Article

Rules of teeth development align microevolution with macroevolution in extant and extinct primates

Check for updates

Fabio A. Machado [1] ✉, **Carrie S. Mongle** [2,3], **Graham Slater**[4], **Anna Penna** [5], **Anna Wisniewski**[4], **Anna Soffin**[6], **Vitor Dutra**[7] & **Josef C. Uyeda**[6]

Macroevolutionary biologists have classically rejected the notion that higher-level patterns of divergence arise through microevolutionary processes acting within populations. For morphology, this consensus partly derives from the inability of quantitative genetics models to correctly predict the behaviour of evolutionary processes at the scale of millions of years. Developmental studies (evo-devo) have been proposed to reconcile micro- and macroevolution. However, there has been little progress in establishing a formal framework to apply evo-devo models of phenotypic diversification. Here we reframe this issue by asking whether using evo-devo models to quantify biological variation can improve the explanatory power of comparative models, thus helping us bridge the gap between micro- and macroevolution. We test this prediction by evaluating the evolution of primate lower molars in a comprehensive dataset densely sampled across living and extinct taxa. Our results suggest that biologically informed morphospaces alongside quantitative genetics models allow a seamless transition between the micro- and macroscales, whereas biologically uninformed spaces do not. We show that the adaptive landscape for primate teeth is corridor like, with changes in morphology within the corridor being nearly neutral. Overall, our framework provides a basis for integrating evo-devo into the modern synthesis, allowing an operational way to evaluate the ultimate causes of macroevolution.

'Macroevolution' is the field of study that aims to understand how the diversification of life occurred on our planet over large timescales[1]. Like any other historical science, it seeks to make sense of patterns over time ingrained in the fossil record and phylogenetic trees by referencing well-understood processes known from direct observations and experimentation[2]. In the case of evolutionary biology, this knowledge comes mainly from fields such as ecology and genetics, which tend to map

evolutionary phenomena that take place during shorter timescales. For this reason, these studies are sometimes called 'microevolution' and are designed to understand how population-level phenomena can produce evolutionary change. However, despite the presumed direct relationship between micro and macro levels, quantitative studies have struggled to explain most macroevolutionary patterns in terms of microevolutionary processes[3–7]. Nevertheless, empirical results have

[1]Department of Integrative Biology, Oklahoma State University, Stillwater, OK, USA. [2]Department of Anthropology, Stony Brook University, Stony Brook, NY, USA. [3]Turkana Basin Institute, Stony Brook University, Stony Brook, NY, USA. [4]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA. [5]Department of Anthropology, University of Texas at San Antonio, San Antonio, TX, USA. [6]Department of Biology, Virginia Tech, Blacksburg, VA, USA. [7]Department of Anthropology, Florida Atlantic University, Boca Raton, FL, USA. ✉e-mail: fabio.machado@okstate.edu
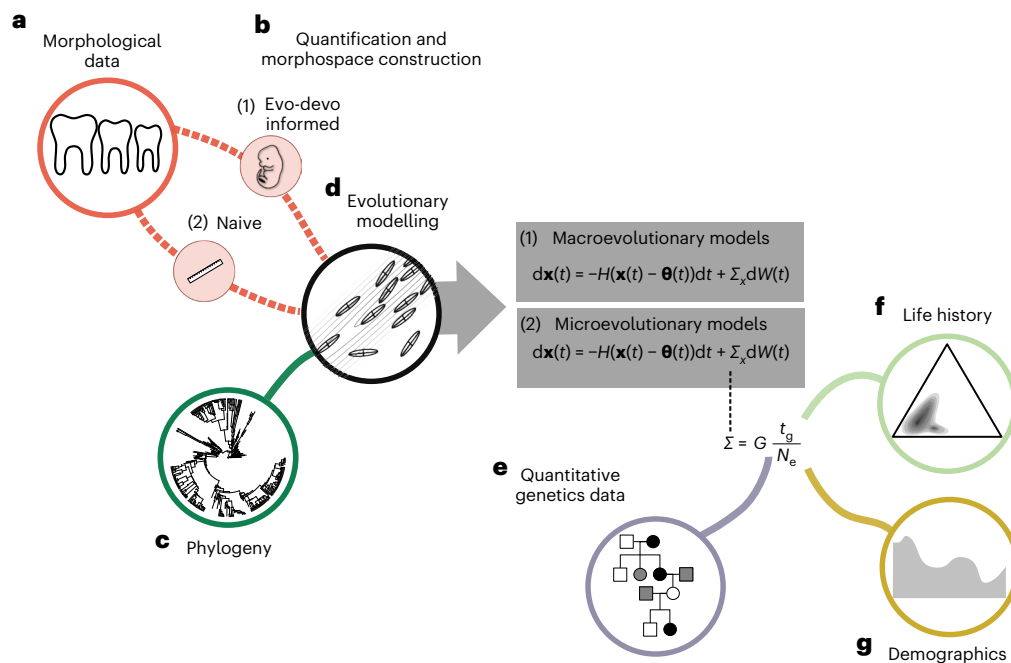
Fig. 1 | Integrative evolutionary modelling framework used in the present study. a, Sources of morphological data, such as direct measurements with calipers, measurements extracted from photographs and data obtained from the literature. **b**, Process of quantification, which can be either evo-devo inspired (1) or 'naive' in relation to developmental processes (2). **c**, Phylogeny of the group under study. **d**, Evolutionary modelling used to infer adaptive landscapes (isolines) where species (ellipses) have evolved. Evolutionary models can be either macroevolutionary (1) or microevolutionary (2). Both models belong to the Ornstein–Uhlenbeck class of models with the same parameters (see Methods for a full description of the parameters) with the single difference being that on the microevolutionary models, the stochastic rate parameter Σ is constrained to be equal to the rate of genetic drift. The rate of drift is modelled as being proportional to Gt_g/N_e. **e**, Quantitative genetics data, specifically additive genetic variance–covariance matrices (G), estimated from pedigreed populations. **f**, Life history data, specifically the generation time used to estimate time of divergence in generations (t_g). **g**, Demographics data, specifically effective population size (N_e). In our framework, the morphological data (**a**) are used to construct naive or biologically informed morphospaces (**b**) and, together with a phylogeny (**c**), are used in an evolutionary modelling process (**d**). Different models, including microevolutionary and macroevolutionary ones, can then be directly compared if estimated under the same morphospace.

consistently shown that the availability of additive genetic variation correlates strongly with rates of macroevolution for different traits[6,8–13], suggesting some effect of lower-level microevolutionary processes at the macroevolutionary scale. Whether we can bridge the gap between these two scales is still unclear, with some authors arguing for their essential irreconcilability[3,14,15] and others advocating for reconciliation within the context of the modern synthesis[5,8–10,16–23].

One long-standing suggestion for bridging the gap between micro- and macroevolution has been through the study of developmental biology and ontogeny (that is, evo-devo)[10,19,24–28]. This suggestion, however, has been challenging to implement. In a microevolutionary context, development can often be reasonably assumed to be a smooth genotype-to-phenotype (GP) map; that is, genotypic variation translates to phenotypic variation in a linear way, with traits being influenced by multiple genes of smaller effect. Such a smooth GP map would, in turn, allow the modelling of evolution and adaptation of the adult phenotype using a quantitative genetic framework, precisely because these classes of models entail this simplified, linear GP mapping[18–20,29]. On larger timescales, however, genetic architectures can change, selection can fluctuate and development can be reorganized, generating nonlinearities between genotypic and phenotypic divergence, even if the GP map was originally smooth. On the phenotypic level, these nonlinearities can produce discontinuities, that is, regions of the morphospace less inhabited, or not inhabited at all, by species, impeding a straightforward extrapolation of microevolutionary processes over millions of years[19,27,30–32]. Therefore, in the absence of in-depth knowledge of development and the GP map, it is likely that macroevolutionary studies will find heterogeneity and discontinuities,

even if the underlying genetic changes at the microevolutionary scale are relatively smooth and continuous[19]. Alternatively, if we can use developmental models as the basis of the quantification of morphology, we might smooth out some of these nonlinearities, maximizing our ability to seamlessly connect micro- and macroevolutionary scales[9,10,19,30,33–36].

Here we propose a new framework for the investigation of morphological evolution over macroevolutionary time that explicitly models evolution at this scale as a consequence of underlying microevolutionary processes (Fig. 1). To deal with potential nonlinearities that might arise over long timescales, we suggest the construction of developmentally informed spaces (1 in Fig. 1b), which, coupled with quantitative genetics modelling (Fig. 1e–g) and comparative methods (Fig. 1d), can facilitate a conceptual bridge between micro- and macroscales. Under our proposed framework, we are able to directly compare microevolution-inspired models (henceforth called 'microevolutionary models') with non-microevolution-inspired ones that account for a wider variety of rate- and state-heterogenous evolutionary processes (henceforth called 'macroevolutionary models').

