

GENETIC CONTINUITY OF ANATOMICALLY MODERN HUMAN BETWEEN INDIA AND ISLAND SOUTHEAST ASIA ISEA: LAST GLACIAL DISPERSAL OF MTDNA LINEAGE N22

Isukapatla Arjun Rao^{1*}, Venugopal Pulamaghatta², Adimoolam Chandrasekar²
¹ *Department of Forensic Science, Guru Ghasidas Vishwavidyalaya, Bilaspur C.G., India*
² *Anthropological Survey of India, Southern Regional Centre, Mysore, India*
E mail: arjun4n6@gmail.com

ABSTARCT:

Our complete sequencing of 220 mtDNA genomes from the Savara and Porja of east coastal India reveals about 25 per cent genomes belongs to European macro haplogroup N. For the first time we identified mitochondrial DNA from one south Indian Savara individual that shares seven specific mutations with the N22 lineage observed in the Orang Asli group of Aboriginal Malaya, Cuyonin from Palawan and one single mutation sharing with Mindanao of Philippines which forms N22b sub-lineage. The coalescence time of N22 lineage is 20,600 ± 7000 kybp shifts the focus of “Two layer” hypothesis of human settlement in ISEA.

Keywords: Human Settlement, mtDNA, Savara, Porja, Phylogenetic

1. INTRODUCTION:

Anatomically modern human settlement in Island Southeast Asia (ISEA) is hypothesized by “two-layer” Settlement model of Australo-Melanesians ~ 50,000 years ago followed by “Mongoloid” Austranesians ~ 4,000 years ago – clearly does not capture the complexity of demographic history in the region. The substantial genetic data from mitochondrial DNA lineages dating between ~ 15,000 years and 5,000 years ago brought considerable shift in the focus of prehistory of ISEA (Hill et al., 2007). From our data set search we find a new evidence for the last glacial dispersal (18,000 to 15,000 yrs ago) of modern humans from South Asia to ISEA. However, Southeast Asian populations and Indigenous Australians acquire all deep rooted clades of both haplogroups M and N, there was a view that, only a single dispersal from Africa (Macaulay et al., 2005) based on the skeletal evidence the modern human populations dispersed out rapidly along the coastlines of southern Asia, south-east Asia and Indonesia to Australia (Connell et al., 2004; Stringer et al., 2000) at least by 45 thousand years before present (KYBP), best represented by the anatomically modern human skeleton from the site of Lake Mungo 3 in New South Wales (Field et al., 2006; Bowler et al., 2003).

An early phylogenetic link between Indians and Malaysian Aborigines has also been suggested by observations not only based on morphology but also culturally and linguistically (Lim et al. 2010; Fix 1995). Genetic admixture between Indians and Malaysians happened during the early colonization of Malays (Hatin et al., 2011) and was first shown by using some biochemical markers (Teng and Tan 1979). However, disperse of modern human from out of Africa to other parts of the world has never been clearly resolved. In the out of Africa scenario at present from the mtDNA lineage-analysis patterns the strong evidence is southern route hypothesis, across the mouth of the Red Sea, and then either northward through Arabia or eastward along the South Asian coastline to Australasia. (Metspalu et al., 2004; Quintana-Murci et al., 2004; Forster et al., 2005; Kivisild et al., 2006; Mellars 2006).

The mtDNA variation shows strong evidence for indigenous origins of the Orang Asli within the Malay Peninsula dating back to ~ 60,000 years ago – probably within only a few thousand years of the dispersal from east Africa. N21 and N22 which appear to be largely restricted to aboriginal Malaya which may represent gene flow among South-east Asia and suggests a large offshore component, dating back to mid/late Holocene in the ancestry of the aboriginal Malaya. Temuan and Orang Asli population of the Malay Peninsula, Cuyonin

(Palawan non-Negrito) from Palawan also belongs to N22 Clade which again is rare but more diverse in Indonesia and Philippines (Table 1) (Macaulay et al., 2005, Hill et al., 2006; 2007; Tabadda et al., 2010, Scholes et al., 2011).

