

Modern Human Ancestry at the Peripheries: A Test of the Replacement Theory

Milford H. Wolpoff,^{1*} John Hawks,² David W. Frayer,³ Keith Hunley¹

The replacement theory of modern human origins stipulates that populations outside of Africa were replaced by a new African species of modern humans. Here we test the replacement theory in two peripheral areas far from Africa by examining the ancestry of early modern Australians and Central Europeans. Analysis of pairwise differences was used to determine if dual ancestry in local archaic populations and earlier modern populations from the Levant and/or Africa could be rejected. The data imply that both have a dual ancestry. The diversity of recent humans cannot result exclusively from a single Late Pleistocene dispersal.

Two conflicting evolutionary models of modern (*Homo*) human origins have emerged in the past decade (2): complete replacement, in which modern humans are a new species that replaced all archaic populations, and multiregional evolution, in which modern humans are the present manifestation of an older worldwide species with populations connected by gene flow and the exchange of ideas, resembling each other because of similar adaptations to ideas and technologies that spread across the inhabited world and because of the dispersals of successful genes promoted by selection. One place where the models have quite different predictions is at the peripheries of the human range, where the replacement theory stipulates that the ancestry of early modern populations is not among their local archaic predecessors, who often are regarded as different human species (3). Multiregional evolution, a reticulate theory, is compatible with a dual ancestry for early moderns from the peripheries. Here we examine ancestry in Australia and Central Europe, two peripheral areas with the best fossil record for the so-called transitional period, considering an early modern Australian cranium, Willandra Lakes Hominid (WLH) 50, and the two adult male crania from Mladeč, in Moravia, Czech Republic.

WLH-50 has been dated to about 15,000 to 13,000 years by gamma spectrometric U-series analysis (4) and to about double that time by electron spin resonance on bone (5). The calvarium exhibits many features that

closely resemble earlier Indonesian hominids, including the Ngandong fossils of Java (6). WLH-50 is regarded as a modern human in all origins models (6, 7), whereas Ngandong fossils are often assessed as archaic humans or late surviving *Homo erectus* (8). The other possible ancestral population we tested for this male is the earlier modern human sample from Africa and the Levant of western Asia, which under the replacement theory must represent the only ancestors (Fig. 1) of the first modern humans in Australia.

In Central Europe, over 100 fragmentary specimens unequivocally associated with an early Aurignacian industry (9) were recovered from the two Mladeč Caves in Moravia, Czech Republic (10). The most complete male calvaria are from the Quarry Cave. Mladeč 5 and 6 have been likened to the temporally earlier Neandertals (11–13) [although it is recognized that they are not themselves Neandertals (10, 11)]. Neandertals, then, are the local archaic predecessors of the Mladeč folk and are potentially their ancestors. The other potential ancestors are from Qafzeh and Skhul, the earlier modern human remains in western Asia (Fig. 2), geographically and temporally the closest source of a replacing population.

The complete replacement model requires a unique relation between the early modern humans at these two peripheries and the earlier Levantines and/or Africans regarded as modern human. If evidence shows significant local ancestry for the peripheral samples, complete replacement must be wrong. To examine this issue, we tested for refutations of the hypotheses of dual ancestry (14) for WLH-50 and the Mladeč males. Obtaining a valid statistical test requires overcoming some formidable obstacles. There are only a few early modern crania complete enough to analyze, their dates are uncertain, the comparative samples are small and haphazardly preserved (15), and the traits may depend on

