

With respect to the application of cladistic analysis to the later Pleistocene hominid fossil record, Lieberman is trying (along with others) to impose an inappropriate methodology. Cladistics might work reasonably well for analyses at a supraspecific or, better, a suprageneric level of taxonomy. However, it is very uncertain whether it can be justifiably employed to sort relationships between samples which, if not conspecific, certainly derive from very closely related species. Most of the conceptual and methodological difficulties in so doing have been discussed previously (e.g., Trinkaus 1990, 1992). However, two points need to be emphasized.

First, evolution operates in terms of shifting patterns of variation, and cladistic methodology does not allow one to deal with this effectively. Indeed, the methodological convolutions employed by Lieberman for traits which differ between samples in their frequencies rather than in presence/absence only illustrate the inappropriateness of this approach to normal distributional data. The multivariate techniques employed in phenetic studies have the potential to deal effectively with this problem; cladistics does not. The same applies to assessments of the intercorrelation of traits.

Second, cladistic analyses require a determination of trait polarity, usually by determining the most parsimonious cladogram available (if one such "most parsimonious" cladogram does indeed exist) or by identifying an outgroup (as in the early *H. erectus* employed here). However, among closely related species, high levels of homoplasy (reversals, parallelisms, etc.) are common (see McHenry 1994 regarding high levels of homoplasy in a closely related hominid species group, one for which cladistic analyses have become *de rigueur*), making such assessment of polarity largely meaningless.

Finally, Lieberman, along with many others, has focused these considerations on only two extreme phylogenetic models of modern human origins, the "replacement" and "regional continuity" ones. However, even a casual perusal of the recent scientific literature by those paleontologists who actually study the later Pleistocene hominid fossil record will reveal a spectrum of phylogenetic scenarios ranging between these extremes. This spectrum of scenarios exhibits variation along several axes (Trinkaus and Shipman 1993). Most important, it encompasses considerable geographical variation in the phylogenetic processes (including replacement, continuity, and gene flow) inferred to have taken place, something which the simple "replacement" vs. "regional continuity" dichotomy cannot accommodate.

Consequently, Lieberman has made a significant contribution in emphasizing the need to assess carefully the biological bases of the morphological traits employed in paleontological phylogenetic considerations of modern human origins. However, he partly negates the value of this by then emphasizing a methodology which ignores the fact that the traits in question exist in biological organisms which made up portions of biological populations and lineages operating dynamically under rules of biological evolution through space and time.

MILFORD H. WOLPOFF AND
TRACEY L. CRUMMETT

*Department of Anthropology, University of Michigan,
Ann Arbor, Mich. 48109-1382, U.S.A. 10 x 94*

We appreciate the opportunity to comment on this paper, as it discusses some of our work. Its stated goal is "not to produce a 'best' tree but to propose some improvements in the way in which models of human origins are tested." However, in figure 4 Lieberman does produce his "best tree" ("two most parsimonious cladograms"), and he falls short of his other goal as well. We focus on three reasons why.

First, Lieberman says that "cladistic theory . . . is often misunderstood or misused." We concur. For example, he contends that we can infer evolutionary relationships among fossil and modern fossil human populations only from shared-derived characters. Cladistic analysis creates hypotheses about the genealogy of species-entities, but can it validly work out the genealogy of populations, races, or subspecies? The problem is that these are not distinct lineages with beginnings and ends but rather indistinct groups that can exchange genes and continuously merge and then separate again. Character states can only be validly ascertained at the species level, because all characteristics that vary systematically between human populations also vary within them (Lewontin 1984). Determining character states of features that vary within a species is not valid because, as Wiley (1981) has pointed out, variable features should not be used in phylogenetic analyses. The valid determination of character states for fossil and living humans requires the assumption that different geographic samples of earlier humans had separate evolutionary histories, which may be true under the recent-African-origin theory but is certainly not the case for multiregional evolution, whereby populations were interconnected across regions through time. It is not that transformations do not occur within species but rather that sharing them may not mean recent common ancestry; it could denote significant or persistent genic exchanges. Determining character states and using them to assess relationships therefore requires *assuming* that multiregional evolution is wrong because there were several contemporary Middle and Late Pleistocene human species.

