



News and Views

## Neandertal taxonomy reconsidered...again: a response to Harvati et al. (2004)

James C.M. Ahern<sup>a,\*</sup>, John D. Hawks<sup>b</sup>, Sang-Hee Lee<sup>c</sup>

<sup>a</sup>Department of Anthropology, University of Wyoming, 3431, 1000 E. University Ave., Laramie WY 82071, USA

<sup>b</sup>Department of Anthropology, University of Wisconsin, Madison, USA

<sup>c</sup>Department of Anthropology, University of California, Riverside, USA

Received 19 February 2004; accepted 20 October 2004

*Keywords:* Late Pleistocene; Species; Systematics; Human evolution; Modern human origins

### Introduction

The debate concerning the fate of the Neandertals has been ongoing since shortly after the discovery of the Feldhofer 1 skeleton in 1856. Were the Neandertals an extinct side-branch in human evolution or were they ancestral to modern humans? If they were an extinct side-branch, reproductively isolated from modern humans, then they should be classified as a distinct species, *Homo neanderthalensis*. On the other hand, if Neandertals were not reproductively isolated from modern humans, they represented a temporospatial population of *Homo sapiens*.

Harvati et al. (2004) have made a welcome contribution to the literature regarding the taxonomy of the Neandertals. They used patterns of three-dimensional morphometric variation within

and between selected living catarrhine species to examine whether or not Neandertals and modern humans should be considered different species. They examined more than 1000 specimens from extant taxa, and included five Neandertal and five Upper Paleolithic specimens. Based on the morphological distances between paired intraspecific and interspecific groups in their study, Harvati et al. (2004) concluded that Neandertal specific distinction from modern humans was strongly supported. Although we are particularly impressed with the breadth of their study, we do not believe that their results refute the conspecificity of Neandertals and modern humans. An alternative interpretation suggests potentially more productive approaches to the problem.

### Problems with the methods

We find three aspects of Harvati et al.'s methods problematic: 1) sample sizes were not

\* Corresponding author. Tel.: +1 307 766 4911; fax: +1 307 766 2473.

E-mail addresses: [jahern@uwyo.edu](mailto:jahern@uwyo.edu) (J.C.M. Ahern), [jhawks@wisc.edu](mailto:jhawks@wisc.edu) (J.D. Hawks), [shlee@ucr.edu](mailto:shlee@ucr.edu) (S.-H. Lee).

adequate for the statistical technique employed, 2) the sample comparisons confused within- and between-population variation, and 3) the compositions of the fossil samples themselves, which appear to have been sex-biased.

The issue of sample size affects all fossil research, but especially multivariate methods. The Mahalanobis distance is effectively a weighted Euclidean distance, meant to compare different variables that possibly covary. But estimation of a covariance matrix requires an exceptionally large sample size, which grows geometrically as more variables are considered. A covariance matrix of  $d$  variables contains  $d(d-1)/2$  cells, since the matrix has to be symmetrical. Therefore, a good estimate for the covariance matrix with  $d$  variables has  $d(d-1)/2$  degrees of freedom, thus requiring  $d(d-1)/2$  observations. In Harvati et al.'s case, the total of 15 characters requires 105 cases to have a good estimate for the covariance matrix. Few populations or species in their data set met this requirement. Can this method ever be fruitfully applied to fossil samples? We think that the possible strength of multivariate techniques may merit the attempt, but some assumptions about the covariance matrix must be made and justified. A sample of five Neandertal and five Upper Paleolithic specimens, as employed in Harvati et al.'s study, cannot be expected to yield valid results. It is unclear whether or not Harvati et al.'s employment of bootstrapping effectively dealt with the small sample size issue, since each one of the bootstrapped samples was not large enough to accurately calculate the covariance matrix.

One problem that stems from the nature of primate variation is a possible confusion of within-population and between-population variation. For example, gorillas have high within-population metric variation because of their large degree of sexual dimorphism. In contrast, chimpanzees have low within-population variation. While morphometric scaling eliminates size differences, it does not eliminate shape differences related to size. Likewise, the use of the Mahalanobis metric scales differences in terms of within-population variances and covariances, but does not consider higher moments of the distribution. Gorillas are highly bimodal in size and shape, while chimpanzees are

unimodal, except for their canines, which Harvati et al. excluded. Thus, comparing two gorilla subspecies results in a larger mean distance, and comparing two chimpanzee subspecies results in a smaller mean distance *merely because the two species differ in within-population variation*.

