

## **HUMAN ANCESTRY AND PEOPLE OF INDIA: GENETIC AND NON-GENETIC EVIDENCE**

**Niraj Rai, Gyaneshwer Chaubey and Kumarasamy Thangaraj**

### **ABSTRACT**

Due to its important geographical location and juxtaposition to the transcontinental admixture and migrations to and from both East and West, India has played a major role in shaping the peopling of old world, in both prehistoric and historic times. The archaeological evidence suggests the presence of hominin in this region since the Middle Pleistocene. The issue of ancient settlers and impact of various migrations are highly debatable; however there is a consensus between different disciplines of humanity that modern humans arrived to India from Africa. Linguistic, Archaeology and Genetic studies of the subcontinent largely portray different pictures. In recent years autosomal studies in human populations have revealed new insights about the prehistoric origin and migrations of anatomically modern humans (AMH). In this review we have analyzed the genetic and non-genetic evidence available so far to build a consensus on peopling of South Asia.

The geographic position and multiple line of evidence suggest that South Asia played a pivotal role in the colonization of the old world (Field *et al.*, 2007; Petraglia and Allchin, 2007; Atkinson *et al.*, 2008; Petraglia *et al.*, 2010). Therefore, studying South Asia may provide the opportunity to re-evaluate and refine the models related to the evolutionary history of the modern humans.

### **ARCHAEOLOGICAL EVIDENCE**

The archaeological evidence on hominid populations are abundant throughout the Indian subcontinent during the middle and late Pleistocene, however, the fossil record of Palaeolithic anatomically modern humans is poor (Mishra, 2001). The so far discovered South Asian distinctive archaeological records challenge many of the models and theoretical framework that have emerged on the basis of findings made in other continents (Allchin and Allchin, 1997; Petraglia *et al.*, 2007). The modern human occupation in South Asia has been suggested just after the Out-of-

---

**Niraj Raj**, Post Doctoral Fellow, Centre for Cellular and Molecular Biology (CCMB), Hyderabad, INDIA, **Gyaneshwer Chaubey**, Evolutionary Biology Group, Estonian Biocentre, Riia 23, Tartu, Estonia-51010 and **Dr. Kumarasamy Thangaraj**, Centre for Cellular and Molecular Biology (CCMB), Hyderabad, INDIA, E-mail: [thangs@ccmb.res.in](mailto:thangs@ccmb.res.in)

Africa migration, however, there is scarcity of fossil evidence due to unfavorable climate for fossilization (James and Petraglia, 2005). The archaeological support for coastal route is vague, likely due to rise in sea levels aftermath LGM (last glacial maximum) (Field *et al.*, 2007). However, the presence of lithic components throughout the region indicates occupation of the subcontinent by hominid populations during late Pleistocene (Kennedy, 2000; Mishra, 2001). The oldest known fossil hominid in India was a partial cranium recovered from Hathnora (Central India) in the Narmada Basin (Sonakia, 1984), which was calibrated by faunal correlation to ca. 300,000–250,000 years ago (Kennedy, 2000), and was suggested to belong to *Homo heidelbergensis* (Rightmire, 2001). The recent studies however, found it considerable younger (160-85 Kya) (Patnaik *et al.*, 2009; Patnaik and Chauhan, 2009).

The finding of Petraglia and colleague based on continuity of middle Paleolithic technology suggested the presence of human before and after the Toba eruption (Petraglia *et al.*, 2007) in South Asia. Though, there is no such old fossil record available so far discovered. This contradicts the 'single southern coastal route' model Out-of-Africa migration supported by the mtDNA and Y chromosome analysis of Eurasian and Oceanian lineages coalesce to a common ancestor to Africa 40-60 thousands years ago. Recently, analysis of ancient genome from Australia suggested that in addition to the main wave to Eurasia there was an earlier one, genetic signal of which is today found in the Australian Aborigines and PNG highlanders (Rasmussen *et al.*, 2011). This opens up a new question which needs a high number of complete genomes to test.

