An enlarged postcranial sample confirms *Australopithecus afarensis* dimorphism was similar to modern humans

Philip L. Reno[1],*, Melanie A. McCollum[3,4], Richard S. Meindl[2,3] and C. Owen Lovejoy[2,3],*

[1]*Department of Developmental Biology, Stanford University School of Medicine, 279 Campus Drive, Beckman 300, Stanford, CA 94305-5329, USA*
[2]*Department of Anthropology, and* [3]*School of Biomedical Sciences, Kent State University, Lowry Hall Room 226, Kent, OH 44242-0001, USA*
[4]*Department of Cell Biology, University of Virginia, Charlottesville, VA 22902, USA*

In a previous study, we introduced the template method as a means of enlarging the *Australopithecus afarensis* postcranial sample to more accurately estimate its skeletal dimorphism. Results indicated dimorphism to be largely comparable to that of *Homo sapiens*. Some have since argued that our results were biased by artificial homogeneity in our *Au. afarensis* sample. Here we report the results from inclusion of 12 additional, newly reported, specimens. The results are consistent with those of our original study and with the hypothesis that early hominid demographic success derived from a reproductive strategy involving male provisioning of pair-bonded females.

Keywords: A.L. 333; taphonomy; monogamy; skeletal dimorphism; modelling

1. INTRODUCTION

Accurately inferring early hominid sexual dimorphism is an important element in interpreting their paleobiology. We previously concluded that skeletal size dimorphism in *Australopithecus afarensis* was significantly lower than that of gorillas and could not be statistically distinguished from that of modern humans (Reno *et al.* 2003, 2005). These findings, which contrast with previous assessments (Zihlman & Tobias 1985; McHenry 1991; Lockwood *et al.* 1996), were achieved through the use of the 'template method'. This method relied on the A.L. 288-1 partial skeleton ('Lucy'), as a source of simple ratios between femoral head diameter (FHD) and other skeletal dimensions. These ratios were then used to obtain estimates of FHD for skeletal dimensions that were also measurable in A.L. 288-1. Postcranial variation within the (thus maximized) *Au. afarensis* sample from the Middle Awash region of Ethiopia ('Combined Afar', CA) and that within the temporally and geographically constricted *Au. afarensis* sample from Afar Locality 333 were then compared with bootstrapped samples of modern humans, chimpanzees and gorillas. This method was specifically designed to overcome problems inherent in calculating sexual dimorphism from a small number of specimens whose sexes must be judged *a priori* on the basis of size (e.g. Zihlman & Tobias 1985; McHenry 1991).

Any given assemblage of *Au. afarensis* fossils was formed by a combination of random sampling and various taphonomic processes. The effect of these processes on sample variation can be modelled by bootstrapping from taxa of known dimorphism. Humans (*Homo sapiens*), chimpanzees (*Pan troglodytes*) and gorillas (*Gorilla gorilla*) represent the three genera most closely related to early hominids and essentially encompass the entire range of primate skeletal sexual dimorphism. Because the sex of any *Au. afarensis* element is essentially unknown, sampling with regard to sex of extant taxa is allowed to vary freely. That is, the sex ratio in each iteration is allowed to vary by simple probability (i.e. the binomial expansion). In sufficiently small samples, this can occasionally result in samples composed of only one sex. In order to simulate the *Au. afarensis* assemblages as precisely as possible (and limit the variation introduced by sampling different anatomical locations), bootstrapped samples of living hominoids were required to exactly match the anatomical compositions of the A.L. 333 and the CA samples (e.g. the number of proximal tibias included in each bootstrapped sample was required to exactly match the number of proximal tibias represented in the *Au. afarensis* sample being simulated). For each iteration, each postcranial metric was converted to a FHD based on ratios calculated from a template specimen that was also randomly chosen to serve as the equivalent of A.L. 288-1.

A.L. 333 probably represents a simultaneous death assemblage (White & Johanson 1989; Behrensmeyer *et al.* 2003). In our previous analysis (Reno *et al.* 2003), two separate simulations of A.L. 333 were generated. In one, each of 22 postcranial

* Authors for correspondence (preno@stanford.edu; olovejoy@aol.com).

metrics preserved at the site was randomly drawn from our complete samples of extant taxa ($N \approx 50$). This exactly modelled A.L. 333 in being composed of as many as 22 separate individuals. However, it is unlikely that each A.L. 333 specimen in fact represents one of 22 different individuals. Based on mandibular dentitions, the minimum number of individuals (MNI) at the site is nine (White & Johanson 1989). Therefore, in a second simulation, we randomly selected nine individuals to serve as the source of all 22 metrics. This ensures that many individuals are multiply represented in the sample of metrics. Our procedures assume only that each 'death' assemblage (fossil sample or extant simulation) was a random sample of its parent population—the biological species from which each was derived (i.e. *Au. afarensis, H. sapiens, P. troglodytes* and *G. gorilla*).

These samples have been challenged as not being representative of the *Au. afarensis* size distribution (Plavcan *et al.* 2005; Scott & Stroik 2006). The rationale has been that because 'Lucy-sized' individuals are absent from the A.L. 333 assemblage, it must over-sample large, presumably male, adult individuals. If true, then our lower estimates of skeletal dimorphism in A.L. 333 may have been flawed by biased sampling during the accumulation, fossilization and/or recovery of the assemblage.

