PNAS
www.pnas.org

Supplementary Information for

Canine sexual dimorphism in *Ardipithecus ramidus* was nearly human-like.

Gen Suwa[1], Tomohiko Sasaki[1], Sileshi Semaw, Michael J. Rogers, Scott W. Simpson, Yutaka Kunimatsu, Masato Nakatsukasa, Reiko T. Kono, Yingqi Zhang, Yonas Beyene, Berhane Asfaw, Tim D. White

[1] equal contribution

corresponding author: Gen Suwa

Email: gsuwa@um.u-tokyo.ac.jp

This PDF file includes:

- Supplementary text
- Figures S1 to S16
- Table S1
- Legends for Datasets S1 to S4
- SI References

Other supplementary materials for this manuscript include the following:

- Datasets S1 to S4

Supplementary Information Text

Comparative modern primate datasets

The Plavcan (1990) dataset (1) includes a wide range of taxa delineated geographically and/or taxonomically at the subspecies level, and is therefore appropriate for our comparisons. As was done in Suwa et al. (2009) (2), we first confined specimens to those scored in the little or moderately worn categories in Plavcan (1). We excluded the relatively worn specimens, not only specimens scored as having heavy wear for canines, but also specimens scored as having heavy wear for incisors and/or molars. This was done to conservatively preserve canine height dimensions as much as possible. Next, we systematically excluded outliers via a Mahalanobis-distance-based protocol, setting the exclusion criteria at $p<0.001$ distance from the sex means. The details of this screening protocol is described in Sasaki et al. (2021) (3). Finally, we chose taxa that had both male and female sample sizes of 10 or greater, but allowed exceptions for balanced taxon representations. The final comparative dataset we used in the crown height comparisons included 4 hylobatid, 15 colobine, 8 cercopithecin, 13 papionin, 6 atelid, 3 pitheciine, and 10 other platyrrhine species. The extant great ape samples are those used in Suwa et al. (2) and include *Gorilla gorilla gorilla*, *Pan troglodytes troglodytes*, *Pan troglodytes schweinfurthi*, *Pan paniscus* and *Pongo pygmaeus*. The extant anthropoid samples, screened as outlined above, were also used in the validation analyses of real samples of known sex in Sasaki et al. (3).

Metrics

In evaluating canine crown size, we focused on maximum basal crown diameter. In the upper (maxillary) canines, in *Ardipithecus ramidus*, *Australopithecus anamensis* and fossil and extant great apes, this is either the mesiodistal or the maximum oblique diameter. To the contrary, in the Plavcan (1990) dataset (1), the mesiodistal diameter was taken as the maximum diameter. In the lower (mandibular) canines, the maximum diameter can be taken oblique to the crown axis and tooth row, thereby termed the maximum oblique diameter (this study, and ref 2), or it may approximate the direction of the tooth row and termed the mesiodistal diameter as in Plavcan (1). In the other *Australopithecus* and *Homo* canines, both upper and lower canines have weaker basal crown asymmetry and a buccolingually broad crown. In these cases, we used the commonly available buccolingual diameter as the appropriate size parameter that is minimally affected by wear. In order to increase sample size of *Ar. ramidus* canines, we derived basal crown diameter estimates in two specimens via 3d-model-superimposition (Figs. S15, S16).

Canine height is difficult to assess because cusp tip wear obscures its initial unworn condition. Therefore, in order to track species specific crown heights to the extent possible, in using the Plavcan (1990) anthropoid dataset, we applied a conservative protocol to screen out effects of wear to the extent possible (see above). For extant great ape canines and those of *Ar. ramidus* and the other fossils, we followed the Kelley (1995) (4) method of estimating unworn crown heights. This was done on canines with minimally worn cusp tips, by extending the crown face contours toward the cusp tip and compensating for small amounts of enamel loss (usually <~1 to 2 mm). The *Ar. ramidus* ARA-VP-6/500 upper canine is not sufficiently preserved to enable a visually-based height estimate. Therefore, in the present study, we used 3d-model superimposition to estimate its crown height (Fig. S15).

Simulation tests

We assessed the performance of the pdPeak method by using computationally generated population samples. We simulated the source populations based on assumptions of homoscedastic normal distributions (in log-scale) of males and females mixed in the same proportion. The female mean was set to 10 mm, and the male mean was set so that the m/f ratio ranged from 1.0 to 1.3 in 0.03 increments. The within-sex CV (wsxCV) was set to either 5% or 8% to represent conditions common with human and anthropoid canine crown diameters. Using the simulated populations (with set m/f ratios and wsxCVs), we sampled test samples of various sizes (N) to determine the accuracy and robustness of each method; the pdPeak, mean, BDI and CV methods. Under each simulated condition 2,000 sets of test samples of N=10 or N=30 were generated, and the pdPeak estimates of the m/f ratio were derived. Using the same

test samples, the m/f ratio estimates of the mean method, the BDI and CV methods were calculated.

Deviation from assumptions

Results of a broader range of performance tests are reported in Sasaki et al. (3), including the effects of deviations from assumptions of log normality, homoscedasticity, and equal source population sex ratio. We found that the effects of deviations are generally slight, with the exception of a highly unbalanced sex ratio or a skewed distribution in which the subpopulations exhibit negative skew in females and positive skew in males (head-to-head skew). In these cases, the pdPeak m/f ratio estimates tend to underestimate, although the magnitude of this bias was not excessive. We also note that there is no *a priori* expectation for the number of males and females to be biased nor the distributions to be notably skewed in the fossil source populations. However, it is nevertheless recommended to examine skewness in the fossil samples that are being investigated. A skewed sample may be an indication of heteroscedasticity, within-sex skew, and/or an unbalanced sex ratio. The skewness values of the fossil samples that we analyzed are shown in Dataset S2. There were no significant cases of skewness at the p=0.05 level (the D'Agostino 1970, two-tailed test, ref 5).

Figure S1

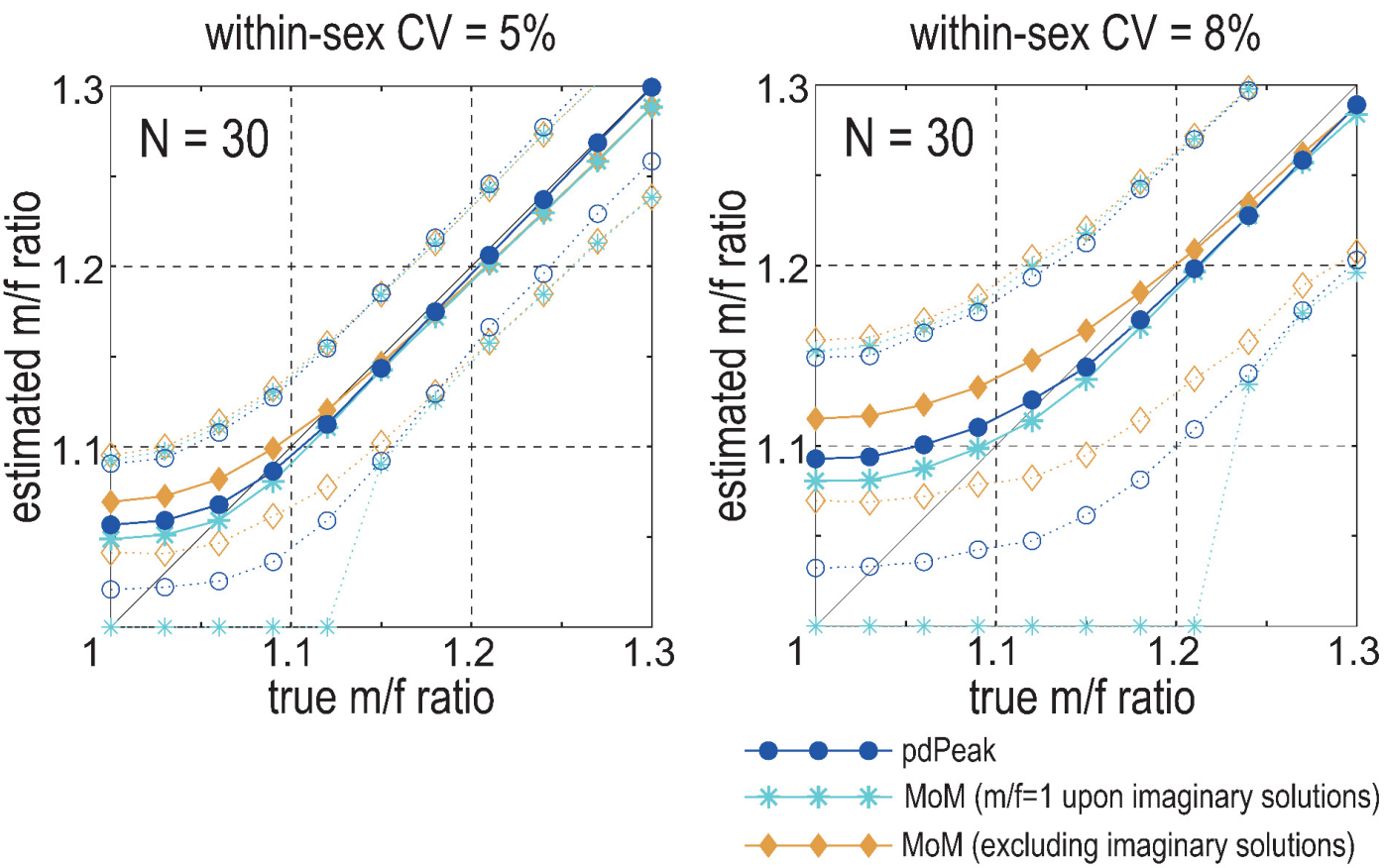


Figure S1. Simulation test comparisons of the pdPeak method and the MoM (method-of-moments).

The x-axis is the m/f ratio (male mean divided by the female mean) of hypothetical populations with either a within-sex CV of 0.05 (left plot) or 0.08 (right plot). The y-axis is the estimated population m/f ratio. Solid lines are the means, and dotted lines are the 5th (lower line) and 95th (upper line) percentiles of the 2,000 N=30 samples extracted from the simulated populations. This was done for simulated populations set at m/f ratios of 0.03 intervals. Blue, pdPeak; cyan, MoM including imaginary solutions; orange, MoM excluding imaginary solutions. In the MoM, imaginary solutions were either excluded from the analysis, or assigned an m/f ratio of 1.0 after Josephson et al. (SI ref 6). In the latter case, although unbiased expected values are seen to extend to <1.10, this is based on the frequent occurrence of imaginary solutions designated 1.0. To the contrary, exclusion of the imaginary solutions results in inability of attaining unbiased m/f ratio estimates less than 1.1 (with a within-sex CV of 0.05) or 1.2 (with a within-sex CV of 0.08). In general, the real-number solutions of the MoM tend to overestimate when the m/f ratio is lower than 1.2, but the degree of overestimation is less than in the MM and BDI.

Figure S2

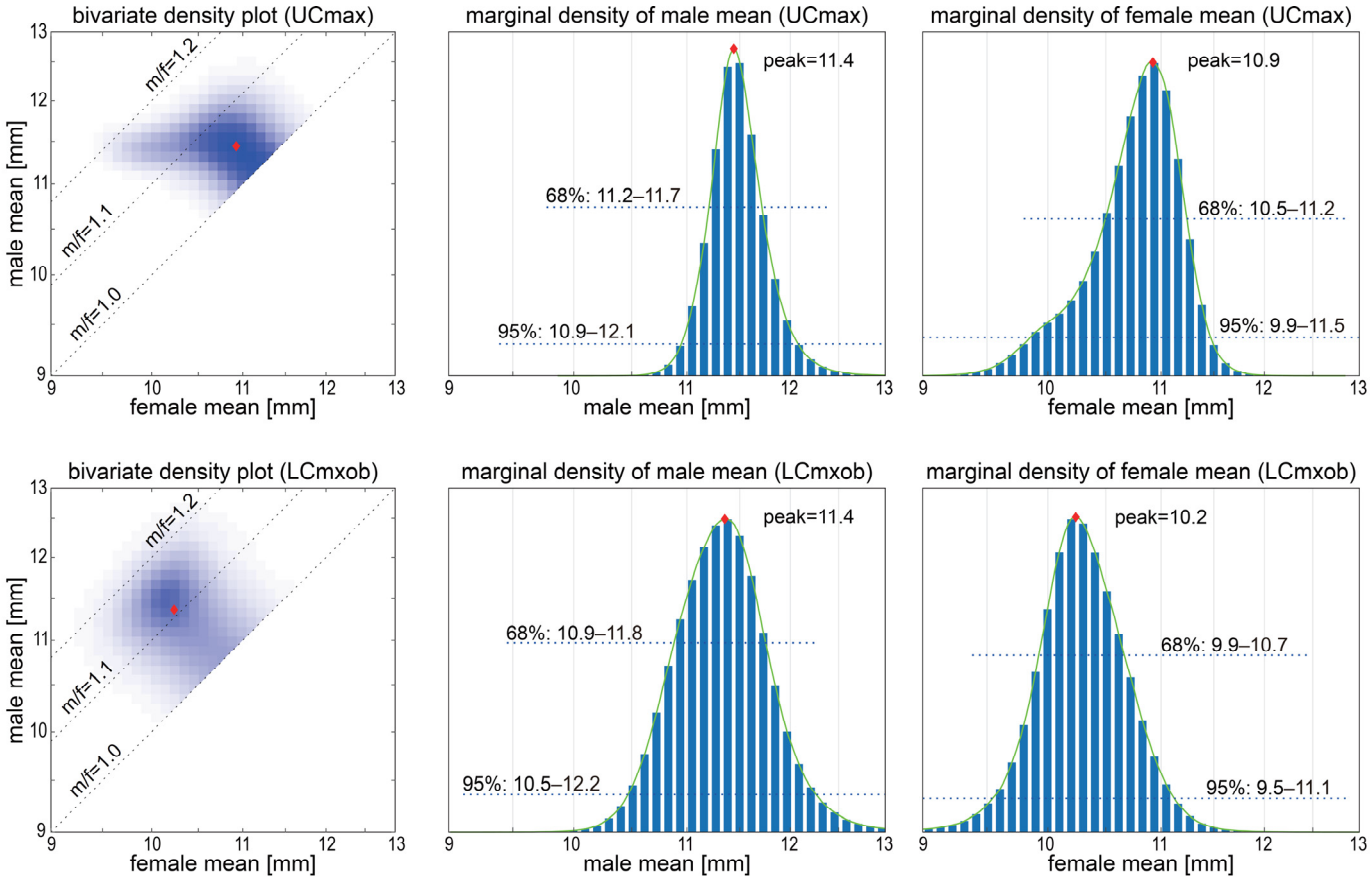


Figure S2. *Ardipithecus ramidus* canine sexual dimorphism shown via male and female mean pdPeak estimates.

