The Race FAQ

John Goodrum

Do biological races exist within the human species? If scientific terms are to be used consistently, this question can only be answered in the broader context of non-human taxonomy. The intent of this paper is to investigate what constitutes a race (or subspecies) in other species, and to answer some questions concerning whether the traditional human races might qualify.

(back to main page)

Q: What is the definition of 'race' or 'subspecies?'

The terms 'race' and 'subspecies' are most often used synonymously [1,2] although the former is normally used when talking about human populations. When a distinction is made, 'race' generally implies a lower level of differentiation, but because this term is not commonly used in the recent non-human literature, 'race' and 'subspecies' are used interchangeably throughout this paper.

Much of the debate over the existence of human races stems from how one chooses to define 'race' (or 'subspecies'). No realistic definition can avoid using qualitative terms, yet these invariably invite disagreement in their application: "a group of individuals in a species showing closer genetic relationships within the group than to members of other such groups"[3]; "essentially discontinuous sets of individuals"[4]; "conspecific populations that differ from each other morphologically"[5]; "genetically non-discrete (confluent) populational entities"[6]; "geographically circumscribed, genetically differentiated populations"[7]; or groups identified "by the usual criterion that most individuals of such populations can be allocated correctly by inspection."[8] Compounding the confusion, still others employ the term 'race' in a way more akin to 'species' than to 'subspecies.'[9]

In response to questionable interpretations of the U.S. Endangered Species Act, and to help ensure the evolutionary significance of populations deemed 'subspecies,' a set of criteria was outlined in the early 1990s by John C. Avise, R. Martin Ball, Jr.[10], Stephen J. O'Brien and Ernst Mayr [11] which is as follows: "members of a subspecies would share a unique, geographic locale, a set of phylogenetically concordant phenotypic characters, and a unique natural history relative to other subdivisions of the species. Although subspecies are not reproductively isolated, they will normally be allopatric and exhibit recognizable phylogenetic partitioning." Furthermore, "evidence for phylogenetic distinction must normally come from the concordant distributions of multiple, independent genetically based traits."[12] This is known as

the phylogeographic subspecies definition, and a review of recent conservation literature will show that these principles have gained wide acceptance.

A number of studies have employed this subspecies definition, and these can be helpful in inferring how the definition is applied in practice. A good example is a paper entitled "Phylogeographic subspecies recognition in leopards (Panthera pardus): Molecular Genetic Variation,"[13] co-authored by Stephen J. O'Brien (one of the definition's co-authors). From the ranges of the revised leopard subspecies (Fig. 1) we can infer that a 'unique geographic locale' does not require that a range be an island, or share no environmental characteristics with another. Rather, it merely requires a subspecies to have a geographical association as opposed to a subset of individuals sharing a trait but drawn from different geographical populations. Conversely, two subspecies will not remain distinct if they occupy the same locale over evolutionary time. Hypothetical human races have been proposed in which members would share a single trait (e.g., lactose tolerance or fingerprint pattern)[14] but not a common geographic locale. These 'races,' therefore, would not be valid under the phylogeographic definition.

Whether a population has had a unique natural history can be inferred from its degree of differentiation with respect to other such populations. The arbitrary division of an interbreeding, genetically unstructured group will result in subgroups that are genetically indistinguishable, whereas populations that evolve more or less independently for some length of time will accumulate genetic differences (divergent gene frequencies, private alleles, etc.) such that they "exhibit recognizable phylogenetic partitioning."

A set of "phylogenetically concordant phenotypic characters" is taken to mean several morphological, behavioral or other expressed traits that tend to co-vary within, but differ among, putative subspecies. This indicates that members of the group have evolved together relative to other groups, and may reflect shared demography, local adaptation, sexual selection or other evolutionary effects.

The need for "concordant distributions of multiple, independent genetically based traits" requires us to recognize that too much inference from a single trait or single genetic locus is unwarranted. For instance, rather than indicating recent co-ancestry, a trait shared by two populations might have evolved independently in response to some environmental variable, while the potential idiosyncrasies of any single gene can limit its reliability to paint an accurate phylogenetic picture. Most population genetics theory relies on loci that have evolved neutrally (i.e., in the absence of natural selection) so a non-neutral locus may give misleading results. The best way to avoid this potential source of error is to examine a large number of independently-evolving loci.

Q: How genetically diverse are humans?

It's become a popular view that the human species is extraordinarily homogeneous genetically when compared to most other species.[15] This notion argues against the existence of human

races, because very little genetic variation within the entire species means there cannot be much variation between major human populations. Before examining this further, we should first inquire about what is meant by 'genetic diversity.'

Because little can be learned from a locus that is the same in every individual, the study of phylogenetics depends on polymorphic loci. Over the past few decades, methods have been developed that allow different kinds of these polymorphic 'markers' to be assayed in individuals. Prior to the 1990s, genetic diversity was usually inferred from classical (non-DNA) polymorphisms, such as blood groups, serum proteins, allozymes and immunoglobins. Later, restriction enzymes were employed to produce a useful class of marker at the DNA level, restriction fragment length polymorphisms (RFLPs). Other loci such as mitochondrial DNA (mtDNA), Alu insertions, minisatellites, single nucleotide polymorphisms (SNPs) and microsatellites (STRPs - short tandem repeat polymorphisms) have also been utilized for population genetic studies. Due to their high polymorphism, rapid mutation rate and random distribution throughout the genome, microsatellites are probably the most important class of marker in use today.[16] Highly variable loci are an advantage in phylogenetics because they can provide the finer resolution necessary for distinguishing closely related populations (such as subspecies).