The earliest fossils of modern humans in South Asia have been unearthed in Sri Lanka at Fa Hien Cave dated to ca. 31,000 years ago and Batadomba-lena dated ca. 28,500 years (Deraniyagala, 1984). The human skeletal finds from this site correspond to the earliest evidence for anatomically modern Humans in South Asia. Moreover, considering abundant microlithic industries at sites such as, Fa Hien Cave Sri Lanka, Patne and Jwalapuram in India bolstered the evidence of early presence of modern human in South Asia (Deraniyagala, 1984; James and Petraglia, 2005).

The initiation of microlithic technologies in human prehistory provided a highly sophisticated tool for hunting. Archaeological studies on African microlithic tools suggested emergence of complex technologies offering greater technological flexibility, as well as more sophisticated hunting strategy and planning abilities (Ambrose, 2008; Powell *et al.*, 2009). Similar but independent technological innovation has also been reported Southern India which signifies that systematic production of small, retouched stone blades was common there by ~28-32 Kya, which was suggested to be associated with Late Pleistocene indigenous population expansion in South Asia (Petraglia *et al.*, 2009). The discovery of microlithic technology by modern humans in South Asia was exciting keeping in mind the varying subsistence pattern before the onset of LGM) (Petraglia *et al.*, 2009; Clarkson *et al.*, 2009).

The Neolithic and Bronze age archaeology of South Asia is also scarce except some remarkable discoveries e.g. Mehrgarh and Harappan civilisation. Mehrgarh in Baluchistan has already demonstrated the origins of this culture, which was gradual from the modest beginning of settled life around 7,500 BC (Jarrige, 1981; Costantini, 1984). The discovery of the Harappan Civilization in the early twentieth century was regarded as one of the most important archaeological discovery in South Asia, not only because this discovery has identified it as one of the earliest civilizations of the world, but also it stretched back the antiquity of social structuring in South Asia by two thousand years at one stroke (Shinde *et al.*, 2009). This civilization was unique on the account of its remarkably large coverage and town planning in comparison to the two contemporary civilizations (*i.e.* Mesopotamian and the Egyptian). Indus Valley which is situated towards north-western part of India has been under the focus of research by archaeologists, geneticists and linguists interested in the continuity and external influences of Indian cultural heritage (Singh, 1971; Mishra, 2001; McElreavey and Quintana-Murci, 2005; Madella and Fuller, 2006; Fuller, 2006; Fuller, 2007; Good *et al.*, 2008). This civilization was at its peak between 2500-1900 B.C. (Lawler, 2008).

The 1400 settlements discovered so far are distributed in an area of about 1250 thousand square kilometers. The genetic study of contemporary Indus populations revealed a mixture of different components from West Eurasia, South Asia and East Asia (Quintana-Murci *et al.*, 2004). It was also noted that the population located at East of Indus shows prevalent South Asian maternal lineages, while populations located West of Indus carry of western Eurasian lineages (Quintana-Murci *et al.*, 2004). Thus, Mehrgarh and Harappa provide an important evidence for the change from hunting-gathering and pastoralism to a subsistence economy revolved around agriculture and domestication of wild animals (Costantini, 1984; , 1997).

The abrupt fall of this civilization is an unresolved issue. Over the past several decades, many theories have been surfaced behind the sudden disappearance of this well flourished civilization. Based on palynological evidence, Singh (Singh, 1971) suggested that increase in the monsoon rainfall directly contributed to the rise of the Indus civilization and a decrease in precipitation contributed to its fall, which was criticized (Madella and Fuller, 2006). The drying up of river Sarasvati and shifting of whole Indus-Valley population towards Ganga plains is another theory given behind the rapid fall of this civilization (Mishra, 2001; Tripathi *et al.*, 2004; Gupta *et al.*, 2006). Madella and Fuller, (Madella and Fuller, 2006) have considered strategic local shifts in agriculture in response to prolonged droughts at 2200 BC, which have contributed to the de-urbanization process. However, so far none of them have been really conclusive.