This argument is now subject to a simple test. Additional postcranial elements of *Au. afarensis* have now been reported for both A.L. 333 and other Middle Awash localities (Kimbel *et al.* 2004; Drapeau *et al.* 2005; Harmon 2006). Inclusion of these additional 12 specimens raises our postcranial sample to 41 and more than doubles the number of individuals represented from non-A.L. 333 localities. This expanded sample provides an opportunity to more accurately assess size dimorphism in *Au. afarensis* and determine whether smaller Lucy-sized individuals were disproportionately lacking from A.L. 333.

2. MATERIAL AND METHODS

In addition to the 29 fossil specimens in our original study (Reno *et al.* 2003), we have now added four specimens from A.L. 333 and eight from other Middle Awash localities (Kimbel *et al.* 2004; Drapeau *et al.* 2005; Harmon 2006; table 1). Metrics from these specimens (some as yet undescribed), as well as their homologues from the A.L. 288-1 partial skeleton, were provided by William Kimbel.

Details of the template method and our bootstrapping procedures are described in Reno *et al.* (2003, 2005). Since those publications, we have observed that the template method yields extreme FHD estimates in rare cases where a small or large skeletal metric is paired with a template specimen with an unusual metric to FHD ratio (all within the bounds of natural variation). While such pairings are infrequent, they nevertheless have the potential to confound results by artificially inflating dimorphism statistics in extant samples. As a means of correcting bias from such cases, we now systematically discard estimated FHDs that are more than 10 mm above or below the

observed range of the extant taxon being sampled. However, this correction is potentially quite conservative as the relative size range of many metrics is actually greater than that of FHD (see below).

The three bootstrap simulations reported here were performed separately to model the following enlarged *Au. afarensis* samples: (i) 26 specimens from the A.L. 333; (ii) 15 non-333 specimens from other Hadar localities and Maka; and (iii) 41 specimens in the CA sample. The 15 specimens in the non-333 sample must represent 15 separate individuals. Therefore, 15 metrics were each randomly drawn from the entire chimpanzee, human or gorilla samples. In contrast, it is unlikely that 26 different individuals contributed to the A.L. 333 sample. Therefore, for each iteration, a separate random subsample of nine chimpanzee, human or gorilla individuals (based upon an MNI from mandiblar dentitions (White & Johanson 1989)) served as the pool from which 26 metrics were then drawn. For the CA simulations, a 'hybrid' was created for each iteration in which 26 metrics representing A.L. 333 were sampled from nine randomly selected individuals. These were combined with an additional 15 drawn from the entire sample to represent non-333 individuals.

Plavcan *et al.* (2005) argued that only five to eight individuals contributed to the A.L. 333 postcranial sample, and it is certainly hypothetically possible that some of the (at least) nine known adult individuals did not contribute to the postcranial sample. However, our simulations already permit sampling of fewer than nine individuals because not all nine individuals selected for each iteration are necessarily randomly sourced for the 26 metrics used to simulate A.L. 333. Therefore, it was unnecessary to perform additional simulations from isolated comparative samples artificially restricted to less than nine *potential* contributors.

Both the coefficient of variation (CV) and the binomial dimorphism index (BDI) were calculated in each simulation. The CV was calculated using the small sample correction (Sokal & Rohlf 1995). The BDI was defined specifically for the calculation of sexual dimorphism in samples of unknown sex. It rests upon three assumptions: (i) both sexes are present in each sample; (ii) every specimen has an equal probability of being male or female, but (iii) when any two specimens are potentially of a different sex, the larger is always male. Using this algorithm, a sample of n yields a total of $n - 1$ possible sex allocations and therefore $n - 1$ skeletal dimorphism estimates. The BDI is then the *weighted* average of the $n - 1$ dimorphism values based on the probability of each sex allocation occurring under the binomial expansion. Note that in light of the assumption that males are always larger than females, the BDI tends to overestimate dimorphism in minimally dimorphic species.

Two estimates of dimorphism (actual DM: male mean/female mean based on known sex) were calculated for each extant sample drawn. The first used estimated FHD dimensions estimated for each specimen by the template method (template sexual dimorphism: TSD) to measure dimorphism, and the second used the original FHDs for each randomly

Table 1. *Australopithecus afarensis* sample used for simulations.

