



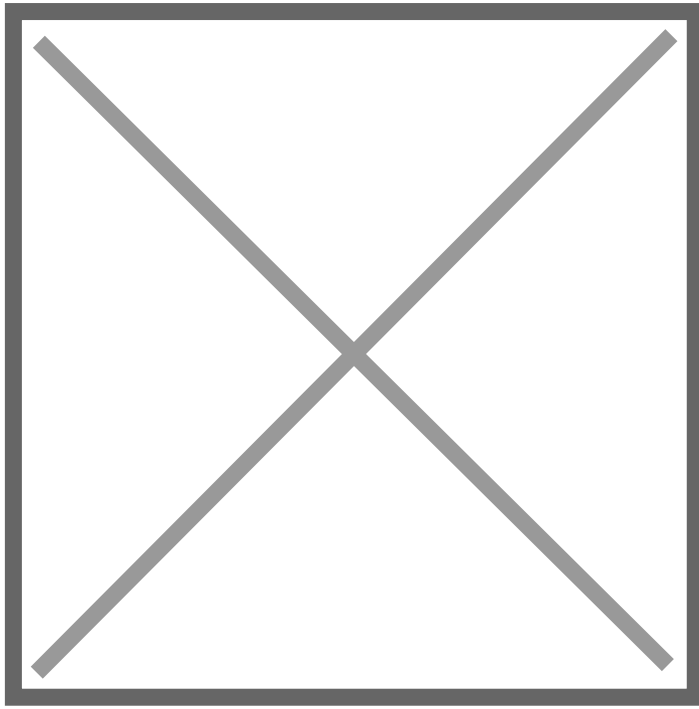
BioSpect

the business of Bio

24 years of leveraging the business of Bio
www.biospectrumindia.com

Indian gene decoded

07 May 2008 | News



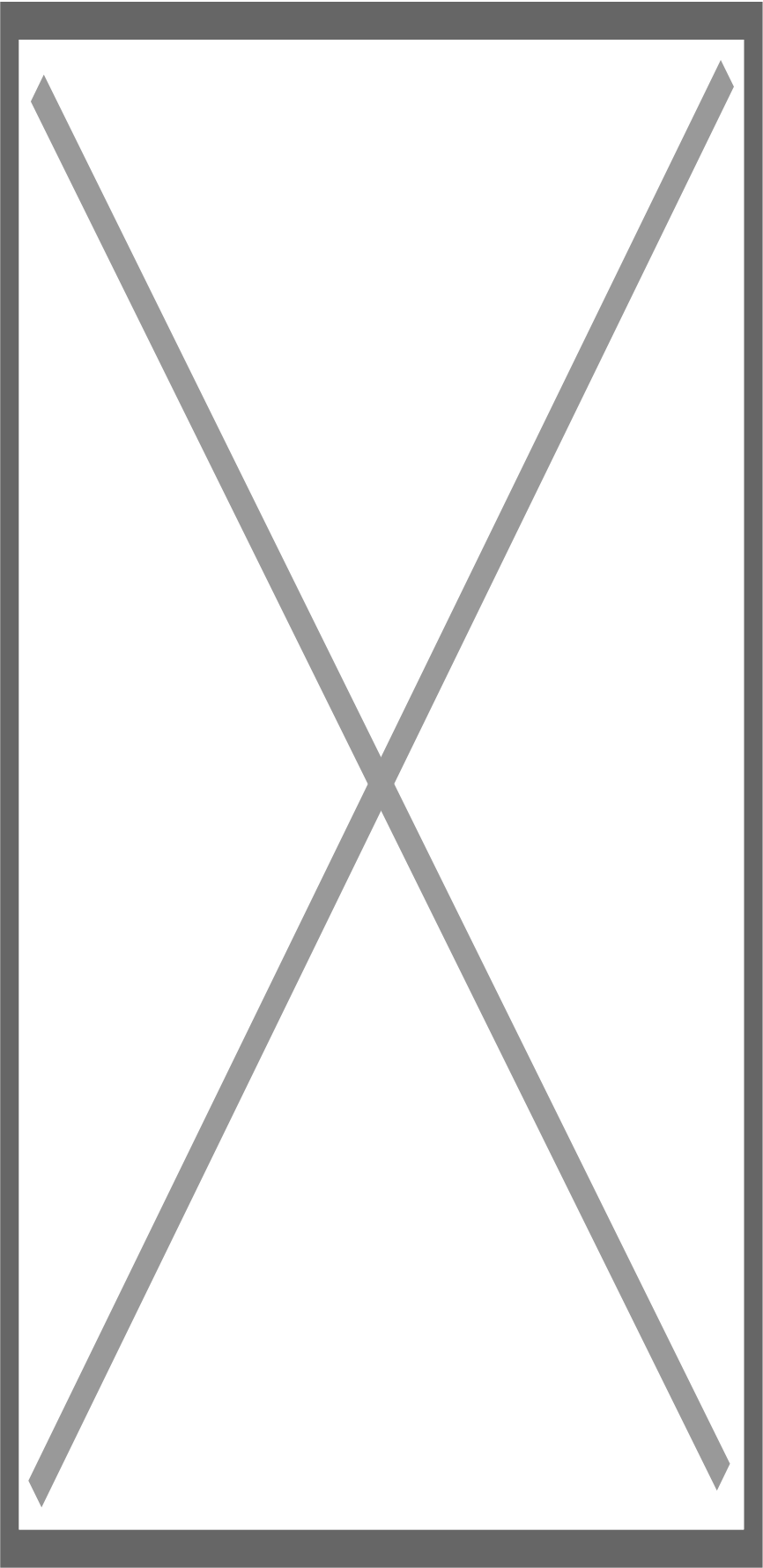
Indian gene decoded

The largest ever study to understand the genetic diversity of India's one billion people by a team of top genetic experts has thrown up some interesting data and concludes that India has one of the most genetically diverse populations in the world. BioSpectrum analyzes the highlights of the study, which has been published in the April issue of the prestigious Journal of Genetics, published by the Bangalore-based Indian Academy of Sciences.

The world's second largest country in terms of population, teeming with over 1.1 billion people, next only to China's 1.4 billion, was treated as a genetically homogenous place, and referred simply as the "Indian" population. A rigorous scientific study has indicated that it is not so.

"Our study reveals a high degree of genetic differentiation among Indian ethnic groups and suggests that pooling of endogamous populations without regard to ethno-linguistic factor will result in false inferences in association studies," said Dr Samir K Brahmachari, one of India's leading geneticists who led the study under the aegis of the Indian Genome Variation Consortium. Dr Brahmachari is also the director-general of the Council of Scientific and Industrial Research (CSIR), which runs a chain of 38 publicly funded research institutions in India.

