How Much Can Fossils Tell Us about Regional Continuity?[1]

ALAN R. ROGERS
Department of Anthropology, University of Utah, Salt Lake City, Utah 84112, U.S.A. (rogers @anthro.utah.edu). 3 II 95

The term *regional continuity* has figured heavily in the debate on modern human origins. It refers to the genetic contribution of earlier populations to later populations within regions. For example, regional continuity would imply that the Asian population of 30,000 B.P. was descended in part from earlier Asian populations. Some paleontologists argue that the human fossil record provides evidence of regional continuity throughout the Pleistocene in various parts of the Old World (Frayer et al. 1993). Morphological features of later populations are held to resemble morphological features of earlier populations within the same region. This view is disputed by others, who see evidence for a replacement of archaic populations by a population of modern humans that expanded throughout the world some 50,000 years ago (Stringer and Andrews 1988, Stringer 1994, Cann, Stoneking, and Wilson 1987, and Lahr 1994).

In view of this disagreement, it would be useful to evaluate the reliability of the argument for regional continuity. This is difficult, however, since paleontological arguments are seldom couched in statistical terms. Here I propose to evaluate the argument by means of a thought experiment. Rather than studying real fossils, I will study hypothetical data with many unrealistic assumptions. Each of these assumptions will serve to make the hypothetical data unrealistically good, increasing the likelihood that regional continuity will be detected. I will show that, even with these unrealistically good data, regional continuity is difficult to detect.

Why the correlation between early and late character values measures regional continuity. I deal with a hypothetical skeletal character, which may be either continuous (such as one might measure on a bone with a pair of calipers) or dichotomous (taking only the values "+" and "−"). I maximize the information that this character can provide by making a simple genetic assumption. If the character is continuous, then I assume it to be completely heritable, with an additive genetic basis. If it is dichotomous, then I assume it to be determined by a single genetic locus, with the "+" phenotype exhibited by one homozygote and half the heterozygotes and

---

[1] I am grateful for comments from Elizabeth Cashdan, Richard Fowles, David Frayer, Henry Harpending, Kristen Hawkes, Steven Josephson, Elisabeth Manderscheid, Dennis O'Rourke, and Milford Wolpoff. This work was supported in part by a grant from the National Science Foundation (DBS-9310105).

the "−" phenotype exhibited by everyone else.[2] These assumptions eliminate the environmental and nonadditive genetic noise that can obscure genetic relationships among populations and therefore make the hypothetical data unrealistically good. In addition, I assume (unrealistically) that this character is selectively neutral. With real characters, which may not be selectively neutral, later populations may resemble earlier ones even without any genetic relationship between the two. Thus, historical inferences are less reliable with characters that may have been affected by selection. I improve the hypothetical data by assuming this problem away.

With data from K regions and two time periods, I denote by x_i the mean character value within the ith region in period 1 and by y_i the corresponding value in period 2. It seems natural to hope that the correlation, ρ, between x_i and y_i (over all i) might measure the strength of regional continuity. I will show that this is indeed the case under a model of population structure general enough to accommodate as special cases both the multiregional and the replacement model of modern human origins.

I assume that between periods 1 and 2 the population of region 1 expands and partially replaces the populations of the other regions. In period 2, a fraction u of each regional population descends from these invaders. In addition, a fraction q of each regional population descends from the earlier population of that same region. Finally, a fraction v $(= 1 - u - q)$ is drawn at random from the population of period 1 as a whole. Figure 1 illustrates this model in the case in which $v = 0$. When $q = 0$, the entire population of period 2 descends from the earlier population of region 1. In this case, the model reduces to the replacement hypothesis of modern human origins (Stringer and Andrews 1988). If, instead, $u = 0$ but q and v are positive (a possibility not shown in the figure), then the model includes only gene flow and regional continuity. In other words, it reduces to the multiregional hypothesis of modern human origins (Wolpoff 1989). In either case, q measures regional continuity—the contribution of earlier populations to later populations within regions.

This model is unrealistic in implying that in period 2 the regional populations are statistically independent. In real data, correlations between regions would make the effective number of regions smaller than the actual number. Thus, it will be unrealistically easy to detect regional continuity with the hypothetical data.

The model of population structure implies that $\rho = q$,[3] whatever the values of the other parameters. Thus, the correlation between character values of earlier and later populations measures regional continuity. Since the model is general enough to include both multiregional and replacement hypotheses, it follows that ρ measures regional continuity regardless of which hypothesis is correct.

