loanwords extend from horse culture (colors, number of rounds in chariot races), to the throne names of Mitanni kings, and a brief list of four Vedic deities (Mitra, Varuṇa, Aryaman, Nāsatya) attached to a treaty with the Hittites (1380 BCE).

Importantly, the early IA language of the Mitanni texts is older than that of the RV:[48] summarizing the technical points, the sounds *-az-* in the middle of a word have changed to *–e-* in the Rigveda. Thus, a word like Mitanni *mazdha* – later famous in Iran as the deity Ahura Mazda – changed to *medha* in Vedic and Sanskrit. The same early change is also seen in verb forms, like *sazdai* > Vedic *sede*.[49] This pre-Vedic language can be dated at around 1400 BCE if not

slightly earlier. This allows to date the Ṛgveda itself: it must have been composed after ca. 1400 BCE but before the onset of the Iron Age in the Greater Punjab area and at Akra, in Bannu (ca. 1000 BCE).[50]

As the geneticists tell us now, the Indo-Aryan immigration into the Swat area (the Ṛgvedic *su-vāstu* 'good land') is dated around 1250 BCE. Thus, there is a fairly narrow time span between roughly 1400 and 1000 BCE when to place the Rigveda.

In the present instance, archaeology (Gandharva grave culture, Akra), linguistics (pre-Ṛgvedic), texts (Śambara's fortresses in the mountains), and genetics (Swat aDNA) overlap perfectly, putting to lie the fantastic dating of the RV at 3000 BCE and its 'complete identity' with the Harappan (Indus-Saraswati) civilization.

## 7. A period of great changes

After the continuing ex-Africa populations (mtDNA M, N), and some agricultural Iranian influx (J2), we now witness the embedding of Central Asian elements in population (aDNA, Swat), language (IA), Indo-European poetics (!), religion (IIr., even IE), ritual (IIr. *Soma*, fire, etc.) and social organization (3–4 IE classes).

To visualize this, we can follow the Ehret model,[51] which entails not an 'Aryan Invasion' but a gradual immigration by various IA-speaking groups, occupying largely empty space in the Greater Punjab. This would not have occurred entirely without fights: the Ṛgveda speaks of the resistance of the chieftain Śambara in his mountain fortresses (of E. Afghanistan) that lasted for 40 years. Likewise, there will also have been some resistance in the plains.[52]

Both populations, the pastoral IAs and the settled remnant Indus one, underwent gradual acculturation: this is indicated by many loanwords found already in the Ṛgveda, and by overlaps in low-level religion. The ultimately dominant IA superstrate is now supported by genetic evidence of the Swat aDNA: the older 'Indus' (the modern ANI) DNA was overlaid by that of an intrusive group.

In general, we may ask: why not? This kind of scenario is not limited to South Asia: for example, Europe has now been shown to have three similar levels of immigration.

| EUROPE | INDIA |
|---|---|
| 40 kya: Out of Africa hunters | 50 kya: Out of Africa hunters |
| 10 kya Near Eastern agriculturalists (Basque!) | 7 kya Iranian agriculturalists |
| 2.5 kya "Ukranian" IE immigration: 75% replacement of older population survives; later reemerges and mixes with the immigrants | 1250 BCE: IA immigration. In Swat some Indus populations survive and mix |

To militate against the combined evidence of genetics, linguistics, texts, and rituals, and to insist on a *pristine* settlement of 'Indians', unchanged since they first arrived from Africa ca. 65 kya, is based on a politically motivated dream of primordial 'Aryans'. We witnessed this already when the first DNA results came out in the early 2000s, based just on modern mtDNA, that indicated a persistence of the female haplogroups M and N ever since the arrival of ex-Africans around 65 kya.

