Supplementary Information: Drimolen cranium DNH 155 documents microevolution in an early hominin species

Jesse M. Martin[1*], A. B. Leece[1*], Simon Neubauer[2], Stephanie E. Baker[3], Carrie S. Mongle[4,5], Giovanni Boschian[3,6], Gary T. Schwartz[7], Amanda L. Smith[8], Justin A. Ledogar[9], David S. Strait[3,10†], Andy I. R. Herries[1,3]

1. Palaeoscience, Dept. of Archaeology and History, La Trobe University, Melbourne Campus, Bundoora, 3086, VIC, Australia

2. Max Planck Institute for Evolutionary Anthropology, Dept. of Human Evolution, Leipzig, Germany

3. Palaeo-Research Institute, University of Johannesburg, Auckland Park, Gauteng, South Africa.

4. Division of Anthropology, American Museum of Natural History, New York, NY, USA.

5. Turkana Basin Institute, Stony Brook University, Stony Brook, NY, USA.

6. Department of Biology, University of Pisa, Italy.

7. Institute of Human Origins, School of Human Evolution and Social Change, Arizona State University, Tempe, AZ, USA

8. Department of Organismal Biology & Anatomy, The University of Chicago, Chicago, IL, USA.

9. Department of Evolutionary Anthropology, Duke University, Durham, NC, USA.

10. Department of Anthropology, Washington University in St. Louis, St. Louis, MO, USA

* These authors contributed equally to this work.

† Author to whom correspondence should be sent.

CONTENTS

SUPPLEMENTARY TEXT

Description of DNH 155. DNH 155 was recovered from a single small area of partially decalcified sediment. A small part of the cranium was still calcified and attached to the abutting breccia wall demonstrating that the specimen had not moved from its original depositional context during decalcification. The cranium was found "upside down" such that the palate and maxillary teeth were facing the surface of the excavation. The face had fractured in roughly the midsagittal plane and the right side of the face had "fallen" inferiorly in the sediment, such that the face was sheared. Below the face, the neurocranium and cranial base were stacked in pieces organized into a few larger

fragments that internally preserved some anatomical relationships. Following excavation by Samantha Good, A B Leece, Stephanie Baker and Andy Herries, all fragments were carefully extracted from the sediment by Jesse Martin, who subsequently refit smaller fragments into six major cranial fragments. These are:

1. Left face
2. Right face
3. Left cranial base
4. Midline neurocranium
5. "Left" neurocranium
6. Occipital squama

These pieces were then assembled into two major segments, the face and neurocranium. These segments refit along fractures in the bones of the orbit, so the face can be confidently positioned relative to the neurocranium. As a generalization, preservation is superb. There is minimal plastic deformation, and fine osteological details are preserved. Several standard craniometric measurements were recorded following published standards[59] (Supplementary Table 6). Visual comparison with other notable *P. robustus* crania is provided in Supplementary Figure 1.

In the face, the superolateral angle (supraorbital corner) of the left orbit is nearly rounded, similar to (although less rounded than) that of DNH 7 and unlike in other adult *P. robustus* crania that exhibit "squared" supraorbital corners. Concomitantly, the superior orbital margins slope inferolaterally whereas the margin is more horizontal in other *P. robustus* adults. Overall, the orbits are taller than they are wide and in overall shape are

rhomboidal grading towards oval-shaped in frontal view. The inferolateral margin of the orbit is rounded as in most other *P. robustus* (inferolateral rounding is not obvious in DNH 7). The frontal process of the zygomatic is "flexed" in lateral view such that the profile of the orbit is concave facing anteriorly. Moreover, the frontal process faces anterolaterally and is notably wider inferiorly than superiorly, and the zygomatic angle is notched as in other *P. robustus* (although the distinctiveness of this notch varies between specimens). This angle is positioned above the inferior margin of the orbit. The zygomatic arch is tall, and does not flare extensively laterally, although the arch is not oriented in a parasagittal plane as in some *P. robustus*.

The zygomatic is inflated and projects anteriorly, but asymmetrically. On the left side, the zygomatic substantially obscures the nasal aperture in lateral view, but on the right side the zygomatic projects only at or slightly beyond the aperture. This asymmetry is not obviously caused by distortion, although the left zygomatic is better preserved. On the left, there are four zygomatic foramina distributed in a line across the body of the zygomatic running superolaterally from the zygomaticomaxillary suture that are located below or roughly at the inferior orbital rim. The inferomedial-most of these foramina is found within a zygomaticomaxillary fossa. On both sides, the contour of the midface grades smoothly from the zygomatic into a maxillary trigon, with the trigon being more strongly expressed on the left side. Unlike *P. robustus* from other sites, the superolateral margin of the trigon is not demarcated with a zygomaticomaxillary step that coincides with the zygomaticomaxillary suture. Rather, on the left, the inflection of the transverse contour as one moves medially from the zygomatic body into the maxillary trigon is

found above the suture on the zygomatic bone itself. Thus, the left zygomaticomaxillary suture is found within the trigon. On the less well-preserved right side, the suture is closer to the inflection of the transverse contour, but the suture itself does not sharply define a change in contour, and thus the right side also lacks a step. This gentle gradation between prominence and trigon is similar to what is seen in DNH 7, although the trigon is more weakly expressed in the latter, to the point of almost being absent. There is a blunt transverse rise running across the trigon just above the infraorbital foramen dividing it into superior and inferior depressions in both DNH 7 and 155, making the trigon shallow in this area.

Weak, broad and somewhat flattened anterior pillars are externally visible, as in other *P. robustus.* Computed tomography reveals that the pillars are solid as in other *P. robustus*. The margins of the nasal aperture are rounded inferiorly but become sharp and everted superiorly. The nasal bones are tall and generally narrow. They are widest superiorly and inferiorly, and narrowest in the middle of their height. As preserved, they project just above the frontomaxillary suture. The infraorbital foramen in DNH 155 is positioned in the lower half of the malar region, as in most other *Paranthropus,* just superior to a transverse plane passing through zygomaxillare. There is a furrow descending from the infraorbital foramen with an especially well-defined, everted medial margin. As the furrow descends, it crosses a sub-foramen divide and empties into a furrow-like maxillary fossula. Rak[12] recognizes a distinction between a maxillary furrow and a maxillary fossula, but the morphology of these depressions is quite variable in the southern African australopiths, possibly representing a complex interaction between palate protrusion and

the position of the zygomatic root. Thus, we conflate these two morphologies here. The zygomaticoalveolar crest in DNH 155 is straighter than the more arched condition seen in DNH 7. The masseteric origin is high, unlike in DNH 7, but in both specimens the masseteric tubercle is anterior to sellion. The tubercle and the anterior surface of the zygomatic root lie above P^4/M^1, as in DNH 7 but more posterior than in other *Paranthropus*. The center of the root and the masseter origin both lie above M^1. Unlike adult *P. robustus* from other sites, the zygomatic root is anteroposteriorly thin as it arises from the body of the maxilla above the alveolus. When seen in palatal view, the posterior surface of the root is only slightly inflated, as opposed to being extensively inflated (Extended Data Fig. 3). The root in DNH 7 is likewise anteroposteriorly thin. The lateral midfacial profile is steep and similar to other *P. robustus*, but the palate protrudes anterior to sellion to a greater extent than in *P. robustus* from other sites (Supplementary Table 2).

The shallow nasoalveolar gutter is subdivided into right and left portions by an elevated median sagittal ridge, as in DNH 7 and unlike most other *Paranthropus*. The floor of the nasoalveolar gutter is oriented more horizontally than in most other *P. robustus* and the incisor roots are more vertically oriented than the clivus floor. Thus, the surface of the floor meets the incisor roots at a blunt angle, and judging by the swellings of the roots in their alveolae, the tips of the roots curve posteriorly at their apices. Curved incisor roots are common in the DMQ sample. DNH 155 resembles other *Paranthropus* in that the palate is very thick (minimally 12 mm), the floor of the nasal cavity is smooth, and there is extensive overlap between the palate and nasoalveolar clivus. In basal view, the palate

is shallow anteriorly as in other *P. robustus.* The postcanine tooth rows diverge distally. The premaxillary suture is faintly visible in palatal view.

The incisor crowns project just anterior to, but overlap with, the bicanine line, as in DNH 7. The incisor row describes a gentle arch, as do the lingual margins of the incisor alveoli. With respect to the latter, DNH 155 resembles some other *P. robustus.* The tooth crowns are heavily worn and do not preserve details of occlusal morphology. The molars are large and buccolingually wider than they are mesiodistally long, although the mesiodistal dimensions have certainly been reduced by interstitial wear. The premolars are large, especially in comparison to the canines and incisors, which are small.

DNH 155 preserves part of the orbit, which exhibits a "foramen-shaped" superior orbital fissure. This trait has not previously been observed in any *P. robustus* specimen. The glenoid fossa is very deep and anteroposteriorly broad. The temporomandibular joint is positioned high above the occlusal plane. An extensive preglenoid plane extends anterior to the medial half of the articular eminence. A Eustachian process is present and prominent. The postglenoid process is very small, barely visible and fused to a vertical tympanic plate. The external auditory porus is laterally positioned, nearly circular and large (10.1 mm high, 9.1 mm long). An ossified styloid process is present. It is difficult to assess whether or not a prominent vaginal process was absent or if it was present and has been broken off. The petrous is oriented at 62 degrees relative to a mediolateral axis. This is a more sagittal orientation than is observed in any other *Paranthropus* specimen, but is similar to that estimated by us in DNH 7 (57 degrees). Thus, petrous orientation in

these specimens is like that of *Au. afarensis* and *Au. africanus* in being intermediate between the sagittal orientation seen in extant non-human apes and the more coronal orientation seen in other *Paranthropus* and most *Homo* (Supplementary Table 3). Inspection of the intracranial surface of the petrous pyramid reveals that its posterior face is roughly vertical, implying that the cerebellum was tucked anteriorly. A transverse-sigmoid sinus system is present, and there is no sign of an occipital sinus. Damage precludes direct assessment of the occipito-marginal sinus system in DNH 7 but this specimen exhibits sigmoid sinuses bilaterally. The margins of the foramen magnum in DNH 155 are preserved on three occipital fragments that are not presently joined with each other (and not all shown in Fig. 1), but it appears that a heart-shaped foramen is absent. The insertion of the longus capitis muscle is large and deeply excavated into the surface of the basioccipital, unlike the condition seen in *P. robustus* specimens SK 47, SK 48 and SKW 18. The supraglenoid gutter is wide.

A fracture through the frontal of DNH 155 reveals a frontal sinus. The face was hafted high on the cranium. DNH 155 exhibits strong anteromedial incursion of the temporal lines, creating the appearance of a "supraorbital rib" with a sharp crest along its posterior margin. As one moves posteriorly on the frontal of DNH 155, the temporal lines slowly converge and meet in the midline in a vertically short sagittal crest. DNH 7 exhibits only moderate incursion of the temporal lines and concomitantly lacks a sagittal crest, so these differences likely represent sexual dimorphism. It is likely that the anterior-most extent of the sagittal crest lies at or posterior to bregma, meaning that the crest arises on the parietal bone (note that the superior temporal lines converge on the frontal bone but do

not rise to form a crest). This is unlike the condition seen at SWM1HR, where on specimens SK 48 and SK 83 the crest arises on the frontal bone[11]. Thus, the sagittal crest is somewhat more posteriorly positioned in DNH 155 than in other putative male *P. robustus*. Immediately after having formed a crest, the temporal lines in DNH 155 start to diverge, but the divergence is associated with a subtle increase in the height of the crest, which becomes bifid. The crest is highest on the posterior half of the neurocranium. The bifid crests continue to diverge very slowly from each other across much if not all of the length of the parietal bone. The crest in OH5 is also bifid in this area, but taller. At no point does the crest reach the height seen in *P. robustus* specimen SK 46. Divergence of the temporal lines implies that there was a "bare area" and that a compound temporonuchal crest, if present, could only have been partial. Postorbital constriction in DNH 155 is comparable to that of SK 48 and SK 52, and somewhat more pronounced than in DNH 7 (Supplementary Table 7). The overlap of the parietal and temporal at the squamosal suture appears to have been extensive. There is a modest bevel at the inferior margin of the parietal, but a rugose arched line extends about 20mm superior to the inferior margin. That arch runs anteriorly to align with the preserved portion of the temporal squama. If the temporal squama did not extend up to this arched line, then minimally there must have been extensive connective tissues holding the suture together. An extensive overlap is evident in DNH 7 and DNH 152 as the parietals can be viewed endocranially. A parietal tuber is absent.