The bivariate probability distribution of the male (y-axis) and female (x-axis) means are shown in gray scale. The probability densities were obtained with the log-transformed data and the axis labels were back-transformed to original scale. The red dots indicate the pdPeak estimates of the male and female means. Upper row, upper (maxillary) canine; lower row, lower (mandibular) canine. To the right of the bivariate distribution plots, the marginal probability densities of the male and female means (of log-transformed data), their pdPeak values, and credible intervals are shown. The m/f ratios calculated from the male and female pdPeak values are 1.046 for the upper canine and 1.117 for the lower canine. These differ slightly, but not significantly, from the direct pdPeak m/f ratio estimates of 1.06 and 1.13, respectively, for the upper and lower canines.

Figure S3

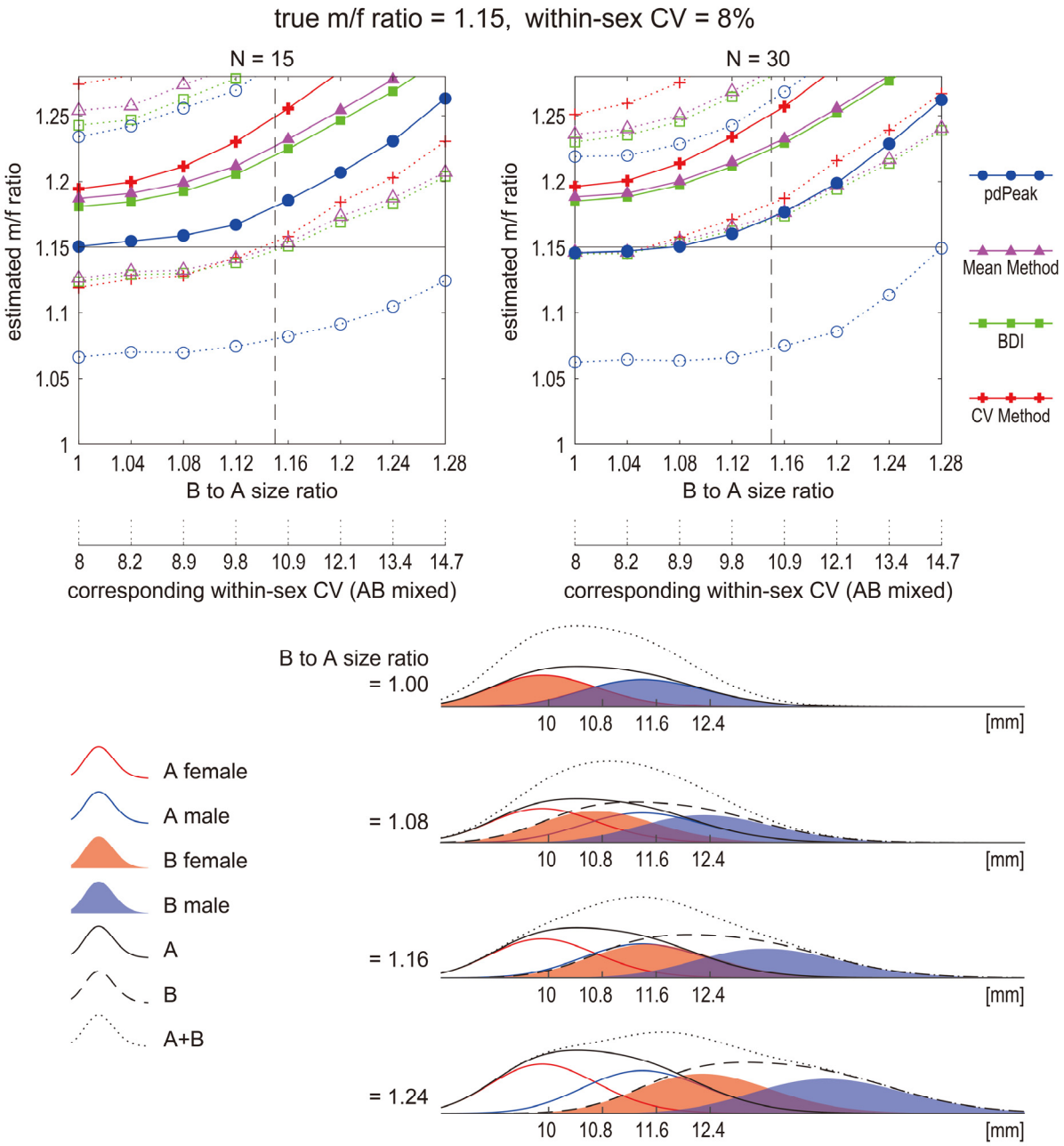


Figure S3. Mixed population effects on the pdPeak m/f ratio estimates.

In this example, groups A and B both have m/f ratios of 1.15, within-sex CV of 0.08, and equal population sizes. Group B (both sexes) is successively simulated to have greater means by the ratio indicated in the x-axis of the upper plots. The y-axis is the estimated population m/f ratio. Solid lines are the expected means; dotted lines are the 5th (lower line) and 95th (upper line) percentiles of the 2,000 simulated samples (of N=15 or N=30) of the A+B combined-group test populations. Blue, pdPeak; purple, mean method; green, BDI; red, CV (Plavcan 1994) method (SI ref 7). Note that, with the pdPeak method, at an x-axis value of 1.08, inflation of the m/f ratio estimate is minimal; even at an x-axis value of 1.15, m/f ratio estimate inflation remains modest. An x-axis value of 1.15 corresponds to the group B female mean being as large as the group A male mean. Simulations under other hypothetical m/f ratios and CVs show that inflation of the m/f ratio estimate via group-mixing is less when within-sex CV is lower and when population m/f ratio is higher. When representation of groups A and B is uneven, the estimated m/f ratio inflation is less.

Figures S4 to S11

Figures S4–S11. Bivariate probability density plots of canine sexual dimorphism in *Ardipithecus ramidus* and other taxa (pdPeak estimates of the m/f ratio and within-sex CV).

Fig. S4. *Ar. ramidus*;
Fig. S5. *Australopithecus anamensis* and *Au. afarensis*;
Fig. S6. *Au. africanus* and *Au. robustus*;
Fig. S7. *Au. boisei* and early *Homo*;
Fig. S8. Atapuerca SH and European Neanderthals;
Fig. S9. European Upper Paleolithic and *Nacholapithecus kerioi*;
Fig. S10. *Hispanopithecus laietanus* and *Ouranopithecus macedoniensis*;
Fig. S11. *Oreopithecus bambolii* and *Gigantopithecus blacki*.

Sample distributions and basic statistics are shown at the left. At the far right, the bivariate probability distributions of the m/f ratio (male mean divided by the female mean) (y-axis) and within-sex CV (x-axis) are shown in gray scale. Probability densities were obtained with the log-transformed data, and the axis labels were back-transformed to original scale (see SI ref 3 for details). Yellow and red squares, respectively, are the modern human (Dataset S1, n≥20) and extant great ape population or taxon values (SI ref 2). The dotted contour lines indicate combined-sex CV levels. The marginal posterior probability densities of the logarithm of the m/f ratios are shown in the middle part of the figure. The pdPeak values and credible intervals are indicated, together with the m/f ratio estimates by the MM, BDI and CV (Plavcan 1994) method (SI ref 7). For analyses of n<10, histograms are not shown because the posterior density was directly calculated (i.e., not by MCMC sampling).

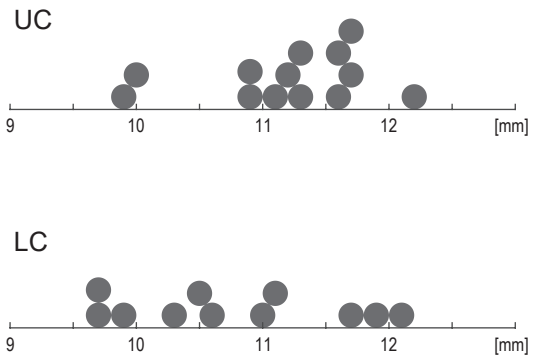
When applicable, population subsets of each species were analyzed, the results of which are tabulated in Dataset S2. Whereas geographically and/or temporally confined subset samples were envisioned to express lower levels of variance, conceptually with tighter pdPeak results, this was not necessarily the case. This may stem from a combination of: 1) chance bias associated with smaller sample sizes of the subdivided species/population samples, and 2) lower analytical resolution of the pdPeak method with smaller sample sizes. However, especially with *Au. afarensis*, because temporal and/or taxonomic heterogeneity has been suggested by many workers, we conducted a wide range of subset analyses. We found these results to be in accord with the entire species results, given analytical uncertainties. Attempts to evaluate subtle differences are complicated by the lower resolutions of the pdPeak estimates with smaller sample sizes.

Figure S4_a

Ar. ramidus

UC:max, LC:mxob

sample distributions

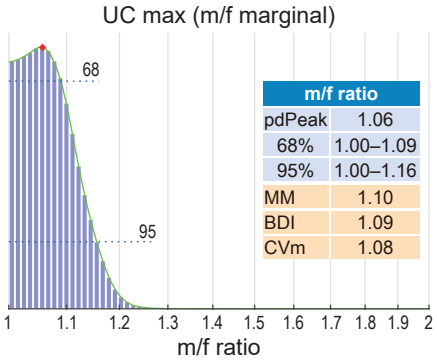


sample statistics

	UC	LC
N	13	11
mean	11.2	10.8
s.d.	0.66	0.86
CV*	6.0%	8.2%

* Sokal and Braumann (1980) correction

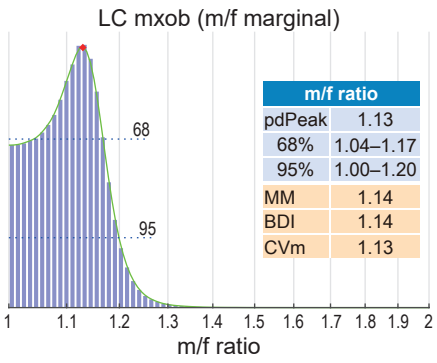
Posterior probability distributions



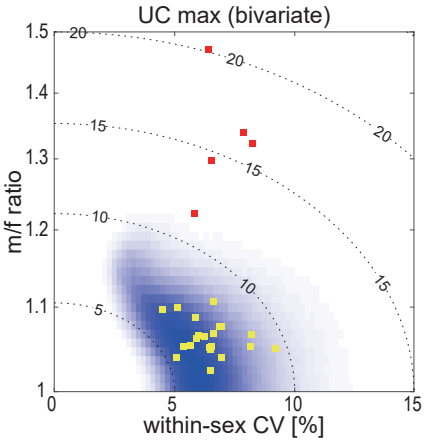
m/f ratio	
pdPeak	1.06
68%	1.00–1.09
95%	1.00–1.16
MM	1.10
BDI	1.09
CVm	1.08

within-sex CV for UC	
pdPeak	5.4%
68%	3.8–7.2%
95%	2.6–9.6%

within-sex CV for LC	
pdPeak	6.5%
68%	4.0–9.1%
95%	2.9–13.2%



m/f ratio	
pdPeak	1.13
68%	1.04–1.17
95%	1.00–1.20
MM	1.14
BDI	1.14
CVm	1.13



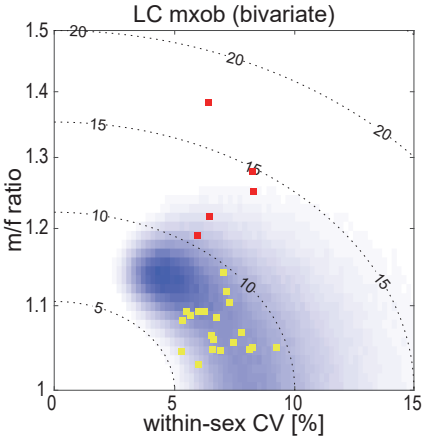
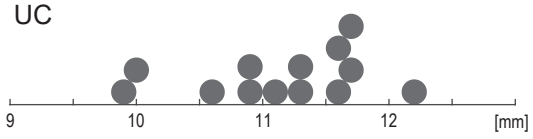


Figure S4_b

Ar. ramidus

UC:mxob

sample distribution

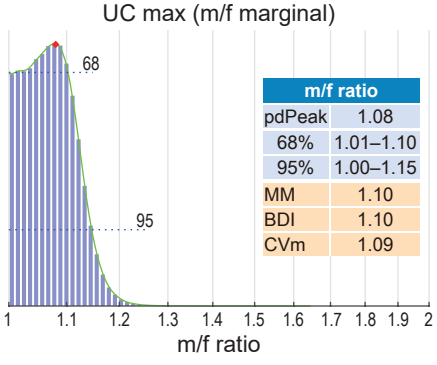


sample statistics

	UC	LC
N	13	
mean	11.1	
s.d.	0.68	
CV*	6.2%	

* Sokal and Braumann (1980) correction

Posterior probability distributions



m/f ratio	
pdPeak	1.08
68%	1.01–1.10
95%	1.00–1.15
MM	1.10
BDI	1.10
CVm	1.09

within-sex CV for UC	
pdPeak	5.5%
68%	3.8–7.3%
95%	2.8–9.8%

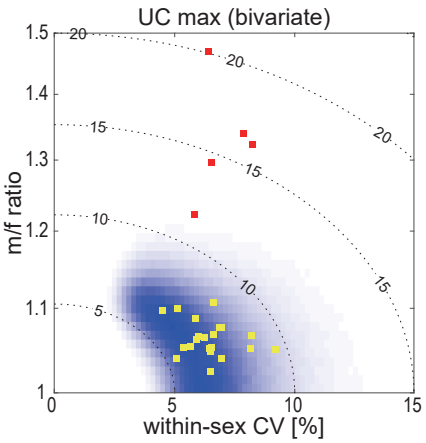
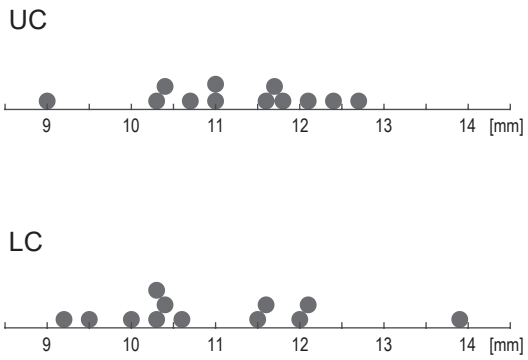