The study, conceptualized by Dr Brahmachari and Dr Lalji Singh, who heads the Center for Cellular and Molecular Biology (CCMB), Hyderabad, included 32 large populations of sizes above 10 million people and 23 isolated tribal populations. Together, these groups represented the vast ethnic, linguistic and geographical diversity of India. Unveiling the findings of a study titled, "Genetic Landscape of People of India: A Canvas for Disease Gene Exploration", Kapil Sibal, India's minister for science and technology and earth sciences, said that the study led by Dr Brahmachari, was the largest scientific endeavor in the field of biology after green revolution.



"The objective was to create a basal data for disease profile and not to map genetic diversity. This is the first time anywhere in the world that a single country has been mapped to this extent", added Dr Brahmachari.

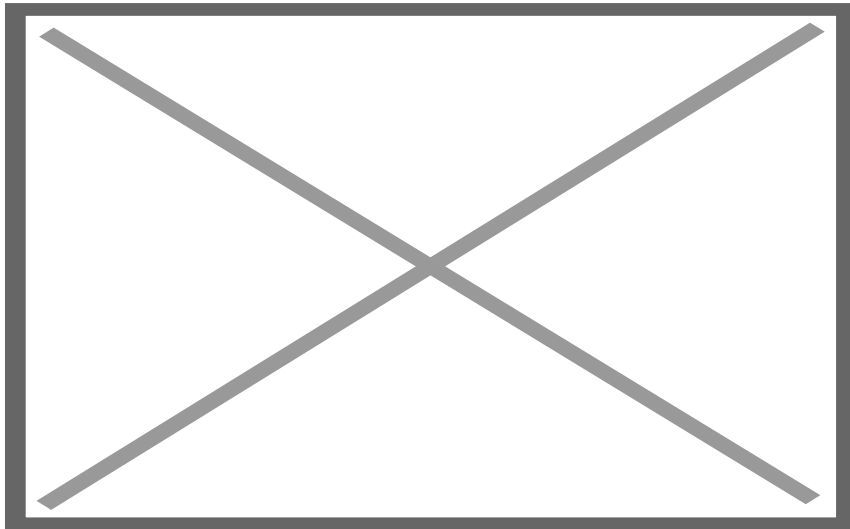
The Indian Genome Variation Consortium, a CSIR network project was started independently in 2003 after it was decided that 45 samples cannot represent the total diversity of India to build the next generation map of human genome initiated by the International HapMap Consortium.

