PROCEEDINGS B

rspb.royalsocietypublishing.org

Research

Cite this article: Du A, Zipkin AM, Hatala KG, Renner E, Baker JL, Bianchi S, Bernal KH, Wood BA. 2018 Pattern and process in hominin brain size evolution are scale-dependent. *Proc. R. Soc. B* **285**: 20172738.
http://dx.doi.org/10.1098/rspb.2017.2738

Received: 7 December 2017
Accepted: 31 January 2018

Subject Category:
Palaeobiology

Subject Areas:
evolution, palaeontology

Keywords:
hominin evolution, endocranial volume, phenotypic evolution, evolutionary mode, microevolution, macroevolution

Author for correspondence:
Andrew Du
e-mail: andrewdu@uchicago.edu

†Present address: Department of Organismal Biology and Anatomy, University of Chicago, 1027 E 57th Street, Chicago, IL 60637, USA.

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4001082.

THE ROYAL SOCIETY PUBLISHING

Pattern and process in hominin brain size evolution are scale-dependent

Andrew Du[1,†], Andrew M. Zipkin[1,2], Kevin G. Hatala[1,3], Elizabeth Renner[1,4], Jennifer L. Baker[1,5], Serena Bianchi[1], Kallista H. Bernal[1] and Bernard A. Wood[1]

[1]Center for the Advanced Study of Human Paleobiology, Department of Anthropology, The George Washington University, 800 22nd Street, NW, Washington, DC 20052, USA
[2]Department of Anthropology, University of Illinois at Urbana-Champaign, 109 Davenport Hall, 607 S. Mathews Avenue, Urbana, IL 61801, USA
[3]Department of Biology, Chatham University, Woodland Road, Pittsburgh, PA 15232, USA
[4]Psychology, University of Stirling, Stirling, Scotland FK9 4LA, UK
[5]Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institute of Health, 12 South Drive, MSC 5635, Bethesda, MD 20892, USA

AD, 0000-0002-7381-5089; AMZ, 0000-0002-3813-8051; KGH, 0000-0001-9131-5304; ER, 0000-0002-3363-4347; JLB, 0000-0002-5192-2377; BAW, 0000-0003-0273-7332

A large brain is a defining feature of modern humans, yet there is no consensus regarding the patterns, rates and processes involved in hominin brain size evolution. We use a reliable proxy for brain size in fossils, endocranial volume (ECV), to better understand how brain size evolved at both clade- and lineage-level scales. For the hominin clade overall, the dominant signal is consistent with a gradual increase in brain size. This gradual trend appears to have been generated primarily by processes operating within hypothesized lineages—64% or 88% depending on whether one uses a more or less speciose taxonomy, respectively. These processes were supplemented by the appearance in the fossil record of larger-brained *Homo* species and the subsequent disappearance of smaller-brained *Australopithecus* and *Paranthropus* taxa. When the estimated rate of within-lineage ECV increase is compared to an exponential model that operationalizes generation-scale evolutionary processes, it suggests that the observed data were the result of episodes of directional selection interspersed with periods of stasis and/or drift; all of this occurs on too fine a timescale to be resolved by the current human fossil record, thus producing apparent gradual trends within lineages. Our findings provide a quantitative basis for developing and testing scale-explicit hypotheses about the factors that led brain size to increase during hominin evolution.

1. Introduction

The large brain of *Homo sapiens* is a defining hallmark of humankind. The brains of ancient hominins do not fossilize, but the estimated volume of a fossil cranium (i.e. endocranial volume or ECV) is a reliable proxy for brain size [1–3]. From studying fossil ECV trends through time, we know that brain size in the hominin clade increased more than threefold from *Australopithecus*—the earliest unambiguous hominin taxon—to our own species. Nevertheless, researchers currently disagree about the manner by which hominin brain size increased. They claim ECV increased either gradually through time [4–7] or via brief episodes of rapid increase separated by extended periods of stasis (i.e. the punctuated equilibrium model) [8–11]. Based on these inferred patterns, researchers then attempt to draw conclusions about what drove hominin ECV increase (e.g. episodes of climate change drove the rapid ECV increase in the punctuated equilibrium model [12,13]).

Caution must be exercised, however, when developing and evaluating these cause-and-effect hypotheses. For one, such hypotheses need to specify the taxonomic level at which the effect is observed because evolutionary patterns are scale-dependent. Evolution is hierarchical, and trends at lower taxonomic levels can look very different when combined at higher taxonomic levels (i.e. Simpson's

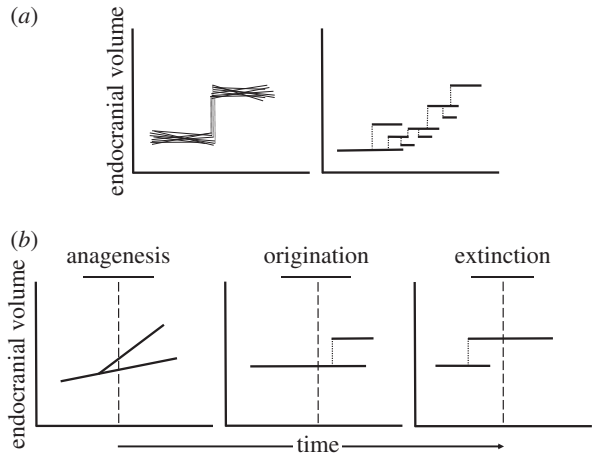

Figure 1. Theoretical plots demonstrating (*a*) how patterns within lineages may not necessarily hold when combined and examined at the level of the entire clade and (*b*) how clade-level brain size can increase via anagenesis, origination or extinction. (*a*) On the left, lineages exhibit gradual trends, but the clade-level pattern shows a punctuated equilibrium pattern. This is because the direction and magnitude of change within each group of co-occurring lineages cancel each other out so that, on average, an emergent stasis pattern is generated. On the right, each lineage experiences stasis, but because origination events produce more and larger ECV increases than decreases on average, a gradual clade-level trend is produced. (*b*) The vertical dashed line represents a bin edge separating two bins. In all three cases, clade-level brain size is increasing, and the increase is attributed solely to the process being illustrated. In 'anagenesis', we see two lineages survive from the earlier time bin into the later time bin, and both lineages exhibit an increase in ECV over time. In 'origination', a lineage-splitting event creates a daughter species with a larger ECV in the later time bin, leading to an increase in overall clade-level ECV. In 'extinction', the smaller-brained lineage goes extinct in the earlier time bin, causing clade-level ECV to increase in the subsequent time bin.