metric	specimen(s)	estimated FHD	estimated GMEAN	FHD/GMEAN ratio
HHD: max. diameter of humeral head	A.L. 333-107[a]	39.4	32.6	1.21
OLCB: ML width of humerus measured tangent to the superior margin of the olecranon fossa	A.L. 137-48A	32.6	27.0	1.21
	A.L. 137-50[b]	38.3	31.6	1.21
	A.L. 223-23[b]	35.3	29.2	1.21
	A.L. 322-1	27.9	23.1	1.21
	A.L. 333-29	33.2	27.4	1.21
	A.L. 333w-31	34.3	28.4	1.21
	Mak VP 1/3	37.8	31.3	1.21
CAPD: max. diameter of capitulum	A.L. 333w-22	39.5	32.7	1.21
	A.L. 444-14[b]	37.2	30.4	1.22
RHD: max. diameter of the radial head	A.L. 333x-14[c]	44.3	36.6	1.21
	A.L. 333x-15[c]	44.5	36.8	1.21
ULB: ML width of ulna immediately distal to radial facet	A.L. 333x-5	37.1	30.6	1.21
	A.L. 333w-36	29.8	24.6	1.21
	A.L. 438-1a[b]	40.9	33.8	1.21
FHD: max. diameter of femoral head	A.L. 152-2[b]	33.1	27.4	1.21
	A.L. 288-1ap	28.6	23.6	1.21
	A.L. 333-3	40.9	33.8	1.21
	A.L. 827[b]	38.1	31.5	1.21
TRCD: max. femoral shaft diameter immediately below lesser trochanter	A.L. 211-1	36.4[d]	30.1	1.21
	A.L. 333-95[c]	35.3[d]	29.1	1.21
	Mak VP 1/1	34.4[d]	28.4	1.21
FNKH: femoral neck height normal to long axis at midpoint	A.L. 333-117	38.7	32.0	1.21
	A.L. 333-123[b]	33.0	27.2	1.21
	A.L. 333-142[b]	30.1	24.8	1.21
GSTB: AP femoral width immediately above gastrocnemius tubercles	A.L. 333-4	35.2	29.1	1.21
	A.L. 333w-56	33.6	27.8	1.21
	A.L. 333-140[b]	30.2	24.9	1.21
CNDC: ML distance between centers of medial and lateral tibial condyles	A.L. 129-1b	27.9	23.0	1.21
	A.L. 333x-26	38.5	31.8	1.21
	A.L. 333-42	36.7	30.3	1.21
PRXTB: max. ML tibial bicondylar breadth	A.L. 330-6[b]	37.4	30.9	1.21
DSTTB: AP articular length at ML mid-point of articular surface of distal tibia	A.L. 333-6	37.2	30.8	1.21
	A.L. 333-7	42.9	35.5	1.21
	A.L. 333-96	38.4	31.7	1.21
	A.L. 545-3[b]	31.9	26.1	1.22
FIBD: max. ML diameter of distal fibula	A.L. 333-9A	42.8	35.4	1.21
	A.L. 333-9B	38.9	32.1	1.21
	A.L. 333w-37	37.8	31.3	1.21
	A.L. 333-85	40.6	33.5	1.21
TAL: max. AP length of talus	A.L. 333-147[b]	36.0	29.8	1.21

[a]Because of slight eccentricity in this specimen the average of the mediolateral (ML) and anteroposterior (AP) diameters was used instead.
[b]Specimens new to this analysis.
[c]These specimens lack epiphyseal fusion and are not strictly adults. They were included because they constitute three of the largest fossils in the sample and their omission would further *decrease* fossil dimorphism estimates.
[d]These values are based on a slightly different metric than the one used in the previous analysis (AP subtrochanteric diameter reported by Lovejoy *et al.* (1982)) and therefore differ from those reported in Reno *et al.* (2003). The new metric corresponds more closely with that taken from the comparative samples and better reflects the relative size of these specimens (i.e. 333-3 is clearly larger than any of these three specimens).

selected individual (direct sexual dimorphism: DSD) to determine it. Comparison of TSD and DSD assesses the effect of using a template specimen to estimate dimorphism. Because calculation of dimorphism statistics (BDI and CV) for *Au. afarensis* requires use of a template specimen (A.L. 288-1), these can be assessed only by comparison with TSD produced by the simulations.

3. RESULTS
Table 1 lists each *Au. afarensis* specimen used in the current analysis, the metric from which its FHD was

calculated using the template (A.L. 288-1) and the resulting estimated FHD. Also included in table 1 are estimated geometric means (GMEAN) of all included metrics that could also be calculated by the template method in addition to the FHD. These are included to illustrate that the results of the template method do not depend on the choice of FHD to measure sample dimorphism. Note that ratios between estimated FHD and estimated GMEAN are always identical. Thus, any measure of sample variation (i.e. CV or BDI) will be identical and any scalar metric from the template will produce the same result. Therefore, any species-specific allometric
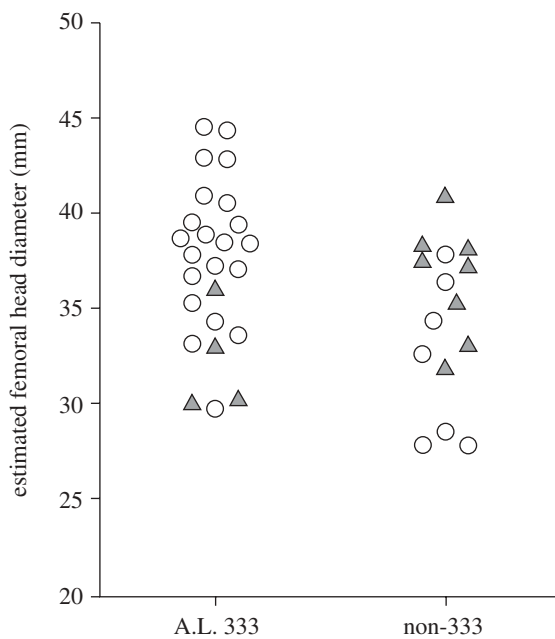
Figure 1. Estimated FHD for individual *Au. afarensis* specimens included in this analysis. Circles, original specimens; triangles, specimens new to this analysis.

relationships with FHD have no effect on the outcome of the procedure.

Estimated FHDs for each of the original 29 specimens of Reno *et al.* (2003) plus the 12 additional specimens included in the present analysis are shown in figure 1. As it demonstrates, the new specimens increase representation of small individuals at A.L. 333 but not to the extreme range represented by the smallest individuals of the non-333 sample. On the other hand, the new specimens expand the upper size range of the non-333 sample (although not to the extent observed in A.L. 333) and appreciably increase the representation of intermediate-sized individuals in the non-333 sample such that there is no longer any potential demarcation between large and small specimens in the combined CA sample. These novel specimens provide little reason to conclude that A.L. 333 under-represents small Lucy-sized individuals. To the contrary, given the large number of intermediate-sized specimens, it is quite possible that such extremely small individuals may actually be over-represented in the non-333 sample.