The majority of population genetic studies over the past decade have investigated the various regions of mitochondrial DNA, a molecule that resides in the cytoplasm outside a cell's nucleus. mtDNA contains 37 genes and is comprised of 16,569 base pairs in humans. Because it is haploid and maternally inherited, mtDNA has an effective population size about one-quarter that of the autosomes (the non-sex chromosomes). It's easy to collect, has a relatively high mutation rate, and in particular, its lack of recombination allows for a straightforward assessment of the relationship between haplotypes. Lack of recombination also means that all parts of the molecule are completely linked, which prevents independent evolution of mtDNA's 37 genes and non-coding control region. For this reason, mtDNA is considered a single genetic locus for phylogenetic purposes. Humans have relatively low mitochondrial diversity compared to the other great apes, and reports of this are mostly responsible for the belief that humans have low genetic diversity. However, mtDNA makes up just a few millionths of the human genome,[17] and as a single locus, carries little statistical weight.

When allele frequency data are used to estimate genetic diversity within a population, a frequently reported statistic is the average number of alleles per locus (A), but because rare alleles do not contribute much to overall diversity, the most informative statistic is average heterozygosity (H). This is estimated from both the number of alleles and the frequencies at which they occur, and is generally defined as the percentage of individuals in a population that are heterozygous (have two different alleles) at a random locus. In general, genetic diversity is synonymous with mean heterozygosity.

Table 1. Comparative figures for the genetic diversity of humans and a variety of other large mammals (sampled across much or all of their range except as noted), based on autosomal microsatellites (He and Ho = expected and observed heterozygosity, respectively):

Species	He	Ho
Humans [18]	--	0.776
Humans [19]	--	0.70-0.76
Humans [20]	--	0.588-0.807
Chimpanzees [21]	0.78	0.73
Chimpanzees [22]	--	0.630
African buffalo [23]	0.759	0.729
Leopards [24]	0.36-0.80	--
Jaguars [25]	0.739	--
Polar bears [26]	0.68	--
Brown bears (N. America) [27]	0.26-0.76	0.30-0.79
Brown bears (Scandinavia) [28]	0.709	0.665
Canada lynx [29]	--	0.66
Bighorn sheep [30]	0.681	0.566
Coyote [31]	0.675	0.583
Gray wolf (N. America) [32]	0.620	0.528
Pumas [33]	--	0.52
Bonobos [34]	0.59	0.48
Dogs (42 breeds) [35]	0.616	0.401
African wild dogs [36]	0.643	--
Australian dingo [37]	0.47	0.42
Wolverines (N. America) [38]	0.42-0.68	--
Wolverines (Scandinavia) [39]	--	0.27-0.38
Elk (North America) [40]	0.26-0.53	--

In addition to microsatellites, a 2001 study [41] reviewed the literature on protein variation for 321 mammal species and reported mean expected heterozygosity of 5.1%. In comparison, Takahata (1995) reports an unbiased estimate of protein heterozygosity in humans of 10-14%.[42] Also, Nei's 1987 text Molecular Evolutionary Genetics gives an estimate of mean heterozygosity for classical protein polymorphisms of 0.148 in humans, and has this to say about the general level of genetic diversity in other organisms:

> "In the last two decades, the extent of protein polymorphism has been studied for numerous organisms ranging from microorganisms to mammals by using electrophoresis. In most of these studies, the extent was measured by average gene diversity or heterozygosity. In early days, the estimate of heterozygosity was based on a small number of loci, so that its reliability was low. In recent years, however, most authors are examining a fairly large number of loci (20 loci or more).
>
> Average heterozygosity or gene diversity varies from organism to organism. In general, vertebrates tend to show a lower heterozygosity than invertebrates. If we consider only those species in which 20 more loci are studied, H is generally lower than 0.1 in vertebrates and rarely exceeds 0.15. In invertebrates, a large

> fraction of species again show an average heterozygosity lower than 0.1, but there are many species showing a value between 0.1 and 0.4. In plants, the number of loci studied is generally very small, so that the estimates are not very reliable. However, if we consider only those species in which 20 or more loci are studied, the average heterozygosity is generally lower than 0.15 except in Oenothera, where permanent heterozygosity is enforced by chromosomal translocations (Levin 1975; Nevo 1978; Hamrick et al. 1979; Nevo et al. 1984). The highest level of gene diversity so far observed is that of bacteria (H=0.48 based on 20 loci in Escherichia coli, Selander and Levin 1980; H = 0.49 based on 29 loci in Klebsiella oxytoca, Howard et al. 1985)." [43]

Obviously, humans are not at the low end of the genetic diversity spectrum, particularly in relation to other mammals.

We might wonder how humans could have accumulated so much genetic diversity when we are such an evolutionarily 'young' species, but this assumes that the human species arose by an extreme founding event - a time at which the entire species' diversity resided in just a few individuals - and that all humans today are descended from those few founders. This supposed event is often conflated with the concept of "mitochondrial Eve," a woman who lived roughly 200,000 years ago and is the most recent common ancestor of all human mtDNA. This conflation is incorrect, however, because the coalescence of mtDNA to a single ancestor back in time does not imply a demographic bottleneck, but is expected even in a population of constant size.[44] Avise (2001) has noted that in a hypothetical population with 15,000 breeding females (about three times the long-term human estimate), reasonable variances in reproductive success would likely see mtDNA coalesce to a single founding lineage in 300,000 years (~15,000 human generations), without any change in population size.[45] Thus, the coalescence time of human mtDNA doesn't necessarily have anything to do with a population bottleneck or speciation event, but rather is more or less a function of long-term effective population size, with a large standard error.[46] Variants of nuclear autosomal genes, having a four-fold greater effective population size than mtDNA, generally coalesce in the neighborhood of 800,000 years ago.[47] This indicates that a substantial amount of our existing genetic variation originated in the population ancestral to modern humans.

In sharp contrast to the shallow genealogy of human mtDNA, some alleles of the major histocompatibility complex appear to coalesce over 30 million years ago, long before the emergence of the hominid lineage.[48,49] Some MHC genes are known to have over two hundred alleles,[50] maintained by balancing selection at loci where heterozygosity confers some fitness advantage. Several researchers have demonstrated that humans retain too much ancestral MHC diversity for a severe bottleneck to have ever occurred during human evolution.[51-53] There's fairly wide agreement that the long-term effective population size of humans has been roughly 10,000,[54] making it unlikely that the sum total of our genetic diversity has ever resided in fewer than several thousand individuals.

