

A SINGLE LINEAGE IN EARLY PLEISTOCENE *HOMO*: SIZE VARIATION CONTINUITY IN EARLY PLEISTOCENE *HOMO* CRANIA FROM EAST AFRICA AND GEORGIA

Adam P. Van Arsdale^{1,2} and Milford H. Wolpoff³

¹Department of Anthropology, Wellesley College, 106 Central Street, Wellesley, Massachusetts 02481

²E-mail: avanarsd@wellesley.edu

³Department of Anthropology, University of Michigan, 1085 South University Avenue, Ann Arbor, Michigan 48109

Received May 3, 2012

Accepted September 18, 2012

The relationship between *Homo habilis* and early African *Homo erectus* has been contentious because *H. habilis* was hypothesized to be an evolutionary stage between *Australopithecus* and *H. erectus*, more than a half-century ago. Recent work re-dating key African early *Homo* localities and the discovery of new fossils in East Africa and Georgia provide the opportunity for a productive re-evaluation of this topic. Here, we test the hypothesis that the cranial sample from East Africa and Georgia represents a single evolutionary lineage of *Homo* spanning the approximately 1.9–1.5 Mya time period, consisting of specimens attributed to *H. habilis* and *H. erectus*. To address issues of small sample sizes in each time period, and uneven representation of cranial data, we developed a novel nonparametric randomization technique based on the variance in an index of pairwise difference from a broad set of fossil comparisons. We fail to reject the hypothesis of a single lineage this period by identifying a strong, time-dependent pattern of variation throughout the sequence. These results suggest the need for a reappraisal of fossil evidence from other regions within this time period and highlight the critical nature of the Plio-Pleistocene boundary for understanding the early evolution of the genus *Homo*.

KEY WORDS: African evolution, fossils, *Homo habilis*, *Homo erectus*, phylogenetics.

Background

Significant uncertainty exists surrounding the evolutionary relationships of fossil crania attributed to the genus *Homo* in the ~1.9–1.5 million year time period. A number of differing interpretations have developed during the first half-century of discoveries and additional specimens have not resulted in a convergence of views. This time period is when the first evidence for a range expansion outside of Africa is found and includes a large number of fossil crania from key East African localities in Olduvai Gorge (Leakey et al. 1964, 1971; Rightmire 1979; Johanson et al. 1987; Tobias 1991, 2003; Antón 2004), the Lake Turkana Basin (Wood 1991; Spoor et al. 2007; Leakey et al. 2012), as well as early Eurasian cranial material from Dmanisi, Georgia

(Gabunia et al. 2002; Lordkipanidze et al. 2006; Rightmire et al. 2006).

Debate about the appropriate taxonomy of individual crania and the evolutionary relationship of the sample as a whole date back to the initial publication of *Homo habilis* material from Olduvai Gorge (Leakey et al. 1964; Tobias 1964; Robinson 1965, 1966). Writing at the time, John Robinson famously disagreed with the classification of *H. habilis* at Olduvai Gorge. While conceding that the Bed I and Bed II material from Olduvai might represent a lineal relationship, he was the first to frame the hypothesis we examine here:

... the Bed I [*Homo habilis*] material may represent an advanced form of *Australopithecus* and the Bed II specimens an



early *H. erectus* and at the same time the latter may be a lineal descendant of the former (Robinson 1966, p. 123).

The subsequent discovery of a large number of hominid remains from this time period in localities in the Lake Turkana region of Kenya stimulated additional discussion (Boaz 1979; Trinkaus 1984; Wood 1992; Rightmire 1993), much of which focused on ER 1470, a seemingly anomalous cranium when dated to 2.6 Ma but now known to have been incorrectly dated (Feibel et al. 1989, who established provenience and date estimates for virtually all of the East African specimens). A number of researchers who analyzed variation in early *Homo* crania from this time period concluded that the variation within this sample exceeds that which might be expected from a single species, even beyond the expectations of greater sexual dimorphism, and subsequently divided the sample into two or more sympatric evolutionary lineages (Donnelly 1996; Grine et al. 1996; Wood 1985; Stringer 1986; Lieberman et al. 1988; Wood 1991; Wood 1993; Wood and Richmond 2000; though see Miller 1991; Tobias 1991; Kramer 1993; 2000; Miller et al. 2004, for dissenting views).

Discussions focused on this time period have been reignited by the discovery of new fossil material from East Africa (Asfaw et al. 2002; Tobias 2003; Spoor et al. 2007; Baab 2008; Leakey et al. 2012) as well as a particularly well-preserved set of cranial remains from Dmanisi, Georgia (Gabunia et al. 2002; Vekua et al. 2002; Lordkipanidze et al. 2005; Lordkipanidze et al. 2006; Rightmire et al. 2006). Yet, rather than clarifying evolutionary relationships that exist during this time period, the new East African and Georgian remains seemed to broaden the observed range of variation, expanding variation accepted within *Homo erectus* if not pushing it to the limit (Rightmire 1990; Bilsborough 2000; Antón 2003) and further complicating the various issues of relationships and evolutionary pattern.

A 21st century reappraisal of the dating for several key localities within the Koobi Fora sequence (Gathogo and Brown 2006; Suwa et al. 2007; Gathogo et al. 2008) is the final element in the case for reexamining the relations between the specimens. While not without controversy (Feibel et al. 2009), the redating prompted Gathogo and Brown (2006, p. 478) to note:

Crania KNM-ER 1813 and KNM-ER 1470 are quite complete, and were previously thought to be contemporary . . . They have been at the heart of the debate over whether *H. habilis sensu lato* is in fact composed of two species: *H. habilis* and *H. rudolfensis*. As shown above, KNM-ER 1813 is ~1.65 myr in age, ~0.25 myr younger than KNM-ER 1470. Thus, those who advocate including both specimens within *H. habilis* no longer need to accommodate the diverse facial morphologies of these specimens within contemporary intraspecific variation.

One research group that systematically reexamined the relationship between *H. habilis* and *H. erectus* (Suwa et al. 2007)

Table 1. Early *Homo/Australopithecus boisei* cranial sample.

