PRINT: ISSN 0972-3757 ONLINE: 2456-6360
Int J Hum Genet, 11(3): 199-201 (2011)
DOI: 10.31901/24566330.2011/11.03.08

Genomic Efficiency of Endogamy in India

Vipin Gupta

South Asia Network for Chronic Disease, C-1/52, Safdarjung Development Area, New Delhi 110 016, India
Telephone: +91-9899119242; E-mail: udaiig@gmail.com

KEYWORDS Endogamy. Genomics. Association

ABSTRACT The genetic underpinnings of Indian population structure and its high correlation with the prevalent social endogamy have now been reconfirmed, which firmly support the conventional assertions of the biological anthropologists.

Indian caste population is composed of different endogamous groups; and their mating pattern is defined by the restricted marriages between the clans of that particular endogamous group within the fold of a particular caste (Bhasin 2006). Thus, endogamy is a defining property of social structure of majority of the Indian populations, which is most of the times limited by geography in this subcontinent. Now, the pertinent question is whether this endogamy also points towards the conserved gene pools of these reasonably restricted groups? Let me start with a statement by Chakravarti (2009a), who said that "anthropologists have conventionally assessed kinship by asking people about their social relationships with others", but in contrast to this, anthropologists assess the social homogeneity and define a particular group by deeply studying its existing mating pattern for few generations along with comprehensive analysis of their cultural aspects during their long and contextual field-works. Further, they also confirm its genetic homogeneity by studying classical biological markers like blood groups, serum proteins, red cell enzymes etc. (Roychoudhury et al. 1985; Das et al. 1986; Mastana and Paphia 1994; Bhasin 2009) or mitochondrial and Y-chromosome markers (Basu et al. 2003;Wooding et al. 2003; Zerjal et al. 2007) in that specific group. This provides a real picture of biological relatedness of a particular group of people living in ecological settings relative to other groups.

The entire Indian population structure can broadly be divided along the lines of caste and tribe status, religion and geography (Bhasin 2009). The appropriate selection of homogenous population group in genetic based association studies is always a pertinent problem. Its solution has been seen by the geneticists to select geography along with race as two important variables for collecting their case and control samples. Some also consider belongingness to one linguistic family as one of the important indicators of homogeneity of their samples (Roychoudhury et al. 2001; Basu et al. 2003). But in India, due to its high cultural diversity between different endogamous groups, the selection of population only on the basis of race, linguistic affiliation and even geography has definitely proved to be the weak criteria of defining the unit of study (Indian Genome Variation Consortium 2008; Reich et al. 2009). In India, the ethnicity, which is a sociocultural homogeneity within the group, along with the geography are most robust measurement of the population homogeneity. The recent Indian Genome Variation Consortium (2008) on the basis of 470 single nucleotide polymorphisms (SNPs) and 1871 individuals revealed a high degree of genetic differentiation among the Indian ethnic groups and suggested that combining of endogamous populations without considering ethno-linguistic factors will result in spurious results in disease association studies. They also criticized heavily the reference of people of India as 'Indian' in many population genetics studies and suggested that blanket usage of the term should not be used in Indian milieu. However, for the first time in anthropological tone, they also suggested that it is possible to identify large clusters of ethnic groups (referring to endogamous groups) that have substantial genetic homogeneity. This first large attempt to detect genetic differentiation among population living in India clearly stated that "it is surprising that inspite of such a high levels of admixtures, the contemporary ethnic groups of India still exhibit high levels of differentiation and sub structuring". This study confirmed the conventional assertions of anthropologists (which were consistently denied by molecular geneticists) and defi-

nitely strengthened the utility of anthropological models of studying populations for Indian subcontinent.