15. C. H. Stirling, D.-C. Lee, J. N. Christensen, A. N. Halliday, *Geochim. Cosmochim. Acta* **64**, 3737 (2000).
16. Supplementary material is available at www.sciencemag.org/cgi/content/full/291/5502/290/DC1.
17. G. Paulay, *Henderson Island: Biogeography and Evolution at the Edge of the Pacific Plate*, E. C. Dudley, Ed., *The Unity of Evolutionary Biology*, Proceedings of 4th International Congress of Systematics and Evolutionary Biology (1991), pp. 304–313.
18. S. G. Blake, *Bio. J. Linn. Soc.* **56**, 43 (1995).
19. J. M. Pandolfi, *Bio. J. Linn. Soc.* **56**, 63 (1995).
20. Following (24), MIS 9 samples with the lowest $\delta^{234}\text{U}(T)$ are considered to have undergone the least postdepositional exchange of U and Th. It follows that corals with $\delta^{234}\text{U}(T)$ values coincident with $149 \pm 1\%$ have accurate ^{230}Th -ages [Web fig. 1 (16)]. The $\delta^{234}\text{U}(T)$ changes by $\sim 2.5\%$ per 1000 years as a result of diagenesis, and a cumulative $\delta^{234}\text{U}(T)$ shift of 8‰ should shift the ^{230}Th -ages beyond their ± 2 - to 3-ka age uncertainties. Therefore, we reject samples with $\delta^{234}\text{U}(T)$ lying outside the range of $149 \pm 8\%$. Despite placing strict limits on acceptable $\delta^{234}\text{U}(T)$, it is recognized that this criterion, although essential, is not necessarily a sufficient condition to guarantee reliability. Given their relatively large age and $\delta^{234}\text{U}(T)$ uncertainties, all MIS 15 samples are considered reliable.
21. The remoteness of the island has reduced the input of detrital contamination, and the effect of foreign ^{230}Th , ^{234}U , and ^{238}U on the ^{230}Th -ages has been negligible. With only two exceptions (HEN 2-7 and FH 175), all samples display detrital ^{232}Th concentrations within the 0.5 part per billion limit specified by modern pristine corals from oceanic islands (2, 4). Further, relatively low rainfall [~ 1300 to 1500 mm/year (26)], low humidity, and relatively cool, stable subtropical air temperatures (15° to 30°C) appear to have minimized postdepositional mobilization of the U-series isotopes.
22. G. M. Henderson *et al.*, *Earth Planet. Sci. Lett.* **115**, 65 (1993).
23. B. Hamelin *et al.*, *Earth Planet. Sci. Lett.* **106**, 169 (1991).
24. C. D. Gallup, R. L. Edwards, R. G. Johnson, *Science* **263**, 796 (1994).
25. R. Grün, F. McDermott, *Quat. Sci. Rev. (Quat. Geochronol.)* **13**, 121 (1994).
26. G. T. Shen, R. B. Dunbar, *Geochim. Cosmochim. Acta* **59**, 2009 (1995).
27. M. Brooke, *Pitcairn Island Expedition Report* (Cambridge Univ. Press, Cambridge, 1992).
28. A. Berger, M. F. Loutre, *Quat. Sci. Rev.* **10**, 297 (1991).
29. J. Galewsky, E. A. Silver, C. D. Gallup, R. L. Edwards, D. C. Potts, *Geology* **24**, 819 (1996).
30. K. R. Ludwig *et al.*, *Science* **258**, 284 (1992).
31. Because Henderson Island is a remote outpost locality, it is possible that a small lag exists between the onset of the sea-level high-stand and the start of reef development due to coral dispersal and recolonization limitations. However, any delay should be minimized for the intense MIS 9 interglacial, and in the absence of information to the contrary, we assume that (i) the “330 ka” reefs started to colonize Henderson Island shortly (< 1 ka) after sea level stabilized at the start of the interglacial period, and (ii) reef growth ceased at the termination of the interval when sea levels began to fall to glacial values. Even allowing for small shifts in the ^{230}Th -ages due to diagenetic alteration and/or delayed reef colonization, these data still shortly postdate the insolation maximum, and will still be concordant with Northern Hemisphere Milankovitch forcing of climate.
32. J. F. McManus, D. W. Oppo, J. L. Cullen, *Science* **283**, 971 (2000).
33. We thank two anonymous reviewers for constructive comments that helped improved the manuscript and M. Johnson, C. Hall, G. Mortimer, and L. Kinsley for technical assistance. We are grateful to J. Pandolfi for providing access to samples. This research was primarily supported by the U. S. Department of Energy and the NSF and Swiss NF grants.

13 September 2000; accepted 29 November 2000

¹Department of Anthropology, University of Michigan, Ann Arbor, MI 48109–1382, USA. ²Department of Anthropology, University of Utah, Salt Lake City, UT 84112–0060, USA. ³Department of Anthropology, University of Kansas, Lawrence, KS 66045–7556, USA.