Table 2 presents samples sizes, CVs, actual DM and the BDI for each metric. Dimorphism in humans is intermediate between non-dimorphic chimpanzees and highly dimorphic gorillas for nearly all characters (only the chimpanzee capitulum (CAPD) BDI and CV are slightly greater than humans). However, within each taxon, the extent to which skeletal metrics differ between the sexes varies extensively. Significantly, variation in FHD in all three hominoid taxa is low in comparison to that observed for most other skeletal metrics (thus, FHD will have a smaller relative range). Given these findings, the template method can be expected to overestimate the means and dispersions of direct dimorphism values. Although the BDIs correlated well with the actual dimorphism observed for each metric, they tended to overestimate size dimorphism in the minimally dimorphic species (compare values in chimpanzees and gorillas; table 2). As noted above, this is an expected finding because males are always assumed to be larger than females.

Frequency histograms of dimorphism values generated by simulating the A.L. 333, CA and non-333 samples are provided in figure 2. As expected, human dimorphism values were found to be intermediate between those of chimpanzees and gorillas. Also as expected (see discussion above), template-derived size dimorphism statistics tended to overestimate direct dimorphism values (table 3). For each iteration, a Pearson correlation coefficient was computed between the resulting template-derived estimated FHDs and the directly measured FHDs. The means and standard deviations of these correlation coefficients for all simulations are presented in table 4. The strength of the correlation between template and direct values varied among species as a direct consequence of their relative dimorphism. As expected, in non-dimorphic chimpanzees, the error in estimating FHD was relatively high compared with the size range of the species. In contrast, in highly dimorphic gorillas, it was relatively low. The patterns of correlation observed between template FHD and direct FHD in the extant taxa verify that the template method satisfactorily reflects actual dimorphism levels in these samples.

As in our original analysis, BDI and CV calculated for *Au. afarensis* were most similar to those of humans (figure 2). This was true not only of the A.L. 333 and CA samples, but also for the non-333 sample. Table 5 presents exact counts of the number of iterations that fell above or below the *Au. afarensis* value in each simulation. As these data demonstrate, dimorphism within the expanded A.L. 333 sample increased from a BDI of 1.167 in our previous analysis to a value of 1.195 here, which places *Au. afarensis* dimorphism in the middle of the distribution of human values. However, because of its small sample size (modelled as representing nine individuals), it is statistically indistinguishable from any of the three hominoids.

Unlike the results for A.L. 333, dimorphism within the expanded CA sample decreased from a BDI of 1.222 in our original study to 1.209 here, a value that is significantly different from that of both the extremely dimorphic gorillas and the minimally dimorphic chimpanzees. The slightly higher dimorphism value of 1.213 calculated for the non-333 sample also differed significantly from that of gorillas using a directional test (which is appropriate considering gorillas set the upper range of primate dimorphism).

Because of its reliance on estimated rather than actual FHDs, the template method contributes an additional source of error to estimates of dimorphism. In order to ensure that this error is not a function of the size of the template specimen—a potential concern given the unusually small size of A.L. 288-1—we compared dimorphism values generated by different sized templates. Template size has no systematic effect (figure 3), and therefore our results are not biased by the small size of A.L. 288-1.

Table 2. Sample sizes and dimorphism statistics for individual metrics measured directly from chimpanzee, human and gorilla specimens. Metrics are explained in table 1.