Figure S5_a

Au. anamensis

UC:max, LC:mxob

sample distributions



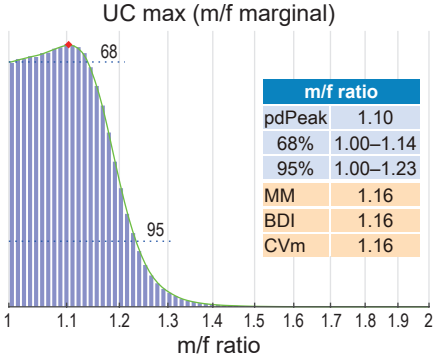
sample statistics

	UC	LC
N	12	12
mean	11.2	11.0
s.d.	1.04	1.32
CV*	9.5%	12.3%

* Sokal and Braumann (1980) correction

Posterior probability distributions

UC max (m/f marginal)

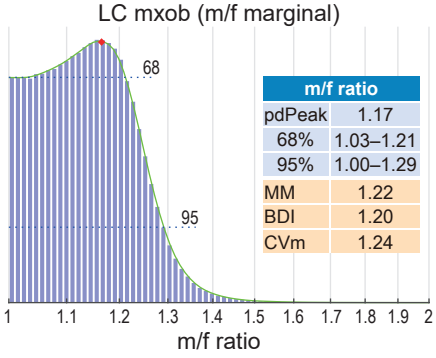


m/f ratio	
pdPeak	1.10
68%	1.00–1.14
95%	1.00–1.23
MM	1.16
BDI	1.16
CVm	1.16

within-sex CV for UC	
pdPeak	8.7%
68%	6.4–11.8%
95%	4.7–16.1%

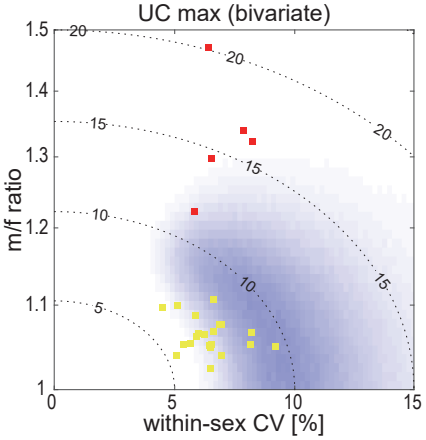
within-sex CV for LC	
pdPeak	10.4%
68%	7.2–14.2%
95%	5.0–19.4%

LC mxob (m/f marginal)



m/f ratio	
pdPeak	1.17
68%	1.03–1.21
95%	1.00–1.29
MM	1.22
BDI	1.20
CVm	1.24

UC max (bivariate)



LC mxob (bivariate)

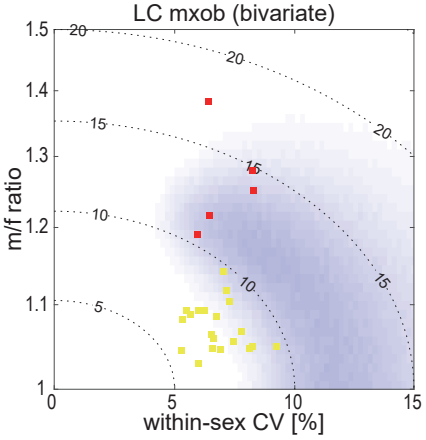
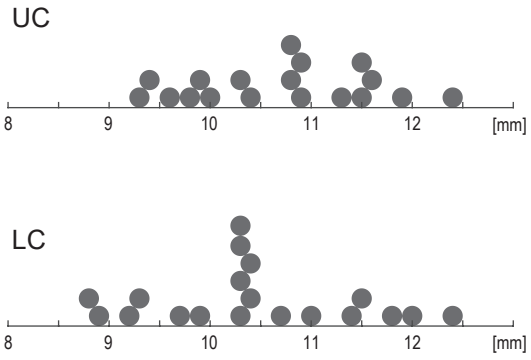


Figure S5_b

Au. afarensis

UC:bl, LC:bl

sample distributions



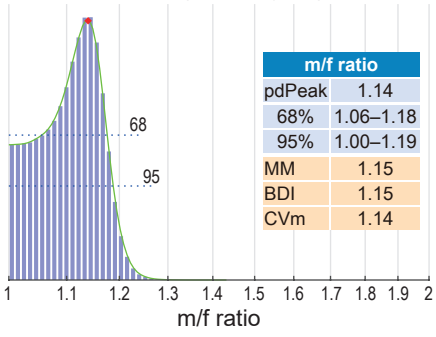
sample statistics

	UC	LC
N	18	19
mean	10.7	10.5
s.d.	0.91	1.04
CV*	8.6%	10.1%

* Sokal and Braumann (1980) correction

Posterior probability distributions

UC bl (m/f marginal)

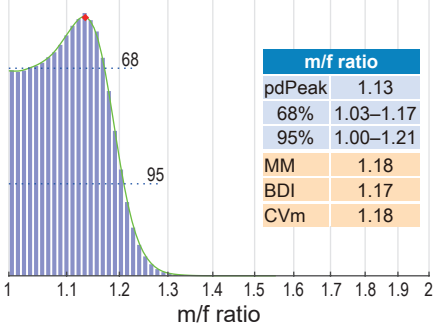


m/f ratio	
pdPeak	1.14
68%	1.06–1.18
95%	1.00–1.19
MM	1.15
BDI	1.15
CVm	1.14

within-sex CV for UC	
pdPeak	7.0%
68%	4.5–8.8%
95%	3.6–11.6%

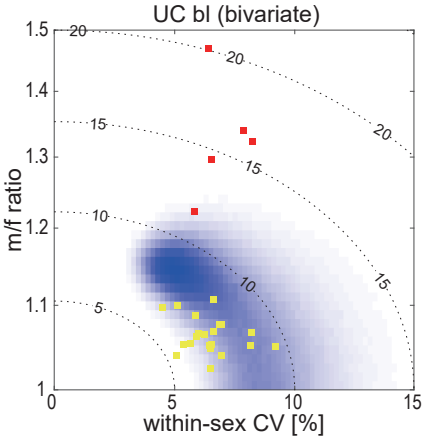
within-sex CV for LC	
pdPeak	8.8%
68%	6.7–11.1%
95%	5.2–13.9%

LC bl (m/f marginal)



m/f ratio	
pdPeak	1.13
68%	1.03–1.17
95%	1.00–1.21
MM	1.18
BDI	1.17
CVm	1.18

UC bl (bivariate)



LC bl (bivariate)

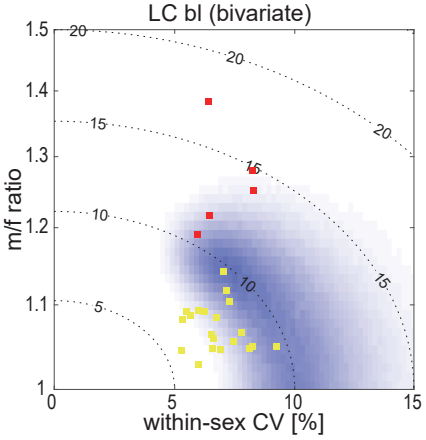
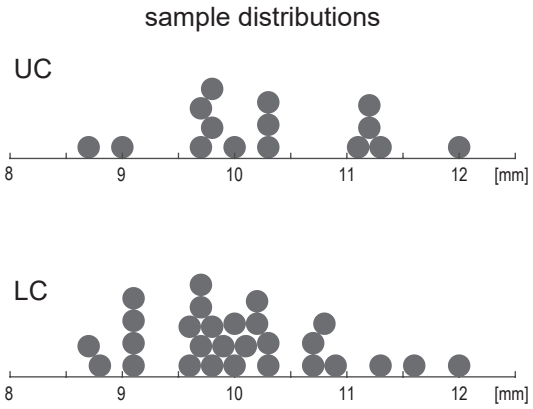


Figure S6_a

Au. africanus

UC:bl, LC:bl

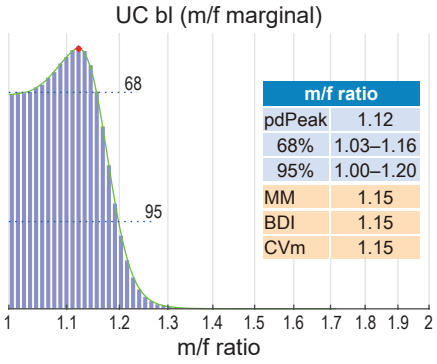


sample statistics

	UC	LC
N	15	28
mean	10.3	10.0
s.d.	0.91	0.82
CV*	9.0%	8.2%

* Sokal and Braumann (1980) correction

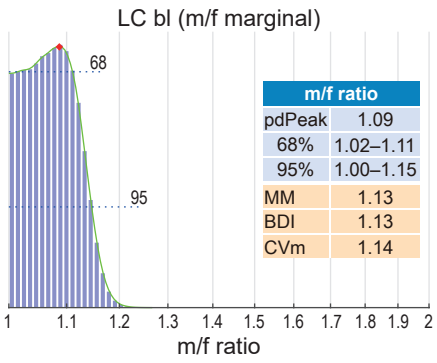
Posterior probability distributions



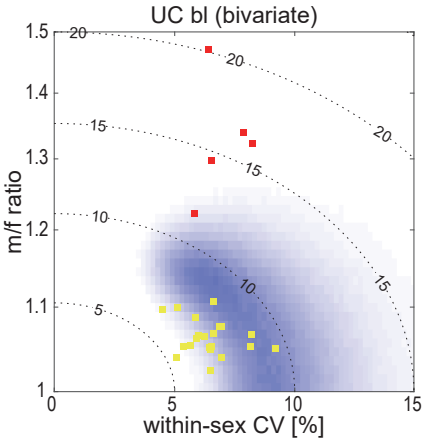
m/f ratio	
pdPeak	1.12
68%	1.03–1.16
95%	1.00–1.20
MM	1.15
BDI	1.15
CVm	1.15

within-sex CV for UC	
pdPeak	7.9%
68%	5.7–10.3%
95%	4.2–13.4%

within-sex CV for LC	
pdPeak	7.4%
68%	6.0–8.8%
95%	4.8–10.4%



m/f ratio	
pdPeak	1.09
68%	1.02–1.11
95%	1.00–1.15
MM	1.13
BDI	1.13
CVm	1.14



: extant great apes

: modern humans

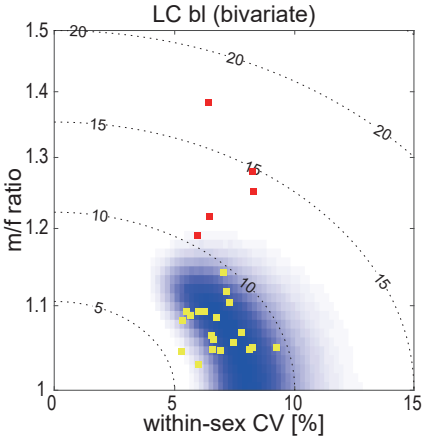
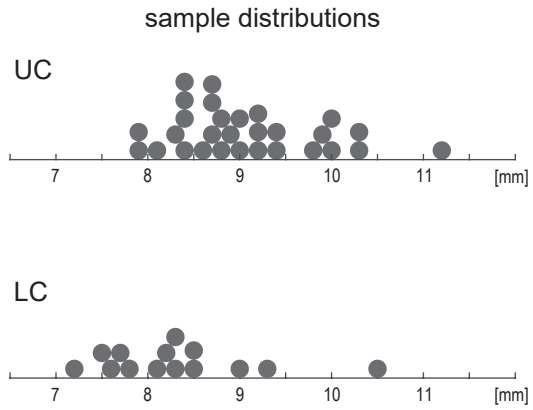


