

## Castes of genes? Representing human genetic diversity in India

YULIA EGOROVA<sup>1</sup>

### Abstract

This paper explores the historical and social context of population genetic research conducted in India by focusing on a study by Reich et al which aimed to reconstruct Indian population history. The paper addresses two themes. First, it considers the agendas and modes of thinking about Indian populations and the caste system on which this study appears to be based. Second, it reflects on the medical implications of this study as they were presented in Reich et al's findings. I suggest that while genetic mapping of Indian populations appears to have inherited many of the problems characteristic of population genetic research conducted in the USA and globally, the specificity of this research in India involves a peculiar interplay of the postcolonial pursuit of genomic sovereignty, desire by the Indian state to become a player in the global realm of biotechnology, and age-old discourses naturalising caste and regional differences. My argument is that, although the study has offered conceptual space for a wide range of interpretations, it has a strong potential not just for naturalising caste and regional differences in India, but also for pathologising them without necessarily bringing tangible healthcare benefits in the foreseeable future.

### Introduction

In September 2009 *Nature* published a paper entitled 'Reconstructing Indian Population History'.<sup>2</sup> One of the objectives of the paper was to fill a gap in genomic research of human diversity, where India is supposedly underrepresented. The study on which the paper was based analysed 25 different groups on the sub-continent and provided evidence that contemporary Indians were the descendants of two ancient populations, labelled 'Ancestral North Indians' (ANI) and 'Ancestral South Indians' (ASI). The two populations were found to be genetically distinct from one another and the former were supposed to be 'genetically close' to Middle Easterners, Central Asians, and Europeans. The study also suggests that due to thousands of years of endogamy, Indian populations demonstrate strong founder effects. It makes a prediction regarding the spread of recessive conditions on the sub-continent. The authors of the paper are based in reputable research institutions in India and the USA and are sponsored by US, UK and Indian funders, including the National Institutes of Health in the USA and the Council of Scientific and Industrial Research of the Government of India. The study was widely reported in the mass media both in India and internationally, giving rise to diverse interpretations of its findings.

In this paper I focus on this study to provide an in-depth analysis of its aims, results, media representations and possible societal implications. This will allow me to explore in more detail mechanisms of knowledge production in Indian population genetics and the multifaceted results of this process. I aim to situate Reich et al's study in the wider context of historical debates about the nature of Indian populations and the origin of the caste system, as well as in the context of Science and Technology Studies (STS) debates

about the sociocultural underpinnings and implications of population genetic research. More specifically, I would like to explore two themes.

First, I consider the agendas and modes of thinking about Indian populations and the caste system on which Reich et al's study appears to be based and discuss what kind of contribution the study makes to discourses about caste and regional population diversity in India. Does the study mirror or challenge any of the existing theories of caste formation and human variation in India? Do its findings 'naturalise' the Indian social hierarchy by linking castes to specific genetic profiles or do they argue against racial explanations for the existence of the caste system?

Second, I reflect on the medical implications of this study as they were presented in Reich et al's article. What is the basis for the authors' prediction about the spread of recessive conditions in India? What is the background to their interest in discerning the possible healthcare benefits of this research? I will explore how Reich et al's study, which claims to have made a significant contribution to genetic mapping of Indian populations that were allegedly underrepresented in wider studies of human diversity, employs the trope of 'genomic sovereignty.' My paper will also discuss in what ways scientists' rhetoric about the medical implications of their findings both universalises Indian populations and describes them as 'genetically' different from the rest of humanity. I argue that though Reich et al's study has offered conceptual space for a wide range of interpretations, it has a strong potential not just for 'naturalising' caste and regional differences in India, but also for pathologising them.<sup>3</sup> Finally, the paper will also show that the study in question 'geneticises' healthcare problems facing Indian populations without offering any treatment or cures.

## **Methods**

This paper is based on an analysis of the following materials: 1) Reich et al's article and other scientific publications in population genetics addressing the issue of the caste system; 2) mass media representations of these publications, focusing on representations of Reich et al; and 3) material from nine in-depth interviews conducted in India, the UK, and the USA with the key scientists involved in population genetic research on South Asia.

Scientific publications were identified through a mapping exercise on population genetics in India, conducted in the course of a previous project on the impact of DNA studies on historical debates.<sup>4</sup> I collected newspaper and internet articles devoted to Reich et al's study by undertaking an extensive search for commentaries on genetic studies on the caste system in the mass media, in academic and popular academic publications, and on the internet, in the period from October 2009 to January 2010. Altogether I assembled 25 articles.

The interviews were conducted in the period from December 2008 to January 2010. My interviewees were identified through scientific papers on population genetics in South Asia. In the interviews my respondents were invited to reflect on their involvement in population genetic research and on the social implications of that research. Interviews

were recorded with informants' permission and transcribed. Analysis occurred in several stages. First, I examined scientific publications, interview transcripts and mass media articles and identified emerging themes and patterns. Second, I explored the way these themes related to each other by looking for commonalities and contradictions. For instance, I compared the way scientists represent their research in scientific publications, in the mass media and in the interviews. Finally, my analysis attended to the relationship between these studies and earlier historical conceptualisations of the notions of 'caste' in the South Asian context and to contemporary debates about the origin of early Indian cultures.

For the purposes of maintaining anonymity I do not disclose the names of my interviewees or the institutions at which they are based. I consider it important to maintain this anonymity because my informants were interviewed on matters that are politically sensitive. All interviewees work in the fields of genetics, biological anthropology and/or statistics, and have been involved in studies in population genetics aimed at reconstructing the history of human migrations in the context of South Asia. All are scientists occupying permanent positions as lecturers, and researchers in universities and research institutes in India, the UK and the USA. In referring to these interviews throughout the paper I avoid providing information about the status or country of residence of each informant, because the circle of scientists involved in genetic research on South Asian populations is relatively narrow, and providing any additional information would make them easily identifiable.

