

ARTICLE


Y-chromosome SNP analysis of the Indian population reveals a distinct genetic structure and a deep divergence from other populations. The study identifies a novel Y-chromosome haplogroup, designated as P1a, which is found at high frequencies in the Indian population. This haplogroup is characterized by a unique set of SNPs, including a mutation at position 16319 (G>A) and a deletion at position 16311 (16311del). The P1a haplogroup is found in all major ethnic groups in India, suggesting a common ancestry. The study also identifies a distinct genetic structure within the Indian population, with a deep divergence from other populations. The results suggest that the Indian population is a distinct genetic entity, with a unique genetic history. The study provides valuable insights into the genetic diversity and evolution of the Indian population.

Gutala Venkata Ramana¹, Bing Su¹, Li Jin¹, Lalji Singh², Ning Wang¹, Peter Underhill³ and Ranajit Chakraborty^{*,1}

¹Human Genetics Center, University of Texas, School of Public Health, Houston, Texas, TX 77030, USA; ²Center for Cellular and Molecular Biology, Hyderabad, India; ³Department of Genetics, Stanford University, Stanford, CA 94305, USA

Full text available at www.nature.com/nature. Y-chromosome STR analysis of the Indian population reveals a distinct genetic structure and a deep divergence from other populations. The study identifies a novel Y-chromosome haplogroup, designated as P1a, which is found at high frequencies in the Indian population. This haplogroup is characterized by a unique set of SNPs, including a mutation at position 16319 (G>A) and a deletion at position 16311 (16311del). The P1a haplogroup is found in all major ethnic groups in India, suggesting a common ancestry. The study also identifies a distinct genetic structure within the Indian population, with a deep divergence from other populations. The results suggest that the Indian population is a distinct genetic entity, with a unique genetic history. The study provides valuable insights into the genetic diversity and evolution of the Indian population.

caste cluster.⁵ There are over 400 tribal populations in India in addition to other religious groups like Muslims, Sikhs, Christians, Jains and migrant groups such as the Parsees and Siddis.⁶ Earlier studies from India, based on Y chromosome short tandem repeat (STR) polymorphisms have shown that there is either negligible or no male gene flow among populations of India.^{7,8} In contrast, mtDNA d-loop sequence variation⁷ showed higher levels of female gene flow between related caste groups. In this research article, we provide new data on 27 Y-chromosome SNP sites in three castes, three tribes, and Siddis (a migrant population of African ancestry) of Andhra Pradesh, South India, and demonstrate that while these SNP markers reveal a substantial genetic variation among these groups, they also detect an evidence of male gene flow among these population groups.

M a e a a d, e d
oth6(ands)-7 

The variance decomposition (AMOVA) analysis of the SNP haplotype frequencies provides a quantitative support of the same trend of genetic affiliation of these populations. With the seven populations divided into three groups (caste, tribe, and the migrant), and using the phylogeny of the 11

alone does not explain this. Excluding them from the analysis, while the numerical value between group differences becomes larger ($V_g=12.5\%$), it still remains non-significant.

Haplotype sharing and frequency differences of haplotypes can be examined in the light of these observations. It is true that the caste populations (both Brahmin groups and the Kammas) can be distinguished from the two tribal groups (Bagata, Poroja), since the caste populations exhibit the haplotypes H4, H4A,

STR level as well within the H14 lineage (Table 3), some of which are at least two mutation steps different from each other. The non-significant caste-tribe group difference of the STR-haplotypes of the H14 lineage supports the gene flow hypothesis rather than the antiquity of the haplotypes.

Our data also suggests that Siddis have assimilated considerable non-African Y chromosomes (haplotypes H4, H5, H14, and H16) from the local Indian populations. The arrival of the Siddis in India dates back to AD 1100¹⁵⁻¹⁷ and they have had social contacts with several local Indian populations. From the combined frequencies of the haplotypes of African (H1 and H2) and non-African haplotypes (H4, H5, H14, and H16), data shown in Table 1 indicates that at least 56% of the male genes of the Siddis could be of Indian origin, consistent with our estimate based on five STR loci we reported elsewhere.¹⁸

Ac c e p t e d

All DNA samples analysed here are anonymised, and were collected with informed consent of the subjects. Research supported by the US Public Health Services Research Grant GM 41399 from the National Institutes of Health.

Refe e n c e

- 1 Jobling MA, Tyler-Smith C: Father and sons: The Y chromosome and human evolution. *Trends in Genetics* 1995; **11**: 449–456.
- 2 Jobling MA, Tyler-Smith C: New uses for new haplotypes the human Y chromosome, disease and selection. *Trends in Genetics* 2000; **16**: 356–363.
- 3 Underhill PA, Jin L, Lin A et al: Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Research* 1997; **7**: 996–1005.
- 4 Underhill PA, Shen P, Lin AA et al: Y chromosome sequence variation and the history of human populations. *Nature Genetics* 2000; **26**: 358–361.
- 5 Malhotra KC: Population structure among the Dhargar caste-cluster of Maharashtra, India. In: Lukacs JR (ed) *The people of*

- 10 Excoffier L, Smouse P, Quattro J: Analysis of molecular variance inferred from metric distances among DNA haplotypes. Application to human mitochondrial DNA restriction data. *Genetics* 1992; **131**: 479–491.
- 11 Schneider S, Kueffer JM, Roessli D et al: Arlequin ver 1.1 A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland, 1997.