



The Cladistic Race Concept: A Defense

ROBIN O. ANDREASEN

Department of Philosophy, University of Delaware, 24 Kent Way, Newark, DE 19716, USA;
E-mail: robina@udel.edu

Received 31 January 2003; accepted 26 June 2003

Key words: Cladism, Classification, Natural kinds, Race, Realism, Social construction, Subspecies, Systematics

Abstract. Many contemporary race scholars reject the biological reality of race. Elsewhere I have argued that they have been too quick to do so. Part of the reason is that they have overlooked the possibility that races can be defined cladistically. Since the publication of the cladistic race concept, a number of questions and objections have been raised. My aim in this paper is to address these objections.

Introduction

Willi Hennig introduced the idea of cladistic classification in 1950 with the German language edition of his book *Phylogenetic Systematics*. Cladistics is a school of classification that defines taxa solely in terms of common ancestry. Evolution often takes the form of a branching pattern, which can be represented as a phylogenetic tree (see Figure 1). Provided that each object in the tree has a unique immediate ancestor, cladism defines its taxa as ‘monophyletic groups’ – i.e., groups composed of an ancestor and all of its descendants. Cladistic classification was originally introduced for defining higher taxa. The branching structure used to identify higher taxa is constructed out of sets of well-defined species. The smallest monophyletic units represent genera, the next largest units represent families, and so on up the taxonomic hierarchy.

Although cladistic classification was originally introduced for defining higher taxa, as I have argued elsewhere, it can be adapted for defining race (Andreasen 1998, 2000). According to the cladistic race concept, races are ancestor-descendant sequences of breeding populations that share a common origin. Rather than relying on sets of well-defined species, the branching structure used to define cladistic races would be constructed out of human ‘breeding populations’. The nodes in the tree would represent breeding populations and the branches would represent ‘the births of new breeding

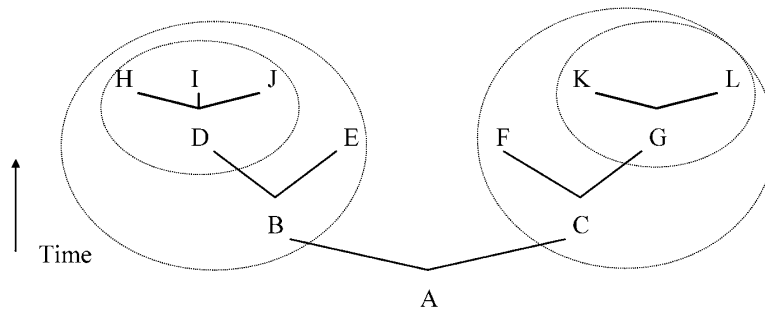


Figure 1. In this figure, the nodes (A, B, C, etc.) represent species and the branches represent speciation events. The circles represent a nested hierarchy of monophyletic groups.

populations'. A 'breeding population' is a set of local populations that are reproductively connected to one another and is reasonably reproductively isolated from other such sets. For example, a gaggle of geese living in Carousel Farms would constitute a local population. When there is interbreeding among two or more local populations, this totality makes up a breeding population. A breeding population is 'born' when a local subpopulation becomes separated from its parent population and is, for the most part, reproductively isolated from that population. Separation often results from the introduction of a geographic barrier, however, in the case of humans, it can also be due to socio-cultural factors (Kitcher 1999). Referring again to Figure 1, the tips of the tree would represent current breeding populations and the nested hierarchy of monophyletic groups would represent a nested hierarchy of races.

I have also argued that current research in human evolution illustrates that it is possible to apply the cladistic race concept to humans. For some time now, human evolutionists have been gathering data on the genetics of living populations. Their goal is to reconstruct a human family tree that accurately represents the patterns and processes of human evolution. One example comes from research done by Cavalli-Sforza and his colleagues (Cavalli-Sforza 1991; Cavalli-Sforza et al. 1994; Cavalli-Sforza et al. 1995).¹

The 'trunk' of this tree represents a racially undifferentiated population of modern humans evolving in Africa approximately 200,000 years ago. The first major split represents a local subpopulation of Africans migrating out of Africa and becoming reasonably reproductively isolated from its 'parent' population. Each subsequent branching point represents further migrations and reproductive isolation among breeding populations. This tree is an abridged version of a much larger human phylogeny. In the larger version,

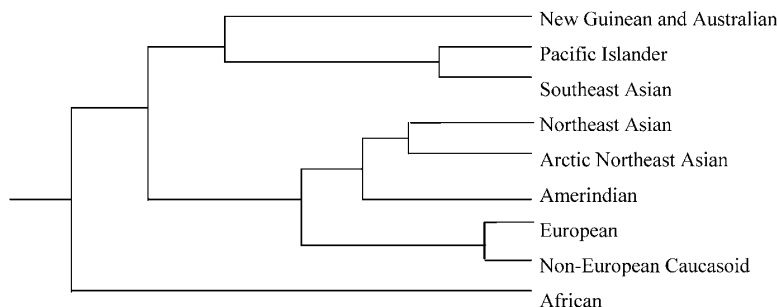


Figure 2.

the tips of the tree represent current breeding populations and the nested hierarchy of monophyletic groups represent a nested hierarchy of cladistic races. In this tree, each major branching point can be understood as the birth of a new race. For example, the group composed of European and non-European Caucasoids is a cladistic race, so is the branch containing only Africans, and so are many other groupings.