### **Caste and genes**

Many studies in population genetics conducted on the subcontinent have attempted, whether directly or indirectly, to address the vexed question of the origin of Indian castes. As I have discussed elsewhere,<sup>5</sup> DNA studies conducted on the subcontinent so far have not developed a consensus of opinion on the genetic history of the caste system. Some of them have argued that caste rank was proportionate to affinity to Europeans, with the upper castes being more similar to Europeans.<sup>6</sup> Others on the contrary, have stressed the indigenosity of genetic diversity in South Asia and have suggested that the Eurasian contribution to the gene pool of the subcontinent was minor.<sup>7</sup> I have also demonstrated that genetic research on caste has been widely publicised and appears to have offered rhetorical ammunition for different political groups to support their narratives of identity.<sup>8</sup> Thus, research supporting the theory of Aryan migration was welcomed by the Dalit movement, who saw it as a proof of the idea that 'upper castes' were alien to the subcontinent. At the same time, the Hindu right favoured the studies which suggest that the 'European' contribution to the South Asian gene pool was minimal, and hence could be construed as supporting the Aryan indigenosity hypothesis.

The caste system can be broadly described as a hierarchical structure divided into a number of endogamous groups, each pursuing one traditional occupation. When looking at the origin of Indian castes some studies have also explicitly or implicitly engaged with the controversial 'theory of Aryan migration'. This theory, which prevailed in historical accounts of the origin of the caste system up until the middle of the 20th

century, suggested that the system of four 'classes' reflected the ancient encounters between the so-called Aryans – who came to the sub-continent from beyond the Indo-Iranian borderlands – with the indigenous Dravidian population. According to this theory, the 'Aryans' were of the same stock as those groups who went west into Europe, and it was they who formed higher castes, while Dravidian 'natives' comprised the lower ones and the untouchables (considered to be 'outcaste').

Recently, historians have argued that such a simplistic account of early Indian history may be viewed as a European construction governed by the complex political and cultural interests of the colonial authorities.<sup>9</sup> Nevertheless, in contemporary India the theory of Aryan migration has become a focus of vigorous political debate about the origin of the Hindu culture. The idea that important elements of Hinduism may have been brought from outside of the subcontinent is not acceptable to the ideologues of Hindutva movement. This movement emerged at the beginning of the 20th century and unites a number of groups which demand a primacy of citizenship for Hindus in India. When the Hindu nationalist Bharatiya Janata Party was in power from 1998 to 2004 its officials actively sought to revise textbooks on Indian history. One of the imposed changes was the notion that the 'Aryans' originated on the sub-continent, which was construed as the cradle of civilisation.<sup>10</sup>

This idea was not by any means new. It was championed in some Hindu reform circles in the later stages of British rule. For instance, Dayananda Saraswati, the leader of the Hindu revivalist Arya Samaj movement, describes Aryans as the chosen people to whom the Vedas were revealed by God. They descended from Tibet at the beginning of time and settled in Aryavarta, which was supposed to have been located on the territory of the Punjab. From there Aryans dominated the world until the epic war of the Mahabharata broke out leading to the decline of their civilisation.<sup>11</sup> At the same time, other nationalist thinkers were all too keen to support the theory about the 'foreign' origin of Aryans, and to use it to justify caste inequality.<sup>12</sup> Thus, debates about Aryan migration and its relation to the caste system were closely linked to the wider discussion about the relationship between caste and 'race', which permeated colonial ideology,<sup>13</sup> nationalist thinking, and post-independence politics, oscillating between discourses that tended to 'naturalise' caste and those that saw it as a purely social institution.

### **Genes and race**

Reich et al's study should also be considered against the backdrop of the wider field of research in population genetics, which, in one way or another, engages with concepts of race and ethnicity. While some such studies attempt to reconstruct the history of human migrations, others try to trace 'molecular' differences between populations for medical purposes. What they have in common is that they inevitably construct and re-inscribe differences between human groups on the basis of alleged variations in their genetic profiles.

In a seminal paper published in 1972 Richard Lewontin demonstrated that the major part of human genetic variation occurs between individuals within local geographic populations, which suggests that the concept of race is an invalid taxonomic

construction.<sup>14</sup> Luigi Luca Cavalli-Sforza, the initiator of the Human Genome Diversity Project, is credited with using his project to dismantle the idea of race further and to refute Herrnstein and Murray's notorious claims made in *The Bell Curve* about the alleged relationship between race and intelligence.<sup>15</sup> And yet, the objective of the Human Genome Diversity Project was to create 'genetic maps' of human populations and to calculate genetic distances between them. This endeavour may have helped to debunk the idea of dividing humanity into a limited number of major races, but it reinforced the notion that individuals could be divided into discrete populations on the basis of their genetic make-up.<sup>16</sup>

Critics of population genetic research have observed that the process of looking for genetic differences between populations is mediated from the outset by pre-existing historical and sociocultural ideas about collectives of individuals constituting separate populations which are supposed to differ from each other on the level of biology. As Troy Duster notes:

It is possible to make arbitrary groupings of populations (geographic, linguistic, self-identified by faith, identified by others by physiognomy, etc.) and still find statistically significant allelic variations between those groupings. For example, we could examine all people in Chicago, and all those in Los Angeles, and find statistically significant differences in allele frequency at *some* loci.<sup>17</sup>

Quite apart from that, population genetic research is increasingly used in the name of improving health care provision in general and for 'minority' groups in particular. Some studies have aimed to include underrepresented communities in clinical trials to ensure that medical research does not rely entirely on people of European descent, which may potentially disadvantage other communities. Steven Epstein has suggested that this is "a victory worth savouring in a long struggle to bring medical attention to the excluded and underserved",<sup>18</sup> but has also cautioned that this approach could reinforce racial categorisation. Other projects have attempted to seek a genetic basis for conditions found at higher rates among particular populations. For instance, there has been a steady interest in connecting heart disease and hypertension in African Americans to the genetics of race, despite existing evidence that these complex conditions are associated with poverty and stress, among other factors, which could also explain why they are common in a community suffering from structural inequalities.<sup>19</sup>