Additionally, the genetic profile of humans is much different from that of other large mammals that are believed to have experienced a recent demographic bottleneck. The cheetah, for example, is thought to have had a severe population contraction sometime during the late

Pleistocene. While cheetahs apparently have had time to accumulate a moderate amount of variation at some rapidly evolving loci, current populations display very little allozyme or MHC variability.[55] Another example is the moose. Old World and New World subspecies are estimated to have diverged at least 120,000 years ago, but sometime before divergence a bottleneck must have occurred that reduced both allozyme and MHC diversity to a fraction of that found in humans.[56]

Q: Haven't human populations been separated for too short a time for distinct races to have evolved?

Although there is some evidence of non-African archaic contributions to the modern gene pool,[57] it appears likely that current human populations derive largely from a single African population, and diverged something less than 150,000 years ago. While time of separation is important in evolutionary divergence, effective population size can be an equally important factor.[58] While the overall size of the human species has probably never been reduced to a handful of individuals, populations that migrated out of Africa may well have remained relatively small for thousands of years, before beginning to expand toward their current numbers.[59,60] If so, genetic divergence due to drift would have occurred rapidly in the absence of high gene flow.[61]

An example of this has been observed in a North American elk herd re-established from a small number of founders. Between 1915 and 1924, 34 animals from two large herds in the western U.S. were released in north-central Pennsylvania. The herd remained at about this size for 50 years and now numbers about 550. Very low microsatellite heterozygosity (0.222) and very large genetic distance from the source populations (pairwise FST = ~0.45) now characterize this herd.[62]

It has also been proposed (originally by Darwin) that sexual selection (mate choice) may promote the retention of physical features in populations, long after neutral genetic variation had been replaced by gene flow, and that this might help explain the prominent morphological variation among human groups.[63]

At any rate, divergence times for major subdivisions within the human species, while relatively shallow, are certainly not unique when compared to subdivisions within many other mammal species. An appendix to Avise et al. (1998)[64] lists eleven mammal species with major phylogroups that diverged between 100,000 and 500,000 years ago, based on mtDNA sequence divergence. Being a single genetic locus, mtDNA is subject to selection effects and a large amount of random variation, so these times are probably not terribly reliable. For example, mtDNA has indicated 2-3 million years of isolation between western and eastern gorilla subspecies in Africa, but a recent study of multiple nuclear loci provided little support for that time depth.[65] A related situation exists in chimpanzee taxonomy, particularly with regard to the distinctiveness of the eastern (P.t. schweinfurthii) and western equatorial (P.t. troglodytes) subspecies. Studies utilizing nuclear loci,[66,67] as well as more thorough sampling of mtDNA, are calling into question earlier mtDNA results that indicated long separation. As some of these

chimp researchers point out, “The current volatile state of chimpanzee molecular taxonomy is largely due to the fact that studies to date have relied heavily on only a handful of genetic loci.”[68]

Q: Isn’t there actually more genetic distance between populations within the traditional human races than between the major races themselves?

In 1972, Richard Lewontin studied global variation at seventeen protein polymorphisms,[69] and found that about 85% of genetic variation existed between individuals within a given population. The next largest portion, about 8%, was found between populations within continents, and only the remaining 6% of variance could be attributed to differences between the major human races (Fig. 2). While the ~85% within-population figure has been affirmed numerous times, the rest of the pattern for that data set is unusual. Many data sets have been assembled since 1972 for classical polymorphisms and all other genetic markers, many of them larger than Lewontin’s (hence, more reliable - e.g., see Ref. 59) and as a general rule, populations within continents are more closely related to one another than they are to populations of other continents. This pattern can be seen in any matrix of global genetic distances, such as those assembled by Cavalli-Sforza et al. in The History and Geography of Human Genes. Despite this, the results of Lewontin’s study are still frequently cited as the normal pattern of human variation.[70-72]

Population genetic studies often report AMOVA statistics (Analysis of MOlecular VAriance), which show the hierarchical proportions of variance between aggregates of the individuals sampled. The following is a discussion of worldwide data on autosomal microsatellites and RFLPs, Alu insertions, mtDNA and Y chromosome STRPs:

> “The hierarchical AMOVA analysis shows that, with the exception of Y STRPs, all systems show much less
> differentiation between populations within continents than between continents. This result is expected when there is greater gene flow between populations that are in close geographic proximity to one another. The autosomal values…are especially small, ranging from 1.3% for the RSPs to 1.8% for the Alu polymorphisms. This is in agreement with the small continental GST values shown in table 4…they are highly consistent both with one another and with previous analyses of worldwide variation in autosomal microsatellites and RFLPs, which also show considerably greater differentiation between continents than between populations within continents... The
> fact that there is little differentiation between populations within continents has important implications in the forensic setting, in that it supports the current practice of grouping reference populations into broad ethnic categories when autosomal STRP data are used...” [73] (Fig. 3)

Q: How genetically differentiated are human continental populations (the major races) from one another compared to populations of other species?

Before the advent of conservation biology and modern phylogenetics, subspecies were normally delineated by morphological characteristics. The "seventy-five percent rule" goes back to 1949, stating that subspecies classification is merited if at least 75% of individuals can be correctly assigned to their group by inspection.[74] This rule isn't in common use today, but the importance of genetically-based morphological differences is still apparent in many recent phylogenetic studies. Some biologists argue that a 70 or 75 percent rule should still be a standard criterion in taxonomy, as applied to individuals outside of hybrid zones where the ranges of subspecies overlap.[75]

On the basis of morphology, we can compare the traditional human races (as well as some minor races) to chimpanzee subspecies. Individuals of the former can be correctly assigned at much greater than 75% accuracy,[76] while the latter are morphologically indistinct, and difficult or impossible to classify when raised in captivity.[77,78]

