

# Human population genomics of past and present: The current scenario of South Asia on the global scale



Dr. Niraj Rai

Birbal Sahni Institute of Palaeosciences

Lucknow, UP, India

E-mail: [nirajrai@bsip.res.in](mailto:nirajrai@bsip.res.in)

Dr. Rai is working as a Scientist at the Birbal Sahni Institute of Palaeosciences, Lucknow, India. His research interests include the genetic investigations of the archaeological human remains with an aim to decipher their ethnic origin, migration and evolutionary affiliations.

**Citation:** Rai N. Human Population Genomics of past and present: The current scenario of South Asia on the global scale. *Polymorphism* 2019;2:114-117.

## INTRODUCTION

In order to reconstruct and explain the patterns of genetic diversity in modern humans, an understanding of both the past and present population dynamics is crucial. While genomes from modern individuals can inform about the present-day population structure, the ancient genomes provide unprecedented insights into the past demographic events that have shaped the present-day gene pool. Population genetics has recently witnessed an explosion in studies on ancient human population histories, primarily from Europe and the Americas. South Asia has no representation in the ancient genomics literature and not much work has been done on the modern day Indian populations despite the wealth of population diversity and the archaeological richness in the form of human skeletal remains that exist in collections all over the country. Having a fifth of the present-day humanity representation, understanding the demographic history of South Asia should not only be a prerequisite but also an urgent need to understand its genetic variations on the global scale. Though the overall picture is still at best emerging, new archaeological and genetic information from the region has started to reveal a more complex scenario of the ancient human migrations and admixtures than what was known until now.

Here, I will discuss the novel questions about deep ancestry components of South Asians at the Eurasian and global scales and how genetic makeup of South Asians can be well understood from studying the ancient genomes. Reconstructions of the past effective population size strongly suggest that South Asia has been populated almost since the Early Upper Palaeolithic (Atkinson et al., 2008; Petraglia et al., 2009; Xing et al., 2010). One of the probable explanations of this pattern relates to the fact that South Asia was among the first geographic regions to be peopled by the modern humans after their African exodus (Petraglia et al., 2009; Xing et al.,

2010). Today, various ethnic groups of South Asia comprise tribes, castes, and populations identifying themselves by different religions, being largely endogamous and hence revealing complex and multilayer genetic differentiation (Reich et al., 2009; Metspalu 2011; Moorjani et al., 2013; Chaubey 2014; Basu et al., 2016). From such a complex structure, several questions have stood out that could not be addressed a few years ago, but can now be using the latest technology and resolution. The novelty of the approaches rely largely, though not solely, on extensive high-coverage sequencing of specifically targeted modern and ancient samples, autosomal SNP genotyping of specific populations, extensive Y chromosome resequencing of the selected population groups (tribal and caste populations from different geographical areas) and statistical analysis of the results in the context of human genomic variations worldwide. Recently, ancient DNA (aDNA) studies (though not directly on South Asian ancient remains) have suggested that the one ancestral component (Ancestral North Indian or ANI) component in South Asia is closely related with Caucasus Hunter-Gatherer (CHG) (Jones et al., 2015) or Neolithic Iranian (Broushaki et al., 2016), or that it might represent an amalgamation of Neolithic Iranian and Steppe ancestries (Lazaridis et al., 2016). However, surprisingly, the proposed Steppe ancestry was much more widely distributed geographically than it was suggested by the linguists and Indologists (Rebftew, 1991, Lal, 1997; Witzel, 2015). Therefore, aDNA data from India is of utmost importance to resolving these complexities. Now with the help of technological advancements, ancient DNA data from South Asia is needed to establish a plausible and well-supported outline for South Asian demographic prehistory. It also merits pointing out that the structure of the South Asian population has rightfully been described as 'pockets of endogamy' (Cavalli et al., 1994), presenting a challenge to medical genetics, and this project will certainly provide valuable empirical data to be translated to disease genetics.

## Major questions that need research

With the technology-driven revolution in (human) population genetics, the complete human genomes of modern and ancient samples and the complete MSY (male-specific region of the Y chromosome) sequences are now becoming readily accessible to test a number of hypotheses.

**1:** Recent studies have identified an excess of highly derived modern human haplotypes and deeper split times for Papuans from Africans (Pagani et al., 2016), suggesting that there was an earlier migration in addition to the main wave of peopling Eurasia. If such migration took place, South Asia would be by far the most likely candidate region for finding its remnants in Eurasia.

**2:** Recent genome-wide analyses suggest two major ancestral components of South Asian populations (Reich et al., 2009; Metspalu 2011). Additionally, based on *f4* ancestry estimation, it has been reported that with the exception of the Andamanese Negrito (Onge) population, all Indian populations carry ANI and ASI ancestries (Reich et al., 2009; Moorjani, 2013). However, off-cline (Reich et al., 2009) populations, i.e. Austroasiatic and Tibeto-Burmans, were not included in such analyses. Interestingly, recently published work on Andaman Islanders showed their highest affinity to Austroasiatics when compared to all Indian populations. This relationship was stable even after masking the East/Southeast Asian specific component among Indian Austroasiatic (mainly South Munda) populations.