Some genetic research has focused on groups which are at higher risk of recessive diseases, particularly Tay Sachs, sickle cell anaemia, and cystic fibrosis.<sup>20</sup> Such studies have undoubtedly contributed to a better understanding of these conditions and in some cases were even used by the 'affected' communities to improve health care provision for their members. For instance, an Orthodox Jewish group in the USA made use of research on Tay-Sachs to organise pre-marital genetic screening for their young men and women.<sup>21</sup> However, the danger of this type of research is that it opens the way for stigmatisation of groups where a recessive condition is found at higher frequency. As Wailoo and Pemberton observe, the emergence of these conditions in the public

consciousness in the USA transformed them into “vehicles for thinking about race and ethnicity”.<sup>22</sup> Not surprisingly, some communities protested against being associated with a genetic disorder. For instance, Ashkenazi Jewish women expressed concern about being identified as being at high risk for breast cancer.<sup>23</sup> Many African Americans were wary of genetic screening and counselling for sickle cell anaemia, which for them seemed to be part of a disturbing history of racial discrimination in health care.<sup>24</sup>

Of particular relevance to my discussion of Reich et al’s paper is the context and problematics of population genetic research in postcolonial countries. Genomics and biotechnology promise to become a measure of political and economic success for any country in the 21st century. Therefore it is not surprising that more and more governments outside Europe and North America are prepared to invest significantly in the development of the life sciences in an attempt to establish genomic sovereignty.<sup>25</sup> Although these policies are designed to promote academic and economic independence for ‘local’ hubs of science and technology, they are also embedded in the global networks and processes of knowledge production. They ‘naturalise’ national populations in the name of postcolonial empowerment, but at the same time borrow practices and conceptual tropes from the wider context of ‘genetic labelling’ to reinforce already existing categories. Ruha Benjamin has argued that:

In the context of national genomics initiatives the work of calibrating scientific and socio-political classifications is not haphazard conflation, but a deliberate interpretation of genomic data to match the socio-historical record and a re-imaging of historical and cultural narratives to make sense of genomic findings.<sup>26</sup>

In the following sections I demonstrate how Reich et al’s data are interpreted by multiple and diverse commentators in light of existing historical narratives and specific political agendas. I also attempt to discern what kind of assumptions about the nature of the caste system and the possible practical value of mapping the ‘genetics’ of Indian populations may have informed the paper itself.

### **The caste and the North-South divide**

Reich et al’s paper begins with the statement that “India has been underrepresented in genome-wide surveys of human variation”,<sup>27</sup> implying that the objective of the authors is to rectify this situation by providing a more comprehensive picture of human genetic diversity on the subcontinent. From the very first sentence, Reich’s paper firmly positions itself in the history of biological anthropological research conducted on Indian populations. In the first paragraph of the paper the authors establish a connection between their research and earlier studies of human variation in India, from the first surveys that used anthropometric data to the more recent genetic studies. The paragraph concludes that:

The most comprehensive survey of genetic variation in India so far analysed 405 single nucleotide polymorphisms (SNPs) in 55

groups and identified distinct clusters correlated to language and geography, while another study analysed 1,200 polymorphisms in 15 Indian American groups. However, neither study analysed enough data to more finely discern patterns of genetic variation.<sup>28</sup>

Thus, the authors proclaim the ‘genealogical’ connection between their investigations and prior research aimed at categorising Indian populations along biological lines, suggesting that their study has continued the work that has been on-going for a long time, but that they have used more advanced technological means.

The paper offers an analysis of 132 samples from 25 groups representing 15 states of India and six language families.<sup>29</sup> It aims to answer five questions about Indian population history,<sup>30</sup> three of which are informed by old historical and political debates about the relationship between South Asian populations and populations from outside the subcontinent, about the nature of the human genetic diversity of India, and about the origin of the caste system.

The first question: “Does India contain more substructure than Europe?” addresses the issue of the genetic composition of Indians in comparison with Europeans. It seeks to determine whether Indian populations are more diverse than European groups, a conclusion that would combat an old Eurocentric tendency to ‘lump together’ various Indian communities. The second question: “Has endogamy been more long-standing in Indian groups?” attempts to cast light on the history of the caste system. Answering this question would allow one to make an inference about how old the caste system is and thus to challenge or support the theory of Aryan migration and suggestions about the role of the British in caste development. The third: “Do nearly all Indians descend from a mixture of populations?” again engages with the theory of Aryan migration, as well as with the issue of the historical relationship between the Northern and the Southern populations of India. In popular discourse and in political debates, the populations of the North and of the South of India are perceived as both culturally and ‘biologically’ different, with ‘northerners’ (and upper castes in the South) being allegedly descended from Eurasians and having fairer skin, and southerners being the dark-skinned descendants of ancient Dravidians.

The study demonstrates that samples taken from people from the same state and caste level have a fixation index which indicates a relatively high genetic proximity of their members - the determined fixation index proved to be 0.0069, higher than the analogous 0.0018 in Europe when comparing within regions.<sup>31</sup> The authors suggest that this finding could be explained “if many groups were founded by a few individuals, followed by a limited gene flow”.<sup>32</sup> It is argued that six tested Indo-European and Dravidian-speaking groups demonstrate evidence of these founder events going back to more than 30 generations ago and in one case (the Vysya caste of Andhra Pradesh) to more than 100 generations ago.<sup>33</sup> This finding is immediately used to assert the view that the caste system is an ancient institution:

Some historians have argued that ‘caste’ in modern India is an ‘invention’ of colonialism in the sense that it became more rigid under colonial rule.<sup>34</sup> However, our results indicate that many current distinctions among groups are ancient and that strong endogamy must have shaped marriage patterns in India for thousands of years.<sup>35</sup>

It is noteworthy that in this paragraph Reich et al make a reference to an anthropologist and historian of India, Nicholas Dirks, who made a significant contribution to debates about the relationship between caste and race. Dirks suggested that the contemporary caste system could be considered as a modern phenomenon and in some ways as a product of British rule. He does not argue that caste was invented by the British ‘from scratch’, but rather that under colonial rule it turned into a powerful label capable of subsuming other forms of social organisation.<sup>36</sup> By pointing out that Dirks’s account allegedly goes contrary to the newly discovered genetic ‘evidence’, Reich et al appear to be providing ammunition to those who support ‘racialist’ explanations for the existence of the caste system.