paradox; figure 1*a*) [14–17]. Furthermore, different taxonomic levels imply different evolutionary mechanisms [14,16,17]. Patterns within lineages (i.e. hypothesized ancestor-descendant sequences) are determined by microevolutionary, population-level processes (i.e. anagenesis). The most popular hypotheses related to hominin ECV increase—which include links to the appearance and development of uniquely human traits such as socio-cultural complexity, symbolic behaviour and language [2,12,13,18]—are in this category. Patterns observed at the clade level, on the other hand, are the result of processes operating across multiple lineages and therefore can also be shaped by macroevolutionary processes (i.e. origination by lineage-splitting and extinction) in addition to microevolutionary ones. For example, from one time period to the next, the mean ECV of a clade can increase because (i) existing lineages evolve larger brains anagenetically, (ii) a new lineage originates with a brain size greater than the pre-existing clade mean or (iii) a lineage with a brain size smaller than the clade mean goes extinct (figure 1*b*). We must first understand how hominin brain size has increased across taxonomic scales in order to properly infer the evolutionary processes and potential drivers involved.

Existing research seldom specifies taxonomic scale when examining whether hominin ECV data more closely match a pattern of gradualism or one of punctuated equilibrium. Some studies have analysed changes in hominin brain size within a lineage [5,7,9–11], while others have done so at the level of the entire hominin clade [4,6,18–20]. It is perhaps not surprising then that both models (i.e. gradualism and punctuated equilibrium) have found support from various analyses with no consensus being reached. Other researchers have argued that the data better support a model that combines elements of both gradualism and punctuated equilibrium [18,21,22], or they claim that the two models cannot be distinguished [23,24]. The thesis of our investigation is that taxonomic scale must be made explicit when describing patterns of hominin ECV evolution and attempting to infer the processes that generated them.

Here, we address the issue of taxonomic scale dependence by examining patterns of hominin ECV change at each of three successively lower taxonomic levels. We first estimate the evolutionary mode of hominin ECV change at the clade level. We then partition clade-level ECV change into anagenetic (microevolution) and origination and extinction (macroevolution) components. Finally, because hypotheses for hominin ECV increase are presently dominated by microevolutionary explanations, we compare the rate of within-lineage ECV change to that expected from an extrapolated microevolutionary model to assess whether observed increases can be solely accounted for by population-level, directional selection. By understanding how hominin ECV patterns change across different taxonomic levels, we can gain a more nuanced understanding of hominin brain size evolution. This work specifically addresses when and at what scale ECV change most likely occurred, and lays a foundation for future research to orient hypothesized evolutionary mechanisms with scale-specific phenotypic patterns.

2. Material and methods

Our dataset is composed of published absolute ECVs for specimens whose geological age ranges encompass dates anywhere between 3.2 and 0.5 Ma (i.e. beginning with the earliest uncontroversial hominin taxon, *Australopithecus*, and ending when the ECVs of fossil crania begin to overlap with the range of modern humans) (figure 2 and electronic supplementary material, table S1). ECV measurements were $\log_{10}$-transformed prior to all analyses because our methods assume linear changes in ECV, and we were interested in proportional differences (e.g. a doubling from 400 to 800 cm^3 was considered more biologically meaningful than an equivalent 400 cm^3 increase from 1000 to 1400 cm^3) [26]. We analysed absolute and not relative brain size because the use of relative brain size, which requires estimation of fossil hominin body masses, introduces additional error into analyses [18].

Clade-level ECV increase was analysed using the 'paleoTS' R package [27] to fit six statistical models of evolutionary modes to the observed ECV data (i.e. random walk, gradualism, stasis, stasis-stasis (i.e. clade-level 'punctuated equilibrium'), stasis-random walk and stasis-gradualism) (electronic supplementary material, figure S1) [28–31]. Random walk and gradualism were modelled as having bin-to-bin ECV transitions drawn from a normal distribution, where the mean equals zero in the random walk model (i.e. increases and decreases are equiprobable) and a non-zero value in the gradualism model (i.e. biased towards increases or decreases) [28,30]. Stasis was modelled as normally distributed variation around an optimal ECV mean, both of which stay constant through time [28,30,32]. The remaining complex models were modelled as permutations of the three simpler models separated by an estimated age point where the transition occurs [31]. To operationalize these models, we grouped all ECV data into 0.2 Myr bins (to maximize the number of binned samples within the time series), fit the models to the data by maximizing

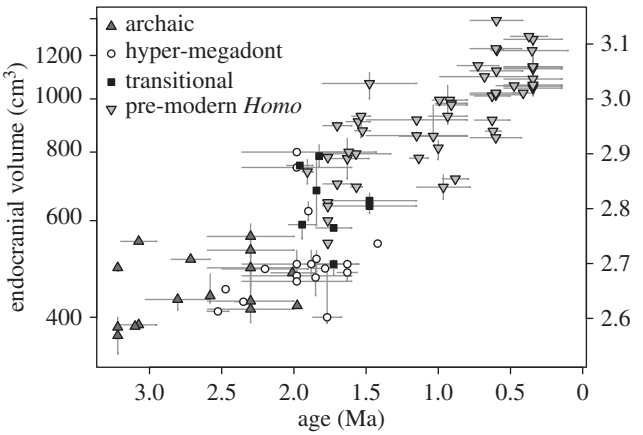

Figure 2. Time series of hominin ECV included in our analyses. Points represent ECV and age midpoints, and bars represent ranges of error on the estimate (or mean ± 3 s.d. for dates with normally distributed error (see electronic supplementary material, appendix S1)). Points are coded by hominin grade [25]. Specifically, archaic species include *Australopithecus afarensis*, *Australopithecus africanus* and *Australopithecus sediba*; hyper-megadont and megadont species include *Australopithecus garhi*, *Paranthropus aethiopicus*, *Paranthropus boisei* and *Paranthropus robustus*; transitional species include *Homo habilis sensu stricto* and *Homo rudolfensis*; and pre-modern *Homo* species include *Homo erectus sensu stricto*, *Homo ergaster*, *Homo georgicus* and *Homo heidelbergensis*. The left *y*-axis is on a logarithmic scale, while the right *y*-axis' tick labels are $\log_{10}$-transformed values.

likelihood, and assessed relative model fit using the bias-corrected Akaike information criterion (AICc) transformed into weights (which sum to one across all models, with higher weights representing more model support) [33]. We resampled age estimates to examine the possible influences of dating error on our model selection results (electronic supplementary material, figures S2 and S3), and we also explicitly incorporated inter-observer error in estimating ECV for individual fossils into our model (electronic supplementary material, figures S2 and S3 and table S2). We conducted sensitivity analyses to evaluate potential effects of bin location and/or bin size on the apparent clade-level pattern (see electronic supplementary material, appendix S1).