### 7.1 *The horse question*

It is generally agreed[53] that the horse is not an Indian, but a steppe animal. It was not present in the ancient near-east either. The typical southern equids are the donkey and the so-called half-ass (onager) that still survives in the Thar desert (Rann of Cutch). The half-ass looks like a horse with a donkey's tail. Their bones are not easily distinguishable. A skeleton may just be that of a big donkey or a smaller horse. Instead, one has to look at the phalanges only then one can decide which kind of animal it is. So if some archaeologists mention that they have found horse bones in the Harappan civilization, one has to be very suspicious. We still do not have a good collection of ancient bones anywhere in India that enables to compare such bones with each other. In short, the horse was introduced in the near-east after 1900 BCE and in India around 1800 BCE.

We have to add the question of the chariot. This is commonly defined as a light (30 kg), spoke-wheeled vehicle, drawn by two (or more) horses; it is quite different from the older, heavy wagon with four full wheels. To call, e.g., the heavy, four-wheeled wagon of Jagannāth in Puri a chariot is simply confusing terms. Chariots were possible only after ca. 2000 BCE in the near-east, and in India after roughly 1700 BCE.

To come back to horses and to put it facetiously: which language 'did these horses speak'? The IAs may very well not have been the first who brought the horse to India but instead some border tribes before them, just like the Guti, Lulubi, and Kassite (and Mitanni) at the rims of Mesopotamia about the same time. We might think of some Hindukush tribes mentioned in the Avesta: the *Muža, Taugra*, etc.

## 8. The post-Ṛgvedic period

After the IA occupation of the Greater Punjab (excluding Sindh, UP, etc.), with five dominant tribes (Anu, Druhyu, Yadu, Turvaśa, and Pūru) it was the victory of the Bharata subclan in the so-called 'the Kings' Battle' (RV 7.18) that settled the political framework for the following centuries: the Bharata-Kuru realm was dominant[54] in the post-Ṛgveda Saṃhitās and Brāhmaṇa texts.

Their particular IA model of society, religion, ritual, and texts gradually expanded, from the Kurukṣetra area by osmotic spread eastwards, as to include most of UP, up to Kausambi, but not to Allahabad (the ancient *Prayāga*).

This model is attested not only in the oldest Vedic prose texts[55] but also in archaeology: the IA pastoralists overlapped symbiotically with the population of the Painted Gray Ware culture. It must be noted, however, that the PGW settlements were hardly inhabited by IAs themselves, who continued to be on the move with their cattle (Ārya-āvarta). Instead, the PGW villages and small market towns were occupied by the local post-Indus populations. Even a relatively late Brāhmaṇa text, the Jaiminīya Br., advises not to stay in a village, or at maximum for just one night (due to inherent 'pollution'), and when traveling through such *dasyu* ('enemy'!) territory, one to take along a Kṣatriya: then the inhabitants 'would come smiling'.

When the post-Ṛgvedic Kuru polity spread eastwards as to include most of UP, the Pañcāla kingdom was formed (their kings in part intermarrying with the Kurus). Again, their tribal differences are not just mentioned in the Vedic texts but can actually be detected in the elite PGW pottery of the UP area: the Kuru and the Pañcāla use different designs.

These vessels were made by local potters who lived in symbiosis with the IAs in their treks (*grāma*), always on the move in Āryāvārta.[56] But in addition to these designations of *grāma* occupants, the first 'outcasts' appear: first, apparently the very low class *caṇḍāla* in a sacred text, the Chāndogya Upaniṣad,[57] with a stark characterization equating them with dogs and pigs.

Due to the close interaction with various artisan classes, more and more substrata words were taken over, especially in agriculture and artisanship.[58]

## 9. Cultural spread to the east

The Kuru-Pañcāla pattern spread further east in the late Vedic period, as reported in a late Brāhmaṇa text,[59] first to the Kosala in eastern UP, and then crossing the Gaṇḍakī (Sadānīrā) river into northern Bihar, into the country of the Videha.

While the PGW culture was limited more or less to Kurukṣetra and UP, the Black and Red Ware culture covered eastern UP and northern Bihar. Its character differed remarkably from that of the western PGW, although we only have few actual indications of this in the texts, such as that the grave mounds of the easterners were round (*parimaṇḍala*), like the *stūpa* grave mound of the Buddha.