We test this workflow to investigate the evolution of primate molars, which is an ideal model system for the present investigation. First, there is a simple yet powerful evo-devo model that describes the development and evolution of mammalian molars, the inhibitory cascade model (ICM). The ICM models teeth size (that is, molar row form) as the result of a balance between inhibition and activation factors[35]. Specifically, it predicts that the sizes of the first, second and third molars (m1, m2 and m3, respectively) will either be the same (m1 = m2 = m3), increase (m1 < m2 < m3) or decrease (m1 > m2 > m3)
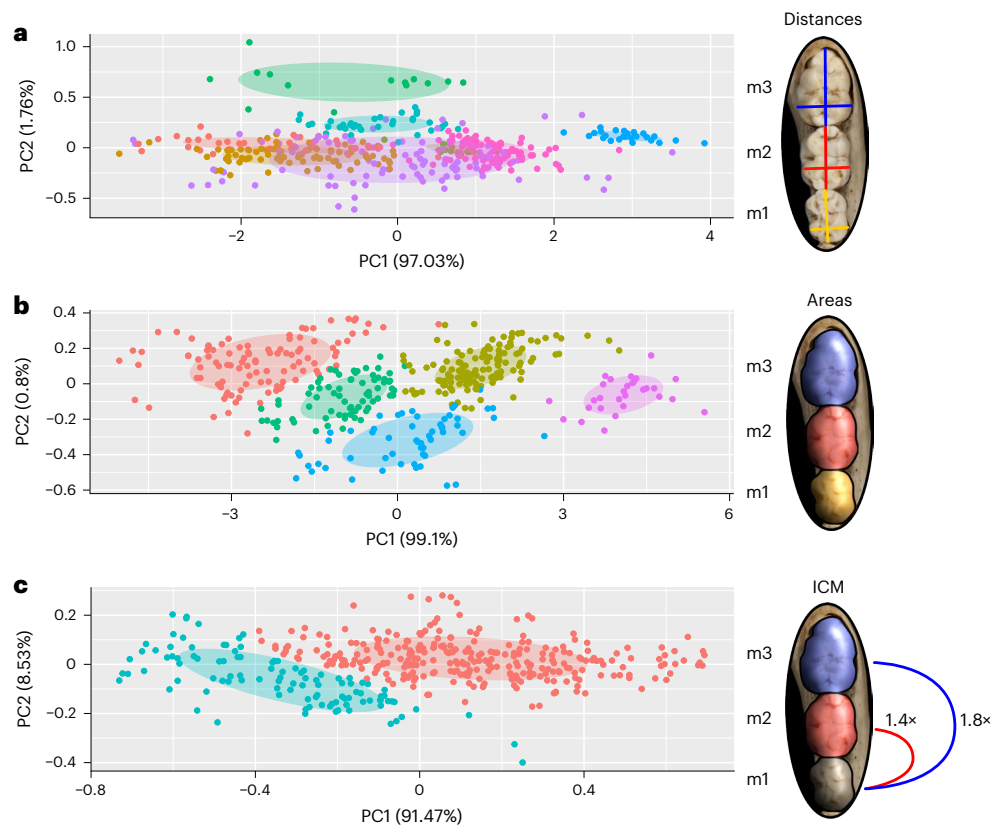
Fig. 2 | Principal component analysis of the full-sample covariance matrix for the three morphospaces. a–c, Linear-distance morphospace (**a**), area morphospace (**b**) and ICM ratio morphospace (**c**). Dots represent species averages (*n* = 480 species), and colours represent groups identified in the clustering analysis. These groups are based only on morphometric proximity and do not represent any taxonomic group. The more groups a morphospace has, the more patchy and discontinuous it is considered. Axes are not depicted to scale for convenience, so distances in the graph should not be considered representative of the metric of the underlying space. PC, principal component.

along the molar row. A corollary of this prediction is that there will be a positive relationship between the ratios of the areas of the last two molars in relation to the first one (m2/m1 and m3/m1), which thus establishes a natural morphospace to investigate this developmental process. This model was initially proposed for rodents[35] and later verified for multiple mammalian species[37,38], including Primates[11,39,40]. Second, there are several studies characterizing aspects of additive genetic variation in molars of Primates[41–44], as well as large-scale life history and demographic information for the group[45,46], parameters that are essential to model microevolutionary processes such as drift and selection (Fig. 1e–g). Third, tooth enamel is the most mineralized substance in vertebrate tissues, making teeth especially resistant to taphonomic processes and abundant in the fossil record (Fig. 1c). The use of a dense fossil record allows us to bridge some phylogenetic gaps between extant species, ensuring that heterogeneities along the tree are more likely due to differences in processes rather than incomplete sampling. This extensive availability of palaeontological and neontological data enables unprecedented power to evaluate evolutionary dynamics through deep time using data-hungry phylogenetic comparative methods[47–49]. We apply our framework to an expansive dataset of both extant (232 taxa) and extinct (248 taxa) species summarized from more than 250 different sources, integrated with a newly published comprehensive phylogeny[50]. To address our hypotheses, we use a model-fitting approach based on information theory (Bayesian information criteria or BIC) and model simplicity (minimizing the parameter number). We expect that variables devised to quantify developmental processes (ICM variables) will favour microevolutionary models, whereas data embedded in biologically 'naive'

spaces (those with no direct correspondence to any developmental model) will favour complex macroevolutionary models.

Results

To test our hypothesis that the use of developmental models to quantify morphological variation would provide a better bridge between micro- and macroevolution, we used three morphospaces (Extended Data Fig. 1). The first is based on the linear distances taken directly from the teeth ('distance space'), and the second is based on the occlusal areas of each molar ('area space'). These two spaces are considered naive because they make no assumptions about underlying developmental processes (2 in Fig. 1b). As an evo-devo-inspired space (1 in Fig. 1b), the third morphospace was constructed based on the relation between the relative occlusal area of m2 and m3 in relation to m1 (m2/m1 and m3/m1, respectively), using the ICM description of molar development[35].

We performed model-based clustering analyses of each set of measurements to test our prediction that development will generate discontinuous morphospaces. If, as explained above, complex developmental processes generate a patchy and discontinuous morphospace, then we expect to evaluate a high number of clusters on naive spaces. However, if the evo-devo-informed space corrects this issue, we will observe fewer clusters on the ICM morphospace. As expected, our clustering analysis shows a tendency of the ICM space to find fewer groups than the naive spaces, suggesting the former is less patchy than the latter (Fig. 2 and Extended Data Fig. 2).

Our model-fitting approach showed that the use of biologically naive morphospaces favours evolutionary model complexity.

Table 1 | Model comparison for the primate lower molar row evolution fit through maximum likelihood and ranking according to the BIC

Traits[a]	Model[b]	N_p[c]	logLik[d]	BIC[e]
Linear distances	BM	27	2,585.03	−5,003.38
	OU	54	2,684.57	−5,035.75
	$BM_{\Sigma \propto P}$	7	2,108.57	−4,173.93
	$OU_{\Sigma \propto P}$	34	2,174.45	−4,138.99
	$BM_{\Sigma \propto G}$	7	1,480.63	−2,918.04
	$OU_{\Sigma \propto G}$	34	1,510.57	−2,811.23
	Three-regime BM[f]	69	2,823.03	**−5,484.50**
Areas	BM	9	415.80	−776.03
	OU	18	413.98	−716.83
	$BM_{\Sigma \propto P}$	4	92.39	−160.09
	$OU_{\Sigma \propto P}$	13	275.16	−470.06
	$BM_{\Sigma \propto G}$	4	74.41	−124.13
	$OU_{\Sigma \propto G}$	13	258.02	−435.78
	Three-regime BM[f]	21	471.42	**−813.19**
ICM	BM	5	589.16	−1,147.46
	OU^D	9	604.56	<u>−1,153.55</u>
	$BM_{\Sigma \propto P}$	3	538.31	−1,058.10
	$OU_{\Sigma \propto P}$	8	584.68	−1,119.98
	$BM_{\Sigma \propto G}$	3	575.69	−1,132.86
	$OU^D_{\Sigma \propto G}$	7	598.54	**−1,153.86**
	Three-regime OU[f]	26	597.35	−1,034.18

[a]Morphospace used to quantify molar form variation, either a biologically naive space (linear distances or areas) or an evo-devo-inspired space (ICM). [b]Model type, either a global BM or OU model or a mixed model, which allows model and parameter heterogeneity. BM and OU can also incorporate the microevolutionary assumption that the evolutionary rate matrix (Σ) is proportional to G or P ($\Sigma \propto G$ or P models). 'D' indicates OU models with a diagonal H. [c]Number of model parameters. [d]Log likelihood of the model. [e]BIC used for model comparison. [f]Results for the mixed model for linear distances and ICM are based on the best regime combination found for the area morphospace. Bold indicates the best models. Underline indicates the model with BIC two units away from the best model.
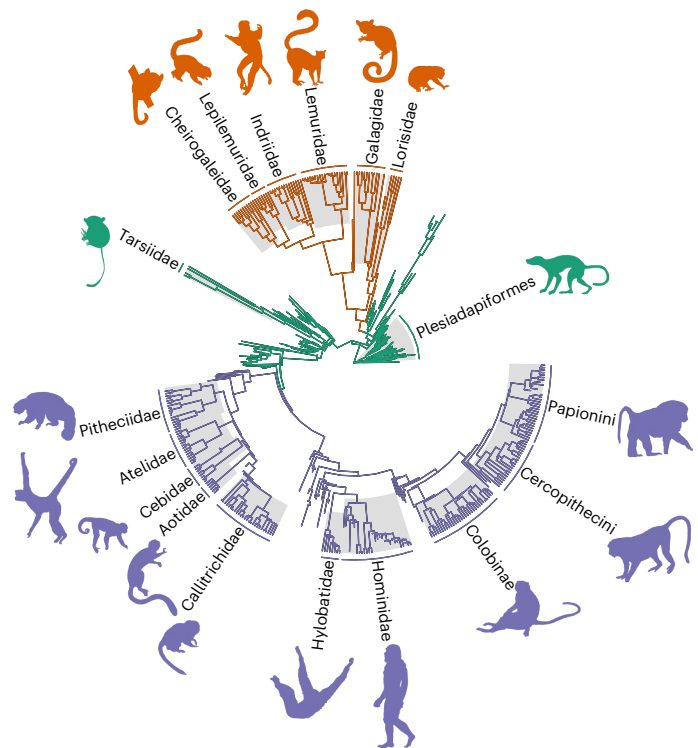

Fig. 3 | Primate phylogenetic tree of the 480 species included in this study painted by the best regime combination found on the phylogenetic mixed-model search for the individual molar areas. A multi-regime model allows each different part of the tree (regimes) to have a different model and/or parameter combination. For areas and linear distances, the best model overall is a multi-regime BM, meaning the different highlighted clades will have different rates of stochastic evolution (Σ). For ICM ratios, the best mixed model is a multi-regime OU, meaning that each clade will have different rates of stochastic (Σ) and deterministic (H) evolution and optima (θ). However, for ICM ratios, single-regime microevolutionary models outperform all mixed models (Table 1). Silhouettes reproduced from PhyloPic under Creative Commons licences CC0 1.0 and CC BY 3.0.