To investigate how clade-level ECV patterns were driven by mechanisms operating at lower taxonomic levels, we divided the ECV data into 0.3 Myr time bins—the smallest bin size that yielded no temporal gaps within lineages—and then additively partitioned between-bin, clade-level ECV changes into their anagenetic, origination and extinction components ([34]; see electronic supplementary material, appendix S1). If lineages survive from one time bin to the next, clade-level change attributed to anagenesis is defined as the difference in mean ECV of surviving lineages between adjacent time bins. Clade-level change due to origination is calculated as the mean of all ECVs (including non-surviving lineages) in the later time bin minus the mean of ECVs from surviving lineages in the same bin, and extinction is the mean ECV of surviving lineages in the earlier time bin minus all ECVs in the same bin. Clade-level changes between time bins with no surviving lineages cannot be partitioned into separate origination/extinction components. Instead of determining which evolutionary mode best characterizes each lineage, this method quantifies the relative importance of microevolution (anagenesis) and macroevolution (origination and extinction) in driving the broader clade-level pattern during a specific time period. We should emphasize that origination and extinction in this case refer, respectively, to observed first and last appearances of species' *cranial* specimens (from which ECV could be estimated) and not the observed appearances of the species as a whole nor their true, unknown ages of origination and extinction [35]. We refer to 'first appearances' and 'last appearances' when describing these phenomena. Because our partitioning method required assigning specimens to anagenetic lineages, we conducted these analyses using taxonomies that recognized smaller and larger numbers of lineages ('less' and 'more' speciose, respectively) to determine if taxonomic philosophy had any influence on our results (electronic supplementary material, figure S4). Similar to our clade-level analyses, we also resampled age and ECV estimates to examine the effect of dating and inter-observer measurement error on our results (electronic supplementary material, figure S3), and conducted sensitivity analyses to evaluate the effects of bin location and size (see electronic supplementary material, appendix S1).

Finally, we examined the rate of ECV change within lineages. Increases in ECV are usually hypothesized as being caused only by natural selection at the population level. However, these hypotheses are difficult to test given the large timescale discrepancy between microevolutionary processes (i.e. generational timescales of 10^1 years in extant great apes [36]) and the timescales of hominin fossil assemblages (i.e. typical temporal resolution of 10^5 years). Furthermore, as mentioned previously, macroevolutionary processes (i.e. origination and extinction) can independently influence ECV, which confounds attempts to estimate how ECV changes at the population level. To address this issue, we used the results from our partitioning analyses to isolate the component of ECV evolution attributable only to within-lineage change (i.e. anagenesis). We then used a simple exponential growth model to investigate whether published estimates of the magnitude and rate of natural selection on generational timescales can be extrapolated to predict observed within-lineage ECV change on geological timescales. Proportional change over multiple generations in some trait for a given population can be calculated with the equation: $(1 + e_\mu \beta_\mu)^t$, where β_μ is selection strength (mean-standardized selection gradient [37]), e_μ is a measure of evolutionary potential quantifying the expected trait response to selection (mean-standardized evolvability [38]), and t is the number of generations [38]. Using a database of morphological traits, Hereford *et al*. found median β_μ is 0.28 for multi-trait studies that take into account trait covariance [37]. This is an appropriate assumption here since hominin brain and body size are known to be highly correlated [39], so selection strength acting on brain size will likely be diminished due to selection on other correlated traits (including, but not limited to, body size) acting in different directions. As a result, univariate selection gradients are spuriously too high and are not appropriate for predicting actual evolutionary change for traits that are correlated with others [37]. Hominin brain size has an e_μ of 0.006%, as estimated from modern human phenotypic data ([39]; M. Grabowski 2016, personal communication). The bin size in our additive partitioning analyses is 0.3 Myr, which translates to 11 110 generations, assuming a hominin generation time of 27 years [36]. If ECV change from the extrapolated exponential model matches observed anagenetic change in the fossil record, this would parsimoniously suggest that hypothesized drivers of natural selection (e.g. socio-cultural complexity, symbolic behaviour and complex language) were operating consistently, rather than episodically, to increase ECV within hominin populations.

Some fossil cranial specimens are particularly incomplete, distorted or crushed. The inclusion of such specimens (i.e. *Australopithecus garhi* [BOU-VP-12/130], *Paranthropus aethiopicus* [Omo L.338y-6], *Paranthropus boisei* [Omo 323-1976-896] and *Homo erectus s.l.* [KNM-OL 45500, Lantian, OH 12, Sangiran 3, Sangiran 31, Yunxian and Zhoukoudian VI]) could potentially bias ECV estimates and thus affect our results and inferences. To address this potential shortcoming and further evaluate the robustness of our results, we repeated our clade-level model selection and lower-level additive partitioning analyses with these 10 specimens removed from the dataset.