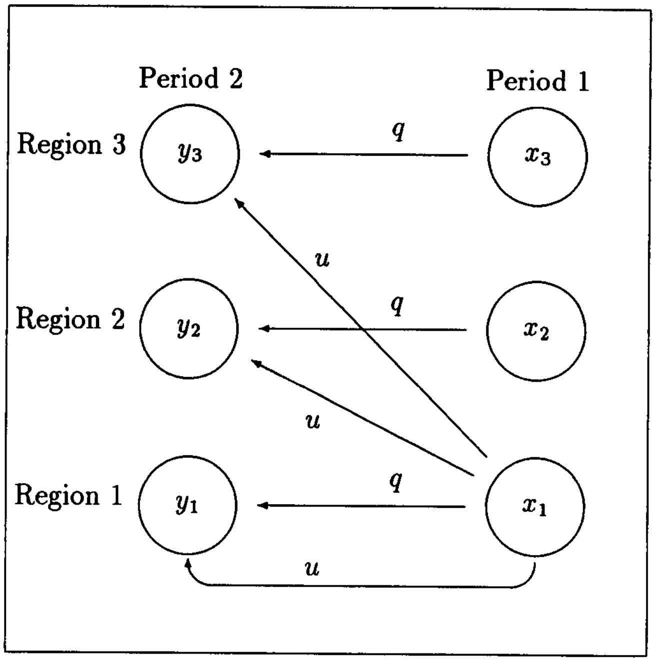

FIG. 1. *Model of population structure in the case in which* v = 0, *so that, apart from the invaders from region 1, there is no gene flow between regions. In the general case, a fraction* v *of each region in period 2 is drawn at random from the population as a whole.*

Testing for regional continuity with a single character. To demonstrate regional continuity (i.e., that $q > 0$), it would be necessary to reject the hypothesis that $\rho = 0$. I improve that hypothetical data still further by assuming that infinite samples are available from each region so that each x_i and y_i is known without error. These are still random variables because of the variation introduced by the evolutionary process itself and have approximately normal probability distributions.[4] Consequently, we can test the hypothesis that $\rho = 0$ by calculating the sample correlation coefficient, r, using the ob-

---

[2] The expected frequency of the "+" phenotype is therefore equal to frequency of the "+" allele.

[3] The assumptions imply that

$$y_i = qx_i + vE[x] + ux_1 + \epsilon_i, \quad i = 1, 2, \ldots, K.$$

Here E denotes the expectation and ϵ_i is a deviation caused by genetic drift, which has expectation zero and is uncorrelated with the other quantities. The expectation of y_i is $E[y] = (q + v)E[x] + ux_1$. Thus, the covariance over i between x_i and y_i is $C_{XY} \equiv E[x(y - E[y])] = qV_X$, where $V_X \equiv E(x^2 - E[x]^2)$ is the variance of x_i. The correlation between x_i and y_i is $\rho \equiv C_{XY}/\sqrt{V_X V_Y} = qV_X/\sqrt{V_X V_Y}$. At equilibrium between migration and genetic drift, $V_X = V_Y$, and $\rho = q$.

[4] The normality of neutral continuous characters with additive genetic bases is well known. In the dichotomous-character case, approximate normality follows from the observations that (1) within each region the mean character value equals the frequency of a neutral allele and (2) the probability distribution of a neutral allele is approximately normal if the number of generations separating periods 1 and 2 is smaller than population size.

served x_i and y_i. With data from $K = 5$ regions (perhaps Africa, Europe, the Middle East, the Far East, and Australia), we would need $r > 0.88$ in order to reject the hypothesis that $\rho = 0$.[5] Thus, if inference were based on a single character, we would be unlikely to detect regional continuity unless $q > 0.88$. This number could be reduced by increasing either the number of regions or the number of characters. Paleontological samples are grouped into major geographic regions because the samples within smaller regions are small. With real data, analysis at a finer geographic scale would require larger samples than are currently available. In the short run, therefore, power can be increased only by increasing the number of characters.

Testing for regional continuity with multiple characters. How many characters might a paleontologist employ? There is presumably no limit to the number of things that one can measure on a bone, so it is tempting to suppose that the number of characters can be large. Were we interested only in an accurate estimate of ρ, we could indeed increase accuracy indefinitely by incorporating additional characters. But to test the hypothesis that $\rho = 0$ it is also necessary to estimate the standard error of our estimate, and this requires estimating the variances of and the covariances between the contributions of different characters. Each new character burdens the model with additional parameters: the mean and variance of the new character and its covariance with each preceding character. To maximize the information provided, four proves to be the optimal number of characters.[6] I improve the hypothetical data once again by assuming these characters to be independent.