## GENETIC EVIDENCE

The extensive research covering all the disciplines of humanity unanimously suggests an African origin of modern humans. All non-African populations display only a subset of genetic diversity present in Africa (Underhill *et al.*, 2000; Behar *et*

*al.*, 2008). Considering mtDNA macrohaplogroups M and N are derived from a single branch (L3) exclusively present in Africa (Behar *et al.*, 2008). Similarly, paternal lineages are also derived from an African branch carrying CT-M168 polymorphism (Underhill and Kivisild, 2007). The number of dispersal events are however is a disputed issue (Lahr and Foley, 1994; Rasmussen *et al.*, 2011). Yet, mtDNA and Y chromosomal studies are in support of a one wave exit of modern humans along the southern route (Thangaraj *et al.*, 2005; Mellars *et al.*, 2013). Similar scenario has also been supported from archaeological studies (Mellars, 2006; Mellars *et al.*, 2013). However, it is also disputed in another archaeological study which suggest the presence of modern humans in South Asia prior to 74Kya (Petraglia *et al.*, 2007; Petraglia *et al.*, 2010).

The relations of different language groups of South Asia and their origin are highly debated issues. Except for Dravidian, none of the current language family in India has in-situ origin in the subcontinent. But it should be taken into consideration that none of the language family in world is as old as mtDNA or Y chromosomal age of any haplogroup and therefore should not be related with linguistic affiliations. Based on results of haploid genetic markers e.g. mtDNA and Y-chromosome, it is widely accepted that the South Asian genepool is largely autochthonous. However, there is a serious dispute on autosomal studies. Population level and individual level variation now is gaining a common place for association studies and analysis of human evolutionary histories. The beginning of genome wide DNA-era was a major breakthrough to compare principally genomic segments in different individuals (Li *et al.*, 2008; Reich *et al.*, 2009; Xing *et al.*, 2010; Metspalu *et al.*, 2011; Shah *et al.*, 2011; Moorjani *et al.*, 2013; Chaubey *et al.*, 2015).

In spite of the fact that India comprises one fifth of the world population, neither the HapMap, nor the global panels that have been used for whole genome analysis included any population from India (International HapMap Consortium, 2005). Some of the studies have added Pakistani populations which was able to provide a glimpse of South Asian prehistory with some degree (Li *et al.*, 2008; Jakobsson *et al.*, 2008). Study based on autosomal microsatellite markers had inferred that Indian populations show low levels of genetic differentiation (Rosenberg *et al.*, 2006), which was likely due to limited sampling and low geographical coverage (Indian Genome Variation Consortium, 2008). In a pioneer study the Indian Genome Variation Consortium studied 1871 individuals from 55 diverse populations and provided a high degree of genetic differentiation among Indian ethnic groups (Indian Genome Variation Consortium, 2008). Using this genetic data, they identified five clusters among these 55 populations. First and second clusters comprise North Indian Indo-European and Sino-Tibetan populations group respectively, third and fourth clusters consist western Indo-European and Dravidian populations respectively, while fifth cluster harbors Munda populations (Indian Genome Variation Consortium, 2008).

Initially, couple of studies came up with few Indian samples showing caste-tribal differences and north-south gradient (Auton *et al.*, 2009; Xing *et al.*, 2009). The South



In conclusions, patterns of modern human population structure discussed here can be used to guide construction of historical models of migration and admixture that will be useful in inferential studies of human genetic history in South Asia.