All analyses were done in R 3.0.3 [40]. See the electronic supplementary material for more details on how we collected

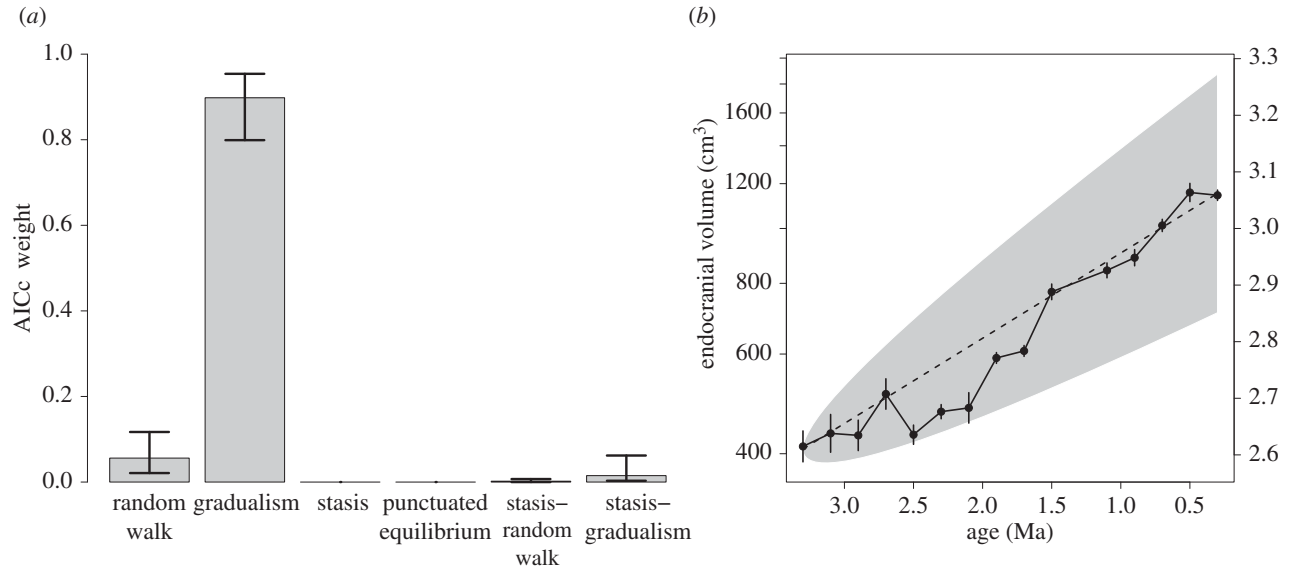

Figure 3. (*a*) Model selection results of the clade-level analysis testing six evolutionary modes. Bias-corrected Akaike information criterion (AICc) weights sum to one across all models, with higher weights representing more model support. Bars represent AICc weight medians, and error bars represent first and third quartiles from resampling age estimates (see electronic supplementary material, appendix S1). (*b*) Gradualism model fit for the clade-level ECV time series using 0.2 Myr bins. Binning here was done using observed (not resampled) age midpoints for plotting purposes only (see electronic supplementary material, appendix S1). Points are mean ECV estimates, and error bars are $\pm$ 1 s.e. Dotted line represents the expected evolutionary trajectory of the fitted gradualism model surrounded by the 95% probability envelope in grey. *Y*-axes as in figure 2.

and analysed the data, as well as the R script for executing analyses.

3. Results

Results from fitting evolutionary models at the clade level show gradualism is by far the best fit for describing hominin ECV change over time (figure 3*a* and electronic supplementary material, table S3). All mean ECV estimates of the observed time series fall within the 95% probability envelope predicted by the gradualism model (figure 3*b* and electronic supplementary material figure S1), and model R^2 is 0.676 (electronic supplementary material, table S4). Multiple sensitivity analyses demonstrate that support for the gradualism model is robust to bin size or location (electronic supplementary material, figure S5).

The second set of analyses shows that clade-level changes between bins—which are the sum of their anagenetic and first and last appearances components—are always positive (figure 4), thus corroborating the gradualism result described above. When we decompose clade-level changes using the less speciose taxonomy, the gradualism pattern is mainly driven by brain size increase within lineages, accounting for 88% of total clade-level change (figure 4 and electronic supplementary material, figure S6). Before 2.6 Ma, however, the extent to which anagenetic change outpaces macroevolutionary change is obscured by dating uncertainty in both the less and more speciose taxonomies (figure 4). Despite the dominance of anagenetic (microevolutionary) processes in the less speciose taxonomy, there are still times when macroevolutionary processes are important. The non-trivial influence of first appearances between 2.3 and 2 Ma and last appearances between 2 and 1.7 Ma (figure 4 and electronic supplementary material, figure S6) is driven by the appearance of *Homo* and disappearance of southern African *Australopithecus-Paranthropus*, respectively, at *ca* 2 Ma (electronic supplementary material, figure S7). The disappearance of eastern African *Paranthropus* at *ca* 1.4 Ma (electronic supplementary material, figure S7) appears to drive the clade-level pattern between 1.4 and 1.1 Ma (figure 4 and electronic supplementary material figure, S6).

Using the more speciose taxonomy, anagenesis is still the principal driver of clade-level ECV increase (64% of clade-level change). However, because more lineages are included, macroevolutionary processes inevitably take on a greater role relative to that observed in the less speciose taxonomy (figure 4 and electronic supplementary material, figure S6). Examples include the first appearances of *Paranthropus robustus*, early *Homo* and *Homo ergaster ca* 2 Ma for the time period between 2.3 and 2 Ma; the disappearances of *Australopithecus africanus* and *P. robustus ca* 2 Ma and the first appearance of *Homo erectus sensu stricto ca* 1.7 Ma for the time period between 2 and 1.7 Ma; the disappearances of eastern African *Paranthropus* and early *Homo ca* 1.4 Ma for the time period between 1.4 and 1.1 Ma; and the disappearance of *H. ergaster ca* 0.8 Ma for the time period between 0.8 and 0.5 Ma (figure 4, and electronic supplementary material, figures S6 and S7).

To summarize the results from both hypothesized taxonomies, within-lineage ECV increase is the primary driver of clade-level change at 1.7–1.4, 1.1–0.8 and 0.5–0.2 Ma. Macroevolutionary processes are important at 2–1.7 Ma (though anagenesis also contributes here in the less speciose taxonomy) and at 1.4–1.1 Ma (qualitatively similar to the results of ref. [18]). These patterns are robust to taxonomic philosophy, dating error, inter-observer ECV measurement error, bin location and bin size (figure 4, electronic supplementary material, figures S6, S8 and S9).