We analyzed 220 unrelated Austro-Asiatic and Dravidian linguistic populations of coastal belt of Andhra Pradesh, India (i) Savara, who live in hilltop regions with flora and fauna of Srikakulam, Andhra Pradesh. Traditionally shifting cultivation, agricultural wage labor, Hunting, Fishing and the collection of minor forest produce are their sources of livelihood; (ii) Porja, who depends upon the shift cultivation and food gathering. The family is the smallest social unit with both patriarchal and patrilineal descent. The marriages are often polygamous among them.

2. MATERIAL AND METHODS

The Indian populations are organized into 4365 communities (Singh KS 1997), which include self defined castes, tribes and religious groups. About 450 tribes constitute 8.08% by 2001 census of the total Indian population. They speak more than 750 dialects, which can be broadly classified into Austro-Asiatic, Dravidian, Tibeto-Burman and Indo-European language families. The tribes are endogamous in nature and socio-culturally distinct. They inhabit mostly in the forests and hilly terrain areas. Government of India has notified 75 tribes as the most primitive group among the original inhabitants of India. Out of 75 tribal groups, we selected Savara and Porja tribes, inhabiting the southern part of India (Figure 1), representing Austro-Asiatic and Dravidian linguistic family and collected 250 blood samples from two communities for the present study.

The Ethical Committee of the Anthropological Survey of India had approved the project, 5–8 ml of blood was drawn from healthy and unrelated individuals after obtaining written consent. Samples were collected in Vacutainer as per standard protocols, and extraction of DNA was performed according to the enzymatic extraction procedure followed by phenol purification, which was standardized at Anthropological Survey of India, Southern Regional Centre DNA laboratory, Mysore.

3. QUALITY CONTROL

Out of 250 samples, 220 samples were selected for the present study. Sequence reactions were carried out with a Big-Dye terminator cycle sequencing FS ready reaction kit (Applied Bio-systems) to produce even signal intensities and to reduce false negatives. It enabled more accurate automated mixed base identification. Sequencing data that were generated on Applied Bio-systems 3730 DNA analyzer were analyzed in SeqScape software V 2.5. KB base caller V 1.4 was used in the analysis protocol. KB base caller process fluorescence signal assigns a base to each peak and assigns quality value (QV) to each base. The QV predicts the probability of a base call error. KB base caller generated QV from 1 to 99. Typically high quality pure bases will have QV ranging from 20–50 (Probability of Error is 1% to 0.001%). Mixed bases were identified if the secondary peak height threshold value was 25%. To set clear range of the sequence quality value method (Remove base from the ends until fewer than 4 bases out of 20 have QVs 20) was used. Filter setting values used were: Maximum mixed bases = 20, Minimum sample score = 25. Depending on the sequence quality and the criteria specified for filtering the data prior to assembly, the samples were not assembled. These unassembled samples were re-sequenced until it satisfied the quality. Editing of data and scoring of mutations were done by two independent groups of researchers. The Phylogenetic tree was reconstructed from median joining networks rooted L3 using network 4.6.1.1 software (Bandelt et al., 1999).

Complete mtDNA genome sequence generated in this study was submitted to GenBank (accession number KC461930).

4. RESULTS AND DISCUSSION

For the first time we find sharing of macro haplogroup N, lineage N22 specific for ISEA with India from the present study (Figure 2). The other major branches in the region are N21, R22 and R9. N22 mostly constrained to Aboriginal Malaya's (Hill et al., 2006) and Philippines (Tabadda et al., 2010). Our complete mtDNA sequencing of 220 individuals from 2 relic populations of India identified one individual sharing seven specific mutations **C150T, A942G, A7158G, A9254G, T11365C, C16168T, T16249C** with the N22 haplogroup, which is specific to Orang Asli group of Malaya (Jinam et al., 2012 and Macaulay et al., 2005), Mindanao Philippines (Tabadda et al., 2010) and Cuyonin Palawan (Scholes et al., 2011) (Fig 2).

Further a sample of Savara, an Indian Tribe shares all the basal seven mutations of N22 and a mutation at nucleotide position 13980 with Mindanao sample of Philippines forming a new sub-lineage N22b. The coalescent time of N22 lineage is calculated as 20,600 ± 7000 years. The coalescent time of Indian and

Philippine lineage N22b is dates back to 15400 ± 6300 yrs (Table I & Figure 2). N22 lineage dispersal represents the last glacial human migrations in south and South East Asia. The coalescence age estimates were calculated as per (Kivisild et al., 2006) using sequence positions between nucleotides 577 to 16023 considering one base substitution per 5,140 years, excluding insertions and deletions.