Feeding biomechanics in early hominins. Some of the features differing between the DMQ robust australopiths and those from Kromdraai B and SWM1HR appear to be

functionally related to feeding biomechanics. This can be deduced for three characters (the anteroposterior placement of the zygomatic root, the anteroposterior placement of the sagittal crest, and palate protrusion) using basic principles of jaw lever biomechanics[e.g., 83,84]. The key metric in this regard is mechanical advantage, which is a measure of muscular efficiency. As mechanical advantage increases, so, too, does the bite force produced by any given set of muscle forces. Mechanical advantage increases as the lever arms of the muscle forces increase relative to the load arm of the bite force, and mechanical advantage decreases as the lever arms decrease relative to the load arm (or, put another way, mechanical advantage decreases as the load arm increases). A simple way of thinking about this is that lever arms increase as the insertions of the masticatory muscles move further away from the temporomandibular joint, and load arms increase as the bite point moves further away from the joint.

Because the masseter muscle originates from the zygomatic arch, the lever arm of the masseter is constrained by the anteroposterior placement of the zygomatic root. A more anteriorly placed root therefore allows for a more anterior origin of this muscle, which should in turn increase masseter lever arm. Rak[12] noted that the zygomatic root is anteriorly placed relative to the tooth row in *P. boisei* and *P. robustus,* and Demes and Creel[84] and Eng *et al.*[85] calculated that these species have elevated masseter leverage relative to a second molar bite point compared to non-human great apes and gracile australopiths. Moreover, Ledogar *et al.*[14] performed a modeling experiment using finite element analysis and found that anterior displacement of the zygomatic root and masseter

origin resulted in higher bite force overall. The root is more posteriorly placed relative to the tooth row in DNH 7 and DNH 155 compared to *P. robustus* at other sites.

It is more difficult to estimate the mechanical advantage of the temporalis muscle, but Demes and Creel[84] found that *P. robustus* specimen SK 48 from SWM1HR had the highest temporalis mechanical advantage of any other early hominin or great ape that they examined. However, coarse measurements of mechanical advantage based on a few bony landmarks do not take into account variation in the architecture of the temporalis muscle. By virtue of their position, the anterior fibers of the temporalis muscle have longer lever arms than more posterior fibers, so increasing the size of the anterior temporalis should disproportionately increase the ability of the muscle to create bite force. The size of the anterior temporalis may be coarsely inferred by the course of the superior temporal line and the positioning of the anterior aspect of the sagittal crest, and as noted above, the sagittal crest in DNH 155 arises further posterior than does the crest in *P. robustus* from SWM1HR. DNH 152 resembles DNH 155 in this regard, and in DNH 152 the left superior temporal line on the frontal bone is clearly visible just off of and slightly oblique to the midline. In DNH 7 there is only moderate anteromedial incursion of the temporal lines. Thus, one may infer that the anterior temporalis was slightly smaller in robust australopiths from DMQ than in those from SWM1HR. This in turn implies that DMQ specimens would have had slightly less capacity to generate high bite force using this muscle.

According to jaw lever biomechanics[83,84], mechanical advantage is equal to the ratio of the lever arm to the load arm, so varying the tooth upon which one bites affects the efficiency of bite force production. It is obvious, therefore, that for any given skull, a bite on a premolar would occur with a lower mechanical advantage than a bite on a molar[e.g., 13,86]. The same basic principle applies when comparing two skulls, one of which has a more protruding palate than the other (implying that the entire tooth row is positioned farther away from the temporomandibular joint in the prognathic than the orthognathic skull). In this scenario, bites on the same tooth (e.g., M^1) would involve different load arms, with the longer load arm in the more prognathic skull. In this regard, the index of palate protrusion (which measures the degree to which the palate protrudes anterior to sellion) should have coarse biomechanical significance. We refrain from estimating this index in DNH 7 because of distortion, but as noted above the index is greater in DNH 155 than any *P. robustus* specimen from SWM1HR, implying a slightly longer load arm in that specimen for any given bite point on homologous teeth, and thus lower mechanical advantage.

The effect on mechanical advantage of the three traits described above is evident in finite element analyses of chimpanzees and gracile and robust australopiths[13,86]. Chimpanzees, with the most projecting faces, the most posteriorly positioned zygomatic roots, and the least well developed attachment sites for the anterior temporalis exhibit the lowest mechanical advantage for molar bites. Robust australopiths (represented by *P. boisei*), with the least protruding palate, the most anteriorly positioned zygomatic root, and the most extensively developed anterior temporalis marking exhibits the highest mechanical

advantage for such a bite. Gracile australopiths exhibit intermediate morphology and have intermediate values of molar mechanical advantage.

Moreover, the principles described above can be demonstrated quantitatively by manipulating standard indices used to describe facial morphology[12]. By dividing the Index of anterior position of the masseter by the Index of M^3 position, one can calculate the ratio of the masseter muscle lever arm divided by the load arm associated with a bite on the M^3. This value, when divided by two (assuming that the muscle can be said to act halfway along its origin), is equivalent to the leverage of the masseter muscle during an M^3 bite. DNH 155 exhibits leverage less than that of any measured *P. boisei* or *P. robustus* specimen (Supplementary Table 2).

A fourth biomechanical difference between the DMQ robust australopiths and those at SWM1HR and Kromdraai B concerns the posterior inflation of the zygomatic root. Mechanical modeling suggests that when a bite occurs directly below the zygomatic root, then the inferior surface of the root experiences high stress as the inferiorly directed pull of the masseter muscle combined with a reaction force at the biting tooth creates shear in the root. This would occur, for example, during molar biting in chimpanzees but premolar biting in *Paranthropus* (Supplementary Figure 2). However, when a bite occurs on a tooth that is positioned posterior to the root, then much of the lateral surface of the maxilla posterior to the root experiences elevated stress. Furthermore, a modeling experiment[14] in which the shape and position of the zygomatic root was changed in an *Au. africanus* model based largely on specimen Sts 5 shows that molar bites induce high

strains in the maxilla when the root is positioned anteriorly (Supplementary Figure 3). It is possible that a posteriorly inflated root structurally reinforces the lateral aspect of the maxilla against this stress by structurally stiffening the maxilla against shear in the sagittal plane between the root and the posterior alveolus.

The angle between the supraorbital torus and the frontal process of the zygomatic has been hypothesized to have biomechanical significance. A configuration in which these two "beams" are at right angles to each other, as when the supraorbital corner is squared, is said to be less effective at resisting stresses imposed by contraction of the masseter muscle[12]. In contrast, when these beams are more nearly aligned, as when the supraorbital corner is rounded or sloped, those structures should then be better equipped to bear tensile stresses produced by the masseter. DMQ robust australopiths exhibit the sloped or rounded morphology, while those at SWM1HR and Kromdraai B exhibit the squared morphology (Extended Data Fig. 5). Biomechanical modeling, however, does not find high strain magnitudes in this region during even maximal bites, and the shapes in question are not associated with meaningful differences in strain (Supplementary Figure 4). Thus, there is no support for the hypothesis that this trait is functionally related to feeding.

Dates for taxa in tip-dated Bayesian analysis. There is some slight uncertainty regarding the age ranges of *Paranthropus* taxa. We accept a FAD of *P. aethiopicus* as 2.7 Ma based on isolated teeth from Omo Shungura C[40] and a maxilla (EP 1500/01) from the Laetoli Ndolanya Beds[87]. The LAD of the species is difficult to define as Suwa[40,88]

identifies specimens from the lower segments of Omo Shungura G whose dental morphology is a mosaic of derived traits characteristic of the later *P. boisei* and earlier *P. aethiopicus*. We heuristically assign these specimens to *P. aethiopicus*[e.g., 89], leaving the LAD of this species as 2.3 Ma. This date might also serve as the FAD of *P. boisei*[40,88], although a fragmentary maxilla from Malema (specimen RC 911) has been assigned to *P. boisei* based on estimated tooth size and dates to 2.5 – 2.3 Ma[90]. Its taxonomic allocation should be viewed cautiously. The LAD for *P. boisei* is provided by specimens from Konso (including the partial skull KGA 10-525) dated to approximately 1.4 Ma[91]. The likely FAD for *P. robustus* at DMQ is ~2.04 Ma, but error associated with uranium-lead dates make possible a maximum FAD of 2.28 Ma. The LAD for *P. robustus* at DMQ is 1.95 Ma, as set by a paleomagnetic reversal in the DMQ sediments[3]. The likely FAD for *P. robustus* from other sites corresponds to SWM1HR and Gondolin and is likely near 1.8 Ma but a maximum older date of 2.3 Ma cannot be completely ruled out[26,92]. The LAD for *P. robustus* corresponds to Swartkrans Member 3, which has been dated using cosmogenic nuclides to be 0.96 Ma[93].

Given these uncertainties, we follow the convention set by other fossil tip-dating studies by accounting for some degree of error in the range estimates[e.g., 78,80]. Specifically, for each fossil taxon, we used maximal age ranges as the limits of a uniform distribution in the tip-dating analysis. As described above, the maximal range of *P. aethiopicus* was judged to be 2.7 – 2.3 Ma. Although the likely range of *P. boisei* is widely considered to be 2.3 – 1.4 Ma, we used a maximal range of 2.5 – 1.4 Ma for this analysis. Likewise, despite a narrow probable range of 2.04 – 1.95 Ma for *P. robustus* from DMQ , we incorporate the uncertainty associated with uranium-lead dating and use the maximal

range of 2.28 – 1.95 Ma. Similarly, while the likely range of *P. robustus* from sites other than DMQ was 1.8 – 0.96 Ma, we allow for uncertainty in this estimate and assign a maximum range of 2.3 – 0.96 Ma.

The tip-dated Bayesian inference analysis used a parsimony-based reconstruction of the last common ancestor of the *Homo* + *Paranthropus* clade as an outgroup taxon. This hypothetical taxonomic unit obviously does not have an age range that can be determined with confidence from the fossil record. We heuristically assigned the tree root age prior for this analysis as an offset exponential distribution with a mean of 4.0 Ma.

Maximum parsimony analysis. Maximum parsimony analysis of the same character data used in the Bayesian inference analysis yielded a single most parsimonious tree (Supplementary Figure 6a). That tree is nearly identical to the Bayesian most credible tree (Fig. 3a) except that the position of *K. platyops* is resolved such that it is the sister taxon of *Homo,* and the position of *Ardipithecus ramidus* is resolved such that it is the sister taxon of all later hominins. Bootstrap analysis reveals that some of the nodes in the most parsimonious tree are unstable, although clade support increases when two taxa that are missing data with respect to several characters (*Kenyanthropus platyops* and *Au. garhi*) are removed (Supplementary Figure 6b).