Of course, the domestic dog demonstrates that morphological difference doesn't necessarily correlate with underlying genetic difference, so let's look at population differentiation from a genetic perspective. Many measures of divergence or 'genetic distance' are in use today, the most common being FST, originally developed by the late population geneticist Sewall Wright. FST is a statistic that describes the proportion of variance within a species that is due to population subdivision. It can be estimated in a variety of ways (e.g., by AMOVA [79] or theta [80]), but the general expression is FST = (Ht-Hs)/Ht where Ht is the genetic diversity within the total population, and Hs the average diversity within subpopulations. Its value can be considered inversely proportional to gene flow, or indicative of the length of time two populations have been evolving separately, and may vary according to which locus or family of loci are under study. As mentioned earlier, haploid loci like mtDNA and the NRY have effective population sizes one quarter that of autosomal loci, making them much more sensitive to drift and thus to the effect of population subdivision. Other types of loci have their own unique evolutionary characteristics, so we need to remember that an FST value based on one class of loci may or may not be representative of the overall evolutionary distinctiveness of the populations in question. For these reasons, values based on several types of loci should be considered before drawing any firm conclusions.

Keeping the preceding caveats in mind, these are qualitative guidelines suggested by Sewall Wright for interpreting FST:

> "The range 0 to 0.05 may be considered as indicating little genetic differentiation.
> The range 0.05 to 0.15 indicates moderate genetic differentiation.
> The range 0.15 to 0.25 indicates great genetic differentiation.
> Values of FST above 0.25 indicate very great genetic differentiation." [81]

Table 2.
Here are some comparative figures for humans and other species (again, sampled across most or all of their ranges except as noted), based on autosomal microsatellites:

Species	FST
Gray wolves (North America) [82]	0.168
Pumas [83]	0.167 (mean pairwise)
Humans (14 populations) [84]	0.155 (AMOVA)
Asian dogs (11 breeds) [85]	0.154
European wildcats (Italy) [86]	0.13
Humans (44 populations) [87]	0.121 (AMOVA)
Coyotes (North America) [88]	0.107
Wolverines (North America) [89]	0.067 (mean pairwise)
Jaguars [90]	0.065
African buffalo [91]	0.059
Polar bears [92]	0.041 (mean pairwise)
Canada lynx [93]	0.033
Humpback whales [94]	0.026 (mean pairwise)

Additionally, Uphyrkina et al. (2001) employed mtDNA and microsatellites to identify nine leopard subspecies by our phylogeographic criteria. Unfortunately for the sake of comparison, the authors reported microsatellite RST rather than FST values. RST is an FST analogue, but their values can be quite different numerically. However, if the RST/FST ratio for leopards is similar to those of other felids [95,96] the maximum reported RST value of 0.363 would correspond roughly to an FST of 0.14-0.15, very similar to the human value at microsatellite loci. The mean proportion of private (population-specific) microsatellite alleles for the nine revised leopard subspecies was found to be 6.3%, compared to a mean value of 7.1% for three major human continental populations [97] while the mean Nei's genetic distance DS for allozymes between the leopard subspecies identified by Miththapala et al. (1996) is 0.019 (range 0.002-0.047) [98] and can be compared to the protein distances between three major human races (mean 0.037; range 0.028-0.048). [99]

Wolverines, polar bears, Canada lynx and humpback whales have not traditionally been divided into subspecies, while two or more subspecies (or 'breeds' in the case of the Asian dogs) have been named in all of the remaining non-human species listed above. The overall FST value for African buffalo is not particularly large, but the mean value of 0.095 between the central African population and other populations was considered large enough to support their traditional subspecies status. Based on cranial morphology and geography, 24 subspecies of the gray wolf in North America were reduced to five in 1995, while North American coyotes are considered to have eastern and western subspecies.

For our purposes, the studies of population structure in the big cats are especially informative, since these used phylogeographic criteria to suggest possible taxonomic revision. Jaguars have traditionally been divided into eight subspecies, but Eizirik et al. (2001) considered the population structure too weak (FST = 0.065) to warrant naming any. In contrast, distinct phylogroups were readily apparent within both pumas and leopards, although somewhat fewer than classically described (6 vs. 32 in pumas, and 8 or 9 vs. 27 in leopards).

It should be noted that high phenotypic diversity in some domestic animals (such as the Asian dogs) is mostly the result of selective breeding for quantitative traits, rather than the long-term allopatry or local adaptation that leads to morphological distinctiveness in "natural" populations. As expected, the average microsatellite distance between these dog breeds as measured by Nei's genetic distance DA (0.194) [100] is correspondingly smaller than the average distance between fourteen human populations (0.322). [101]

Human FST values of 12-15% are typical not just for microsatellites, but also for classical protein polymorphisms,[102] autosomal RFLPs[103] and Alu insertions.[104] Values for mitochondrial DNA and the Y chromosome are substantially higher. It would seem, then, that the level of genetic differentiation among human populations is not especially small, and in fact is entirely adequate for race designation, particularly when coupled with consistent morphological differences.

Q: Which human populations qualify as major races?

The construction of reliable evolutionary trees involves a number of technical issues, such as sampling design, mutation mechanisms, genetic distance measures and particularly, tree-building algorithms. Nonetheless, the topology of human trees (Figs. 4, 5) is remarkably consistent regardless of which class of loci are considered, and principal component analysis of genetic data also produces predictable clustering (Fig. 6). Either method gives a good visual overview of the general relatedness of the world's populations.

By analysis of classical markers, Nei & Roychoudhury (1993) identified five major human clades: sub-Saharan Africans, Caucasians, Greater Asians, Australopapuans and Amerindians. Evolutionary trees constructed with autosomal RFLPs,[105] microsatellites[106] and Alu insertions[107] show similar topology. Frequently, Amerindians are grouped together with Asians, indicating four major clades, and it has been suggested that this should be a minimum.[108] Obviously, additional structure exists within each of these groups, but as we've seen, it's relatively weak compared to the differentiation among the ones listed here. For this reason alone, the term 'race' applies well to these major groupings.