It is important to note that the cladistic race concept does not depend on the accuracy of Cavalli-Sforza's tree. Different research groups have proposed slightly different human phylogenies (Nei et al. 1982; Cann et al. 1987; Vigilant et al. 1991; Wilson et al. 1992; Nei et al. 1993), and it is not yet clear which one best represents human evolutionary history.² What is important for our purposes is that many human evolutionists agree that it is possible to accurately represent human evolution as a branching pattern.³ As long as this is possible, it is possible to define races cladistically.

Earlier I developed the cladistic view of race and defended it against some objections (Andreasen 1998, 2000). I will not repeat those arguments here. My aim in this paper is to address a number of questions that have been raised since the publication of my view (Zack 2002; Kaplan and Pigliucci 2003; Root, personal correspondence):^{4,5}

1. Some arguments against the biological reality of race suggest that there is too much human variation – both phenotypic and genetic – for races to be biologically real. Assuming that these empirical findings are accurate, how is it possible for cladistic races to exist?
2. Biological definitions of race are sometimes used to underwrite inferences from race membership to socially significant phenotypes such as behavioral and psychological traits. To what extent, if any, does the cladistic concept justify such inferences?
3. The cladistic race concept deviates from folk conceptions of race in a number of important ways. Why should we call cladistic races 'races'?

Cladistic races and the received view

Many contemporary race scholars maintain that races are biologically unreal.⁶ What they usually mean by this is that folk racial categories (such as blacks, whites, and Asians) are biologically uninteresting. They are explanatorily and predictively weak in the natural sciences – especially when it comes to explaining or predicting biologically based socially significant traits. Several arguments have been used in support of this view. Two of the most persuasive are what I call the ‘independent variation argument’ and the ‘genetic argument’.

Defenders of ‘the independent variation argument’ maintain that many of the traits traditionally used to define races – skin color, hair type, eye shape, blood type, propensity towards disease – vary independently (Livingstone 1964; Gould 1977; Diamond 1994). For example, a classification based on skin color may cross-classify one based on blood type. Both classifications may disagree with one that is based on propensity towards a certain disease. This problem is said to be compounded as more and more traits are added to the classification scheme. If one trait is used (e.g., skin color), it may be possible to provide an unambiguous classification scheme.⁷ If two traits are used (e.g., skin color and blood type), it may still be possible – but there will be some problem cases. As the number of traits increases, racial classification becomes increasingly difficult. Theorists conclude that because there is no non-arbitrary way to choose one classification scheme over another, we ought to abandon biological racial classification altogether.

Defenders of ‘the genetic argument’, on the other hand, deny the biological reality of race by arguing that folk racial categories (blacks, whites, Asians) are often defined in terms of skin color and that skin color groupings fail to represent significant genetic differences. Support for this view comes from current work in human genetics, which indicates that there is more genetic variation within than among the races (Lewontin 1982; Lewontin et al. 1984). Research indicates that approximately 85% of human genetic variation is among individuals in the same local population; approximately 8% is among populations within races; and approximately 7% is among races.

In order to better understand this argument, let us look more closely at how the above figures are calculated. Geneticists frequently study human genetic variation by examining ‘protein’ and ‘DNA polymorphisms’. ‘Polymorphic proteins’ are alternative forms of a single protein that are coded by alternative forms of a single gene (also called alleles). An example of such a protein is the hemoglobin protein, which comes in a sickle cell and a wild type (normal) form. Because DNA is a direct causal factor in the production of proteins (with mRNA as an intermediate), biologists can usually infer allelic

frequency differences from differences in the distribution of polymorphic proteins. 'DNA polymorphisms' are differences in base-pair sequences for the same stretch of DNA between two individuals. Unlike protein polymorphisms, DNA polymorphisms provide a direct measure of allelic frequency differences among individuals or populations.

Over the past 30 years, geneticists have gathered extensive data on genetic and protein polymorphisms in contemporary populations from all over the world. Using these data, they have estimated that humans are 99.9% genetically similar. This means that the total genetic variation among humans (T) is 0.1%. Researchers estimated T by estimating the probability that a randomly chosen allele taken from a single individual is a different allele than that taken from another randomly chosen individual within the species as a whole. They then asked what proportion of T is accounted for by variation among races (A_R) as compared to the amount that is accounted for by variation among local populations within a race (A_P) and variation among individuals within a population (W_P). They estimated the variation among races by estimating the probability that a randomly chosen allele taken from an individual is a different allele than that taken from another randomly chosen individual within the same race (W_R). The proportion of variation that is due to racial difference (A_R) is calculated by taking the difference between T and W_R . In a manner similar to the estimation of T and W_R , researchers estimate the amount of variation among individuals within the same population (W_P). They then subtract this measure from W_R to get the proportion of total variation that is due to difference among populations (A_P). As already mentioned, these studies revealed that the variation among major racial groups (7%) is slightly less than the variation among local populations within the same race (8%). Moreover, the variation among individuals within a population is (85%) substantially greater than the variation among populations or races. Many races theorists have concluded that biology lends no justification to common sense (CS) racial groupings.