1.9–1.8	1.8–1.7	1.7–1.6	1.6–1.5	<1.5
OH 7	OH 16	OH 13	ER 1808	OH 9
OH 24	ER 1590	ER 730	ER 3883	OH 12
ER 1470	D 2280	ER 1805	ER 42700	Daka
ER 3732	D 2282	ER 1813*	WT 15000	
ER 3735	D 2700	ER 3733		
ER 62000	D 3444	ER 3891		
	ER 13750	WT 17400	ER 405	ER 733
	OH 5		ER 406	Ches 1
			ER 407	Ches 303
			ER 732	
			ER 23000	

consequently concluded that all known specimens assigned to *H. habilis/H. rudolfensis* within the Turkana basin predate known specimens assigned to *H. erectus/H. ergaster*. They failed to reject the null hypothesis of a single lineage of *Homo* spanning this time period (though see Spoor et al. 2007; Leakey et al. 2012, for differing views).

We believe this null hypothesis is a straightforward, statement of relationship that can be tested. Using cranial remains, we expand this follow-up of Robinson's hypothesis to include crania from Ethiopia, Kenya, Tanzania, and Georgia dated to this time period and propose a novel statistical method to formalize our hypothesis-testing approach.

Methods and Materials

To test the hypothesis of a single, evolving lineage of *Homo* in the ~1.9–1.5 Ma window, we divide the cranial sample into a sequence of five time intervals and focus our efforts on predictions relating to the pattern of variation within and between time intervals. Such an approach is necessary to distinguish the critical role that time sequencing plays in discriminating between the null hypothesis and potential alternative hypotheses, such as parallel lineages or diverging lineages. Explicit in the null hypothesis of a single, evolving lineage, is that the pattern of variation changes through time in characteristic and identifiable ways. Our work is based on comparisons of homologous measurements from complete or partially complete crania within each time interval (see Table 1). A sample of crania from East Africa post-dating this time frame, consisting of Olduvai hominins 9 and 12 (Antón 2004), as well as the Daka specimen from Bouri, Ethiopia (Asfaw et al. 2002), have been included to provide a posterior framing perspective on cranial variability for this time period. Additionally, to validate the efficacy of our novel approach, we provide a complimentary set of analyses that includes data from cranial remains of *Australopithecus boisei* (*Au. boisei*), a

Table 2. Summary results for different analytical treatments of data.

Pairwise specimen comparisons	ER 1813 in interval one	ER 1813 in interval three
<i>Homo</i> -only	$P = 0.100$	$P = 0.012^*$
<i>Homo/Australopithecus</i>	$P = 0.220$	$P = 0.110$
Pairwise measurement comparisons	ER 1813 in interval one	ER 1813 in interval three
<i>Homo</i> -only	$P = 0.001^{**}$	$P = 0.001^{**}$
<i>Homo/Australopithecus</i>	$P = 0.03^*$	$P = 0.04^*$

contemporary hominid lineage (Table 1). This second analysis, involving a mixed *Homo/Australopithecus* sample, is intended to demonstrate the ability of our approach to distinguish between single versus multiple lineage hypotheses by contrasting a hypothetical early *Homo* lineage with an uncontested and identifiable contemporary hominid lineage.

A complete set of the 134 measurements used in the study can be found in the online supplemental appendix (Appendix S1). Efforts were made to include cranial measurements that could be well sampled within these fossils and that provided broad coverage of the preserved cranial anatomy. The complete dataset includes additional measures not usually reported for complete specimens that are specifically intended to allow for the comparison of fragmentary remains, thus expanding the set of remains available for comparison. Most measurements were recorded directly off the original fossil material, but have been supplemented with data gathered from the literature when necessary (see Appendix S1 for clarification).

The placement of specimens into given time intervals is based on current understandings of the stratigraphic placement of early *Homo* fossils drawn from the literature. The Dmanisi dates are from Lordkipanidze et al. (2007). For the temporal assessment of the most recent discoveries from the Turkana Basin, we rely on the dates provided by Spoor et al. (2007) and Leakey et al. (2012). For the earlier cranial material recovered from Olduvai and Turkana, we rely on the revised dating from Feibel et al. (1989), with the exceptions presented by Gathogo and Brown (2006) and Suwa et al. (2007). We believe these revisions have clarified the evolutionary relationships within the sample, but a few problems require further discussion.

Gathogo and Brown (2006, Table 1) place ER 1813 at 1.65 Myr and “approximately the same age as ER 3733” (p. 478). Suwa et al. (2007, Table 2) suggest a wider possible date range for the specimen, 1.55–1.80 Myr, which is less precise but not contradictory. Feibel and colleagues (2009) also reevaluate specimens from Koobi Fora Area 123 and estimate an age of $1.86 \pm$

0.08 Myr for ER 1813, even older than the Suwa et al. (2007) range.

We are not in a position to resolve this issue based on the geologic associations of the ER 1813 specimen, and are inclined to accept the more constrained Gathogo and Brown estimate because of the close similarity of ER 1813 to OH 13, a second female also the age of ER 3733 according to Suwa et al. (2007, Table 2). The OH 13 age is an independent indication that this female anatomy is consistent with the age of ER 3733. However, to avoid prejudicing our analysis on the basis of an unresolved dating dispute, and to further examine the robustness of our results, we have conducted all of our analyses twice, once with the Gathogo and Brown placement between 1.6 and 1.7 Myr and a second analysis using the Feibel placement of ER 1813 in the 1.8–1.9 Myr interval (following Feibel et al. 2009).

As a result, our final analyses include four treatments of the data, including an analysis based solely on the putative early *Homo* sample and one based on a mixed *Homo/Australopithecus* sample, each with ER 1813 in one of two time intervals.

To define our testing criteria, we propose that if these samples come from a single evolutionary lineage, reflecting the evolving changes the pattern of variation in the cranial sample should be time-dependent. In particular, the time-dependent nature of the assemblage, coupled with the evolutionary change observed across this time series, should create a pattern of variation whereby pairs of individuals from within any time interval, on average, are less variable than pairs of individuals drawn from across time intervals. Our null hypothesis makes explicit the notion that early *Homo* is an evolving lineage and not static with respect to the pattern of variation it preserves. This is a novel approach to the question of taxonomic variation in early *Homo*, but one that is necessitated by the insufficiency of approaches that use static models of variation, such as standing variation in recent humans or nonhuman primates, to test hypotheses regarding a lineage undergoing evolutionary change. A failure to reject the time-dependent pattern of variation within our sample, as well as the pattern of directional change within that lineage, suggests an appropriate parsimonious interpretation is that of a single evolutionary lineage based on the existing cranial evidence.

