Ind. J. Phys. Anthrop. & Hum. Genet. Vol. 34. No. 2, (2015) : 301-307

GENETIC BASIS OF UNITY
(Early Palaeolithic Antiquity and Continuity of the Contemporary Indian Populations)

Vadlamudi Raghavendra Rao*

ABSTRACT

Out of Africa expansion of anatomically modern humans and Palaeolithic continuity of the present day Indian populations is the most parsimonious explanation as of now, based on maternal and paternal haploid DNA lineages and high density autosomal DNA markers. Archaeological dating corroborate with DNA clock of expansion to about 1,60,000 ybp (years before present). The route of this expansion, whether north via Levant or south via horn of east Africa to Indian coast to east- Asia to euro-Asia is an intense debate. However, large amount of empirical data from India generated on complete mtDNA sequences of more than 3000 samples from 37 tribal populations by Anthropological Survey of India, irrevocable in support of the southern route. Now, it is increasingly believed that the southern-route is the only expansion by which modern humans moved out of Africa and peopled all other non- African continents. Further support to this tantalizing proposition, even pointing out that Indian populations have ancestral foot prints to Chinese, has come from high density DNA mapping of large number of populations from Asia- Pacific region by international collaborative study.

INTRODUCTION

While situating people of India as representatives of earliest anatomically modern human expansion that resulted in the peopling of the world, the Indian genetic data is also forth-with in explaining the socio-cultural and historical paradigms like Tribe, Caste, language categories, large scale north-west Indo-Aryan invasion, etc. The results are emphatic that ancient genetic substratum and continuity resulted in sharing and Tribe-Caste continuum; language is later super-imposition and shifting of languages is a common phenomenon; large scale invasion which could have resulted in non-Indian genetic lineages in the hierarchical structure of Indian populations does not exist.

Vadlamudi Raghavendra Rao, Formerly Director-In-Charge, Anthropological Survey of India, Kolkata; Professor of Anthropology, University of Delhi, Delhi and Presently "ICMR Emeritus Scientist, Dept of Genetics, Osmania University, Hyderabad". Research Professor (Hon.), Genome Foundation, Hyderabad, INDIA, *E-mail: profraovr@gmail.com*

The article is based on lectures delivered at Ramakrishna Mission and Asiatic Society, Kolkata.

DNA dating for Palaeolithic continuity in Indian scenario starts from 60-65 kybp (thousand years before present) and glacial, inter-glacial climatic fluctuations could have largely affected the ancient anatomically modern humans largely surviving on hunter-gatherer subsistence. We do have populations like Jarawas, Onge and Sentinelese of Andaman and Nicobar Islands as direct descendents of earliest human expansion, while all other populations share this ancient substratum.

HOLOCENE SCENARIO

Continuity of Palaeolithic substratum into Holocene starting from about 13,000 ybp to the present, that witnessed domestication of plants/animals, agricultural expansion, more arid post glacial climatic conditions and the onset of monsoons, has brought in tremendous changes in the genetic profile of Indian populations, especially with respect to change from protein diet of hunter gatherers to starch metabolism as a result of agriculture. The copy number variation of 'amylase gene' for starchy food and regulatory sequence variation of 'lactase gene' for milk consumption are well studied genetic paradigms. Even the genes responsible for present day non-communicable diseases like diabetes, coronary heart diseases, cancer etc have a bearing on the ecological-environmental and cultural conditioning of Indian populations in the Holocene. The most significant outcome of this period is geographical deep rooting of individual populations, with far reaching medical implications, still not realised by the Indian Scientific community, while west has already realised the commercial importance of this human genetic resource.

Without delimiting the medical importance of Holocene genetic architecture of Indian populations, the objective of present presentation is to re-construct the genetic history of people of India during Holocene, particularly with reference to the existing paradigms:

1) Large scale migration from Central Asia or Western Europe pushing the already existing populations to south thereby crediting the Indus Valley Civilization to this exodus.