### REFERENCES

- Allchin B and Allchin FR. (Ed.), *Origins of a civilization Viking Adult* (1997).
- Ambrose S. Small Things Remembered: Origins of Early Microlithic Industries in Sub-Saharan Africa. 2008; 12: 9-29.
- Atkinson QD, Gray, RD and Drummond AJ. mtDNA variation predicts population size in humans and reveals a major Southern Asian chapter in human prehistory. *Mol. Biol. Evol.* 2008; 25: 468-474.
- Auton A, Bryc K, Boyko AR, Lohmueller KE, Novembre J, Reynolds A, Indap A, Wright MH, Degenhardt JD, Gutenkunst RN, King KS, Nelson MR and Bustamante CD. Global distribution of genomic diversity underscores rich complex history of continental human populations. *Genome Res.* 2009; 19: 795-803.
- Behar DM, Villems R, Soodyall H, Blue-Smith J, Pereira L, Metspalu E, Scozzari R, Makkan H, Tzur S, Comas D, Bertranpetit J, Quintana-Murci L, Tyler-Smith C, Wells RS and Rosset S. The Dawn of Human Matrilineal Diversity. *The American Journal of Human Genetics.* 2008; 82: 1130-1140.
- Chaubey G, Kadian A, Bala S and Rao VR. Genetic Affinity of the Bhil, Kol and Gond Mentioned in Epic Ramayana. *PloS one.* 2015; 10: e0127655.
- Clarkson C, Petraglia M, Korisettar R, Haslam M, Boivin N, Crowther A, Ditchfield P, Fullar D, Miracle P, Harris C, Connell K, James H and Koshy J. The oldest and longest enduring microlithic sequence in India: 35 000 years of modern human occupation and change at the Jwalapuram Locality 9 rock shelter. *Antiquity.* 2009; 83: 326-348.
- Costantini L. 1984. The beginning of agriculture of the Kachi plain: the evidence of Mehrgarh, in: Allchin, B. (Ed.), *South Asian Archaeology* UK, 1981: pp. 346.
- Deraniyagala SU. Mesolithic stone tool technology at 28,000 B.P. in Sri Lanka. *Ancient Ceylon.* 1984; 5: 105-108.
- Field J, Petraglia M and Lahr M. The southern dispersal hypothesis and the South Asian archaeological record: examination of dispersal routes through GIS analysis. *BMC Biol.* 2007; 26: 88-108.
- Fuller D. Non-Human Genetics, Agricultural Origins and Historical Linguistics in South Asia, in: Petraglia, M., Allchin, B. (Ed.), *Vertebrate Paleobiology and Paleoanthropology.* Springer, 2007: pp. 393-443.
- Fuller D. Agricultural Origins and Frontiers in South Asia: A Working Synthesis. *J. World Prehist.* 2006; 20: 1-86.
- Good I, Kenoyer J and Meadow R. New Evidence for Early Silk in the Indus Civilization. *Nature Precedings* (2008).
- Gupta A, Anderson D, Pandey D and Singhvi A. Adaptation and human migration, and evidence of agriculture coincident with changes in the Indian summer monsoon during the Holocene. *Curr. Sci.* 2006; 60: 1082-1090.

- Indian Genome Variation Consortium. Genetic landscape of the people of India: a canvas for disease gene exploration. *J. Genet.* 2008; 87: 3-20.
- International HapMap Consortium. A haplotype map of the human genome. *Nature.* 2005; 437: 1299-1320.
- Jakobsson M, Scholz SW, Scheet P, Gibbs JR, VanLiere JM, Fung HC, Szpiech ZA, Degnan JH, Wang K, Guerreiro R, Bras JM, Schymick JC, Hernandez DG, Traynor BJ, Simon-Sanchez J, Matarin M, Britton A, van de Leemput J, Rafferty I, Bucan M, Cann HM, Hardy JA, Rosenberg NA and Singleton AB. Genotype, haplotype and copy-number variation in worldwide human populations. *Nature.* 2008; 451: 998-1003.
- James H and Petraglia M. Modern human origins and the evolution of behavior in the Late Pleistocene record of South Asia. *Curr. Anthropol.* 2005; 46: S3-S27.
- Jarrige JF. Chronology of the earlier periods of the Greater Indus as seen from Mehrgarh, Pakistan, in: Allchin, B. (Ed.), *South Asian Archaeology*. Cambridge University Press, Cambridge, 1981: pp. 21-28.
- Kennedy KAR. *God Apes and Fossil Men: Palaeoanthropology in South Asia* University of Michigan Press, Ann Arbor (2000).
- Kenoyer JM. (Ed.) *Ancient cities of the Indus Valley civilization* Oxford University Press (1998).
- Lahr M and Foley R. Multiple dispersals and modern human origins. *Evol. Anthropol.* 1994; 3: 48-60.
- Lawler A. Unmasking the Indus. Indus collapse: the end or the beginning of an Asian culture? *Science.* 2008; 320: 1281-1283.
- Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, Cann HM, Barsh GS, Feldman M, Cavalli-Sforza LL and Myers RM. Worldwide human relationships inferred from genome-wide patterns of variation. *Science.* 2008; 319: 1100-1104.
- Madella M and Fuller D. Palaeoecology and the Harappan Civilisation of South Asia: A reconsideration. *Quaternary Science Reviews.* 2006; 25: 1283-1301.
- McElreavey K and Quintana-Murci L. A population genetics perspective of the Indus Valley through uniparentally-inherited markers. *Ann. Hum. Biol.* 2005; 32: 154-162.
- Mellars P. Going east: new genetic and archaeological perspectives on the modern human colonization of Eurasia. *Science.* 2006; 313: 796-800.
- Mellars P, Gori KC, Carr M, Soares PA and Richards MB. 2013. Genetic and archaeological perspectives on the initial modern human colonization of southern Asia. *Proc. Natl. Acad. Sci. USA.* 2013; 110: 10699-10704.
- Metspalu M, Romero IG, Yunusbayev B, Chaubey G, Mallick CB, Hudjashov G, Nelis M, Mägi R, Metspalu E, Remm M, Pitchappan R, Singh L, Thangaraj K, Vilems R and Kivisild T. Shared and unique components of human population structure and genome-wide signals of positive selection in South Asia. *Am. J. Hum. Genet.* 2011; 89: 731-744.
- Mishra VN. Prehistoric human colonization of India. *J. Biosci.* 2001; 26: 491-531.
- Moorjani P, Thangaraj K, Patterson N, Lipson M, Loh PR, Govindaraj P, Berger B, Reich D, and Singh L. Genetic evidence for recent population mixture in India. *Am. J. Hum. Genet.* 2013; 93: 422-438.