For the final set of analyses, we find hominin ECV within lineages is expected to increase via directional selection by 0.08 $\log_{10}cm^3$, or a factor of 1.2, over the course of 0.3 Myr (the bin size of our additive partitioning analyses). When comparing this expectation (black, horizontal line in figure 4) to observed within-lineage increases in fossil hominin ECV, we see the observed increases are always below the expected increase derived from the exponential model (except for the 1.7–1.4 Ma time period; figure 4 and electronic supplementary

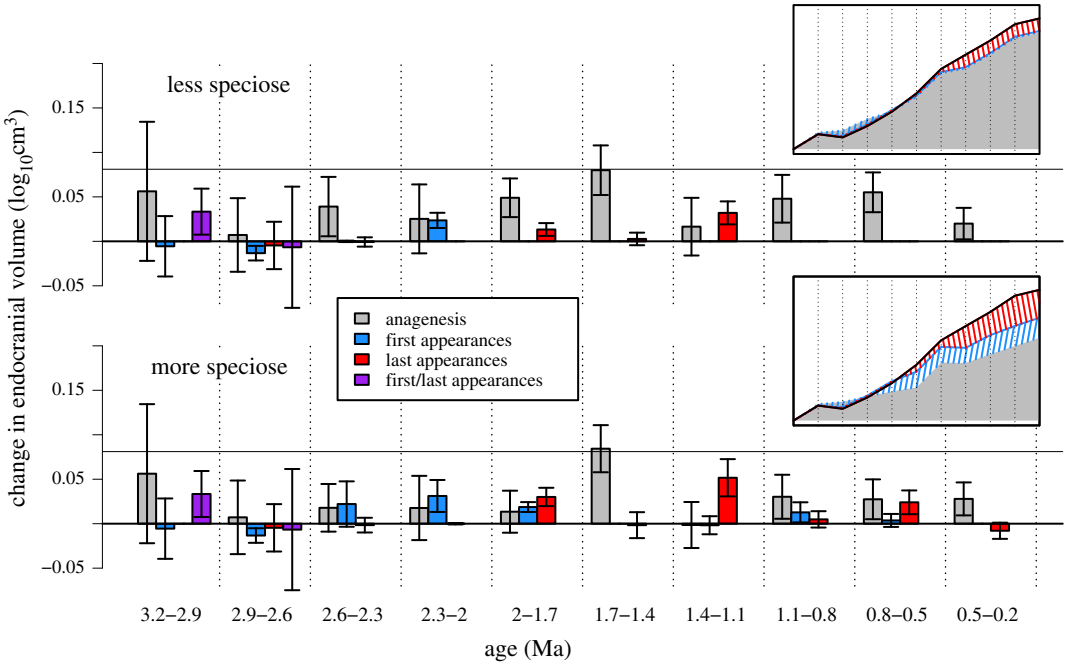

Figure 4. Additive partitioning of clade-level ECV transitions into their anagenetic and observed first and last appearances components. Ages separated by hyphens indicate over which two age bins (represented by their midpoints) the ECV transition is measured. 'First/last appearances' represents macroevolutionary change that cannot be partitioned into separate first or last appearances components (see 'Methods'). The sum of all partition means within a given time period equals the mean clade-level change, which in this case are all positive. Error bars are $\pm$ 1 s.e. calculated by randomly resampling age estimates (see electronic supplementary material, appendix S1). The horizontal black line represents the expected amount of within-lineage ECV increase over 0.3 Myr given our knowledge of how quickly natural selection operates from microevolutionary studies. Insets depict the cumulative effect of each component's mean (excluding 'first/last appearances') on the net clade-level trend (black line). Vertical dotted lines in the inset correspond to the vertical dotted lines in the main figure. The top graph uses a taxonomy that recognizes fewer taxa, while the bottom graph uses a taxonomy that recognizes a larger number of taxa.

material, figure S6). These findings are consistent with previous research showing that evolutionary rates in fossil lineages are slower than those in modern microevolutionary studies, even when fossil lineages show evidence of directional trends [41].

Our results are unaffected by the removal of especially incomplete or damaged specimens (electronic supplementary material, figure S10 and table S5).

4. Discussion

Our results show that fossil hominin ECV data at the clade level are most consistent with a gradual pattern of ECV increase through time. Understanding how this pattern emerged from processes that operate at lower taxonomic levels is more complicated. Our analyses are consistent with microevolutionary mechanisms as the dominant driver of clade-level change (64 or 88% of change using a more or less speciose taxonomy, respectively), alternating with secondary macroevolutionary mechanisms. This implies changing selective pressures and shifts in the relative importance of different evolutionary processes through time.

To date, most explanations for hominin brain size increase have focused on microevolutionary mechanisms. These hypotheses can explain anagenetic patterns but may not be relevant for patterns caused by origination and extinction [14,42]. For example, some researchers argue extinction is an emergent phenomenon because species do not go extinct for the same reasons individual organisms die [42]. ECV increase via anagenesis and lineage splitting are likely different enough processes that it makes sense to understand how each independently influenced ECV increase. Our results emphasize that origination and extinction were also important in shaping ECV patterns at the clade level, and both micro- and macroevolutionary change influenced hominin brain size to different extents at different times (figure 4 and electronic supplementary material, figure S6). Therefore, we must construct new, comprehensive theories to explain potential influences on hominin brain size evolution.

Periods when macroevolutionary processes drove clade-level ECV increase were, by definition, characterized by a combination of factors promoting origination and extinction. These factors may have included large-scale climate and environmental change [12,43], habitat fragmentation and vicariance [44], interspecific interactions [45] etc. It is worth repeating here that inferred periods of macroevolutionary importance were estimated using observed first and last appearance dates of lineages' cranial specimens, and these dates are very likely to shift as new fossil specimens are discovered (specifically, first appearances would become older, and last appearances would become younger) [35].

It is noteworthy that almost all the first and last appearances are associated with an increase in average clade-level brain size, and the importance of each is staggered in time (i.e. appearance of large-brained species mostly from *ca* 2.3 to 1.7 Ma and disappearance of smaller-brained species from *ca* 1.7 to 0.5 Ma) (figure 4 and electronic supplementary material, figure S6). The connection between first appearances and larger brain size is an example of directional speciation [14], where there is a consistent, biased shift in the phenotypes of daughter lineages relative to that of their ancestral lineages. Directional speciation may be caused by developmental or evolutionary constraints that bias phenotypic

change towards larger ECV, or selection for larger brains in peripatric populations [14,16,17]. The association between last appearances and smaller-brained species may signal some kind of extinction selectivity, either directly or indirectly related to ECV (e.g. extinction rates may be correlated with geographic range size which in turn is correlated with body and brain size). This result is corroborated by smaller-brained species having shorter lineage durations (mean species' ECV versus mean species' duration: Spearman's $\rho = 0.41$), and therefore higher mean extinction rates [26] (this also suggests greater persistence of larger-brained species).