In terms of our phylogeographic definition, each of the major human clades has a geographical association (slightly less clear today than 500 years ago, but only slightly); each has a distinguishing set of phenotypic traits; phylogenetic partitioning is apparent and consistent at multiple genetic loci; and substantial intergroup genetic distances (i.e., FST) indicate unique natural histories on an evolutionary timescale.

The criticism can be made that the placement of some populations located between the "cores zone" of these major races (e.g., Europe or East Asia) is ambiguous. However, in non-human taxonomy this would not normally invalidate the subspecies status of well-differentiated core populations.[109,110] In fact, zones of intergradation have traditionally been taken as evidence that core groups are indeed subspecies rather than different species.[111] While some clinal variation in the genetic traits of subspecies is generally the rule, human variation tends to show

extensive zones where clinal gradients are relatively flat, separated by short zones of steeper gradient. This pattern can be seen on the dust jacket illustration of The History and Geography of Human Genes.

In conclusion…

Some will find provocative the idea that humans display a subspecies-like population structure, but given that the major human subdivisions revealed by modern genetics had already been recognized as early as 1775,[112] it shouldn't be as provocative as the alternative notion, i.e., that human races don't exist.

So if we do belong to different biological races, what, if anything, does this mean? Subspecies are closely related by definition, and human races appear to be less distant than the major phylogroups of many other species.[113] While FST values for neutral variation are not negligible from a population genetics point of view, it's significant that the overwhelming majority of genetic variation is found within populations, reaffirming the importance of treating people as individuals. It's also significant that the FST value for the most prominent racial trait - skin color - has been estimated to be about 0.60,[114] which means that the visible variation between races greatly exaggerates overall genetic differences. Admixture in some populations further clouds the picture. The average European contribution to the gene pool of American blacks has been found to be about 20%,[115] and admixture between the major races in some other regions is substantially higher.

Nevertheless, when the taxonomic term is used consistently across species, it's difficult to see any justification for the common assertion that human races are merely 'social constructs.' The motivation behind the assertion is a positive one, but denying biological realities at the outset is unlikely to lead to productive social dialogue on coping with human differences.

Supplement

F_{ST} Follies

In 1998, *American Anthropologist* published a paper by Alan Templeton entitled "Human Races - A Genetic and Evolutionary Perspective" [116] which seems to have had broad influence on the race question within anthropology and the social sciences. In the first section of the paper, Templeton cites a 1997 article from *Herpetological Review* entitled "Subspecies and Classification."[117] Templeton asserts that, according to this paper, an FST value of .25 or .30 between populations is a "standard criterion" for subspecies classification. He then provides a graph showing FST (or FST analogue) values for humans and 12 other species of large mammals (Fig. 7). (The human value of 0.156 is from a 1997 paper, "An Apportionment of Human DNA

Diversity"[118] in *Proceedings of the National Academy of Sciences*.) Two of the non-human values listed are lower than that for humans, but the other ten values are substantially higher, and appear to support Templeton's claim that human populations are only weakly differentiated.

There are several curious things about this. First, there is little, if any, corroboration in the recent literature for an FST value of .25 or .30 being a standard criterion for subspecies designation. Secondly, if you actually read the paper by Smith et al., they never mention anything about FST values. Rather, they say that "overlap [of differentiae] exceeding 25-30% does not qualify for taxonomic recognition of either dichopatric populations or parapatric populations outside of their zones of intergradation." What the authors are referring to here is not an FST value, but simply the long-standing 75 (or 70) percent rule discussed earlier.[119] Templeton's misinterpretation is all the more obvious when you consider that this subspecies rule and FST have an inverse relationship, i.e., a 75 percent rule implies greater differentiation than does a 70 percent rule, whereas an FST value of 0.25 indicates lesser differentiation than does a value of 0.30. Additionally, FST is generally used to assess neutral genetic variation in these kinds of studies, which, as we've seen, can be quite different from expressed morphological variation.

The most interesting thing, however, about Templeton's FST comparison is the fact that he uses a human value (0.156) based on *autosomal* loci (microsatellites and RFLPs), while nine of the ten largest non-human values, including the eight highest, are based on *mitochondrial* DNA. This is quite misleading, because FST values for mtDNA are *expected* to be much higher than autosomal values.[120,121] The primary mechanism causing populations to diverge is usually genetic drift, and the magnitude of the effects of drift is inversely proportional to population size, as shown by Bodmer and Cavalli-Sforza (1976) through computer simulations (reproduced in Ref. 17, p.14). The four-fold greater effective population size of autosomal loci vs. mtDNA virtually ensures that FST values based on the latter will be substantially greater than values based on the former, and in fact this is nearly always observed in population studies. Since mtDNA is maternally inherited, sex-biased dispersal can also play a role in elevating FST for species in which males disperse over greater distances than do females.

In the present paper, every attempt has been made to use comparable data.

Some typical comparative FST values for autosomal and mitochondrial loci, respectively, for similar or identical samples:

Jaguar[122]	0.065 vs. 0.295
Puma [123]	0.167 vs. 0.467
Gray wolf [124,125]	0.168 vs. 0.76

References

1. Templeton, A.R. 1998. Human races: An evolutionary and genetic perspective. *American Anthropologist* 100:632-650.

2. Bodmer, W.F & Cavalli-Sforza, L.L. 1976. *Genetics, Evolution, and Man.* WH Freeman and Company. San Francisco. p561

3. Hartl, D., Clark, AG. 1989. *Principles of population genetics.* Sinauer Associates, Sunderland, MA. p301.

4. Mayr, E. (1963) *Animal Species and Evolution.* Belknap, Cambridge, MA. Cited in Ref. 82.

5. Mayr, Ernst. "What Is A Species and What Is Not?," *Philosophy Of Science,* 63:262-277. June 1996.

6. Smith, H.M., Chiszar, D., Montanucci, R.R., 1997. Subspecies and Classification. *Herpetological Review* 28:13-16.

7. Templeton 1998.

8. Wright, S. 1978. *Evolution and the Genetics of Populations, Vol. 4, Variability Within and Among Natural Populations.* Univ. Chicago Press, Chicago, Illinois. p439.