The argument given by Chakravarti (2009a) that "the global picture of relatedness that is emerging from DNA studies stands to shatter many of believes about ourselves", is no more viable in Indian context as comprehensively demonstrated by the landmark work of Reich et al. (2009) on 132 Indian samples using more than half a million SNPs. This study strongly proves that our beliefs about endogamy in India were not incorrect and our notion of population needs not to be redefined. Their results explicitly indicated (with high F_{ST} values) that strong endogamy must have shaped the marriage pattern in India from thousands of years (after lots of different founder events followed by limited gene flow), and again very firmly confirms the outcome of decades long research efforts of Indian biological anthropologists, that the caste / endogamy along with geography are the precious factors for studying the genetic structure of Indian population (Bhasin and Walter 2001). Reich and colleagues (2009) also estimated the age of founder events (beginning of endogamy) in Indian population which ranges between 30 to 100 generations ago. The genomic efficiency of endogamy has also been finally approved by the argument that the "the degree of differentiation from the neighbor endogamous population is at least as large as between northern and southern Europeans and sufficient to cause false positive associations to disease, if uncorrected" (Reich et al. 2009). Although Reich et al. (2009) proved that founder events may cause more recessive diseases than consanguinity in India even in Indo-European population groups, but practically the prevalence of recessive diseases in India still has to be addressed. Given the lethal nature of recessive diseases; existing strict cultural beliefs during marriages; and prevalence of consanguinity, mainly in south India that too represents a fraction of their overall number of marriages (Krishnamoorthy and Audinarayana 2001; Nath et al. 2004) expects low overall prevalence of recessive diseases in India (esp. northern region). This is partly explained by the culture of between-clan marriages in different endogamous groups within a particular caste (that is, caste endogamy but clan exogamy). Moreover, the ingenious model of "Indian cline" proposed by Reich et al. (2009) reflects the proportions of ancestry from just two ancestral populations (that is, ancestral north Indians and ancestral south Indians) would also suggest that how well the natural mating experiment of endogamy has maintained the ancestral gene-pool in India. After judging the empirical evidence from Reich and his colleagues (2009), the relevance of endogamy in India has also been finally accepted by Chakravarti et al. (2009b).

Thus, the anthropological model of endogamy could be recommended for genetic epidemiology research via genome–wide association studies in order to alleviate the impact of population stratification and detecting true association related to the studied complex trait. Further, minimizing the limitations of disease association studies during field-work by collecting case-control samples from same endogamous groups within a geographical boundary will be more useful, rather than incomplete post-hoc corrections for population stratification using different statistical tools (like Principal Component, Structure, and Multi Dimensional Scaling etc.). Hence, a good field-work strategy among different endogamous groups by truly representing their cultural and genetic diversity could be more powerful method in mapping genes for complex human traits.

ACKNOWLEDGEMENTS

Grant: South Asia Network for Chronic Disease, New Delhi

REFERENCES

Basu A, Mukherjee N, Roy S, Sengupta S, Banerjee S, Chakraborty M, Dey B, Roy M, Roy B, Bhattacharyya NP, Roychoudhury S, Majumder PP 2003. Ethnic India: A genomic view, with special reference to peopling and structure. *Genome Res*, 13: 2277–2290.

Bhasin MK 2006. Genetics of castes and tribes of India: Indian population milieu. *Int J Hum Genet*, 6: 233-274

Bhasin MK 2009. Morphology to molecular anthropology: Castes and tribes of India. *Int J Hum Genet*, 9: 145-230.

Bhasin MK, Walter H 2001. *Genetics of Castes and Tribes of India.* Kamla-Raj Enterprises, New Delhi.

Chakravarti A 2009a. Kinship: Race relations. *Nature*, 457: 380-381.

Chakravarti A 2009b. Tracing India's invisible threads. *Nature*, 461: 487-488.

Das BM, Das PB, Das R, Walter H, Danker-Hopfe H 1986. Anthropological studies in Assam, India. *Anthropol Anz*, 44: 239–248.

Indian Genome Variation Consortium 2008. Genetic landscape of the people of India: A canvas for disease gene exploration. *J Genet*, 87: 3-20.

Krishnamoorthy S, Audinarayana N 2001. Trends in consanguinity in South India. *J Biosoc Sci*, 33: 185-197.

Mastana SS, Paphia PP 1994.Genetic structure and microdifferentiation among four endogamous groups of Maharashtra. *Ann Hum Biol*, 21: 261-262.

Nath A, Patil C, Nai VA 2004. Prevalence of consanguineous marriages in a rural community and its effect on pregnancy outcome. *Ind J Comm Med*, 1: 41-43.

Reich D, Thangaraj K, Patterson N, Price AL, Singh L 2009. Reconstructing Indian population history. *Nature*, 461: 489-494.

Roychoudhury AK, Nei M 1985. Genetic relationships between Indians and their neighboring populations. *Hum Hered*, 35: 201–206.

Roychoudhury S, Roy S, Basu A, Banerjee R, Vishwanathan H, Usha Rani MV, Sil SK, Mitra M, Majumder PP 2001. Genomic structures and population histories of linguistically distinct tribal groups of India. *Hum Genet*, 109: 339-350.

Wooding S, Ostler C, Ravi Prasad RV, Watkins S, Sung S, Bamshad M, Jorde LB 2003. Directional migration in the Hindu castes: Inferences from mitochondrial, autosomal and Y-chromosomal data. *Hum Genet*, 115: 221-229.

Zerjal T, Pandya A, Thangaraj K, Ling EY, Kearley J, Bertoneri S, Paracchini S, Singh L, Tyler-Smith C 2007. Y-chromosomal insights into the genetic impact of the caste system in India. *Hum Genet*, 121: 137–144.