*To whom correspondence should be addressed at Department of Anthropology, University of Michigan, 500 South State, Ann Arbor, MI 49109–1382, USA. E-mail: wolpoff@umich.edu

REPORTS

each other in ways we cannot observe. In order not to compound these problems with extraneous assumptions, we used a simple method to address the relation among individuals that is independent of group assignments and the statistical structure of the groups. We collected details on certain nonmetric traits, scored as presence or absence to make unweighted sums of differences potentially informative (16). Using the scores on these nonmetric traits, we calculated the pairwise differences between specimens in the two early modern samples from the peripheries and each of the specimens in the earlier samples of potential ancestors. This kind of pairwise difference analysis is commonly applied to DNA sequence data to derive information about past population demography (17) and to estimate closeness of relation (18). Its use for nonmetric traits is tolerant of missing data—because we treat such cases as the absence of differences—as long as the missing data are either randomly distributed or are more strongly represented in the African or Levantine sample of potential ancestors predicted by the replacement theory. Additionally, we minimized missing data by appropriate choices of specimens and traits. The chance of a type I error is high with the pairwise test only if an early modern specimen is more similar in size to the local archaic sample and if the traits that are correlated with cranial size do not reflect geographic differences. We can reject both these conditions, as discussed below. If the early modern specimen is exclusively related to the earlier Africans and/or Levantines, then the probability of it looking more like the archaic sample should be effectively zero.

It may be claimed that because our analysis is a phenetic procedure, it is inappropriate to test the relations of these early modern humans with various putative ancestors because we do not take the polarity of the character states into account. We chose this procedure because we were unwilling to assume that the potential ancestors were different species; indeed, to do so would be to assume our conclusions. Character polarity is meaningless if the evolution of these groups included gene flow, which is one of the hypotheses under consideration. Moreover, although cladistic procedures might provide an implicit test of the replacementist assumption that archaic and modern peoples represent different species, the broad distribution of most of the nonmetric traits across groups, with very few traits unique to any group, leaves only a small number of potential synapomorphies required for phylogenetic analysis. The problem can be illustrated in the WLH-50 analysis. There are only eight traits unique to one potentially ancestral group or another in our nonmetric data. Of these eight character states that are uniquely in one sample, five are polymorphic within their samples, and only three are shared with WLH-50 (all three are shared between WLH-50 and Ngandong). The degree to which these nonmetric traits are shared among groups may itself suggest the conspecificity of the specimens involved, but we question the validity of this approach, which is why we did not take it.

For the WLH-50 analysis, we considered all of the specimens in the two earlier samples, regardless of sex, because we cannot reliably determine the sexes of the Ngandong

calvaria. However, in the Mladeč analysis we restricted our comparisons to the males. We believe that our identification of males in Mladeč and in the earlier samples of potential ancestors is reliable, and note that females are poorly represented in the Skhul/Qafzeh cranial sample. The comparative samples are detailed in Tables 1 to 3.

Most nonmetric observations were taken on the original fossil crania by at least one and often two of us; observations on casts of Laetoli Hominid 18 and the Omo crania were also used (see legends to Tables 1 and 2). In our scoring system, the presence of a trait was scored as 1, and its absence as 0. Some of these traits may reflect robustness, and for all such traits the more robust condition was scored as 1. In cases where a trait could not be observed on a specimen, we scored the trait as missing for that individual and treated it as not different in all comparisons involving that individual. We could unambiguously score 16 nonmetric traits on WLH-50. We avoided duplicating features that seemed to reflect the consequences of the same anatomical variation. These were scored in the Levant sample of four specimens and in six Late Pleistocene Africans, and were present in most cases. The pattern of missing data is random for the WLH-50 comparisons (Kruskal-Wallis test, $\chi^2 = 0.126$, $P = 0.939$). In the Mladeč analysis, we examined 30 nonmetric traits from all parts of the Mladeč 5 cranium (22 could be scored on the less complete Mladeč 6). These nonmetric traits were scored in four Neandertal and five Skhul/Qafzeh males and were present in virtually all instances. In this case, the pattern of missing data in the comparative samples is

Fig. 1. Ngandong 1 (left) compared with WLH-50 (center, cast) and Qafzeh 9 (right, cast), shown in lateral view to the same scale. In its simplest form, the issue we address is which two are alike and which one is different. WLH-50 is the youngest of the three.



Fig. 2. Mladeč 5 (center) compared with Qafzeh 9 (right, cast) and Spy 2 (left), in lateral view. As in Fig. 1, we ask which two are alike and which one is different. The three specimens shown here are males, and Mladeč is the youngest of the three.



REPORTS

not random. The Skhul/Qafzeh crania have more missing data than the Neandertals. This makes the analysis conservative because if all else is equal, the results will be biased to show more similarities with the Skhul/Qafzeh remains (i.e., will support replacement).