Now let us turn to the question that I raised earlier: Assuming that the empirical findings cited in the above arguments are accurate, how is it possible for cladistic races to exist? The independent variation argument presents the cladist with the following related challenges. First, she must explain how it is possible for cladistic races to be biologically real when the traits commonly used to define races vary independently. Second, she must explain why her choice of shared history as the basis of racial classification is non-arbitrary (Zack 2002). The genetic variation argument presents the cladist with a third challenge – namely, to explain how it is possible to designate cladistic races when there is more genetic variation within than among races. I will take each challenge in order.

How is it possible for cladistic races to be biologically real when the traits commonly used to define race vary independently? The answer to this question is simple. The independent variation of so-called racial traits – skin color, hair type, bone structure, and so forth – is irrelevant to the existence of cladistic races. The reason is that the cladistic concept defines race in terms of a *single* relational property (genealogy) rather than in terms of shared intrinsic properties (phenotypic and genetic similarity). Cladistic races are ancestor-descendent sequences of breeding populations that share a common origin. It is a consequence of this definition that two populations, A and B, can be more closely related to one another than either is to a third, C, even if A and C are more morphologically similar to one another than either is to B.⁸ In fact, research indicates that overt morphology is sometimes a poor indicator of human racial ancestry (Cavalli-Sforza et al. 1995). Some of the traits used to identify CS racial groupings – such as skin color and various aspects of bone structure – are a product of recent and rapid evolution by natural selection. Referring again to Figure 2, Northeast Asians and Amerindians are more closely related to one another than either is to Southeast Asians. However, Northeast and Southeast Asians appear to be more morphologically similar to one another than either is to Amerindians. To understand how this is possible, let us recall that the branching points among these groups is determined by their degrees of reproductive isolation. Rapid evolution within one of two closely related populations – such as Northeast Asians and Amerindians – could cause them to appear morphologically very different. Likewise convergent evolution among two distantly related populations – Northeast and Southeast Asians – could cause them to appear quite similar. Research indicates that this is exactly what has occurred in the case of some human racial traits (Cavalli-Sforza et al. 1995).

But why should we use genealogy rather than similarity for defining races – especially when CS conceptions of race often appeal to overt morphology (Zack 2002)? I have several things to say in response to this question. First, as I have argued elsewhere, we should not reject the biological reality of race merely because the cladistic concept deviates from a certain CS conception of race (Andreasen 2000). Second, defining races in terms of genealogy is theoretically interesting. One of the primary aims of evolutionary biology is to reconstruct evolutionary history by mapping the genealogical relations among populations, subspecies, species, and so forth. Another aim of evolutionary biology is to understand the process of speciation and this too requires a mapping of the genealogical relations among populations, subspecies, species, etc. Third, genealogy is the relevant criterion for defining species and higher taxa (Hull 1976, 1978; Ridley 1986, 1993). Species are often defined as sets of interbreeding natural populations that are reproductively

isolated from other such sets (Mayr 1969). Although competing definitions exist, most require that species be defined genealogically (Ereshefsky 1992).⁹ Higher taxa are also defined genealogically. They are frequently defined as monophyletic groups of species that share a common ancestor (Ridely 1986, 1993). Thus, it makes sense to define biological races in terms of genealogy as well. Furthermore, the debate over the biological reality, or lack thereof, of race often takes place within systematic biology. Defenders of biological realism usually assume that races are subspecies – i.e., taxonomic subdivisions of a species – and defend an application of the subspecies concept to humans. Opponents of biological realism usually reject the biological reality of race either by arguing that the subspecies concept is superfluous or by revealing problems with its application to humans.

Turning our attention to the genetic challenge, how is it possible to designate cladistic races when there is more genetic variation within major racial groups than there is among them? I will not disagree with the empirical data used to support this argument. Yet, I maintain that these data are consistent with the biological reality of cladistic races. In Andreasen (1998), I argued that a consequence of the cladistic view is that races are dynamic categories – races once existed, but may be on their way out – and that a possible explanation of these data is that races have become more genetically similar as degrees of reproductive isolation have lessened. The point that I want to make here is that even if this is not the case, the genetic argument does not undermine the existence of cladistic races. Cladistic classification merely requires that evolution take the form of a branching process; it does not require that there be significant phenotypic or genetic differences among clades (Edwards 2003). Racial evolution will take this form when a species splits into two or more breeding populations that experience different evolutionary forces under a significant degree of reproductive isolation. Current work in human evolution indicates that this condition was met in the past (contra Kaplan and Pigliucci 2003 and Zack 2002). The genetic data used to reconstruct human phylogenies indicate that Old World human populations had low levels of genetic contact for a substantial portion of time (Nei et al. 1993; Cavalli-Sforza et al. 1994; Cavalli-Sforza et al. 1995; Jorde et al. 1997; Perez-Lezaun et al. 1997).¹⁰ In addition, it is possible for there to be significant reproductive isolation without significant phenotypic or genetic change. This will occur when evolution within breeding populations is relatively slow, or when the evolution of distinct breeding populations is moving in the same direction. In fact we should not expect huge phenotypic and genetic differences among cladistic races, in part, because human evolutionary history is reasonably short when measured on the scale of evolutionary time.