| species	HHD	OLCB	CAPD	RHD	ULB	FHD	TRCD	FNEKH	GSTB	CNDC	PRXTB	DSTTB	FIBD	TALUS	GMEAN
chimpanzees															
male (*n*)	23	23	23	22	23	23	23	23	23	23	23	23	22	19	18
female (*n*)	25	25	25	25	25	25	25	25	25	25	25	25	25	24	24
CV	6.64	9.42	9.49	6.65	9.03	6.48	7.80	7.27	7.31	7.40	5.98	8.20	6.72	6.23	5.32
actual DM	1.066	1.073	1.090	1.038	1.059	1.049	1.054	1.032	1.029	1.036	1.041	1.003	1.052	1.029	1.043
BDI	1.115	1.163	1.165	1.113	1.155	1.114	1.133	1.123	1.126	1.127	1.103	1.147	1.110	1.106	1.093
difference	0.049	0.090	0.075	0.075	0.096	0.065	0.079	0.091	0.097	0.091	0.062	0.144	0.058	0.077	0.050
humans															
male (*n*)	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
female (*n*)	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
CV	8.92	11.08	9.17	8.87	13.07	8.81	9.64	11.00	8.67	10.17	7.73	12.07	9.67	8.02	8.02
actual DM	1.170	1.176	1.124	1.151	1.203	1.157	1.142	1.178	1.100	1.173	1.134	1.183	1.116	1.135	1.151
BDI	1.169	1.202	1.161	1.160	1.237	1.161	1.165	1.199	1.151	1.186	1.140	1.214	1.166	1.141	1.153
difference	−0.001	0.026	0.037	0.009	0.034	0.004	0.023	0.021	0.051	0.013	0.006	0.031	0.050	0.006	0.002
gorillas															
male (*n*)	24	25	25	24	25	25	25	25	25	25	25	25	23	21	19
female (*n*)	25	25	25	25	25	25	25	25	25	25	25	25	25	22	22
CV	14.92	18.09	15.02	14.60	16.74	12.66	12.66	14.00	14.98	14.23	13.62	13.80	16.06	10.98	13.29
actual DM	1.325	1.390	1.300	1.296	1.324	1.258	1.225	1.258	1.297	1.275	1.279	1.257	1.330	1.208	1.279
BDI	1.305	1.378	1.297	1.290	1.324	1.252	1.242	1.268	1.299	1.284	1.271	1.259	1.325	1.208	1.264
difference	−0.020	−0.012	−0.003	−0.006	0.000	−0.006	−0.003	0.010	0.002	0.009	−0.008	0.002	−0.005	0.000	−0.015
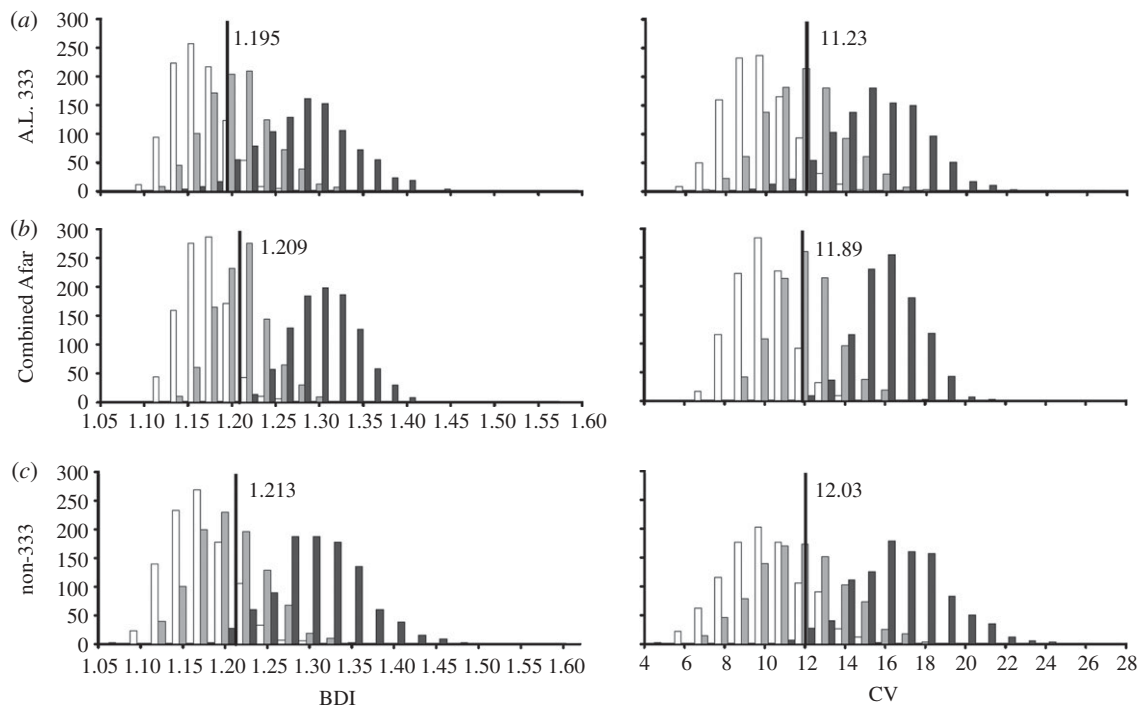
Figure 2. Frequency histograms of dimorphism values generated by simulating the (*a*) A.L. 333, (*b*) Combined Afar and (*c*) non-333 assemblages using chimpanzee (white bars), human (grey bars) and gorilla (black bars) comparative samples (1000 iterations each). The vertical line in each plot indicates dimorphism for the *Au. afarensis* sample.

Table 3. Summary statistics from each of the extant hominoid simulations.

	chimpanzee		human		gorilla	
	TSD	DSD	TSD	DSD	TSD	DSD
A.L. 333 simulation: 1000 iterations						
actual DM						
mean	1.045	1.053	1.152	1.160	1.290	1.258
s.d.	0.050	0.046	0.052	0.040	0.060	0.046
BDI						
mean	1.155	1.102	1.198	1.140	1.272	1.206
s.d.	0.029	0.027	0.037	0.036	0.053	0.051
CV						
mean	9.29	6.05	11.50	8.21	14.94	11.78
s.d.	1.59	1.31	1.90	1.74	2.26	2.07
Combined Afar simulation: 1000 iterations						
actual DM						
mean	1.049	1.051	1.155	1.159	1.297	1.258
s.d.	0.036	0.030	0.038	0.026	0.041	0.029
BDI						
mean	1.162	1.108	1.203	1.150	1.292	1.227
s.d.	0.025	0.014	0.031	0.025	0.037	0.032
CV						
mean	9.52	6.24	11.55	8.50	15.49	12.21
s.d.	1.33	0.82	1.54	1.16	1.54	1.18
non-A.L. 333: 1000 iterations						
actual DM						
mean	1.057	1.050	1.162	1.156	1.305	1.258
s.d.	0.053	0.033	0.054	0.029	0.061	0.032
BDI						
mean	1.161	1.107	1.194	1.148	1.294	1.223
s.d.	0.037	0.019	0.042	0.025	0.052	0.033
CV						
mean	9.67	6.43	11.36	8.67	16.14	12.54
s.d.	2.00	0.94	2.18	1.29	2.31	1.39

Table 4. Means and standard deviations of the correlations between actual and estimated FHD computed for each of the 1000 iterations.

	chimpanzee	human	gorilla
A.L. 333			
mean	0.453	0.620	0.819
s.d.	0.179	0.145	0.087
Combined Afar			
mean	0.498	0.642	0.831
s.d.	0.132	0.102	0.056
non-333			
mean	0.565	0.675	0.832
s.d.	0.190	0.147	0.091