However, there is a clear indication of the *intentional* spread of the Kuru pattern to Bihar, in the legend of the eastward march of the legendary King Videgha Māthava, from the Kuru heartland in Kurukṣetra to Videha, bringing the sacred fire (and its Brahmins) to 'sweeten' the marshy land of Videha (northern Bihar). The quest of Videgha's 'missionary' Gautama Rahūgaṇa is echoed by many other details in the late Vedic texts. They stress an intentional 'Sanskritzation' of the eastern area by fixing the sacred texts,[60] and by initiating many discussions among immigrant Brahmins for a prize set out by the Videha king Janaka.

Although he is praised for this in the late Brāhmaṇa and early Upaniṣad texts, his famous kingdom seems to have vanished a few generations later: the early Buddhist texts (in Pāli) merely mention the oligarchic Vajji (Vṛji) confederation in the area, of which the Videha were, by then, a non-monarchial part.

Even the Buddha's Śākya realm (in the Vajji area) was not a 'kingdom' as is usually thought: once his father is mentioned as plowing his own field – not exactly a royal occupation. As so often, we have to revise our well-established ('inherited') concepts or prejudices in view of new evidence, often staring us into the face, but overlooked. (This includes the revised date of the Buddha around 400 BCE, not around 500 BCE as believed for more than a hundred years.)[61]

Additionally, though there are some clear differences between the Kuru-Pañcāla realm and the East, we cannot go as far as Bronkhorst[62] as done in his book on 'Greater Magadha', as we hardly know anything about (northern) Bihar – forgetting about *actual* Magadha area south of the Ganges, which never was a Vedic territory.[63]

However, what we clearly do witness is the gradual, osmotic spread of the Kuru model all over northern, and later also over southern, India. It informed what later on would become the model of Hindu society and religion.

The Kuru model thus has its effects until today: notably with the dominant Brahmin-Ksatriya 'combine'. The *puruṣa sūkta* of Ṛgveda 10.90 with its division of society into four classes still is the unwritten but effective constitution of India.

**Endnotes**

1. The term 'Aryan', used in India nowadays, in fact refers to the *speakers* of the old Indo-Aryan language of the Veda, similar to but preceding that of Pāṇini's Sanskrit. Their religion, ritual, poetry, poetics and societal setup are usually subsumed under this cover term. All of this has no relationship with 'race' (merely a cultural construct), though even the last, monarchial constitution of Nepal still spoke of the King as being of the 'Aryan race'.
2. Even recently: Kennedy KAR 2000 *God-apes and fossil men: Paleoanthropology of South Asia* (Ann Arbor: University of Michigan Press), spoke of 'no invasion' and a (actually spurious) similarity of Indian paleontological data with those of eastern Central Asia (Xinjiang).
3. Narasimhan VS *et al.* The genomic formation of South and Central Asia, http://dx.doi.org/10.1101/292581.
4. Ali I *et al*. 2009 Latest archaeological explorations in the Chitral Valley. *Pakistan Heritage* 2.
5. Blažek V 2006 Was there Australian substratum in Dravidian? *Mother Tongue* **XI**, 275–285.
6. Masica CP 1979 Aryan and non-Aryan elements in North Indian agriculture; in *Aryan and non-Aryan in India* (eds) MM Deshpande and PE Hook (Ann Arbor:

Center for South and Southeast Asian Studies) pp 55–151.