Specifically, the best model for these spaces was a multi-regime multivariate Brownian motion (BM) model (Table 1, Supplementary Tables 1 and 2, and Extended Data Fig. 3). BM is a stochastic model in which divergence accumulates linearly with time and is associated with genetic drift under a strictly microevolutionary interpretation or random selection under a macroevolutionary interpretation. For this model, the main parameter is the rate matrix Σ, which controls the traits' stochastic rate of evolution. Because the preferred model was a multi-regime one, our tree is subdivided into different 'regimes', which are parts of the tree with different model parameters (rates of evolution for BM). For molar occlusal areas, the best model had three main regimes (Fig. 3). The first regime covers most fossil groups (thus named 'ancestral regime'), including Plesiadapiformes, stem−Haplorhini, part of stem−Simiiformes and Tarsiidae. The second regime refers to Strepsirrhini, both crown and stem, and the third refers to crown Simiiformes (monkeys and apes, including humans). This three-regime model was also considered a better fit than any global model (microevolution inspired or not) for the morphospace defined by linear distances, even though it was not the best solution found (Supplementary Table 2). In both morphospaces, the ancestral regime accumulated more variance over time than any derived regime, suggesting a weaker constraint on the former (Extended Data Fig. 4). It may be tempting to assign interpretations that are either biological (for example, higher divergence rates after the Cretaceous−Palaeogene extinction event resulting

from ecological opportunity) or statistical in nature (for example, increased phylogenetic uncertainty of fossil placement resulting in upwardly biased rates[51]). However, we find that such patterns do not appear universally across morphospaces and are absent from the developmentally informed one, thus making any interpretation of these partitions premature.

By contrast, the investigation of evo-devo-inspired variables based on the ICM paints a strikingly different picture (Fig. 2c). Instead of favouring more complex and heterogeneous models, the ICM morphospace favours a single global Ornstein−Uhlenbeck (OU) process (Table 1). OU models, like BM, have a Σ that governs the stochastic rate of evolution. Differently from BM, however, OU variance does not scale linearly with time, as the evolving species are under the influence of a phenotypic attractor, θ, to which species converge with a rate governed by the parameter H. Under a macroevolutionary interpretation (1 in Fig. 1d), OU models any evolutionary process with a constraint, be that selective, developmental, genetic and so on. Under a strict microevolutionary interpretation (2 in Fig. 1d), Σ is considered the rate of evolution due to random drift, and θ and H govern the optimum and the shape of the adaptive landscape, respectively. To achieve this interpretation, our microevolutionary model assumes a Σ which is proportional to the additive genetic covariance matrix of the traits (G matrix; equations (1) and (2), and Fig. 1e), with a value within a range governed by empirical estimates of demography and the life history of Primates (Fig. 1f,g)[45,46]. This implies that, instead of optimizing
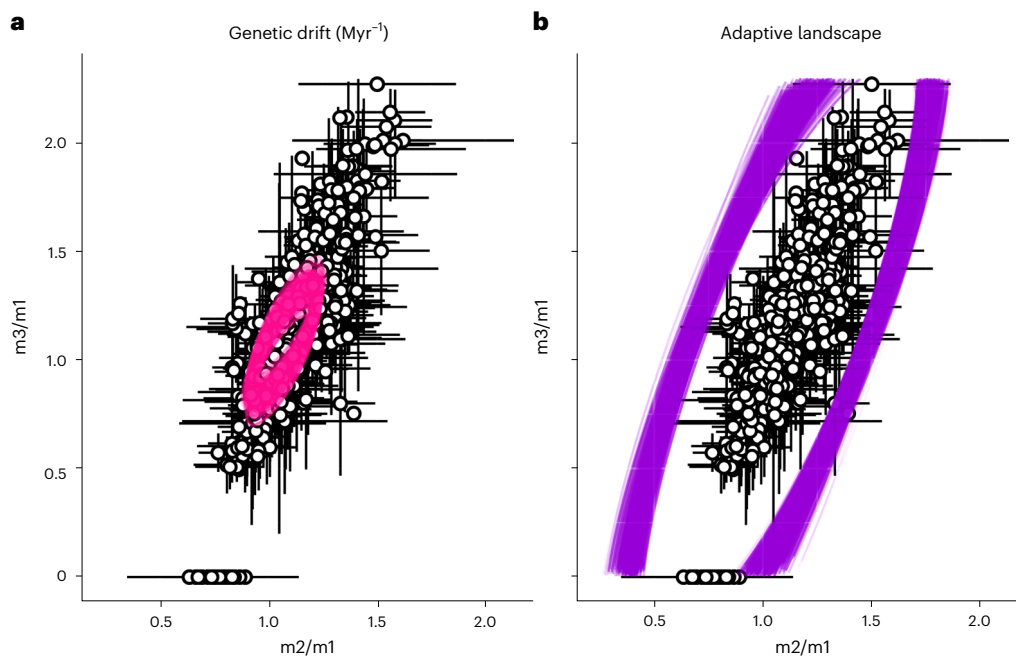
Fig. 4 | Graphical representation of the best selected model ($\mathrm{OU}^D_{\Sigma \propto G}$) for the ICM variables based on molar ratios m2/m1 and m3/m1. Dots are averages for the 480 species studied, and horizontal and vertical lines depict ±1.96 s.d. ($n = 6{,}142$ individuals). **a**,**b**, Ellipses are covariance matrices representing the following parameters of the best model: stochastic rate matrix Σ attributed to the amount of genetic drift introduced in the system every 1 Myr (**a**) and individual adaptive landscape Ω based on model estimates for the rate of adaptation towards the optima (H) (**b**). See Methods for explanations for these parameters. Multiple ellipses were calculated from parameter value combinations that are sampled along the multi-dimensional likelihood contour 2 log likelihoods away from the peak.

values for each entry of Σ, this model fits only one proportionality parameter (κ), which makes it simpler and more parsimonious than the macroevolutionary one.

Both macroevolutionary (1 in Fig. 1b) and microevolutionary (2 in Fig. 1b) versions of this model (OU and $\mathrm{OU}_{\Sigma \propto G}$, respectively) had essentially the same BICs, suggesting that their information content is effectively the same (Table 1). However, inspecting the confidence intervals for the macroevolutionary OU model reveals that the 95% intervals for its parameters overlap with the values implied by the microevolutionary model (Supplementary Table 3). This suggests that the $\mathrm{OU}_{\Sigma \propto G}$ model can be interpreted in terms of microevolutionary processes not only in terms of patterns but also in terms of the magnitude of variation. So, we choose the microevolutionary $\mathrm{OU}_{\Sigma \propto G}$ as our preferred model not only because it reports the best BIC but also because of its simplicity and biological interpretability.

Following the microevolutionary interpretation of our preferred model, the variation introduced by drift is aligned with the distribution of phenotypes on the ICM morphospace (Fig. 4a), suggesting that the similarity between intra- and interspecific patterns of trait variation[11] is consistent with drift. This is further reinforced by the investigation of node-specific rates of evolution, which shows a huge overlap with rates expected under genetic drift (Fig. 5). However, drift alone would generate more variation than the total observed disparity during the period in which Primates have evolved (Extended Data Fig. 5), suggesting that stabilizing selection played a crucial role in shaping the pattern of evolution in the group as well.