9. E.g., see Graves, Joseph, 2001. *The Emperor's New Clothes: Biological Theories of Race at the Millennium.* Rutgers University Press, Piscataway, New Jersey. p 2. Graves suggests that valid racial categories would exist if "pairs of individuals from different races either had reduced capacity, or no capacity, to produce viable offspring."

10. Avise, J.C., Ball, R.M. 1990. Principles of genealogical concordance in species concepts and biological taxonomy. *Oxford Surveys in Evolutionary Biology* 7:45-67.

11. O'Brien, S.J., Mayr, E. 1991. Bureaucratic Mischief: Recognizing Endangered Species and Subspecies. *Science.* 2 51:1187-1188.

12. Miththapala, S., Seidensticker, J., O'Brien, S.J. 1996. Phylogeographic Subspecies Recognition in Leopards (Panthera pardus): Molecular Genetic Variation. *Conservation Biology* 10:1115-1132.

13. Miththapala et al. 1996.

14. Diamond, J. Race without color. *Discover.* November 1994.

15. E.g., Olson, S. April 2001. The Genetic Archaeology of Race. *Atlantic Monthly.* http://www.theatlantic.com/issues/2001/04/olson-p1.htm

16. Cooper, G., Amos, W., Bellamy, R., Siddiqui, M.R., Frodsham, A., Hill, A., Rubinsztein, D. 1999. An Empirical Exploration of the delta-mu^2 Genetic Distance for 213 Human Microsatellite Markers. *American Journal of Human Genetics* 65:1125-1133.

17. Cavalli-Sforza, LL., Menozzi, P., Piazza, A. 1994. *The history and geography of human genes.* Princeton University Press, Princeton. p83.

18. Wise, C., Sraml, M., Rubinsztein, D., Easteal. S. 1997. Comparative Nuclear and Mitochondrial Genome Diversity in Humans and Chimpanzees. *Molecular Biology and Evolution* 14:707-716.

19. Jorde, L., Rogers, A., Bamshad, M., Watkins, W.S., Krakowiak, P., Sung, S., Kere, J., Harpending, H. April 1997. Microsatellite Diversity and the Demographic History of Modern Humans. *Proceedings of the National Academy of Sciences* 94:3100-3103.

20. Bowcock. A.M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J.R., Cavalli-Sforza, L.L. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 368:455-457.

21. Reinartz, G.E., Karron, J.D., Phillips, R.B. and Weber, J.L. 2000. Patterns of microsatellite polymorphism in the range-restricted bonobo (Pan paniscus): considerations for interspecific comparison with chimpanzees (P. troglodytes). *Molecular Ecology* 9:315-328.

22. Wise et al. 1997.

23. Van Hooft, W.F., A. F. Groen and H. H. T. Prins, Microsatellite analysis of genetic diversity in African buffalo (Syncerus caffer) populations throughout Africa. *Molecular Ecology* (2000) 9, 2017-2025

24. Uphyrkina, O., Johnson, W.E., Quigley, H., Miquelle, D., Marker, L., Bush, M., O'Brien, S.J. 2001. Phylogenetics, genome diversity and origin of modern leopard, Panthera pardus. *Molecular Ecology* 10:2617-2633.

25. Eizirik, E., Kim, J., Menotti-Raymond, M., Crawshaw, P.G. Jr., O'Brien, S.J., Johnson, W.E. 2001. Phylogeography, population history and conservation genetics of jaguars (Panthera onca, Mammalia, Felidae). *Molecular Ecology* 10:65-79.

26. Paetkau, D., Amstrup, S.C,. Born, E.W., Calvert, W., Derocher, A.E., Garner, G.W., Messier, F., Stirling, I., Taylor, M.K., Wiig, Strobeck,,C. 1999. Genetic structure of the world's polar bear populations. *Molecular Ecology* 8:1571-1584.

27. Paetkau, D., Waits, L., Clarkson, P., Craighead, L., Vyse, E., Ward, R., Strobeck, C. 1998. Variation in genetic diversity across the range of North American brown bears. *Conservation Biology* 12:418-429.

28. Waits, L., Taberlet, P., Swenson, J.E., Sandegren, F., Franz_n, R. 2000. Nuclear DNA microsatellite analysis of genetic diversity and gene flow in the Scandinavian brown bear (Ursus arctos). *Molecular Ecology* 9:421-431.

29. Schwartz, M.K., Mills, L.S., McKelvey, K.S., Ruggiero, L.F., Allendorf, F.W. 2002. DNA reveals high dispersal synchronizing the population dynamics of Canada lynx. *Nature* 415:520-522.

30. Forbes, S.H., Hogg, J.T., Buchanan, F., Crawford, A., Allendorf, F. 1995. Microsatellite evolution in congeneric mammals: domestic and bighorn sheep. *Molecular Biology and Evolution* 12:1106-1113.

31. Garcia-Moreno, J., Matocq, M., Roy, M., Geffen, E., Wayne, R.K. 1996. Relationships and Genetic Purity of the Endangered Mexican Wolf based on Analysis of Microsatellite Loci. *Conservation Biology* 10:376-389.

32. Garcia-Moreno et al. 1996.

33. Culver, M., Johnson, W.E., Pecon-Slattery, J., O'Brien, S.J. Genomic Ancestry of the American Puma (Puma concolor). 2000. *Journal of Heredity* 91:186-197.

34. Reinartz et al. 2000.

35. Garcia-Moreno et al. 1996.

36. Girman, D. J., C. Vil_, E. Geffen, S. Creel, M. G. L. Mills, J. W. Mcnutt, J. Ginsberg, P. W. Kat, K. H. Mamiya and R. K. Wayne. 2001. Patterns of population subdivision, gene flow and genetic variability in the African wild dog (Lycaon pictus). *Molecular Ecology* 10:1703-1723.