2) Related issue of paramount contemporary social importance i.e independent origin of Tribes and Castes.

A deeper understanding of these paradigms require corroborating evidences in the form of scientific dating of ancient events from fields like astronomy, geology, archaeology, flora, fauna, folk tales, etc. and Genetics has provided an intense supportive evidence in view of its recent DNA technology with the advantage of coalescence dating and high density markers.

In the present paper, as it is meant for experts from other fields than Genetics, I intend to first explain the DNA technology as is being used for reconstructing the past (archaeogenetics) in simple terms, then proceed with examining the recent findings that is available in reputed scientific journals contributing scientific evidences for a debate on the above mentioned two paradigms. While examining

the DNA evidence addressing issues that are of Holocene period (13000 ybp–present), I have briefly touched upon the existing 'Out of Africa Expansion' theory in view of the fact that this theory is crucial in understanding the peopling not only India, but also all other continents, except Africa.

DNA LINEAGES AND DATING

DNA is the blue print of life. All life forms are composed of DNA. It is just a chemical molecule that has an inbuilt mechanism of self replication and codes for proteins, the building blocks of life. The DNA is composed of nucleotides and each nucleotide is a composition of a nucleotide base (commonly known as bases, A, T, G, C), a sugar and phosphate molecules. In Humans, it is found that there are 3 billion pairs of these, A, T, G, Cs. Ultimately the individual uniqueness is in terms of particular arrangement of these 3 billion base pairs. For example, if we consider (hypothetical) there are 10 nucleotides, then A, C, T, T, T, C, C, A, G, G could be one combination (array), that defines one individual, where as G, T, A, A, G, C, C, T, C, G could be another combination that defines another individual. Like that we will have 10 to the power 10 number of combinations with each defining an individual (uniqueness). Imagine the immense possibility of this uniqueness if we consider 3 billion. During the life time of an individual, this combination of nucleotides does not change. But, when he or she passes on this combination to his or her child, the child will inherit a further combination (reshuffled) of these nucleotides from two sources i.e. father and mother, a further source of variation and uniqueness. Not only reshuffling but also there can be changes (mutation) introduced as the DNA passes from generation to generation. It is estimated that approximately there will be one nucleotide change in 2.2 million years. That much is the conservation of the genetic code. This is the principle on which *DNA dating* is based.

In the reconstruction of ancestry, it is not possible to follow one change (mutation) from generation to generation if reshuffling due to mating (parental) occurs. For that reason, scientists found DNA that does not participate in reshuffling, but passed on from generation to generation, as the changes (mutations) go on accumulating. These DNA are called haploid markers or non-recombinant DNA. In Humans we have two haploid marker systems that are called a) mtDNA (mitochondrial DNA) that passes on from mother to all children b) non-recombinant portion of Y chromosome (nY), that passes on from father to sons only. Hence, we have a system which can infer maternal inheritance (mtDNA) and paternal inheritance (nY).

mt DNA is 16569 nucleotides (also know as base pairs) long and nY a million. By sequencing (determining the exact combination, array of nucleotides also know as haplotype), it is possible to determine an individual specific array, which is called *DNA lineage*. So, we have literally maternal and paternal lineages depending upon which DNA we are using to determine lineages. Suppose we have an ancestral lineage, in which down the time line (say 2.2 million years) a mutation occurred (mutation A) and the lineage is called Haplogroup A (Hg A). Subsequently down

the line (2.2 million years) another mutation(mutation B) occurs and the lineage is for convenience called Hg B. Ultimately at a particular point in time (that is, suppose today) if we analyse mt DNA in a population, we may find some people with ancestral lineage, some with Hg A and some with Hg B. As far as dating is considered we can date the mtDNA lineage based on the number of mutations it has accumulated.