Probability of DMQ and SWM1HR robust australopiths being drawn from statistically identical populations. The robust australopiths recovered from DMQ appear to differ morphologically from those recovered from SWM1HR (Supplementary

Table 1), but before conclusions about taxonomy or evolutionary processes can be drawn, it is necessary to assess the probability that those differences are real or are sampling artifacts. The differences would be real if it was very unlikely that the DMQ and SWM1HR samples were drawn from populations that were the same with respect to statistical parameters like mean morphology, while the differences would be artifacts if it was reasonably likely that such differences could be drawn from populations with the same statistical parameters. The possibility of the differences representing sampling artifacts should be considered seriously given that sample sizes from DMQ and SWM1HR are small, and that only two samples are available.

An examination of the characters that vary within *Paranthropus* (Supplementary Table 1) shows that 13 such characters can be observed in both the DMQ and DWM1HR samples, and that the samples differ morphologically without overlap with respect to 10 of them. A test of the null hypothesis that there are no differences between the DMQ and SWM1HR populations would entail determining the probability of sampling 10 out of 13 differences in small samples drawn from these populations. This probability can be calculated using basic principles of probability theory[e.g., 94].

Let us adopt very conservative assumptions concerning samples. Most of what is known about cranial morphology among the DMQ robust australopiths derives from two specimens, DNH 7 and DNH 155. More specimens preserving cranial morphology have been recovered from SWM1HR, but those specimens are more fragmentary than those at DMQ. Let us simplify the example and consider two fossil samples, each containing

only two individuals. By any standard, these are small samples. Let us further assume that the two samples are drawn from statistically identical populations in which balanced polymorphisms exist with respect to the thirteen characters such that half of the population exhibits state 0 with respect to any given character, and half of the population exhibits state 1. Assuming populations with this structure maximizes the probability of observing the pattern seen in the DMQ and SWM1HR samples, in which they differ with respect to 10 out of 13 characters. This therefore allows a conservative test of the null hypothesis.

If Sample X consisted of two individuals drawn randomly from Population X, and if Sample Y consisted of two individuals drawn randomly from Population Y, and if each individual could be scored for 13 characters (A – M), and if each character could exhibit either of two states (0 or 1), then for any given character there are 16 permutations of how that character could be expressed in the four specimens collectively present in Samples X and Y. For example, if with respect to Sample X the specimens both exhibited state 0 for Character A, then there are four permutations of how character A could be expressed in the specimens in Sample Y (00, 01, 10, 11). But there are four different permutations of how Character A can be expressed in Sample X (also 00, 01, 10, 11). For each of those permutations in Sample X, there are four permutations in Sample Y, and hence a total of 16 possible permutations for the expression of Character A in Samples X and Y. Note, however, that there are only 2 of these 16 permutations in which Samples X and Y are morphologically different without overlap (Sample X is 00 and Sample Y is 11, and Sample X is 11 while Sample Y is 00). Thus, assuming random

sampling, there is only a $1/8^{th}$ chance for any given character that Sample X and Sample Y will be different in the manner that the DMQ and SWM1HT samples are different.

The probability *p* of sampling at least 10 out of 13 characters in which Samples X and Y differ can then be calculated as:

$$p = C_{(13,13)}(1/8)^{13}(7/8)^{0} + C_{(13,12)}(1/8)^{12}(7/8)^{1} + C_{(13,11)}(1/8)^{11}(7/8)^{2} + C_{(13,10)}(1/8)^{10}(7/8)^{3}$$

where $C_{(n,r)} = n!/r!(n-r)!$

In this case, $p \approx 0.0000002$. Using conventional standards of statistical significance, one can reject the null hypothesis that Samples X and Y and, by extension, DMQ and SWM1HR were drawn from populations with identical statistical parameters. From a taxonomic standpoint this leaves only two possible interpretations: DMQ and SWM1HR sample different species, or they sample populations of the same species that differ from each other morphologically. The latter implies that there has been evolution within the species.

A caveat to the above analysis is that this model assumes that all 13 characters are independent of each other. This may not be true (although character independence is a common assumption in paleobiology), in which case the calculated value of *p* would be greater. However, this possibility would be offset if the distribution of the character states in a given population were uneven such that one state (i.e., 0) is more common than

the other state (i.e., 1). If this were the case, then the calculated value of *p* would be less. The basic fact remains that sampling a rare event many times is vanishingly improbable.

Sexual dimorphism, ontogeny, social behavior and taphonomy. Keyser[1] noted that the DMQ *P. robustus* cranial sample differed from that of from Kromdraai B and SWM1HR insofar as DNH 7 lacks a sagittal crest, has weakly expressed or absent anterior pillars, has rounded rather than squared orbital margins, and is appreciably smaller overall than the smaller specimens from the other sites (e.g., SK 48, TM 1517). Similarly, Moggi-Cecchi et al.[42] observed statistically significant size differences between the DMQ maxillary dental assemblage and those from SWM1HR and Kromdraai B, with the DMQ sample being consistently smaller. Collectively, these cranial and dental differences have been hypothesised to be reflective of a highly sexually dimorphic population, with females being over-represented at Drimolen, and males being over-represented at SWM1HR and Kromdraai B[1,11,42].

Lockwood et al.[11] examined all SWM1HR and Kromdraai B maxillary specimens from which a set of several facial measurements could be taken (SK 12; SK 46; SK 48; SK 83; TM 1517) and found them all to be males, with DNH 7 from DMQ being the only female among the better preserved faces. The additional qualitative character traits that differ between the DMQ robust australopiths and those from SWM1HR and Kromdraai B were also argued to be effects of sexual dimorphism[11]. The degree of sexual dimorphism within the combined SWM1HR, Kromdraai B, and DMQ sample was found to exceed

the degree of sexual dimorphism present in extant *Gorilla gorilla*[11]. Consequently, Lockwood et al.[11] suggested that the extreme sexual dimorphism apparently evinced by *P. robustus* sample may be the result of bimaturism, whereby growth is extended into adulthood as the result of selection favoring large size in the presence of high male-male competition. In support of this hypothesis, Lockwood et al.[11] noted that the largest male in the combined sample from Swartkrans (SK 12) also appeared to be the oldest based on dental wear patterns.

While the hypotheses concerning the degree of sexual dimorphism within *P. robustus* can, at least theoretically, explain size variation within *P. robustus* as a whole, they do not explain why DMQ should contain mostly females, while SWM1HR and Kromdraai B should contain mostly males. In order to explain this extreme collection bias, Lockwood et al.[11] and Moggi-Cecchi et al.[42] have proposed a nested set of hypotheses. Specifically, Lockwood et al.[11] considered that SWM1HR represented a carnivore accumulation after Brain[21] and thus concluded that young male members of *Paranthropus robustus* must have been living independently from the broader family unit making them more vulnerable to predation. Thus, size differences between the DMQ and SWM1HR samples can be attributed to sexual dimorphism combined with taphonomic processes resulting in "a relative abundance of male specimens" at SWM1HR[42:404].

The key components of this hypothesis are that *P. robustus* exhibits a greater degree of sexual dimorphism than *Gorilla gorilla*; *P. robustus* exhibits extreme sexual dimorphism as a result of extended male growth or bimaturism as a response to high male – male

competition in a harem like social structure; young male *P. robustus* would have been especially vulnerable to predation as a result of a social structure that pushed young males out on their own; SWM1HR is a carnivore accumulation and thus preserves predominately male *P. robustus* specimens. In other words, the hypothesis that SWM1HR represents a carnivore accumulation informed interpretations that "the social structure of a highly dimorphic species" resulted in "nondominant males" becoming isolated and so at a greater risk of predation[42:404; see also 11,95].

The hypothesis that males are over-represented at SWM1HR relies critically on the premise that DNH 7 is generally representative of females across the species. However, the presence of an old male at DMQ (DNH 155) that is smaller than all of the better-preserved putative males at SWM1HR and Kromdraai B suggests that perhaps there has simply been a phyletic increase in size within the *P. robustus* lineage. Maxillary dental metrics are compatible with this view. Thus, taphonomic processes related to inferred patterns of social behaviour and ontogeny[11] are not needed to explain the differences between DMQ and SWM1HR. Under this scenario, it might be possible that some of the putative males at SWM1HR are, in fact, females.

The Unified Species Concept, palaeoanthropology, and *P. robustus*. The Unified Species Concept[39, see also 38] suggests that most of the species "concepts" that have dominated debate about species delimitation (e.g., Biological Species Concept[55], Evolutionary Species Concept[37], Recognition Species Concept[97], Phylogenetic Species Concept[98], Cohesion Species Concept[99]) are not in fact concepts but rather identify

criteria (e.g., reproductive isolation, specific mate recognition systems, diagnosability, cohesion) used to define boundaries between groups that are then called species. Because different species criteria necessarily identify different groups, these species concepts appear to at least occasionally be in conflict with each other. However, the Unified Species Concept points out that all competing "concepts" share a common element, namely, that species are independent lineage segments linking metapopulations from different time periods via ancestry and descent. Thus, at its core, the Unified Species Concept is the Evolutionary Species Concept[37] but with an additional epistemological insight regarding the role of species criteria. Namely, species originate as lineages diverge from each other and eventually that divergence becomes so complete that there is little disagreement about recognizing ancestral and descendant species (Supplementary Figure 7). However, the initial stages of lineage divergence involve the sequential evolution of species criteria (diagnosability, reproductive isolation, etc.) in each new lineage segment, and those criteria may evolve in any order and over an unspecified length of time. Thus, the early stages of lineage differentiation represent a "grey area" where different species "concepts" might delimit species differently, leading to disagreement. Yet, ontologically, a species is an independent metapopulation lineage segment. Thus, the question of whether or not there is more than one species in a given sample is in fact the question of whether or not there is more than one lineage segment. Put in a hypothetico-deductive framework, a single species (i.e., a single lineage segment) is the null hypothesis and multiple species are present when the null hypothesis can be rejected. The various species criteria each represent distinct lines of evidence that can be used to reject the null hypothesis. According to this logic, why would the

Phylogenetic Species Concept[98], the concept used most extensively (but only rarely acknowledged[100]) by palaeoanthropologists, be problematic? After all, diagnosability is a species criterion, so diagnosably distinct populations or samples would seem to represent distinct lineage segments. However, the logic of the Unified Species Concept makes clear that diagnosability does not allow rejection of the single species null hypothesis under all circumstances.

When diagnosably distinct populations are contemporaneous (as are all populations of organisms extant today), then such *horizontal diversity* allows rejection of a single lineage segment null hypothesis because lineage divergence is the phenomenon that best explains fixed, population-level morphological differences. However, when diagnosably distinct populations are not contemporaneous, then such *vertical diversity* does not allow rejection of the null hypothesis because the idea of the species as a lineage segment specifically incorporates the prediction that early and later parts of a lineage segment may be diagnosably distinct from each other. Thus, if diagnosability evolves in the "grey area," then it follows that the parts of the lineage segment before and after this event are morphologically different from each other. There is no specific prediction about how long it takes for diagnosability to evolve, so there is no reason to suppose that early and late members of a lineage segment cannot be sampled in the fossil record. Moreover, there is more than one manner in which diagnosability might evolve. Early populations in a lineage segment could be characterized either by polymorphisms that eventually become fixed, or by the sequential evolution in stages of the derived traits that will eventually come to characterize the later parts of the lineage segment. Thus, in the

specific context in which palaeopopulations do not overlap in time, diagnosability alone does not allow a straightforward rejection of the null hypothesis that only a single species lineage segment is present. Indeed, it has long been recognized that the Phylogenetic Species Concept may overestimate the number of distinct species under these circumstances[101].