4. DISCUSSION

The present study is based on dimorphism estimates generated from 41 fossils representing a minimum of 20 separate individuals. While we look forward to the potential of adding more fossils when available, it is likely that the sample is now reaching a 'critical mass' such that additional specimens are unlikely to appreciably change the dimorphism estimates. Results confirm our previous conclusions that dimorphism is only minimal to moderate in *Au. afarensis*. Skeletal variation in the CA sample differs significantly from those of gorillas and chimpanzees but cannot be statistically distinguished from that of modern humans (table 5). Significantly, the dimorphism values calculated for the non-333 sample demonstrate that the results obtained for the combined sample are not biased in any way by the composition of A.L. 333, a finding that renders moot all criticisms of our original study which relied on this argument (i.e. Plavcan *et al.* 2005; Scott & Stroik 2006). Incorporation of four additional individuals of small to intermediate body size did indeed increase the dimorphism in the A.L. 333 sample, just sufficient to prevent statistical significance in its difference from gorillas (table 5). However, as is confirmed by both the lower dimorphism values actually calculated for this sample (figure 2), and the nearly equivalent ranges of variation observed between the A.L. 333 and non-333 samples (figure 1), this finding reflects A.L. 333's small sample size ($n = 9$), as sample size has a profound impact on adequately inferring dimorphism (Koscinski & Pietraszewski 2004). Indeed, the A.L. 333 locality, which represents one of the most complete and taphonomically unbiased hominid sites ever found, still probably provides the most accurate sample of *Au. afarensis* dimorphism. It should also be noted, in addition, that an upper limit to dimorphism within this species is set by combining specimens from geographically and temporally distinct sites (i.e. the total CA sample), as this practice must enhance the variance beyond that typical of local demes.

The inclusion of new specimens also reinforces the fact that, while A.L. 333 preserves a number of large specimens, and multiple small individuals have been recovered from non-333 localities, the majority of *Au. afarensis* specimens are intermediately sized (figure 1). It is thus also noteworthy that the CA sample, which includes both large and small size extremes, can reject both low chimpanzee and high gorilla-like dimorphism. This also stresses the need to maximize sample size to include the numerous intermediate sized specimens, as this tends to ensure that more complete yet extreme-sized individuals (i.e. A.L. 288-1, A.L. 128/129 and A.L. 333-3) do not unduly influence dimorphism estimates (e.g. Gordon *et al.* 2008).

It is clear that our method of assessing skeletal dimorphism within the A.L. 333 assemblage is appropriate regardless of any sex bias due to sampling error (e.g. as argued by Plavcan *et al.* (2005) and Scott & Stroik (2006)). Moreover, the presence of small juvenile specimens preserved at A.L. 333 suggests that no systematic size sorting occurred during the formation of the assemblage. As noted above, because the sex of each postcranial element in the *Au. afarensis* sample is unknown, the numbers of males and females included within the bootstrapped samples were allowed to vary freely. Therefore, the bootstrapped simulations produced all possibilities with respect to sex composition. Indeed, some of our simulations generated samples containing only one sex.

That *Au. afarensis* displayed only moderate size dimorphism is consistent with the minimal size dimorphism observed in *Ardipithecus ramidus* (Suwa *et al.* 2009; White *et al.* 2009). Indeed, given the absence of appreciable skeletal size dimorphism in both *Pan* and *Ardipithecus*, there is now strong evidence that the last common ancestor of chimpanzees, bonobos and humans also displayed minimal skeletal dimorphism and that it probably increased in hominids subsequent to 4.4 Ma.

Recently, Lawler (2009) established that ecological factors (e.g. substrate preference or feeding niche) often produce dimorphism ratios that differ substantially from those predicted by simple sexual selection theory (e.g. the 'tournament sex' of Devore & Lovejoy (1985)). Lovejoy (1981, 1993, 2009) has argued that a provisioning model favours the selection of large males by females because greater body mass increases both mobility and predator resistance in males. Also, selection of small females by males reduces that female's fat/protein requirements and thereby lowers competition with the male's offspring for nutrient-rich foods. In addition, the obviously minimal intermale aggression in *Ar. ramidus*, as now established by the multiple trait shifts in its sectorial canine complex (including those of size, crown form, eruption time and upper/lower canine differences (Suwa *et al.* 2009; White *et al.* 2009)), makes it even more unlikely that extreme dimorphism would evolve so rapidly in *Au. afarensis* via direct male–male competition for mates. Instead, moderate dimorphism appears to be an ecologically driven feature in the hominid lineage that probably continued into later taxa (i.e. *Au. africanus* (Harmon 2009)), and although most probably the result of sexual selection, it was probably not driven by direct male–male agonistic competition for mates, but rather by ecologically driven male and female choice. Indeed, it would seem that there are now two competing explanations for the increase in skeletal dimorphism from 4.4 Ma (*Ar. ramidus*, White *et al.* 2009) to 3.2 Ma

Table 5. Simulations of *Au. afarensis* dimorphism from three different fossil assemblages. These are exact counts of values that fall less than or greater than the *Au. afarensis* value. Each count can be transformed into a proportion by dividing by 1000.