Figure S6_b

Au. robustus

UC:bl, LC:bl

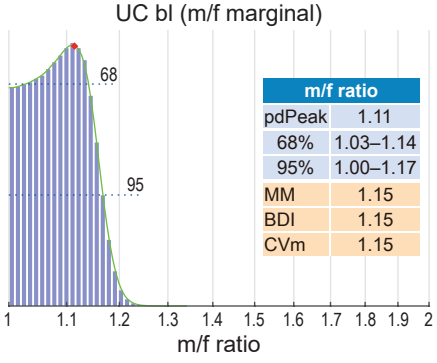


sample statistics

	UC	LC
N	29	14
mean	9.1	8.3
s.d.	0.79	0.85
CV*	8.8%	10.4%

* Sokal and Braumann (1980) correction

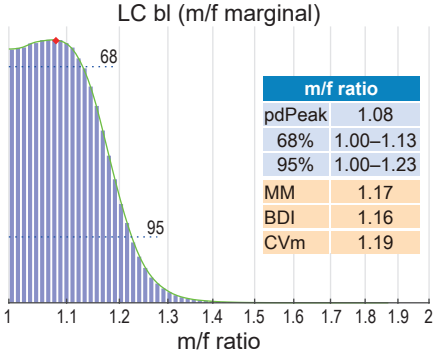
Posterior probability distributions



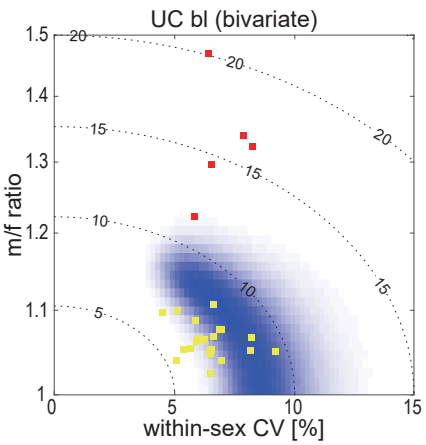
m/f ratio	
pdPeak	1.11
68%	1.03–1.14
95%	1.00–1.17
MM	1.15
BDI	1.15
CVm	1.15

within-sex CV for UC	
pdPeak	7.8%
68%	6.0–9.3%
95%	4.6–10.9%

within-sex CV for LC	
pdPeak	9.0%
68%	6.9–11.8%
95%	5.2–15.5%



m/f ratio	
pdPeak	1.08
68%	1.00–1.13
95%	1.00–1.23
MM	1.17
BDI	1.16
CVm	1.19



: extant great apes

: modern humans

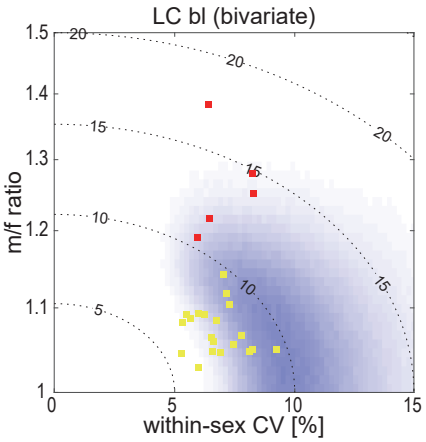
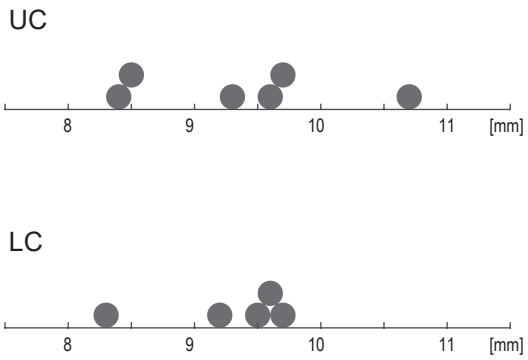


Figure S7_a

Au. boisei

UC:bl, LC:bl

sample distributions

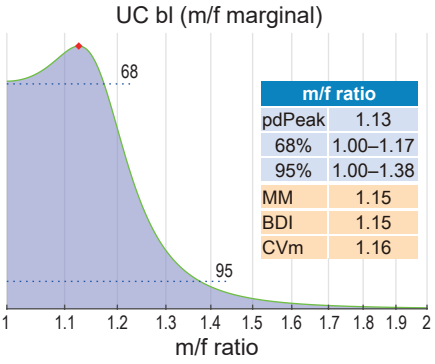


sample statistics

	UC	LC
N	6	5
mean	9.4	9.3
s.d.	0.85	0.57
CV*	9.5%	6.4%

* Sokal and Braumann (1980) correction

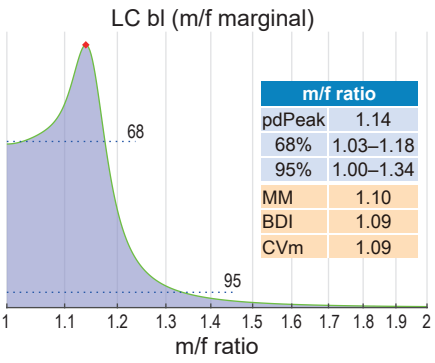
Posterior probability distributions



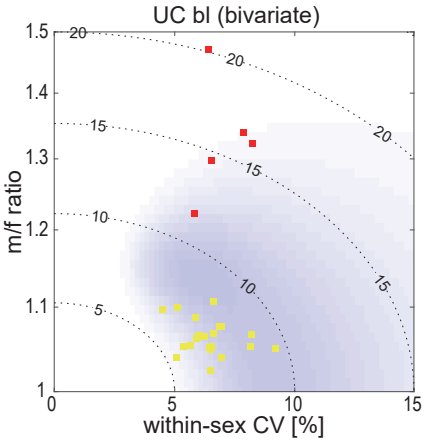
m/f ratio	
pdPeak	1.13
68%	1.00–1.17
95%	1.00–1.38
MM	1.15
BDI	1.15
CVm	1.16

within-sex CV for UC	
pdPeak	7.9%
68%	4.6–14.0%
95%	3.1–28.7%

within-sex CV for LC	
pdPeak	2.4%
68%	1.4–9.8%
95%	1.0–28.4%



m/f ratio	
pdPeak	1.14
68%	1.03–1.18
95%	1.00–1.34
MM	1.10
BDI	1.09
CVm	1.09



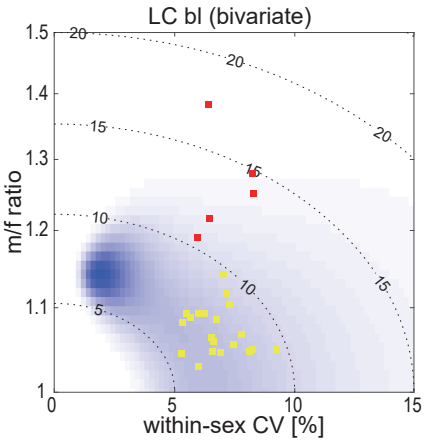
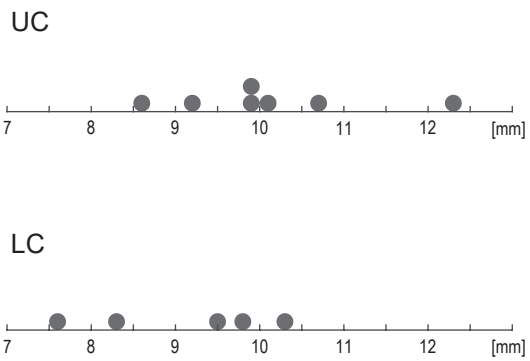


Figure S7_b

Early *Homo*

UC:bl, LC:bl

sample distributions

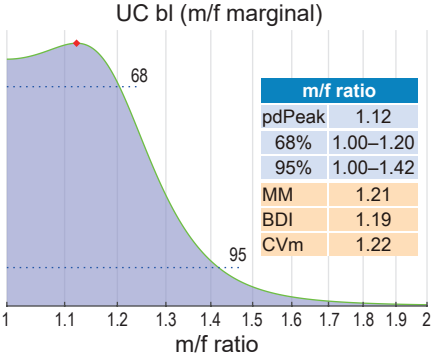


sample statistics

	UC	LC
N	7	5
mean	10.1	9.1
s.d.	1.18	1.12
CV*	12.1%	12.9%

* Sokal and Braumann (1980) correction

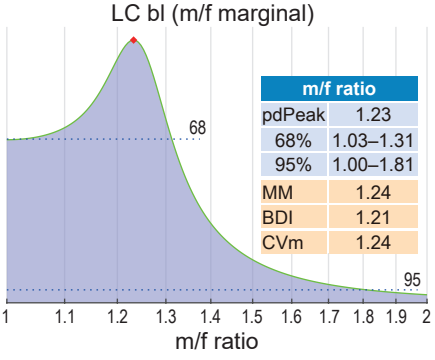
Posterior probability distributions



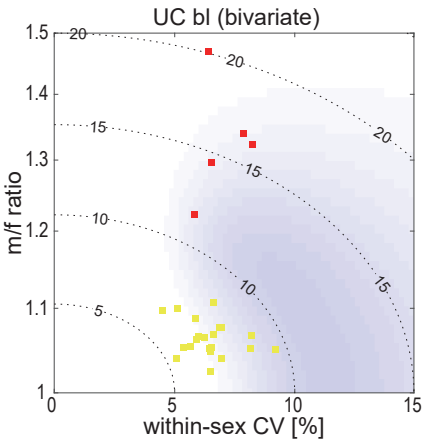
m/f ratio	
pdPeak	1.12
68%	1.00–1.20
95%	1.00–1.42
MM	1.21
BDI	1.19
CVm	1.22

within-sex CV for UC	
pdPeak	10.2%
68%	6.7–16.4%
95%	4.7–29.1%

within-sex CV for LC	
pdPeak	10.2%
68%	3.3–21.0%
95%	2.3–63.0%



m/f ratio	
pdPeak	1.23
68%	1.03–1.31
95%	1.00–1.81
MM	1.24
BDI	1.21
CVm	1.24



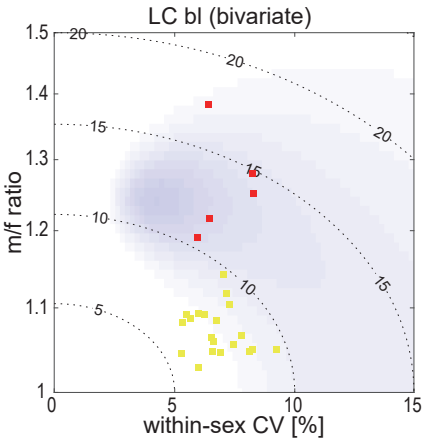
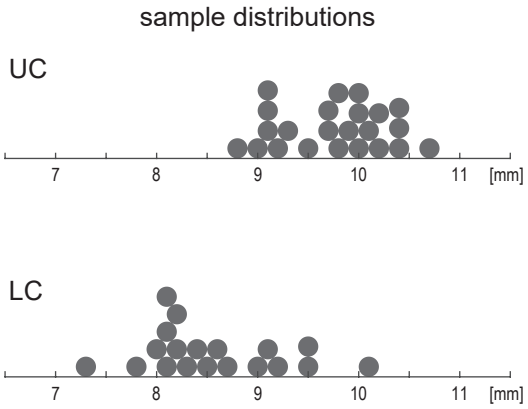


Figure S8_a

Atapuerca-SH

UC:bl, LC:bl

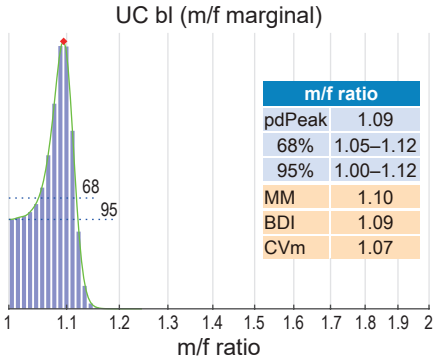


sample statistics

	UC	LC
N	23	19
mean	9.8	8.6
s.d.	0.53	0.69
CV*	5.5%	8.1%

* Sokal and Braumann (1980) correction

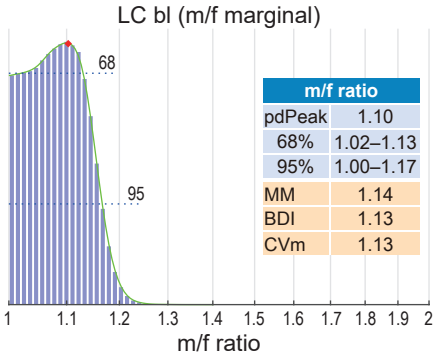
Posterior probability distributions



m/f ratio	
pdPeak	1.09
68%	1.05–1.12
95%	1.00–1.12
MM	1.10
BDI	1.09
CVm	1.07

within-sex CV for UC	
pdPeak	3.1%
68%	2.6–5.2%
95%	2.2–6.8%

within-sex CV for LC	
pdPeak	7.2%
68%	5.3–9.0%
95%	3.9–11.1%



m/f ratio	
pdPeak	1.10
68%	1.02–1.13
95%	1.00–1.17
MM	1.14
BDI	1.13
CVm	1.13

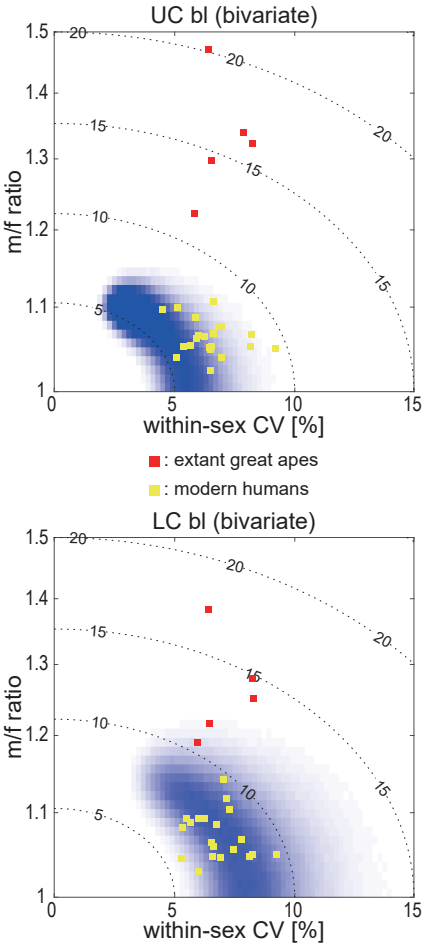
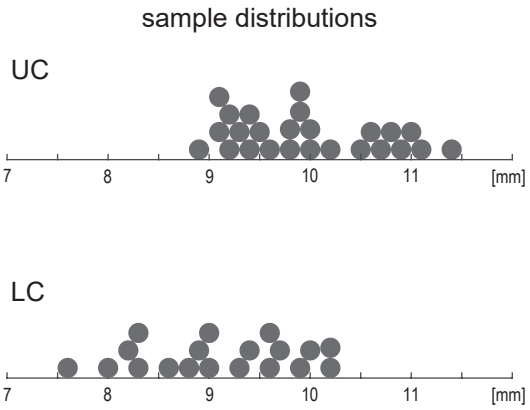