37. Wilton, A.N., Steward, D.J., Zafiris, K. 1999. Microsatellite variation in the Australian dingo. *Journal of Heredity* 90:108-111.

38. Kyle, C J. and C. Strobeck, Genetic structure of North American wolverine (Gulo gulo) populations. *Molecular Ecology*
(2001) 10, 337-347

39. Walker, C.W., Vil_, C., Landa, A., Lind_n M., Ellegren, H. 2001., Genetic variation and population structure in Scandinavian wolverine (Gulo gulo) populations. *Molecular Ecology* 10:53-63.

40. Polziehn, R.O., Hamr, J., Mallory, F.F., Strobeck, C. 2000. , Microsatellite analysis of North American wapiti (Cervus elaphus) populations. *Molecular Ecology* 9:1561-1576.

41. Makarieva, A., 2001. Variance of protein heterozygosity in different species of mammals with respect to the number of loci studied. *Heredity* 87:41-51.

42. Takahata, N. 1995. A Genetic Perspective on the Origin and History of Humans. *Annu. Rev. Ecol. Syst.* 26:343-372.

43. Nei, Masatoshi. *Molecular Evolutionary Genetics.* Columbia University Press, 1987 pp.192-193.

44. Nei, M., Livshits, G., 1990. Evolutionary relationships of Europeans, Asians, and Africans at the Molecular Level. *Population Biology of Genes and Molecules.* Eds. N. Takahata and J.F. Crow. Baifukan, Tokyo. p259.

45. Avise, John C. 2000. *Phylogeography — The History and Formation of Species.* Harvard University Press, Cambridge, MA. p114.

46. Nei and Livshits 1990. p259.

47. Nachman, M.W., Bauer, V.L., Crowell, S.L., Aquadro, C.F., 1998. DNA Variability and Recombination Rates at X-Linked Loci in Humans. *Genetics* 150:1133-1141.

48. McAdam, S., Boyson, J., Liu, X., Carber, T., Hughes, A., Bontrop. E., Watkins, D. 1995. Chimpanzee MHC Class I A Locus Alleles Are Related to Only One of the Six Families of Human A locus Alleles. *The Journal of Immunology* 154: 6421-6429.

49. Takahata, N. 1993. Allelic genealogy and human evolution. *Molecular Biology and Evolution* 10:2-22.

50. MHC Consortium. *Nature* 401:921-923. Oct 28, 1999.

51. Li, W. and Sadler, L. 1992. DNA variation in humans and its implications for human evolution. *Oxford Surveys in Evolutionary Biology*, 8:111-134.

52. Takahata 1993.

53. Ayala, F.J., Escalante, A., O'hUigen, C., Klein, J., 1994. Molecular genetics of speciation and human origins. *Proceedings of the National Academy of Sciences* 91:6787-6794.

54. Harpending, H. and Eller, E. 1999. Human diversity and its origins. In *The Biology of Biodiversity*, ed. M. Kato, pp.301-14. Tokyo: Springer-Verlag.

55. Menotti-Raymond, M., O'Brien, S.J. 1993. Dating the genetic bottleneck of the African cheetah. *Proceedings of the National Academy of Sciences* 90:3172-3176.

56. Mikko, S., Andersson, L. 1995. Low major histocompatibility complex class II diversity in European and North American moose. *Proceedings of the National Academy of Sciences* 92:4259-4263.

57. Harding, R.M., Fullerton, S.M., Griffiths, R.C., Bond, J. Cox, M.J., Schneider, J.A., Moulin, D.S., Clegg, J.B., 1997. Archaic African and Asian lineages in the genetic ancestry of modern humans. *American Journal of Human Genetics* 60: 772-789.

58. Kalinowski, S.T. 2002. Evolutionary and statistical properties of three genetic distances. *Molecular Ecology* 11:1263-1273.

59. Zhivotovsky, L., Bennett, L, Bowcock, A., Feldman, M. 2000. Human Population Expansion and Microsatellite Variation. *Molecular Biology and Evolution* 17(5):757-767.

60. Nei, M. and Roychoudhury, A. 1993. Evolutionary Relationships of Human Populations on a Global Scale. *Molecular Biology and Evolution* 10(5):927-943.

61. Mills, L.S., Allendorf, F.W. 1996. The One-Migrant-per-Generation Rule in Conservation and Management. *Conservation Biology* 10:1509-1518.

62. Williams, C.L., Serfass, T.L., Cogan, R., Rhodes, O.E. Jr. 2002. Microsatellite variation in the reintroduced Pennsylvania elk herd. *Molecular Ecology* 11:1299-1310.

63. Harpending and Eller 1999.

64. Avise, J.C., Walker, D., Johns, G.C. (1998) Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proc. R. Soc. Lond. B* 265, 1707-1712.

65. Jensen-Seaman, M. June 2000. Western and Eastern Gorillas: Estimates of the Genetic Distance. *Gorilla Journal* Vol. 20.

66. Kaessmann, H., Wiebe, V., Paabo, S. 1999. Extensive Nuclear DNA Sequence Diversity Among Chimpanzees. *Science* 286:1159-1162.

67. Gonder, M. K. 2000 *Evolutionary genetics of chimpanzees (Pan troglodytes) in Nigeria and Cameroon.* PhD thesis, The City University of New York, New York, USA. (Cited in Ref. 68.)

68. Gagneux, Pascal, M. Katherine Gonder, Tony L. Goldberg and Phillip A. Morin. 2001. Gene flow in wild chimpanzee populations: what genetic data tell us about chimpanzee movement over space and time. *Phil. Trans. R. Soc. Lond. B.* 356:889-897.

69. Lewontin, R. C. 1972. The apportionment of human diversity. *Evolutionary Biology* 6:381-398. Cited in Ref. 35.

70. Ruvolo, M. 1997. Genetic diversity in hominoid primates. *Annual Review of Anthropology* 26:515-540.

71. Marks, J. 2002. *What it Means to be 98% Chimpanzee — Apes, People and Their Genes.* University of California Press, Los Angeles. p82.