Our procedure tests the null hypothesis by examining the relationship between all possible pairwise comparisons of each cranium. To assess variation, an index of relative difference was calculated from each available homologous measurement comparison based on the absolute value of the log-transformed difference between measurement values (see below). This index was calculated for the comparison of every homologous measurement shared between each of the crania within our study sample. These values were then converted into a composite value for each pairwise fossil comparison that we refer to as an Average Index of Relative Difference.

Average Index of Relative Difference

$$= \frac{\sum_{x=1}^n \{abs[\ln(x_A) - \ln(x_B)]\}}{n}.$$

The complete dataset, including early *Homo* and *Au. boisei* specimens, includes 630 Average Index of Relative Difference values that represent the relative size difference between each of the specimens in our study. These data were culled to eliminate individuals that preserved fewer than five homologous measurements for comparison. The remaining set included 451 values, 222 of which represented putative early *Homo* comparisons and 229 of which represented pairwise comparisons between early *Homo* and *Au. boisei* specimens or between *Au. boisei* specimens. The number of measurements available for the calculation of the Average Index of Relative Difference ranged from five to 124, with the total dataset the product of more than 10,700 individual measurement comparisons.

Another potential source of bias within our data comes from the differences in preservation between individual specimens. Although some specimens are well preserved and contain many potential measurement comparisons, others are more fragmentary or have had only a limited number of measurements published. Although we have eliminated comparisons made on fewer than five homologous measurements, it remains possible that specimens with relatively fewer measurements might bias the data. To assess this possibility, we also tested our hypothesis using the entire set of individual measurement comparisons rather than the composite, average index of pairwise difference outlined above. Our approach is intended to mitigate the potential bias by testing the null hypothesis under conditions in which individual measurements are the primary unit of comparison and alternatively under conditions in which individual specimens are the primary unit of comparison. In the analysis based on whole specimen comparisons, each pairwise specimen comparison is given equal weight, regardless of how many homologous measurement comparisons contributed to that value (excepting those with fewer than five such measures). The analysis based on individual measurement comparisons instead treats every homologous measurement equally, thereby weighting the final result toward specimens with greater preservation. Ideally, both sets of data will produce equivalent results with respect to our hypothesis. The complete set of fossil pairwise comparisons that were made, as well as the number of homologous measurements that contributed to each comparison can be found in the Supporting Information (see Appendix S2).

THE PREDICTION OF INCREASING VARIANCE AS THE BREADTH OF TIME BEING SAMPLED INCREASES

Our hypothesis test is based on the variance of pairwise difference values across time intervals. The hypothesis of a single, evolving lineage predicts that the variance generated from a set of pair-

wise comparisons should increase as the breadth of time being sampled increases, “regardless of whether the pairwise sample consists of individuals earlier or later in time.” If the sample is a single evolving lineage, the variance will increase “from whatever point in time the sample is being viewed,” looking forwards or backwards through time. Sample variance should be lowest for pairwise samples drawn from the same time interval, regardless of what time interval they are drawn from, and increase as the time being sampled by a pair increases.

A commonly argued alternative hypothesis for early *Homo* is that at least two taxa are present throughout the sequence; suggesting two or more parallel or divergent lineages are hypothesized to be present. The prediction for a pattern consistent with two lineages diverging, such as might be expected from contemporaneous lineages of *Homo*, is that the variance of sampled pairs will increase “from the point of divergence but decrease toward the point of divergence.” This distinction can be seen between scenarios B and C of Figure 1, with particular reference to the matrix of expectations at the bottom of each.

Our hypothesis can be mathematically represented by a matrix of average pairwise difference variance values that increase away from the diagonal (see Fig. 1). The more time that separates any two intervals, the larger the expectation of variance in a sample of pairwise comparisons because of the greater effect of directional change in either direction through time. The observed pattern of variation within our sample, specifically the stepwise increase in the variance of pairwise comparisons away from the diagonal (within time interval comparisons), can be used as a test statistic to examine the time-dependent nature of variation within these fossils as well as the observed patterns concordance with the expectations of a single, evolving lineage. This prediction contrasts with the predictions for pairwise variance for two alternative explanatory models, that of static, sympatric lineages through time, and sympatric lineages diverging from a recent common ancestor (see Fig. 1).

Including *Australopithecus* data allows us to directly address whether our approach is capable of distinguishing different patterns of variation in a mixed taxa assemblage, because this is a separate, diverging lineage. The addition of *Au. boisei* should disrupt the time-dependent nature of the overall pattern of variation. It should be noted that the available *Au. boisei* cranial sample is considerably smaller than that of early *Homo*, and consists largely of more fragmentary specimens. Therefore, this is a strong test, in that if the sample of both genera can be distinguished from the *Homo* sample, it took a minimal number of specimens to identify the difference.

To assess the significance of the observed pattern of variation, we created a distribution of expected values for the stepwise increase in variance away from the diagonal by repeatedly randomly reordering our sample and creating an equivalent matrix of