Figure S8_b

European neanderthal

UC:bl, LC:bl

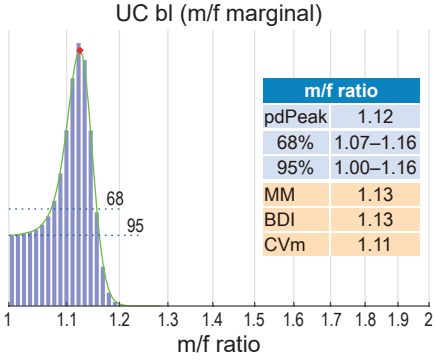


sample statistics

	UC	LC
N	25	19
mean	10.0	9.1
s.d.	0.72	0.77
CV*	7.3%	8.6%

* Sokal and Braumann (1980) correction

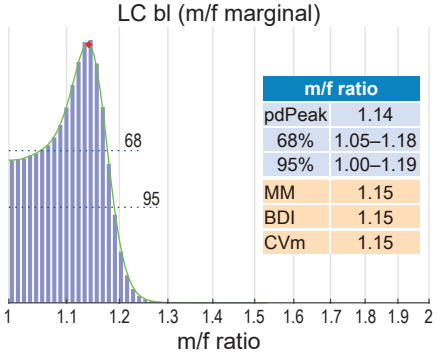
Posterior probability distributions



m/f ratio	
pdPeak	1.12
68%	1.07–1.16
95%	1.00–1.16
MM	1.13
BDI	1.13
CVm	1.11

within-sex CV for UC	
pdPeak	4.2%
68%	3.4–6.5%
95%	3.0–8.7%

within-sex CV for LC	
pdPeak	7.4%
68%	4.8–9.1%
95%	3.8–11.7%



m/f ratio	
pdPeak	1.14
68%	1.05–1.18
95%	1.00–1.19
MM	1.15
BDI	1.15
CVm	1.15

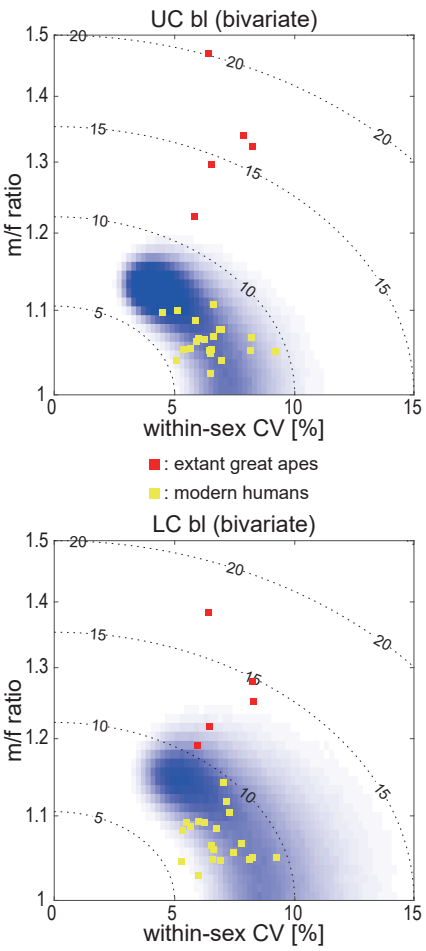
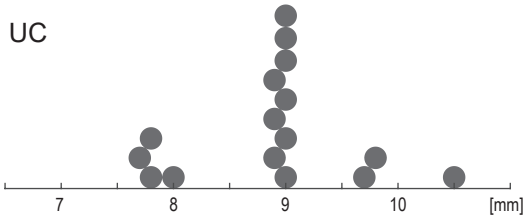


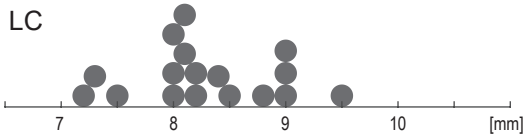
Figure S9_a

European UPL *H. sapiens*

UC:bl, LC:bl

sample distributions





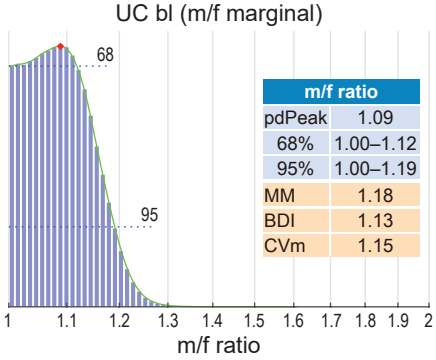
sample statistics

	UC	LC
N	16	17
mean	8.9	8.3
s.d.	0.76	0.63
CV*	8.7%	7.8%

* Sokal and Braumann (1980) correction

Posterior probability distributions

UC bl (m/f marginal)

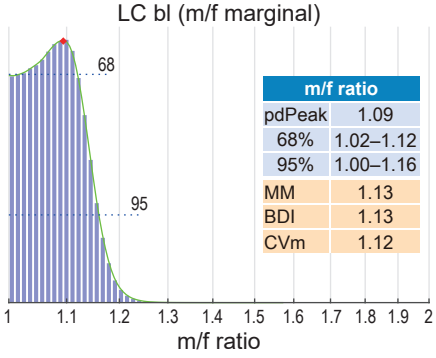


m/f ratio	
pdPeak	1.09
68%	1.00–1.12
95%	1.00–1.19
MM	1.18
BDI	1.13
CVm	1.15

within-sex CV for UC	
pdPeak	7.9%
68%	6.0–10.1%
95%	4.3–12.9%

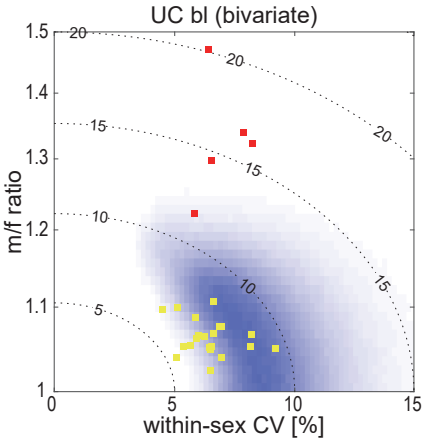
within-sex CV for LC	
pdPeak	6.9%
68%	5.2–8.8%
95%	4.0–11.2%

LC bl (m/f marginal)



m/f ratio	
pdPeak	1.09
68%	1.02–1.12
95%	1.00–1.16
MM	1.13
BDI	1.13
CVm	1.12

UC bl (bivariate)



LC bl (bivariate)

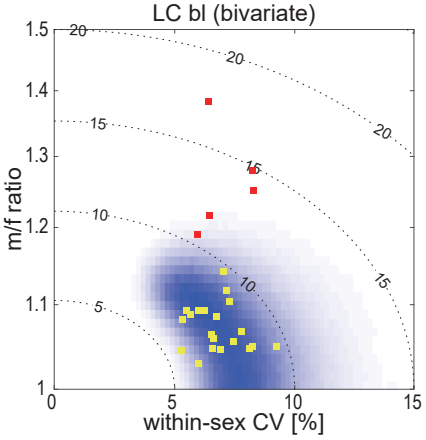
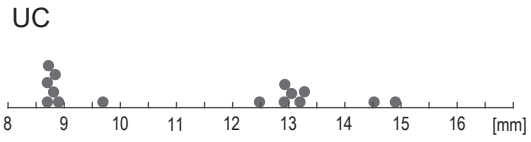


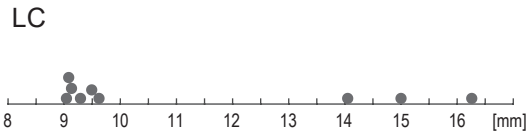
Figure S9_b

Nacholapithecus kerioi

UC:max, LC:mxob

sample distributions





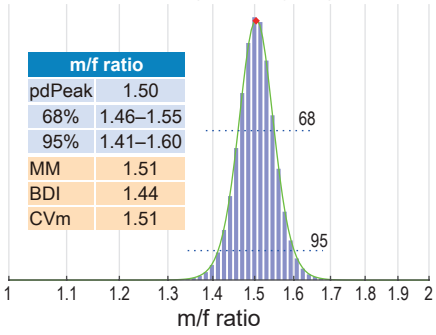
sample statistics

	UC	LC
N	15	9
mean	11.3	11.2
s.d.	2.41	2.97
CV*	21.7%	27.2%

* Sokal and Braumann (1980) correction

Posterior probability distributions

UC max (m/f marginal)

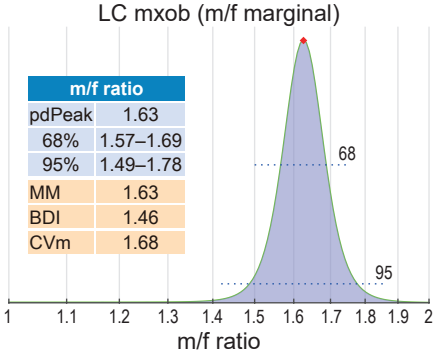


m/f ratio	
pdPeak	1.50
68%	1.46–1.55
95%	1.41–1.60
MM	1.51
BDI	1.44
CVm	1.51

within-sex CV for UC	
pdPeak	5.2%
68%	4.3–6.5%
95%	3.6–8.5%

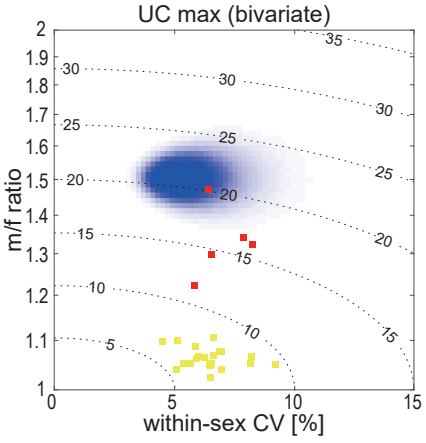
within-sex CV for LC	
pdPeak	4.5%
68%	3.4–6.3%
95%	2.6–9.9%

LC mxob (m/f marginal)



m/f ratio	
pdPeak	1.63
68%	1.57–1.69
95%	1.49–1.78
MM	1.63
BDI	1.46
CVm	1.68

UC max (bivariate)



LC mxob (bivariate)

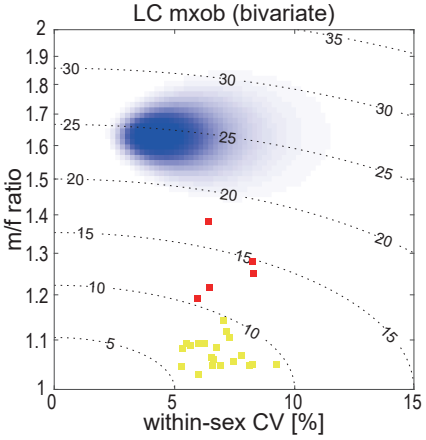
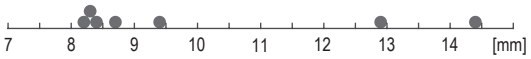


Figure S10_a

Hispanopithecus laietanus

UC:max

sample distribution

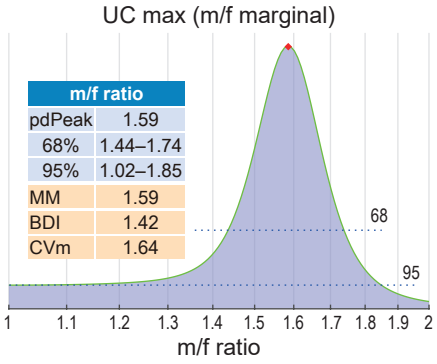


sample statistics

	UC	LC
N	7	
mean	10.0	
s.d.	2.53	
CV*	26.1%	

* Sokal and Braumann (1980) correction

Posterior probability distributions



within-sex CV for UC	
pdPeak	6.0%
68%	3.7–13.2%
95%	3.1–40.0%

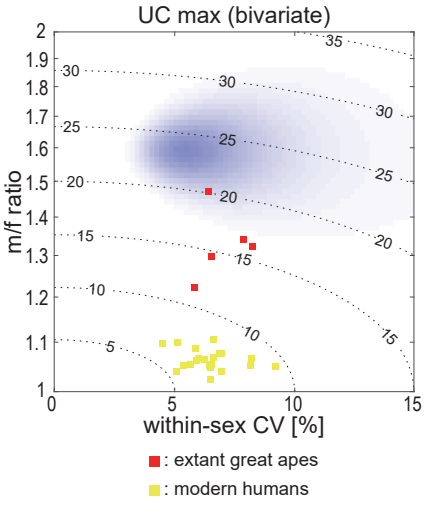
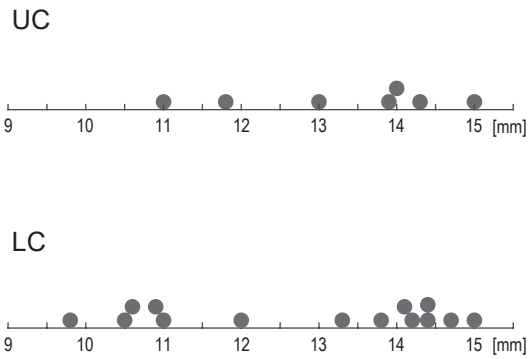


