Philosophy of Race *Versus* Population Genetics: Round 3

Geoffrey K Chambers*

School of Biological Sciences, Victoria university of Wellington, PO Box 600, Wellington, 6140, New Zealand

Abstract: The Social Constructionist and Biologically Realist views of human races are often presented as mutually exclusive alternatives. Surprisingly, this debate has its origin in work on blood group population genetics. A finding that the greater part of human genetic variation lies within populations, rather than between races, has led some to deny the reality of geographically limited biological clusters. An extension of this view maintains that any differences that may exist are distributed in clinal fashion along ancient human migration routes precluding reliable delineation of racial clusters.

Here I argue that new genetic data adequately demonstrate that statistically significantly differentiated human subgroups (aka biological races) do exist. Further, that the analytical methods used to reconstruct the history of these human clusters are themselves, in part, social constructs. Therefore, these two contrasting philosophical viewpoints may be seen as capable of working together. Indeed, they are a sub-element of a larger metaphysical debate regarding the reality or otherwise of race, social or biological. I conclude by stressing the importance of racial definitions as regards the collection of reliable census information, the formation of equitable social policies and better informed medical decisions, particularly those involving the prescription of pharmaceuticals.

Keywords: Biological realism, Social constructionism, Genetic variation, Human races.

1. INTRODUCTION

Ideas about human races have a long history and now occupy an entire sub-discipline of Philosophy. Consequently, this subject has a huge literature and no single article can capture every facet of the rich history of the field. This abundance is reflected in the extensive reference list at the end of this article, which is chosen so that readers may as far as possible gain access to the entire area. A great deal of recent discussion has centred around the academic tension between those who think that human races can be mapped directly on to some biological property or properties (*Biological Realists; BR*) versus those who argue that the idea of races emerges from societal processes and our imaginations (*Social Constructionists, SC*). In this article I will argue that at best this is a form of intellectual pseudo-competition because these ideas cannot legitimately be seen as genuinely exclusive alternatives. Rather, in my view, it is better to ask if each in turn is true (and I will argue that they are) and then if they can live together (and I will argue that they can). From here, I will take up Lemeire's point [1] that this is all part of a bigger question; *i.e*. are human races (social or biological) real or not?

The whole BR vs. SC debate is made even more difficult by problems around communication having to do with the meaning of the word '*race*' itself in this context. First, the term is used as a descriptor at all and every level of human population structure; *i.e*. from the totality of the human race in all its aspects all the way on down to specific populations, such as the Maori race or the Indian race, etc. Second, over the last one hundred years common practical usage of the term has slowly moved from an ancestry-based concept using phenotypic characters, through ethnicity (which combines ancestry with cultural practice) to today's democratic standard of ethnic affiliation by self-declaration [2, 3]. There can be very little doubt that this last definition is a social construct and indeed the consensus of received wisdom from Sociologists who seem to be resolutely in the Social Construction camp. However, empirical genetic data [4] show that even this system of classification has some biological basis. Finally, the contest between definitions is frequently bundled with the larger question: Does race qualify as a real thing (or real kind)? Some may feel that in part this comes down to how one might choose to define race. Thus, many writers appear to treat biological races as real and social constructs as not real. I think that doing so may be premature because in my view the SC and BR positions at best only represent a pseudo-contest distinct from the real or not question.

The proximal origin of the present BR vs. SC debate can be traced to a single paper [5] by Lewontin which provides a preliminary quantitative estimate of genetic variation within and between human populations. His finding that the former is much larger than the latter has been widely adopted as a platform for the SC position; a tradition that continues even to the present, for example see [6]. Remarkably, this position has been

---

*Address correspondence to this author at the School of Biological Sciences, Victoria university of Wellington, PO Box 600, Wellington, 6140, New Zealand; Tel: +644-463-6091; Fax: +644-463-5331;
E-mail: geoff.chambers@vuw.ac.nz

E-ISSN: 2410-2806/17

maintained despite Edwards' clear explanation [7] why correlated genetic signals still allow reliable partitioning of humans into clusters. I begin my discussion with a brief retelling of this story and proceed to show how this argument has developed up to the present day. I conclude by making a claim to demonstrate that the SC and BR schools do not represent a dichotomy and, further, that genetic differences between population groups are too important to ignore, particularly in regard to their medical significance. Throughout this work I will make reference to key papers that represent especially influential contributions and from which readers should be able to reconstruct the wider literature.

2. THE DEBATE

2.1. Round 1: Edwards *Versus* Lewontin

In 1972 Lewontin published his analysis of blood group data from around the world [5] and was able for the first time to provide hierarchical quantitative estimates for the distribution of this form of genetic variation. His choice of target was astute because the genetic bases of these polymorphisms was well established and abundant, geographically widespread data were available from hospitals worldwide. His estimates have been included in numerous publications, but they are well worth repeating here. He found (p. 396) that 85.4% of total species diversity is contained within populations with a further 8.3% accounted for by differences between populations within a race and only 6.3% is accounted for by racial classification. This allowed him to make the following claim about:

"the largest part by far of human variation being accounted for by the differences between individuals"

Which in turn lead to the bold conclusion:

"Human racial classification is of no social value …. no justification can be offered for its continuance"

Today's investigators might offer many criticisms of Lewontin's experimental design because human taxonomy has advanced considerably since then. For instance, his population categories include one called Oceanian. This is worrying on at least three counts. First, like Amerind, it is derivative of other Asian categories. Second, it is formed by reticulation and admixture and (between Mongoloid and Australoid groups). Third, it contains representatives from Polynesian and Melanesian populations who have different admixture fractions [8, 9]. In addition, the placement of some national groups into racial categories is questionable. For example, Malayans and Filipinos are grouped with Mongoloids when they have greater affinity with the two Oceanic subgroups listed above. Ditto Madagascans who are grouped with Black Africans. There is no doubt that these complications do add noise to the signal. However, it probably matters little because others have often replicated these findings in various types of genetic survey since this time and consistently validated the general conclusion; see [10] for a list of subsequent studies. So, although perhaps contemporary commentators should be discouraged from quoting these actual figures, Lewontin's general observation [5] remains robust and his paper continues as a genuine celebration of human individuality.