Throughout the paper references are made to the possibility of using these findings in health care. The authors suggest that the history of founder events in India is medically significant because it was likely to lead to higher rates of recessive diseases and that it was responsible for an even higher volume of recessive conditions in India than consanguinity. When writing about the possibility of using the findings of population genetics in health care, the authors stress the importance of analysing datasets derived specifically from Indian populations, rather than relying on DNA samples collected elsewhere. They give an example of a genetic marker which increases heart-failure risk by about sevenfold and which occurs at 4 per cent in India and is practically absent elsewhere. It is concluded that it is therefore imperative to conduct a full gene-mapping in India to identify “clinically significant” alleles that cannot be discovered by studying genetic variation outside of India.<sup>37</sup> Thus, the paper constructs Indian populations as genetically different from the rest of humanity and explicitly advocates the importance of conducting genetic population surveys in India, rather than relying on DNA data from other countries.

To get an insight into genetic ancestry found only in India the authors conducted a principal component analysis of European and Chinese alongside 22 Indian groups. On the basis of this analysis it is concluded that the tested Indian groups have inherited different proportions of the genetic material of what is referred to in the paper as the ‘Ancestral North Indians (ANI)’ and the ‘Ancestral South Indians (ASI)’. The former are supposed to be genetically close to Middle Easterners, Central Asians and Europeans, and the latter are described to be “as distinct from ANI and East Asians as they are from each other”.<sup>38</sup> The authors also suggest on the basis of their findings that there is a correlation between the proportion of the ANI-ASI ancestry and the rank of the caste, an argument which contributes to the idea that caste differences are observable on the level of DNA :

ANI ancestry is significantly higher in Indo-European than



Dravidian speakers (P50.013 by a one-sided test), suggesting that the ancestral ASI may have spoken a Dravidian language before mixing with the ANI. We also find significantly more ANI ancestry in traditionally upper than in lower or middle caste groups (P50.0025), and find that traditional caste level is significantly correlated to ANI ancestry even after controlling for language (P50.0048), suggesting a relationship between the history of caste formation in India and ANI–ASI mixture.<sup>39</sup>

Interestingly, Reich et al's admit that the ANI-ASI divide constructed in their study is first and foremost a model, which was used to simplify the data. However, they immediately reiterate that it can nevertheless be used in reconstructing the history of Indian populations:

We warn that 'models' in population genetics should be treated with caution. Although they provide an important framework for testing historical hypotheses, they are oversimplifications. For example, the true ancestral populations of India were probably not homogeneous as we assume in our model, but instead were probably formed by clusters of related groups that mixed at different times. However, modelling them as homogeneous fits the data and seems to capture meaningful features of history.<sup>40</sup>

Thus, the caste hierarchy and the North-South divide are de-naturalised in one sentence of the paper only to be 're-biologised' in another. This ambiguity was picked up by the mass media, with some articles emphasising the 'unifying' potential of the study, and others focusing on the genetic differences between castes that the study allegedly highlighted. Thus, *The Deccan Herald*, a major English-language newspaper published in the South Indian state of Karnataka, in an article entitled 'Genetically, as good as chalk and cheese' asserted that, according to the study, "the great North-South divide existed even in pre-historic India, at least genetically".<sup>41</sup> Adam Rutherford, writing in the *Guardian*, observes that the study's conclusion about the different composition of higher and lower castes may prove to have problematic political consequences. Kanishk Tharoor, commenting on Rutherford's article, suggests that although the 'discovery' of the genetic diversity of Indians vis-a-vis Europeans is worth celebrating, what it points to is "an uncomfortable reality of a stratified society".<sup>42</sup>

At the same time, an article published in *Jai Bihar* (an online magazine devoted to news about the Northern state of Bihar) asserts that Reich et al's paper has dismantled the myth of Aryan invasion, as well as the notion of the North-South divide in India. According to this interpretation, the findings about the ANI and the ASI populations imply that "although the country's 1.2 billion people belong to about 4,600 religions, castes and linguistic communities, the population shares a deep genetic heritage [of just two ancient groups]".<sup>43</sup> The authors of the paper were quoted as saying that the study had answered the politically sensitive question about the North-South divide by demonstrating that there was no need to speak separately about Aryans and Dravidians, as nobody in India was even close to having just ANI or ASI ancestry. The fact that,

according to the study, the degree of the ANI component in populations corresponds to their status in the caste hierarchy somehow was not seen to be divisive. Nor was it noted how similar the argument about the genetic proximity between the ANI and Eurasians, made in the paper, was to some elements of the theory of Aryan migration. *Times Online* quotes Aravinda Chakravarty from John Hopkins School of Medicine as saying that “the common shared ancestry and rampant ANI-ASI mixture may be the strong, invisible thread that binds all Indians”.<sup>44</sup>

### **Biologising the caste**

As discussed at the beginning of this paper, population genetics dismantled the concept of biologically distinct races only to reinforce the idea of biologically distinct (but smaller) populations. A similar process appears to be at work here. Commenting on Reich et al’s study, two scientists who were familiar with the study and have been involved in population genetic research in South Asia themselves, stressed in interviews with me that it indicates that practically all Indians have a component of both ANI and ASI ancestry – a message that, as was shown above, has been picked up by some of the mass media as well. However, as was noted in the previous section, the paper also clearly states that the proportion of ANI (and therefore the gradient of proximity to Eurasians) of the tested groups correlates with their position in the caste system and regional/linguistic affiliation.