Figure S10_b

Ouranopithecus macedoniensis

UC:max, LC:mxob

sample distributions

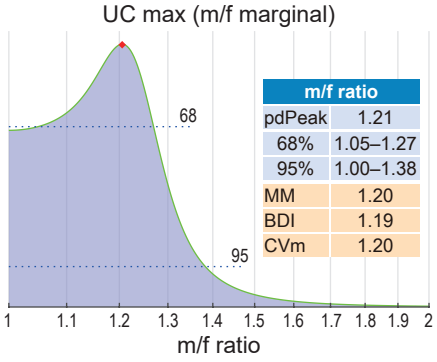


sample statistics

	UC	LC
N	7	14
mean	13.3	12.8
s.d.	1.44	1.86
CV*	11.2%	14.8%

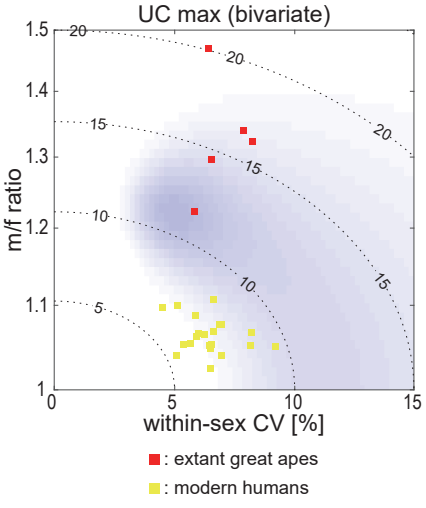
* Sokal and Braumann (1980) correction

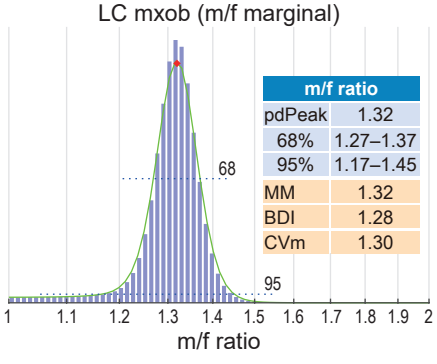
Posterior probability distributions



within-sex CV for UC	
pdPeak	9.2%
68%	4.3–14.4%
95%	3.0–26.5%

within-sex CV for LC	
pdPeak	5.3%
68%	4.1–7.1%
95%	3.2–13.7%





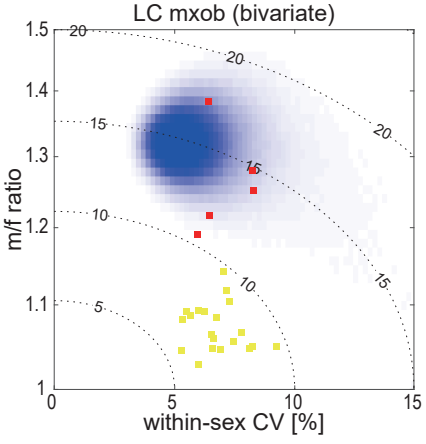


Figure S11_a

Oreopithecus bambolii

UC:max

sample distribution

UC

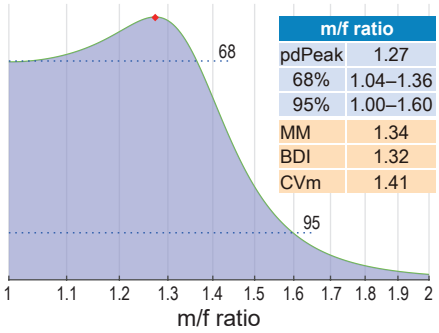
sample statistics

	UC	LC
N	8	
mean	8.6	
s.d.	1.57	
CV*	18.8%	

* Sokal and Braumann (1980) correction

Posterior probability distributions

UC max (m/f marginal)



m/f ratio	
pdPeak	1.27
68%	1.04–1.36
95%	1.00–1.60
MM	1.34
BDI	1.32
CVm	1.41

within-sex CV for UC	
pdPeak	15.1%
68%	9.3–23.0%
95%	6.4–37.9%

UC max (bivariate)

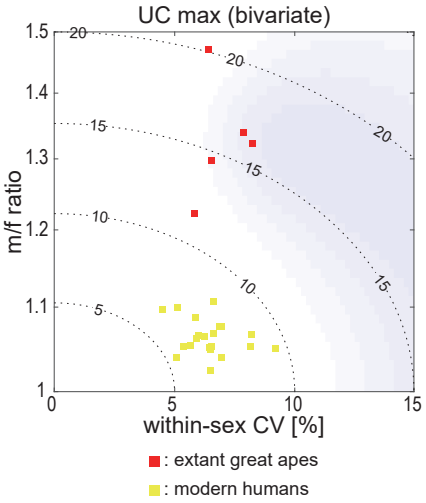


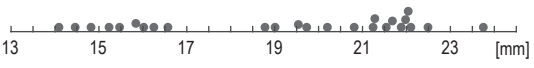
Figure S11_b

Gigantopithecus blacki

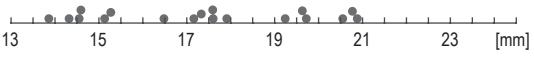
UC:max, LC:mxob

sample distributions

UC



LC



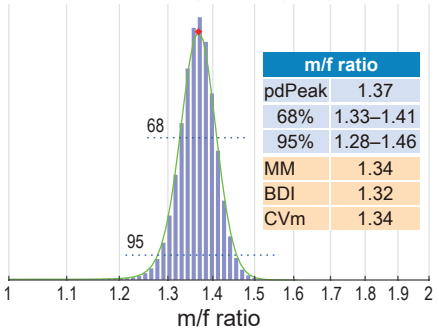
sample statistics

	UC	LC
N	25	18
mean	19.1	17.4
s.d.	3.04	2.38
CV*	16.1%	13.9%

* Sokal and Braumann (1980) correction

Posterior probability distributions

UC max (m/f marginal)

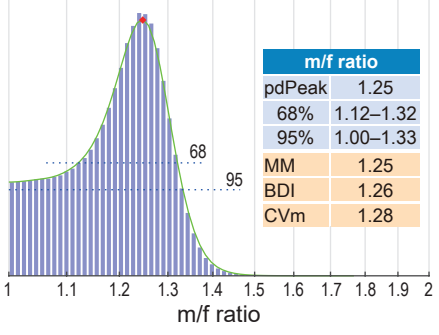


m/f ratio	
pdPeak	1.37
68%	1.33–1.41
95%	1.28–1.46
MM	1.34
BDI	1.32
CVm	1.34

within-sex CV for UC	
pdPeak	6.3%
68%	5.3–7.5%
95%	4.6–9.6%

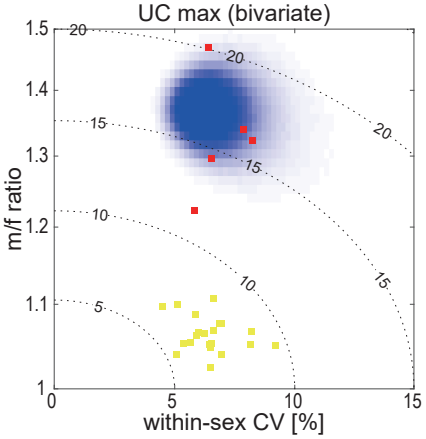
within-sex CV for LC	
pdPeak	8.3%
68%	6.7–13.5%
95%	5.5–18.4%

LC mxob (m/f marginal)



m/f ratio	
pdPeak	1.25
68%	1.12–1.32
95%	1.00–1.33
MM	1.25
BDI	1.26
CVm	1.28

UC max (bivariate)



LC mxob (bivariate)

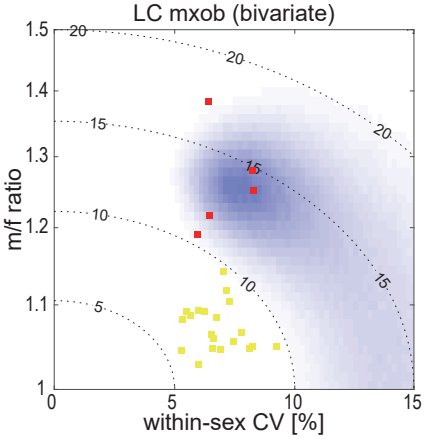


Figure S12

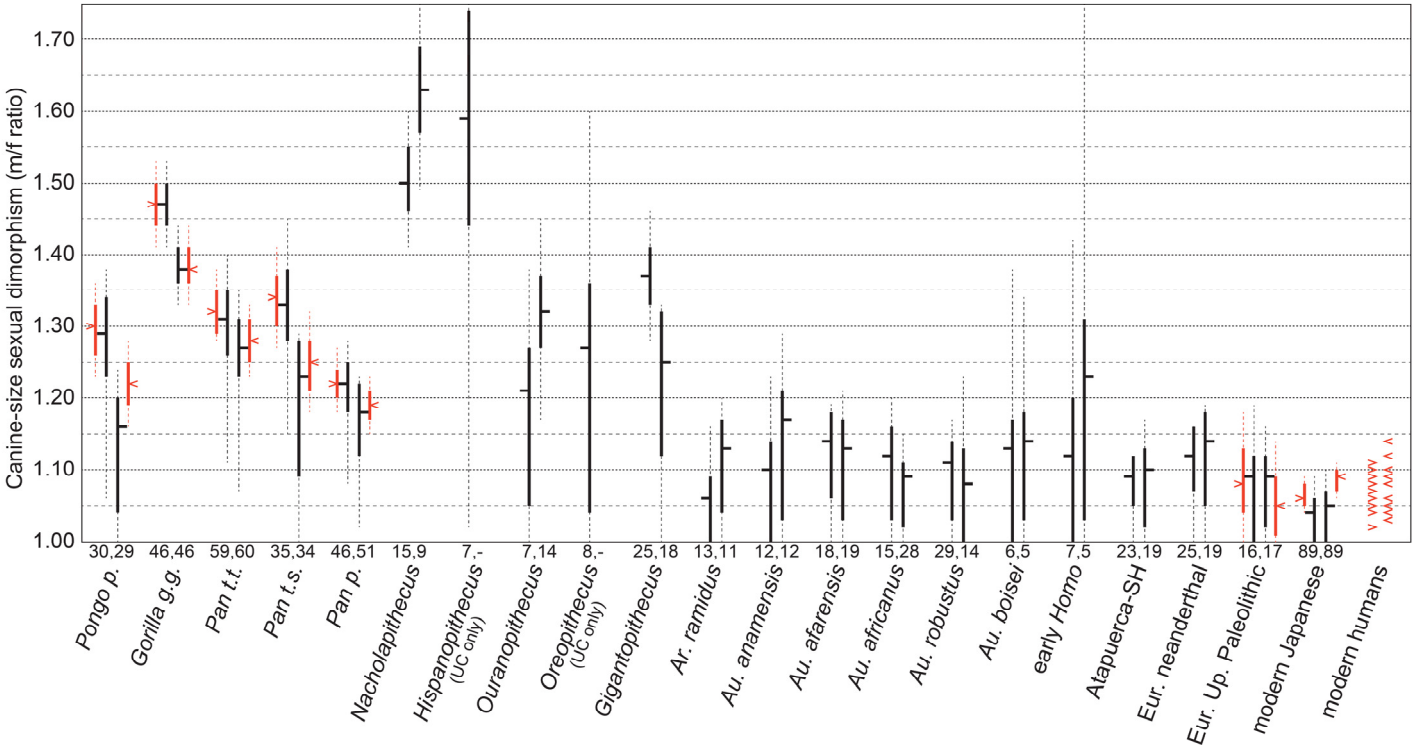


Figure S12. Comparison of the pdPeak estimates of the extant samples with their sample m/f ratios and CIs, and with the fossil taxa/populations.

The pdPeak method was applied to the extant samples by treating all specimens as unknown sex. The results are shown in black lines and ticks, together with the pdPeak estimates of the fossil taxa/populations (same as main text Fig. 4). The m/f ratios (male mean divided by the female mean) are shown by the horizontal tick, and the solid and dashed vertical lines indicate the 68% and 95% credible intervals, respectively. For each taxon/population, the upper and lower canines are shown on the left and right sides, respectively. Sample sizes are indicated below the vertical lines. The extant great ape and modern Japanese samples are from SI ref 2 and SI ref 3, and the other modern humans are from Dataset S1. *P. t. t.*, *P. t. s.* and *P.p* are *Pan troglodytes troglodytes, P. t. schweinfurthi*, and *Pan paniscus*, respectively. For the extant great ape and modern Japanese, the known-sex sample m/f ratios (arrowheads) and the 68% and 95% confidence intervals (CIs) are shown in red (solid and dashed vertical lines, respectively), and the same with the European Upper Paleolithic sample based on sex determination of associated pelves. The other modern humans are based on population means and shown as point values (red arrowheads).

Note that the credible intervals and CI ranges are either broadly equivalent, or the former (credible intervals) tend to be wider (i.e., more conservative). Note also that the pdPeak 95% credible intervals of the extant samples encompass the sample m/f ratios in all 14 known-sex cases that we examined, and in 11 of 14 cases with regards to the 68% credible interval. Overestimation seen in the European Upper Paleolithic sample reflects overestimation bias of the pdPeak method at low population m/f ratios, depending on within-sex CV level and sample size (as in main text Fig. 1). Underestimation seen in the *Pongo* lower canine and the modern Japanese samples is explained by either chance factor or a combination of chance factor and deviations from model assumptions (log-scale normality and homoscedasticity, equal source population sex ratio). The distributional properties of the *Pongo* and modern Japanese samples were examined in SI ref 3 Dataset S3, and a probable contributing factor is the “head-to-head” skew (male and female distributions that are positively and negatively skewed, respectively) especially in the *Pongo* sample (discussed briefly in SI ref 3). Note also that the generally narrower pdPeak credible intervals of the extant samples largely relate to larger sample sizes. The effects of sample size are discussed in Fig. S13.