- Patnaik R, Chauhan PR, Rao M, Blackwell B, Skinner A, Sahni A, Chauhan M and Khan H. New geochronological, paleoclimatological, and archaeological data from the Narmada Valley hominin locality, central India. *J. Hum. Evol.* 2009; 56: 114-133.
- Patnaik R and Chauhan P. India at the cross-roads of human evolution. *J. Biosci.* 2009; 34: 729.
- Petraglia MD and Allchin B. Human evolution and culture change in the Indian subcontinent, in: Petraglia, M.D., Allchin, B. (Ed.), *The evolution and history of human populations in South Asia*. Springer Verlag, 2007: pp. 464.
- Petraglia MD, Haslam M, Fuller DQ, Boivin N and Clarkson C. Out of Africa: new hypotheses and evidence for the dispersal of *Homo sapiens* along the Indian Ocean rim. *Ann Hum Biol.* 2010; 37: 288-311.
- Petraglia M, Clarkson C, Boivin N, Haslam M, Korisettar R, Chaubey G, Ditchfield P, Fuller D, James H, Jones S, Kivisild T, Koshy J, Lahr MM, Metspalu M, Roberts R and Arnold L. Population increase and environmental deterioration correspond with microlithic innovations in South Asia ca. 35,000 years ago. *Proc. Natl. Acad. Sci. USA.* 2009; 106: 12261-12266.
- Petraglia M, Korisettar R, Boivin N, Clarkson C, Ditchfield P, Jones S, Koshy J, Lahr MM, Oppenheimer C, Pyle D, Roberts R, Schwenninger JL, Arnold L and White K. 2007. Middle Paleolithic assemblages from the Indian subcontinent before and after the Toba super-eruption. *Science.* 2007; 317: 114-116.
- Powell A, Shennan S and Thomas MG. Late Pleistocene demography and the appearance of modern human behavior. *Science.* 2009; 324: 1298-1301.
- Quintana-Murci L, Chaix R, Wells RS, Behar DM, Sayar H, Scozzari R, Rengo C, Al-Zahery N, Semino O, Santachiara-Benerecetti AS, Coppa A, Ayub Q, Mohyuddin A, Tyler-Smith C, Qasim Mehdi S, Torroni A and McElreavey K. Where west meets east: the complex mtDNA landscape of the southwest and Central Asian corridor. *Am. J. Hum. Genet.* 2004; 74: 827-845.
- Rasmussen M, Guo X, Wang Y, Lohmueller KE, Rasmussen S, Albrechtsen A, Skotte L, Lindgreen S, Metspalu M, Jombart T, Kivisild T, Zhai W, Eriksson A, Manica A, Orlando L, De La Vega FM, Tridico S, Metspalu E, Nielsen K, Ávila-Arcos MC, Moreno-Mayar JV, Muller C, Dortch J, Gilbert MTP, Lund O, Wesolowska A, Karmin M, Weinert LA, Wang B, Li J, Tai S, Xiao F, Hanhara T, van Driem G, Jha AR, Ricaut FX, de Knijff P, Migliano AB, Gallego Romero I, Kristiansen K, Lambert DM, Brunak S, Forster P, Brinkmann B, Nehlich O, Bunce M, Richards M, Gupta R, Bustamante CD, Krogh A, Foley RA, Lahr MM, Balloux F, Sicheritz-Pontén T, Vilems R, Nielsen R, Wang J and Willerslev E. An Aboriginal Australian genome reveals separate human dispersals into Asia. *Science.* 2011; 334: 94-98.
- Reich D, Thangaraj K, Patterson N, Price AL and Singh L. Reconstructing Indian population history. *Nature.* 2009; 461: 489-494.
- Rightmire G. Comparison of Middle Pleistocene hominids from Africa and Asia, in: Barham, L., Robson-Brown, L., Robson-Brown, K. (Ed.), *Human Roots: Africa and Asia in the Middle Pleistocene*. Western Academic Press, Bristol, 2001: pp.123-135.
- Rosenberg NA, Mahajan S, Gonzalez-Quevedo C, Blum MG, Nino-Rosales L, Nini V, Das P, Hegde M, Molinari L, Zapata G, Weber JL, Belmont JW and Patel PI. Low levels of



