

# GENOME RESEARCH

## Indian Caste Origins: Genomic Insights and Future Outlook

Partha P. Majumder

*Genome Res.* 2001 11: 931-932

Access the most recent version at doi:[10.1101/gr.192401](https://doi.org/10.1101/gr.192401)

---

### References

This article cites 2 articles, 1 of which can be accessed free at:  
<http://www.genome.org/cgi/content/full/11/6/931#References>

### Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#)

---

### Notes

---

To subscribe to *Genome Research* go to:  
<http://www.genome.org/subscriptions/>

---



# Indian Caste Origins: Genomic Insights and Future Outlook

Partha P. Majumder

*Anthropology and Human Genetics Unit, Indian Statistical Institute, Calcutta 700035, India*

**T**he main feature of Indian society, seen at its strongest in the rural areas, is caste. A caste is a collection of people who share similar cultural and religious values and practices. Members within a caste generally marry among themselves; intercaste marriages are a cultural taboo. These social regulations governing the institution of marriage have resulted in a substructuring of the Indian gene pool. There are also elaborate social regulations of avoidance of marriages within castes, and thus there is genomic substructuring even within a caste.

The origins of the castes in India remain an enigma. Many castes are known to have tribal origins, as evidenced from various totemic features that manifest themselves in these caste groups (Kosambi 1964). The caste system in northern India may have developed as a class structure from within tribes: As agriculture spread from the Indus River valley to the Gangetic basin, knowledge and ownership of the means of food production may have created hierarchical divisions within tribal societies (Kosambi 1964). Karve (1961) has also argued that “something very like castes were in India” even before Aryan speakers entered India.

The Aryan world comprised three classes (varnas): priests, nobles, and commoners. Aryans as the conquering people possibly placed their three classes on the indigenous Indian society. The varna organization is hierarchical. Initially, the system had names for two ranks, Brahma (Brahmin) and Kshatra

(Kshatriya), Brahmin being of a socially higher rank than Kshatriya. The third rank was made up of Vis, that is, all the subjects. To this society, a fourth rank was added: Shudra, who had no rights to Aryan ritual. In southern India, the menial workers, the so-called “untouchables”, were placed in a new varna, Pan-chama (meaning fifth). It is conceivable that the Aryan speakers had greater contact, including genetic admixture, with the Brahmins, who were professionally the torchbearers and promoters of Aryan rituals. The Aryan contact should have been progressively less as one descended the varna ladder. The genetic expectation, therefore, is that the proportions of those genes (or genomic features, such as haplotypes or haplogroups) that “characterized” the Aryan speakers should progressively decline from the highest varna to the lowest and a reverse trend should be observed with respect to those genes that “characterized” the indigenous Indians.

Although some previous studies have sought to test this expectation, the observed trends were equivocal. The primary reason was the lack of data on a large uniform set of markers from populations of India and central/west Asia (the region from which the Aryans speakers who entered India originated). The study by Bamshad et al. (2001), who have also sought to test the above expectation, is clearly a landmark. Using a very large battery of genomic markers and DNA sequences, spanning autosomal, mitochondrial, and Y-chromosomal genomic regions, they have shown that the observed trend of genetic admixture estimated from castes belonging to different varnas is congru-

ent with expectations. This trend was observed in each of the three data subsets. The only exception was in respect of mtDNA restriction site haplotypes, which was also noted in a recent study conducted by us (Roychoudhury et al. 2000). However, after combining these haplotype data with DNA sequence data, Bamshad and colleagues were able to capture the expected trend. Thus, this study not only provides a wonderful genomic view of the castes and of their origins, but also underscores the need for careful statistical analysis of genomic data for drawing appropriate inferences.

The use of “upper”, “middle”, and “lower” to designate caste hierarchy is much more recent than the use of varna. Whereas varnas are traditionally defined, different anthropologists have used different definitions of upper, middle, and lower castes, in terms of the castes that they included in each of these clusters. Sometimes these differences in definitions have stemmed from socio-cultural similarities or differences as noted or perceived by different anthropologists, and sometimes ranked caste-cluster compositions were altered for convenience, such as pooling to adjust for small sample sizes. As noted earlier, in studies such as Bamshad et al.’s, the most appropriate classification is by varna. As the reader will note, the authors have analyzed their data using different compositions of hierarchical caste-clusters and have obtained homologous results. However, it needs to be emphasized that traditional varna system is the only unequivocally accepted hierarchical system. In studies pertaining to the origins of castes, one is liable to draw incorrect inferences by in-

**E-MAIL** [ppm@isical.ac.in](mailto:ppm@isical.ac.in); **FAX** 91-33-577 3049.

Article and publication are at [www.genome.org/cgi/doi/10.1101/gr.192401](http://www.genome.org/cgi/doi/10.1101/gr.192401).

cluding castes belonging to different varnas in the same ranked cluster.

Bamshad et al. have chosen to study caste populations drawn from a restricted geographical region of India. They have rightly emphasized the need to replicate their findings. This is absolutely essential because, as Karve (1961) has noted, "it is not generally realized that the caste society in a sense was a very elastic society." Indeed, a caste bearing the same name may have very different origins in different geographical regions. There are examples in which a tribe dispersed over a large geographical region, took up different occupations

in different sub-regions, and "fitted" itself into the caste hierarchy on different rungs. Karve's work has also indicated that each of the different Brahmin castes (Chitpavan, Sarasvat, etc.) in Maharashtra probably has a different origin. Thus, the origin of caste populations may not be uniform over the entire India geographical space, and it is crucial to undertake studies to replicate Bamshad et al.'s findings. Finally, I would also like to suggest that in future studies bearing on the origins of the Indian castes, it would be a good idea to include tribal populations inhabiting the same region along with the caste populations.

## REFERENCES

- Bamshad, M., Kivisild, T., Watkins, W.S., Dixon, M.E., Ricker, C.E., Rao, B.B., Naidu, J.M., Prasad, B.V.R., Reddy, P.G., Rasanayagam, A., et al. 2001. *Genome Res.* **11**: 994–1004.
- Karve, I. 1961. *Hindu Society – An Interpretation*. Deshmukh Prakashan, Poona.
- Kosambi, D.D. 1964. *The Culture and Civilisation of Ancient India in Historical Outline*, 1991 Reprint. Vikas Publishing House Pvt. Ltd., New Delhi.
- Roychoudhury, S., Roy, S., Dey, B., Chakraborty, M., Roy, M., Roy, B., Ramesh, A., Prabhakaran, N., Usha Rani, M.V., Vishwanathan, H., et al. 2000. *Curr. Sci.* **79**: 1182–1192.