Figure S13

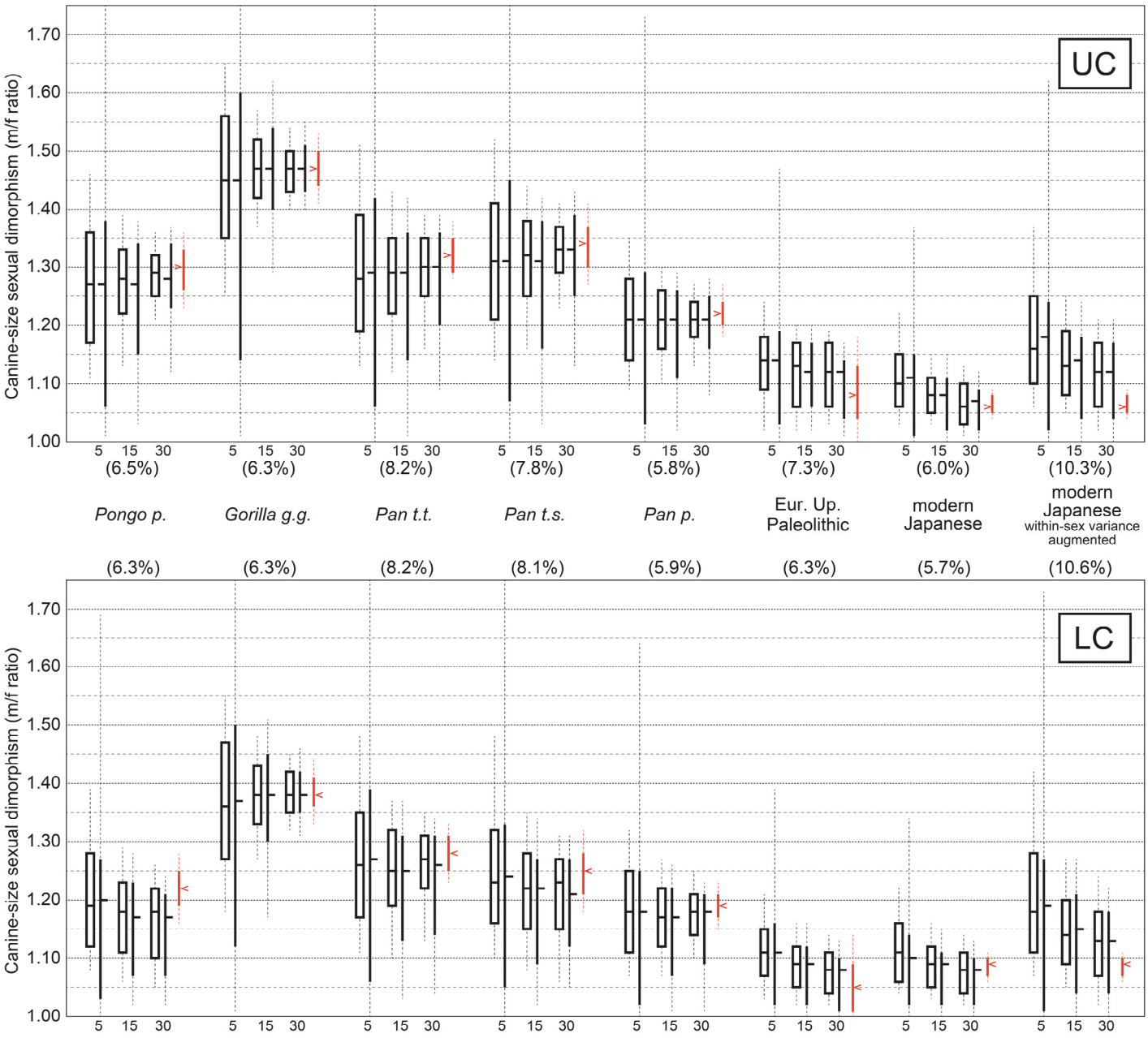


Figure S13. Sample pdPeak estimates and uncertainty level in relation to sample size. The effect of sample size (N) on the pdPeak estimates was evaluated by bootstrap resampling the extant great ape, European Upper Paleolithic, and modern Japanese canine samples. The upper panel is for upper canines and the lower panel is for lower canines. N was set at 5, 15 or 30, as indicated below the plots. The within-sex CVs (average of males and females) are shown in the parentheses. In deriving the pdPeak estimates, all specimens were treated as sex unknown. In this way, results of the pdPeak method by sample size can be cross-checked with the actual known-sex sample values and confidence intervals (CIs). The following resampling procedures were taken. For each taxon/population sample, sex assignments were made for N specimens with the probability of 0.5 for either sex (cases of uniform sex

(continued)

Figure S13 (continued from previous page)

were discarded and redone). Bootstrap samples were obtained by random sampling with replacement in each sex, according to the determined number for each sex. This was repeated 1,000 times. A small random value (normally distributed with SD of 0.1 mm) was added to avoid exact duplicate values in the sampled sets. First, using the 1,000 bootstrap samples, the pdPeak m/f ratio was estimated 1,000 times, and the 2.5%, 16%, median, 84%, and 97.5% percentiles were determined; these are shown by the box and whisker plots. Next, the credible intervals were derived for a smaller number of 100 bootstrap samples (because of computation load), and the averages of the (100 sets of) 68% and 95% boundary points were calculated; these are shown as the endpoints of the vertical solid and dashed lines, respectively, and with the horizontal tick indicating the average pdPeak m/f ratio. The red arrowheads show the actual known-sex sample m/f ratio values, and the red solid and dashed lines are the 68% and 95% CIs, respectively.

It can be seen that sample size clearly affects the uncertainty level (represented by the width between the credible interval boundary point averages, or by the pdPeak bootstrap percentile ranges). When sample size is 5, the uncertainty level is high in all the examined cases, but this is the case especially when the m/f ratio is large, as expected from the simulation results of main text Fig. 1. Note that with a large sample size of 30, both the credible interval boundary point averages and the bootstrap pdPeak estimates correspond generally well with the sample m/f ratios and CIs (the few slight discrepancies are discussed below). A sample size of 15 tends to exhibit intermediate levels of uncertainty. It is also seen that, when the m/f ratio is similar, the uncertainty level tends to be higher with larger within-sex variation. This is also the pattern expected from the simulations of main text Fig. 1. To examine the effect of a large within-sex variance, a hypothetical case was generated (far right plots) by elevating the within-sex CV of the modern Japanese sample to ~10% (by adding random values at an SD of 0.7 mm). The results show that the pdPeak m/f ratio estimate of a population with a modern human level of dimorphism but with a large within-sex CV would tend to lie between 1.1 and 1.2. The slight overestimation seen in the European Upper Paleolithic sample is best attributed to overestimation bias of the pdPeak method when the m/f ratio is <~1.1, within-sex CV moderately high, and sample size not large (as in main text Fig. 1). Underestimation seen in the *Pongo* lower canine sample cautions that m/f ratios as large as ~1.2 may at times return pdPeak estimates well below 1.2. However, even in these cases, note that the credible interval correctly includes an m/f ratio of >1.2.

Figure S14

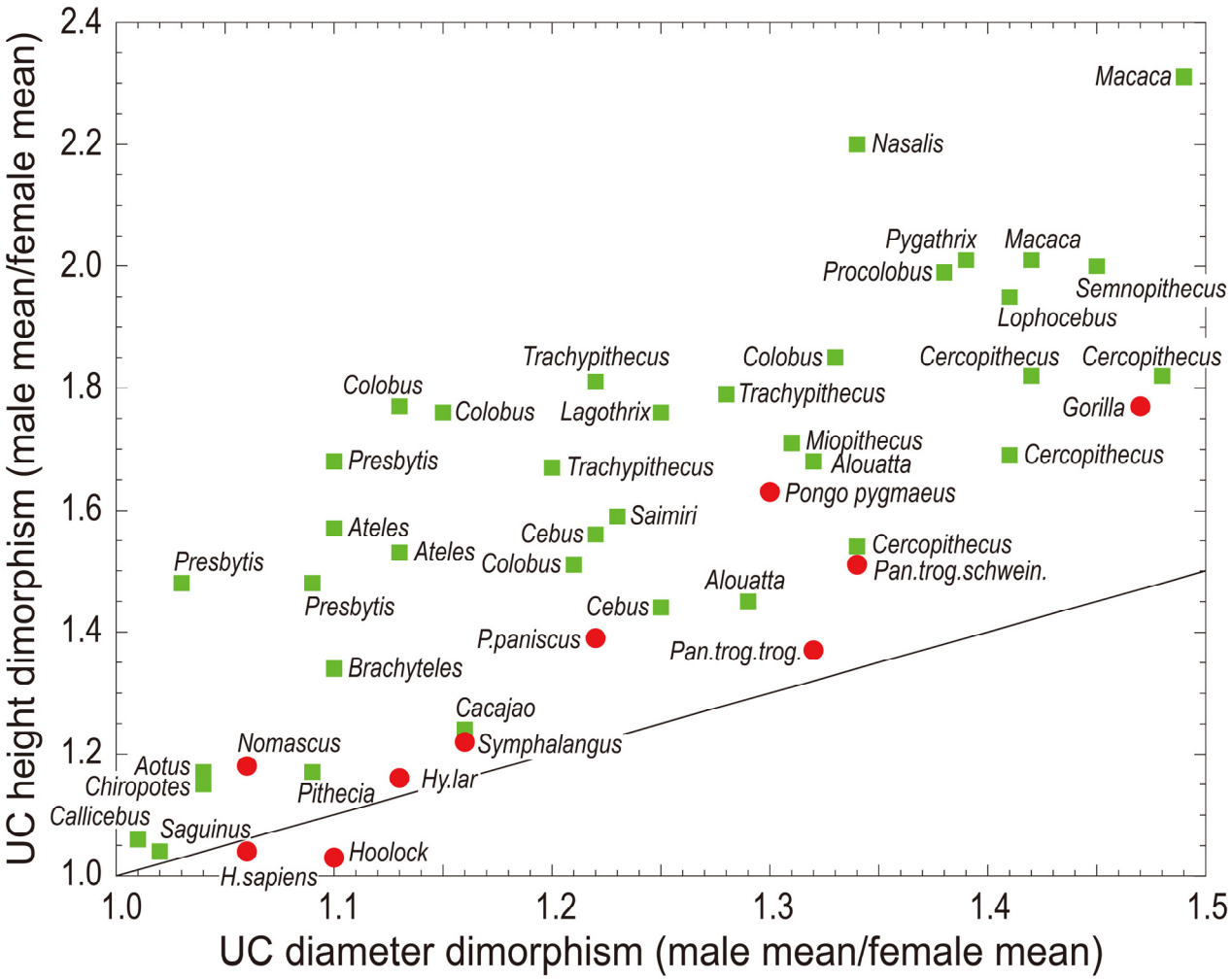


Figure S14. Bivariate plot of upper (maxillary) canine crown height dimorphism vs basal crown diameter dimorphism.

The ratios between male and female means of maxillary canine height (y-axis) and maximum basal crown diameter (x-axis) are plotted. Red circles, hominoids; green squares, New and Old World monkeys. The diagonal line indicates equivalent crown height and diameter dimorphisms. All input data are shown in Dataset S3; *Gorilla*, *Pan* and *Pongo* data are from SI ref 2, and the other non-human anthropoids are from SI ref 1. In order to obtain accurate as possible crown heights, the Plavcan 1990 dataset (SI ref 1) was conservatively screened, excluding worn specimens and outliers (see SI Text). Note that, in species with dimorphic canines, the canine is much more dimorphic in crown height than in basal (mesiodistal or maximum) diameter. This is observed in extant great apes in relatively less extreme form. All papionin species exhibit extremely enhanced crown height dimorphism, but most are not shown here because basal crown dimorphism is greater than 1.5. Many colobines and *Ateles* exhibit weak dimorphism in canine diameter, but strong dimorphism in crown height. In colobines, this may stem from their short faces linked to masticatory function, perhaps constraining crown diameter size but not crown height. The hylobatids and callithrichids have weak dimorphism from male-like female canines. Pitheciines have large robust canines in both sexes, probably related to diet.

Figure S15

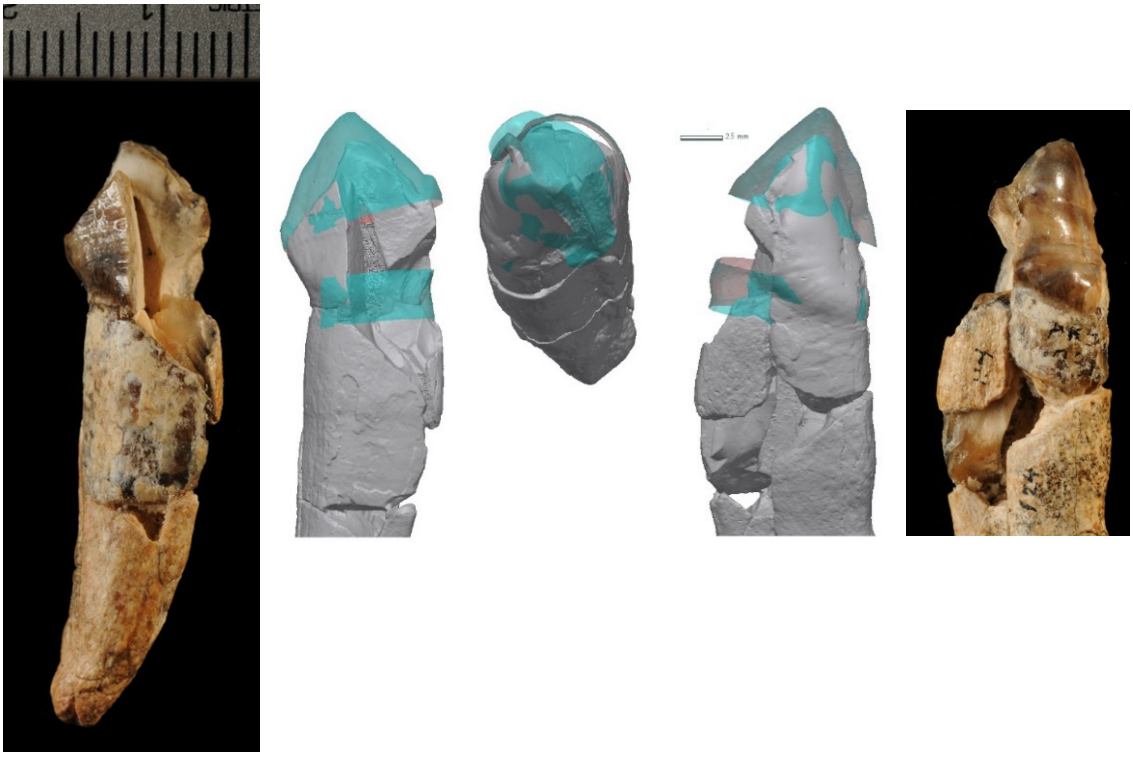


Figure S15. Estimating the ARA-VP-6/500 left upper canine crown diameter and height.