For some, there might be a worry lurking in the background. I have just made the ontological point that cladistic races can exist even when there is more variation within than among races. The reader might agree with this claim, but might raise the following epistemic question: How is it possible to infer the existence of cladistic races when there is more variation within than among races? One could argue, for example, that even though the cladistic account does not *define* races by appeal to similarity, systematic biologists often use similarity as *evidence* of shared ancestry. Thus, it might not be possible to reconstruct human evolutionary history.

I would like to begin by noting that the genetic argument was never considered as a point about phylogenetic inference. It is supposed to show that CS racial categories – defined in terms of skin color and other morphological features – are poor predictors of genetic similarities and differences among individuals. That is, if I tell you that two individuals are members of the same race, I have very little information about what they are like at the genetic level. The cladistic race concept, however, is a population concept and, thus, does not need to answer this standard about prediction of individual traits. Still, it is worth considering whether the pattern of more variation within than among races will undermine phylogenetic reconstruction.

Let us distinguish two basic approaches to phylogenetic inference – pheneticism and cladism. According to the ‘phenetic approach’, similarity is used as evidence of common ancestry. If two taxa, A and B, are more similar to one another than either is to a third, C, this is evidence that A and B are more closely related to one another than either is to C. According to the ‘cladistic approach’, taxa are arranged in a phylogeny according to the principle of parsimony. The ‘principle of parsimony’ is the idea that the best estimate of the true phylogeny for a given set of taxa is the one that requires the smallest number of evolutionary changes. In addition, cladists treat only some types of similarity as evidence of common ancestry. Any character shared between taxa can be sorted into one of two types – homologies or analogies. An ‘homology’ is a character that is shared by two or more taxa because each inherited it unmodified from a common ancestor. An ‘analogy’ is a character that is shared by two or more taxa because it evolved independently in each lineage. Homologies, in turn, can be divided into those that are ancestral and those that are derived. An ‘ancestral homology’ (or symplesiomorphy) is a homology that is present in the most recent common ancestor of a given set of taxa, but is also shared with other more distantly related taxa. A ‘derived homology’ (or synapomorphy) is a homology that first evolved in the common ancestor of two or more taxa and is not shared with other more distantly related taxa.

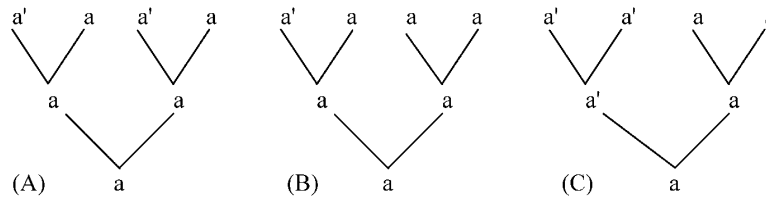


Figure 3. ¹¹ (A) a' is an analogy. (B) a is an ancestral homology. (C) a' is derived homology.

Of these different types of similarity, cladists use only shared derived homologies as evidence of common ancestry. Homologies are preferred to analogies because cladists assume that evolutionary change is improbable. Ancestral homologies are preferred to derived homologies because the former often make a difference to how parsimonious competing hypotheses are, whereas the latter do not.

It is possible to use molecular data to infer the existence of cladistic races using both phenetic and cladistic methods – and this is so, even given the assumption that there is more variation within than among races. For example, according to the phenetic approach, phylogenetic reconstruction can be based on between group variation alone. One tool for reconstructing phylogenies is a quantitative measure called ‘genetic distance’. This is a measure of the difference in gene frequencies between two breeding populations. All other things being equal, the smaller the genetic distance between two populations, the closer their ancestral relation.¹² The reasoning at work here is that all humans share a single common ancestor. As local subpopulations migrated from their original location and formed new populations, mutations occurred and gene frequency differences accumulated. Thus, if there were no genetic differences among human breeding populations, it would not be possible to reconstruct a human phylogeny. However, this is not the case. Genetic data merely reveal that the differences among races are less significant than the differences within a race.

Readers may be surprised by my discussion of pheneticism in the context of cladistic classification. It is important to note, however, that the terms ‘cladism’ and ‘pheneticism’ serve double duty (Felsenstein 1984). They can be used to refer to theories of classification or to different approaches to phylogenetic reconstruction. Moreover, the question of how to reconstruct a phylogeny is logically distinct from the question of how one ought to construct a classification scheme. One should not confuse the two, since doing so can lead to a number of misunderstandings about the cladistic view of race. See Zack (2002) for this error. For example, one might use phenetic methods for constructing a human phylogeny, but use cladistic

methods for constructing a racial classification scheme. It does not follow from this, contrary to Zack's suggestion, that cladistic races are defined solely in terms of genetic clusters. I would also like to add that a cladist need not reject distance statistics entirely. They are an important method in phylogenetic inference, since they can supply a large amount of data and there are circumstances in which they are almost as reliable as parsimony (Ridley 1993).