OUT OF AFRICA HYPOTHESIS AND THE ROUTE OF EXPANSION

Rebecca Cann and her co-workers (1987), based on the analyses of mtDNA sampled from contemporary populations from across the world found more diversity (accumulation of more mutations) in African populations, than other continental populations. This finding was further supported by more empirical data by Watson *et al.,* (1997) and the hypothesis that a small number of anatomically modern humans possessing mtDNA, L3 lineage migrated out of Africa around 1,60,000 ybp, and expanded forming into continental populations of Asia, Europe, Americas was formulated. Several studies from various populations all over the world with astounding massive empirical data of both mtDNA (maternal) and Y (paternal) markers, support postulated ancient migration histories of populations based on mutations and further supported the hypothesis, which is now famously known as 'Out of Africa' hypothesis.

The route of the out of Africa expansion to other continents, whether via Europe, Mediterranean to Asia or via horn of east Africa to South Asia by coastal migration to western, central and eastern Asia to Europe is an intense debate even now.

However, extensive population data on Indian Tribal populations by the Anthropological Survey of India publications (Barik *et al.,* 2008; Chandrasekhar *et al.,* 2009; Satish *et al.,* 2009) are equivocal in terms of deep rooted 'in situ' origin of mtDNA lineages and especially the scientific evidence in support of Andaman populations as the direct descendents of out of Africa expansion and sharing of DNA lineages between Indian main land tribes and Australian Aborigines.

The most forthcoming evidence for the importance of South Asia as the corridor of ancient human expansion is from South Asia Pacific Consortium study (HUGO Pan-Asian SNP Consortium *et al.,* 2009), wherein with advanced high density markers, it is conclusive that Indian sub-continent is the place from where other continents excluding Africa were populated.

CHALLENGING THE EXISTING PARADIGMS

Paradigm 1: Large scale migration from Central Asia or Western Europe pushed the already existing populations to south thereby crediting the Indus Valley Civilization to this exodus', is being questioned based on the recent DNA studies from Central Asia, Western Europe, Indian sub-continent and other continental populations.

The notable studies and their conclusions in this respect are:

(a) *Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages* (Kivisild *et al.*, 1999) study summarised that 'the extensive deep late Pleistocene genetic link between contemporary Europeans and Indians, provided by the mtDNA haplogroup U, which encompasses roughly a fifth of mtDNA lineages of both populations estimated to be split and close to the suggested time for the peopling of Asia and the first expansion of anatomically modern humans in Eurasia (Kivisild *et al.*, 1999; Kivisild *et al.*, 2000; Barik *et al.*, 2008; Satish *et al.*, 2009) and likely pre-dates their spread to Europe'.

(b) *An Indian Ancestry: a Key for Understanding Human Diversity in Europe and Beyond* (Kivisild *et al.*, 2000) concluded that *'Through the analyses of about 1000 mtDNA genomes and 400 Y chromosomes from various locations in India we reached the following conclusions, relevant to the peopling of Europe in particular and of the Old World in general. First, we found that the node of the phylogenetic tree of mtDNA, ancestral to more than 90 per cent of the present-day typically European maternal lineages, is present in India at a relatively high frequency. Inferred coalescence time of this ancestral node is slightly above 50,000 BP. Second, we found that haplogroup U is the second most abundant mtDNA variety in India as it is in Europe. Summing up, we believe that there are now enough reasons not only to question a 'recent Indo-Aryan invasion' into India some 4000 BP, but alternatively to consider India as a part of the common gene pool ancestral to the diversity of human maternal lineages in Europe. Our results on Y-chromosomal diversity of various Indian populations support an early split between Indian and east of Indian paternal lineages, while on a surface, Indian (Sanskrit as well as Dravidic speakers) and European Y-chromosomal lineages are much closer than the corresponding mtDNA variants'.*

(c) Polarity and Temporality of High-Resolution Y-Chromosome Distributions in India Identify Both Indigenous and Exogenous Expansions and Reveal Minor Genetic Influence of Central Asian Pastoralists (Sengupta *et al.*, 2006) opined that 'using high-resolution data on 69 informative Y-chromosome binary markers and 10 microsatellite markers from a large set of geographically, socially, and linguistically representative ethnic groups of South Asia, we found that the influence of Central Asia on the pre-existing gene pool was minor. The ages of accumulated microsatellite variation in the majority of Indian haplogroups exceed 10,000–15,000 years, which attests to the antiquity of regional differentiation'.