The investigation of the adaptive landscape implied by the best model shows that stabilizing selection is aligned with the interspecific distribution of phenotypes (Fig. 4b). An examination of the half-lives ($t_{1/2}$, the time necessary for a species to reach halfway between the ancestral state and the regime optimum) in different directions of this adaptive landscape shows that $t_{1/2}$ is higher along the activation–inhibition gradient direction of the ICM and lower in directions that would lead to deviations from the ICM (Extended Data Fig. 6). These results
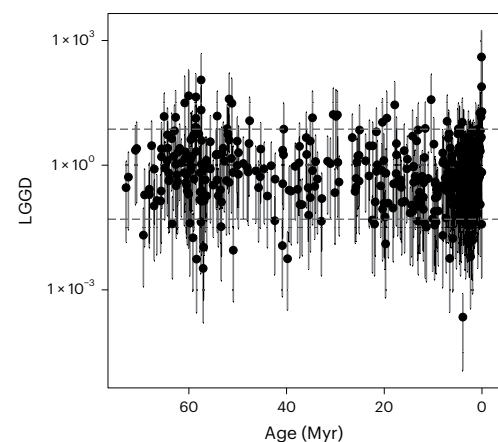

Fig. 5 | LGGD used to measure the node-specific rate of evolution throughout Primate divergence and diversification. Dashed lines represent the expected rates under genetic drift. For each of the 479 nodes, a distribution of values was calculated by integrating over the variation in heritability, effective population size and generation time. Dots represent the median values, and vertical lines are the 95% confidence intervals based on 10,000 replicates.

indicate that the macroevolution of primate molars is being shaped by a strong stabilizing selection against deviation from the ICM pattern while allowing evolution to occur along the activation–inhibition gradient, in a corridor-like manner.

Discussion

Previous work has usually highlighted that larger-scale morphological evolution tends to conform to the expectation of microevolutionary models qualitatively but rarely (if ever) in terms of magnitudes of change[6]. In other words, while macroevolution seems to follow directions with more genetic variation, as expected due to neutral

change[6,11–13,22,52], rates of evolution tend to fall below those expected under genetic drift[6–8,53]. This paradox has been used to argue for a fundamental mismatch between micro- and macroevolution, as simplistic quantitative genetics models seem unlikely to represent million-year evolutionary processes[7,54,55]. Here we constrained the proportionality parameter for our preferred microevolutionary-inspired model to be within realistic values for Primates (equation (2) and 2 in Fig. 1d). This constraint results in an estimated rate matrix compatible with drift around a stationary adaptive peak not only in patterns of trait association but also in magnitude. The key modelling choice that led to this conclusion was the quantification of developmentally informed traits (1 in Fig. 1b), which smoothed out transitions between micro-evolutionary and macroevolutionary data—defining and identifying a neutral subspace aligned with a conserved developmental process. The resulting morphospace of this modelling choice, the ICM space, lacks discontinuities along the diversity of primate molars (Fig. 2), which probably reduces the need for heterogeneous rates along the phylogenetic tree. Furthermore, by focusing on relative shape changes, which are governed by the balance of inhibitory and activation factors, this space limits the influence of other factors, such as static allometry, possibly allowing a closer match between genetic and macroevolutionary variation.

The differences we observe among morphometric representations might be partly due to how different spaces codify size variation. Both naive spaces contain size information, whereas the ICM variables do not. By using m1 size as a scaling factor, the ICM variables still include information regarding allometric variation (technically, they are unscaled versions of the Mosimann shape ratios[56]). Both area and distance spaces are log scaled, meaning they fit a power-law allometric model of variation[57]. Therefore, a higher heterogeneity in size variation in the naive spaces might favour more complex models, while the same is not true for ICM variables. While this suggests size correction can smooth out much of the heterogeneity in this case, this appears to derive from the fact that using shape ratios can provide a means to quantify localized ontogenetic effects[58]. Nevertheless, without actual knowledge of developmental systems, it is hard to know beforehand that shape ratios will necessarily lead to better conformity between micro- and macroevolutionary scales. In fact, depending on the system, raw measurements and shape ratios might produce similar results[8]. Thus, studying a well-understood system such as molar development allows us to piece apart the possible role of ontogenetic models, helping us connect micro- to macroscales.

Previous work in Primates has suggested that some traits have evolved with rates consistent with those expected under drift[53,59], including some dental features[60,61]. These works have largely been focused on hominin species, which could bias interpretations regarding the dental evolution of the whole order. Our results partly agree with these results and extend this phenomenon to the group's origin (Fig. 5). Although at face value this suggests that drift guided over 70 Myr of dental evolution in Primates, our model-fitting approach tells otherwise. Within the microevolutionary-inspired models, the OU models outperformed the BM models (Table 1), suggesting a crucial role of stabilizing selection in shaping macroevolutionary patterns. Considering the amount of variation introduced by drift every million years (Fig. 4a), a purely neutral process would result in overdispersion of tip values and higher phylogenetic signals (Extended Data Fig. 5). Instead, the patterns of stabilizing selection seem to be essential in shaping the ICM pattern by both constraining variation that deviates from the ICM pattern and facilitating evolution along the activation–inhibition gradient (Extended Data Figs. 6 and 7).

Even though these two results might seem contradictory—rates of evolution consistent with drift and the best model including stabilizing selection—we foresee at least two possibilities of how they both might be true: one has to do with the topography of the inferred adaptive landscape and the other with the estimates of the evolutionary parameters.

Regarding the adaptive landscape, the shape of the landscape implied by the preferred model is almost corridor like (Fig. 4b). If this corridor is relatively smooth internally (no great selection differentials within its limits), this would mean that species are free to explore this landscape neutrally, within the bounds of the corridor. In addition, because the matrix of additive genetic covariances G is aligned with the corridor as well (Fig. 4a), this means that most neutral changes will happen in accordance with the landscape and will not result in great stabilizing selection. The other possibility is based on the precision of the rate-parameter estimates. Even though the OU model was the preferred one, estimated rates of evolution of the BM models are remarkably similar to the ones of the full model (Supplementary Tables 6–9). Considering that node-specific rates of evolution are calculated under the assumption of a BM model[62,63], this could mean that a dense fossil sample in a comprehensive phylogenetic framework might allow for a good estimation of rates of evolution, even under model violation. Irrespective of which is true (or even if both are), the observation that most evolutionary rates are compatible with drift is a pattern rarely seen for macroevolutionary data[6–8,53].

Together, these results point to the interplay of genetic variation, selection and development leading to a homogeneous macro-evolutionary process within a defined subspace. It has been argued that selection can mould genetic patterns of trait association and variation[29,64,65], specifically by altering developmental pathways and genetic interactions[66,67]. Conversely, development has also been argued to impose direct selective pressures (that is, internal selection) by reducing the viability of non-conforming phenotypes[29], which could, in turn, trickle down to the organization of genetic variation. While in the present case we can observe this triple alignment between genetics, ontogeny and selection, its origins are harder to decipher. The ICM was originally described in rodents and later verified in many other mammalian groups[37–40,68], suggesting that it is the ancestral condition for molar development in the group. In this case, ontogeny is viewed as the organizing factor behind both selective patterns and the organization of genetic variation[11]. Furthermore, this explains the near-neutral quality of primate dental evolution, as conformity to the developmental process would be the main selective pressure on relative tooth sizes[69]. However, some mammalian groups have been shown to deviate from the predictions of the ICM to different degrees, suggesting that the ontogenetic process itself could be malleable[37,38,68,70,71]. Indeed, it has been argued that molar tooth eruption timing in Primates is shaped by biomechanical demands at different ontogenetic stages[72], revealing a possible mechanism through which external selection could shape development and, indirectly, the morphology of the molar row.

Conclusions

To what degree microevolution can be extended to macroevolution is a central question in evolutionary biology[4]. While there is little doubt that the fundamental causes at both levels are the same (for example, selection, drift, mutation), efforts to model the connection have generally failed beyond the qualitative alignment of patterns. When it comes to morphological evolution, the consensus has been overwhelmingly to reject any straightforward connection between both levels, specifically because of the fact that empirical evolutionary rates are orders of magnitude inferior to the ones expected by genetic drift[6,7]. The results presented here reject this consensus, as we show that microevolutionary models can fit well into the data, as long as we choose the proper morphometric representation. Even the relatively simple task of characterizing the multivariate dimensions of three molars poses a large number of choices for measurement[10,35,73]. Our results suggest that phenotypic quantification based on evo-devo models maximally narrows the gap between both levels of analysis and allows for the discovery of the underlying subspaces that both qualitatively and quantitatively align macroevolutionary patterns with microevolutionary processes.

While primate molar seems unique in both the presence of a well-constrained ontogenetic model and abundance of data, other systems might also fit the requirements of the methods described here. The existence of evolutionary stable developmental pathways and modules suggests a long history of similarly stable selective pressures[25,66,74,75]. This makes developmental modules good systems to investigate adaptive landscapes in deep time[32,58,76]. Furthermore, assuming that these pathways are shaped by natural selection to optimize the generation of adaptive variation[64–66,74], they are a likely place to identify simple connections between micro- and macroevolutionary scales[32,58]. So, other evolutionary stable systems are the probable candidates to verify the connection between scales of organization. Good examples are modules built on serially homologous structures, such as limbs, phalanges and vertebrae[77–79]. For more complex structures formed by the interaction of multiple tissues, it might be harder to devise simple models that sufficiently describe the system ontogeny and variation. However, works that focus on the mammalian skull and used individualized bone measurements have had a good track record of modelling multivariate evolution of these structures under microevolutionary models[12,22,59,80,81], going even further than the simple alignment between variation and evolutionary rates[8,53,59]. Since vertebrate skull bones are elements with deep individualized history, measuring them individually might represent a good first approximation of the multiple morphogenic fields that interact to form the complete structure. This perspective contrasts with the regular practice of constructing morphospaces as comprehensive, phenomenological and statistical descriptors of biological form without a clear connection to underlying biological processes[19,36,82,83]. However, given that different morphometric methods seem to point to similar overall patterns of trait variation[84], finding the correct quantification protocol might be a matter of proper scaling of morphometric variables than a radical departure from classically established measurement practices.