These considerations allow a simple morpho-chronological test for species delimitation. In particular, for a given time period, the number of temporally overlapping, diagnosably distinct palaeopopulations represents the minimum number of species / lineage segments. Note that this test does not identify all of the species within a given clade, but rather it identifies those in which one may have epistemological confidence. The test is not novel in palaeoanthropology[71,102] but has rarely been formalized[103], and when applied to the human fossil record this test provides clarity with respect to the recognition of at least certain hominin species. For example, Neanderthals and anatomically modern humans are diagnosably distinct and contemporaneous[104], so the null hypothesis that they represent a single species can be rejected even though they empirically are not reproductively isolated from each other[105]. Rather, they represent distinct lineage segments in the grey area of divergence before all species criteria have evolved, and should be recognized as *Homo neanderthalensis* and *H. sapiens*, respectively.

In the absence of contemporaneity, does morphology provide any clue about species delimitation? Yes, because the Unified Species Concept implicitly limits the manner in which ancestral and descendant metapopulations should differ from each other. Ideally,

members of the same lineage segment should share most of a unique suite of primitive and derived characters, with the exception that early members could exhibit primitive states with respect to some traits while later members could exhibit derived states. Moreover, early members of a lineage segment should not exhibit obvious autapomorphies that are absent in later members[41,106], because it is more parsimonious to assume that such traits evolved once in a divergent lineage than to have evolved and then were lost in a single lineage. Temporally successive fossil assemblages that do not meet these two conditions would not be good candidates to be considered conspecific, and could be used as a basis for provisionally rejecting a single species null hypothesis.

Other considerations might be relevant to assessing the likelihood that diagnosably distinct and temporally non-overlapping palaeopopulations represent different species. First, it goes without saying that a palaeopoulation in an early time period can belong to the same lineage segment as only one palaeopopulation in the next successive time period. Namely, if that later time period contains multiple, contemporaneous, diagnosably distinct palaeopoulations, then only the one that best matches the two morphological conditions described above should be considered a candidate as a continuation of the lineage segment. Phylogenetic analysis might also inform species delimitation insofar as it reveals branching patterns that may or may not be compatible with the hypothesis that two or more operational taxonomic units belong to the same lineage segment. Monophyly among palaeopopulations is obviously compatible with a single species null hypothesis, but some systematists argue that paraphyly is compatible as well[38]. Polyphyly is presumably incompatible with the null hypothesis. Note,

however, that phylogenetic relationships are inferences and are not data in the same sense that morphology and chronological age are data. Thus, the rejection of a single species hypothesis on the basis of phylogenetic relationships is somewhat more conditional than a rejection based on diagnosability and time.

A recent fossil hominin discovery provides an excellent example of how the Unified Species Concept can be applied to palaeoanthropology in a hypothesis-testing context. Up until just recently[71], it has been argued that because the gracile australopiths from Kanapoi, Allia Bay, Laetoli and Hadar exhibit a pattern in which the later samples possess more derived states than the earlier samples such that the samples are paraphyletic in parsimony analysis, then *Au. anamensis* and *Au. afarensis* are effectively chronospecies of the same lineage segment[41]. A chronospecies is not, however, ontologically valid under the Unified Species Concept, so if these assemblages legitimately sample a single evolving lineage segment, then they should all be recognized as *Au. afarensis*[41] even though current taxonomic convention retains them in distinct species. However, the recent description of a nearly complete fossil cranium of *Au. anamensis* has led to the suggestion that this taxon and *Au. afarensis* overlapped in time, and that *Au. anamensis* possessed some derived features that might not be expected in a direct ancestor of *Au. afarensis*. The implication is that the fossils typically attributed to these taxa do not, in fact, simply represent early and late members of the same lineage segment[71]. Ultimately, future fossil discoveries will determine the validity of this interpretation (about which we are neutral), but its underlying logic is compatible with the rejection of a single species null hypothesis under the Unified Species Concept.

Regarding Drimolen, the robust australopiths from DMQ temporally predate those from all other South African sites (SWM1HR, Swartkrans Members 2 and 3, Kromdraai B Member 3, Coopers D, Sterkfontein Member 5, Gondolin). They further share most of a unique suite of primitive and derived traits with the later specimens, and differ only in that they exhibit putatively primitive states with respect to a few characters in which the later South African robust australopiths are presumably derived. Moreover, the DMQ robust australopiths lack obvious autapomorphies discriminating them from the later robust specimens. Finally, when treated as an independent unit in phylogenetic analysis, the DMQ robust australopiths are either paraphyletic or monophyletic with the other South African robust australopiths (Fig. 3). There is therefore no basis on which to reject the null hypothesis that the DMQ robust australopiths represent an early palaeodeme of the *P. robustus* species lineage, despite the fact that the DMQ sample is diagnosably distinct from the collective sample drawn from the later sites.

The logic employed above can be applied to other hominins. Late Pleistocene Neanderthals exhibit a suite of derived traits that are variably present in a mosaic fashion in the archaic humans that immediately precede them in Europe[104]. Those Middle Pleistocene Europeans are sometimes recognized as a distinct species, *H. heidelbergensis*, but their morphology and temporal range are compatible with them representing an early portion of the Neanderthal lineage segment in which derived Neanderthal traits are present as polymorphisms[104]. One therefore cannot reject a single species null hypothesis, and many of those specimens (minimally, those possessing at

least some Neanderthal apomorphies) should be allocated to *H. neanderthalensis.* Similarly, it would be difficult to falsify the hypothesis that *Ardipithecus kadabba* is the direct phyletic ancestor of *Ar. ramidus* within a single lineage segment, implying that the former should be subsumed into the latter. Moreover, the Unified Species Concept implies interesting and open questions as to whether one should re-evaluate whether australopiths from Bouri typically allocated to *Au. garhi* are distinct at the species level from temporally earlier *Au. afarensis,* that australopiths from Malapa typically allocated to *Au. sediba* are distinct from temporally earlier *Au. africanus* (a possibility compatible with a recent study of facial growth[107]), and that eastern African robust australopiths typically allocated to *P. aethiopicus* are distinct from temporally later *P. boisei.* Certainly, the Malapa hominins are a better candidate to be a member of the *Au. africanus* species lineage than are the temporally overlapping South African robust australopiths, which do not share most of an obvious suite of primitive and derived traits with *Au. africanus,* which possess some features that are plausibly more primitive than those of the latter taxon (i.e., an anteriorly shallow palate), and which are polyphyletic with *Au. africanus* in phylogenetic analyses[e.g.,70].

Note that our argument regarding species delimitation is independent of the long-running debate in palaeoanthropology regarding the appropriateness of taxonomic "lumping" vs. "splitting" because that debate is not formally grounded in species concept theory. Rather, that debate concerns how best to use empirical patterns of morphological variation in living species to interpret variation in the fossil record[e.g.,108]. This approach has a long pedigree in palaeontology[109] and is especially useful when interpreting

differences between contemporaneous fossil assemblages. However, there is no explicit prediction concerning the magnitude of morphological change that can accumulate within a single evolving lineage segment, so it is unclear how to apply such information when assessing temporally non-overlapping fossil assemblages. Of course, one could apply an arbitrary criterion such that change beyond a certain magnitude (informed by variation in extant species) would be interpreted as evidence that a new, independent lineage segment had evolved, but the ontological basis for such an inference is unclear. Taken to one extreme, the logical extension of this viewpoint is that the evolution of any novel characteristic is evidence of a new species[e.g.,110] and that anagenesis within a species does not occur (or, at least cannot meaningfully be observed). We contend, however, that the logic of the Unified Species Concept ensures that change within lineage segments should occur and be potentially detectable. If this is true, then the magnitude of morphological difference between temporally non-overlapping fossil assemblages may not by itself be useful for species delimitation (although, as noted in the main text, the patterning of primitive and derived morphological differences may be informative). Traditionally, palaeoanthropologists have placed great importance on the amount of variation expected within a species, but we suggest that this taxonomic criterion cannot easily be applied to temporally patterned morphological variation.

There is no doubt that a consideration of lineage segments in species delimitation is more complicated than defining species as minimally diagnosable units. Yet, biological processes are messy, and speciation is no exception. It is illogical to define species using an operationally simple standard when that standard lacks a fully valid ontological basis.

Thus, if we wish to understand the patterns and processes of human evolution, we must grapple with the fact that there are limitations to our knowledge of species diversity in the past and that some of our inferences will be better supported than others. Morphological diversity is not necessarily equivalent to species diversity. Critically, we must consider hominin species as hypotheses to be tested rather than groups to be identified.

Possible evidence of peramorphosis. Specimen SK 52 is a large subadult preserving two features that more closely resemble adult robust australopiths from DMQ than adults from SWM1HR, the unit from which it was recovered. SK 52 lacks a squared supraorbital corner and instead exhibits a sloped or rounded contour (Supplementary Figure 8a) that resembles that seen in DNH 7, DNH 152 and DNH 155 rather than that in SK 46 and SK 48 (Extended Data Figure 5). SK 52 also exhibits an anteroposteriorly thin zygomatic root with very little inflation of the alveolus above the M^2 and M^3. In this regard, SK 52 also resembles DNH 7 and DNH 155 more than it does contemporaneous adults SK 12 and SK 48 (Extended Data Fig. 3). If SK 52 is representative of subadult *P. robustus* from SWM1HR, then one might infer that the ontogenetic trajectory of *P. robustus* from SWM1HR incorporates the adult form of its putative ancestors, which would be an example of peramorphosis[111,112]. The processes producing peramorphosis, such as acceleration or hypermorphosis, might also explain the larger facial size in robust australopiths from SWM1HR relative to those from DMQ. Ultimately, more fossil evidence is needed to evaluate this possibility.

SUPPLEMENTARY DATA

Data used in the principal component and cluster analyses (Extended Data Fig. 6) are provided in Supplementary Data Martin et al.xlsx. Data from fossil specimens other than DNH 7 and DNH 155 were taken from Wood[59].

REFERENCES

83. Smith, R. J. Mandibular biomechanics and temporomandibular joint function in primates. *Am. J. Phys. Anthropol.* **49,** 341-350 (1978).
84. Demes, B. & Creel, N. Bite force, diet, and cranial morphology of fossil hominids. *J. Hum Evol.* **17,** 657-670 (1988).
85. Eng, C. M., Lieberman, D. E., Zink, K. D. & Peters, M. A. Bite force and occlusal stress production in hominin evolution. *Am. J. Phys. Anthropol.* **151,** 544-557 (2013).
86. Ledogar, J. A. et al. Mechanical evidence that *Australopithecus sediba* was limited in its ability to eat hard foods. *Nature Comm.* **7,** 10596.
87. Harrison, T. In *Paleontology and Geology of Laetoli: Human Evolution in Context. Vol. 2.* (ed. Harrison, T.) 141-188 (Springer, Dordrecht, 2011).
88. Suwa, G. In *Evolutionary History of the "Robust" australopithecines.* (ed. Grine, F. E.) 199-222 (Aldine de Gruyter, New York, 1988).
89. Strait, D. S., Grine, F. E. & Moniz, M. A. A reappraisal of early hominid phylogeny. *J. Hum. Evol.* **32,** 17-82 (1997).

90. Kullmer, O. et al. The first *Paranthropus* from the Malawi Rift. *J. Hum. Evol.* **37,** 121-127 (1999).

91. Suwa, G. et al. The first skull of *Australopithecus boisei.* *Nature* **389,** 489-492 (1997).

92. Adams, J. W., Herries, A. I. R., Conroy, G. C., Kuykendal, K. L. Taphonomy of a South African cave: geological and hydrological influences on the GD 1 fossil assemblage at Gondolin, a Plio-Pleistocene paleocave system in the Northwest Province, South Africa. *Quart. Sci. Rev.* **26,** 2526-2543 (2007).

93. Gibbon, R. J. et al. Cosmogenic nuclide burial dating of hominin-bearing Pleistocene cave deposits at Swartkrans, South Africa. *Quart. Geochronol.* **24,** 10-15 (2014).

94. Strait, P. T. *A First Course in Probability and Statistics with Applications.* (Harcourt Brace Jovanovich, San Diego, 1983).