To return to the argument made by Troy Duster, the very fact that Indian populations in the study were sampled by groups which had already been ‘historically’ and ‘sociologically’ known as castes, made it possible for the scientists to highlight ‘statistically significant’ genetic differences precisely between these groups/castes. It appears that the thinking that went into designing the methodology for this research was very much informed by prior knowledge and assumptions about the structure of the Indian caste system and by the categories developed in physical anthropological research conducted among Indian populations in the past. Indeed, as noted above, Reich’s paper begins by establishing its connection to earlier biological research aimed at categorising Indian populations. In the interviews, five of my informants involved in DNA studies on the subcontinent explicitly suggested that they were continuing the anthropometrical work of exploring the physical profiles of various populations that had been going on in biological anthropology for a century, but were using the more ‘advanced’ techniques of DNA research. As one of them put it, commenting on the relationship between his work in population genetics and anthropometric studies dating back to the first half of the 20th century, “It is nice to be able to maintain the continuity.”

Nevertheless, in their interviews all my respondents also argued that it is the newspapers that tend to ‘geneticise’ caste. As one of them put it:

You wake up in the morning and find that the headlines read ‘Your caste is in your genes’, which is a pathetic misinterpretation of genetic findings. If you can find distinctive genetic features in caste, the interpretation is not that your caste

is in your genes, it is just that there are certain genetic signatures that are predominant among caste populations and not among tribal populations. If you say your caste is in your genes that basically [means that] the caste system is genetic, it is transmitted in a Mendelian fashion from one generation to another, and that's definitely not true.

At the same time, it may be suggested that this very 'geneticisation' of caste has a liberating potential. Writing in *The Guardian*, Tharoor notes that he has always been suspicious of 'genetic history', which seemed to him "an upgraded version of 19th century racial 'science', a politer, more polished mode of measuring skulls... But in India, the revelations of this latest study are stark reminders of how much division underlies the fabric of the nation, and how much work remains in perfecting our democratic society".<sup>45</sup> He therefore sees Reich et al's research as a potential weapon to fight the caste system. But how likely is this potential to translate itself into concrete developments? I asked one of the Indian scientists involved in Reich's study whether he could see his work resulting in more inter-caste marriages, since, as implied in his paper, the caste system was supposed to lead to an increase in recessive diseases. He thought it was highly unlikely:

People may wish to re-think what they are doing in light of genetic research, but, on the whole, they want to keep their [marriage patterns] intact. For instance, in the South [of India] people have to pay a lot of dowry money and they want to keep it within their own groups.

Coincidentally, during our interview somebody contacted him about the implications for their prospective marriage that the scientist's findings could have, which indicates that the lay public might take the results of the study on board when choosing marriage partners. However, he did not think people were going to start marrying outside of their traditional groups even out of fear of recessive diseases. In his view, what was more likely to develop was a situation in which prospective brides and grooms would start using genetic screening for recessive alleles before agreeing to marriage. "We could test the prospective bride and groom, and see on the basis of their genetic profiles whether their marriage would be recommended or not."

Similarly, another scientist interviewed in India observed that Reich et al's paper or similar studies highlighting the 'medical' implications of endogamy were not going to have any impact on the way people in India choose their marriage partners:

Social values are something that you deal with on an everyday basis. Genes are not something that you deal with on an everyday basis. They are intangible. So I think the social values will transcend whatever implications genes may have.

### **The 'smokescreen' of genetic medicine**

In this respect it is important to reflect on the medical implications of the study as they were presented in the paper and in mass media comments. The whole

discussion of the possible medical relevance of the study presented in Reich et al's paper appears to be rather speculative. The paper suggests that "the widespread history of founder events in India is ... medically significant because it predicts a high rate of recessive diseases". This prediction is made on the basis of an observation made in Finland, where a high rate of recessive conditions was due to a founder effect.<sup>46</sup> Hence the 'public health' recommendation of the paper is that its "results highlight the value of systematically surveying Indian groups to identify those with the strongest founder effects, and prioritizing them for studies to identify recessive diseases and map genes".<sup>47</sup> In one publication Reich is quoted as saying that, "Further studies of these groups should lead to the rapid discovery of genes that cause devastating diseases, and will help in the clinical care of individuals and their families who are at risk."<sup>48</sup> Some may be somewhat surprised by the implication that research should happen in this order rather than other way round. One might suggest that it would be easier to 'spot' a disease through its symptoms first and look for its genetic basis afterwards, rather than to have to predict outbreaks of such illnesses on the basis of genetic mapping. Thus, Reich et al's paper 'pathologises' the caste system without detecting any diseases among the tested populations or offering any cures or other medical intervention.

In our conversations about the medical implications of this research and of other studies in population genetics conducted in India, all the scientists I spoke to admitted that they were not expecting any immediate medical benefits from such studies, as there was not enough data available that would allow for existing genetic information to be tied to specific conditions and reactions to drugs. It was their hope, though, that such data would appear in the foreseeable future and that the current studies will prove to be medically useful. Many suggested that it was a very important, though a long term endeavour. As one respondent put it:

We do have a framework of using such knowledge [*coming from population genetics*] in the field of medicine or to benefit human health, and the framework is most likely right. But what we do not have is the data that is required to tie the genetic make-up of the individuals with how you deal with drugs against specific diseases or for the enhancement of health. So, until that knowledge is beginning to crystallise, and once that knowledge crystallises we will know whether this framework that we work under will or will not work. It is my belief that it will.

Most of my informants admitted that they had to try to tie their research to medicine, because otherwise it would be very hard for them to have their work funded. "Evolutionary research is becoming very difficult to do now, especially because of lack of funding. There is hardly any funding that is available for pure evolutionary research. We are necessarily having to tie this research to certain medical goals," said one.