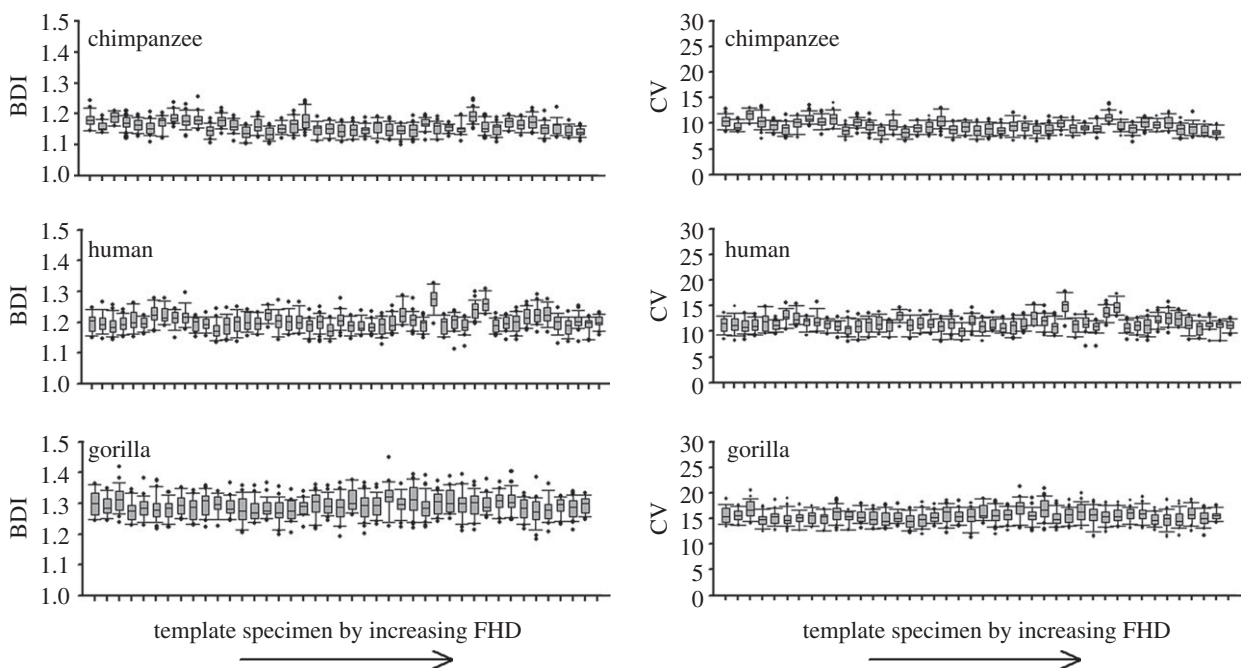|	fossil assemblage dimorphism	chimpanzee		human		gorilla	
		less than	greater than	less than	greater than	less than	greater than
A.L. 333							
BDI	1.195	905	95	484	516	74	926
CV	11.23	878	122	447	553	51	949
Combined Afar							
BDI	1.209	960	40	607	393	8	992
CV	11.89	954	46	597	403	7	993
non-333							
BDI	1.213	916	84	698	302	50	950
CV	12.03	870	130	627	373	35	965

Figure 3. Box and whisker plots showing range of sample dimorphism values generated for each template specimen. Template specimens are arrayed by increasing FHD. Boxes indicate interquartile range, whiskers 95% interval; circles are outliers.

(A.L. 333, Kimbel *et al.* 1994): (i) an increase in male–male agonism for mate selection or (ii) the enhancement of male resistance to predation in response to occupation of novel environments by the more ecologically expansive *Australopithecus* radiation, including the invasion of new predator-rich environments such as lake margins, savannas and veldts. Given that the former of these two choices would likely depress sub-adult survivorship and increase parenting load on females, when coupled with the now clear adaptive radiation of *Australopithecus* that followed *Ar. ramidus*, the latter of these two seems far more likely.

5. CONCLUSIONS

The template method is a robust technique for estimating size variance in early hominids and is the only method currently available with which sample sizes sufficient for statistical reliability are likely to be generated from rare early hominid fossils. It should be noted, moreover, that this method is fully applicable to other species of fossil hominoids, so long as a partial skeleton and a sufficiently large series of unassociated fossils with homologous anatomical sites are available (e.g. *Proconsul* (Walker & Teaford 1989; Ward *et al.* 1993), *Ar. ramidus* (White *et al.* 2009) and *Au. africanus* (Clarke 1999)). For South African *Australopithecus*, however, special consideration of taphonomic variables will have to be made, since cave assemblages are probably the result of carnivore kills (Brain 1981). Because no specific size-sorting mechanism has been identified for A.L. 333, this site remains an appropriate venue for examination of skeletal dimorphism in *Au. afarensis*.

REFERENCES

Behrensmeyer, A. K., Harmon, E. H. & Kimbel, W. H. 2003 Environmental context and taphonomy of the A.L. 333 Locality, Hadar, Ethiopia. Abstracts of the Paleoanthropology Society. See http://www.paleoanthro.org/abst2003.htm.

Brain, C. K. 1981 *The hunters or the hunted*. Chicago, IL: University of Chicago Press.

Clarke, R. J. 1999 Discovery of complete arm and hand of the 3.3 million-year-old *Australopithecus* skeleton from Sterkfontein. *S. Afr. J. Sci.* **95**, 477–480.

Devore, I. & Lovejoy, C. O. 1985 The natural superiority of women. In *The physical and mental health of aged women* (eds M. R. Haug, A. B. Ford & M. Sheafor), pp. 27–38. New York, NY: Springer.