95. Moggi-Cecchi, J. The life histories of fossil hominins. *J. Anthropol. Sci. – Riv. Antropol. JASS* **87**, 3-5 (2009).

96. Mayr, E. *Systematics and the Origin of Species.* (Columbia University Press, New York, 1942).

97. Patterson, H. E. H. In *Species and Speciation* (ed. Vrba, E. S.) 21-29 (Transvaal Museum, Pretoria, 1985).

98. Cracraft, J. Species concepts and speciation analysis. *Curr. Ornithol.* **1**, 159-187 (1983).

99. Templeton, A. R. In *Speciation and its Consequences.* (eds. Otte, D. & Endler, J. A.) 3-27 (Sinauer, Sunderland MA, 1989).

100. Kimbel, W. L. & Rak, Y. In *Species, Species Concepts, and Primate Evolution* (eds. Kimbel, W. L. & Martin, L. B.) 461-484 (Plenum Press, New York, 1993).

101. Frost, D. R. & Kluge, A. G. A consideration of epistemology in systematic biology, with special references to species. *Cladistics* **10,** 259-294 (1994).

102. Leakey, M. G. et al. New hominin genus from eastern Africa shows diverse middle Pliocene lineages. *Nature* **410**, 433-440 (2001).

103. Gilbert, W. H. In Homo erectus *Pleistocene Evidence from the Middle Awash, Ethiopia* (eds. Gilbert, W. H. & Asfaw, B.) 349-372 (University of California Press, Berkeley, 2008).

104. Hublin, J. J. The origin of Neanderthals. *Proc. Natl. Acad. Sci. USA* **106,** 16022-16027 (2009).

105. Fu, Q. et al. An early modern human from Romania with a recent Neanderthal ancestor. *Nature* **524,** 216-219 (2015).

106. Smith, A. B. *Systematics and the Fossil Record* (Blackwell, Oxford, 1994).

107. Kimbel, W. H. & Rak, Y. *Australopithecus sediba* and the emergence of *Homo:* Questionable evidence from the cranium of the juvenile holotype MH 1. *J. Hum. Evol.* **107,** 94-106 (2017).

108. Tattersall, I. Species recognition in human paleontology. *J. Hum. Evol.* **15**, 165-175 (1986).

109. Simpson, G. G. The Fort Union of the Crazy Mountain field, Montana and its mammalian faunas. *USNM Bull.* **169**, 1-287 (1937).

110. Rosen, D. E. Fishes from the upperlands and intermontane basins of Guatemala: revisionary studies and comparative geography. *Bull. Am. Mus. Nat. Hist.* **162**, 267-376 (1979).

111. Gould, S. J. *Ontogeny and Phylogeny.* (Harvard University Press, Cambridge, 1977).

112. Godfrey, L. R., Sutherland, M. R. Paradox of peramorphic paedomorphosis: heterochrony and human evolution. *Am. J. Phys. Anthropol.* **99,** 17-42 (1996).

113. Dean, M. C. & Wood, B. A. Metrical analysis of the basicranium of extinct hominoids and *Australopithecus.* *Am. J. Phys. Anthropol.* **54,** 63-71 (1981).

114. De Ruiter, D. J., Steininger, C. M. & Berger, L. R. A cranial base of *Australopithecus robustus* from the Hanging Remnant of Swartkrans, South Africa. *Am. J. Phys. Anthropol.* **130,** 435-444 (2006).

115. Smith, A. L. et al. Biomechanical implications of intraspecific shape variation in chimpanzees crania: moving towards and integration of geometric morphometrics and finite element analysis. *Anat. Rec.* **298,** 122-144 (2015).

Supplementary Table 1. Morphological differences among *Paranthropus* samples.

Character	*P. robustus* Drimolen Main Quarry	*P. robustus* Kromdraai B	*P. robustus* Swartkrans Member 1 Hanging Remnant	*P. aethiopicus*	*P. boisei*
Placement of anterior aspect of sagittal crest in presumed males	Approximating or posterior to bregma	?	Anterior to bregma	Approximating or posterior to bregma	Approximating or posterior to bregma
Posterior extent of sagittal crest	Bare area	?	Bare area	Extensive posteriorly	Bare area
Endocranial capacity	Lower end of *Paranthropus* range	?	Upper end of *Paranthropus* range	Lower end of *Paranthropus* range	Upper end of *Paranthropus* range
Position of zygomatic root relative to the toothrow	More posterior	More anterior	More anterior	Most anterior	Most anterior
Index of palate protrusion anterior to sellion	More protruding	Less protruding	Less protruding	Strongly protruding	Less protruding
Maximum anteroposterior thickness on maxilla of zygomatic root	Thin	Thick	Thick	Thick	Thick
Petrous orientation	More sagittal	?	More coronal	More coronal	More coronal
Zygomaticomaxillary step	Absent	Present	Present	Present	Absent

Zygomaticomaxillary fossa	Present	Present	Present	Absent	Absent
Maxillary fossula	Present	Present	Present	Absent	Absent
Supraorbital "corner"	More rounded	?	Squared	?	More rounded
Shape of the superior orbital fissure	Foramen	?	?	Fissure	Fissure
Size of *longus capitis* insertion	Large	?	Small	?	Small
Size of maxillary postcanine teeth	Smaller	?	Larger	Largest	Largest

The dental sample from Kromdraai B is too small to allow a statistical assessment of maxillary tooth size. One juvenile specimen from Swartkrans (SK 52) exhibits a thin zygomatic root and a more rounded supraorbital corner. In these respects, it is similar to adults from DMQ but unlike adults from Swartkrans, possibly providing insights into the ontogeny of these features in the Swartkrans population. One subadult specimen of *P. boisei* (KNM-WT 17400) evidently possessed a small endocranial capacity (although this measurement cannot be made with high confidence)[19].

Supplementary Table 2: Indices of the masticatory apparatus (%)

Species/specimen	Index of palate protrusion anterior to sellion	Index of palate protrusion anterior to masseter	Index of position of anterior part of masseter	Index of M^3 position	Overlapping index	Masseter leverage at M^3 bite
Australopithecus afarensis						
AL 444-2	55	59	97	57	26	85
AL 417-1d	59	48				
Australopithecus africanus						
Sts 5	68	65	103	66	22	78
Sts 71	43	47	93	35	40	133
Paranthropus aethiopicus						
KNM-WT 17000	82	53	148	77	28	96
Paranthropus robustus						
DNH 155 (Drimolen)	60	35	122	60	41	102
TM 1517 (Kromdraai)		42			37	
SK 48 (Swartkrans)	46	31	119	40	50	149
SK 52 (Swartkrans)		34			45	
Paranthropus boisei						
OH 5	43	32	113	38	57	149
KNM-ER 406	53	26	130	49	50	133
KNM-ER 732	41	41	100	36	43	139

Indices defined as in Rak[12]. All data are from Kimbel et al.[15], except those for DNH 155. All indices are ratios of distances parallel to Frankfort Horizontal projected into a sagittal plane. Index of palate protrusion anterior to sellion = (sellion to prosthion)/(Prosthion to M^3). Index of palate protruision anterior to masseter = (zygomatic tubercle to prosthion)/(Prosthion to M^3). Index of position of anterior part of masseter = (articular eminence to zygomatic tubercle)/(articular eminence to sellion). Index of M^3 position = (articular eminence to M^3)/(articular eminence to sellion). Overlapping index = (M^3 to zygomatic tubercle)/(articular eminence to prosthion). Note that none of these measures directly measure biomechanical variables related to bite force production (e.g., bite force load arm, muscle force lever arm), but rather describe the relative positioning of different parts of the masticatory system. However, by dividing the Index of position of the anterior part of masseter by the Index of M^3 position, one obtains a ratio (articular eminence to zygomatic tubercle)/(articular eminence to M^3) that, when divided by 2, is a fair approximation of the leverage of the masseter muscle during an M^3 bite.

Supplementary Table 3. Petrous orientation in australopith specimens.

Species/specimen	Petrous angle
Au. afarensis:	
A.L. 444-2	59
A.L. 333-45	53
Au. africanus:	
MLD 37/38	60
Sts 5	65
Sts 25	72
P. robustus:	
DNH 7	57
DNH 155	62
SK 47	45
SKW 18	42
TM 1517	42
P. boisei:	
OH 5	45
KNM-ER 406	44
KNM-ER 407	49
KNM-ER 23000	50
P. aethiopicus:	
KNM-WT 17000	50

Petrous orientation is measured as defined in Dean and Wood[113]. Data from specimens other than DNH 7 and DNH 155 are published elsewhere[15,114].

Supplementary Table 4: Summary statistics for adult maxillary dental metrics in the DMQ and SWM1HR samples.

Site	Element	Measurement	n	Mean (mm)	Std Dev
DMQ	I^1	MD	4	8.15	0.67
		BL	3	6.98	0.66
	I^2	MD	4	5.67	0.54
		BL	4	6.20	0.82
	C	MD	6	8.42	0.94
		BL	8	8.97	0.82
	P^3	MD	5	9.14	0.32
		BL	5	12.86	0.4
	P^4	MD	5	9.74	0.39
		BL	5	13.84	0.31
	M^1	MD	7	12.44	0.97
		BL	7	14.32	0.63
	M^2	MD	7	13.06	1.09
		BL	7	14.92	0.92
	M^3	MD	7	13.49	0.95
		BL	7	15.27	0.88
SWM1HR	I^1	MD	19	8.94	0.75
		BL	18	7.33	0.51
	I^2	MD	12	6.37	0.68
		BL	12	6.54	0.67
	C	MD	23	8.33	0.64
		BL	23	9.13	0.81
	P^3	MD	22	9.84	0.54
		BL	18	14.17	0.75
	P^4	MD	26	10.63	0.61
		BL	24	15.16	0.88
	M^1	MD	22	13.48	0.71
		BL	21	14.79	0.61
	M^2	MD	21	14.11	0.90
		BL	21	15.96	0.85
	M^3	MD	23	14.90	1.08
		BL	23	16.79	0.73

Supplementary Table 5. Comparisons between the adult, maxillary dental samples from DMQ and DWM1HR using two-tailed Mann-Whitney U tests.

Element	Measurement	U	*p*
I^1	MD	15	0.061
	BL	19	0.420
I^2	MD	8.5	0.059
	BL	19.5	0.584
C	MD	68	0.957
	BL	80	0.587
P^3	MD	12.5	0.008
	BL	8.5	0.006
P^4	MD	15	0.007
	BL	10.5	0.004
M^1	MD	29.5	0.015
	BL	47.5	0.167
M^2	MD	33.5	0.034
	BL	27	0.013
M^3	MD	24.5	0.006
	BL	13.5	0.001

Supplementary Table 6: Linear dimensions (in mm) of DNH 155 compared to other robust australopith crania.