Testing cladistic hypotheses also requires the assumption that homology is a far more likely cause of synapomorphies than homoplasy. This assumption is not obviously correct within species because of the marked similarity of gene pools (similar gene pools have a good chance of responding to similar sources of selection in the same ways) that maximizes the potential for homoplasies and because genic exchange is a potential source of shared features that would reflect not population *divergence* but population *contacts*. Two sets of populations that diverged at the same time may be quite different in number of similarities if one set had many genic exchanges and the other had few. The situation is even

TABLE I
Reality Check

| Lieberman's Assertion | Reality |
|---|--|
| The WLH 50 frontal is artificially deformed, according to Brown (1981). | Examination reveals no evidence of deformation, and Brown (1981) did not include WLH 50 in his list of deformed specimens. |
| The suprainiac fossa "may just be an additional area for the attachment of the upper neck (nuchal) muscles." | Dissection or a review of the literature would quickly dispel this claim. |
| Irhoud 1 and 2 are the only transitional fossils that have the globular braincase of modern humans. | Apart from the fact that many modern humans do not have this globular braincase (are they not modern?) and non-African transitional specimens such as Jinniushan do, the Irhoud specimens' parietal curvature in the sagittal and coronal planes is within the Neanderthal range. |
| "Sangiran 17, the only preserved archaic Australasian face, is heavily reconstructed." | The Sangiran 17 face is separated from the vault, and reconstructions differ on how to attach it, but there are good contacts between most of the facial elements and it has never been taken apart with the goal of reconstructing it differently from its present form. It is true, however, that the specimen is heavy. |
| "A high degree of facial prognathism . . . characterizes all archaic <i>Homo</i> and <i>H. erectus</i> fossils." | This is impossible to reconcile with the facial anatomy of archaic East Asian specimens such as Dali (Wu and Wu 1985) and Jinniushan (Lü 1991). |
| "Most (87%) [of the characters studied] clearly do not support polycentric evolution." | This 87% is not the number of features that <i>disprove</i> the MR hypothesis but the number that he did not use in his analysis because they did not meet his criteria; 87% of the characters also do not support the RA theory. |
| "The MR hypothesis must invoke high levels of gene flow among all regions of the world in all directions throughout the Late Pleistocene, which seems unlikely." | Templeton, who supports the multiregional model, writes (1993:69) that "mtDNA analysis strongly and clearly indicates that these Old World human populations had low but recurrent levels of genetic contact with one another for a substantial portion and perhaps all of this one-million-year time period." |
| "the lack of identifiable specifically African regional features . . . is one of the best sources of evidence in support of" the RA hypothesis. | The RA hypothesis specifically predicts that modern humans first attained African regional features and then spread around the world (Stringer and Andrews 1988, Bräuer 1989). ^a In this piece of reverse logic, a refutation of the theory is turned into a pillar of its support. |
| Table 3 shows Neandertals as having a wide cranial base (feature 25), while in the discussion of feature 2 they are said to have a narrow cranial base. | Neandertal cranial bases are broad. |
| In his discussion of homology, Lieberman writes that "a rounded forehead is not a real shared-derived character." However, in his discussion of features 6 and 7, he writes that "high rounded frontals . . . support the RA hypothesis." | Lieberman asserts that hypotheses of relationship can only be supported by shared-derived features. |

^a According to Bräuer (p.139), "the ancestors of the inhabitants of Europe and Western Asia of some 30,000 years B.P. consisted of modern Africans with some admixture of Neandertals."

more complicated if one of the populations exchanged genes with yet another population.

Lieberman further asserts that "the argument that regional human clades are best characterized by *combinations* of derived and primitive characters rather than by any *specific* derived characters is illogical." Apart from the fact that no one has made such an argument (Frayer et al. do not identify character states, for the reasons stated above), this allegation ignores decades of forensic literature providing combinations of features (not single features) for the "racial" identification of human remains. What Lieberman does is attempt to define regional continuity out of existence by offering criteria for identifying geographic variation so stringent that they cannot distinguish living human groups. First he calls the prediction of "a set of unique regional features" linking ancient and modern humans in different parts of the world "the most important prediction of the multiregional evolution hypothesis." In fact, it is not a set of unique features but a unique *set* of features that multiregional evolution addresses, and this is not a *prediction* but *the observation on which it is based*. But he does not test this "important prediction." Instead, he fabricates the criterion that regional features must be *unique* and *universal* to be valid, and failing to find such features he dismisses regional continuity. For example, he discusses the absence of a supratatorial sulcus, part of the character complex that helps distinguish continuity in Australasia, and dismisses it as a valid regional feature because some African specimens also lack a sulcus. Lieberman uses *his* criterion to dismiss one regional feature after another. The surprise is that he finds any at all.

Secondly, there are many instances of misunderstanding or misuse of data (see our table 1). We discuss one that we find especially disturbing; the absence of further discussion should not be taken to mean that we agree with Lieberman's assessments. Lieberman provides two sets of data for shoveling, one which treats the morphology as present or absent (table 3) and another which separates the shape into several components (table 5). Both are misleading. His presence/absence treatment violates his first criterion, homology, as it lumps together varying sets of specimens with a variety of different shapes and varying shape frequencies. By distinguishing different morphologies he avoids one homology problem but creates another by compiling incomparable data. "Slight," "moderate," and "prominent," the categories he uses, are subjective terms for discriminating these shapes. Comparisons of samples defined by these terms as used by different researchers mean nothing (Crummett 1994) because each defines the categories differently. For instance, relying on two different observers Lieberman compares Brown and Walker's (1993) description of WT 15000 with Crummett's (n.d.) description of Chinese *Homo erectus* to show that these morphologies were the same in order to dismiss shovel shaping as a support of multiregional evolution. Yet, had he reported the result of the same

comparison from the observations of a single researcher (for instance, Crummett's [1994] descriptions of *both* WT 15000 and Chinese *H. erectus*, which were provided to him), he would have shown clear *differences* between these incisors and the presence of a regional pattern in Asia.

Thirdly, as the geneticist Templeton (1994) has recently noted, discussion on the mitochondrial DNA side of the modern-human-origins debate has focused on *hypothesis compatibility* rather than *hypothesis testing*. This is increasingly characteristic of anatomical studies as well (Waddle 1994, Lahr 1994). With the Lieberman paper added to these, we believe there have been far too many instances of data's being described as "compatible with" or "supportive of" various hypotheses; in fact, all three of these anatomically oriented studies fail to draw a conclusion, a not-unexpected consequence of trying to determine which hypothesis is better supported instead of determining which is wrong. Valid testing of hypotheses, not analyses of their compatibility, is what finally counts. Refutation is the accepted mode of examining the validity of hypotheses, and this is especially true of the cladistic procedures that Lieberman purports to use, which were developed to provide a way of applying the refutatory approach to the historical field of paleontology.

Lieberman's own data, in fact, provide a clear refutation of the recent-African-origin (and Eve) theories. Recent African origin predicts, as Lieberman says, that anatomically modern humans "share derived regional features with no more than one population of archaic *Homo*, probably in Africa." However, using these data, it is possible to refute this prediction in two regions, where characters appear in both the archaic and modern populations but are not found in the African archaic sample. According to Lieberman's data both facial size (character 14) and malar orientation (character 18) provide refutatory evidence in East Asia, while nasal size (character 16) and nasal-subnasal margin (character 23) provide refutatory evidence in Australasia.

Reply

DANIEL E. LIEBERMAN
Cambridge, Mass., U.S.A. 21 XI 94

The variety of constructive and generally positive responses to this article demonstrates that the debate over modern human origins is still dynamic and far from resolution, but it also illustrates the challenges that we face in integrating biology into palaeoanthropology. Many anthropologists still do not agree on what methods to use to analyze the same fossil skulls. I will, therefore, first address the issues raised about the methods we should use to interpret craniodental data. Many of the respondents also provide useful comments on spe-