Drapeau, M. S., Ward, C. V., Kimbel, W. H., Johanson, D. C. & Rak, Y. 2005 Associated cranial and forelimb remains attributed to *Australopithecus afarensis* from Hadar, Ethiopia. *J. Hum. Evol.* **48**, 593–642. (doi:10.1016/j.jhevol.2005.02.005)

Gordon, A. D., Green, D. J. & Richmond, B. G. 2008 Strong postcranial size dimorphism in *Australopithecus afarensis*: results from two new resampling methods for multivariate data sets with missing data. *Am. J. Phys. Anthropol.* **135**, 311–328. (doi:10.1002/ajpa.20745)

Harmon, E. H. 2006 Size and shape variation in *Australopithecus afarensis* proximal femora. *J. Hum. Evol.* **51**, 217–227. (doi:10.1016/j.jhevol.2006.01.009)

Harmon, E. 2009 Size and shape variation in the proximal femur of *Australopithecus africanus*. *J. Hum. Evol.* **56**, 551–559. (doi:10.1016/j.jhevol.2009.01.002)

Kimbel, W. H., Johanson, D. C. & Rak, Y. 1994 The first skull and other new discoveries of *Australopithecus afarensis* at Hadar, Ethiopia. *Nature* **368**, 449–451. (doi:10.1038/368449a0)

Kimbel, W. H., Rak, Y. & Johanson, D. C. 2004 *The skull of Australopithecus afarensis*. New York, NY: Oxford University Press.

Koscinski, K. & Pietraszewski, S. 2004 Methods to estimate sexual dimorphism from unsexed samples: a test with computer-generated samples. *Anthropol. Rev.* **67**, 33–55.

Lawler, R. R. 2009 Monomorphism, male–male competition, and mechanisms of sexual dimorphism. *J. Hum. Evol.* **57**, 321–325. (doi:10.1016/j.jhevol.2009.07.001)

Lockwood, C. A., Richmond, B. G., Jungers, W. L. & Kimbel, W. H. 1996 Randomization procedures and sexual dimorphism in *Australopithecus afarensis*. *J. Hum. Evol.* **31**, 537–548. (doi:10.1006/jhev.1996.0078)

Lovejoy, C. O. 1981 The origin of man. *Science* **211**, 341–350. (doi:10.1126/science.211.4480.341)

Lovejoy, C. O. 1993 Modeling human origins: are we sexy because we're smart, or smart because we're sexy? In *The origin and evolution of humans and humanness* (ed. D. T. Rasmussen), pp. 1–28. Boston, MA: Jones and Bartless Publisher.

Lovejoy, C. O. 2009 Reexamining human origins in light of *Ardipithecus ramidus*. *Science* **326**, 74e1–74e8. (doi:10.1126/science.1175834)

Lovejoy, C. O., Johanson, D. C. & Coppens, Y. 1982 Hominid lower-limb bones recovered from the Hadar formation—1974–1977 collections. *Am. J. Phys. Anthropol.* **57**, 679–700. (doi:10.1002/ajpa.1330570411)

McHenry, H. M. 1991 Sexual dimorphism in *Australopithecus afarensis*. *J. Hum. Evol.* **20**, 21–32. (doi:10.1016/0047-2484(91)90043-U)

Plavcan, J. M., Lockwood, C. A., Kimbel, W. H., Lague, M. R. & Harmon, E. H. 2005 Sexual dimorphism in *Australopithecus afarensis* revisited: how strong is the case for a human-like pattern of dimorphism? *J. Hum. Evol.* **48**, 313–320. (doi:10.1016/j.jhevol.2004.09.006)

Reno, P. L., Meindl, R. S., McCollum, M. A. & Lovejoy, C. O. 2003 Sexual dimorphism in *Australopithecus afarensis* was similar to that of modern humans. *Proc. Natl Acad. Sci. USA* **100**, 9404–9409. (doi:10.1073/pnas.1133180100)

Reno, P. L., Meindl, R. S., McCollum, M. A. & Lovejoy, C. O. 2005 The case is unchanged and remains robust: *Australopithecus afarensis* exhibits only moderate skeletal dimorphism. A reply to Plavcan *et al.* (2005). *J. Hum. Evol.* **49**, 279–288. (doi:10.1016/j.jhevol.2005.04.008)

Scott, J. E. & Stroik, L. K. 2006 Bootstrap tests of significance and the case for humanlike skeletal-size dimorphism in *Australopithecus afarensis*. *J. Hum. Evol.* **51**, 422–428. (doi:10.1016/j.jhevol.2006.06.001)

Sokal, R. R. & Rohlf, F. J. 1995 *Biometry*. New York, NY: W. H. Freeman and Co.

Suwa, G., Kono, R. T., Simpson, S. W., Asfaw, B., Lovejoy, C. O. & White, T. D. 2009 Paleobiological implications of the *Ardipithecus ramidus* dentition. *Science* **326**, 94–99. (doi:10.1126/science.1175824)

Walker, A. & Teaford, M. 1989 The hunt for *Proconsul*. *Sci. Am.* **260**, 76–82. (doi:10.1038/scientificamerican0189-76)

Ward, C. V., Walker, A., Teaford, M. F. & Odhiambo, I. 1993 Partial skeleton of *Proconsul nyanzae* from Mfangano Island, Kenya. *Am. J. Phys. Anthropol.* **90**, 77–111. (doi:10.1002/ajpa.1330900106)

White, T. D. & Johanson, D. C. 1989 The hominid composition of Afar Locality 333: some preliminary observations. In *Hominidae*, pp. 97–101. Milan, Italy: Jaka Book.

White, T. D., Asfaw, B., Beyene, Y., Haile-Selassie, Y., Lovejoy, C. O., Suwa, G. & WoldeGabriel, G. 2009 *Ardipithecus ramidus* and the paleobiology of early hominids. *Science* **326**, 75–86. (doi:10.1126/science.1175802)

Zihlman, A. L. & Tobias, P. V. 1985 *Australopithecus afarensis*: two sexes or two species? In *Hominid evolution: past, present and future*, pp. 213–220. New York, NY: Alan R. Liss, Inc.