The rationale for the medical genetic mapping in India is explained at length in a paper published in *Nature Reviews Genetics*, reflecting on the findings of the Indian Genome Variation Consortium (IGVC). The Consortium is a government-funded collaborative

network among seven institutions, which developed the first large-scale database of human genomic variation in India. The objective of this initiative was to “facilitate research on disease predisposition, adverse drug reactions, and population migrations”.<sup>49</sup> The creation of IGVC could be viewed as an example of India’s attempt to establish genomic sovereignty. Hardy et al’s overview observes that:

Genomic sciences and related technologies can add value to India’s local health-care system by emphasizing prediction and prevention, and possibly decreasing the cost of health care through better diagnosis, early detection and improved treatment and management.<sup>50</sup>

Reich et al’s paper also gives an example of the perceived importance of achieving genomic sovereignty by referring to a gene mutation (a 25-base-pair deletion in MYBPC3) which increases the heart failure risk and is found mainly on the subcontinent. Scientists use this case to demonstrate how relying solely on genetic mapping conducted abroad could disadvantage Indian people. The authors suggest:

It has recently been shown that the power to discover disease risk variants can be increased by modelling Indian genetic variation using a reference panel of European and Chinese chromosomes. However, the example of MYBPC3 shows that this is an imperfect solution, because clinically significant alleles that are rare outside of India cannot be imputed by studying non-Indian variation.<sup>51</sup>

Interestingly, in this respect the authors also stress the genetic unity rather than diversity of Indian populations. They argue that cases like MYBPC3 are expected to occur at increased frequencies among Indians, because of their “shared descent from a common Indian ancestral population”.<sup>52</sup> Thus in this example the authors construct the population of India as relatively homogenous and different from other populations of the world.

And yet, the article from *Nature Reviews Genetics* offers an answer to the question about the rationale for developing a predictive population genetic database, which builds on a very different view of the genetic relationship between Indian and world populations. Apparently, to use a phrase coined by Adriana Petryna, such a database would make India a more attractive locale for “when experiments travel”.<sup>53</sup> As Hardy et al note:

Multinationals have been increasingly conducting early stage clinical trials in India because of the cost savings and readily accessible, largely drug-naive population resource. India has thus positioned itself as a global hub for conducting clinical trial research by investing in capacity and infrastructure. Some domestic companies subsidize their research and development (R&D) platforms by providing contract services for clinical trials for multinational and foreign companies. Accordingly, a ‘predictive population database’ could help maintain India’s

competitive edge by improving the selection specificity through stratification of the test population, thereby further reducing the time and cost associated with conducting clinical trials in India.<sup>54</sup>

Kaushik Sunder Rajan<sup>55</sup> has observed that although in some respects India is still a developing country with rather low human resources indices, it is prioritising the development of science and technology as a way of becoming a power to reckon with in the international arena. In the process of developing its involvement in the life sciences and biotechnology it has both displayed characteristics of a ‘colonised’ state actor *and* has framed itself as a ‘global player’ in biotech, he suggests. It appears that one of the meta-objectives of the studies in human genetic variation in India has been to facilitate India’s emergence as a market actor in the realm of global bionetworking. Interestingly, such studies try to achieve this objective while simultaneously addressing a time-old agenda of measuring and categorising Indian populations. Thus, paradoxically, they construct Indian communities as genetically distinct from each other for the purposes of categorisation, and as distinct from the rest of humanity for the purposes of establishing genetic sovereignty, while at the same time presenting them as universal for the purposes of clinical trials.

### The ‘pathology’ of caste

Reich et al’s study provides an intriguing attempt at reconstructing ancient history, while contributing to what Nikolas Rose and Carlos Novas have referred to as “the political economy of hope”.<sup>56</sup> In his interview with me, one of the scientists based in India gave an example of a genetic condition which prevents patients from responding well to anaesthesia. In India this condition appears to be found at a high frequency in a particular caste group. Although the importance of exploring the genetic basis of this disease and predicting which *individuals* would suffer from it is undeniable, it is hard to imagine what benefit linking this condition to a specific *caste* group would be, as not every member of this caste would carry the disease and the condition is bound to be found at lower frequency among other groups as well.

Nevertheless, Reich et al suggest that their results “highlight the value of systematically surveying Indian groups to identify those with the strongest founder effects, and prioritising them for studies to identify recessive diseases”.<sup>57</sup> It is quite likely that these ‘Indian groups’ will again be surveyed by caste and language and that such a survey will result in associating genetic conditions with caste and region. As a result, this research has a potential not only to reaffirm discourses about the alleged ‘biological’ basis of the caste system and of Indian regional/linguistic diversity, but also to pathologise caste and regional communities.

As I noted in the beginning of the paper, in the USA ‘minority’ communities who had a history of medical discrimination protested against being identified with a genetic condition. In this respect, it may be suggested that in India research linking particular castes with illness might be viewed with suspicion by those whose caste groups become associated with genetic disease. Introducing genetic screening may provide medical practitioners with another tool for controlling family size in vulnerable groups. Quite

apart from that, the possibility of genetic mapping decreasing the cost of health care in the whole of India appears to be rather remote, as it is not clear yet what kind of conditions will be connected to genetics as a result of this research and whether they will be solely genetic or multifactorial. It is hard not to wonder whether health care would benefit from a redirection of funding from genetic research to the treatment of already known illnesses.

### **Conclusion: Whose Genomic Sovereignty?**

Laura Nader has suggested that: “The politicization of science is unavoidable, not only because politicians, corporations and governments try to use what scientists know, but because virtually all science has social and political implications.”<sup>58</sup> Writing about biological research on race, John Hartigan observed that Nader’s suggestion could hardly be more evident anywhere other than in the study of race, explaining that “[t]here simply is very little basis to regard the recent efforts to establish a scientific basis for the study of race from a neutral stance, because *the prior belief in races* – as in the earlier era of scientific racism - remains a component of current research.”<sup>59</sup> In this respect, he invokes David Hess’s notion of technototemism, which Hess introduced to describe how class, race, and gender are articulated and interpreted through technoscientific configurations in the process of the coproduction of technical and social difference.<sup>60</sup> Studies in population genetics provide one such example of technoscientific totems that are used to continue the political, social and cultural work of dividing individuals into groups on the basis of perceived differences in their physicalities. In the case considered here, the genetic study was bound to give rise to socially relevant interpretations, just by virtue of the fact that it set out to engage with issues of social and political significance. It sampled Indian populations by caste and regional communities and thus was based on the *prior* assumption that these communities may in some ways be genetically different.