The more important question is whether the conclusion he draws from it is correct. Edwards [8] makes a strong claim that Lewontin's observation is a red herring in the race debate. He carefully explains that because the frequencies of allelic variants at different loci are correlated, then these data can reliably be used to cluster groups regardless of the overall distribution of this variation. This concept has been well illustrated in theoretical diagrams and lengthy explanations [11,12]. Thus for example, if one has loci A and B with alternate alleles A_1 and A_2 and B_1 and B_2 and both the A_1 and B_1 alleles are 10x more common in population **X** than population **Y** then scoring these markers together can be used as a diagnostic to assign individuals to these groups; here an $A_1A_1;B_1B_1$ homozygote is 10, 000 times more likely to be an **X** rather than a **Y**. The above references have much more elaborate expositions of the mathematic logic behind the wider concept. Later papers from Lewontin and colleagues [13,14] do not entirely dispute this view, but do reaffirm their earlier findings based on newer microsatellite data and suggest that the clusters may not be as clearly defined as some scholars might think. This view has been taken up at length by Templeton [15] and I will return to this point in Round 3 below.

Ultimately Edwards [7] is drawn to the conclusion that:

"it is not true to say … that you can't predict someone's race from their genes"

One might think that this struck a death blow to the SC school, but far from it. Instead, an extended debate sprang up regarding the very nature of population genetic analysis and some [16, 17] still cling to their original interpretation regarding the significance of Lewontin's partition.

2.2. Round 2: The Sociological Establishment Fights Back

It is ironic that SC has such strong foundations in population genetics when its response to later developments in the field has been to go on the defensive. Indeed, their position was already strongly entrenched with SC endorsements coming from both the American Anthropological Association [18] and the American Sociological Association [19] and see [20]. The language used by some members of the SC school is colourful indeed with biological race called "*fiction*", "*a unicorn*" and "*the phlogiston of our time*" ultimately destined to "*follow the flat Earth into oblivion*" – see Sesardic [21] for an account. Sussman has devoted an entire book [22] to present race as "*The troubling persistence of an unscientific* idea". Those social scientists who have strayed from the orthodox path have been chastised because they were arguing for "*the utility of genetic data in sociological analyses*"; see [23] for a list of the culprits. This is not a conspiracy per se by my estimation, but more of a social and intellectual programme motivated in large part by a perceived need to protect against an imagined re-emergence of Eugenics and the genuinely worst excesses of racist policies. So, the stakes are high and the response is vigorous.

In turn, the SC programme has been cited as an outstanding example of left-wing thinking and political correctness taken to excess via "*groupthink*" and "*confirmation bias*" [24]. This last descriptor is a euphemism for the suppression of opposing viewpoints. This author goes on to quote the example of a letter to the New York Times signed by more than one hundred sociologists denouncing Wade's book, *A troublesome inheritance* [25] as inaccurate, but failing to provide any concrete examples. A key issue was Wade's view that variation has developed in 8% of the human genome since the first Out of Africa migrations and this partitions humans into 5 easily distinguishable races which have arisen in response to regional environmental challenges. These people are Africans, East Asians, Caucasians, Native Americans and the Aboriginal peoples of Australia and Papua New Guinea (herein after termed a K = 5 partition solution). Whatever one may more generally think about Wade's idea as expressed in this work, the K = 5 model must be viewed as pretty reasonable and many others have come up with more or less the same partitions (see later). However, should one prefer to reject it, then this should be done supported by reasoned argument, rather than simply dismissing it out of hand.

Finally, I note in passing a further interesting feature of social behaviour in this controversy. The views of population geneticists who advocate the SC position seem to be granted considerable weight; *e.g.* particularly scholars such as Long and Templeton [15, 26]. Often this seems to be disproportionate regarding their particular contributions regarding the question. Whereas other scholars working in the same area rarely seem to get a mention in the SC canon. I term this phenomenon *'artificial inflation of the academic authority of presumed defectors'*. Appiah [27] and Fullerton [28] present detailed accounts of the attacks by SC proponents on the BR programme.

Thus, sadly, rather than embrace new genetic findings the SC school has elected to challenge them in a multiplicity of ways. Frank [23] shows how the battle lines were drawn by Duster who in his 2005 presidential address [29] to the American Sociological Association called upon members to

> *"prepare to defend against the genomic data juggernaut heading their way down the pike" [23]*

Many authors have contributed to the debate which followed, but the writing of Kaplan, Winther and colleagues probably comes closest to the SC/BR interface in their detailed examination of the way that population genetic analysis works. I will now explore these contributions in some detail together with my analysis and contrasting views from other commentators such as Sesardic and Spencer. Kaplan and Winther [30] begin by crediting Mills [31] with the origin of the constructionist/realist classification and present his taxonomic view of the metaphysics of race. They next embark on an extended historical tour through the Edwards/Lewontin debate and seek resolution by considering genetic variation as diversity, differentiation and heterozygosity , including a valuable plain language explanation of the mathematics involved. Ultimately they arrive at what they recognise as *Objectivist Constructionism* to use one of Mills' original categories. Here, they do recognise that there must be some genetic structure in human populations, but that it is insufficient to allow classification and thereby avoid becoming '*prisoners of abstraction*' because, they claim that '*using biological theory to ground race is a pernicious reification*'. Fine words, but wide of the mark, I think, because they do not adequately dispose of Edwards' facts [7] about correlation among allele frequencies at different loci. There are further problematic issues with their otherwise interesting paper which are covered by Spencer in a detailed and more formal philosophical analysis [32].