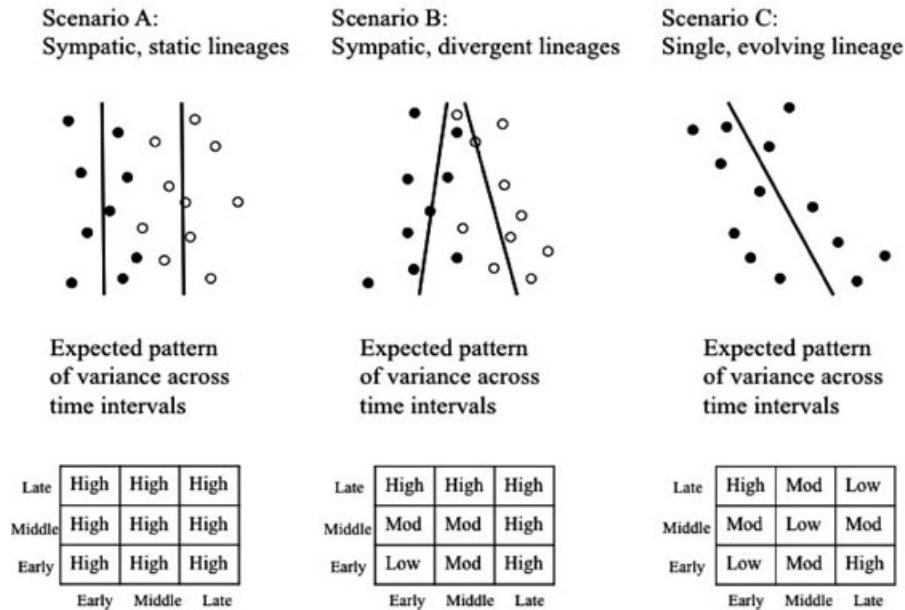


Figure 1. Expectations for pattern of pairwise variance. This figure displays the expected pattern of variation in pairwise comparisons across three evolutionary scenarios. The pattern observed within our sample, that of consistent increases in variance away from the diagonal, aligns most closely with that of scenario C, a single, directionally evolving lineage through time.

within and between time interval variance values. The goal of this procedure to create an expected distribution of variance values under a scenario where the time ordering of the fossils within our sample does not matter. The generated expected distribution of values allows us to test the significance of our observed values by generating a functional *P*-value, assessing whether the time ordering we observe in the fossil record is important in explaining the variation within our sample. This is an explicit test of directional evolutionary change within a hypothetical early *Homo* lineage. A significant departure from expectations would suggest that our observed pattern of variation is strongly connected to the observed time ordering of our fossil sample. All analyses were conducted using custom code written in the Matlab software package.

Results

Our research findings fail to disprove the simplest explanation of the cranial variation, that East African and Georgian fossil crania attributed to *Homo* and variously described as *H. rudolfensis*, *H. ergaster*, *H. georgicus*, and *H. erectus* sample a single evolving lineage. We fail to reject the hypothesis that the pattern of variance in pairwise comparisons of *Homo* fossils spanning the 1.9–1.5 Myr range is highly time-dependent and that the observed pattern is most consistent with a single lineage experiencing evolutionary change through time.

Our analyses produced eight sets of results. These include results generated from a *Homo*-only sample, results generated from a mixed *Homo/Australopithecus* sample, and each of these with

ER 1813 in either time interval one or time interval three. Additionally, these four results are generated using pairwise specimen comparisons (i.e., average index of pairwise difference) as well as pairwise measurement comparisons. The results can be seen in Tables 2–4.

ER 1813 IN TIME INTERVAL THREE, *HOMO*-ONLY SAMPLE

With ER 1813 placed in time interval three (as we believe it should be, for independent reasons), the pattern of variation observed in the average pairwise variance values across time intervals within our sample was significantly ordered relative to a randomly generated expected distribution. Our observation that variance increases as the time breadth increases in 11 of 16 possible comparisons (see Table 3C) was observed in just 12 of the 1000 simulated distributions, giving us an estimated functional *P*-value of approximately 0.012.

Figure 2 demonstrates this result more clearly, displaying a contour illustration of the observed distribution of variance alongside a contour illustration of the expected pattern in a randomly ordered sequence. Although the random expectation is more or less flat, with moderate variance, our observed pattern shows a trough along the diagonal with elevated values moving away from the diagonal. This saddle-shaped pattern of variance, with variance increasing as the time frame sampled increases, is exactly the pattern expected for our null hypothesis as indicated by scenario C in Figure 1. Additionally, we do not observe the pattern expected were we sampling two divergent lineages, in which case

Table 3. Matrix of average pairwise variance values, pairwise specimen comparisons.

	Time 1	Time 2	Time 3	Time 4	Time 5
(A) <i>Homo</i> -only, ER 1813 in time interval one					
Time 1	0.0053	0.0072	0.0080	0.0087	0.0168
Time 2	0.0072	0.0022	0.0050	0.0048	0.0084
Time 3	0.0080	0.0050	0.0111	0.0052	0.0121
Time 4	0.0087	0.0048	0.0052	0.0065	0.0073
Time 5	0.0168	0.0084	0.0121	0.0073	0.0123
(B) <i>Homo/Australopithecus</i> , ER 1813 in time interval one					
Time 1	0.0053	0.0112	0.0075	0.0070	0.0163
Time 2	0.0112	0.0030	0.0046	0.0038	0.0074
Time 3	0.0075	0.0046	0.0087	0.0053	0.0133
Time 4	0.0070	0.0038	0.0053	0.0055	0.0104
Time 5	0.0163	0.0074	0.0133	0.0104	0.0075
(C) <i>Homo</i> -only, ER 1813 in time interval three					
Time 1	0.0058	0.0072	0.0086	0.0106	0.0192
Time 2	0.0072	0.0022	0.0046	0.0048	0.0084
Time 3	0.0086	0.0046	0.0089	0.0043	0.0098
Time 4	0.0106	0.0048	0.0043	0.0065	0.0073
Time 5	0.0192	0.0084	0.0098	0.0073	0.0123
(D) <i>Homo/Australopithecus</i> , ER 1813 in time interval three					
Time 1	0.0058	0.0118	0.0082	0.0083	0.0181
Time 2	0.0118	0.0030	0.0044	0.0038	0.0074
Time 3	0.0082	0.0044	0.0074	0.0046	0.0114
Time 4	0.0083	0.0038	0.0046	0.0055	0.0104
Time 5	0.0181	0.0074	0.0114	0.0104	0.0075

Values in bold and bold italic represent significant results.

we would expect to observe an increasing pattern of variance moving forward from the point of species divergence (scenario B in Fig. 1).

These results are further supported by the analysis utilizing pairwise measurement comparisons (see Table 4c), where 10 of the 16 observed stepwise comparisons are in line with the expectations for a directionally evolving lineage. The sensitivity of our test is increased using pairwise measurement data owing to the larger available dataset (>10,700 homologous measurement comparisons, vs. >200 pairwise specimen comparisons), but the same pattern of results is produced. The observed pattern was found in only one of 1000 simulated distributions, providing a functional P -value of $P = 0.001$ (Table 2).