Character	DNH 155 *P. robustus*	DNH 7 *P. robustus*	SK 48 *P. robustus*	ER 406 *P. boisei*	ER 732 *P. boisei*	OH 5 *P. boisei*	WT 17000 *P. aethiopicus*
8. Minimum frontal breadth	64	(63)	70	61	60	69	65
9. Maximum parietal breadth	91			104	90	134	
11. Biporionic breadth	(111)			126	110	134	
21. Bregma – left pterion	57			62			
23. Bregma – right pterion	56			62			
25. Parietal sagittal length	102			80		75	
43. Superior facial height	88	68	80	88		112	99
45. Alveolar height	30	26	30	35		42	33
49. Superior facial breadth	98	(88)	94	115	90	115	
50. Biorbital breadth	88	(78)	94	100	82	96	
51. Bijugal breadth	116	(101)	121	135	107	130	
52. Bizygomatic breadth	(135)	(125)	145	179	143	168	
53. Bimaxillary breadth	93	(96)	105	134	100	122	
54. Outer alveolar breadth	65	62	66	77	62	81	
55. Anterior interorbital breadth	20	(18)	28	27	23	23	23
56. Orbital breadth (Left)	33		33	37		41	36
56. Orbital breadth (Right)	33	29	33	39	32	39	
57. Orbital height (Left)	36		30	36		33	41
57. Orbital height (Right)		33	30	36	30	34	
58. Orbitale-zygomaxillare (Left)	30		30			50	
58. Orbitale-zygomaxillare (Right)	31	31	30	36	29	46	
59. Minimum malar height (Left)	30		29			48	

59. Minimum malar height (Right)	30	29	29	34	29	43
60. Malar thickness (Left)	16		13			
60. Malar thickness (Right)	14	14	13	17	13	17
61. Width temporal gutter (Left)	25	15	29	30		34
61. Width temporal gutter (Right)	25	15	29	30	25	28
66. Frontal torus breadth	82	(62)	105	111	86	112
68. Maximum nasal width	27	21	29	28	28	32
69. Nasal height	64	50	54	58		70
70. Rhinion - nasospinale	32	27	23	36		34
72. Sagittal length of nasal bones	31		31	24		36
73. Superior breadth of nasal bones	(6)		11	16	12	13
74. Inferior breadth of nasal bones		9		6	7	7
87. Maxillo-alveolar length	69	60	69	80		86
88. Maxillo-alveolar breadth	65	62	62	79		82
90. Palate length	45	(40)	49	60		71
91. Palate breadth	26	21	27	32		38
92. Incisive canal – palatomaxillary suture	29	26		57		61
93. Internal alveolar breadth M^3	24	22	27	31		34
94. $I^1 – I^2$ alveolar length	14	12	12			15
95. Canine alveolus breadth	9	6.8	8.3			7.6
96. $P^3 – P^4$ interalveolar length	20	17	15	23		21

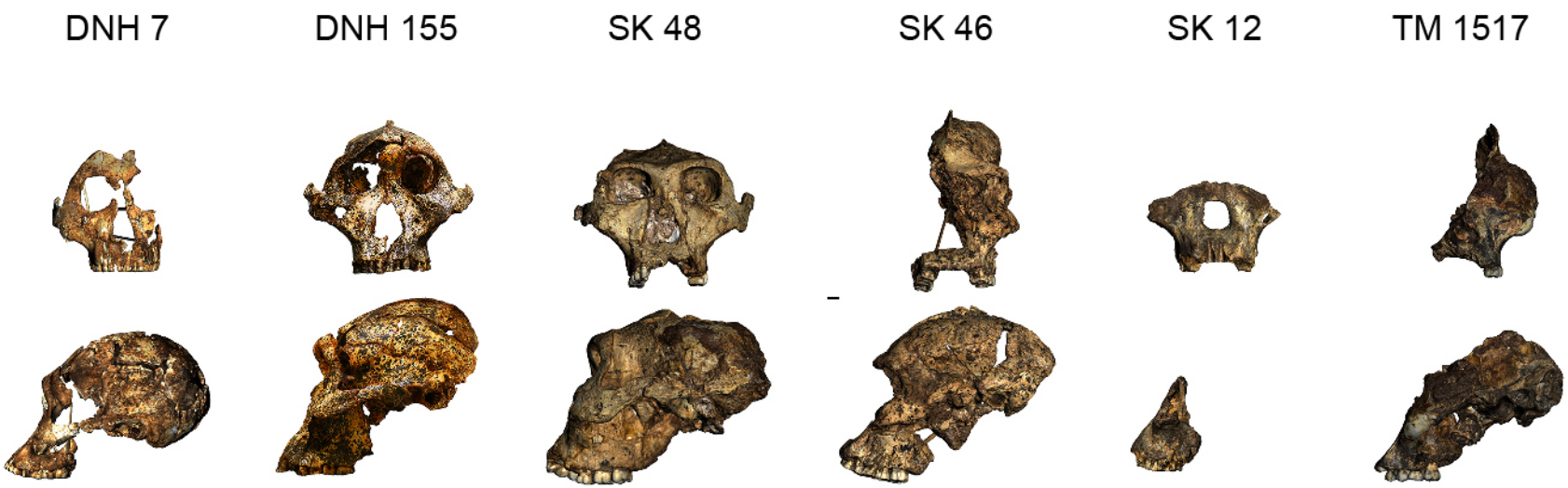
97. $M^1 – M^3$ interalveolar length	36	34	39	46	45
98. Intercanine distance	25	23	28		31
99. P^3 interalveolar distance	24	24	27	31	30
100. P^4 interalveolar distance	27	27	30	31	33
101. M^2 interalveolar distance	34	33	37	34	40
102. M^3 interalveolar distance	37	36		38	42
103. Palatal height	9	8	14	19	20

Measurements numbered and defined as in Wood[59]. Data for specimens other than DNH 7 and DNH 155 are from Wood[59]. Values in parentheses are estimated.

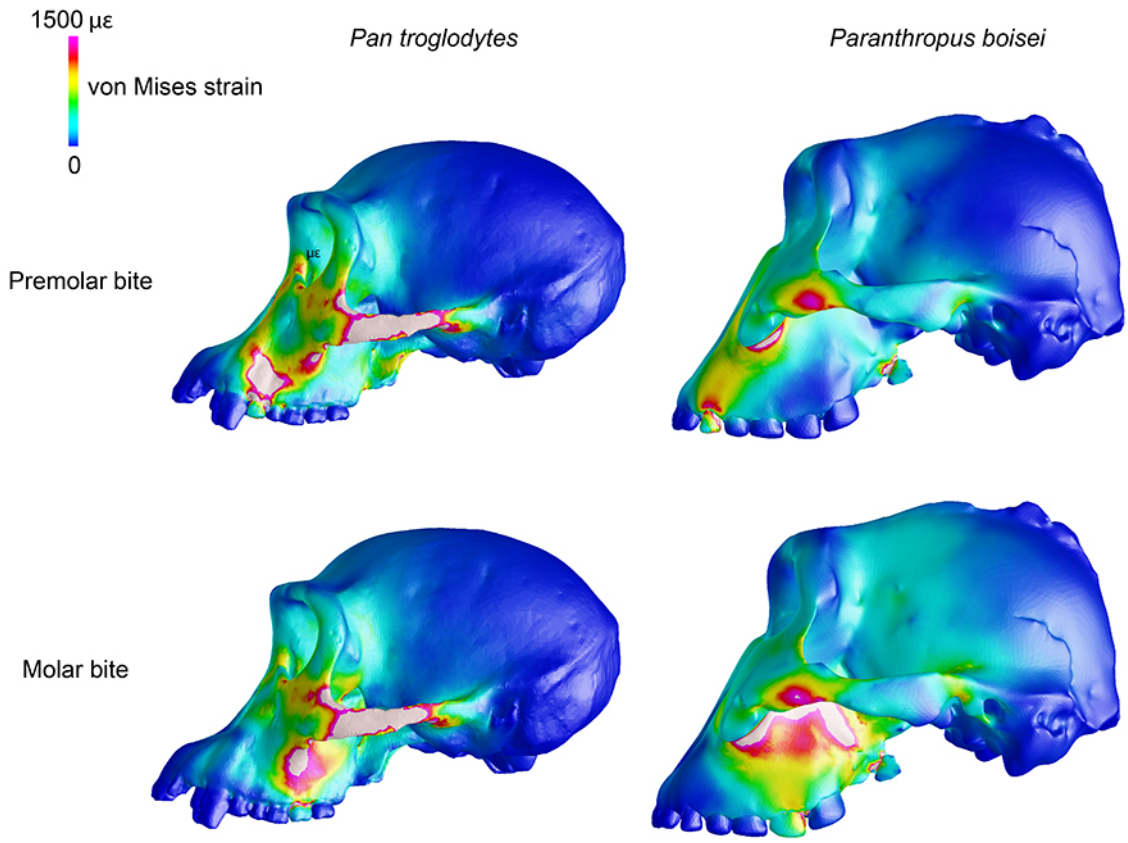
Supplementary Table 7. Postorbital constriction index in australopith and pre-australopith specimens.

Species/specimen	Postorbital constriction index
S. tchadensis	
TM-266-01-060-1	59.1
Ar. ramidus	
ARA-VP-6/500	60.2
Au. anamensis	
MRD-VP-1/1	62.8
Au. afarensis:	
A.L. 444-2	66.4
Au. africanus:	
Sts 5	69.1
Sts 71	71.7
P. robustus:	
DNH 7	71.6
DNH 155	65.3
SK 48	66.4
SK 52	67.3
P. boisei:	
OH 5	60.0
KNM-ER 406	53.0
KNM-ER 732	66.7

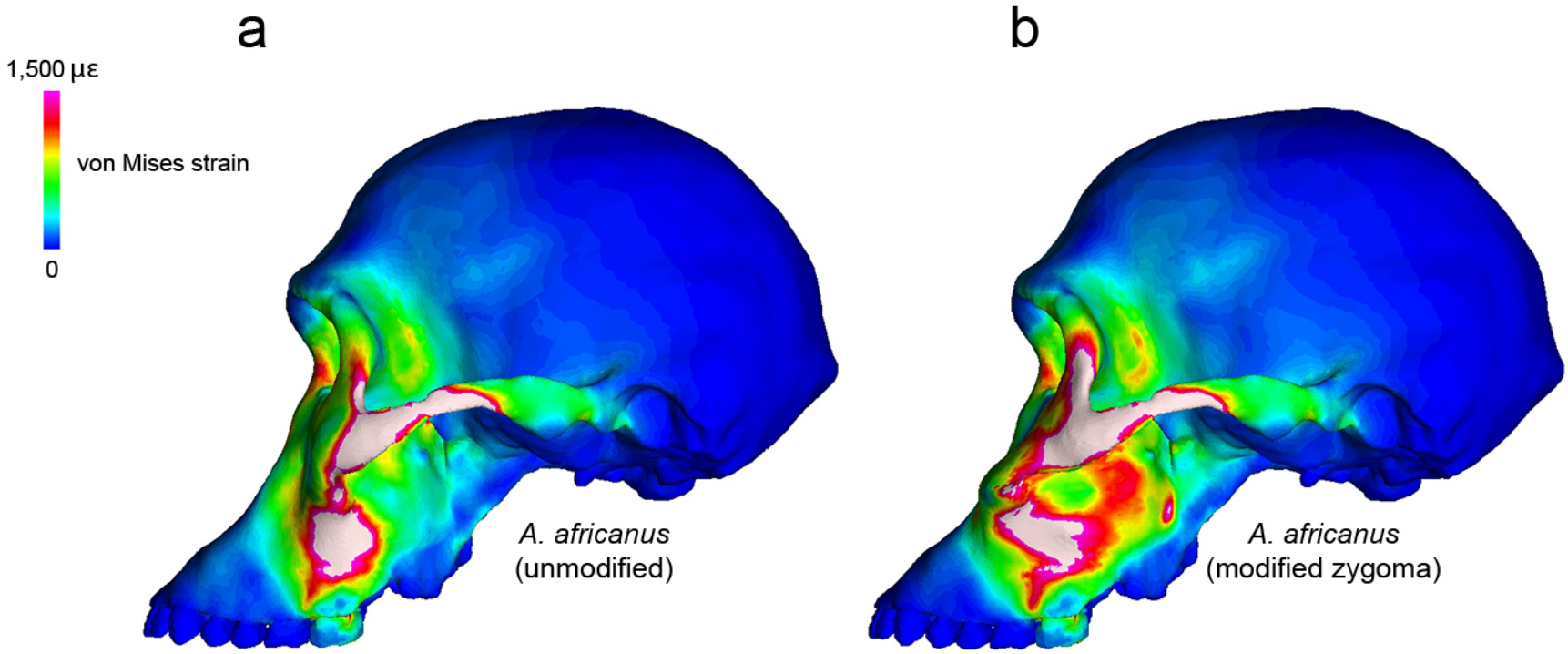
The index is measured as the ratio of minimum frontal breadth to superior facial breadth, multiplied by 100. Data from specimens other than DNH 7 and DNH 155 are published elsewhere[59,71].