Our investigation also provides a new framework in which developmental biology can be more fully incorporated into macroevolutionary modelling (Fig. 1). It has long been considered that developmental biology was left out of the evolutionary synthesis[26,27,58], and indeed, such data are rarely incorporated into comparative analyses. Recent efforts have had different degrees of success, with many pointing out how complexities of the ontogenetic systems can lead to core violations of the modern synthesis[26,30,31,58,85]. By reframing the question of microevolutionary model adequacy into a problem of quantification of biological phenomena[82,83], we show how evo-devo is essential for a fully unified view in the context of the evolutionary synthesis.

Methods

Sample and morphometrics
We used the standard mesiodistal length (MD) and buccolingual breadth (BL) as basic descriptors of each molar. MD and BL were obtained for each tooth of the lower molar row (m1, m2 and m3). We obtained raw measurements from available datasets in the literature ($n = 6,142$) and from newly measured museum specimens using a caliper ($n = 150$). For rare species, we took measurements from images that were either published or provided to us ($n = 266$). All photos used had a scale and were digitized using the Fiji software[86]. Only adult and not heavily worn teeth were used in our sample, and each specimen was measured once. See Supplementary Information for a full list of data source references. In total, we compiled a sample of 6,558 individuals distributed among 480 species, divided between 232 extant and 248 extinct species. To evaluate the evolution of these traits on a phylogenetic framework, we used the most comprehensive phylogeny available that included both living and fossil primate species[50]. Our sample covered all genera and 52.98 % of the species diversity included in ref. 50, spanning the full 75 Myr of the group's evolution.

We constructed three distinct morphospaces to quantify molar variation (Extended Data Fig. 1). For our biologically naive

representation of tooth form, we used a 'distance space' based on linear distances obtained from each tooth and an 'area space' based on each tooth's occlusal area. The occlusal molar area was approximated using a crown index (BL × MD)[38,68]. Both areas and distances were log transformed to normalize the data and reduce the effect of large-sized outliers. For our evo-devo-informed space, we used the ratios of areas of the second and third molars in relation to the third (m2/m1 and m3/m1, respectively), as defined by the ICM of molar development[35]. We call this last morphospace the 'ICM space'. On each morphospace, we calculated species averages for comparative analyses. Measurement error was accounted by calculating the standard error of each measurement for each species. When a species had a sample of $n = 1$, we assigned a standard error equal to the pooled within-group standard deviation calculated for all species with sample sizes larger than 30. This implies a very high measurement error for species known from single specimens, such as the case of many fossils. The degree of genetic association between traits was approximated both by the intraspecific pooled phenotypic covariance matrix P and by an independently derived additive genetic covariance matrix G obtained from a pedigreed *Papio hamadryas* baboon population[41,42]. Because G for this population was originally estimated for MD and BL linear distance, we produced a Monte Carlo approximation for lower dimensionalities (Supplementary Information and Extended Data Fig. 8).

To evaluate morphospace patchiness, we performed a clustering based on parameterized finite Gaussian mixture models (GMM)[87]. This method tests for a series of nested models, in which groups are modelled as belonging to different multivariate normal distributions with different group averages. Models differ in the treatment of covariance structures. For example, the covariance matrix of different groups might differ in their volume (trace), shape (proportion of eigenvalues) or orientation (direction of eigenvectors). Furthermore, covariance matrices might be either spherical (zero covariances, equal variances), diagonal (zero covariances, different variances) or ellipsoidal (non-zero covariances). In total, the method tests 14 different covariance models and finds the best partition of the data and the best covariance models according to the BIC.

Phylogenetic comparative methods
To model morphological evolution, we used a maximum-likelihood model-selection approach, which fits different BM and OU models under a mixed Gaussian phylogenetic model (MGPM) framework implemented under the R packages PCMbase and PCMfit[88]. This method shares some similarities with the GMM clustering method used above to measure morphospace patchiness. Both GMM and MGPM model the data and allow different groups to have different parameter values. However, while GMM fits the data to normal distributions in a non-phylogenetic context, MGPM fits the data according to evolutionary models along a phylogenetic tree. In other words, while the GMM is a non-phylogenetic clustering method based on species phenotypic proximity in the morphospaces, MGPM groups species according to shared evolutionary models and phylogenetic history.

Under the MGPM framework, the evolution of a p-dimensional multivariate trait is modelled as an OU process as follows:

$$d\mathbf{x}(t) = -H(\mathbf{x}(t) - \mathbf{\theta}(t))dt + \Sigma_x dW(t) \qquad (1)$$

where H is the $p \times p$ selective rate matrix, $\mathbf{x}(t)$ is a p vector of trait values at time t, $\theta(t)$ is a p vector of trait evolutionary optima at time t, Σ_x is the Cholesky factor of the $p \times p$ stochastic rate matrix Σ (sometimes called evolutionary rate matrix) and $W(t)$ denotes the p-dimensional standard Wiener process.

Under a strict quantitative genetics interpretation[89], the diagonal of H contains the rate of adaptation to the optima of each trait (α_p) and the off diagonal measures the shape of co-selection among traits. Conversely, the diagonal of Σ contains the rate of evolution due to drift,

with its off-diagonal elements containing the amount of coevolution due to genetic covariation. If H is a matrix of zeros, the model collapses into a multivariate BM model.

Under this microevolutionary perspective, Σ is not an entirely free parameter. Instead, if Σ is the genetic drift parameter, then it has to be proportional to the additive genetic covariance matrix G of those traits[90], as follows:

$$\Sigma = G\frac{t_{\mathrm{g}}}{N_{\mathrm{e}}} \qquad (2)$$

where t_{g} is the time in generations and N_{e} is the effective population size. Because t_{g} and N_{e}, and even the size of G, are hard to estimate at evolutionary timescales, some have argued for treating $t_{\mathrm{g}}/N_{\mathrm{e}}$ as a nuisance parameter, reducing the investigation of drift at the macroevolutionary scale to a simple evaluation of the proportionality between Σ and G[12,22,54]. Consistent with these suggestions, here we implement a series of proportionality models, or κ models[11], in which Σ is set to be equal to a target matrix times a scaling factor κ. We used both the intraspecific pooled phenotypic covariance matrix P and G as target matrices. These models are implemented in the package PCMkappa (https://github.com/MachadoFA/PCMkappa).

Because the proportionality models are more tightly connected to a microevolutionary interpretation of the OU model, we call them 'microevolutionary models'. Full models (models in which all parameters are estimated freely) are called 'macroevolutionary models' because they do not have explicit microevolutionary assumptions.

We fitted two macroevolutionary BM and OU models and two microevolutionary models, using either P or G as a target matrix, for both BM and OU, totalling six global models (full BM and OU, $\mathrm{BM}_{\Sigma\propto P}$, $\mathrm{BM}_{\Sigma\propto G}$, $\mathrm{OU}_{\Sigma\propto P}$, $\mathrm{OU}_{\Sigma\propto G}$) for each morphospace. For the OU models, we investigated the confidence intervals of the parameters (Supplementary Tables 6–9) to evaluate if the model could be further reduced. Specifically, if the confidence interval of the off-diagonal elements of H overlapped with 0, another model was fit, setting H to be a diagonal matrix[88].

In addition, we performed an MGPM search for the combination of regimes, models and model parameters that best fit the data[88]. For both the mixed-model search and model comparison, we used the BIC, which minimizes parameter inflation due to large samples and is most appropriate for our model-selection question[91] (that is, asymptotically identifying the data-generating process as opposed to minimizing trait prediction error). For the mixed Gaussian models, we only fit full BM and OU models, and no κ model due to software restrictions. Therefore, the mixed models are also considered macroevolutionary models. All searches were conducted setting the minimum clade size to five species.

To ensure that the κ models were compatible with microevolutionary processes, we constrained the κ parameter to be within the range of expected values under drift, as expressed in equation (2). Because Σ is given in the tree (Myr) scale, we found approximations for t_{g} and of N_{e} for Primates to infer the expected scaling factor κ. t_{g} was estimated as $t_{\mathrm{g}} = 1\,\mathrm{Myr}\,g_{\mathrm{t}}^{-1}$ where g_{t} is the average generation time in years obtained from ref. 45. Because we lack good estimates of g_{t} for fossil species, we used the phylogenetic average ± s.d. throughout the phylogeny. This was done by trimming the dataset to only the species with g_{t} data and obtaining the ancestral value and standard deviation at the base through maximum likelihood[92]. For N_{e}, we used 20,000–1,000,000 as the range of possible values consistent with the genomics estimates for multiple primate species and hypothetical common ancestors[46]. While g_{t} and N_{e} are expected to vary over the tree, we assumed that the effect of this variation would be at least partially cancelled out by the fact that these two quantities are generally inversely related to each other.