ER 1813 IN TIME INTERVAL THREE, MIXED HOMO/AUSTRALOPITHECUS SAMPLE

Results based on pairwise specimen comparisons showed clear differentiation between the *Homo*-only analysis and the mixed *Homo/Australopithecus* analysis (Table 2). The pattern of variance in the mixed *Homo/Australopithecus* sample with ER 1813 in time interval three shows a more chaotic pattern of variation, with a reduction in the number of stepwise variance comparisons con-

Table 4. Matrix of average pairwise variance values, pairwise measurement comparisons.

	Time 1	Time 2	Time 3	Time 4	Time 5
(A) <i>Homo</i> -only, ER 1813 in time interval one					
Time 1	0.0183	0.0168	0.0208	0.0242	0.0423
Time 2	0.0168	0.0114	0.0136	0.0198	0.0316
Time 3	0.0208	0.0136	0.0379	0.0190	0.0270
Time 4	0.0242	0.0198	0.0190	0.0247	0.0259
Time 5	0.0423	0.0316	0.0270	0.0259	0.0357
(B) <i>Homo/Australopithecus</i> , ER 1813 in time interval one					
Time 1	0.0183	0.0374	0.0206	0.0267	0.0448
Time 2	0.0374	0.0364	0.0135	0.0286	0.0354
Time 3	0.0206	0.0135	0.0371	0.0247	0.0293
Time 4	0.0267	0.0286	0.0247	0.0259	0.0300
Time 5	0.0448	0.0354	0.0293	0.0300	0.0243
(C) <i>Homo</i> only, ER 1813 in time interval three					
Time 1	0.0246	0.0168	0.0204	0.0279	0.0529
Time 2	0.0168	0.0114	0.0115	0.0198	0.0316
Time 3	0.0204	0.0115	0.0234	0.0183	0.0257
Time 4	0.0279	0.0198	0.0183	0.0247	0.0259
Time 5	0.0529	0.0316	0.0257	0.0259	0.0357
(D) <i>Homo/Australopithecus</i> , ER 1813 in time interval three					
Time 1	0.0246	0.0351	0.0204	0.0293	0.0532
Time 2	0.0351	0.0364	0.0116	0.0286	0.0354
Time 3	0.0204	0.0116	0.0249	0.0232	0.0283
Time 4	0.0293	0.0286	0.0232	0.0259	0.0300
Time 5	0.0532	0.0354	0.0283	0.0300	0.0243

Values in bold and bold italic represent significant results.

sistent with a single evolving lineage (see Table 3D). When this sample is tested against expectations generated from randomly reordering the data, we fail to reject a pattern of temporal randomness in our data, with an estimated functional P -value of approximately $P = 0.11$. This is the case despite the fact that the *Au. boisei* sample is relatively modest, including no representatives from time interval one, a single individual from time interval two and just two individuals from time interval three. The paucity of available comparative specimens for *Au. boisei* thus limits the ability of including these fossils to disrupt the pattern observed in the sample of *Homo*, making this approach conservative with respect to our hypothesis. Nevertheless, adding this sample is sufficient to distinguish between a positive and negative result for our test.

Looking instead at the pairwise measurement data, we continue to differentiate the results from the mixed *Homo/Australopithecus* data and the *Homo*-only sample. Although we observe a weakly significant difference away from a pattern of randomness in the mixed sample ($P = 0.04$), this value is distinguishable from the highly significant result observed in the *Homo* sample ($P = 0.001$, see Table 2). Our test appears to successfully identify a time-dependent pattern in the early *Homo*

data, whereas it fails to identify such a pattern in the mixed *Homo/Australopithecus* data, suggesting the test we propose can effectively discriminate between competing evolutionary explanations.

ER 1813 IN TIME INTERVAL ONE, COMPARISONS WITHIN *HOMO*

With ER 1813 placed in time interval one, the same relationship between the results for the *Homo*-only analysis and mixed *Homo/Australopithecus* analysis is observed, although the strength of the observed pattern is weakened (Table 3A, B). The estimated P -value in comparison with a randomly ordered sequence is approximately $P = 0.1$ for the *Homo*-only sample, with the observed pattern of variation present in 100 of the 1000 simulated results. Counterintuitively, the change in result with the differing time position of ER 1813 is largely the result of an increased level of variation observed within time interval three in the absence of ER 1813. Although ER 1813 does not change the overall level of variability observed in time interval one, where it closely matches OH 24 in many of its preserved metrics, its removal from time interval three leaves that subsample dominated by comparisons between the large, well-preserved ER 3733 specimen and the more fragmentary remains of the diminutive OH 13 and enigmatic (though small) ER 1805.

When pairwise measurement, rather than pairwise specimen data are examined, the significant difference between the observed pattern of variation and expectations under a model of randomness is more apparent. The observed pattern of pairwise measurement variation is found only one time in 1000 simulated results (Table 2).

ER 1813 IN TIME INTERVAL ONE, MIXED *HOMO/AUSTRALOPITHECUS* SAMPLE

Just as with the placement of ER 1813 in time interval three, when ER 1813 is placed in time interval one a clear distinction is visible between the results of the analysis of only *Homo* and the mixed *Homo/Australopithecus* sample. The pairwise specimen data produce a nonsignificant result, with 220 of 1000 simulations matching or exceeding the observed value. The measurement pairwise data produce a significant result ($P = 0.03$), but not nearly as significant as that observed in utilizing the *Homo*-only data. For both sets of data, there is separation between the *Homo*-only results and those from the mixed *Homo/Australopithecus* sample. An additional important observation regarding the *Homo* versus *Homo/Australopithecus* comparisons is that the overall level of variation, as measured by average variance of pairwise comparisons, is substantially elevated (~16–18%) in the mixed taxa comparison.