Paradigm 2: Independent Origin of Tribes and Castes is being questioned by DNA studies that showed sharing of lineages among tribes and caste populations across different geographical and linguistic categories.

(a) The Genetic Heritage of the earliest settlers persists both in Indian Tribal and Caste populations (Kivisild *et al.*, 2003). In mtDNA phylogenetic analyses, the Chenchu and Koyas coalesce at Indian-specific branches of haplogroups M and N that cover populations of different social rank from all over the continent.

Haplotype R1a, previously associated with the putative Indo-Aryan invasion, was found at its highest frequency in Punjab and also at a relatively high frequency (26%) in the Chenchu tribe.

(b) The tribes of south extending to central Indian region and tribes of eastern region harbour mtDNA M2 lineages, M2a and M2b with coalescent time depth of 37000 years before present (Kumar *et al.,* 2008).

These findings, taken together show that Indian tribal and caste populations derive largely from the same genetic heritage of Pleistocene and have received limited gene flow from external regions.

CONCLUSION

To conclude, the antiquity of contemporary Indian populations is evidenced by the fact that Indian mtDNA diversity is large next only to African populations. Several populations in main land share major mtDNA lineage M, which is one of the two main lineages at the root of mtDNA L3 and they are deep rooted in Indian sub-continent with coalescent dates of 65-70 thousand years ago. The continuity of this antiquity is evidenced by sharing of mtDNA M31 lineage of the Andaman tribes with tribes representing ancient substratum in the contemporary Indian populations. Further, the DNA evidence is forth coming in questioning the existing paradigms of Indo-Aryan invasion and independent origin of Tribes and castes.

The article is based on the authors' contributions in seminars on 'Ancient Dates' held at Delhi and on 'Cultural Unity' at Kolkata.

ACKNOWLEDGEMENTS

The author acknowledges support from Ministry of Culture in implementing a National Project on DNA Polymorphisms while heading the Anthropological Survey of India, Kolkata, INDIA.

REFERENCES

Barik SS, Sahani R, Prasad BVR, Endicott P, Metspalu M, Sarkar BN, Bhattacharya S, Annapoorna PCH, Sreenath J, Sun D, Sanchez JN, Ho SYW, Chandrasekar A and Rao VR. Detailed mtDNA Genotypes Permit a Reassessment of the Settlement and Population Structure of the Andaman Islands. *American Journal of Physical Anthropology.* 2008; 136: 19-27.

Cann RL, Stoneking M and Wilson AC. Mitochondrial DNA and human evolution. *Nature.* 1987; 325: 31-36.

Chandrasekar A, Kumar S, Sreenath J, Sarkar BN, Urade BP, Mallick S, Bandopadhyay SS, Barua P, Barik SS, Basu D, Kiran U, Gangopadhyay P, Sahani R, Prasad BVR, Gangopadhyay S, Lakshmi GR, Ravuri RR, Padmaja K, Venugopal PN, Sharma MB and Rao VR. Updating phylogeny of mitochondrial DNA macrohaplogroup m in India: dispersal of modern human in South Asian corridor. PloS one. 2009; 4(10): e7447.