Winther moves on to consider the use of genetic data in phylogenetic and clustering analyses [6], which is more directly to the point in relation to trying to define biological races. He makes the important point that, although these are directed towards the same ends, their methodologies are different and include different assumptions. The fact that operators must make decisions about setting parameters in these analytical models and assign diagnostic limits to define clusters is highly significant in my view and a point that I will return to later. One should also note that these two general approaches are different. Clustering analyses (unsupervised fuzzy genetic clustering, multidimensional scaling and principal components and principal co-ordinates) return differentiated groups without indicating any preferred hierarchical ancestral history linking them. Phylogenetic analysis has many different forms (including phenetic distance, cladistics, maximum likelihood, Bayesian, Hadamard and network methods), but all of them return the history, even if they may not always agree with one another or give a clear clustering scheme; Long and Kittles [26] cover many of these same points with worked examples [26]. This puts a premium on the investigator's ability to interpret the output. One example of a study which uses both types of analysis is that of Friedlaender *et al.* [33] on Pacific Islanders and I will refer to this again too. Winther also highlights the concerns expressed here about the subjective human factor in input assumptions for models and output interpretation [6]. Kaplan and Winter [34] further suggest that even having more data from genomics is not necessarily going to make things better and also stress the clines *versus* clusters dilemma [35] with regards to interpretation or phylogenetic patterns.

Finally, the wider debate has been examined in detail by Spencer [36] who recognises four principal semantic and metaphysical objections that philosophers of race have put forward against biological realists who rely on the new genetic clustering studies to support their own position. In my view this is a particularly valuable contribution and a helpful source of reference material. So Spencer's ideas clearly merit consideration in depth in the context of this article while noting that they are primarily directed towards US racial groups. He begins from the position that human population structure really does look racial (a central theme of this commentary) and surveys a number of studies that illustrate this view including [33]. He does note that various studies of worldwide populations have returned different numbers of major clusters (aka K values = statistically supported data partitions) when examined using unsupervised fuzzy genetic (UFG) methods; clustering algorithms such as STRUCTURE, frappe, admixture, etc. and see [35] and [37] for lists of names and explanations of how they work). However, many such analyses do converge on what I earlier termed as a 'K = 5 solution' viz N Africa, S Africa, W Eurasia, E Eurasia and Americas plus Oceania [38] and recognising that this differs from the scheme in [25]. These five major groups map closely with the official US racial groups. Spencer also notes that biologists routinely use these methods to classify within and between species of clusters of other organisms without raising concerns about what they are doing or why. Beyond this commentary the present author notes that clustering and admixture analyses of human populations have now reached an exceptionally powerful stage of their development and have been shown to be effective at a national scale [39] or across continental Europe [40] and Island Southeast Asia [41]: also see [42, 43] for recent global perspectives based on very large datasets. However, it must also be recognised that other workers have presented K = 3 and various K = 4 solutions for humans as a whole. And, further, that when using tree-building methods, it is the analysts themselves that finally decide the best K number equivalent solution by lumping and splitting partitions albeit guided by a statistical test.

In Spencer's view [36] the first (semantic) objection to a genetic clustering diagnosis of race is *discreteness i.e*. that clinal variation and/or admixture obscures clusters [35]. Proponents of this view worry about such things as the hypodescent or 'one-drop' rule that has in the past been used to classify a person with any fraction of African descent as Black. Spencer diagnoses this (correctly in my opinion) as a signal to noise problem and not one that creates significant practical difficulty quoting the study of Guo *et al.* [4] as supporting evidence. I think one must agree with this assessment. The second (semantic) objection is *visibility i.e*. that (US) racial groups are recognisable by their phenotypic characters. While this idea [44] may be valid under some forms of racial discourse it is both retrograde in the sense of changing definitions discussed earlier and unhelpful. Spencer explains that many observers might not easily distinguish Africans from Melanesians (another recurrent theme of the present article).

The third objection (metaphysical) is parochial to some extent, but nonetheless called the *very important objection.* This states that for US racial groups (*e.g*. as captured by US Census Bureau, USCB) to be biologically real, they must form a very important biological classification sometimes equated to sub-species status. Objectors such as Templeton [45] and maintain that fixation index (F_{ST}) values between the human K = 5 groups are smaller than those for chimpanzee subspecies, therefore there are no human subspecies and thus the K = 5 groups are not biologically important. Spencer argues that importance per se is not the issue. What matters is if the groups are real. In my view, the objection is simply unfair and I will return to this point in my discussion of human taxonomy later. The fourth (also metaphysical) objection is the *objectively real* objection *i.e*. that human races do not exist independent of human thought processes. Spencer argues that a willingness even to raise this objection shows that we should not continue to contrast biological realism with social constructionism. Putting it another way, stars are stars regardless of whether we choose to gaze upon them or not. The debate begins once we ask if using the signs of the zodiac is the best way to organise our thinking about them.

2.3. Round 3: A Proposed Road to Redemption and Reconciliation

In the paragraphs that follow I will attempt to show that 1) human races are real biological entities theoretically accessible via more or less objective means, but 2) practically inaccessible without the intervention of human agency and therefore also social constructs to some greater or lesser extent.

I begin by returning to the value of term 'race' as a descriptor. In previous works my colleagues and I have rejected the use of this term entirely [2, 3]. We do not do this just because the term is emotionally and politically loaded. We do it, first because it is used to describe units at all levels of the human taxonomic hierarchy, and second because the definition has been shifting over time. In this latter regard, usage has progressed from one based on ancestry, appearance and behaviour, via the idea of ethnicity (being some combination of ancestry and cultural practice) through to self-declared ethnic affiliation. This last form has found widespread appeal and application and I think most would agree that it is for the best that people should be known by the collective term that they prefer to use themselves. This definition is not entirely ancestry free, but it can be. Consider, for example the Cherokee Freedmen who are the descendants of African slaves, but who identify as Amerindians [46]. There are two added complications, first that different agencies collect race/ethnicity data using different questions and so collect information from different groups of people under a single heading and second that official forms sometimes do and sometimes do not allow responders to nominate multiple categories. For instance, it is a common experience for someone meeting a US citizen for the first time to learn early on in their relationship that their new friend is actually Italian, German or Irish. Worse still, when official categories require, or seem to require, a knowledge of ancestry, then responses may be unreliable. Family histories may be romanticised, people may be uninformed adoptees or have missed paternity or simply not understand the question.

I now return to the beginning of the entire discussion taking my lead from Lewontin [5] and Spencer [36] to claim that there seems to be something intuitively obvious about the biological reality of human races. Presented with a simple K = 3 thought experiment most people would probably agree that one could distinguish an indigenous African Zulu from a Scandinavian and from a Han Chinese, just based on appearance. Again most, if not all, scientists would anticipate confirmation by genetic testing. However, once we begin to recruit a few other volunteers for our imaginary identity parade, then things get more complicated *e.g*. try adding an Arab, an Amerindian or a Polynesian into the line-up. However, despite the addition of one, or even all, of them one would still expect genetic diagnosis to work. The key question becomes; How far may we continue this process. In other words; How many human groups can we reliably recognise using genetic methods? One should note along the way that this answer will probably depend on how many samples one can analyse, how many loci one can examine and how extensive and reliable are the data for reference populations.

I am going to argue that these questions have parallels with what is known to biologists and philosophers as '*The species problem*' – see [47] for an earlier discussion of this topic. I begin like [12] by asking if humans are polytypic? Answering this question starts right at the top and discussion around this problem will be used help to draw one of the major conclusions in this work concerning the intrusion of subjectivity in phylogenetic analyses. This matter comes down to needing to decide where to draw barriers between clades and how one can make these decisions. It is like deciding just how many celestial bodies should make up the star sign Virgo. For humans, one must decide if extinct closely related hominids including, but not necessarily limited to, Neanderthals and Denisovans should be included in *Homo sapiens*; see [2, 3] for a discussion. Nobody to my knowledge has suggested that the even older *H. erectus* should be lumped with *H. sapiens*. However, there may be some die hard multi-regionalists who might even consider this prospect seriously. There is no doubt that each of these three more modern and more closely related taxa named above forms a distinct clade in its own right in any and all types of phylogenetic exercise and do so despite fairly extensive previous interbreeding. The issue is just where to draw the bounds around each cluster. So, here we have a genuine example of the interface between BR and SC contributions to our understanding of just what it is to be human. We are real, as are our extinct cousins, and thus BR entities, but do we want to admit them into our family as part of a SC community?

Let’s now move on down a step and suppose for the sake of progressing the debate that we have decided to exclude our extinct relatives. When we examine the *H. sapiens* tree (or more correctly trees) one can see at once that it has obvious internal clade structure that maps fairly well on geographical distributions and having all African lineages as basal. So should we subdivide the tree? This is not a new idea as [48] had a K = 4 solution with *H. s. europeus*, *H. s. asiaticus*, *H. s. americanus* and *H. s. afrer*; and see [49] for further discussion around this topic. Here it is correctly recognised that in Zoology the first taxonomic division of species is into sub-species and one might even consider applying the well-known 75% diagnosability rule [50] to aid in decision making. Now we are in real trouble. First, nobody seems to want to even think seriously about splitting humans into subspecies (except to run the straw man argument in comparison with chimpanzees). Second, many subclades of humans are better than 75% diagnosable – there are up to 38 of such things alive in the world today; *i.e.* human taxa recognisable as distinct individual species under the Phylogenetic Species Concept (PSC) according to [12] based on Cavali-Sforza *et al.*’s sub-continental populations [51, 52]. I confidently suggest that nobody wants this solution, defensible or otherwise. In zoological taxonomy the rank below subspecies is often known as ‘a race’ although some, like [45], treat race and subspecies as more or less equivalent terms. A zoological race is generally taken as being a group or population of animals or birds with a recognisable (usually phenotypic) character, such as particularly dark skin or plumage colouration. Such characters often have variable expression and do not meet the 75% diagnosability requirement. Confusingly, these are also sometimes known as ‘morphs’, particularly when they are intermixed with other types within populations. Neither of these terms is easily transferable to humans and the use of the term race in this way by biologists is one further reason why Callister *et al.* [2, 3] prefer not to use it in connection with humans.

Before leaving this topic and considering the analytical methods themselves some mention must be made of the fairly widespread claims that human genetic variation is clinal and not discrete [35]. There can be little doubt that this proposition is true to some extent, but the critical question would seem to be; Does this obscure the boundaries of the major clades to such an extent that they are ambiguous? If this is the case, then the clades or clusters are in the eye of the beholder and cannot be much other than social constructs. If human evolution had occurred simply by a limited number of discrete migrations out of Africa along geographic lines of ancestor-descendant lineages, *i.e.* in a strictly tree like fashion, then phylogeneticists would have little or no problem reconstructing their history from DNA sequences. Clearly, this is not the case and the more we discover about human evolution, the more it is not the case. We now have examples of back migration, semi-continuous gene flow along lineages and reticulation (sometimes gender-biased) between lineages [15]. The process is somewhat akin to the formation of what are known as ‘*ring species*’ in birds (*e.g.* some types of seagulls and warblers). Here adjacent populations may be more or less indistinguishable, but those at the ends of the distribution (*i.e.* at the point where the ring closes) are distinct and no longer able to interbreed. Of course, the human situation is not quite this extreme, but the ring species model is a useful guide when thinking about clinal variation.

These features all add noise to the phylogenetic signal, but not sufficient to obscure it. The model study of Friedlaender *et al.* [33] provides a good illustration of the power of contemporary analytical methods to identify 5 local clusters and understand their relationships. Here, a Polynesian population (NZ Maori) were shown as a discrete cluster, but with a phylogenetic placement linking them to both Austronesian (Mongoloid) and Papuan (Australoid) peoples. The other studies referred to earlier in Europe [40] and SE Asia [41] are even more impressive examples of how this all works. Two other lessons may be taken from the Friedlaender *et al*. study cited above. First, their result was achieved with only relatively small sample sizes and using relatively few loci compared with contemporary standards. Second, it illustrates how UFG and phylogenetic analyses may be combined to advantage to resolve complex historical patterns of relationship. In their study the ancestral Polynesian lineage is reticulate and formed by extensive gender-biased admixture as recently as around 3000 ybp.

I turn, at last, to consider subjective elements in cluster and phylogenetic analysis procedures. This is a fairly well-recognised issue among specialists and I am certainly not the first to mention it in relation to human evolutionary studies. First, come decisions about which populations/datasets to sample, how many individuals to include, which loci to collect data from and finally which reference populations to include. Many of the decisions will be driven by the multi-factorial contingency that is the reality of research work everywhere. Nonetheless, it is undeniable that all require human input. These factors are not denied by investigators, but often go unnoticed or unremarked. Next, come the analyses, the UFG methods are relatively innocent in this regard. One puts the data in and the truth emerges – or so one might hope. One still has to pick a method, decide on how and where to place divisions between clusters and then to allocate names to them. Using multiple methods and hopefully obtaining congruent patterns is a good way to try to avoid biases due to analytical method. The various versions of the STRUCTURE algorithm have an internal likelihood routine that picks a value for K (the number of partitions best supported by the data). All well and good and driven by the data. However, in practice the analysis often returns clustering patterns with K + 1 and/or K – 1 partitions that are as good as or almost as good as the K partition arrangement. In fairness to investigators it should be recognised that they do usually present these alternatives either in the report itself or include them in online supplementary material. This ensures that nothing is hidden and the relative merits of all alternatives can be scrutinized. But in the end a human agent must decide which is best.

Phylogenetic tree construction algorithms are objective, but require humans to set a whole range of parameters before they can operate. These include; tree searching method, definition of the best tree, model of DNA sequence evolution and statistical testing of alternatives. Finally, one must decide which branches to lump into a clade and if a consistent rule is to be applied to this process across the whole tree. This is a particularly difficult task because it often amounts to trying to impose a discrete categorization on a semi-continuously distributed output. Some of the other problems can be reduced *e.g*. by using a computational routine like Model Test to select the model of evolution for use in maximum likelihood analyses. Also using multiple tree building methods helps to build confidence when they converge on identical or near identical solutions. In summary, everything in this and the preceding paragraph goes to build my case that what may very well be biologically real entities in terms of distinct genetic cluster of humans can only be discovered by processes involving more or less direct human intervention. In other words, although races may be biologically real entities, they are necessarily social constructs too.

3. WHY IT ALL MATTERS SO MUCH AND TO SO MANY

By my estimation there are three principal reasons why increasing our understanding of human races is important. The first is to gain insights into human evolution and has already been explored fairly thoroughly above. The second is to promote social equity and combat discrimination. One of the reasons that social scientists seem to abhor the BR model of race is that they fear it may work against these ends [16, 53]. This view seems to be founded in racial constructs popular in the C. 19th. Here biological races are claimed to exhibit characteristic phenotypes (mainly through their physical appearance) and behaviours. This leads to the formation of prejudicial stereotypes; so, if you are Chinese, then you must like table tennis and be very good at playing it or be addicted to gambling and smoke copious amounts of opium etc. These sorts of statements stem from common type of transition from sense to nonsense founded on some easily recognised lapses in logic. I describe these below, but first some facts about life. There really is a clear genetic link between ancestry, appearance and behaviour. One should never be afraid to recognise these relationships. All human physical, intellectual and behavioural characters do have some greater or lesser genetic component. For individuals, these are shaped by our package of allelic variants, but also influenced by our environment and the interaction between these variants and the environment (aka GxE effects). Recent advances in the field of epigenetics show how environmental influences can alter gene expression, closing the circle of cause and effect. Different traits have different contributions from each of these factors. Their relative contributions are unknown and, for technical reasons, unknowable without making assumptions in analysis of twin and adoption study data. The much criticised idea of *biological determinism* is said to arise from holding the view that heritage is paramount. This view seems to be reified in the popular phrase "*It's in our DNA*", but nothing could be further from the truth as the usage implies the meaning "*It is traditional for our social group to do this sort of thing*".

In short, Biological Realism is at heart an essentialist view, but not in the sense of equating to Biological Determinism. It should not be invoked to explain causality. Thus speaking Swahili may fairly be said to run in families, but not for exclusively genetic reasons. There is a genetic component in the sense of general and specific language ability. There is also an environmental component in the form of exposure to written or spoken Swahili. Finally, there is a GxE component that relates to one's response to teaching methods and circumstances *e.g*. childhood exposure to native speakers vs. missionary training school. Different people will respond differently to these various experiences and teachers. Thus the BR view is determinist, but only conditionally determinist at best.

It is my view that we should never fear or deny empirically observable differences between racial groups, but we should always be very careful about jumping to conclusions about why they exist. In New Zealand many fast food outlets specializing in the popular delicacy known as fish and chips are run by citizens of Greek descent in Wellington, but by Chinese and Vietnamese in Christchurch. Many dairies (our equivalent of the US convenience store) are owned and staffed by Indians. It is highly doubtful that either of these commercial situations has much, if any, genetic basis. The failures of logic that would occur in stereotyping all Indians as shopkeepers comes from regarding genetics (in the sense of ancestry) as fate and imposing group observations on individuals. Only in special instances can genes decide fate. This occurs, for example, in the case of Huntingdon's Disease where receiving just one copy of a mutated gene leads to tortuous physical and mental decline from mid-life onwards. However, this type of situation is rare and most human traits have bewilderingly multi-factorial genetics. Thus, if you want to be a basketball star it certainly helps to be tall. There is no question that human stature has some genetic determinants. Equally, there are nutritional and environmental determinants too. Sadly, just being tall does not mean that you will eventually be drafted to join an NBA team, regardless of the encouragement you might receive from your high school coach. Tagging all tall people as potential basketball players shows the second type of logical lapse in stereotyping – imposing perceived group values on individuals. This is what leads to discrimination and usually involves a great deal of false inference about cause and effect. The outcome is a sort of '*blame game*' where real societal disadvantage is projected back onto the disadvantaged group(s) and seen as a product of their intrinsic moral or biological failings.

At this point the realism debate takes centre stage [1]. What matters is whether race is real in a metaphysical sense. It makes scant little difference if this is biological realism or social realism. It is not easy to answer this question. One requires a two-step philosophical programme 1) to decide and describe the set of properties that something needs to be a real kind, 2) test how well human races (*i.e*. as defined the BR and SC models) match the requirement(s) for being real things. Indeed, it seems to be the very idea (correct or not) that races are real which fuels racism. Thus, it serves to rationalise what seems to be an inherent, intuitive and adaptive xenophobia. Judgemental comparisons such as "*These people are just cockroaches*" are tacked on the side for good measure. Here race or ethnicity is just used as a flag to signal out group members. Perhaps it would be better to turn the entire problem on its head and ask; If human races genuinely are real kinds, then how can we prevent ourselves from becoming racist (in the pejorative sense of this term).

The third reason why this larger debate is important concerns medicine. Health care is becoming an increasingly genetic science. Hence, being able to interpret statistical data about self-identified race/ethnicity in relation to their patients is increasingly important for practitioners [54, 55]. The field of Pharmacogenomics has arisen through the realisation that new medicines researched by and tested in one ethnic population may not work as well in others. Events surrounding the drugs Bidil in the US and Iressa in Asia were taken as good first illustrations of this approach, but have since faced strong criticism from some [29]. Many otherwise beneficial pharmaceuticals are recognised as poisonous foreign compounds (which in truth they often are) by the human body. This means that medication is often quickly scheduled for detoxification in the liver and excretion as soluble metabolites via the kidneys. One of the principal detox systems uses the hepatic cytochrome P450 catalytic complex; see for example our own work on these systems in New Zealanders [56, 57]. This is encoded by multiple genes and which all have many variants thereby providing a selective advantage in wide-ranging biochemical response. As one might suppose, the frequencies of these variants differ between populations meaning that clearance rates will differ between them. This means that dose rates may need to be adjusted to suit patients based in part on their ancestry. And here it is ancestry that really matters, rather than ethnicity or ethnic affiliation; noting that in BR terms race and ancestry are exactly congruent concepts. This can be a tricky matter. Mersha and Abebe [55] Explain how the US Hispanic or Latino category includes Puerto Rican and Mexicans with the former having 18-25% African admixture and the latter 35-64% Native American admixture based on the figures from [58]. These difficulties are not insurmountable and can best be solved by taking a careful family history from patients according to Braun *et al*. (2007), but noting earlier caveats regarding the reliability or otherwise about a person’s stated knowledge of their ancestry.

CONCLUSIONS

In the account presented here I have argued that the BR and SC viewpoints on race are both valid and mutually compatible. I claim that the division between them derives in large part from the commentaries by others on the Edwards/Lewontin debate. Social scientists have, in my view, been strongly drawn to the SC conceptualisation of race in part driven by concerns about racist stereotyping and in part by fear of perceived biological determinist agendas and eugenic thinking. Their position has been strengthened by the more or less widespread adoption of self-declared ethnic affiliation as an official reporting category for race. I have presented the view that this does not make their position exclusively correct. Rather, I have attempted to show that an avalanche of new genetic and genomic data makes biological races as ancestral geographic population clusters real to the point where their denial is perverse. However, I hold back from taking an extreme BR position, because by my account recovering these clusters from the data is itself, at least partly, a social construction.

Thinking around this issue finds a ready application in New Zealand. The ancestors of Europeans and Maori first separated around 70 to 100,000 ybp as migrants from Africa. Since then their genomes have been subject to all the usual population level forces, mutation, genetic drift and natural selection etc. Their gene pools will have become different as a consequence. This expectation is now a fact. True much of their genetic material remains identical and thus is shared in common. But, there are plenty of differences too, and these can and have been measured in their descendants. Therefore, there is a genuine sense in which these groups absolutely must be real things. Surprisingly, it is not these real things them selves which feature in official statistics. The figures record how individuals see themselves, which adds noise to the ancestry signal. Worse still, many of these figures arise from processes, which depend on how others see the individual rather than vice versa which adds a further layer of complication.

I believe that it is important that this debate should continue in the interest of improved and consistent official reporting of race/ethnicity data [2, 3]. It is of vital concern that this should be done in medicine, so that all might have access to treatments best suited to their individual needs. Some may see this proposal in a negative light as calling for “*race-based medicine*”. I reject this opinion and describe my recommendation as being for genetically informed medicine. Duster provides a detailed and valuable history and overview of this area [16], but concludes that this form of medical practice is a bad thing because it leads to a “*reinscription of race*”. Not all commentators agree with this conclusion; see Duster [17] for a list of criticisms and an energetic response. For the present, allele frequencies and allelic associations in ethnic groups serve only as proxies to guide practitioners. Thus, it is that a diagnosis of sickle-cell anaemia is only *more likely* for patients with some degree of West African ancestry, than it is for Europeans. The disease does occur in Europeans and may also be present due to undisclosed or unknown ancestral admixture. We look to a future where these considerations will be pushed aside by the advent of personalised genomic medicine. For the present, it should be enough to have the courage to recognise what we truly are and where we have come from.

ACKNOWLEDGEMENTS

Geoff Chambers is grateful to Victoria University of Wellington for alumnus scholar support and to academic colleagues particularly Paul Callister and Robert Didham for providing comments, discussions and references.

CONFLICT OF INTEREST STATEMENT

The author affirms that he has no conflict of interest.

REFERENCES

[1] Lemeire O. Behind the realism debate: The metaphysics of 'racial' distinctions. Stud Hist Phil Sci Part C 2016; 59: 47-56. https://doi.org/10.1016/j.shpsc.2016.08.001

[2] Callister P, Didham R, Chambers GK. 'Marrying' demographic and genetic measures: An exploration of opportunities. NZ Sociol 2015a; 30: 229-233.

[3] Callister P, Didham R, Chambers GK. 'Marrying' demographic and genetic measures: New tools for understanding New Zealand Population sub-groups 2015b An extended online working paper (43 pp.) can be found at URL http://callister.co.nz/papers.htm

[4] Guo G, Fu Y, Lee H, Cai T, Mullan Harris K, Li Y. Genetic bio-ancestry and social construction of racial classification in social surveys in the contemporary United States. Demog 2014; 51: 325-329. https://doi.org/10.1007/s13524-013-0242-0

[5] Lewontin RC. The apportionment of human diversity. Evol Biol 1972; 6: 381-398. https://doi.org/10.1007/978-1-4684-9063-3_14

[6] Winther RG. The genetic reification of "Race"?, A story of two mathematical methods. Crit Phil Race 2014; 2: 204-223.

[7] Edwards AWF. Human genetic diversity: Lewontin's fallacy. Bio Essays 2003; 25: 798-801. https://doi.org/10.1002/bies.10315

[8] Kimura R, Ohashi J, Matsumura Y, Nakazawa M, Inaoka T, Ohtsuka R, *et al*. Gene flow and natural selection in oceanic human populations inferred from genome-wide SNP typing. Mol Biol Evol 2008; 25: 1750-1761. https://doi.org/10.1093/molbev/msn128

[9] Wollstein A, Lao O, Becker C, Brauer S, Trent RJ, Nürnberg P, *et al*. Demographic history of Oceania inferred from genome wide data. Curr Biol 2010; 20: 1983-1992. https://doi.org/10.1016/j.cub.2010.10.040

[10] Edge MD, Rosenberg NA. Implications of the apportionment of human genetic diversity for the apportionment of human phenotypic diversity. Stud Hist Phil Biol Biomed Sci 2015; 52: 32-45. https://doi.org/10.1016/j.shpsc.2014.12.005

[11] Sesardic N. Race: a social destruction of a biological concept. Biol Phil 2010; 25: 143-162. https://doi.org/10.1007/s10539-009-9193-7

[12] Woodley MA. Is Homo sapiens polytypic? Human taxonomic diversity and its implications. Med Hypoth 2010; 74: 195-201. https://doi.org/10.1016/j.mehy.2009.07.046

[13] Feldman MW, Lewontin RC. Race, ancestry and medicine. In: Koenig B, Lee S, Richardson S, editors. Revisiting Race in a Genomic Age; New Brunswick, NJ, Rutgers University Press 2008; p: 89-101.

[14] Fujimura JH, Bolnick DA, Rajagoplan R, Kaufman, JS, Lewontin RC, Duster T, *et al*. Clines without classes: How to make sense of human variation. Sociol Theor 2014; 32; 208-227. https://doi.org/10.1177/0735275114551611

[15] Templeton A. Biological races in humans. Stud Hist Phil Biol Biomed Sci 2013; 44: 262-271. https://doi.org/10.1016/j.shpsc.2013.04.010

[16] Duster T. A post-genomic surprise: The molecular reinscription of race in science, law and medicine. Brit J Sociol 2015a; (66): 1-27. https://doi.org/10.1111/1468-4446.12118

[17] Duster T. Response to comments on 'A post-genomic surprise'. Brit J Sociol 2015b; 66: 83-92.

[18] American Anthropological Association. AAAstatement on race; 1998. Downloaded from www.americananthro. org (accessed on 03-NOV-17).

[19] American Sociological Association The importance of collecting data and doing social scientific research on race 2003. American Sociological Association, Washington DC downloaded from www.asanet.org (accessed on 03- NOV-17).

[20] Byrd WC, Hughey MW. Biological Determinism and racial essentialism: The ideological double helix of racial inequality. Annal AAPSS 2015; 661: 8-22. https://doi.org/10.1177/0002716215591476

[21] Sesardic N. Confusions about race: A new instalment. Studies in the Hist Phil Biol Biomed Sci 2013; 44: 287-293. https://doi.org/10.1016/j.shpsc.2013.03.005

[22] Sussman RW. The myth of Race. Cambridge, MA, Harvard University Press 2016.

[23] Frank R. Back to the future? The emergence of a geneticized conceptualization of race. Annal AAPSS 2015; 661: 51-64. https://doi.org/10.1177/0002716215590775

[24] Tierney J. The real war on science. City Journal 2016 online at www.city-journal.org/html/real-war-science-14782.html (accessed on 11-NOV-17)

[25] Wade N. A troublesome inheritance. New York, The Penguin Press 2014.

[26] Long JC, Kittles RA. Human genetic diversity and the nonexistence of biological races. Human Biology 2003; 75: 449-471. https://doi.org/10.1353/hub.2003.0058

[27] Appiah A. Race, Culture, Identity: Misunderstood Connections In Appiah, KA, Gutman, A. editors. Colour Conscious: The Political Morality of Race. Princeton NJ, Princeton University Press 1996; p: 30-105.

[28] Fullerton SM. On the absence of biology in philosophical considerations of race. In Sullivan S, Tuana N editors. Race and Epistemiologies of Ignorance. New York, NY, State University of New York Press 2007; p: 241-258.

[29] Duster T. Comparative perspectives and competing explanations: Taking on the newly configured reductionist challenge to sociology. Am Sociol Rev 2006; 71: 1-15. https://doi.org/10.1177/000312240607100101

[30] Kaplan JM Winther RG. Prisoners of abstraction? The theory and measure of genetic variation, and the very concept of "Race". Biol Theor 2013b; 7: 401-412.

[31] Mills CW. But what are you really? The metaphysics of race. In Mills. CW editor. Blackness visible: Essays on philosophy and race. Ithaca, NY, Cornell University Press 1988; p: 41-66.

[32] Spencer Q. Biological Theory and the metaphysics of race: A reply to Kaplan and Winther. Biol Theor 2013; 8: 114-120. https://doi.org/10.1007/s13752-013-0095-1

[33] Friedlaender J, Friedlaender F, Reed FA, Kidd KK, Kidd, JR, Chambers, GK *et al*. The genetic structure of Pacific islanders. PLoS Genet 2008; 4: 173e190.

[34] Kaplan JM, Winther RG. Ontologies and politics of biogenomic 'Race'. Theoria 2013a; 60: 54-80.

[35] Kopec M. Clines, clusters and clades in the race debate. Phil Sci 2014; 81: 1053-1065. https://doi.org/10.1086/677695

[36] Spencer Q. Philosophy of race meets population genetics. Stud Hist Phil Biol Biomed Sci 2015; 52: 46-55. https://doi.org/10.1016/j.shpsc.2015.04.003

[37] Padhukasahasram B. Inferring ancestry from population genomic data and its applications. Front Genet 2014; 5: 204. https://doi.org/10.3389/fgene.2014.00204

[38] Rosenberg NA, Pritchard JK, Weber JL, Cann HM, Kidd KK, Zhivotovsky L A, *et al*. Genetic structure of human populations. Science 2002; 298: 2381-2385. https://doi.org/10.1126/science.1078311

[39] Leslie S, Winney B, Hellenthal G, Davison, D, Boumertit A, Day T, *et al*. The fine structure of the British population. Nature 2015; 519: 309-314. https://doi.org/10.1038/nature14230

[40] Lazaridis I, Patterson N, Mittnick A, Renaud G. Mallick S, Kirsanow K, *et al*. Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature 2014; 513: 409-413. https://doi.org/10.1038/nature13673

[41] Lipson M, Loh PR, Patterson N, Moorjani P, Ko YC, Stoneking M, *et al*. Reconstructing Austronesian population history in Island Southeast Asia. Nature Communications 2014; 5: 4689. https://doi.org/10.1038/ncomms5689

[42] Creanza N, Ruhlen M, Pemberton TJ, Rosenberg N, Feldman M, Ramachandran S. A comparison of phonemic and genetic variation in human populations. Proc Nat Acad Sci USA 2015; 112: 1265-1272. https://doi.org/10.1073/pnas.1424033112

[43] Duda P, Zrzavy J. Human population history revealed by asupertree approach. Sci Rep 2016; 6; 29890. https://doi.org/10.1038/srep29890

[44] Glasgow J. A theory of race. New York, NY, Routledge 2009.

[45] Templeton A. Evolution and notions of human race, In Losos JB, Lenski RE editors. How evolution shapes our lives. Princeton, NJ, Princeton University Press 2016; p: 346-361.

[46] Sturm C. Race, sovereignty and civil rights: understanding the Cherokee Freedmen debate. Cult Anthropol 2014; 29: 575-598. https://doi.org/10.14506/ca29.3.07

[47] Chambers GK. The species problem: Seeking new solutions for philosophers and biologists. Biol Phil 2012; 27: 755-765. https://doi.org/10.1007/s10539-012-9314-6

[48] Linnaeus C. Systema naturae per regna trianaturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis Vol. 1, 10th ed. Stockholm, Laurentius Salvius 1758.

[49] Batai K, Kittles RA. Race, genetic ancestry and health. Race Soc Prob 2013; 5: 81-87. https://doi.org/10.1007/s12552-013-9094-x

[50] Amadon D. The seventy-five percent rule for subspecies. Condor 1949; 51: 251-258. https://doi.org/10.2307/1364805

[51] Cavalli-Sforza LL. Genes, Peoples and Languages. San Francisco, CA, North Point Press 1991.

[52] Cavalli-Sforza, LL, Menozzi P, Piazza A. The History and Geography of Human Genes. Princeton, NJ, Princeton University Press 1994.

[53] Duster T. The molecular reinscription of race: unanticipated issues in biotechnology and forensic science. Patt Prejud 2007; 40: 427-441. https://doi.org/10.1080/00313220601020148

[54] Lorusso L, Bacchini F. A reconsideration of the role of self-identified races in epidemiology and biomedical research. Stud Hist Phil Biol Biomed Sci 2015; 52: 56-64. https://doi.org/10.1016/j.shpsc.2015.02.004

[55] Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: Its potential impact on understanding health disparities. Hum Genom 2015; 9: 1. https://doi.org/10.1186/s40246-014-0023-x

[56] Lea RA, Chambers GK. Pharmacogenetics in admixed populations in Polynesia. In Suarez-KurtzG. editor. Pharmacogenomics Austin, TX, Landes Bioscience; 2007; p: 164-189

[57] Lea RA, Roberts RL, Green MR, Kennedy MA, Chambers GK. Allele frequency differences of cytochrome P450 polymorphisms in a sample of New Zealand Maori. NZ M J 2008; 121: No 1272; ISSN 1175/8716.

[58] Risch N, Choudry S, Via M, Basu A, Sebro R, Eng C, *et al*. Ancestry-related assortative mating in Latino populations. Genom Biol 2009; 10: R132. https://doi.org/10.1186/gb-2009-10-11-r132

[59] Hammonds E, Braun L, Fausto-Sterling A, Fullwiley D, Nelson A, Quivers W, *et al*. Racial categories in medical practice: how useful are they? PLoS Med 2007; 4: e271. https://doi.org/10.1371/journal.pmed.0040271

Received on 10-10-2017 Accepted on 06-11-2017 Published on 21-12-2017

http://dx.doi.org/10.15379/2410-2806.2017.04.02.01