72. Smedley, A. The American Anthropological Association Statement on "Race." http://www.aaanet.org/stmts/racepp.htm

73. Jorde, L.B., Watkins, W.S., Bamshad, M.J., Dixon, M.E., Ricker, C.E., Seielstad, M.T., Batzer, M A. 2000. The Distribution of Human Genetic Diversity: A Comparison of Mitochondrial, Autosomal, and Y-Chromosome Data. American *Journal of Human Genetics* 66:979-988.

74. Amadon, D. 1949. The seventy-five percent rule for subspecies. *Condor* 51: 251-258.

75. Smith et al. 1998.

76. Wright 1978. p439.

77. Stone, A.C., R. Griffiths, S. Zegura, M. Hammer. 2002. High levels of Y chromosome nucleotide diversity in the genus Pan. *Proceedings of the National Academy of Sciences* 99:43-48.

78. Kaessmann et al. 1999.

79. Excoffier, L., Smouse, P.E., Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479-494.

80. Weir, B.S., Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38:1358-1370.

81. Hartl and Clark 1989. p118-119.

82. Roy, M.S., Gefen, E., Smith, D., Ostrander, E.A., Wayne, R.K. 1994. Patterns of Differentiation and Hybridization in North American Wolflike Canids, Revealed by Analysis of Microsatellite Loci. *Molecular Biology and Evolution* 11:553-570.

83. Culver et al. 2000.

84. Barbujani, G., Magagni, A., Minch, E., Cavalli-Sforza, L.L. 1997. An apportionment of human DNA diversity. *Proceedings of the National Academy of Sciences* 94: 4516-4519.

85. Kim, K. S., Y. Tanabe, C. K. Park, and J. H. Ha. Genetic Variability in East Asian Dogs Using Microsatellite Loci Analysis. *The Journal of Heredity* 92:398-403. 2001.

86. Randi, E., Pierpaoli, M., Beaumont, M., Ragni, B., Sforzi, A. 2001. Genetic Identification of Wild and Domestic Cats (Felis silvestris) and Their Hybrids Using Bayesian Clustering Methods. *Molecular Biology and Evolution* 18:1679-1693.

87. Jorde et al. 2000.

88. Roy et al. 1994.

89. Kyle and Strobeck 2001.

90. Eizirik et al. 2001.

91. Van Hooft et al. 2000.

92. Paetkau et al. 1999.

93. Schwartz et al. 2002.

94. Valsecchi, E., Palsboll, P., Hale, P., Glockner-Ferrari, D., Ferrari, D, Clapham, P., Larsen, F., Mattila, D., Sears, R., Sigurjonsson, J., Brown, M., Corkeron, P., Amos, B. 1997. Microsatellite Genetic Distances Between Oceanic Populations of the Humpback Whale (Megaptera novaeangiae). *Molecular Biology and Evolution* 14:355-362.

95. Randi et al. 2001.

96. Eizirik et al. 2001.

97. Jorde et al. 2000.

98. Miththapala et al. 1996.

99. Nei, M. and G. Livshits. 1989. Genetic relationships of Europeans, Asians and Africans and the origins of modern Homo sapiens. *Human Heredity* 39:276-281.

100. Kim et al. 2001.

101. Nei, M. and Takezaki, N. 1996. The root of the phylogenetic tree of human populations. *Molecular Biology and Evolution* 13:170-177.

102. Cavalli-Sforza et al. 1994. p117.

103. Barbujani et al. 1997.

104. Romualdi, C., Balding, D., Nasidze, I., Risch, G., Robichaux, M., Sherry, S., Stoneking, M., Batzer, M., Barbujani, G. 2002. Patterns of Human Diversity, within and among Continents, Inferred from Biallelic DNA Polymorphisms. *Genome Research* 12:602-612.

105. Nei and Takezaki 1996. p173.

106. Bowcock et al. 1994. p457.

107. Stoneking, M., Fontius, J.J., Clifford, S.L., Soodyall, H., Arcot, S.S., Saha, N., Jenkins, T., Tahir, M.A., Deininger, P.L., Batzer, M.A. 1997. Alu Insertion Polymorphisms and Human Evolution: Evidence for a Larger Population Size in Africa. *Genome Research* 7:1061-1071.

108. Wright 1978. p439.

109. Avise and Ball 1990. p59.

110. Smith et al. 1997. p13.

111. Wright 1978. p4.

112. Blumenbach, J.F. 1775. *De generis humani varietate nativa*. M.D. thesis, University of Gottingen. (Cited in Ref. 17, p17.)

113. Nei 1987. p241.

114. Relethford, J.H. 1992. Cross-cultural analysis of migration rates: effects of geographic distance and population size. *American Journal of Physical Anthropology* 89:459-66.

115. Parra, E., Marcini, A., Akey, J., Martinson, J., Batzer,M.A., Cooper, R., Forrester, T., Allison, D.B., Deka, R., Ferrell, R.E., Shriver, M.D. 1998. Estimating African American Admixture Proportions by Use of Population-Specific Alleles. *American Journal of Human Genetics* 63:1839-1851.

116. Templeton 1998.

117. Smith et al. 1997.

118. Barbujani et al. 1997.

119. Chiszar, D. Personal communication, Feb. 2002.

120. Whitlock, M. and McCauley, D. 1999. Indirect measure of gene flow and migration: FST [does not equal] 1/(4Nm+1). *Heredity* 82: 117-125

121. Takahata 1993.

122. Eizirik et al. 2001.

123. Culver et al. 2000.

124. Roy et al. 1994.

125. Wayne, R.K., Lehmann, N., Allard, M.W., Honeycutt, R.L. 1992. Mitochondrial DNA Variability of the Gray Wolf: Genetic Consequences of Population Decline and Habitat

Fragmentation. *Conservation Biology* 6:559-569.

The citation for this paper is:

Goodrum, J. The Race FAQ. July 2002. http://www.goodrumj.com/RaceFaq.html

(back to main page)