If species with larger ECVs are found to have higher diversification rates (origination minus extinction rates), this may suggest that species sorting also caused clade-level ECV to increase [46,47]. Just as natural selection operates via differential birth/death of individuals associated with a given trait, species sorting operates via differential origination/extinction of species, in this case, associated (directly or indirectly) with brain size. Species sorting is implied by the increased variation in brain size between 2.0 and 1.5 Ma (figure 2), generated by the addition of larger-brained lineages via directional speciation; this increased variation was later culled by selective extinction of smaller-brained lineages (figure 4 and electronic supplementary material, figure S6). If species sorting is borne out, it would suggest that all three mechanisms known to influence phenotypic evolution within a clade (i.e. anagenesis, directional speciation and species sorting) were acting in concert at multiple taxonomic scales to produce the directional ECV trend observed at the hominin clade level (as proposed elsewhere for hominin body mass and stature [48]). Moreover, the potential influence of species sorting requires a reorientation of how we think about hominin brain size evolution, since oft-proposed microevolutionary mechanisms are not necessary and may not be sufficient to generate higher-level sorting.

Inferring the potential drivers for periods of anagenetic change is more difficult. Within-lineage trends are typically explained as being caused by *only* directional selection. However, the observed rate of within-lineage ECV increase is too slow to be consistent with uniform directional selection, given our knowledge from empirical microevolutionary studies [49] and theoretical models (like the exponential model above) about the rate at which natural selection operates. In fact, to 'force' the model prediction downwards to match the observed data (i.e. mean anagenetic increase in the less speciose taxonomy; figure 4), one would need to decrease the mean-standardized selection gradient by 50% from 0.28 to 0.14 (if evolvability is held constant). In their compilation of mean-standardized selection gradients, Hereford *et al*. found that such low estimates are so small as to not be significantly different from zero (fig. 3 in ref. [37]). Our finding is consistent with those from other researchers who have shown that rates of hominin phenotypic evolution are consistent with, or even slower than, random genetic drift alone [50–52]. The potential prevalence of genetic drift should perhaps not be surprising given the rarity of hominins in the fossil record [53], which implies small population sizes, but drift is difficult to reconcile with the strongly directional ECV pattern we find within hominin lineages.

We hypothesize that the seemingly too slow within-lineage evolutionary rate is caused by the dynamics of the selective pressures themselves (i.e. the adaptive landscape) over evolutionary time [41,54,55]. Microevolutionary studies have shown that populations can respond rapidly to selection pressures on generational timescales (i.e. populations rapidly climb the adaptive peak and stay at the summit). Such high rates, however, need not characterize the tempo at which the adaptive peak *itself* moves over geological time [41]. Therefore, selection was for larger ECV on average but must have fluctuated and included episodes of stasis and/or drift. All of this occurs on too fine a timescale to be resolved by the current hominin ECV fossil record, resulting in emergent directional trends within lineages. If this is the case, the microevolutionary question of interest shifts to what caused ECV selection pressures to fluctuate but still ultimately select for larger ECVs on average at a geologically gradual pace [41]?

5. Conclusion

Traditionally, a gradual trend has been interpreted as the result of consistent directional selection at the population level for larger brains. However, when taxonomic scale is accounted for, we find the gradual, clade-level trend was generated primarily by within-lineage mechanisms that likely involved both directional selection *and* stasis and/or drift. In addition to these within-lineage processes, directional speciation producing larger-brained lineages, higher extinction rates of smaller-brained lineages and potentially higher-level species sorting all worked together to generate the strongly trended, emergent clade-level pattern. Our findings illustrate the complicated, multi-causal nature of hominin ECV evolution and the need for future hypotheses and models to recognize and incorporate this hierarchical complexity. There is no one canonical scale at which to conduct evolutionary research, and different questions can and should be asked when studying ECV increase at different scales. The analytical framework we suggest can be used to generate more precise hypotheses pinpointing when and at what taxonomic level ECV increase occurred, thus enabling stronger tests of proposed explanations. For example, predictions of the rate and magnitude of ECV increase from models invoking microevolutionary processes (e.g. stone tool innovation, major dietary shifts) can be benchmarked against the anagenetic partitions in figure 4 and electronic supplementary material, figure S6.

Within just the past few years, it has been made clear through fossil discoveries, or through comprehensive analyses permitted by those discoveries, that many of the so-called 'defining' characteristics of modern humans emerged through evolutionary processes that were significantly more complicated than had previously been appreciated. For example, we now have direct fossil evidence of a widely diverse set of adaptations for bipedalism in Pliocene hominins [56] and direct archaeological evidence that stone tool technologies were potentially manufactured and used by hominins well before the emergence of the 'handy man', *Homo habilis* [57,58]. This trend of falsifying and then refining hypotheses after the emergence of new fossil data is inevitable in palaeobiology. However, it has proved easier to accumulate, but more difficult to reject, hypotheses for why ECV increased in the hominin clade. Certain hypotheses may actually prove unfalsifiable if they explain evolutionary patterns in ways that are too imprecise and overly general [59]. Palaeoanthropologists, as with other practitioners of historical science, cannot conduct experiments with their data, so they must rely on theory to develop precise, falsifiable predictions to elucidate the mechanisms

underlying observed patterns [59–61]. Informed by micro- and macroevolutionary theory, our taxonomically scale-explicit analyses provide a revised, quantitatively rigorous framework for both developing and testing hypotheses and models related to the evolution of hominin brain size. This moves us closer to identifying and understanding what ultimately drove the evolution of large brains in the human clade.

Data accessibility. The datasets supporting this article have been uploaded as part of the electronic supplementary material and are also available via the Dryad Digital Repository at https://doi.org/10.5061/dryad.c30g9 [62].

Authors' contributions. All authors contributed to the formulation of the project. A.M.Z., K.G.H., E.R. and A.D. collected data. A.D. formulated the research design and performed analyses. A.D. wrote the paper with contributions from B.A.W. and K.G.H.

Competing interests. We have no competing interests.

Funding. Research was supported by National Science Foundation IGERT DGE-080163 and SMA-1409612.