The four independent characters provide four independent tests of the hypothesis that $\rho = 0$. These can be combined into a single test using a method outlined by Bulmer (1967:164).[7] It turns out that the hypothesis of no continuity can be rejected only if $r > 0.77$ for each character.[8] Since r estimates q, this result implies that continuity is unlikely to be detected unless $q > 0.77$.

Discussion. No matter how many characters are measured, paleontology is unlikely to detect regional continuity unless $q > 0.77$, and this number is unrealisticaly low because our hypothetical data are unrealistically good. With real data, statistical significance would require an even larger value. Even where it exists, regional continuity is unlikely to be reflected in fossils.

Why are these data so lacking in statistical power? The answer lies in the one factor that I did not increase to unrealistically high values: the number of regions in the analysis. With only five regions, the correlation must be extremely high to differ significantly from zero. This problem will remain until paleontological samples are much larger.

This analysis has shown that paleontological inference requires large values of q but has not yet shown that such values are implausible. Elsewhere, Manderscheid and I (1994) ask what hypotheses about q can be rejected on the basis of the uniformity of human mitochondrial DNA. Assuming that mitochondrial variation is selectively neutral, we find that if the effective human population size has been at least 320,000 for the past 30,000 years, then the hypothesis that $q > 0.03$ can be rejected at the 0.05 significance level. If our neutrality assumption is correct, then the paleontological case for regional continuity cannot be statistically significant.

References Cited

BULMER, M. G. 1967. *Principles of statistics.* New York: Dover.

CANN, R. L., M. STONEKING, AND A. C. WILSON. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31–36.

FRAYER, D. W., M. H. WOLPOFF, A. G. THORNE, F. H. SMITH, AND G. G. POPE. 1993. Theories of modern human origins: The paleontological test. *American Anthropologist* 95:14–50.

LAHR, M. M. 1994. The multiregional model of modern human origins: A reassessment of its morphological basis. *Journal of Human Evolution* 26:23–56.

MANDERSCHEID, E., AND A. R. ROGERS. 1994. Genetic admixture in the late Pleistocene. MS.

MENDENHALL, W., AND R. L. SCHEAFFER. 1973. *Mathematical statistics with applications.* North Scituate: Duxbury Press.

STRINGER, C. B. 1994. "Out of Africa: A personal history," in *Origins of anatomically modern humans.* Edited by M. H. Nitecki and D. V. Nitecki, pp. 149–72. New York: Plenum Press.

STRINGER, C. B., AND P. ANDREWS. 1988. Genetic and fossil evidence for the origin of modern humans. *Science* 239:1263–68.

WOLPOFF, M. H. 1989. "Multiregional evolution: The fossil alternative to Eden," in *The human revolution: Behavioral and biological perspectives on the origins of modern humans.* Edited by P. Mellars and C. Stringer, pp. 62–108. Princeton: Princeton University Press.

---

[5] The hypothesis that $\rho = 0$ can be rejected at the 0.05 significance level if $Z \equiv \sqrt{K - 3}\ \ln(1 + r/1 - r)/2 > 1.96$ (Mendenhall and Scheaffer 1973:421). Solving for r with $K = 5$ gives $r > 0.88$.

[6] With L characters from 5 regions, we have $5L$ observations, and the model includes the following parameters: q, u, the mean of each character, and a covariance matrix comprising $L(L + 1)/2$ variances and covariances. (v doesn't count because $v \equiv 1 - q - u$.) Thus, the model has $2L - (2 + L + L(L + 1)/2)$ degrees of freedom. Degrees of freedom are maximal when L is 3 or 4.

[7] If p_i is the p-value of the ith of K independent tests, then $M = -2 \sum_{i=1}^{K} \ln p_i$ is a χ^2 random variable with $2K$ degrees of freedom. Given a set of p_i values, we can reject the hypothesis that $\rho = 0$ if M exceeds some critical value found in a χ^2 table. Replacing each p_i with $\bar{p}$ will inflate M's value (a consequence of Jensen's inequality). Thus, I improve the hypothetical data once again by assuming all the p_i's to be equal.

[8] With $K = 4$ characters we have 8 degrees of freedom. A χ^2 table shows that the hypothesis can be rejected when $M > 15.503$, which requires that $p < 0.144$. Consulting a table of the normal distribution, I find that this requires $Z > 1.46$, where Z is as defined in n. 5. Finally, solving $Z > 1.4$ with $K = 5$ gives $r > 0.77$.