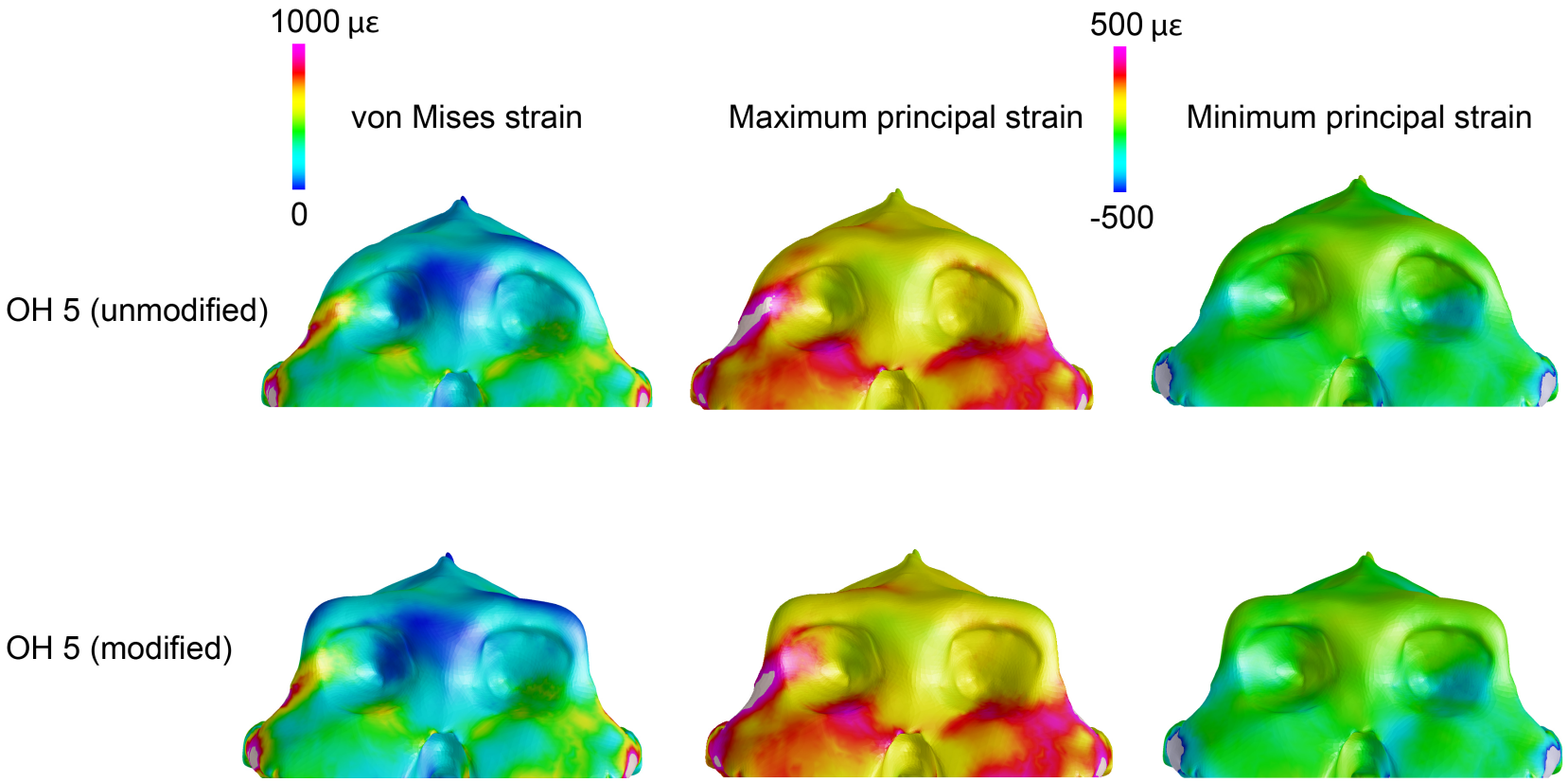
Supplementary Figure 1. Comparison among crania of *P. robustus*. Crania aligned in frontal (top row) and lateral (bottom row) view. Note that the face of DNH 7 is detached from the specimen's neurocranium and neurocranial distortion makes it difficult to join the two pieces. The placement of the face on the neurocranium in lateral view is heuristic and no morphological assessments reported here are dependent on these cranial parts being properly aligned. Moreover, the lateral views of DNH 7 and SK 12 are reflected images of the specimens' right sides.



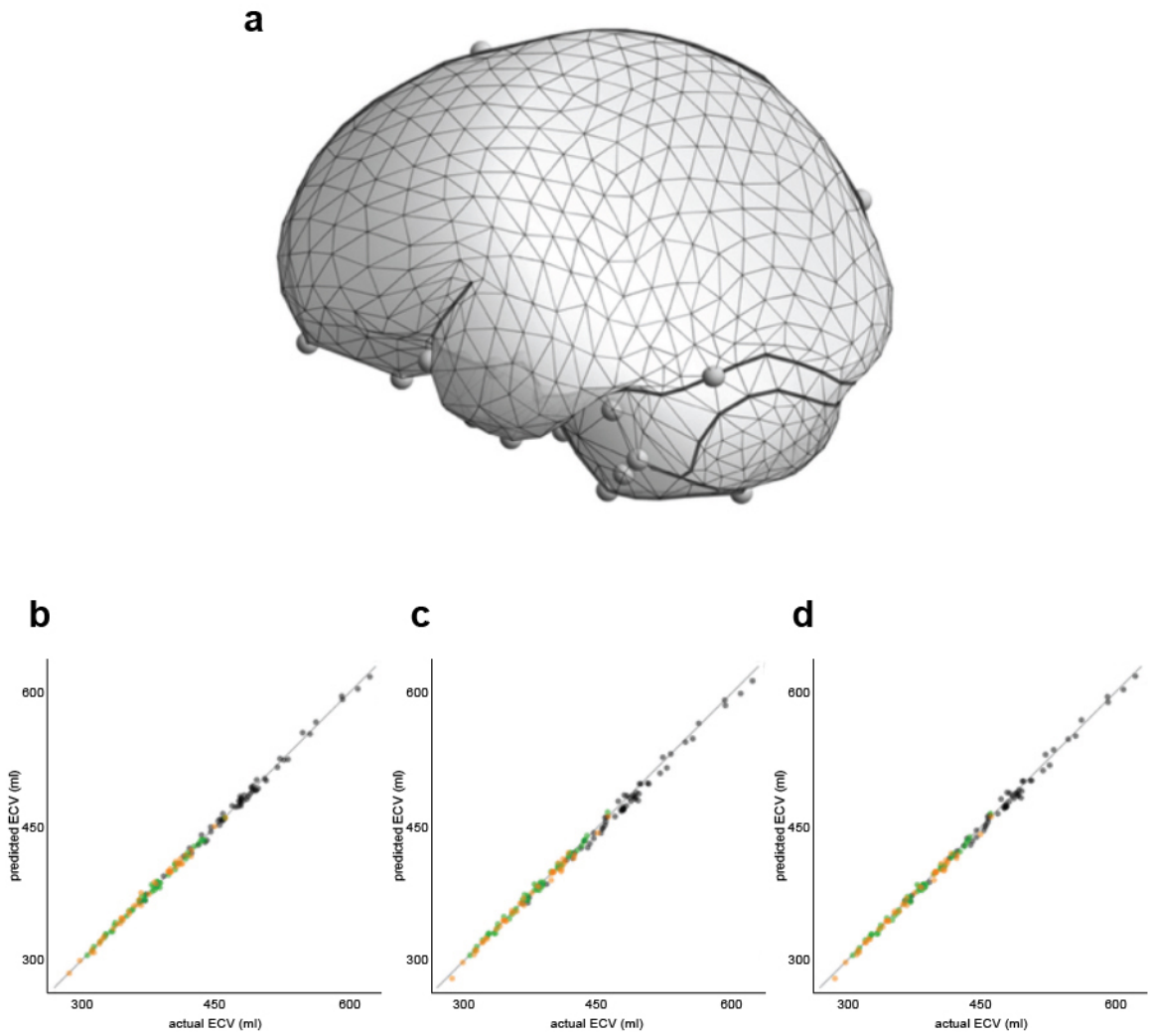
Supplementary Figure 2. Zygomatic root position and its effect on facial strain during premolar and molar bites. Lateral views of finite element models of a chimpanzee (*Pan troglodytes*) and robust australopith specimen OH 5 (*Paranthropus boisei*) simulating maximal bites on the P^3 and M^2. Magnitudes of von Mises strain indicated in color; white regions are those in which strain magnitudes exceed the maximum value of the scale. Full details on model construction and boundary conditions are provided in Smith et al.[13,115]. The root is positioned approximately above the bite point during molar biting in the chimpanzee and premolar biting in OH 5, and in those simulations strain is concentrated on the inferior surface of the root and the alveolus below the root. However, when the bite point is positioned far behind the zygomatic root, as during molar biting in OH 5, then the lateral surface of the maxilla posterior to the root exhibits high strains. Thus, the evolution of an anteriorly positioned root has the perhaps unexpected consequence of weakening the posterior maxilla, at least during molar bites.