7. Witzel 2012 *The origins of the world's mythologies* (New York: Oxford University Press).
8. Except for such data as tooth shape (Sino-dont, Sunda-dont), which are not of importance here.
9. Note the early alleged carbon-14 dating at Bhirrana in the Saraswati river basin in 7th millennium BCE.
10. Fuller DQ 2006 Silence before sedentism and the advent of cash-crops. A status report on early agriculture in South Asia from plant domestication to the development of political economies (with an excuses on the problem of semantic shift among millets and rice); in *Proceedings of the Pre-Symposium of RHIN and 7th ESCA Harvard-Kyoto Round Table* (ed) Osada T (Kyoto: Research Institute for Humanity and Nature (RHIN)) pp 175–213; Fuller, D.Q. Framing a Middle Asian corridor of crops exchange and agricultural innovation. In: 13th Harvard University Round Table. Ethnogenesis of South and Central Asia (ESCA), Kyoto session. Kyoto: Research Institute for Humanity and Nature (RHIN), Kyoto, Japan, 30–31 May 2009, p. 3–11; Fuller, D.Q. *The Archaeobotanist*, 25 June 2009, *http://Archaeobotanist.blogspot.com/indian*-archaeology-watch-lahuradewa.html; Fuller, D. Q. *The Archaeobotanist* 25 Aug 2009; *http://archaeobotanist.blogspot.com/2009/08/millets*- and-mistakes.html.
11. Witzel 2004 Central Asian Roots and acculturation in South Asia: Linguistic and archaeological evidence from Western Central Asia, the Hindukush and Northwestern South Asia for early Indo-Aryan language and religion; in *Linguistics, archaeology and the human past* (ed) T Osada (Kyoto: Indus Project, Research Institute for Humanity and Nature) pp 87–211), 2006 South Asian agricultural vocabulary; In: T Osada (ed.) Proceedings of the Pre-Symposium of RHIN and 7th ESCA Harvard-Kyoto Round Table. Published by the Research Institute for Humanity and Nature (RHIN), Kyoto, Japan 2006 pp 96–120.
12. See Narasimhan 2018 Shahr-i Sokhta etc. on the E. Iran/Afghan border. More Indus data will be coming out soon, by the Reich team at Harvard.
13. Masica 1979
14. Like *kangu*, cf. Bantu *i-kōngō*; see Witzel 2006 South Asian agricultural vocabulary; in. *Proceedings of the Pre-Symposium of RHIN and 7th ESCA Harvard-Kyoto Round Table* (ed.) T Osada (Kyoto, Japan: Published by the Research Institute for Humanity and Nature (RHIN)) pp 96–120; Witzel 2009 The linguistic history of some Indian domestic plants. Indian Academy of Sciences. *Journal of Biosciences* 34(6) 829–833. *http://www.ias.ac.in/jbiosci/dec2009/contents.htm*. Full text: *http://www.ias.ac.in/jbiosci/dec2009/Witzel_fulltext.pdf*
15. Zvelebil K 1970 *Comparative Dravidian phonology* (The Hague). Zvelebil K *Dravidian linguistics: An introduction* (Pondicherry: Pondicherry Institute of Linguistics) – Note also the Vedda language in Sri Lanka. De Silva MW Sugathapala *Vedda language of Ceylon; texts and lexicon* Münchener Studien zur Sprachwissenschaft. Beiheft n.F. 7. München: R. Kitzinger, 1972.
16. Sato Y-I 2006 Rice and the Indus civilization; in *Osada* pp 213–214. Sato Y-I 2006 Crops: What is common and what is different? - Fudo and agriculture; in *Proceedings of the Pre-Symposium of RHIN and 7th ESCA Harvard-Kyoto Round Table* (ed.) T Osada (Kyoto: Published by the Research Institute for Humanity and Nature) pp 73–78.
17. Witzel 1999, 2004.
18. Southworth FC 2005 *Linguistic archaeology of South Asia* (London and New York: Routledge).
19. Note that an Australian trait had been discovered in the modern DNA of a Tamil individual already some 20 years ago by Professor R. Pitchappan. *https://www.hindustantimes.com/…/geneticists…/story-INqikMSk5JRjGsICqHFzMN.htm*. *https://www.ncbi.nlm.nih.gov/pubmed/28287095*); cf. *https://www.researchgate.net/profile/Ramasamy_Pitchappan/research*).
20. There is a list in Witzel 1999, including a map of (north) Indian substrate languages, see: *http://www.people.fas.harvard.edu/%7Ewitzel/MT-Substrates.pdf*.
21. *http://www.aa.tufs.ac.jp/sarva/*; *http://www.aa.tufs.ac.jp/sarva/materials_frame.html*. The dictionary itself: SARVA substrate dictionary (in progress; upper right hand frame)
22. For the mature Harappan period; maximally, the Harappan dates are 3300–1300 BCE.
23. Incidentally, the old name of Harappa cannot be *hariyupīya* (RV), even if the language of the northern Indus area had been Sanskrit: this form would have resulted in something like modern Punjabi Haryua or similar.
24. Actually, both the words *Sindhu* and *Sarasvatī* are not taken from an (unknown) Harappan language but from much later Vedic Sanskrit.
25. See: Steve Farmer, Richard Sproat and Witzel M 2004 The collapse of the Indus-Script Thesis: The myth of a literate Harappan civilization. *Electronic Journal of Vedic Studies* 11–2, 19–57.
26. Such as: the so-called 'granary' and the 'great bath'.
27. See: Tripathi JK *et al*. 2004 'Is river Ghaggar, Saraswati? Geochemical constraints', *Current Science* **87** 1141–1145; Giosan L *et al*. 2012 Fluvial landscapes of the Harappan civilization. *PNAS* **109** E1688–E1694; Clift PD *et al*. 2012 U-Pb zircon dating evidence for a Pleistocene Saraswati river and capture of the Yamuna river. *Geology* **40** 211–214; Maemoku H *et al*. 2012 Geomorphological constraints on the Ghaggar river regime during the mature Harappan period; in *Climates, landscapes, and civilizations* (ed.) Giosan L *et al*. (Washington, D.C.: AGU) pp 97–106;