Elsewhere, I have discussed the way different social groups in India responded to earlier genetic research on the origin of the caste system and the nature of human diversity.<sup>61</sup> I suggested that, on the one hand, ‘genetic evidence’ was ascribed superior cognitive authority by these groups, but, on the other hand, it allowed enough conceptual space for different actors to interpret genetic research in such a way that it could be used to defend their (conflicting) agendas. Reich et al’s paper gave rise to multiple interpretations, and some even saw in it a potential to provide rhetorical ammunition against the caste system. There is no evidence to suggest that any of the scientists involved in this study set out to support a particular political agenda through their research or were under any pressure from their funders or institutions they were based at to produce a particular picture of Indian genetic diversity. However, by focusing their research on the genetic profiles of discrete caste, regional, and linguistic groups they contributed to the tradition of naturalising the caste and regional groups of the sub-continent. Moreover, it opened the way for pathologising such groups, while it is not at all clear whether this work will bring any tangible medical benefits for the people of India in the foreseeable future.

Benjamin has suggested that proponents of genomic sovereignty in postcolonial countries “celebrate its emergence as a form of empowerment without careful attention

to the ways in which genomic sovereignty inherits the perils produced by the ‘geneticization of life’ more broadly”.<sup>62</sup> I have demonstrated here that genetic mapping of Indian populations appears to have inherited many of the problems characteristic of population genetic research conducted in the USA and globally – naturalisation of socially distinct communities and a potential distraction of resources from comprehensive health care. The specificity of this research in India involves a peculiar interplay of the postcolonial pursuit of genomic sovereignty, desire by the state to become a player in the global realm of biotechnology, and time-old discourses naturalising caste and regional differences. These agendas have informed genetic mapping of the sub-continent in such a way as to allow it to produce narratives which both universalise Indian populations and present them as genetically unique, constructing them as homogenous on one level, while re-inscribing them as biologically different on another. How this research affects Indian policies and everyday practices remains to be seen.

### Acknowledgments

The study that this paper is based on was sponsored by the Nuffield Foundation and I am grateful for this support. Parts of this paper were presented at the annual meeting of the Society for the Social Studies of Science (Tokyo, August 2010) and at the international conference on Biohistories: DNA and bones in cultures of remembrance (Zurich, October 2010). I would like to thank the audiences for their feedback, and I am particularly grateful to Kaushik Sunder Rajan, Marianne Sommer and Gesine Kruger for their in-depth discussion of this material. I am also grateful to the peer reviewers for their insightful comments, and to the editors of the journal for their helpful suggestions.

---

<sup>1</sup> Department of Anthropology, Durham University. [yulia.egorova@durham.ac.uk](mailto:yulia.egorova@durham.ac.uk)

<sup>2</sup> D. Reich, K. Thangaraj, N. Patterson, A.L. Price and L. Singh. Reconstructing Indian Population History. *Nature* 2009; 461: 489-495.

<sup>3</sup> By pathologisation of caste I mean the process through which castes become associated with illness. I am grateful to one of the anonymous reviewers for suggesting this term to me.

<sup>4</sup> ‘DNA Evidence? The impact of genetic research on historical debates’ project was sponsored by the British Academy. Some of the outcomes of this project are presented in Y. Egorova. DNA Evidence? The impact of genetic research on historical debates. *BioSocieties* 2010; 5 (3): 348 – 365.

<sup>5</sup> Y. Egorova. De/geneticizing Caste: Population Genetic Research in South Asia, *Science as Culture* 2009; 18(4): 417-434.

<sup>6</sup> M. Bamshad et al. Genetic Evidence on the Origins of Indian Caste Populations. *Genome Research* 2001; 11: 994-1104; R. Cordeax et al. Independent Origins of Indian Caste and Tribal Paternal Lineages, *Current Biology* 2004; 14: 231-5.

<sup>7</sup> T. Kivisild et al. The genetic heritage of India’s early settlers persists both in Indian tribal and caste populations. *American Journal of Human Genetics* 2003; 72: 313-32; S. Sengupta et al. Polarity and temporality of high-resolution y-chromosome distributions in India identify both indigenous and exogenous expansions and reveal minor genetic influence of Central Asian pastoralist. *American Journal of Human Genetics* 2006; 78: 202-221.