The left upper canine of ARA-VP-6/500 (the “Ardi” partial skeleton) was recovered in multiple pieces. Assembled, it lacks the cusp tip, the mesial and distal occlusal crests, and the entire mesiobuccal crown. However, the basal half of the distobuccal crown and much of the lingual crown surfaces are intact. Therefore, by juxtaposing well-preserved equivalent-sized tooth surfaces, it is possible to obtain estimates of crown diameters and height. This was done by the following procedures. 1) scale the near-unworn crown of the ARA-VP-1/300 canine to approximate ARA-VP-6/500 size by a factor of 0.9, 2) extract the well-preserved distobuccal to mesiolingual basal crown and root surfaces of ARA-VP-1/300, 3) 3d-best-fit-register this with the corresponding preserved surfaces of ARA-VP-6/500, 4) fine-tune the fit by translation so that the distobuccal and mesiolingual cervical lines match, 5) redo the 3d-best-fit but by constraining superoinferior translation and tilt (so as to preserve the cervical fits), 6) fine tune the distobuccal alignment by minimal translation without affecting the overall mesiolingual surface match. Next, in order to estimate crown height, 7) scale the unworn ARA-VP-6/1 canine to approximate ARA-VP-6/500 size by a factor of 0.855, 8) align the scaled ARA-VP-1/300 and ARA-VP-6/1, respectively, with the preserved distobuccal and lingual crown surfaces of the ARA-VP-6/500 canine by 3d-best-fit. 9) the thus aligned apical surfaces of the scaled ARA-VP-1/300 and ARA-VP-6/1 canine crowns, respectively, provides alternative probable cusp tip positions of the ARA-VP-6/500 canine. The resulting metrics were as follows. Maximum oblique diameter, 10.0 mm; labial crown heights 13.5 mm and 13.2 mm, respectively, by the ARA-VP-1/300 and ARA-VP-6/1 alignments. Micro-ct volumes of 42 or 30 micron resolutions were used to generate the surface polygon models, and the software Geomagic XOS was used in the surface alignments.

Figure S16

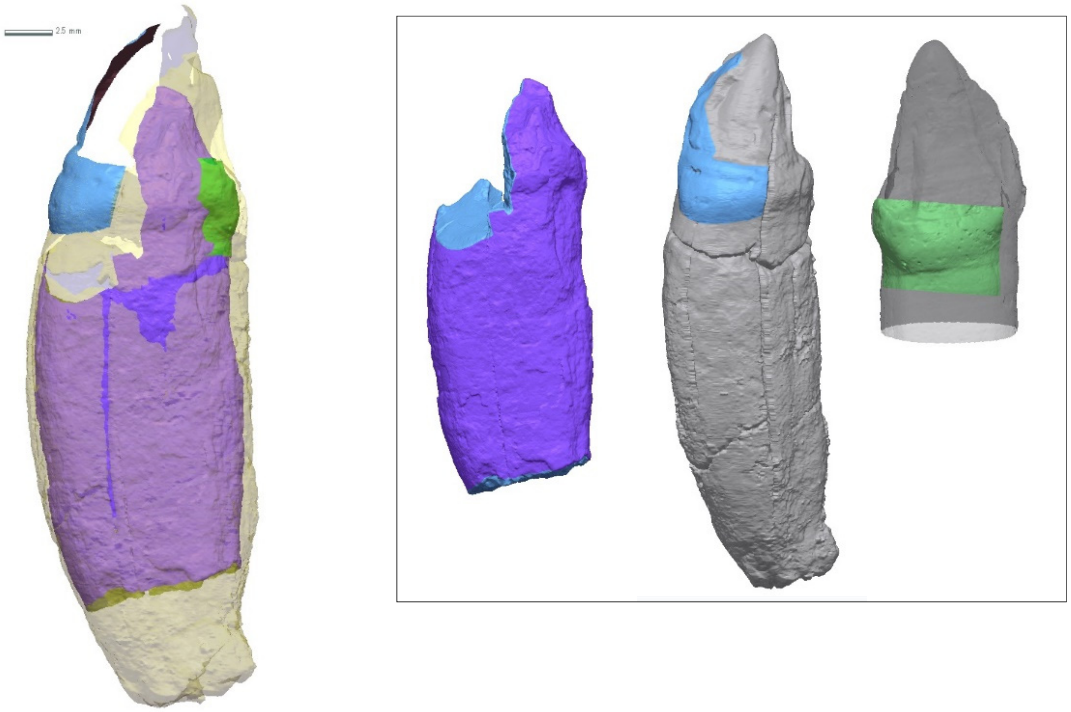


Figure S16. Estimating the ARA-VP-1/3293 left lower canine crown diameter. Much of the crown of the ARA-VP-1/3292 left lower canine (second from left) is not preserved, but the preserved small portion of distolingual crown allows approximation of the distolingual endpoint of the maximum oblique crown diameter metric. The entire buccal crown is also missing, but much of the root is preserved. We therefore estimated the ARA-VP-1/3292 maximum oblique (mxob) diameter by superimposing buccally the better preserved and overall similar-sized ARA-VP-6/500 lower canine (second from right), and distolingually a scaled ARA-VP-1/300 lower canine 3d surface model (far right). The procedures we took were as follows. 1) 3d-best-fit-register the buccal half of the ARA-VP-1/3293 and ARA-VP-6/500 roots, 2) translate this vertically so that the distal cervices match, 3) to compensate surface erosion of the ARA-VP-1/3293 buccal root surface, translate ARA-VP-6/500 buccally so that the model surfaces do not overlap, 4) the thus aligned ARA-VP-6/500 buccal crown surface (light blue of the far left and second from right images) is used in estimating the ARA-VP-1/3293 crown diameter, 5) reduce the ARA-VP-1/300 canine model to approximately ARA-VP-6/500 size by using a factor of 0.85, and 3d-best-fit-register with the ARA-VP-6/500 canine, 6) cut out the distolingual crown surface of the scaled ARA-VP-1/300 canine (green of the far right and left images), 7) translate this to visually best the ARA-VP-1/3293 crown surface, but avoiding overlap to compensate for surface erosion of the preserved ARA-VP-1/3293 distolingual crown. The resulting mxob metric estimate was 9.9 mm. Micro-ct volumes of 42 or 30 micron resolutions were used to generate the surface polygon models, and the software Geomagic XOS was used in the surface alignments.

Table S1. Fossil canine sample sizes of the present study.

Taxon/population	Age	Upper canines	Lower canines	Age and context references
Ardipithecus ramidus	4.3–4.6 Ma	13	11	SI ref 22, 23, 28
Australopithecus anamensis	3.8–4.2 Ma	12 (3)	12 (3)	SI ref 12, 24–27
Au. afarensis	3.0–3.75 Ma	18 (5)	19 (6)	SI ref 14
Au. africanus	2.1–2.6 Ma	15	28	SI ref 19
Au. robustus	~1–2.2 Ma	29 (9)	14 (3)	SI ref 16, 18
Au. boisei	1.3–2.3 Ma	6 (1)	5	SI ref 11
early *Homo*	1.7–2.4 Ma	7 (1)	5 (2)	SI ref 11
Atapuerca SH	~0.4 Ma	23 (23)	19 (19)	SI ref 9, 17
European *H. neanderthalensis*	0.04–0.15 Ma	25	19	approximate age range of this sample
European Upper Paleolithic *H. sapiens*	<0.04 Ma	16 (16)	17 (17)	SI ref 10
Nacholapithecus kerioi	15–16 Ma	15	9	SI ref 13
Hispanopithecus laietanus	9.6–9.7+ Ma	7	4	SI ref 8
Ouranopithecus macedoniensis	~9–9.7 Ma	7 (1)	14 (5)	SI ref 15
Oreopithecus bambolii	7–8 Ma	8	4	SI ref 21
Gigantopithecus blacki	~0.9–2 Ma	25	18	age range of this sample (SI ref 20, 29)

( ) number of specimens taken from the literature not involving the present authors

Datasets

Dataset S1
Modern human canine buccolingual (BL) diameter means, coefficient of variation (CV), and male mean/female mean ratio (m/f ratio).

Dataset S2
Summary sheet of pdPeak estimates, and specimen lists of the fossil samples.

Dataset S3
Extant anthropoid upper canine height metrics used in Figure 6 and Figure S14.

Dataset S4
Fossil canine height and other data used in Figure 7.

SI References

1. J. M. Plavcan, Sexual dimorphism in the dentition of extant anthropoid primates. Ph.D. dissertation, Duke University, Durham, NC (1990).

2. G. Suwa et al., Paleobiological implications of the *Ardipithecus ramidus* dentition. *Science* **326,** 94–99 (2009).

3. T. Sasaki et al., Estimating sexual size dimorphism in fossil species from posterior probability densities. *Proc. Natl. Acad. Sci. USA* **118** (2021).

4. J. Kelley, Sexual dimorphism in canine shape among extant great apes. *Am. J. Phys. Anthropol.* **96**, 365–389 (1995).

5. R. B. D'Agostino, Transformation to normality of the null distribution of g1. *Biometrika* **57**, 679–681 (1970).

6. S. C. Josephson, K. E. Juell, A. R. Rogers, Estimating sexual dimorphism by method-of-moments. *Am. J. Phys. Anthropol.* **100**, 191–206 (1996).

7. J. M. Plavcan, Comparison of four simple methods for estimating sexual dimorphism in fossils. *Am. J. Phys. Anthropol.* **94,** 465–476 (1994).

8. D. M. Alba et al., New dental remains of *Hispanopithecus laietanus* (Primates Hominidae) from Can Llobateres 1 and the taxonomy of Late Miocene hominoids from the Vallès-Penedès Basin (NE Iberian Peninsula). *J. Hum. Evol.* **62**, 231–246 (2012).

9. J. L. Arsuaga et al., Neandertal roots: cranial and chronological evidence from Sima de los Huesos. *Science* **344**, 1358–1363 (2014).

10. D. W. Frayer, Evolution of the Dentition in Upper Paleolithic and Mesolithic Europe. *Univ. Kansas Publ. Anthropol.* **10**, 1–201 (1978).

11. F. E. Grine et al. Complete permanent mandibular dentition of early *Homo* from the upper Burgi Member of the Koobi Fora Formation, Ileret, Kenya. *J. Hum. Evol.* **131**, 152–175 (2019).

12. Y. Haile-Selassie, S. M. Melillo, A. Vazzana, S. Benazzi, T. M. Ryan, A. 3.8-million-year-old hominin cranium from Woranso-Mille, Ethiopia. *Nature* **573**, 214–219 (2019).

13. Y. Kikuchi et al. Sexual dimorphism of body size in an African fossil ape, *Nacholapithecus kerioi*. *J. Hum. Evol.* **123**, 129–140 (2018).

14. W. H. Kimbel, L. K. Delezene, “Lucy” redux: a review of research on *Australopithecus afarensis*. *Yearbook Phys. Anthropol.* **52**, 2–48 (2009).

15. G. D. Koufos, L. de Bonis, D. Kugiumtzis, New Material of the hominoid *Ouranopithecus macedoniensis* from the Late Miocene of the Axios Valley (Macedonia, Greece) with some remarks on its sexual dimorphism. *Folia Primatol.* **87**, 94–122 (2016).

16. J. M. Martin et al., Drimolen cranium DNH 155 documents microevolution in an early hominin species. *Nature Ecol. Evol.* **5**, 38–45 (2021).

17. M. Martinón-Torres, J. M. Bermúdez de Castro, A. Gómez-Robles, L. Prado-Simón, J. L. Arsuaga, Morphological description and comparison of the dental remains from Atapuerca-Sima de los Huesos site (Spain). *J. Hum. Evol.* **62**, 7–58 (2012).

18. T. R. Pickering et al., New early hominin teeth from the Swartkrans Formation, South Africa. *J. Hum. Evol.* **100**, 1–15 (2016).

19. R. Pickering, A. I. R. Herries, “A new multidisciplinary age of 2.61–2.07 Ma for the Sterkfontein Member 4 australopiths” in *Hominin Postcranial Remains from Sterkfontein, South Africa, 1936-1995*, B. Zipfel, B. G. Richmond, C. V. Ward Eds. (Oxford University Press, New York, NY, 2020), pp. 21–30.

20. W. J. Rink, W. Wei, D. Bekken, H. L. Jones, Geochronology of *Ailuropoda–Stegodon* fauna and *Gigantopithecus* in Guangxi Province, southern China. *Quaternary Research* **69**, 377–387 (2008).

21. L. Rook, O. Oms, M. G. Benvenuti, M. Papini, Magnetostratigraphy of the Late Miocene Baccinello–Cinigiano basin (Tuscany, Itlay) and the age of *Oreopithecus bambolii* faunal assemblages. *Palaeogeography Palaeoclimatol. Palaeoecol.* **305**, 286–294 (2011).

22. S. Semaw et al., Early Pliocene hominids from Gona, Ethiopia. *Nature* **433**, 301–305 (2005).

23. S. W. Simpson, N. E. Levin, J. Quade, M. J. Rogers, S. Semaw, *Ardipithecus ramidus* postcrania from the Gona Project area, Afar Regional State, Ethiopia. *J. Hum. Evol.* **129**, 1–45 (2019).

24. C. V. Ward, M. G. Leakey, A. Walker, Morphology of *Australopithecus anamensis* from Kanapoi and Allia Bay, Kenya. *J. Hum. Evol.* **41**, 255–368 (2001).

25. C. V. Ward, F. K. Manthi, J. M. Plavcan, New fossils of *Australopithecus anamensis* from Kanapoi, West Turkana, Kenya (2003–2008). *J. Hum. Evol.* **65**, 501–524 (2013).

26. C. V. Ward, J. M. Plavcan, F. K. Manthi, New fossils of *Australopithecus anamensis* from Kanapoi, West Turkana, Kenya (2012–2015). *J. Hum. Evol.* **140**, 102368 (2020).

27. T. D. White et al., Asa Issie, Aramis and the origin of *Australopithecus*. *Nature* **440**, 883–889 (2006).

28. T. D. White et al., *Ardipithecus ramidus* and the paleobiology of early hominids. *Science* **326**, 75–86 (2009).

29. Y. Zhang, R. T. Kono et al. Evolutionary trend in dental size in *Gigantopithecus blacki* revisited. *J. Hum. Evol.* **83**, 91–100 (2015).