Singh A *et al* Stratigraphic response to Late Quaternary monsoonal fluctuations in a buried valley complex in Ghaggar plains, NW India. Poster presented at the 4th PAGES Open Science Meeting (Goa, India, 13–16 Feb 2013).

28. Note that some Harappan cities were built right inside the up to 10 km widestream bed that was created during the post-glacial meltdown - like the several Central European '*urstromtal*' river beds.
29. See Khatayat G *et al*. 2017 The Indian Monsoon variability and civilization changes in the Indian subcontinent. *Science Advances* **3** 2017. https://doi.org/10.1126/sciadv.1701296. This study is based on data retrieved from ice cores in Himalayan mountain caves; these ice cores have helped to date the changing civilizations well: the gaps between pre-Harappan (pre-Indus) Valley Civilization, the early Vedic period, etc. were marked by a change in climate, which actually ties up well with most of the preserved texts.
30. Note that *Saras-vatī* (like Avestan Harax$^{v}$aitī) means 'having lakes'.
31. Note also the careful water management in Dholavira.
32. Though not infrequently alleged in nationalistic writings and even by Parpola A 1994 *Deciphering the Indus script* (Cambridge: Cambridge University Press).
33. The so-called *gasshō dōgu*.
34. And maybe the temporary Sutlej diversion into the lower Ghaggar-Hakra (Sarasvatī) stream bed: this needs to be studied *in loco*. See Mughal MR 1997 *Ancient Cholistan. Archaeology and architecture* (Rawalpindi-Lahore-Karachi: Ferozsons).
35. RV book 4, e.g. 4.26.3, on the mountain chieftain Śambara, see Stuhrmannm Ṛgvedisch *púr*. EJVS 15, 2008: https://crossasia-journals.ub.uni-heidelberg.de/index.php/ejvs/article/view/394/389.
36. 383 words listed by Kuiper FBJ 1991 *Aryans in the Rigveda* (Amsterdam-Atlanta: Rodopi).
37. Drav. loanwords increased only in the post-Ṛgvedic period, see Witzel 1999, 2004.
38. Instead, there is a substrate in the Punjab (Witzel 1999) that I had called 'Para-Munda', a term often misunderstood: it shares only some grammatical features (prefixes, but not its infixes) with Munda, but is not closely related to the Munda languages.
39. There has been only a minor excavation at Harappa; other evidence comes from some late Harappan cemeteries.
40. Habitually, nearly all old carts are incorrectly called 'chariots' in the press and even in Indian scholarly papers.
41. Just as is the case in other parts of the northwestern subcontinent.
42. See Witzel 1999, 2004.There were several different substrates in the North Indian area, see Witzel 1999 (with map).
43. Painted Gray Ware (PGW), Black and Red Ware (BRW), etc.
44. Extending even to numbers like **ćata* '100', see Rédei K ältesten i. in *lhf*, (ed.) Sinor D, (Leiden: Brill) pp 638–64.