- genetic divergence across geographically and linguistically diverse populations from India. *PLoS Genet.* 2006; 2: e215.
- Shah AM, Tamang R, Moorjani P, Rani DS, Govindaraj P, Kulkarni G, Bhattacharya T, Mustak MS, Bhaskar LVKS, Reddy AG, Gadhvi D, Gai PB, Chaubey G, Patterson N, Reich D, Tyler-Smith C, Singh L and Thangaraj K. Indian siddis: african descendants with Indian admixture. *Am. J. Hum. Genet.* 2011; 89: 154-161.
- Shinde V, Deshpande S, Osada T, and Uno T. Basic Issues in Harappan Archaeology: Some Thoughts. *Ancient Asia.* 2009; 1: 63-72.
- Singh G. The Indus Valley culture (seen in the context of postglacial climate and ecological studies in Northwest India). *Archaeol. Phys. Anthropol. Oceania.* 1971; 6: 177-189.
- Sonakia A. The skull-cap of Early Man and Associated Mammalian Fauna from Narmada Valley Alluvium, hoshangabad Area, Madhya Pradesh (India). Records of the Geological Survey of India, 1984: 159-172.
- Thangaraj K, Chaubey G, Kivisild T, Reddy AG, Singh VK, Rasalkar AA and Singh L. Reconstructing the origin of Andaman Islanders. *Science.* 2005; 308: 996.
- Tripathi J, Tripathi K, Bock B, Rajamani V and Eisenhauer A. "Is River Ghaggar, Saraswati? Geochemical Constraints" *Curr. Sci.* 2004; 87(8); 1141-1145.
- Underhill PA and Kivisild T. Use of Y Chromosome and Mitochondrial DNA Population Structure in Tracing Human Migrations. *Annu. Rev. Genet.* 2007; 41: 539-564.
- Underhill PA, Shen P, Lin AA, Jin L, Passarino G, Yang W, Kauffman E, Bonne-tamir B, Bertranpetit J, Francalacci P, Ibrahim M, Jenkins T, Kidd J, Mehdi S, Seielstad M, Wells R, Piazza A, Davis R, Feldman M, Cavalli-Sforza LL and Oefner PJ. Y chromosome sequence variation and the history of human populations. *Nat. Genet.* 2000; 26: 358-361.
- Xing J, Watkins WS, Hu Y, Huff CD, Sabo A, Muzny DM, Bamshad MJ, Gibbs RA, Jorde LB, and Yu F. Genetic diversity in India and the inference of Eurasian population expansion. *Genome Biol.* 2010; 11: R113.
- Xing J, Watkins WS, Witherspoon DJ, Zhang Y, Guthery SL, Thara R, Mowry BJ, Bulayeva K, Weiss RB and Jorde LB. Fine-scaled human genetic structure revealed by SNP microarrays. *Genome Res.* 2009; 19: 815-825.