"India represents a global genetic mosaic in its population," said Sibal. "The Indian population forms a continuum of genetic spectrum that not only bridges the Caucasians and Oriental Asians and also includes indigenous populations derived mostly from Austroasiatic and Dravidian speaking population." he added. "This means that now clinical trials to test the efficacy of drugs on all major populations can be conducted in India itself given such a spectrum. The results will be used to take up further scientific studies to generate data on the nature and extent of variations related to a large number of genes and identify genomic regions related to susceptibility to specific diseases and response to drugs.

The study, largest of its kind done on Indian population, included 1871 people from diverse groups, with 1,240 men and 631 women.

It has been found that there were no clear geographical groupings of people in India. However, ethnicity or being tribal or non-tribal and languages seems to be the major defining factors of genetic affinities between the populations in the country. This assumption, earlier concluded by other scientific studies analyzing blood group, serum protein and enzyme loci, has been validated by the genomic study.

India's vast population has evolved historically with different evolutionary histories and with multiple cultures. The earliest humans were estimated to have reached the north-western part of the Indian sub-continent nearly 70,000 years ago and then moved southward towards Sri Lanka in the next 20,000 years. There were further migrations into eastern India from Myanmar some 4,500 to 11,000 years ago.



Broadly Indians belong to four major language-speaking groups: Austro-Asiatic (AA), Tibeto-Burman (TB), Indo-European (IE) and Dravidian (DR). Over 80 percent of the practice the distinct Hindu religion and 8 percent of the nation's population comprise ethnic tribals or indigenous people, spread across the country. Almost all the major regions of the world such as Islam, Christianity, Buddhism, Judaism, Zoroastrianism are followed by the people. As a result of the intermingling of people belonging to these diverse religious, over centuries, several distinct endogamous groups have been formed. In fact, studies conclude that Indian population now consists of 4693 communities with several thousands of endogamous groups, 325 functioning languages and 25 scripts.

Another factor leading to the genetic diversity is the caste (socio-cultural) system followed by the Hindus, which sub-divides them into four distinct communities which initially evolved on the basis of occupations. Scientists have concluded that linguistic, geographical, socio-cultural distinctions have prevented free gene flow between the populations and led the formation of thousands of endogamous groups in India.

The four major morphological types -Caucasoid, Mongoloid, Australoid and Negrito, are present in the Indian population. The Caucasoid and Mongoloid populations are mainly concentrated in the north and northeastern parts of the country. The

Australoids are mostly confined to the central, western and southern India, while the Negritos are restricted only to the Andaman Islands

The Austro-Asiatic speakers are exclusively tribals and are dispersed mostly in the central and eastern parts of the country. Molecular diversity studies have revealed that the Austro-Asiatic speakers are likely to have been the most ancient inhabitants of India.

The current genome study has revealed that at the pan-India level, there were significant genetic variations between the tribal and caste populations. Further, in a geographical region too, tribes and caste people sub-classified on the basis of languages, exhibited genetic diversity.

For instance, the isolated populations in the Himalayan belt in the North India, were genetically closer to the Chinese and Japanese populations than the people in the rest of the country. So far, the global genomic studies had classified Indian populations as part of the Asian group which mainly included the Chinese and Japanese people. The Indian study has demonstrated that two major populations segments-those speaking the Dravidian (DR) and Austro-Asiatic (AA) languages who are mainly the tribals -living in the central and southern regions were entirely different from the Asian populations used in the HapMap populations.

Look for a person for my medicine

...

The first dollar for human genome sequencing came in the 1990s. In fact, this June, it's the 10th year since the completion of the DNA sequence for the last of the 24 human chromosomes. More than two human genomes have been sequenced-Craig Venter and James Watson. A Chinese genome has also been sequenced recently and the Panda genome sequence will be completed by August. Besides, almost 1,000 human genomes will be sequenced in the next year. The pace of the sequence has accelerated but the question is, have we conquered the cures?

One has been talking of personalized medicine, but Dr Brahmachari and his team are taking a different approach-they are talking about finding the right person person for my medicine. Recognizing India's unique strengths, CBT/IGIB launched two major initiatives in 1998. Scientists started analyzing genomes and correlating it with disease-based information available from hospitals. In the process, they also developed novel computing techniques--bioinformatics tools--that would help in the analysis of genomes.