Acknowledgements. We thank Gene Hunt for methodological and coding advice and comments on the manuscript, Carl Simpson, Kjetil Voje and Mark Grabowski for methodological advice, and Aida Gómez-Robles and Andrew Barr for comments on an earlier version of the manuscript.

References

1. Count EW. 1947 Brain and body weight in man: their antecedents in growth and evolution: a study in dynamic somatometry. *Ann. NY Acad. Sci.* **46**, 993–1122. (doi:10.1111/j.1749-6632.1947.tb36165.x)
2. Aiello LC, Dunbar RIM. 1993 Neocortex size, group size, and the evolution of language. *Curr. Anthropol.* **34**, 184–193. (doi:10.2307/2743982)
3. Holloway RL, Broadfield DC, Yuan MS. 2004 *The human fossil record, brain endocasts: the paleoneurological evidence*, vol. 3. New York, NY: Wiley-Liss.
4. Lestrel PE, Read DW. 1973 Hominid cranial capacity versus time: a regression approach. *J. Hum. Evol.* **2**, 405–411. (doi:10.1016/0047-2484(73)90020-1)
5. Cronin JE, Boaz NT, Stringer CB, Rak Y. 1981 Tempo and mode in hominid evolution. *Nature* **292**, 113–122. (doi:10.1038/292113a0)
6. Pagel M. 2002 Modelling the evolution of continuously varying characters on phylogenetic trees. In *Morphology, shape and phylogeny* (eds N MacLeod, PL Forey), pp. 269–286. Boca Raton, FL: CRC Press.
7. Lee S-H, Wolpoff MH. 2003 The pattern of evolution in Pleistocene human brain size. *Paleobiology* **29**, 186–196. (doi:10.1666/0094-8373(2003)029<0186:TPOEIP>2.0.CO;2)
8. Gould SJ, Eldredge N. 1977 Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology* **3**, 115–151. (doi:10.1017/S0094837300005224)
9. Rightmire GP. 1981 Patterns in the evolution of *Homo erectus*. *Paleobiology* **7**, 241–246. (doi:10.1017/S0094837300004012)
10. Hofman MA. 1983 Encephalization in hominids: evidence for the model of punctuationalism. *Brain. Behav. Evol.* **22**, 102–117. (doi:10.1159/000121511)
11. Hawks J. 2011 No brain expansion in *Australopithecus boisei*. *Am. J. Phys. Anthropol.* **146**, 155–160. (doi:10.1002/ajpa.21420)
12. deMenocal PB. 2011 Climate and human evolution. *Science* **331**, 540–542. (doi:10.1126/science.1190683)
13. Antón SC, Potts R, Aiello LC. 2014 Evolution of early *Homo*: an integrated biological perspective. *Science* **345**, 1236828. (doi:10.1126/science.1236828)
14. Gould SJ. 2002 *The structure of evolutionary theory*. Cambridge, MA: Belknap Press of Harvard University Press.
15. Novack-Gottshall PM, Lanier MA. 2008 Scale-dependence of Cope's rule in body size evolution of Paleozoic brachiopods. *Proc. Natl Acad. Sci. USA* **105**, 5430–5434. (doi:10.1073/pnas.0709645105)
16. McShea DW. 2004 A revised Darwinism. *Biol. Philos.* **19**, 45–53. (doi:10.1023/B:BIPH.0000013260.40162.dd)
17. Jablonski D. 2007 Scale and hierarchy in macroevolution. *Palaeontology* **50**, 87–109. (doi:10.1111/j.1475-4983.2006.00615.x)
18. Shultz S, Nelson E, Dunbar RIM. 2012 Hominin cognitive evolution: identifying patterns and processes in the fossil and archaeological record. *Phil. Trans. R. Soc. B* **367**, 2130–2140. (doi:10.1098/rstb.2012.0115)
19. De Miguel C, Henneberg M. 2001 Variation in hominid brain size: how much is due to method? *HOMO - J. Comp. Hum. Biol.* **52**, 3–58. (doi:10.1078/0018-442X-00019)
20. Stanyon R, Consigliere S, Morescalchi MA. 1993 Cranial capacity in hominid evolution. *Hum. Evol.* **8**, 205–216. (doi:10.1007/BF02436715)
21. Leigh SR. 1992 Cranial capacity evolution in *Homo erectus* and early *Homo sapiens*. *Am. J. Phys. Anthropol.* **87**, 1–13. (doi:10.1002/ajpa.1330870102)
22. Rightmire GP. 2004 Brain size and encephalization in early to Mid-Pleistocene *Homo*. *Am. J. Phys. Anthropol.* **124**, 109–123. (doi:10.1002/ajpa.10346)
23. Godfrey L, Jacobs KH. 1981 Gradual, autocatalytic and punctuational models of hominid brain evolution: a cautionary tale. *J. Hum. Evol.* **10**, 255–272. (doi:10.1016/S0047-2484(81)80063-2)
24. McHenry HM. 1982 The pattern of human evolution: studies on bipedalism, mastication, and encephalization. *Annu. Rev. Anthropol.* **11**, 151–173. (doi:10.1146/annurev.an.11.100182.001055)
25. Wood B. 2010 Reconstructing human evolution: achievements, challenges, and opportunities. *Proc. Natl Acad. Sci. USA* **107**, 8902–8909. (doi:10.1073/pnas.1001649107)
26. Foote M, Miller AI. 2007 *Principles of paleontology*, 3rd edn. New York: NY: W.H. Freeman.
27. Hunt G. 2015 paleoTS: analyze paleontological time-series. R package version 0.5-1. See cran.r-project.org/web/packages/paleoTS/.
28. Hunt G. 2006 Fitting and comparing models of phyletic evolution: random walks and beyond. *Paleobiology* **32**, 578–601. (doi:10.1666/05070.1)
29. Hunt G. 2008 Gradual or pulsed evolution: when should punctuational explanations be preferred? *Paleobiology* **34**, 360–377. (doi:10.1666/07073.1)
30. Hunt G. 2008 Evolutionary patterns within fossil lineages: model-based assessment of modes, rates, punctuations and process. In *From evolution to geobiology: research questions driving paleontology at the start of a new century. Paleontological society papers* (eds PH Kelley, RK Bambach), pp. 117–131. Pittsburgh, PA: Paleontological Society.
31. Hunt G, Hopkins MJ, Lidgard S. 2015 Simple versus complex models of trait evolution and stasis as a response to environmental change. *Proc. Natl Acad. Sci. USA* **112**, 4885–4890. (doi:10.1073/pnas.1403662111)
32. Sheets HD, Mitchell CE. 2001 Why the null matters: statistical tests, random walks and evolution. *Genetica* **112–113**, 105–125. (doi:10.1023/A:1013308409951)
33. Burnham KP, Anderson DR. 2002 *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn. New York, NY: Springer.
34. Rego BL, Wang SC, Altiner D, Payne JL. 2012 Within- and among-genus components of size evolution during mass extinction, recovery, and background intervals: a case study of Late Permian through Late Triassic foraminifera. *Paleobiology* **38**, 627–643. (doi:10.1666/11040.1)
35. Wang SC, Marshall CR. 2016 Estimating times of extinction in the fossil record. *Biol. Lett.* **12**, 20150989. (doi:10.1098/rsbl.2015.0989)
36. Langergraber KE *et al.* 2012 Generation times in wild chimpanzees and gorillas suggest earlier divergence times in great ape and human evolution. *Proc. Natl Acad. Sci. USA* **109**, 15 716–15 721. (doi:10.1073/pnas.1211740109)
37. Hereford J, Hansen TF, Houle D, Fenster C. 2004 Comparing strengths of directional selection: how strong is strong? *Evolution* **58**, 2133–2143. (doi:10.1111/j.0014-3820.2004.tb01592.x)