To evaluate the fitted model mechanistically under quantitative genetics theory, we generalized the equation for the adaptive landscape[89,93] to the multivariate case as

$$\Omega = H^{-1/2}GH^{-1/2} - P \qquad (3)$$

Rates of evolution

Rates of evolution were used to evaluate whether the evolutionary change conforms to the expectation of genetic drift. To calculate rates of evolution, we employed Lande's generalized genetic distance (LGGD[90])

$$\mathrm{LGGD} = \frac{N_{\mathrm{e}}}{t_{\mathrm{g}}}\Delta z^{t}\mathbf{G}^{-1}\Delta z \qquad (4)$$

where Δz is the phenotypic divergence calculated as the time-standardized phylogentic independent contrasts for each node[8,62]. We produced a distribution of 10,000 values for each node by sampling values of G, N_{e} and t_{g} from a uniform distribution in the range defined above. Confidence intervals for the null hypothesis of drift were generated from simulations based on equation (2)[8]. Values that fall within the bounds of the null distribution are thought to conform to the expectation under genetic drift. Values that fall above or below are thought to be indicative of directional or stabilizing selection, respectively.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All of the data analysed during this study are included in the Supplementary Data.

Code availability

All of the code used for this paper is available at https://github.com/MachadoFA/PCMkappa and https://github.com/MachadoFA/PrimateTeethProject.

References

1. Dobzhansky, T. *Genetics and the Origin of Species* 11 (Columbia Univ. Press, 1982).
2. de Santis, M. D. Misconceptions about historical sciences in evolutionary biology. *Evol. Biol.* **48**, 94–99 (2021).
3. Stanley, S. M. A theory of evolution above the species level. *Proc. Natl Acad. Sci. USA* **72**, 646–650 (1975).
4. Gould, S. J. *The Structure of Evolutionary Theory* (Harvard Univ. Press, 2002).
5. Uyeda, J. C., Hansen, T. F., Arnold, S. J. & Pienaar, J. The million-year wait for macroevolutionary bursts. *Proc. Natl Acad. Sci. USA* **108**, 15908–15913 (2011).
6. Houle, D., Bolstad, G. H., van der Linde, K. & Hansen, T. F. Mutation predicts 40 million years of fly wing evolution. *Nature* **548**, 447–450 (2017).
7. Lynch, M. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *Am. Nat.* **136**, 727–741 (1990).
8. Machado, F. A., Marroig, G. & Hubbe, A. The pre-eminent role of directional selection in generating extreme morphological change in glyptodonts (Cingulata; Xenarthra). *Proc. R. Soc. B* https://doi.org/10.1098/rspb.2021.2521 (2022).
9. Hlusko, L. J. Integrating the genotype and phenotype in hominid paleontology. *Proc. Natl Acad. Sci. USA* **101**, 2653–2657 (2004).

10. Hlusko, L. J., Schmitt, C. A., Monson, T. A., Brasil, M. F. & Mahaney, M. C. The integration of quantitative genetics, paleontology, and neontology reveals genetic underpinnings of primate dental evolution. *Proc. Natl Acad. Sci. USA* **113**, 9262–9267 (2016).

11. Mongle, C. et al. Developmental processes mediate alignment of the micro- and macroevolution of primate molars. *Evolution* **76**, 2975–2985 (2022).

12. Marroig, G. & Cheverud, J. M. Did natural selection or genetic drift produce the cranial diversification of Neotropical monkeys? *Am. Nat.* **163**, 417–428 (2004).

13. McGlothlin, J. W. et al. Adaptive radiation along a deeply conserved genetic line of least resistance in *Anolis* lizards. *Evol. Lett.* **2**, 310–322 (2018).

14. Erwin, D. H. Macroevolution is more than repeated rounds of microevolution. *Evol. Dev.* **2**, 78–84 (2000).

15. Hautmann, M. What is macroevolution? *Palaeontology* https://doi.org/10.1111/pala.12465 (2019).

16. Lande, R. The dynamics of peak shifts and the pattern of morphological evolution. *Paleobiology* **12**, 343–354 (1986).

17. Charlesworth, B., Lande, R. & Slatkin, M. A neo-Darwinian commentary on macroevolution. *Evolution* **36**, 474 (1982).

18. Arnold, S. J., Pfrender, M. E. & Jones, A. G. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* **112**, 9–32 (2001).

19. Polly, P. D. Developmental dynamics and G-matrices: can morphometric spaces be used to model phenotypic evolution? *Evol. Biol.* **35**, 83–96 (2008).

20. Hansen, T. F. Macroevolutionary quantitative genetics? A comment on Polly (2008). *Evol. Biol.* **35**, 182–185 (2008).

21. Melo, D., Porto, A., Cheverud, J. M. & Marroig, G. Modularity: genes, development and evolution. *Annu. Rev. Ecol. Evol. Syst.* **47**, 463–486 (2016).

22. Machado, F. A. Selection and constraints in the ecomorphological adaptive evolution of the skull of living Canidae (Carnivora, Mammalia). *Am. Nat.* **196**, 197–215 (2020).

23. Laland, K. N. et al. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc. R. Soc. B* **282**, 20151019 (2015).

24. Thompson, D. W. *On Growth and Form* Vol. 2 (Cambridge Univ. Press, 1942).

25. Waddington, C. H. *The Strategy of the Genes* (Routledge, 1957).

26. Alberch, P., Gould, S. J., Oster, G. F. & Wake, D. B. Size and shape in ontogeny and phylogeny. *Paleobiology* **5**, 296–317 (1979).

27. Alberch, P. Ontogenesis and morphological diversification. *Am. Zool.* **20**, 653–667 (1980).

28. Love, A. C. Evolutionary morphology, innovation, and the synthesis of evolutionary and developmental biology. *Biol. Phil.* **18**, 309–345 (2003).

29. Cheverud, J. M. Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.* **110**, 155–171 (1984).

30. Salazar-Ciudad, I. & Jernvall, J. A computational model of teeth and the developmental origins of morphological variation. *Nature* **464**, 583–586 (2010).

31. Linde-Medina, M. & Diogo, R. Do correlation patterns reflect the role of development in morphological evolution? *Evol. Biol.* **41**, 494–502 (2014).

32. Hether, T. D. & Hohenlohe, P. A. Genetic regulatory network motifs constrain adaptation through curvature in the landscape of mutational (co)variance. *Evolution* **68**, 950–964 (2014).

33. Pearson, K. & Davin, A. G. On the biometric constants of the human skull. *Biometrika* **16**, 328 (1924).

34. Raup, D. M. Geometric analysis of shell coiling: general problems. *J. Paleontol.* **40**, 1178–1190 (1966).

35. Kavanagh, K. D., Evans, A. R. & Jernvall, J. Predicting evolutionary patterns of mammalian teeth from development. *Nature* **449**, 427–432 (2007).

36. Alba, V., Carthew, J. E., Carthew, R. W. & Mani, M. Global constraints within the developmental program of the *Drosophila* wing. *eLife* https://doi.org/10.7554/eLife.66750 (2021).

37. Asahara, M. Unique inhibitory cascade pattern of molars in canids contributing to their potential to evolutionary plasticity of diet. *Ecol. Evol.* **3**, 278–285 (2013).

38. Halliday, T. J. D. & Goswami, A. Testing the inhibitory cascade model in Mesozoic and Cenozoic mammaliaforms. *BMC Evol. Biol.* **13**, 79 (2013).

39. Bernal, V., Gonzalez, P. & Perez, S. I. Developmental processes, evolvability, and dental diversification of New World monkeys. *Evol. Biol.* **40**, 532–541 (2013).

40. Carter, C. & Worthington, S. The evolution of anthropoid molar proportions. *BMC Evol. Biol.* **16**, 18 (2016).

41. Hlusko, L. J., Maas, M.-L. & Mahaney, M. C. Statistical genetics of molar cusp patterning in pedigreed baboons: implications for primate dental development and evolution. *J. Exp. Zool. B* **302**, 268–283 (2004).

42. Hlusko, L. J., Sage, R. D. & Mahaney, M. C. Modularity in the mammalian dentition: mice and monkeys share a common dental genetic architecture. *J. Exp. Zool. B* **316**, 21–49 (2011).

43. Hardin, A. M. Genetic correlations in the dental dimensions of *Saguinus fuscicollis*. *Am. J. Phys. Anthropol.* **169**, 557–566 (2019).

44. Hardin, A. M. Genetic correlations in the rhesus macaque dentition. *J. Hum. Evol.* **148**, 102873 (2020).

45. Pacifici, M. et al. Generation length for mammals. *Nat. Conserv.* **5**, 89 (2013).

46. Brevet, M. & Lartillot, N. Reconstructing the history of variation in effective population size along phylogenies. *Genome Biol. Evol.* https://doi.org/10.1093/gbe/evab150 (2021).

47. Plavcan, J. M. *Sexual Dimorphism in the Dentition of Extant Anthropoid Primates*. PhD thesis, Duke Univ. (1990).

48. Godfrey, L. R., Samonds, K. E., Jungers, W. L. & Sutherland, M. R. Teeth, brains, and primate life histories. *Am. J. Phys. Anthropol.* **114**, 192–214 (2001).

49. Delson, E., Harcourt-Smith, W. E., Frost, S. R. & Norris, C. A. Databases, data access, and data sharing in paleoanthropology: first steps. *Evol. Anthropol.* **16**, 161–163 (2007).

50. Wisniewski, A. L., Lloyd, G. T. & Slater, G. J. Extant species fail to estimate ancestral geographical ranges at older nodes in primate phylogeny. *Proc. R. Soc. B* **289**, 20212535 (2022).

51. Liam, L. J., Harmon, L. J. & Collar, D. C. Phylogenetic signal, evolutionary process, and rate. *Syst. Biol.* **57**, 591–601 (2008).