The study supports the hypothesis rather simple “two-layer” settlement of ISEA last glacial dispersal from South Asia to ISEA has to be considered for better understanding the demographic history of ISEA. The earliest AMH expansion through South East Islands from India documented in mtDNA macro-haplogroup M, lineages M42 Australian (Kumar et al., 2009) and M31 Andaman Islanders (Barik et al., 2008). Further our results showed a shared mtDNA lineage N22 between Indian and ISEA provides direct genetic evidence that AMH colonized through South East Asia through multiple migrations in concurrence with the recent analyses of skeletal remains and burial patterns in the region, which stress heterogeneity rather than an abrupt transition (Detroit et al., 2004). The last and post glacial dispersal lineages relate more closely to those of the Southeast Asian mainland than to those of modern aboriginal Australians and New Guineans. The most likely driving forces behind such re-dispersals are the huge sea-level rises that flooded much of Greater Sunda land, reducing it to the present day archipelago. The present study supports the earlier genetic data reported by (Hill et al., 2007) indicating a considerable shift of focus and a broadening of perspective with regard to the prehistory of ISEA.

5. CONCLUSION

The reconstructed mtDNA lineage N22 by complete mtDNA sequencing provides a genetic continuity between mainland India and Island South East Asia. Our result provides direct genetic evidence that ISEA was inhabited by modern humans through south Asia following the “Southern Route”.

Table 1.
Coalescence age of N22 haplogroup from ISEA.

Sample	Origin	Haplogroup	Coalescence age
Cuyonin (Non-negrato) (Scholes et al. 2011)	Palawan	N22	18000 ± 6100 kybp
Orang Asli (Jinam et al. 2012)	Malaya Peninsula	N22a	19500 ± 7400 kybp
Orang Asli (Macaulay et al. 2005)	Malaya Peninsula	N22a	19500 ± 7400 kybp
Mindanao (Tabadda et al. 2010)	Philippines	N22b	15400 ± 6300 kybp
Savara (Present Study)	India	N22b	15400 ± 6300 kybp



Fig 1: N22 lineage distribution in India and around Island Southeast Asia: ISEA.

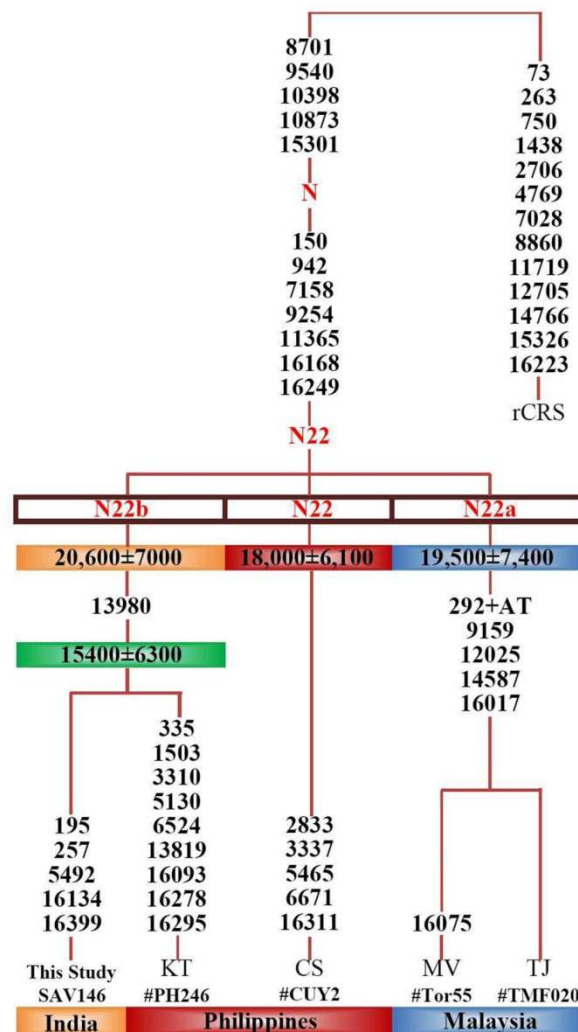


Fig 2

Fig. 2: Phylogenetic tree re-constructed on the basis of complete N22 mtDNA sequences from the Savara of South India. Samples collected from published sources were referred by symbols PH (Tabbada et al., 2010), CUY (Scholes et al., 2011) and Tor, TMM, TMF (Hill et al., 2006) followed by “#” and the original sample code. Suffixes A, C, G and T indicate transversions; ‘d’ denotes deletion and plus sign (+) indicates an insertion, since the variation at 16519 is extremely hypervariable and so not shown here. The coalescent age estimates calculated as per (Kivisild et al. 2006).