We examined whether there might be special similarities between the peripheral samples and their archaic predecessors because of robustness in the peripheral specimens created by their size, an explanation some have suggested (19, 20). To specify the relation, we examined each region for an influence of size on the robustness, as estimated by the sum of the nonmetric scores (20). We used endocranial capacity to estimate cranial size because it is not based on any particular observation in our analysis (21). The question of how robustness and size are related has different answers at different levels of analysis. For instance, within populations there is usually a clear relation between size and robustness because males are on average and in particular larger and more robust than fe-

males. There were no significant within-group associations of cranial size and presence of nonmetric traits in our samples (Africans: $P = 0.527$; Levantines: $P = 0.901$; Ngandong: $P = 0.580$; Mann-Whitney test). Between populations, on the other hand, the relation could be just the opposite. Comparing the African and Levantine samples, there is a positive association, indicating that any relation between cranial size and presence of nonmetric features must depend strongly on the geographic locations of the groups being compared. In the Australasian analysis, the extraordinary robustness of WLH-50 resembles one potential ancestor, and its size resembles the other. Of the 16 nonmetric variables used in the WLH-50 analysis, 6 were correlated with cranial capacity ($P = 0.05$, Kolmogorov-Smirnov test). In each instance where the presence of a nonmetric feature was correlated with cranial capacity, the specimens that possessed the feature had smaller cranial capacities than those lacking the feature. These relations exist because Ngandong is the most robust group in our sample (Fig. 3) and has the smallest mean cranial capacity. The large size of WLH-50 ensures that if there is a significant nonmetric relation of WLH-50 with Ngandong, a rejec-

tion of the replacement hypothesis will be a conservative outcome.

A similar but weaker relation can be found in the Mladeč comparisons where the population with larger crania is not the more robust one. The Neandertal males ($n = 8$) are more robust than the Skhul/Qafzeh males ($n = 5$), but their mean cranial capacity is smaller (1531 cm^3 compared with 1552 cm^3). It follows that the larger size of the early modern descendants cannot be posited as a cause of their robustness, so nonmetric similarities between earlier and later specimens within regions cannot have resulted from cranial size. The most parsimonious explanation for such similarities is ancestry.

We found significant differences among the groups for pairwise differences from WLH-50 (Table 1). Six of the seven Ngandong crania are closer to WLH-50 than are any other specimens, and the seventh is separated from the others by only one individual (Skhul 9). On average, WLH-50 possesses fewer differences from the Ngandong group (3.7 pairwise differences) than from either the African (9.3) or Levantine (7.3) groups. Mean pairwise differences between Ngandong and African, and Ngandong and Levantine groups are statistically significant at the 0.05 level (Mann-Whitney). The Ngandong mean pairwise difference from WLH-50 is significantly lower than both the African mean and the Levantine mean differences at the 0.05 level (Mann-Whitney). WLH-50 is most different from the Africans, although difference in the number of pairwise differences between the African and Levantine samples relative to WLH-50 is not significant. In sum, in its nonmetric traits WLH-50

Table 1. Pairwise comparison of WLH-50 to Indonesian, African, and Levantine specimens. The mean difference with the Ngandong specimens is 3.7; with the Levantine sample, 7.3; and with the Africans, 9.3. Unlike the Mladeč analyses (Tables 2 and 3), this analysis does not include Skhul 4 because it possesses only a few of the nonmetric traits listed below, mostly because of the parts preserved and the poor condition of the surface of the cranium. The features used in the Mladeč comparisons could be observed much more often. The 16 nonmetric features used in this pairwise comparison are as follows: angular torus; coronal keel; sagittal keel on frontal; lateral frontal trigone; *linea obliquus* strongly developed; mastoid crest; sagittal keel on parietal; postlambdoidal eminence; prebregmatic eminence; projectinginion; sulcus dividing the medial and lateral elements of the supraorbital torus or superciliary arches; superior margin of the orbit blunt (as opposed to sharp); suprainiac fossa; supramastoid crest; temporal line forms a ridge; and transversely extensive nuchal torus.

Specimen	Number of differences from WLH 50
Ngandong 5	2
Ngandong 9	3
Ngandong 10	3
Ngandong 11	3
Ngandong 4	4
Ngandong 1	5
Skhul 9	5
Ngandong 6	6
Jebel Irhoud 2	6
Omo 1	6
Qafzeh 6	7
Qafzeh 9	7
Skhul 5	7
Omo 2	8
Jebel Irhoud 1	10
Laetoli 18	10
Singa	12