On that strength, Prof. Brahmachari and his team undertook the study of the Indian Genome Variation Consortium, a CSIR network project started independently in 2003 after it was decided that 45 samples cannot represent the total diversity of India to build the next generation map of human genome initiated by the International HapMap Consortium. "The objective was to create a basal data for disease profile and not to map genetic diversity. This is the first time anywhere in the world that a single country has been mapped to this extent," pointed out Dr Brahmachari.

The Indian Genome Variation Consortium (IGVdb) database provides a genetic map of relatedness of 55 contrasting Indian populations for drug response and risk with respect to cardiovascular disorder, metabolic syndrome, diabetes, infectious disorders, high altitude disorders, asthma, and susceptibility to malaria and tuberculosis. While the information for the infectious disease will come free, that for all drugs in the market will come at a price. It is a monumental study on genetic profile of the country's population and the largest ever by Indian organizations.

The Indian Genome Variation Consortium is a network of six institutes of the Council of Scientific and Industrial Research (CSIR).

Genetic affinities

The study found that genetically Indian populations could be classified into various clusters. The first cluster primarily comprised of tribal communities and Dravidians. The second cluster included Tibetan-Burmese populations irrespective of their geographical regions of habitat. This implies that these people living in different parts of India were genetically similar. Some of the populations living in isolated Himalayan regions also belonged to this category genetically, though many of them were speakers of Indo-European language family.

An interesting revelation is that the native populations of southern India, who spoke four or five distinct languages, formed a separate cluster together based on their genetic composition. This essentially means that the social barriers erected by the caste system has no genetic base as it was believed to be. Most indigenous people in south India belonged to the same cluster.

"The HapMap does not capture the entire diversity of the Indian subcontinent. Thus, it may be difficult to directly use the HapMap data to design genetic epidemiological studies for the entire population of India," the authors wrote in their Journal of Genetics paper.

Another interesting revelation is that the populations in the southern parts of the country were genetically more prone to infections like the HIV than their counterparts in the north.

The Aryan invasion theory

Earlier studies had hypothesized that the Dravidian language speakers (Tamil, Telugu, Kannada and Malayalam), now geographically confined to the southern parts of the country, were spread across the country before the arrival of the Indo-European speakers. And the Dravidian speakers, considered to be the original inhabitants of India, were forced to retreat to only the southern parts after their interaction with the "invaders". Dr Brahmachari's study of genetic heterogeneity among the Dravidian speakers validates the above hypothesis. This hypothesis was also supported by the mitochondrial DNA analysis done by other researchers.

The study indicates that the Indo-European speakers too exhibit genetic variations between various groups. However, and surprisingly, these two groups showed high levels of genetic differentiations between them despite centuries of inter-mingling. Essentially, the studies indicate that the north and south Indian populations, who have been culturally very different, were distinct genetically too. Such conclusions are likely to spark another round of intense debate within India as the distinctions are further accentuated by the vastly different economic profiles of the regions.

Practical applications

Citing that the study clearly elucidates the genetic profile of India, Sibal said, "The study finds the absence of a known genetic marker against HIV-1 in India, which has enormous implications in terms of finding what populations are susceptible to what diseases and help us not only predict the diseases as well as the effectiveness of drugs for these diseases." The study indicates a strong association between the genetic and linguistic profiles in India and also significant genetic differences in the association of disease associated genetic markers.

The study has enormous practical implications. The data is expected to help in the construction of specific drug response maps to aid policy level decision making for drug dosage interventions and disease risk management for complex as well as infectious diseases.

Scientists had tested the response of people to various drug molecules available in the Indian market. Preliminary results indicate that people in some parts of the country do not require some of the common vaccinations because they are genetically-equipped to respond to the infections. Public health agencies may be able use this data to make right and relevant purchases of essential vaccines and divert scarce resources to regions which need them the most.

Pharma companies will also be able to tailor their distribution and marketing plans and ensure supply of appropriate medicines only to different regions. The details of the drug response mechanisms will be available to pharma companies.

For instance the study reveals that Indians are susceptible to HIV and diabetes. Hence knowing which populations stand the risk of a particular disease, management of that disease would be easier. "A large number of projects have been undertaken in the data for predictive marker discovery and XIth plan aimed at utilizing this basal pharmacogenomics," Sibal said. He added, "The project will show the world how we go forward managing risk and discovering drugs."

Narayanan Suresh in Bangalore and Shalini Gupta in Delhi