52. Schluter, D. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766 (1996).

53. Schroeder, L. & von Cramon-Taubadel, N. The evolution of hominoid cranial diversity: a quantitative genetic approach. *Evolution* **71**, 2634–2649 (2017).

54. Ackermann, R. R. & Cheverud, J. M. Discerning evolutionary processes in patterns of tamarin (genus *Saguinus*) craniofacial variation. *Am. J. Phys. Anthropol.* **117**, 260–271 (2002).

55. Hansen, T. F. Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51**, 1341 (1997).

56. Mosimann, J. E. Size allometry: size and shape variables with characterizations of the lognormal and generalized gamma distributions. *J. Am. Stat. Assoc.* **65**, 930 (1970).

57. Jolicoeur, P. The multivariate generalization of the allometry equation. *Biometrics* **19**, 497–499 (1963).

58. Gould, S. J. *Ontogeny and Phylogeny* (Belknap Press, 1977).

59. Weaver, T. D., Roseman, C. C. & Stringer, C. B. Close correspondence between quantitative- and molecular-genetic divergence times for Neandertals and modern humans. *Proc. Natl Acad. Sci. USA* **105**, 4645–4649 (2008).

60. Gómez-Robles, A. Dental evolutionary rates and its implications for the Neanderthal–modern human divergence. *Sci. Adv.* **5**, eaaw1268 (2019).

61. Monson, T. A., Fecker, D. & Scherrer, M. Neutral evolution of human enamel-dentine junction morphology. *Proc. Natl Acad. Sci. USA* **117**, 26183–26189 (2020).

62. Felsenstein, J. Phylogenies and quantitative characters. *Annu. Rev. Ecol. Syst.* **19**, 445–471 (1988).

63. Felsenstein, J. Phylogenies and comparative method. *Am. Nat.* **125**, 1–15 (1985).

64. Melo, D. & Marroig, G. Directional selection can drive the evolution of modularity in complex traits. *Proc. Natl Acad. Sci. USA* **112**, 470–475 (2015).

65. Jones, A. G., Bürger, R., Arnold, S. J., Hohenlohe, P. A. & Uyeda, J. C. The effects of stochastic and episodic movement of the optimum on the evolution of the G-matrix and the response of the trait mean to selection. *J. Evol. Biol.* **25**, 2210–2231 (2012).

66. Riedl, R. *Order in Living Organisms: A Systems Analysis of Evolution* (John Wiley & Sons, 1978).

67. Watson, R. A., Wagner, G. P., Pavlicev, M., Weinreich, D. M. & Mills, R. The evolution of phenotypic correlations and 'developmental memory'. *Evolution* **68**, 1124–1138 (2014).

68. Polly, P. D. Development with a bite. *Nature* **449**, 413–414 (2007).

69. Monson, T. A. et al. Evidence of strong stabilizing effects on the evolution of boreoeutherian (Mammalia) dental proportions. *Ecol. Evol.* https://doi.org/10.1002/ece3.5309 (2019).

70. Varela, L., Tambusso, P. S. & Fariña, R. A. Unexpected inhibitory cascade in the molariforms of sloths (Folivora, Xenarthra): a case study in xenarthrans honouring Gerhard Storch's open-mindedness. *Frühförderung interdisziplinär* **76**, 1–16 (2020).

71. Sadier, A. et al. Bat teeth illuminate the diversification of mammalian tooth classes. *Nat. Commun.* https://doi.org/10.1038/s41467-023-40158-4 (2023).

72. Glowacka, H. & Schwartz, G. T. A biomechanical perspective on molar emergence and primate life history. *Sci. Adv.* **7**, eabj0335 (2021).

73. Monson, T. A. et al. Keeping 21st century paleontology grounded: quantitative genetic analyses and ancestral state reconstruction re-emphasize the essentiality of fossils. *Biology* https://doi.org/10.3390/biology11081218 (2022).

74. Riedl, R. A systems-analytical approach to macro-evolutionary phenomena. *Q. Rev. Biol.* **52**, 351–370 (1977).

75. Houle, D. & Rossoni, D. M. Complexity, evolvability, and the process of adaptation. *Annu. Rev. Ecol. Evol. Syst.* **53**, 137–159 (2022).

76. Olson, E. C & Miller, R. L. *Morphological Integration* (Univ. Chicago Press, 1958).

77. Kavanagh, K. D. et al. Developmental bias in the evolution of phalanges. *Proc. Natl Acad. Sci. USA* **110**, 18190–18195 (2013).

78. Young, N. M., Wislow, B., Takkellapati, S. & Kavanagh, K. Shared rules of development predict patterns of evolution in vertebrate segmentation. *Nat. Commun.* **6**, 6690 (2015).

79. Carraco, G., Martins-Jesus, A. P. & Andrade, R. P. The vertebrate embryo clock: common players dancing to a different beat. *Front. Cell Dev. Biol.* **10**, 944016 (2022).

80. Schroeder, L., Roseman, C. C., Cheverud, J. M. & Ackermann, R. R. Characterizing the evolutionary path(s) to early *Homo*. *PLoS ONE* **9**, e114307 (2014).

81. Schroeder, L., Elton, S. & Ackermann, R. R. Skull variation in Afro-Eurasian monkeys results from both adaptive and non-adaptive evolutionary processes. *Sci. Rep.* **12**, 12516 (2022).

82. Bookstein, F. L. Measurement, explanation, and biology: lessons from a long century. *Biol. Theory* **4**, 6–20 (2009).

83. Houle, D., Pélabon, C., Wagner, G. P. & Hansen, T. F. Measurement and meaning in biology. *Q. Rev. Biol.* **86**, 3–34 (2011).

84. Machado, F. A., Hubbe, A., Melo, D., Porto, A. & Marroig, G. Measuring the magnitude of morphological integration: the effect of differences in morphometric representations and the inclusion of size. *Evolution* **73**, 2518–2528 (2019).

85. Mitteroecker, P. The developmental basis of variational modularity: insights from quantitative genetics, morphometrics, and developmental biology. *Evol. Biol.* **36**, 377–385 (2009).

86. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

87. Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *R J.* **8**, 289–317 (2016).

88. Mitov, V., Bartoszek, K. & Stadler, T. Automatic generation of evolutionary hypotheses using mixed Gaussian phylogenetic models. *Proc. Natl Acad. Sci. USA* **116**, 16921–16926 (2019).

89. Lande, R. Natural selection and random genetic drift in phenotypic evolution. *Evolution* **30**, 314 (1976).

90. Lande, R. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* **33**, 402–416 (1979).

91. Dennis, B., Ponciano, J. M., Taper, M. L. & Lele, S. R. Errors in statistical inference under model misspecification: evidence, hypothesis testing, and AIC. *Front. Ecol. Evol.* https://doi.org/10.3389/fevo.2019.00372 (2019).

92. Schluter, D., Price, T., Mooers, A. Ø. & Ludwig, D. Likelihood of ancestor states in adaptive radiation. *Evolution* **51**, 1699–1711 (1997).

93. Hansen, T. F. & Martins, E. P. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* **50**, 1404–1417 (1996).

Acknowledgements

Author contributions

F.A.M., C.S.M. and J.C.U. conceptualized the project. J.C.U. gathered the necessary funds. F.A.M., C.S.M., A.P., A.S. and V.D. gathered the dataset. A.W. and G.S. conducted the phylogenetic analysis. F.A.M. performed statistical analysis and produced the first draft. F.A.M., C.S.M., A.P., A.W., G.S and J.C.U. wrote the following versions of the draft. All authors approved the last draft.

Competing interests

The authors declare no competing interests.