ACKNOWLEDGEMENT

This work is essentially a part of the Anthropological Survey of India's national project "DNA Polymorphisms in Contemporary Indian Populations" we express our gratitude to the Ministry of Culture, Government of India for supporting the project. We are thankful to a large number of anonymous subjects from different parts of India who voluntarily participated in this study and provided blood sample.

REFERENCES:

- [1] Bandelt, HJ.; Forster, P.; Rohl, A. (1999): Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*, **16**, pp. 37-48.
- [2] Barik, SS.; Sahani, R.; Prasad, BVR.; Endicott, P.; Metspalu, M.; Bhattacharya, S.; Sarkar, BN.; Annapoorna, PCH.; Sreenath, J.; Sun, D.; Sanchez, JJ.; Ho, SYW.; Chandrasekar, A.; Rao, VR. (2008): *Am J Phys Anthropol*. **136**, pp. 19-27.
- [3] Bowler, JM.; Johnston, H.; Olley, JM.; Prescott, JR.; Roberts, RG.; Shawcross, W.; Spooner, NA. (2003): New ages for human occupation and climatic change at Lake Mungo, Australia. *Nature*. **421**, pp. 837-840.

- [4] De'troit, F.;Dizon, E.;Falgue`res, C.;Hameau, S.;Ronquillo, W.; Se´mah, F. (2004): Upper Pleistocene Homo sapiens from the Tabon cave (Palawan, The Philippines): description and dating of new discoveries. *Human Palaeontology and Prehistory*. **3**, pp. 705-712.
- [5] Field, JS.; Lahr, MM. (2006): Assessment of the Southern Dispersal: GIS-Based Analyses of Potential Routes at Oxygen Isotopic Stage 4. *J World Prehist*. **19**, pp. 1-45.
- [6] Fix, AG. (1995): Malayan paleosociology: implications for patterns of genetic variation among the Orang Asli. *Am. Anthropol*.**97**(2) pp. 313-23.
- [7] Forster, P.; Matsumura, S. (2005): Enhanced: Did Early Humans Go North or South? *Science*. **308**, pp. 965-966.
- [8] Hatin, WI.;Nur-Shafawati, AR.;Zahri, M.;Xu, S.; Jin, L.; Tan, S.;Rizman-Idid, M.;Zilfalil, BA.; The HUGO Pan-Asian SNP Consortium. (2009): Population Genetic Structure of Peninsular Malaysia Malay Sub-Ethnic Groups. *Plos One*. **6**(4):e18312.
- [9] Hill, C.;Soares, P.;Mormina, M.; Macaulay, V.; Meehan, W.; Blackburn, J.; Clarke, D.; Raja, JM.; Ismail, P.;Bulbeck, D.; Oppenheimer, S.; Richards, M.(2006):Phylogeography and ethnogenesis of aboriginal Southeast Asians. *MolBiolEvol*.**23**, pp.2480-2491.
- [10] Hill, C.;Soares, P.;Mormina, M.; Macaulay, V.; Clarke, D.;Blumbach, PB.;Vizueté-Forster, M.; Forster, P.;Bulbeck, D.; Oppenheimer, S.; Richards, M. (2007). A mitochondrial stratigraphy for island southeast Asia. *Am J Hum Genet*.**80**, pp. 29-43.
- [11] Jinam, TA.; Hong, LC.;Phipps, ME.;Stoneking, M.;Ameen, M.;Edo, J.; HUGO Pan-Asian SNP Consortium.;Saitou N. (2012). Evolutionary history of continental Southeast Asians: "early train" hypothesis based on genetic analysis of mitochondrial and autosomal DNA data. *MolBiolEvol*. **29** (11), pp. 3513-3527.
- [12] Kumar, S.;Rajasekhara, RR.;Padmaja, K.;Urade, BP.;Sarkar, BN.; Chandrasekar, A.; Rao, VR. (2009). Reconstructing Indian-Australian phylogenetic link. *BMC Evo Bio*. **9**, pp. 173.