Table 2. Pairwise differences between Mladeč 5 and the most complete Neandertal males (mean difference = 14.8) and Skhul/Qafzeh males (mean difference = 14.0). Presence or absence of 30 nonmetric traits is scored in this analysis: anterior temporal fossa border angled; thick parietal at asterion (>9 mm); broad frontal (>125 mm); broad occiput (>120 mm); central frontal boss; cranial rear rounded (as seen from back); external auditory meatus leans forward; frontal keel; long frontal (glabella-bregma length > 113); frontonasal suture arched; glabellar depression; glenoid articular surface flattened; thick occipital at lambda (>8 mm); lateral supraorbital central thinning; mastoid process projects minimally; mastoid tubercle; mastoid-supramastoid crests well separated; medially tall supraorbital (>19 mm); nuchal torus extends across occiput; occipital bun; long occipital plane (>60 mm); occipitomastoid crest prominent; paramastoid crest prominent; retromastoid process prominent; sagittal groove along vault posterior; supraglenoid gutter long; elliptical suprainiac fossa; supraorbital center dips downward; "teardrop" shape (seen from top); and vertical occipital face short.

Specimen	Number of differences from Mladeč 5
Skhul 4	8
Qafzeh 6	10
Skhul 9	11
Spy 2	12
La Chapelle	13
Qafzeh 9	15
La Ferrassie	15
Guattari	15
Skhul 5	23

Table 3. Pairwise differences between Mladeč 6 and the most complete Neandertal males (mean difference = 7.8) and Skhul/Qafzeh males (mean difference = 11.6). Presence or absence of 22 nonmetric traits is scored in this analysis, less than the number for Mladeč 5 because the vault is less complete. The eight traits from Table 2 not found on this specimen are as follows: thick occipital at lambda; prominent occipitomastoid crest; minimal projection of mastoid process; mastoid tubercle; long supraglenoid gutter; central thinning of lateral supraorbital; angled anterior temporal fossa border; and external auditory meatus leans forward.

Specimen	Number of differences from Mladeč 6
Qafzeh 6	6
La Chapelle	7
Guattari	7
Spy 2	8
La Ferrassie	9
Skhul 4	11
Skhul 9	11
Qafzeh 9	14
Skhul 5	16

is closer to the specimens from Ngandong than to any other group. If Ngandong was the expected ancestor under the replacement theory, this would disprove a dual-ancestry hypothesis. However, it is not, and a Ngandong ancestry disproves the replacement theory, so the conservative interpretation of these results is that the dual-ancestry hypothesis cannot be disproved.

We might expect a similar pattern for the Mladeč specimens. Examination of characteristics said to be unique for Neandertals (22) indicates that Mladeč 6 has a suprainiac fossa of elliptical form, extensive lambdoidal flattening, and a short posterior face on the occiput. Mladeč 5 has a Neandertal-like sagittal contour, a well-developed occipitomas-toid crest, minimal mastoid projection, and evidence of midfacial prognathism, insofar as the marked anterior projection of the upper face predicts this.

The average pairwise difference between Mladeč 5 and the Neandertal sample is 14.8, and between Mladeč 5 and the Skhul/Qafzeh sample, 14.0—virtually the same (Table 2). On the other hand, for Mladeč 6 the corresponding comparisons are 7.8 and 11.6, so it is closer to the Neandertal sample (Table 3). A Sample Runs Test (23) was used to examine whether the ordering of Neandertal and Skhul/Qafzeh crania, based on the number of pairwise differences from the Mladeč crania, is random (the null hypothesis). Randomness can be rejected at the $P = 0.05$ level when there are two or less, or 9 or more, runs (24) from the same site, for a sample of this size. There are five runs for Mladeč 5 and three runs for Mladeč 6—randomness in the order of pairwise similarities cannot be rejected (25). Again, these data fail to reject the dual-ancestry hypothesis. They imply that the Mladeč crania are not uniquely related to Skhul/Qafzeh.

Another way to examine this nonmetric variation is in terms of its distribution within and between the samples of putative Mladeč ancestors. Three of the nonmetric variables completely separated the Neandertal and Skhul/Qafzeh samples. Of these, the Mladeč crania were like the Neandertals in two, and

like Skhul/Qafzeh in one. Seven additional traits almost completely separated the putative Mladeč ancestors (26). Of these, the Mladeč crania were like the Neandertals in four, and like Skhul/Qafzeh in two. For the seventh trait, one Mladeč cranium was like each comparative sample. In spite of the predominance of Neandertal resemblances for this subset of 10 traits, the normal approximation of the binomial distribution shows that an ancestry hypothesis in which dual ancestry is assumed to be equal ancestry cannot be rejected at the 0.05 level.