- <sup>8</sup> Egorova op.cit. note 5.
- <sup>9</sup> R.S. Sharma. 1999. *Advent of the Aryans in India*. New Delhi: Manohar; R. Thapar. 2002. *The Penguin History of Early India: From the origins to AD 1300*. London: Penguin; T. Trautmann. 1997. *Aryans in British India*. Berkeley: University of California Press.
- <sup>10</sup> K. Roy et al. 2005. *The Vedas, Hinduism, Hindutva*. Kolkata: Ebong Alap.
- <sup>11</sup> C. Jaffrelot. 1996. *The Hindu Nationalist Movement and Hindu Politics, 1925 to the 1990s*. London: Hurst.
- <sup>12</sup> N. Dirks. 1997. Recasting Tamil Society: The Politics of Caste and Race in Contemporary Southern India. In *Caste Today: SOAS Studies on South Asia. Understandings and perspectives*. C.J. Fuller, ed. Oxford: Oxford University Press: 263-296.
- <sup>13</sup> S. Bayly. 1995. Caste and 'Race' in the Colonial Ethnography of India. In *The Concept of Race in South Asia*. P. Robb, ed. Delhi: Oxford University Press.
- <sup>14</sup> R. Lewontin. The Apportionment of Human Diversity. *Evolutionary Biology* 1972; 6: 391-398.
- <sup>15</sup> R.J. Herrnstein and C. Murray. 1994. *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press.
- <sup>16</sup> For a critical discussion of the HGDP see A. M'charek. 2005. *The Human Genome Diversity Project. An Ethnography of Scientific Practice*. Cambridge: Cambridge University Press; J. Reardon. 2005. *Race to the Finish: Identity and Governance in an Age of Genomics*. Princeton and Oxford: Princeton University Press.
- <sup>17</sup> T. Duster. 1998. Buried Alive. The Concept of Race in Science. In *Genetic Nature/Culture. Anthropology and Science Beyond the Two-Culture Divide*. A.H. Goodman, D. Heath and M.S. Lindee, eds. Berkeley, Los Angeles, London: University of California Press: 258-278.
- <sup>18</sup> S. Epstein. 2007. *Inclusion: the politics of difference in medical research*. Chicago: University of Chicago Press.
- <sup>19</sup> Duster op. cit. note 17; T. Duster. Race and Reification in Science. *Science* 2005; 307: 1050-1051; J. Kahn. How a Drug Became "Ethnic": Law, commerce and the production of racial categories in medicine. *Yale Journal of Health, Policy, Law and Ethics* 2004; 4: 1- 46.
- <sup>20</sup> K. Wailoo and S. Pemberton. 2006. *The Troubled Dream of Genetic Medicine: Ethnicity and innovation in Tay-Sachs, cystic fibrosis, and sickle cell disease*. Baltimore: The Johns Hopkins University Press.
- <sup>21</sup> Ibid, pp.17-20.
- <sup>22</sup> Ibid, p.10.
- <sup>23</sup> Duster op. cit. note 17.
- <sup>24</sup> Wailoo and Pemberton op.cit. note 20.
- <sup>25</sup> R. Benjamin. A Lab of Their Own: Genomic Sovereignty as Postcolonial Science Policy. *Policy and Society* 2009; 28: 341-355.
- <sup>26</sup> Ibid, p.342.
- <sup>27</sup> Reich et al op.cit. note 2, p.489.
- <sup>28</sup> Ibid, p.489.
- <sup>29</sup> Ibid, p.489.
- <sup>30</sup> Ibid, p.489.
- <sup>31</sup> Ibid, p.489.
- <sup>32</sup> Ibid, p.489.
- <sup>33</sup> Ibid, p.489.
- <sup>34</sup> Here Reich et al make a reference to N. Dirks 2001. *Castes of Mind: Colonialism and the Making of Modern India*. Princeton: Princeton University Press.
- <sup>35</sup> Reich et al, op.cit. note 2, p.490.
- <sup>36</sup> Dirks op.cit. note 34.
- <sup>37</sup> Reich et al, op.cit. note 2, p.491.
- <sup>38</sup> Ibid, p.489.
- <sup>39</sup> Ibid, p.493.
- <sup>40</sup> Ibid, p.492.
- <sup>41</sup> K. Ray. Genetically, As Good As Chalk and Cheese. *The Deccan Herald*. 24 September 2009. <http://www.deccanherald.com/content/27013/genetically-good-chalk-cheese.html> (accessed 26 December 2009).

- 
- <sup>42</sup> K. Tharoor. Tracing the Fissures in India's Society, *The Guardian*. 10 October 2009. <http://www.guardian.co.uk/commentisfree/2009/oct/10/india-genome-society> (accessed 26 December 2009).
- <sup>43</sup> Aryan Invasion Theory a Myth: Study. *Jai Bihar* 25 September 2009. <http://jaibihar.com/aryan-invasion-theory-a-myth-study/200912693.html> (accessed 26 December 2009).
- <sup>44</sup> M. Henderson. India's Caste System Descended From Two Tribes, 'Not Colonialism'. *Times Online* 24 September 2009. <http://www.timesonline.co.uk/tol/news/science/genetics/article6846424.ece> (accessed 26 December 2009).
- <sup>45</sup> Tharoor, op.cit. note 42.
- <sup>46</sup> Reich et al, op.cit. note 2, p.491.
- <sup>47</sup> Ibid, p.491.
- <sup>48</sup> Physorg.com. Genomic Research Shows Indians Descended From Two Groups. 23 September 2009. <http://www.physorg.com/news172931737.html> (accessed 26 December 2009).
- <sup>49</sup> B.J. Hardy et al. From Diversity to Delivery: The Case of the Indian Genome Variation Initiative. *Nature Reviews Genetics* October 2008; S9-S14.
- <sup>50</sup> Ibid., S9-S10.
- <sup>51</sup> Reich et al, op.cit. note 2, p.491.
- <sup>52</sup> Ibid, p.491.
- <sup>53</sup> A. Petryna. 2009. *When Experiments Travel: Clinical trials and the global search for human subjects*. Princeton: Princeton University Press.
- <sup>54</sup> Hardy, op.cit. note 49.
- <sup>55</sup> K. Sunder Rajan. 2006. *Biocapital. The constitution of postgenomic life*. Durham and London: Duke University Press.
- <sup>56</sup> N. Rose and C. Novas. 2005. Biological Citizenship. In *Global Assemblages: Technology, Politics, and Ethics as Anthropological Problems*. A. Ong and S. Collier, eds. Malden: Blackwell: 439-63.
- <sup>57</sup> Reich et al, op.cit. note 2, p.491.
- <sup>58</sup> L. Nader. 1996. *Naked Science: Anthropological Inquiry Into Boundaries, Power and Knowledge*. New York: Routledge.
- <sup>59</sup> J. Hartigan. 2010. *Race in the 21st Century: Ethnographic Approaches*. New York, Oxford: Oxford University Press.
- <sup>60</sup> D. Hess. 1995. *Science and Technology in a Multicultural World. The Cultural Politics of Facts and Artifacts*. New York: Columbia University Press.
- <sup>61</sup> Egorova, op.cit. note 5.
- <sup>62</sup> Benjamin, op.cit. note 25.