These results provide statistical strength for the inference that time is an important explanatory variable for the pattern of

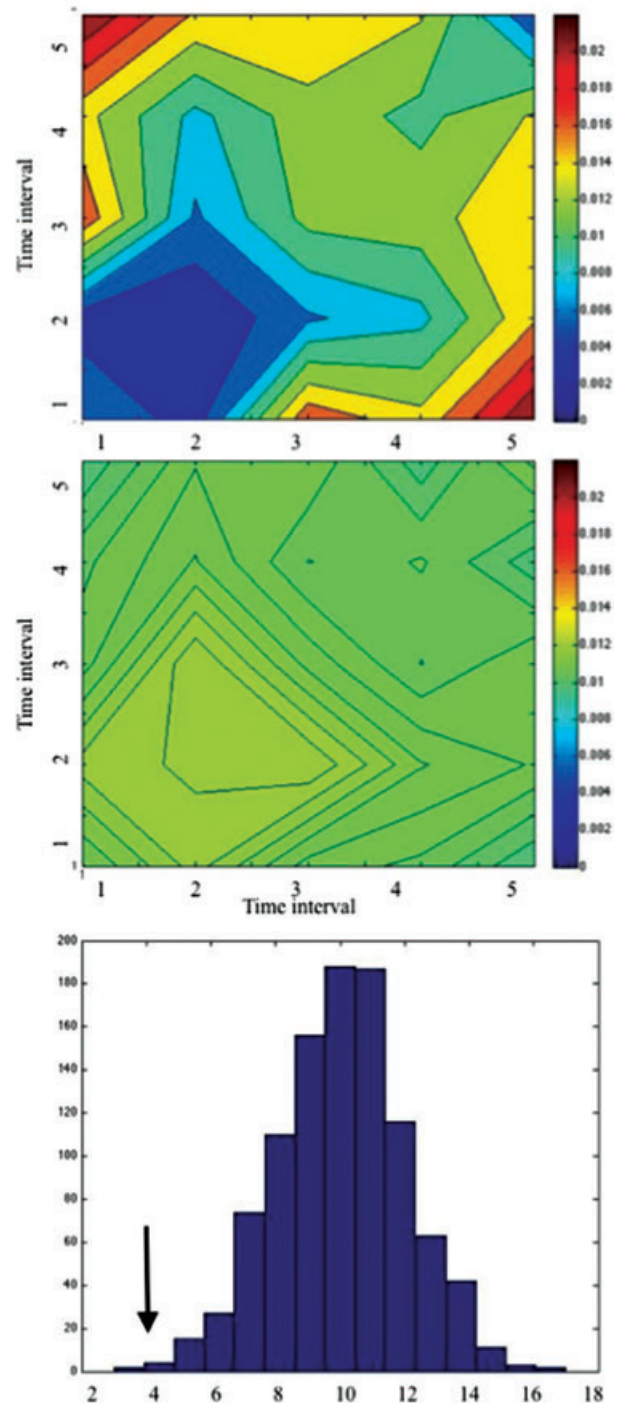


Figure 2. Observed (top) and expected pattern (middle) in the distribution of pairwise IRD value variance. The observed pattern shows low values across the central diagonal (lower-left to upper-right), representing within time interval variance, with elevated values moving away from the diagonal, representing between time intervals of increasing separation. The expected distribution in these values, assuming no time patterning to the data, is a flat distribution with values close to the global sample mean. The histogram (bottom) shows the distribution of expected matrix transitions not consistent with a single evolving lineage (see text), with the arrow indicating where our observed value lies.

change in cranial measurements within our sample. Additionally, the time-dependent pattern that is observed within these cranial remains is consistent with the hypothesis of a single lineage experience directional, evolutionary change.

Discussion

Our results support and significantly broaden the conclusions reached by Suwa and colleagues (2007), because our sample and its geographic range are larger, and because our results are supported by statistical analysis. We show that the hypothesis of a single, evolving lineage cannot be rejected for the 1.9–1.5 Ma sequence of fossils in East Africa and Georgia assigned to *Homo*. The key finding is that the pattern of cranial variation is consistent with expectations for an evolutionary lineage experiencing directional evolutionary change through time. That we are able to observe this pattern, despite using as broad a dataset as possible, allowing for uncertainty in the date of a key specimen, and preserving the ability to reject our null hypothesis even when adding a few, closely related and fragmentary *Au. boisei* specimens, provides strong support in favor of the null hypothesis of a single lineage.

Our data show a robust pattern—largely reflecting an increase in size of the neurocranium, a reduction of the masticatory structures, and related changes to the cranial base, splanchnocranium and cranial vault—throughout our sequence and a corresponding pattern of increase in the variance of pairwise samples as the time breadth being sampled increases. These are not unexpected evolutionary trends; what does seem unexpected is that they are evinced by the entire sample of *Homo*, which thereby fits the description of a species lineage. These results suggest a strong degree of temporal size patterning that extends throughout the entirety of our sequence and across different aspects of the cranium.

Our findings are not consistent with a pattern of sister species evolving away from a recent common ancestor. The notion of parallel evolving lineages, in turn, is both less parsimonious than that of a single evolving lineage and also fails to fit the data (see Figs. 1 and 2).

Our results are not without complexity, much as the evolutionary notion of linearity should not imply simplicity. For example, our analysis supports much prior research that sexual dimorphism in this hominid sample is large, perhaps approaching levels observed in *Gorilla* or *Pongo*, and not human-like, and that the *Homo* lineage evolves through time and the evolutionary change is directional, particularly associated with an expanding neurocranium and reduction of the masticatory apparatus throughout this time sequence (Lee and Wolpoff 2003). The observation of elevated levels of sexual dimorphism in early *Homo* is consistent with the work of other researchers, even those who support

concurrent lineages of *Homo* for this time period (Spoor et al. 2007).

Regardless of the position of ER 1813, we note that several deviations from the expected pattern of variation are associated with comparisons that include the subsample of specimens from time interval two. In particular, these comparisons are associated with a reduction in observed variation relative to the expectations generated from other time interval subsamples. This time interval includes Dmanisi and two important, but fragmentary, East African specimens (ER 1590 and OH 16). The fragmentary nature of the two African specimens in this sample means the sample at this particular interval is dominated by the single locality of Dmanisi. The stratigraphic setting of the Dmanisi remains is much more constrained than contemporary African deposits (Lordkipanidze et al. 2007), and is possibly sampling a reduced range of temporal and geographic variation. As such, the sample provides a fantastic window into variation at this time, but might underestimate the expected level of variation generated when compared with more geographically and temporally dispersed samples.