Cladists have developed a variety of statistical algorithms that assume parsimony for inferring phylogenies using molecular data. I will not go into the details here, but will refer the reader to Cavalli-Sforza et al. (1967) and Felsenstein (1981, 1988) for details. What is important for our purposes is that these algorithms have been successfully applied to human breeding populations (Cann et al. 1987; Vigilant et al. 1991; Wilson et al. 1992), and this illustrates that the distribution of more genetic variation within than among races does not undermine phylogenetic reconstruction using cladistic methods. Moreover, it is worth noting that although many genetic polymorphisms show little difference from one population to the next, there are also a significant number of genes that differ significantly among populations. These data will be especially useful for diagnosing branching.

Cladistic races and belief in biologically based racial superiority

So far, I have argued that two of the standard arguments against the biological reality of race do not apply to the cladistic race concept. Yet, there is a further issue that ought to be addressed. When contemporary race theorists reject the biological reality of race, they are in part denying that knowledge of a person's race allows biologically justified inferences from overt physical features to psychological and behavioral traits. Race scholars are concerned about such inferences because, if biologically justified, they could be used to underwrite claims about biologically based racial superiority. In this section I will discuss the extent, if any, to which the cladistic concept justifies generalizations from one's race to the psychological and behavioral traits that she might possess.

Although it is not possible to deny that the cladistic concept could be used to make inferences from race membership to socially significant biologically based phenotypes, people need to be careful about over-interpreting the data. Human phylogenies are hypotheses representing the history of human populations. A well-confirmed tree accurately represents patterns of migration and degrees of reproductive isolation among human breeding populations. What this has to do with racial similarities and differences (behavioral, psychological, or otherwise) is an open empirical question. As we

have already seen, human phylogenies are constructed using protein and DNA polymorphisms as well as mitochondrial DNA. Psychological and behavioral traits are not represented. Moreover, there is no current evidence to indicate that the DNA used to construct human phylogenies code for psychological or behavioral traits. Part of the reason is that geneticists often rely on selectively neutral DNA for phylogenetic reconstruction, such as 'junk DNA' (i.e., DNA that serves no apparent function) and mitochondrial DNA, neither of which code for psychological or behavioral traits. Even when genes (i.e., functional DNA) are used in phylogenetic reconstruction, often little is known about the roles, if any, that they play in the development of complex psychological and behavioral traits. Not only is there no current evidence indicating that the data used to reconstruct human phylogenies code for socially significant traits, there is no current evidence indicating that they correlate with socially significant traits. Moreover, even if ancestry were a good guide to socially significant biological differences among races, there is no reason to suppose that these differences would be large or that the patterns within the races would be strong – especially if there is more variability within major races than there is among them. Finally, many psychological and behavioral traits are environmentally plastic. This is especially likely for the types of traits that are of interest to the race theorist – such as IQ, propensity towards violence, sexual promiscuity, and so forth (Lewontin 1984; Kitcher 1985; Block 1995). In such cases, there is reason to suppose that existing differences (if any) could be reversed by changes in environment (Block 1995).

Zack (2002) takes it to be a problem for the cladistic view that human phylogenies are often reconstructed using nonfunctional DNA. Part of the reason is that such phylogenies are based on genetic differences that have no known relation to traits that are commonly used to identify races – such as skin color and other 'racial' phenotypes. I will address this objection in the next section when I discuss the relationship between cladistic race taxa and CS racial categories.

Others might have a different worry. The above response raises the possibility that the cladistic concept may be explanatorily and predictively weak, at least when it comes to explaining (or predicting) socially significant biologically based differences among races, and therefore may seem theoretically uninteresting. Yet, it is important to note that this is not how taxonomic categories are expected to earn their living. As already mentioned, cladistic races are useful for explaining the history of human populations, their patterns of migration, and their degrees of reproductive isolation. Not only is this of general interest, it is also of theoretical interest to the evolutionary biologist who aims to understand the genealogical relations among all living things.

Why should we call cladistic races ‘races’?

Some critics argue that cladistic ‘races’ are not really races because the cladistic concept deviates too far from conventional notions of ‘race’ (Zack 2002; Root, personal correspondence). I have already argued that the cladistic concept corresponds to CS in a number of important ways (Andreasen 2000). In what follows, I will offer several new responses to this objection.

Let me start by noting that the cladistic concept largely agrees with accepted biological and anthropological uses of the term. In the past, anthropologists and biologists relied on the ‘geographical subspecies concept’ for defining race. According to this concept, races are geographically circumscribed phenotypically and genetically distinct populations. For quite some time, this was the received biological race concept. Yet, it was eventually revised for reasons specified in Wilson et al. (1953). According to the revised definition, races are genetically differentiated geographically circumscribed breeding populations (Shaffer et al. 1996; Legge et al. 1996; Kitcher 1999; Risch et al. 2002). The main difference between this definition and the former is that the geographical concept allows that phenotypic and genetic differentiation *alone* is sufficient for defining race. According to the revised definition, however, races are defined in terms of genetic differentiation and historical continuity. This latter point is especially important for our purposes because it illustrates that I am not alone in defining races as distinct evolutionary lineages. The cladistic race concept can be thought of as a refinement of the revised geographical concept. Cladistic races are geographically circumscribed breeding populations that share a common origin. Unlike the revised geographical concept, which relies in part on genetic differentiation for defining race, the cladistic concept uses genetic differentiation solely as evidence for race membership. Like the revised geographical concept, however, historical continuity is an essential element in defining cladistic races.