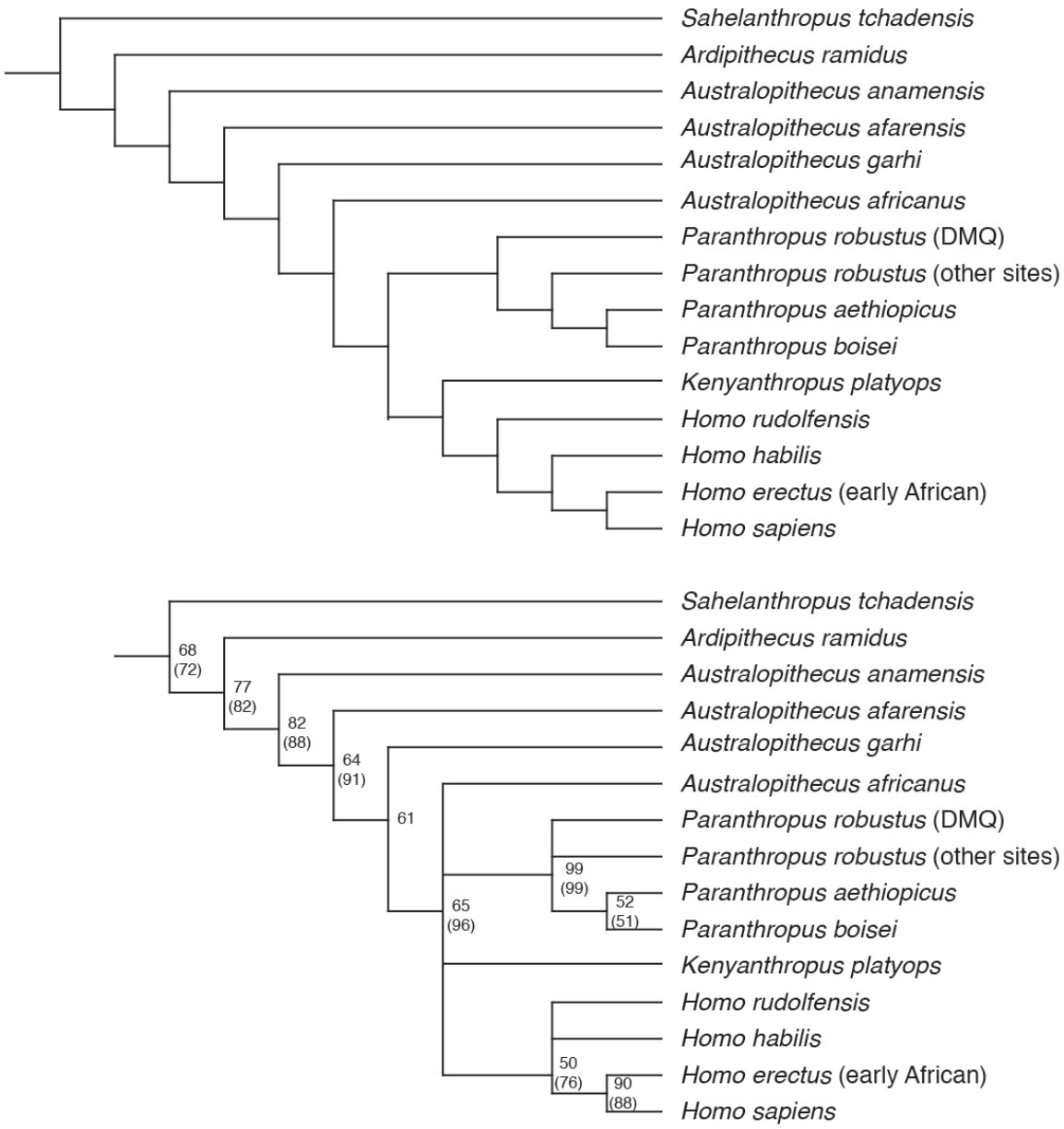
Supplementary Figure 3. Modeling experiment illustrating the biomechanics of zygomatic root position. Lateral views of finite element models of *Australopithecus africanus* composit model simulating maximal bites on the M^2. The composite model is based largely on specimen Sts 5. **a,** Unmodified model with posteriorly positioned zygomatic root. **b,** Model modified using geometric morphometric methods to have an anteriorly positioned zygomatic root. Magnitudes of von Mises strain indicated in color; white regions are those in which strain magnitudes exceed the maximum value of the scale. Full details on model construction and boundary conditions are provided in Ledogar et al.[14]. During molar biting in the unmodified model (**a**), strain is concentrated on the inferior surface of the zygomatic root and the alveolus below the root. However, when the zygomatic root is positioned anteriorly, molar biting produced high strains in the maxilla posterior to the root. This further illustrates that the evolution of an anteriorly positioned zygomatic root has the effect of weakening the posterior maxilla during molar bites.



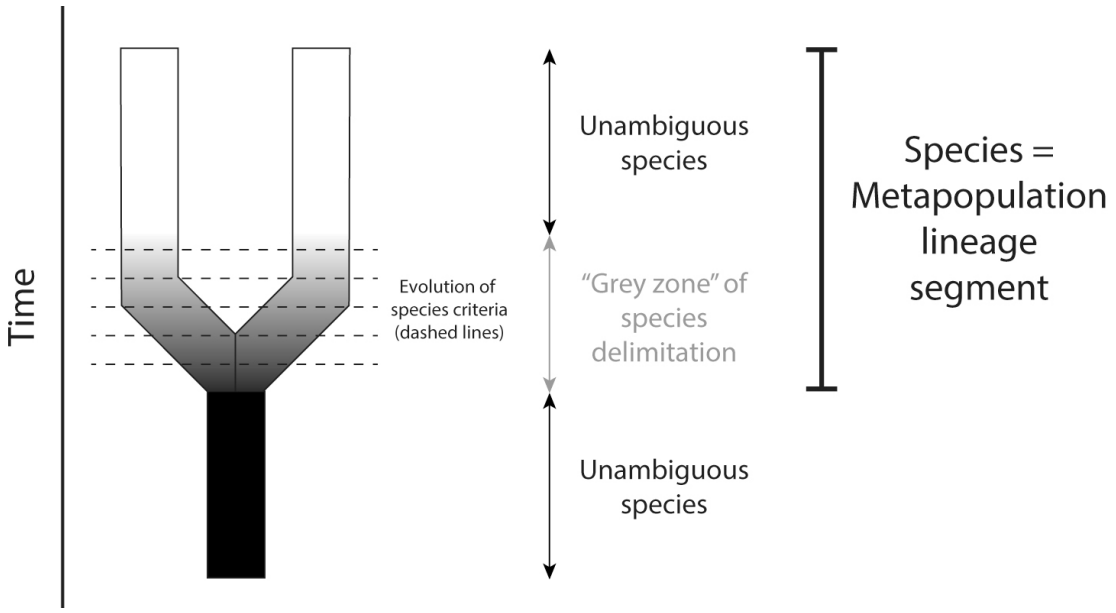
Supplementary Figure 4. Biomechanics of the supraorbital corner. Strain magnitudes in the circumorbital region during simulations of maximal molar biting in a finite element models of *Parathropus boisei* specimen OH 5, and a modified OH 5 model in which the supraorbital corner has been "squared" as in *P. robustus* and African apes. Magnitudes of von Mises, maximum principal, and minimum principal strain are indicated in color; white regions are those in which strain magnitudes exceed the maximum value of the scale. Note that von Mises and the principal strains are calibrated to different scales. Full details on model construction and boundary conditions of the unmodified OH 5 model are provided in Smith et al.[13]. The supraorbital torus and corner experience low strains during even very forceful bites. Alterations in the shape of the supraorbital corner have only very subtle impacts on strain patterns, so the morphology of this region is unlikely to be related to feeding biomechanics.



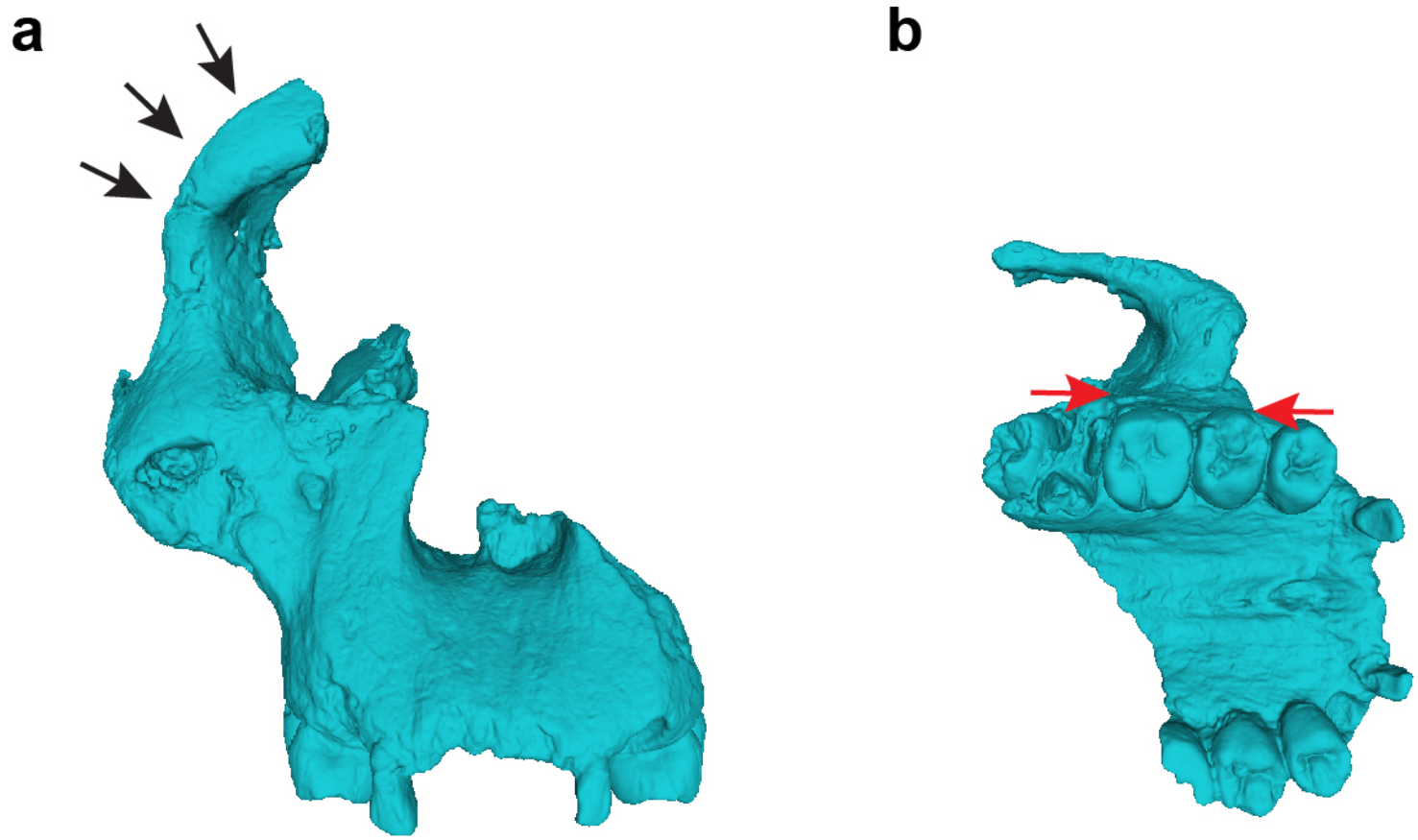
Supplementary Figure 5. Endocranial volume estimation of DNH 155. a, Endocranial landmark set used for ECV estimation. Each vertex of the surface is used as a landmark or semilandmark. Anatomical landmarks are shown as spheres, curve semilandmarks are connected as black lines. **b – d,** Simulations based on reference sample: measured versus predicted ECV [**b**, regression-based, **c**, pooled TPS-based and **d**, species-specific TPS-based estimates] for gorillas (grey), orang-utans (orange), chimpanzees (green).



Supplementary Figure 6. Results of maximum parsimony analysis. a, Most parsimonious cladogram. **b,** 50% majority rule consensus tree derived from bootstrap analysis. Numbers at internal nodes indicate bootstrap support for the corresponding clade. Numbers in parantheses indicate bootstrap support for the clade when *K. platyops* and *Au. garhi* (both of which are missing data with respect to many characters) are excluded from the analysis.



Supplementary Figure 7. The Unified Species Concept. A simple depiction styled after [39] of speciation in which an ancestral species gives rise to two descendant species. Note that other manifestations of speciation are possible (i.e., as when an ancestral species persists to be contemporaneous with a single descendant species). According to the Unified Species Concept[38,39], a species is a lineage segment extending from its initial divergence to eventual extinction. Prior to the divergence, and at some unknown time after divergence, species should be unambiguously recognized. However, when lineages are initially diverging from each other, they evolve species criteria (dashed lines) at unspecified times and in an unspecified order. Such criteria include ecological differentiation, reproductive isolation, the evolution of specific mate recognition systems, phenotypic diagnosability, and so on. During this "grey zone" of speciation, criteria-based species concepts may identify species in different ways, leading to ambiguity in species delimitation. Yet, ontologically, species are lineage segments, so these species criteria should not be used to define species boundaries, but rather to test the null hypothesis that a single lineage segment is present. However, because morphological diagnosability may evolve anywhere in the grey zone, it is expected that early and late portions of a single lineage segment may differ morphologically from each other. Thus, diagnosable differences between temporally non-overlapping fossil assemblages are not necessarily sufficient to falsify a single species null hypothesis.



Supplementary Figure 8. DMQ-like traits in a subadult from SWM1HR. Specimen SK 52 in **a,** frontal and **b,** palatal view. This specimen resembles DNH 7 and DNH 155 in having a rounded rather than squared supraorbital corner (black arrows) and a narrow zygomatic root that is not expanded posteriorly (red arrows).