Additionally, we are sympathetic to the possibility that the reduced variation in this time interval may reflect the fact that it is possibly the only sample that fails to include a male splanchnocranium. If so, the magnitude of cranial sexual dimorphism is underestimated (Rosas et al. 2002). It should be noted that the Dmanisi sample does include a mandible, D2600, which expresses dramatic size differences, probably as a result of sexual dimorphism, with the remainder of the mandibular sample (Van Arsdale 2006; Rightmire et al. 2008; Van Arsdale and Lordkipanidze in press). Given the large size of D2280 relative to the remainder of the Dmanisi sample and the presence of the hyper-robust Dmanisi 2600 mandible, the presence of an adult male face from the site would quite likely, in our estimation, result in a pronounced increase in the variation observed for this time interval. These two factors may account for the reduced level of variation within this time interval; we propose this as a testable hypothesis that can be addressed in future studies.

In sum, our findings are based on statistical analysis of the largest sample of crania from the 1.9 to 1.5 Mya time period analyzed to date. They address the hull hypothesis of a single species lineage of *Homo* and do not disprove it. The implications of this finding will be explored in subsequent papers, and independent analysis of the mandibular and dental remains is now required. Thus, we believe our results provide strong motivation for further analyses of this key time period of human evolution.

Our conclusions imply an acceptance of a large amount of variation within this lineage. As suggested above, an elevated level of sexual dimorphism might account for some or all of this variation, but it is also possible that this elevated variation

is an evolutionary reality of early *Homo* and reflects the mosaic ecological transition from *Australopithecus* to *Homo*, as the latter adapted to a more technologically mediated subsistence strategy with significant interpopulation variation. These considerations also suggest areas of potentially productive future research. The conclusions of this study do not, in our view, represent a 40-year regression to the time period in which Robinson and Tobias were first debating the evolutionary status of *H. habilis*. Rather, they represent the development of a productive and new way of conceptualizing the complexity of linear evolution in the early members of the genus *Homo* spurred by the great amount of work and fossils that have been recovered since that time.

ACKNOWLEDGMENTS

This work would not be possible without the assistance of many friends, colleagues, and curators. We would like to make special note of the assistance of all of the curators of the fossil material examined by the authors, making special note of the assistance from D. Lordkipanidze, T. Jashashvili, and A. Margvelashvili at the Georgia National Museum, and E. Mbua at the National Museum of Kenya. Two anonymous reviewers, the editor, and J. DeSilva provided invaluable feedback in the preparation of our final manuscript. Finally, we would like to thank the many colleagues and students who engaged us in conversation and debate on this issue over the past several years.

LITERATURE CITED

- Antón, S. C. 2003. Natural history of *Homo erectus*. *Yearb. Phys. Anthropol.* 46:126–170.
- . 2004. The face of Olduvai Hominid 12. *J. Hum. Evol.* 46:337–347.
- Asfaw, B., W. H. Gilbert, Y. Beyene, W. K. Hart, P. R. Renne, G. Wolde Gabriel, E. S. Vrba, and T. D. White. 2002. Remains of *Homo erectus* from Bouri, Middle Awash, Ethiopia. *Nature* 416:317–320.
- Baab, K. L. 2008. A re-evaluation of the taxonomic affinities of the early *Homo* cranium KNM-ER 42700. *J. Hum. Evol.* 55:741–746.
- Bilsborough, A. 2000. Chronology, variability and evolution in *Homo erectus*. *Variability Evol.* 8:5–30.
- Boaz, N. T. 1979. Hominid evolution in Eastern Africa during the Pliocene and early Pleistocene. *Annu. Rev. Anthropol.* 8:71–85.
- Donnelly, S. M. 1996. Size and shape variation in early East African *Homo* and the validity of *Homo habilis* as a single species. Ph.D. thesis, University of Tennessee, Knoxville.
- Feibel, C., F. Brown, and I. McDougall. 1989. Stratigraphic context of fossil hominids from the Omo group deposits: northern Turkana Basin, Kenya and Ethiopia. *Am. J. Phys. Anthropol.* 78:595–622.
- Feibel, C. S., C. J. Lepre, and R. L. Quinn. 2009. Stratigraphy, correlation, and age estimates for fossils from Area 123, Koobi Fora. *J. Hum. Evol.* 57:112–122.
- Gabunia, L., M. A. de Lumley, A. Vekua, D. Lordkipanidze, and H. de Lumley. 2002. Discovery of a new hominid at Dmanisi (Transcaucasia, Georgia). *Comptes Rendus Palevol* 1:243–253.
- Gathogo, P. N., and F. Brown. 2006. Revised stratigraphy of Area 123, Koobi Fora, Kenya, and new age estimates of its fossil mammals, including hominins. *J. Hum. Evol.* 51:471–479.
- Gathogo, P. N., F. H. Brown, and I. McDougall. 2008. Stratigraphy of the Koobi Fora Formation (Pliocene and Pleistocene) in the Loiyangalani region of northern Kenya. *J. Afr. Earth Sci.* 51:277–297.
- Grine, F. E., W. L. Jungers, and J. Schultz. 1996. Phenetic affinities among early *Homo* crania from East and South Africa. *J. Hum. Evol.* 30:189–225.
- Johanson, D. C., F. T. Masao, G. G. Eck, T. D. White, R. C. Walter, W. H. Kimbel, B. Asfaw, P. Manega, P. Ndessokia, and G. Suwa. 1987. New partial skeleton of *Homo habilis* from Olduvai gorge, Tanzania. *Nature* 327:205–209.
- Kramer, A. 1993. Human taxonomic diversity in the Pleistocene—does *Homo erectus* represent multiple Hominid species. *Am. J. Phys. Anthropol.* 91:161–171.
- Leakey, L. S. B., J. R. Napier, and P. V. Tobias. 1964. New species of genus *Homo* from Olduvai Gorge. *Nature* 202:7–9.
- Leakey, M. D., R. J. Clarke, and L. S. B. Leakey. 1971. New hominid skull from Bed I, Olduvai Gorge, Tanzania. *Nature* 232:308–312.
- Leakey, M. G., F. Spoor, M. C. Dean, C. S. Feibel, S. C. Antón, C. Kiarie, and L. N. Leakey. 2012. New fossils from Koobi Fora in northern Kenya confirm taxonomic diversity in early *Homo*. *Nature* 488:201–204.
- Lee, S.-H., and M. H. Wolpoff. 2003. The pattern of evolution in Pleistocene human brain size. *Paleobiology* 29:186–196.
- Lieberman, D. E., D. R. Pilbeam, and B. A. Wood. 1988. A probabilistic approach to the problem of sexual dimorphism in *Homo habilis*: a comparison of KNM-ER 1470 and KNM-ER 1813. *J. Hum. Evol.* 17:503–511.
- Lordkipanidze, D., A. Vekua, R. Ferring, J. Agustí, G. Kiladze, A. Mouskhelishvili, M. Nioradze, M. Ponce de Leon, M. Tappen, and C. Zollikofer. 2005. The earliest toothless hominin skull. *Nature* 434:717–718.
- Lordkipanidze, D., A. Vekua, R. Ferring, G. P. Rightmire, C. Zollikofer, M. Ponce de Leon, J. Agustí, G. Kiladze, A. Mouskhelishvili, M. Nioradze, et al. 2006. A fourth hominin skull from Dmanisi, Georgia. *Anat. Rec.* 288A:1146–1157.
- Lordkipanidze, D., T. Jashashvili, A. Vekua, M. Ponce de Leon, C. Zollikofer, G. P. Rightmire, H. Pontzer, R. Ferring, O. Oms, M. Tappen, et al. 2007. Postcranial evidence from early *Homo* from Dmanisi, Georgia. *Nature* 449:305–310.
- Miller, J. A. 1991. Does brain size variability provide evidence of multiple species in *Homo habilis*? *Am. J. Phys. Anthropol.* 84:385–398.
- . 2000. Craniofacial variation in *Homo habilis*: an analysis of the evidence for multiple species. *Am. J. Phys. Anthropol.* 112:103–128.
- Miller, J., G. H. Albrecht, and B. R. Gelvin. 2004. Craniomeric variation in early *Homo* compared to modern gorillas: a population-thinking approach. *Camb. Stud. Biol. Evol. Anthropol.* 40:66–98.
- Rightmire, G. P. 1979. Cranial remains of *Homo erectus* from Beds II and IV, Olduvai Gorge, Tanzania. *Am. J. Phys. Anthropol.* 51:99–115.
- . 1990. The evolution of *Homo erectus*: comparative anatomical studies of an extinct human species. Cambridge Univ. Press, New York.
- . 1993. Variation among early *Homo* Crania from Olduvai Gorge and the Koobi Fora Region. *Am. J. Phys. Anthropol.* 90:1–33.
- Rightmire, G. P., D. Lordkipanidze, and A. Vekua. 2006. Anatomical descriptions, comparative studies and evolutionary significance of the hominin skulls from Dmanisi, Republic of Georgia. *J. Hum. Evol.* 50:115–141.
- Rightmire, G. P., A. P. Van Arsdale, and D. Lordkipanidze. 2008. Variation in the mandibles from Dmanisi, Georgia. *J. Hum. Evol.* 54:904–908.
- Robinson, J. T. 1965. *Homo habilis* and the australopithecines. *Nature* 205:121–124.
- . 1966. The distinctiveness of *Homo habilis*. *Nature* 209:957–960.