38. Hansen TF, Pélabon C, Houle D. 2011 Heritability is not evolvability. *Evol. Biol.* **38**, 258–277. (doi:10.1007/s11692-011-9127-6)
39. Grabowski M. 2016 Bigger brains led to bigger bodies? The correlated evolution of human brain and body size. *Curr. Anthropol.* **57**, 174–196. (doi:10.1086/685655)
40. R Core Team. 2017 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. See http://www.R-project.org/.
41. Geary DH, Hunt G, Magyar I, Schreiber H. 2010 The paradox of gradualism: phyletic evolution in two lineages of lymnocardiid bivalves (Lake Pannon, central Europe). *Paleobiology* **36**, 592–614. (doi:10.1666/08065.1)
42. Simpson C. 2016 The case for species selection. bioRxiv. (doi:10.1101/084046)
43. Shultz S, Maslin M. 2013 Early human speciation, brain expansion and dispersal influenced by African climate pulses. *PLoS ONE* **8**, e76750. (doi:10.1371/journal.pone.0076750)
44. Vrba ES. 1992 Mammals as a key to evolutionary theory. *J. Mammal.* **73**, 1–28. (doi:10.2307/1381862)
45. Rabosky DL. 2013 Diversity-dependence, ecological speciation, and the role of competition in macroevolution. *Annu. Rev. Ecol. Evol. Syst.* **44**, 481–502. (doi:10.1146/annurev-ecolsys-110512-135800)
46. Vrba ES, Gould SJ. 1986 The hierarchical expansion of sorting and selection: sorting and selection cannot be equated. *Paleobiology* **12**, 217–228. (doi:10.1017/S0094837300013671)
47. Jablonski D. 2008 Species selection: theory and data. *Annu. Rev. Ecol. Evol. Syst.* **39**, 501–524. (doi:10.1146/annurev.ecolsys.39.110707.173510)
48. Will M, Pablos A, Stock JT. 2017 Long-term patterns of body mass and stature evolution within the hominin lineage. *R. Soc. open sci.* **4**, 171339. (doi:10.1098/rsos.171339)
49. Kinnison MT, Hendry AP. 2001 The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* **112–113**, 145–164. (doi:10.1007/978-94-010-0585-2_10)
50. Lynch M. 1990 The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am. Nat.* **136**, 727–741. (doi:10.1086/285128)
51. Williams GC. 1992 *Natural selection: domains, levels, and challenges*. New York, NY: Oxford University Press.
52. Grabowski M, Roseman CC. 2015 Complex and changing patterns of natural selection explain the evolution of the human hip. *J. Hum. Evol.* **85**, 94–110. (doi:10.1016/j.jhevol.2015.05.008)
53. Bobe R, Leakey MG. 2009 Ecology of Plio-Pleistocene mammals in the Omo-Turkana basin and the emergence of *Homo*. In *The first humans—origin and early evolution of the genus* Homo (eds FE Grine, JG Fleagle, RE Leakey), pp. 173–184. Berlin, Germany: Springer Netherlands.
54. Arnold SJ, Pfrender ME, Jones AG. 2001 The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* **112–113**, 9–32. (doi:10.1023/A:1013373907708)
55. Hansen TF. 2012 Adaptive landscapes and macroevolutionary dynamics. In *The adaptive landscape in evolutionary biology* (eds EI Svensson, R Calsbeek), pp. 205–226. Oxford, UK: Oxford University Press.
56. Haile-Selassie Y, Saylor BZ, Deino A, Levin NE, Alene M, Latimer BM. 2012 A new hominin foot from Ethiopia shows multiple Pliocene bipedal adaptations. *Nature* **483**, 565–569. (doi:10.1038/nature10922)
57. McPherron SP, Alemseged Z, Marean CW, Wynn JG, Reed D, Geraads D, Bobe R, Béarat HA. 2010 Evidence for stone-tool-assisted consumption of animal tissues before 3.39 million years ago at Dikika, Ethiopia. *Nature* **466**, 857–860. (doi:10.1038/nature09248)
58. Harmand S *et al.* 2015 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature* **521**, 310–315. (doi:10.1038/nature14464)
59. McGill B. 2003 Strong and weak tests of macroecological theory. *Oikos* **102**, 679–685. (doi:10.1034/j.1600-0706.2003.12617.x)
60. Lakatos I. 1978 Introduction: science and pseudoscience. In (eds J Worrall, G Currie), The methodology of scientific research programmes, pp. 1–7. Cambridge, UK: Cambridge University Press.
61. Binford LR. 1981 *Bones: ancient men and modern myths*. San Diego, CA: Academic Press.
62. Du A, Zipkin AM, Hatala KG, Renner E, Baker JL, Bianchi S, Bernal KH, Wood BA. 2018 Data from: Pattern and process in hominin brain size evolution are scale-dependent. Dryad Digital Repository. (https://doi.org/10.5061/dryad.c30g9)