- [13] Kivisild, T.;Shen, P.; Wall, DP.; Do, B.; Sung, R.; Davis, K.;Passarino, G.; Underhill, PA.;Scharfe, C.;Torroni, A.;Scozzari, R.;Modiano, D.;Coppa, A.; de,Knijff, P.; Feldman, M.;Cavalli-Sforza, LL.;Oefner, PJ. (2006). The role of selection in the evolution of human mitochondrial genomes. *Genetics*.**172**, pp. 373-387.
- [14] Lim, LS.;Ang, KC.;Mahani, MC.;Shahrom, AW.;Md-Zain, BM. (2010). Mitochondrial DNA polymorphism and phylogenetic relationships of Proto Malays in Peninsular Malaysia. *Jour Biol Sci*.**10**, pp. 71-83.
- [15] Mellars, P. (2006). Going East: New Genetic and Archaeological Perspectives on the Modern Human Colonization of Eurasia. *Science*.**313**, pp. 796-800.
- [16] Metspalu, M.; Kivisild, T.;Metspalu, E.;Parik, J.;Hudjashov, G.;Kaldma, K.;Serk, P.;Karmin, M.; Behar, DM.; Gilbert, MTP.; Endicott, P.;Mastana, S.;Papiha, SS.;Skorecki, K.;Torroni, A.;Villemis, R. (2004). Most of the extant mtDNA boundaries in South and Southwest Asia were likely shaped during the initial settlement of Eurasia by anatomically modern humans. *BMC Genet*. **5**, pp. 26.
- [17] Macaulay, V.; Hill, C.;Achilli, A.; Rengo, C.; Clarke, D.; Meehan, W.; Blackburn, J.;Semino, O.;Scozzari, R.;Cruciani, F.;Taha, A.;Shaari, NK.; Raja,JM.; Ismail, P.;Zainuddin, Z.; Goodwin, W.;Bulbeck, D.;Bandelt, HJ.; Oppenheimer, S.;Torroni, A.; Richards, M. (2005). Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science*. **308**,pp. 1034-6.
- [18] O'Connell, JF.; Allen, FJ. (2004). Dating the colonization of Sahul (Pleistocene Australia-New Guinea):A review of recent research.*Jour Archaeol Sci*.**31**, pp. 835-853.
- [19] Quintana-Murci, L.;Chaix, R.; Wells, S.; Behar, D.;Sayar, H.;Scozzari, R.; Rengo, C.; Al-Zahery, N.;Semino, O.;Santachiara-Benerecetti, A.;Coppa, A.;Ayub, Q.;Mohyuddin, A.; Tyler-Smith, C.; Mehdi, Q.;Torroni, A.;McElreavey, K. (2004). Where West meets East: the complex mtDNA landscape of the Southwest and Central Asian corridor. *Am J Hum Genet*. **74**, pp. 827-845.
- [20] Singh, KS.(1997). *The Scheduled Tribes*. Oxford, Oxford University Press.
- [21] Stringer, C. (2000). Palaeoanthropology: Coasting out of Africa. *Nature*. **405**, pp. 24-25.
- [22] Scholes, C.;Siddle, K.;Ducourneau, A.;Crivellaro, F.;Järve, M.;Rootsi, S.;Bellatti, M.;Tabbada, K.;Mormina, M.;Reidla, M.;Villemis, R.; Kivisild, T.; Lahr, M.;Migliano, AB. (2011). Genetic diversity and evidence for population admixture in BatakNegritos from Palawan. *Am. J. Phys. Anthropol*.**146**, pp. 62-72.
- [23] Tabbada, KA.;Trejaut, J.; Loo, JH.; Chen, YM.; Lin, M.;Mirazon-Lahr, M.; Kivisild, T.; De Ungria, MC. (2010). Philippine mitochondrial DNA diversity: a populated viaduct between Taiwan and Indonesia? *MolBiolEvol*. **27**, pp. 21-31.
- [24] Teng, YS.; Tan, SG. (1979). Genetic evidence of gene flow from Indians to Malays. *Japanese Journal of Human Genetics*. **24**(1), pp. 1-8.