- Rosas, A., M. Bastir, C. Martinez-Maza, and J. M. Bermudez de Castro. 2002. Sexual dimorphism in the Atapuerca-SH hominids: the evidence from the mandibles. *J. Hum. Evol.* 42:451–474.
- Spoor, F., M. G. Leakey, P. N. Gathogo, F. H. Brown, S. C. Antón, I. McDougall, C. Kiarie, F. K. Manthi, and L. N. Leakey. 2007. Implications of new early *Homo* fossils from Ileret, east of Lake Turkana, Kenya. *Nature* 448:688–691.
- Stringer, C. 1986. The credibility of *Homo habilis*. Pp. 266–294 in B. Wood, L. B. Martin and P. Andrews, eds. *Major topics in primate and human evolution*. Alan R. Liss, New York.
- Suwa, G., B. Asfaw, Y. Haile-Selassie, T. D. White, S. Katoh, G. Wolde Gabriel, W. K. Hart, H. Nakaya, and Y. Beyene. 2007. Early Pleistocene *Homo erectus* fossils from Konso, southern Ethiopia. *Anthropol. Sci.* 133–151.
- Tobias, P. V. 1964. Olduvai Bed I Hominine with special reference to its cranial capacity. *Nature* 202:3–7.
- . 1991. *Olduvai Gorge*. Vol. 4. Cambridge Univ. Press, Cambridge.
- . 2003. Encore Olduvai. *Science* 299:1193–1194.
- Trinkaus, E. 1984. Does KNM-ER 1481A establish *Homo erectus* at 2.0 myr BP? *Am. J. Phys. Anthropol.* 64:137–139.
- Van Arsdale, A. P. 2006. Mandibular variation in early *Homo* from Dmanisi, Georgia. Ph.D. thesis, University of Michigan, Ann Arbor.
- Van Arsdale, A. P., and D. Lordkipanidze. 2012. A quantitative assessment of mandibular variation in the Dmanisi hominins. *Paleoanthropology*. 2012:134–144.
- Vekua, A., D. Lordkipanidze, G. P. Rightmire, J. Agusti, R. Ferring, G. Maisuradze, A. Mouskhelishvili, M. Nioradze, M. Ponce de Leon, M. Tappen, et al. 2002. A new skull of early *Homo* from Dmanisi, Georgia. *Science* 297:85–89.
- Wood, B. A. 1985. Early *Homo* in Kenya, and its systematic relationships. Pp. 206–214 in E. Delson, ed. *Ancestors: the hard evidence*. Alan R. Liss, New York.
- . 1991. *Koobi Fora Research Project*. Vol. 4. Hominid Cranial Remains. Clarendon Press, Oxford.
- . 1992. Early hominid species and speciation. *J. Hum. Evol.* 22:351–365.
- . 1993. Early *Homo*: how many species? Pp. 485–522 In W. H. Kimbel and L. B. Martin, eds. *Species, species concepts, and primate evolution*. Plenum, New York.
- Wood, B., and B. G. Richmond. 2000. Human evolution: taxonomy and paleobiology. *J. Anat.* 197:19–60.

Associate Editor: D. Carrier

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1

Appendix S2