I have just argued that the cladistic concept corresponds with some accepted scientific and anthropological uses of the term. I will now address the worry that it deviates too far from common sense. Many race theorists take skin color and other morphological features to be an essential part of any good definition of race (Zack 2002). This is said to pose a problem for the cladistic view because it defines races by appeal to genealogy alone. I argue, however, that this objection is misleading. It gives the impression that phenotypic and genetic similarities are irrelevant to the cladistic race concept. Yet, this is not the case. Phenotypic and genetic similarities may not define cladistic races, but (as already mentioned) they are used as evidence for race membership. Skin color and other racial phenotypes can be thought of as giving a first approximation of ancestral relations and there is often

considerable agreement between classifications based solely on similarity and those based on genealogy alone. Nonetheless when the two disagree, the cladistic concept favors the classification scheme that is based on genealogy alone.

This brings me to a second worry that is often expressed about the cladistic view: Cladistic races fail to agree extensionally with everyday conceptions of race. In everyday discourse, the term is often used to refer to blacks, whites and Asians. The cladistic concept, however, identifies a nested hierarchy of races – the number of which exceeds the three major races mentioned above. Moreover, cladistic races may cross-classify standard racial groupings because the category ‘Asian’ may not be a cladistic race. The reason is that Northeast Asians appear to be more closely related to Amerindians than they are to Southeast Asians and Pacific Islanders.

One problem with this objection is that it assumes that common sense is unambiguous about the extension of the term. Yet, as many race scholars have noted, there has always been widespread disagreement over how many racial categories there are and who belongs to what category (Omi et al. 1994; Wright 1994; Marks 1995). At times, Americans have treated different national groups – such as Mexicans, Filipinos, and the Irish – and certain ethnic groups, such as Hispanics and Jews, as races. Furthermore, the racial categories identified in the US. census have varied widely from decade to decade. In 1860, three racial categories were identified – viz., white, black, mulatto. In 1870 the categories American Indian and Chinese were added to this list. By 1890 eight racial categories were identified – White, Black, Mulatto, Chinese, American Indian, Quadroon, Octoroon, Japanese. The 1970 census asked people to sort themselves into one of five categories: American Indian or Alaskan Native, Asian or Pacific Islander, Black, White, or other. In 1980, Asian Indians were moved from the ‘white’ race to the ‘Asian and Pacific Islander’ group. In 2000, Native Hawaiians and other Pacific Islanders were separated from the ‘Asian’ category. Moreover, census takers were given the option of marking one or more race to indicate their racial identity. These examples illustrate that the extension of ‘race’ in everyday discourse is far less determinate than the above criticism allows. Once we acknowledge this, the criticism loses much of its bite.

At the most general level, the cladistic concept identifies five major races – namely, Africans, Caucasians (European and Non-European), NE Asians, SE Asians and Pacific Islanders (including New Guineans and Australians), and Native Americans. It also identifies a number of subraces within these racial groups. When we compare these categories to those identified by the 2000 census, we see quite a bit of agreement. The US census identifies similar

major racial groups. It also breaks down the major racial groups into subraces in a manner similar to those recognized by the cladistic view.

The final issue that needs to be addressed is the potential cross-classification between cladistic races and folk racial categories. Before addressing this objection, however, it is important to note that the reconstruction of human evolutionary history is still underway. As already mentioned, several different research groups have proposed human phylogenies, each of which represent slightly different racial classification schemes. The phylogeny proposed by Cavalli-Sforza, as well as those proposed by his competitors, should be thought of as works in progress – and the claim that ‘Asian’ may not be a cladistic race should be met with caution. Yet, I continue to maintain that potential cross-classification with folk racial categories is not a problem for the cladistic view. When we consider changes to Directive 15 and the US census, we see that folk conceptions of race cross-classify one another as well. Directive 15 was developed by the Office of Management and Budget and is used to help standardize the collection and reporting of racial and ethnic statistics among federal agencies, researchers, and businesses. In 1977, the directive described four races – i.e., American Indian or Alaskan Native, Asian or Pacific Islander, Black, and White – but was revised in 1997 to include five: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. The 2000 census was also revised and now matches Directive 15. In this example, Pacific Islanders were once considered part of the Asian race, but after 1997, they were eliminated from this group and are now considered to be their own race. When we consider other changes to the US census, we can see a second example of cross-classification among differing folk conceptions of race. Recall that Asian Indians were considered ‘white’ on the 1970 census, but were considered ‘Asian’ on the 1980 census.

The above arguments illustrate that there is no such thing as *the* extension of the term ‘race’ in everyday discourse. Moreover, although similarity is not used to define cladistic races, it is not irrelevant to the cladistic concept. Thus, it is hard to make the case that the cladistic concept deviates too far from CS. That being said, I continue to hold that the term ‘race’ is ambiguous (Andreasen 1998, 2000). As already noted, when contemporary race scholars deny the biological reality of race, they are arguing that the division of humans into three major races (blacks, whites, and Asians) is explanatorily and predictively weak – at least when it comes to explaining or predicting biologically based socially significant traits. I do not disagree with this claim; races in this sense are social constructs. Nonetheless, I maintain that the cladistic race concept provides a biologically objective definition of race.

Conclusion

The cladistic race concept helps to clarify a number of issues concerning the relationship between biology and race. First, the genetic and independent variation arguments have been widely accepted as showing that races are biologically unreal. Yet, as I have argued above, they do not work against the cladistic view. Second, many scholars have expressed concern about belief in the biological reality of race because it has been, and sometimes continues to be, used in an effort to justify racism. It is important to recognize, however, that the connection between biological realism and racism is a contingent one that depends on empirical details that are currently unknown. The cladistic concept helps to underscore this fact. Finally, some race theorists object to the cladistic race concept on the grounds that it deviates from some folk conceptions of race. Yet, because common sense beliefs about race have been and continue to be in flux, such deviations are not serious enough to undermine existence of cladistic races.

Acknowledgements

I would like to thank Elliott Sober, Christopher Stephens, as well as the editor and the anonymous referees of this journal for helpful comments on an earlier draft of this paper. Part of the research for this paper was carried out with the support of the University of Delaware General University Research fund.

Notes

¹ This branching diagram differs from the phylogenetic tree depicted in Figure 1 in that it depicts branching sequence only and does not represent the relations among ancestral populations.

² See Andreasen (1998) for a discussion of some of the controversies surrounding Cavalli-Sforza's tree.

³ For an opposing view, see Tempelton (1999).

⁴ Many of the objections raised by Zack (2002) have already been addressed in Andreasen (1998, 2000). My focus in this paper will be on objections that have not been addressed or ones that require further clarification.

⁵ There are important differences between the views offered by Kaplan and Pigliucci, Zack, and Root. Kaplan and Pigliucci (2003), for example, do not deny the biological reality race. Rather they argue that races are ecotypes. Zack (1993, 2002) and Root (2000), on the other hand, reject the biological reality of race, but offer different positive accounts. Zack argues for the view that race is an empty term, on a par with terms like 'phlogiston'. Root argues for the social reality of race.

⁶ Some scholars use double quotations around the term 'race' as a way of indicating that races

do not exist. For the most part, I will not follow this convention here. Part of the reason is that even if races are biologically unreal, it does not follow that races do not exist. For example, many race scholars deny the biological reality of race but maintain that races are socially real (Outlaw 1990; Omi et al. 1994; Root 2000; Mallon 2003).

⁷ It is important to note that many so-called racial traits – such as skin color, hair type, and bone structure – are clinal in nature; they vary gradually across geographical regions. Some race theorists have used this to argue that racial classification is beset with line-drawing problems and, therefore, is arbitrary (Gould 1977). This objection, however, is a strawman. Vague boundaries are not a problem for distinguishing between kinds, provided that there are clear differences between the extremes. For example, there are clear differences between being bald and having a full head of hair. These properties are real, even though there is no precise line of demarcation between them.

⁸ For example, lizards and crocodiles are more similar to one another than either is to birds, yet birds are more closely related to crocodiles than either is to lizards. Many other examples of this sort are present in the biological literature (Ridley 1986).

⁹ The main exception is the phenetic species concept, which defines a species in terms of the overall similarity of its members. However, most of today's biologists reject the phenetic view for reasons specified in Ridley (1986, 1993).

¹⁰ For further support for this claim see Andreasen (1998: 199–225).

¹¹ This figure is from Ridley (1993).

¹² Of course, things are not always equal. Many things – convergence, rapid genetic drift in small populations, and interbreeding among previously isolated populations – can disturb the correlation between time and genetic distance. Nonetheless, geneticists have ways of minimizing these potential biases.

References

- Andreasen, R.O.: 1998, 'A New Perspective on the Race Debate', *British Journal for the Philosophy of Science* **49**, 199–225.
- Andreasen, R.O.: 2000, 'Race: Biological Reality or Social Construct?', *Philosophy of Science* **67** (Proceedings), 653–666.
- Block, N.: 1995, 'How Heritability Misleads About Race', *Cognition* **56**, 99–126.
- Cann, R.L., Stoneking, M. and Wilson, A.C.: 1987, 'Mitochondrial DNA and Human Evolution', *Nature* **325**, 31–36.
- Cavalli-Sforza, L.L.: 1991, 'Genes, People, Language', *Scientific American* **265**, 104–110.
- Cavalli-Sforza, L.L. and Cavalli-Sforza, F.: 1995, *The Great Human Diasporas: The History of Diversity and Evolution*, Addison-Wesley, New York.
- Cavalli-Sforza, L.L. and Edwards, A.W.F.: 1967, 'Phylogenetic Analysis: Models and Estimation Procedures', *Evolution* **21**, 550–570.
- Cavalli-Sforza, L.L., Menozzi, P. and Piazza, A.: 1994, *The History and Geography of Human Genes*, Princeton University Press, Princeton.
- Diamond, J.: 1994, 'Race Without Color', *Discover* (November), 83–89.
- Edwards, A.W.F.: 2003, 'Human Genetic Diversity: Lewontins Fallacy', *Bio Essays* **25**, 798–801.
- Ereshefsky, M.: 1992, 'Eliminative Pluralism', *Philosophy of Science* **59**, 671–690.
- Felsenstein, J.: 1981, 'Evolutionary Trees from Gene Frequencies and Quantitative Characters: Finding Maximum Likelihood Estimates', *Evolution* **35**(6), 1229–1242.

- Felsenstein, J.: 1984, 'The Statistical Approach to Inferring Evolutionary Trees and What it Tells Us about Parsimony and Compatibility', in T. Duncan and T. Stuessy (eds.), *Cladistics: Perspectives on the Reconstruction of Evolutionary History*, Columbia University Press, New York, pp. 169–191.
- Felsenstein, J.: 1988, 'Phylogenies from Molecular Sequences: Inference and Reliability', *Annual Review of Genetics* **22**, 521–565.
- Gould, S.J.: 1977, *Ever Since Darwin: Reflections in Natural History*, W. W. Norton, New York.
- Hull, D.: 1976, 'Are Species Really Individuals?', *Systematic Zoology* **25**, 174–179.
- Hull, D.: 1978, 'A Matter of Individuality', *Philosophy of Science* **45**, 335–360.
- Jorde L.B.: 1997, 'Microsatellite Diversity and the Demographic History of Modern Humans', *Proceedings of the National Academy of Science* **94**, 3100–3103.
- Kaplan, J. and Pigliucci, M.: 2003, 'On the Concept of Biological Race and Its Applicability to Humans', *Philosophy of Science*, Supplement, Vol. 70(3), Proceedings of the 2002 Biennial Meeting of the Philosophy of Science Association, Part I: Contributed Papers.
- Kitcher, P.: 1985, *Vaulting Ambition: Sociobiology and the Quest for Human Nature*, MIT Press, Cambridge, MA.
- Kitcher, P.: 1999, 'Race, Ethnicity, Biology, Culture', in L. Harris (ed.), *Racism*, Humanity Books, Amherst, NY.
- Legge, J.T.: 1996, 'Genetic Criteria for Establishing Evolutionarily Significant Units in Cryans Buckmoth', *Conservation Biology* **10**, 85–98.
- Lewontin, R.C.: 1982, *Human Diversity*, Freeman, San Francisco, CA.
- Lewontin, R.C. and Rose, S.: 1984, 'I. Q.: The Rank Ordering of the World', in *Not in Our Genes: Biology, Ideology, and Human Nature*, Pantheon Books, New York, pp. 83–129.
- Livingstone, F.B.: 1964, 'On the Nonexistence of Human Races', in A. Montagu (ed.), *The Concept of Race*, Collier Books, New York, pp. 46–60.
- Mallon, R.: 2003, 'Social Construction, Social Roles, and Stability', in F. Schmitt (ed.), *Socializing Metaphysics*, Rowman and Littlefield.
- Marks, J.: 1995, *Human Biodiversity: Genes, Race, and History*, Aldine De Gruyter, Hawthorne, NY.
- Mayr, E.: 1969, *Principles of Systematic Zoology*, McGraw Hill, New York.
- Nei, M. and Roychoudhury, A.K.: 1982, 'Genetic Relationships and Evolution of Human Races', *Evolutionary Biology* **14**, 1–59.
- Nei, M. and Roychoudhury, A.K.: 1993, 'Evolutionary Relationships of Human Populations on a Global Scale', *Molecular Biology and Evolution* **10**(5), 927–943.
- Omi, M. and Winant, H.: 1994, *Racial Formation in the United States: From the 1960s to the 1990s*, 2nd. ed., Routledge, New York.
- Outlaw, L.: 1990, 'Toward a Critical Theory of "Race"', in D.T. Goldberg (ed.), *Anatomy of Racism*, University of Minnesota Press, Minneapolis, MN, pp. 58–82.
- Perez-Lezaun, A.: 1997, 'Microsatellite Variation with the Differentiation of Modern Humans', *Human Genetics* **99**, 1–7.
- Ridley, M.: 1986, *Evolution and Classification: The Reformation of Cladism*, Longman Scientific and Technical, London.
- Ridley, M.: 1993, *Evolution*, Blackwell, Boston, MA.
- Risch, N.: 2002, 'Categorization of Humans in Biomedical Research: Genes, Race, and Disease', *Genome Biology* **3**(7), 1–12.
- Root, M.: 2000, 'How We Divide the World', *Philosophy of Science* **67** (Proceedings), S628–S639.

- Shaffer, H.B., and McKnight, M.L.: 1996, 'The Polytypic Species Revisited', *Evolution* **50**, 417–433.
- Templeton, A.: 1999, 'Human Races: A Genetic and Evolutionary Perspective', *American Anthropologist* **100**, 632–650.
- Vigilant, L. and Stoneking, M.: 1991, 'African Populations and the Evolution of Human Mitochondrial DNA', *Science* **253**, 1503–1507.
- Wilson, A.C. and Cann, R.L.: 1992, 'The Recent African Genesis of Humans', *Scientific American* **266**, 68–73.
- Wilson, E.O. and Brown, W.L.: 1953, 'The Subspecies Concept and Its Taxonomic Applications', *Systematic Zoology* **2**, 97–111.
- Wright, L.: 1994, 'One Drop of Blood', *The New Yorker* **25** (July), 46–55.
- Zack, N.: 1993, *Race and Mixed Race*, Temple University Press, Philadelphia, PA.
- Zack, N.: 2002, *Philosophy of Science and Race*, Routledge, New York.