These tests show that a dual-ancestry model cannot be rejected for either of the geographically peripheral samples we have considered. There is no evidence suggesting that WLH-50 can be grouped with either Late Pleistocene Africans or Levantines to the exclusion of the Ngandong sample, or that the Mladeč males can be grouped with the Levantines to the exclusion of European Neandertals. This refutes the replacement theory and contradicts the interpretation that local archaic populations represent now-extinct human species (27). It means that Ngandong should not be classified as *H. erectus* (6). This result joins other evidence of sufficient Neandertal features in Upper Paleolithic Europeans to reject interpreting their variation as that of an extinct human species (11, 13, 28–30), because no matter how different Neandertals may seem, diagnostically Neandertal anatomy in later populations is an indication of sufficient Neandertal ancestry to reject such a species hypothesis.

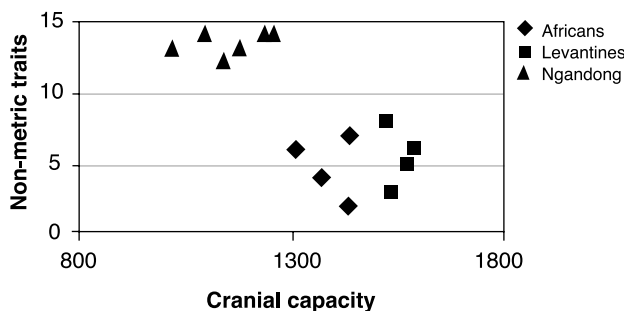
Our analyses, like all discussions of modern human origins, are limited by the small sample sizes and ignorance of the underlying variance and covariance matrices for the data, and we interpret them with this in mind. However, their failure to disprove the hypothesis of dual ancestry for early moderns in two far-flung regions of the world adds to the improbability of a replacement explanation for modern human origins. We do not doubt that many prehistoric groups were replaced by others, but we conclude that the hypothesis that all living humans descended from a single geographically isolated group during

the Late Pleistocene is false, and that the replacement explanation for the origin of these early modern Australians and Europeans can be ruled out.

References and Notes

1. We use this word guardedly; we take "modern" to refer to all living humans and their immediate ancestors, but whether modernity originated at one time and place as a single entity, or whether it reflects the worldwide distribution of shared anatomies and behaviors, is the question examined in this paper.
2. M. H. Wolpoff, R. Caspari, *Race and Human Evolution* (Simon & Schuster, New York, 1997); C. B. Stringer, in *Origins of Anatomically Modern Humans*, M. H. Nitecki, D. V. Nitecki, Eds. (Plenum, New York, 1994), pp. 149–172; G. A. Clark, C. M. Willermet, Eds., *Conceptual Issues in Modern Human Origins Research* (Aldine de Gruyter, New York, 1997).
3. Although Ngandong has been considered *Homo erectus* (8) and Neandertals classified as *Homo neanderthalensis* (27), not all replacement theorists agree that there is a taxonomic difference at the species level. Some have suggested that Neandertals could have interbred with Upper Paleolithic Europeans but did not do so very often [G. Bräuer, in *Continuity or Replacement? Controversies in Homo sapiens Evolution*, G. Bräuer, F. H. Smith, Eds. (Balkema, Rotterdam, 1992), pp. 83–98; C. B. Stringer, *Nature* **356**, 201 (1992)]. However, others note that this is actually a variant of multiregional evolution [J. H. Relethford, *Evol. Anthropol.* **8**, 7 (1999); M. H. Wolpoff, J. Hawks, R. Caspari, *Am. J. Phys. Anthropol.*, **112**, 29 (2000)]. This is because models including population mixing are reticular, and enough mixing to be recognized in the fossil record is a disproof of replacement, not an example of it [A. R. Templeton, in *Conceptual Issues in Modern Human Origins Research*, G. A. Clark, C. M. Willermet, Eds. (Aldine de Gruyter, New York, 1997), pp. 329–360, 437–492].
4. J. J. Simpson, R. Grün, *Quat. Geochr.* **17**, 1009 (1998).
5. D. Caddie, D. Hunter, P. Pomery, H. Hall, in *Archaeometry: Further Australasian Studies*, W. Ambrose, J. Mummary, Eds. (Department of Prehistory, Australian National University, Canberra, 1987), pp. 156–166.
6. J. Hawks et al., *J. Hum. Evol.* **39**, 1 (2000).
7. C. Stringer, *J. Hum. Evol.* **34**, 327 (1998).
8. G. P. Rightmire, *The Evolution of Homo erectus: Comparative Anatomical Studies of an Extinct Human Species* (Cambridge Univ. Press, Cambridge, 1990); C. C. Swisher III et al., *Science* **274**, 1870 (1996).
9. J. Svoboda, *J. Hum. Evol.* **38**, 523 (2000).
10. J. Szobath, *Eiszeit* **2**, 1, 73 (1925); F. H. Smith, in *History of Physical Anthropology. An Encyclopedia*, F. Spencer Ed. (Garland, New York, 1997), pp. 659–660.
11. J. Jelínek, *Anthropologie (Brno)* **21**, 57 (1983).
12. D. W. Frayer, *Anthropologie (Brno)* **23**, 243 (1986).
13. F. H. Smith, *Curr. Anthropol.* **23**, 667 (1982).
14. The dual-ancestry model, which is a multiregional one, depends on reticulation and is visible through evidence of regional continuity. In criticizing the multiregional explanation of modern human origins in Europe, G. Bräuer and H. Broeg [in *Origins and Past of Modern Humans: Towards Reconciliation*, K. Omoto, P. V. Tobias, Eds. (World Scientific, Singapore, 1998), pp. 106–125] state, "during the long period of coexistence, mixing and gene flow between Neandertals and the dispersing modern populations might have occurred to varying degrees in different regions ... such gene flow could have mimicked some degree of continuity" (p. 106). To avoid confusion, we note that multiregional evolution is not mimicked by gene flow, but is based on it.
15. They have unequal numbers of males and females, and because of differences in preservation, the sample represented for one trait is often not the same as for another.
16. A trait, say, scored with four character states could contribute as much to a pairwise difference analysis as three traits scored as present or absent. We realize that our nonmetric characters are not formally equal to each other in complexity or heritability, but con-