HUGO Pan-Asian SNP Consortium, Abdulla MA, Ahmed I, Assawamakin A, Bhak J, Brahmachari SK, Calacal GC, Chaurasia A, Chen CH, Chen J, Chen YT, Chu J, Cutiongco-de la Paz EMC, De Ungria MCA, Delfin FC, Edo J, Fuchareon S, Ghang H, Gojobori T, Han J, Ho SF, Hoh BP, Huang W, Inoko H, Jha P, Jinam TA, Jin L, Jung J, Kangwanpong D, Kampuansai J, Kennedy GC, Khurana P, Kim HL, Kim K, Kim S, Kim WY, Kimm K, Kimura R, Koike T, Kulawonganunchai S, Kumar V, Lai PS, Lee JY, Lee S, Liu ET, Majumder PP, Mandapati KK, Marzuki S, Mitchell W, Mukerji M, Naritomi K, Ngamphiw C, Niikawa N, Nishida N, Oh B, Oh S, Ohashi J, Oka A, Ong R, Padilla CD, Palittapongarnpim P, Perdigon HB, Phipps ME, Png E, Sakaki Y, Salvador JM, Sandraling Y, Scaria V, Seielstad M, Sidek MR, Sinha A, Srikummool M, Sudoyo H, Sugano S, Suryadi H, Suzuki Y, Tabbada KA, Tan A, Tokunaga K, Tongsima S, Villamor LP, Wang E, Wang Y, Wang H, Wu JY, Xiao H, Xu S, Yang JO, Shugart YY, Yoo HS, Yuan W, Zhao G, Zilfalil BA Indian Genome Variation Consortium. Mapping human genetic diversity in Asia. Science. 2009;326 (5959):1541-1545. Kivisild T, Bamshad MJ, Kaldma K, Metspalu M, Metspalu E, Reidla M, Laos S, Parik J, Watkins WS, Dixon ME, Papiha SS, Mastana SS, Mir MR, Ferak V and Villems R. Deep common ancestry of Indian and western-Eurasian mitochondrial DNA lineages. Current Biology. 1999; 9(22): 1331-1334.

Kivisild T, Papiha SS, Rootsi S, Parik J, Kaldma K, Reidla M, Laos S, Metspalu M, Pielberg G, Adojaan M, Metspalu E, Mastana SS, Wang Y, Golge M, Demirtas H, Schnakenberg E, de Stefano GF, Geberhiwot T, Claustres M and Villems R. An Indian Ancestry: a key for understanding Human Diversity in Europe and Beyond. In Archaeogenetics: DNA and population prehistory of Europe. Edt by Colin Renfew and Katie Boyle, Published by Mc Donald Institute for Archaeological Research, Cambridge, 2000:pp 267-275.

Kivisild T, Rootsi S, Metspalu M, Mastana S, Kaldma K, Parik J, Metspalu E, Adojaan M, Tolk HV, Stepanov V, Gölge M, sanga E, Papiha SS, Cinnioğlu C, King R, Cavalli-Sforza L, Underhill PA and Villems R. The genetic heritage of the earliest settlers persists both in Indian tribal and caste populations. *Am. J. Hum. Genet.* 2003; 72(2): 313-332.

Kumar S, Padmanabham PB, Ravuri RR, Uttaravalli K, Koneru P, Mukherjee PA, Das B, Kotal M, Xaviour D, Saheb SY and Rao VR. The earliest settlers' antiquity and evolutionary history of Indian populations: evidence from M2 mtDNA lineage. BMC Evol. Biol. 2008; 8(1): 230.

Kumar S, Reddy R, Koneru P, Urade BP, Sarkar BN, Chandrasekar A, Rao VR. Reconstructing Indian-Australian Phylogenetic link. BMC Evolutionary Biology. 2009:9;173. Sengupta S, Zhivotovsky LA, King R, Mehdi SQ, Edmonds CA, Chow CE, Lin AA, Mitra M, Sil SK, Ramesh A, Usha Rani MV, Thakur CM, Cavalli-Sforza LL, Majumder PP, Underhill PA. Polarity and temporality of high resolution Y-chromosome distribution in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian Pastoralists. *Am. J. Hum. Genet.* 2006; 78: 202-221.

Watson E, Forster P, Richards M and Bandelt HJ. Mitochondrial footprints of human expansions in Africa. *Am. J. Hum. Genet.* 1997; 61: 691-704.