In addition to the size and variation of the samples, it is possible that sex-related sample bias has exaggerated the reported Neandertal vs. Upper Paleolithic and Neandertal vs. modern human comparisons. Although the choice of specimens did reflect a focus on overall cranial measurements, it led to a Neandertal sample that we believe was entirely male, while the Upper Paleolithic sample was three-fifths female. Such sex-bias can drastically alter perceived differences between fossil samples, since males tend to appear more archaic than females within their populations (Ahern et al., 2002). The effects of a male-biased Neandertal sample are easy to see. Harvati et al. used a resampling test, choosing five specimens randomly in each of 10,000 sample replicates, in an attempt to reflect the properties of the Neandertal sample. But if there are any differences at all between males and females in a sample, then 10,000 random replicates from the comparative samples must be closer to the population mean (including both sexes) than 10,000 all-male replicates. In other words, the Neandertal comparisons exaggerated distance, while all other comparisons minimized distance. Without knowing the sexual dimorphism of the scaled landmarks used by Harvati et al., it is impossible to say what effect such bias may have had on the tests of significance, but the effects must have certainly been nonzero.

The study could have been improved by the inclusion of well known specimens. There are additional well preserved Upper Paleolithic crania that would have helped round out the Upper Paleolithic sample (e.g., Arene Candide 1, Barma Grande 2, Barma Grande 5, Bruniquel 24, Cap Blanc 1, Chancelade, Combe-Capelle, Cro-Magnon 2, Dolní Věstonice 3, Dolní Věstonice 15, Dolní Věstonice 16, Gough's Cave 1, Grotte des Enfants 4, Grotte des Enfants 5, Laugerie-Basse 4, St. Germaine la Riviere 4, Oberkassel 1, Oberkassel 2, Ortucchio 2, San Teodoro 1, San Teodoro 2, San Teodoro 3, San Teodoro 7). Had the authors

limited their focus to braincases, as opposed to entire crania, their samples would have been drastically improved (UP:  $n = 35$  vs.  $n = 5$ ; Neandertal:  $n = 11$  vs.  $n = 5$ ). While excluding faceless specimens in the interest of preservation is understandable, excluding the well preserved fossils cannot be justified by appeal to either too much reconstruction or to the problems with measuring casts, since the authors do use the heavily reconstructed Amud 1 and La Ferrassie 1 specimens and casts of Shanidar 1 and the Předmostí crania in their original paper. Unfortunately, even if all of the appropriate fossils were included, sex-bias would have remained a significant issue, since the currently known sample of well preserved Neandertal crania is likely almost exclusively male. Certainly the fossil record poses similar problems for any multivariate study, and they may be currently insurmountable.

### Extant species as referents

We contend that using the *range* of variation in living species is not a guaranteed means of accurately delineating paleontological species. Citing the volume edited by Kimbel and Martin (1993), Harvati et al. (2004: 1147) stated: “there is consensus that species in paleontology should be equivalent to living ones in the variation that they accommodate.” The dialogue in Kimbel and Martin (1993) actually reflects the lack of consensus on this issue (cf. Kelley, 1993; Kimbel and Rak, 1993; Krishtalka, 1993). For example, the Miocene ape species *Lufengpithecus lufengensis* has been interpreted as exhibiting greater variation than that seen in living relatives (Kelley, 1993). The argument that fossil species’ boundaries should be tied to variation within living species results in fossil species that are comparable in their variation to living species, but does it create evolutionarily meaningful species?

In order to delineate species that have evolutionary meaning, as opposed to classificatory utility, the range of variation cannot be the ultimate test of species boundaries (Kelley, 1993). Rather it is the *pattern* of variation that matters (e.g., Ahern, 1998; Kimbel and Rak, 1993). Does the pattern of

variation indicate reticulating (within-species) or cladistic (multiple species) relationships? The persistence of Neandertal-like features among post-Neandertal Europeans (Frayer, 1992; Frayer et al., 1993) suggests that the pattern of genetic relationship between Neandertals and modern humans is tokogenetic rather than cladistic, even if the Neandertal–modern human range of variation exceeds that of living species.

One reason to argue that fossil species may encompass more variation is that they sample temporally dispersed populations (Simpson et al., 1960). Temporal variation may substantially increase the total amount of variation even without directional change, if for example there is oscillating change over time (Vrba, 1980). From Neandertals to recent human populations, a number of directional changes may be expected to have occurred, creating the expectation of greater variation among samples, even if they come from a single evolutionary lineage. Harvati et al. presented a Neandertal sample covering anywhere from 24,000 to 40,000 years, an Upper Paleolithic sample covering approximately 12,000 years, and modern human samples spanning 1,000 years or less, excepting the small ( $n = 8$ ) Iberomaurusian sample. As pointed out by Harvati et al., the comparison of samples covering the shortest time span (36,000–62,000 years), between the Neandertal and Upper Paleolithic samples, yielded the lowest distance, while Neandertal–modern comparisons covering approximately 70,000 years had greater distances. In Fig. 1, the mean  $D^2$  values for each paired comparison that has time depth are plotted against the time spans that the comparisons cover. Mean  $D^2$  increases with the time span of the comparison, and the relationship as plotted is significant ( $p < 0.0001$ ). This result is very tentative, but is consistent with the expectations of evolutionary change among temporally dispersed populations.

How many extant species would make an adequate reference sample for subspecies diversity? The sample of 40 intersubspecific comparisons reported by Harvati et al. seems like a fairly large sample. But all of these comparisons occurred among subspecies within only four species—*H. sapiens*, *Gorilla gorilla*, *Pan troglodytes*, and *Papio hamadryas*. Does this

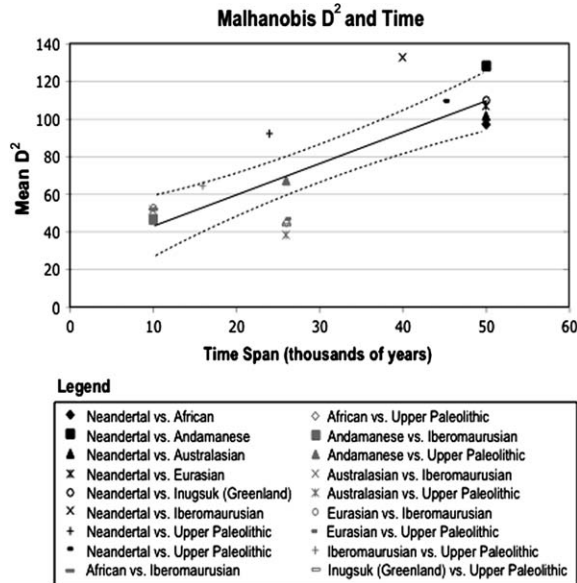


Fig. 1. Mean  $D^2$  values taken from Harvati et al.'s (2004) study for each paired human population comparison that has time depth plotted against the time spans that the comparisons cover. Mean  $D^2$  increases with the time span of the comparison, and the relationship as plotted is significant (least squares linear regression:  $r^2 = 0.6543$ ,  $F = 30.2833$ ,  $df = 1$ ,  $p < 0.0001$ ).

yield a sample size of 40, or of 4? Consider human population structure as an example. Humans have relatively low differences among populations, a reflection of our shared history. Human populations have exchanged many genes recently, a function of large-scale migrations and high rates of dispersal among populations. The human adaptive pattern makes this pattern of dispersal and migration possible, because humans are highly mobile and historically structured in breeding units in which both sexes might disperse across relatively long distances. For gorillas, we might expect quite the opposite pattern, since dispersal is more limited, large-scale migrations are unknown, and the adaptive pattern limits populations to relatively small ranges within forests and forest remnants. Thus, we can expect that the level of differences among subspecies is strongly affected by the adaptive pattern and dispersal rate of the species. Pairwise comparisons within one species are not independent, and furthermore, subspecies differences among close phylogenetic relatives probably

are not independent. Thus, the study involved at most four independent comparisons, of which one (*G. gorilla*) and possibly two (*P. hamadryas*) indicated that Neandertals are not significantly different from humans.

### Where to go from here?

In light of these possible problems, can morphological variation among primate subspecies be a model for the variation among ancient hominids? Historically, hominoid subspecies have been identified based on geography and morphology. More recently, genetic considerations have proven more fruitful than morphology for distinguishing hominoid subspecies, defined as evolutionary lineages. For example, primatologists currently recognize four allopatric chimpanzee subspecies because these geographic groups have been discovered to have deep genetic differences (Gagneux et al., 2001). Likewise, Bornean and Sumatran orangutans exhibit craniometric differences, but these are minor compared to the deep genetic differentiation between them (Zhi et al., 1996). Considering that subspecies are typically defined as evolutionary lineages within a species (Shaffer and McKnight, 1996; Templeton, 1998), some degree of genetic distinctiveness is potentially a good indicator of such a history, even in the absence of strong morphological differences.

But morphological differences may not reflect the degree of genetic relationship, because the two may respond to local selection and partial isolation in different ways. Gorilla subspecies have approximately the same level of genetic differences as other apes (Gagneux et al., 1999; Kaessmann et al., 1999; Jensen-Seaman et al., 2003), but exhibit substantially greater morphological differentiation (Groves, 2001). Indeed, gorilla subspecies differ morphologically more than the two species of *Pan*. Yet, bonobos exhibit significant behavioral and anatomical differences from chimpanzees, despite the fact that the genetic differences between the two species are only slightly greater than between chimpanzee subspecies (Gagneux et al., 1999). Likewise, human populations appear to be

morphologically more different than chimpanzee subspecies, despite the fact that their genetic differences are less than half the chimpanzee levels (Stone et al., 2002). In short, morphological differences among hominoid populations do not reflect either their antiquity or their level of gene flow with other populations.

On the other hand, the genetic similarities across the geographic range of hominoid species can indicate a history of gene flow (Gagneux et al., 2001). Genetic evidence for recent gene flow includes great similarities among some genes, amid a background of strong genetic differentiation among subspecies. From ancient DNA evidence, it would appear that the Neandertal–human difference is smaller than that among subspecies of either chimpanzees or gorillas (Krings et al., 1999; Hawks and Wolpoff, 2001), despite the relatively greater morphological difference. It is currently premature to claim that the known pattern of ancient mtDNA diversity is evidence of gene flow between the ancestors of Neandertals and the ancestors of living people, and indeed such a finding does not preclude the hypothesis that Neandertals speciated after this shared ancestor. But certainly the critical morphological evidence to address whether Neandertals were a distinct species is the pattern of change in both populations over time (Hawks and Wolpoff, 2001). The hypothesis that Neandertals and modern humans were conspecific predicts that they shared a common evolutionary trajectory, regardless of the degree of morphological difference between them. Future multivariate work may pursue the extent to which different characters may provide complementary evidence of such shared trajectories among ancient human populations.

## References

- Ahern, J.C.M., 1998. Underestimating intraspecific variation: the problem with excluding Sts 19 from *Australopithecus africanus*. *Am. J. Phys. Anthropol.* 105, 461–480.
- Ahern, J.C.M., Lee, S.-H., Hawks, J.D., 2002. The late Neandertal supraorbital fossils from Vindija Cave, Croatia: a biased sample? *J. Hum. Evol.* 43, 419–432.
- Frayser, D.W., 1992. The persistence of Neanderthal features in post-Neanderthal Europeans. In: Bräuer, G., Smith, F.H. (Eds.), *Continuity or Replacement: Controversies in Homo Sapiens Evolution*. AA Balkema, Rotterdam, pp. 179–188.
- Frayser, D.W., Wolpoff, M.H., Thorne, A.G., Smith, F.H., Pope, G.G., 1993. Theories of modern human origins: the paleontological test. *Am. Anthropol.* 95, 14–50.
- Gagneux, P., Gonder, M.K., Goldberg, T.L., Morin, P.A., 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.
- Gagneux, P., Wills, C., Gerloff, U., Tautz, D., Morin, P.A., Boesch, C., Fruth, B., Hohmann, G., Ryder, O.A., Woodruff, D.S., 1999. Mitochondrial sequences show diverse evolutionary histories of African hominoids. *Proc. Natl. Acad. Sci. (USA)* 96, 5077–5082.
- Groves, C.P., 2001. *Primate Taxonomy*. Smithsonian Institution Press, Washington, D.C.
- Harvati, K., Frost, S.R., McNulty, K.P., 2004. Neanderthal taxonomy reconsidered: implications of 3D primate models of intra- and interspecific differences. *Proc. Natl. Acad. Sci. (USA)* 101, 1147–1152.
- Hawks, J., Wolpoff, M.H., 2001. Paleoanthropology and the population genetics of ancient genes. *Am. J. Phys. Anthropol.* 114, 269–272.
- Jensen-Seaman, M.I., Deinard, A.S., Kidd, K.K., 2003. Mitochondrial and nuclear DNA estimates of divergence between western and eastern gorillas. In: Taylor, A.B., Goldsmith, M.L. (Eds.), *Gorilla Biology: A Multidisciplinary Perspective*. Cambridge University Press, Cambridge, pp. 247–268.
- Kaessmann, H., Weibe, V., Pääbo, S., 1999. Extensive nuclear DNA sequence diversity among chimpanzees. *Science* 286, 1159–1162.
- Kelley, J., 1993. Taxonomic implications of sexual dimorphism in *Lufengpithecus*. In: Kimbel, W.H., Martin, L.B. (Eds.), *Species, Species Concepts, and Primate Evolution*. Plenum Press, New York, pp. 429–458.
- Kimbel, W.H., Martin, L.B. (Eds.), 1993. *Species, Species Concepts, and Primate Evolution*. Plenum, New York.
- Kimbel, W.H., Rak, Y., 1993. The importance of species taxa in paleoanthropology and an argument for the phylogenetic concept of the species category. In: Kimbel, W.H., Martin, L.B. (Eds.), *Species, Species Concepts, and Primate Evolution*. Plenum Press, New York, pp. 461–484.
- Krings, M., Geisert, H., Schmitz, R.W., Krainitzki, H., Pääbo, S., 1999. DNA sequence of the mitochondrial hypervariable region II from the Neandertal type specimen. *Proc. Natl. Acad. Sci. (USA)* 96, 5581–5585.
- Krishtalka, L., 1993. Anagenetic angst: species boundaries in Eocene primates. In: Kimbel, W.H., Martin, L.B. (Eds.), *Species, Species Concepts, and Primate Evolution*. Plenum Press, New York, pp. 331–344.
- Shaffer, H.B., McKnight, M.L., 1996. The polytypic species revisited: genetic differentiation and molecular phylogenetics of the tiger salamander (*Ambystoma tigrinum*) (Amphibia: Caudata) complex. *Evolution* 50, 417–433.

- Simpson, G.G., Roe, A., Lewontin, R.C., 1960. *Quantitative Zoology*. Harcourt and Brace, New York.
- Stone, A.C., Griffiths, R.C., Zegura, S.L., Hammer, M.F., 2002. High levels of Y-chromosome nucleotide diversity in the genus *Pan*. *Proc. Natl. Acad. Sci. (USA)* 99, 43–48.
- Templeton, A.R., 1998. Human Races: a Genetic and Evolutionary Perspective. *Am. Anthropol.* 100, 632–650.
- Vrba, E., 1980. Evolution, species, and fossils: how does life evolve? *S. Afr. J. Sci.* 76, 61–84.
- Zhi, L., Karesh, W.B., Janczewski, D.N., Frazier-Taylor, H., Sajuthi, D., Gombek, F., Andau, M., Martensen, J.S., O'Brien, S.J., 1996. Genomic differentiation among natural populations of orang-utan (*Pongo pygmaeus*). *Curr. Biol.* 6, 1326–1336.