Fig. 3. Robustness and cranial size. The relation between presence of nonmetric traits used in the WLH-50 analysis, presented here as the sum of their scores for each individual, and cranial size, as measured by cranial capacity. The sum of the nonmetric scores is a measure of robustness because each was scored as 0 or 1, and whenever robustness characterized the difference, the more robust condition was scored as 1. The geographic affinities of the groups being compared provide the main source of variation.



tend that our approach to scoring makes them as comparable as possible.

17. A. R. Rogers, *Evolution* **49**, 608 (1995).
18. M. Krings *et al.*, *Cell* **90**, 1 (1997).
19. P. Brown, *Coobool Creek. Terra Australis* **13** (Department of Prehistory, Research School of Pacific Studies, Australian National University, Canberra, 1989).
20. M. M. Lahr, H. V. S. Wright, *J. Hum. Evol.* **31**, 157 (1996).
21. This helps to avoid spurious correlations that might result from using different measurements of the same thing as independent and dependent variables [B. Solow, *Acta Odontol. Scand.* **24** (suppl. 46), 1 (1966)].
22. J.-J. Hublin, in *Neandertals and Modern Humans in Western Asia*, T. Akazawa, K. Aoki, O. Bar-Yosef, Eds. (Plenum, New York, 1998), pp. 295–310.
23. F. S. Swed, C. Eisenhart, *Ann. Math. Stat.* **14**, 66 (1943).
24. A run is one or more crania from the same group in

- a row. For instance, the Mladeč 6 analysis (Table 3) has three runs.
25. With a larger sample and fewer runs in the WLH-50 analysis, randomness can be rejected at $P = 0.05$. Ngandong is significantly closer to WLH-50, which is compatible with dual ancestry but is a refutation of the replacement hypothesis, so the consequences are the same even though the significance test results were opposite.
26. Meaning that all of one sample were the same for the character state, and only one specimen in the other sample differed from the opposite character state.
27. I. Tattersall, *The Last Neanderthal: The Rise, Success, and Mysterious Extinction of our Closest Human Relatives*. (MacMillan, New York, 1995).
28. D. W. Frayer, *Préhist. Eur.* **2**, 9 (1993).
29. C. Duarte *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7604 (1999); A. E. Mann, J. Monge, M. Lampl, *Am. J. Phys. Anthropol.* **86**, 175 (1991); J. Szilvássy, H. Kritscher, E. Vlček, *Ann. Vienna Nat. Hist. Mus.* **89**, 313 (1987).

30. R. Caspari (University Microfilms, Ann Arbor, MI, 1991).
31. We thank A. Thorne (Australian National University) for allowing M.H.W. to study the original WLH-50 specimen. We are grateful to J. Jelinek (Moravian Museum, Brno), H. Kritscher, H. Poxleitner, L. Seitzl, J. Szilvássy, and M. Teschler (Vienna Natural History Museum) for permission to work on the Mladeč specimens, and thank the many curators and museum directors with responsibility for the comparative samples we discuss here for access to the fossil collections in their care. Supported by NSF grants BNS 75-21756, BNS 76-82729, and BNS 85-09147; grants from the National Academy of Sciences, Eastern European Program; a grant from the Department of Prehistory, Research School of Pacific Studies, Australian National University; and grants from the Rackham Graduate School, University of Michigan.