45. David Anthony and Dorcas Brown December 2017 The dogs of war: A Bronze Age initiation ritual in the Russian steppe, *J. Anthropological Archaeology* p 48. Cf. Anthony 2007 *The horse, the wheel, and language: how Bronze-Age riders from the Eurasian steppes shaped the modern world* (Princeton: Princeton University Press).
46. See Witzel The home of the Aryans; in *Anusantatyai. Fs. für Johanna Narten zum 70. Geburtstag* (ed.) A Hintze and E Tichy (Münchener Studien zur Sprachwissenschaft, Beihefte NF 19) Dettelbach 2000 (J.H. Roell) pp 283–338. http://www.people.fas.harvard.edu/%7Ewitzel/AryanHome.pdf.
47. Lubotsky A 2001 'The Indo-Iranian substratum'; in *Early contacts between Uralic and Indo-European: Linguistic and archaeological considerations* (ed.) Chr. Carpelan *et al.* (Helsinki: suomalais-Ugrilainen seura); Witzel 1999, 2004; Witzel 2015 The central Asian substrate in Old Iranian. *Mother Tongue* **XX** 149–178.
48. Witzel 2014 Mitanni Indo-Aryan *Mazda* and the date of the Ṛgveda; in *The complex heritage of early India. Essays in memory of R.S. Sharma* (ed.) DN Jha (New Delhi: Manohar) pp 73–96.
49. Analogical changes are still going on in the RV and later: Such as the prefect form *mene.*
50. Possehl G and Gullapalli P 1999 The Early Iron Age in South Asia; in *The archaeometallurgy of the Asian Old World* (ed.) V Pigott (Philadelphia: The University Museum) pp 153–175.
51. Ehret 1988 Ch. Language change and the material correlates of language and ethnic shift. *Antiquity* **62** 564–574.
52. Stuhrmann 2016 Die Zehnkönigsschlacht am Ravifluß. *EJVS* **23**. https://crossasia-journals.ub.uni-heidelberg.de/index.php/ejvs/article/view/933/971.
53. Except by some revisionists who bring up the long extinct Ice-Age Siwalik horse, or even the still 3-toed *Equus namadicus*. The Siwalik horse died out in the post-glacial mega-fauna extinctions at the end of the Ice Age at 10,000 or 9000 BCE)
54. Witzel 1997 Early Sanskritization. Origins and development of the Kuru State; in *Recht, Staat und Verwaltung im klassischen Indien. The state, the law, and administration in classical India* (ed.) B Kölver (München: R. Oldenbourg) pp 27–52. EJVS version: http://www.ejvs.laurasianacademy.com/ejvs0104/ejvs0104article.pdf. Summary, Kyoto 1989: http://www.people.fas.harvard.edu/%7Ewitzel/KuruKyoto89.htm.
55. Maitrāyaṇī, Kaṭha, and Taittirīya Saṃhitā of the Kṛṣṇa Yajurveda.
56. This is seen also with the culturally very conservative Kalasha people in Chitral (on the Afghan border) who still worship the god Indr.

57. ChU 5.10.7.
58. *See* Witzel 2006 South Asian agricultural vocabulary.
59. Śatapatha Brāhmaṇa 1.4.1.
60. Such as the Ṛgveda by the Easterner Śākalya, the Śrauta ritual in Śatapatha Brāhmaṇa (and its excerpt, the Vājasaneyi Saṃhitā), see Witzel 1997.
61. See Bechert H 1982 The date of the Buddha reconsidered. *Indologica Taurinensia* **10** 29–36.
62. Bronkhorst J 2007 *Greater Magadha* (Leiden/Boston: Brill).
63. A Magadha man is only brought in Vedic ritual to represent the typical, polluting outsider.