**3:** In spite of the past difficulties, the field of aDNA has now emerged as a reliable and robust research tool due to the recent significant methodological improvements (Jones et al., 2011; Broushaki et al., 2016; Lazaridis et al., 2016). It is now possible to study the human skeletal remains excavated from the biggest Harappan city, Rakhigarhi,

which is dated back to 6,000 years before present (YBP), Mesolithic time period samples from Gangetic plain (8,000 YBP), Chalcolithic samples (4,500 YBP) from Orissa and Chhattisgarh, Megalithic samples (up to 3500 YBP), early historic samples from South of India ( up to 3,000 YBP) and a very well preserved ancient remains found in a lake known as Roopkund, where >1000 individuals died approximately ~1200 years ago.

## Conclusion

Indian diversity is highest after the African continent, which shows the oldest migration of modern humans from Africa. The highest diversity among Indians was also achieved by multi-layered admixture events followed by strict endogamy practices among different population groups. The genetic data of modern and ancient human specimens from different geographical regions of South Asia has the potentiality to understand the deep rooting genetic makeup of South Asians. Considering the fact that Indian populations show maximum founder effect, Geneticist have the opportunity to map the genes, if genetic data are being interpreted holistically using a wide range of analytical procedures (from phylogeography, coalescent theory, and statistics) in an interdisciplinary synthesis in the light of evidence from archaeological, linguistic and genomic data. In search of this, we aim to sequence the complete human genomes of 250 individuals from carefully selected relic populations for e.g. the linguistic isolates Kusunda, Nihali, Vedda and many more (Bengston et al., 1996; Whitehouse, 1997; Chaubey, 2014) and also the ancient populations (Archaeological human remains) with an aim to revisit the two versus one migration scenario for the peopling of Eurasia.

## REFERENCES

- Atkinson, Q. D., Gray, R. D. & Drummond, A. J. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Mol Biol Evol.* 25, 468-474 (2008).
- Basu, A., Sarkar-Roy, N. & Majumder, P. P. Genomic reconstruction of the history of extant populations of India reveals five distinct ancestral components and a complex structure. *Proc Natl Acad Sci U S A.* 113, 1594-1599 (2016).
- Bengston, J. "Nihali and Ainu.". *Mother Tongue*, 51 55 (1996).
- Broushaki, F. et al. Early Neolithic genomes from the eastern Fertile Crescent. *Science.* (2016).
- Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. *The history and geography of human genes* (Princeton University Press, Princeton, N.J., 1994).
- Chaubey, G. Language isolates and their genetic identity: a commentary on mitochondrial DNA history of Sri Lankan ethnic people: their relations within the island and with the Indian subcontinental populations. *J Hum Genet.* 59, 61-63 (2014).
- Jones, E. R. et al. Upper Palaeolithic genomes reveal deep roots of modern Eurasians. *Nature communications.* 6, 8912 (2015).
- Lal, B. *The Earliest Civilization of South Asia: Rise, Maturity, and Decline* (Aryan Books International, 1997).
- Lazaridis, I. et al. Genomic insights into the origin of farming in the ancient Near East. *Nature.* (2016).
- Metspalu, M. et al. Shared and unique components of human population structure and genome-wide signals of positive selection in South Asia. *Am J Hum Genet.* 89, 731-744 (2011).
- Moorjani, P. et al. Genetic evidence for recent population mixture in India. *Am J Hum Genet.* 93, 422- 438 (2013).
- Pagani, L. et al. Genomic analyses inform on migration events during the peopling of Eurasia. *Nature.* 538, 238-242 (2016).
- Petraglia, M. et al. Population increase and environmental deterioration correspond with microlithic innovations in South Asia ca. 35,000 years ago. *Proc Natl Acad Sci U S A.* 106, 12261-12266 (2009).
- Reich, D., Thangaraj, K., Patterson, N., Price, A. L. & Singh, L. Reconstructing Indian population history. *Nature.* 461, 489-494 (2009).
- Renfrew, C. *The coming of the Aryans to Iran and India and the cultural and ethnic identity of the Dasas.* By Asko Parpola. (*Studia Orientalia*, Vol. 64.) pp. 195-302, 33 figs. Helsinki, The Finnish Oriental Society, 1988. *Journal of the Royal Asiatic Society (Third Series).* 1, 106-109 (1991).
- Whitehouse, P. "The External Relationships of the Nihali and Kusunda Languages.". *Mother Tongue.* 3, 4-44 (1997).
- Witzel, M. in *Linguistics, Archaeology and the Human Past.* (ed. Osada, T.) 87-211 (Research Institute for Humanity and Nature, Kyoto, Japan, 2005).
- Xing, J. et al. Genetic diversity in India and the inference of Eurasian population expansion. *Genome Biol.* 11, R113 (2010).