2 October 2000; accepted 27 November 2000

Natal Homing in a Marine Fish Metapopulation

Simon R. Thorrold,^{1*} Christopher Latkoczy,² Peter K. Swart,³ Cynthia M. Jones¹

Identifying natal origins of marine fishes is challenging because of difficulties in conducting mark-recapture studies in marine systems. We used natural geochemical signatures in otoliths (ear bones) to determine natal sources in weakfish (*Cynoscion regalis*), an estuarine-spawning marine fish, in eastern North America. Spawning site fidelity ranged from 60 to 81%, comparable to estimates of natal homing in birds and anadromous fishes. These data were in contrast to genetic analyses of population structure in weakfish. Our findings highlight the need for consideration of spatial processes in fisheries models and have implications for the design of marine reserves in coastal regions.

The exchange of individuals among geographically separated groups, or connectivity, is a critical property of marine populations (1). Connectivity rates determine colonization patterns of new habitats, the resiliency of populations to harvest, and the design of marine protected areas (MPAs). Quantifying exchange rates in marine organisms is, however, extremely difficult because natal origins of adults are almost invariably unknown (2, 3). This lack of knowledge is primarily due to the difficulty of conducting mark-recapture studies in species that are characterized by the production of large numbers of small pelagic offspring that suffer high initial mortality rates. Recently, tagging techniques using natural isotopic and elemental markers have been developed for species that were not

able to be tagged or recaptured using conventional approaches. Use of these natural tags may, therefore, allow questions of natal origins, spawning site fidelity (or philopatry), and return migrations to be addressed in organisms and environments for which we have little information, including migratory songbirds (4), butterflies (5), and fish (6, 7). For instance, fisheries ecologists have yet to determine rates of natal homing in any species of marine fish, even though spawning site fidelity and other complex migratory behaviors are common in birds (8) and anadromous fishes (9). We provide unique estimates of philopatry and population structure in weakfish (*Cynoscion regalis*) using stable isotope and elemental signatures in otoliths of returning spawners as a natural tag of natal origin.

Adult weakfish follow an annual migration pattern along the east coast of the United States that takes them from overwintering grounds south and offshore of Cape Hatteras to spawning locations in estuaries and coastal embayments throughout the species range (Florida to Maine) in the spring and early summer. Larvae are generally retained within natal estuaries through selective tidal stream transport (10), and they reside in these estuaries until migrating to overwintering

grounds in the autumn. Given the lack of larval dispersal, connectivity rates are primarily determined by the propensity for adult fish to return to their natal estuary to spawn. Previous studies of weakfish population genetics using allozymes and mitochondrial DNA (mtDNA) (11, 12) have found no evidence of genetic differentiation. However, genetic approaches may not have sufficient resolution to quantify natal homing unless straying is negligible over evolutionary time scales. The natural tag approach that we took provided estimates of natal homing and population structure in the presence of significant connectivity among groups within the larger metapopulation.

The otoliths of teleost fish are accretionary structures located within the inner ear, and are composed primarily of aragonite deposited on a proteinaceous matrix. The utility of otolith chemistry as a natural tag relies on three properties of otoliths (13). First, the deposition time of otolith material can be estimated by reference to concentric rings in otoliths that are routinely used in age estimation by fisheries biologists. Second, the metabolically inert nature of otoliths ensures that the aragonite mineralogy remains unaltered after deposition. Third, the calcium carbonate and trace elements that make up most (>90%) of the otolith are derived primarily from the ambient water, as modified by temperature. The isotopic and elemental composition of the otolith will, therefore, reflect the environmental characteristics of the water in which the fish lives. Because physical and chemical composition characteristics of water vary spatially, otolith geochemistry records the water mass characteristics specific to a particular natal area. The mechanisms generating variability in the signatures may not necessarily be restricted to environmental differences. Nonetheless, application of geochemical signatures in otoliths as a natural tag does not require the reconstruction of these environmental differences, only that the signatures are sufficiently robust to allow

¹Department of Biological Sciences, ²Laboratory for Isotope and Trace Element Research, Old Dominion University, Norfolk, VA 23529, USA. ³Division of Marine Geology and Geophysics, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149, USA.

*To whom correspondence should be addressed. E-mail: sthorrold@whoi.edu

†Present address: Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA.