J Biosci (2019)44:75
DOI: 10.1007/s12038-019-9885-3

Commentary

On Reconstruction of ancestral footfalls in South Asia using genomic data *By* Saikat Chakraborty and Analabha Basu

Keywords. Ancestry; DNA; frequency; genome; populations

The events of the past that produced the huge diversity of population groups in present-day South Asia is a fascinating subject for population geneticists. Its geographic location, cultural history and the basic fact that South Asia represents a very significant fraction of the world's population means that questions of its history are ones of great importance for our understanding of humanity. Chakraborty and Basu describe the population genetics approaches that have been applied to illuminating the past and producing the understanding that we have today. Population genetics is an imperfect tool but it is an incredibly important one because it can directly connect the people living today to the events of the past. Chakraborty and Basu very nicely lay out the important questions that population genetics can address, but the reader is left hanging to some degree since the emphasis of the presentation is on the questions that have been best answered to date, with a clear bias towards the answers that have come from the paired approaches of mitochondrial and Y chromosome haplogroup analysis. These answers are very nicely laid out and discussed and the explanations of the technical limits that apply are very accessible to the reader. Meanwhile, areas that are less settled and are of ongoing interest and addressable with the newer tools of autosomal markers and ancient DNA are discussed in less detail. Going back to the questions laid out at the beginning of the piece to see where one may hope to have more complete answers emerging from population genetics approaches gives a useful perspective.

The question of when modern humans colonized South Asia has been answered to some degree through the analysis of mitochondrial haplogroups. The coalescence dates of around 72 kya are consistent with dates for arrival of modern humans in Australia derived from a variety of approaches. What are the prospects for improving on the confidence in this answer? Ancient DNA studies are dependent upon the recovery of scarce remains that have been preserved sufficiently to provide DNA for analysis. Results from ancient DNA recovered in South Asia may begin to provide answers to some questions, but there is no good reason to think that they will produce a picture of this time period in South Asia, and so we are likely to be left with the answer that we have at present.

Studies of Y haplogroups give a clear picture of a subsequent migration into the region that largely replaced the male lineage of the first wave of migration. Autosomal marker analysis has provided a great deal more information and much attention has been called to a significant migration that brought the ANI to the subcontinent where they admixed with the descendants of the first migrants. Chakraborty and Basu rightly bring out the role that migration of Tibeto-Burman- and Austro-Asiatic-speaking peoples played in adding to the genetic diversity of South Asia. It seems unlikely that additional studies of autosomal markers in present-day populations will add much to this story. Ancient DNA studies of European remains have helped to provide data on migrations into Europe during similar time frames. One can hope that technical improvements and expansion of the technologies to more laboratories will allow ancient DNA studies to provide more detailed answers in South Asia as well.

The people that first migrated to South Asia came from Africa, as did all of the populations of modern humans. Given the time frames involved, it seems very unlikely that population genetics will ever be able to add to details as to when they arrived in South Asia, and even the most optimistic scenarios for ancient DNA would not promise much hope to provide a detailed picture of the migratory routes of modern humans into South Asia.

The last four questions raised by Chakraborty are the ones that seem most likely to receive increasingly fine-grained answers over the coming years, perhaps in part because of technical advances that are yet to be developed, but certainly through wider and more systematic application of present methods to the populations of South Asia. The first of these four questions, as to how many populations contributed to the present-day mixture, seems to have achieved its answer in broad

Published online: 08 July 2019

strokes. Many detailed questions arise, though, as one looks more closely. Although one can be confident that four populations contributed major components to the genetic diversity of South Asia, it is not clear whether those four groups were actually contemporaneously homogenous or were multiple distinct groups that have not been distinguished from each other with present-day approaches. The Ancestral North Indians seem likely to have been made up of more than one constituent group, as alluded to by Chakraborty and Basu and speculated upon by Narasimhan and co-authors, and further work with ancient DNA could readily provide clearer answers to this question in the future.

The question of when the major admixture events leading to the present-day population structure of South Asia occurred may also receive improved and more fine-grained answers. The study of ancient DNA has only limited ability to address this question, although the answers that they do provide can be tied to a time scale. It also seems possible that the development of newer statistical methods applied to autosomal markers might be able to improve upon the approaches that have been applied to date and to give us improvements in resolution at the time scale of a few thousand years.

The last two questions raised by Chakraborty and Basu are left unanswered but these two questions are the ones that have the most concrete implications for the people of South Asia today and the questions that are likely to receive much more detailed answers in the coming years. The straightforward application of existing methods to the analysis of autosomal markers in larger numbers of individuals will increasingly provide answers. Much of the efforts of population geneticists have to date been applied to clearly defined and recognizably isolated population groups. This is an appropriate approach to answer questions of history but it is less well-suited to answering the last two questions. It is well accepted that endogamy has persisted in South Asia to a much larger degree than in European populations, but to what degree this affects the genetic makeup of the rapidly changing peoples of South Asia today is not clear. The related questions of how much and how rapidly population groups expanded from their founders have enormous implications for how the modern-day concepts of the genetics of disease and of personalized medicine intersect with the delivery of health care to the people of South Asia today. Endogamy and rapid population expansion from a small founder population both serve to increase the frequency of otherwise rare loss of function variants within a population. This has implications for the diagnosis of recessive diseases, and the Finns have long recognized this by the coining of the term ‘Finn Diseases’. The populations of India, to varying degrees, carry similar disease susceptibility in their genomes in a fashion that remains diverse and largely uncharacterized. It sometimes seems as though all of the world’s people will soon have their genomes sequenced and South Asia will not escape this trend. It will be incumbent upon population geneticists to elbow their way in and work together with medical geneticists to provide practical information about genetic diversity, its historical and cultural antecedents, and its medical implications.

Andrew Peterson

Department of Molecular Biology and Metabolic Disease,
Genentech, South San Francisco, USA
(Email, andpete@gmail.com)