Extended Data Fig. 1 | Variables used to construct morphospaces. The Distance-space was built on the mesiodistal length (MD, vertical) and buccolingual breadth (BL, horizontal) taken from each molar. The Area-space was built by estimating the occlusal areas of each molar as the A=MDxBL. The ICM-space was built by calculating the relative area size for m2 and m3 in relation to m1.
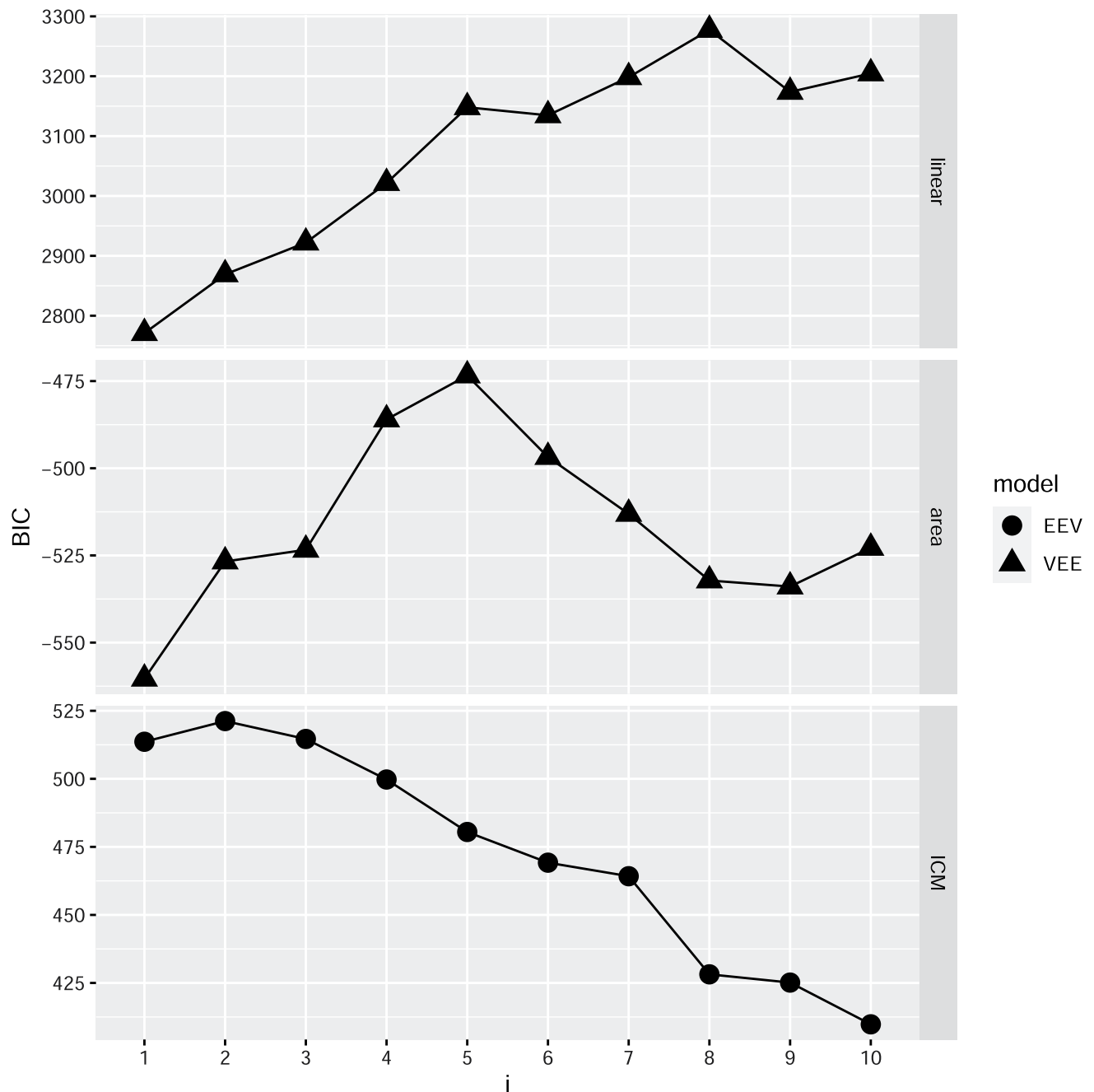
**Extended Data Fig. 2 | Bayesian Information Criterion results for the Gaussian mixture models for different morphospaces.** Negative BIC values for the Gaussian mixture models for the linear distances, areas and ICM spaces for different numbers of clusters (i). EEV- Elipsoidal model with the same shape, same volume and different orientations VEE- Elipsoidal model with the same shape, different volumes and same orientation. Higher values of negative BIC suggest the best model for each morphospace.
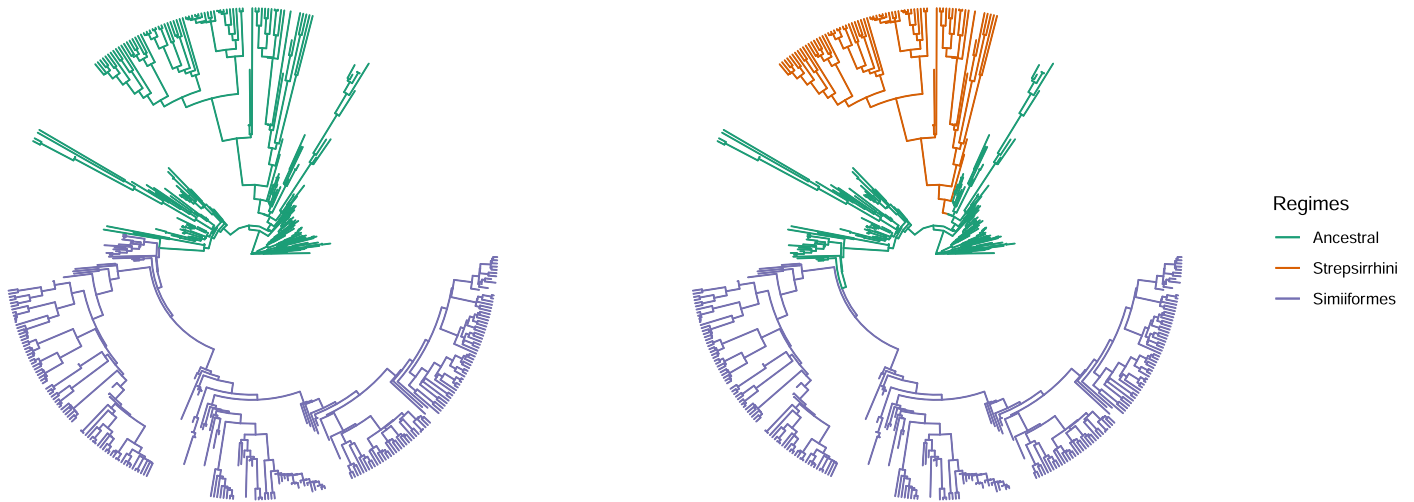
Extended Data Fig. 3 | Regimes for different runs of the heuristic search. Regimes for different runs of the heuristic search for the Distance morphospace. Left- Best model (Search 5). Right- Model compatible with the best model for areas (Fig. 3 on the main text).
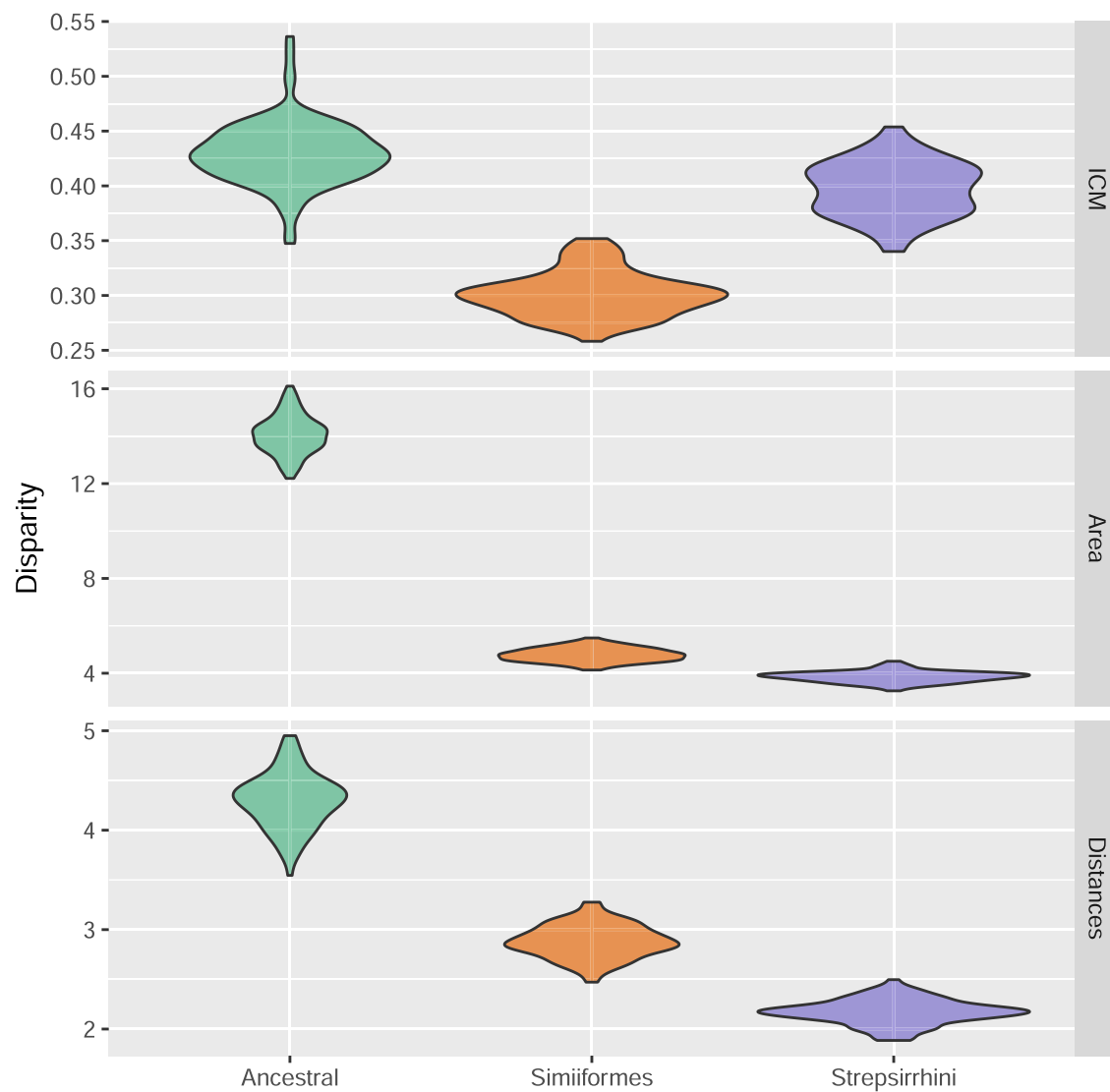
Extended Data Fig. 4 | Regime-specific disparities for Three-regime model on morphospace. Simulated regime-specific disparities for Three-regime model on morphospace. Regimes are described on the main text and illustrated on Fig. 3.
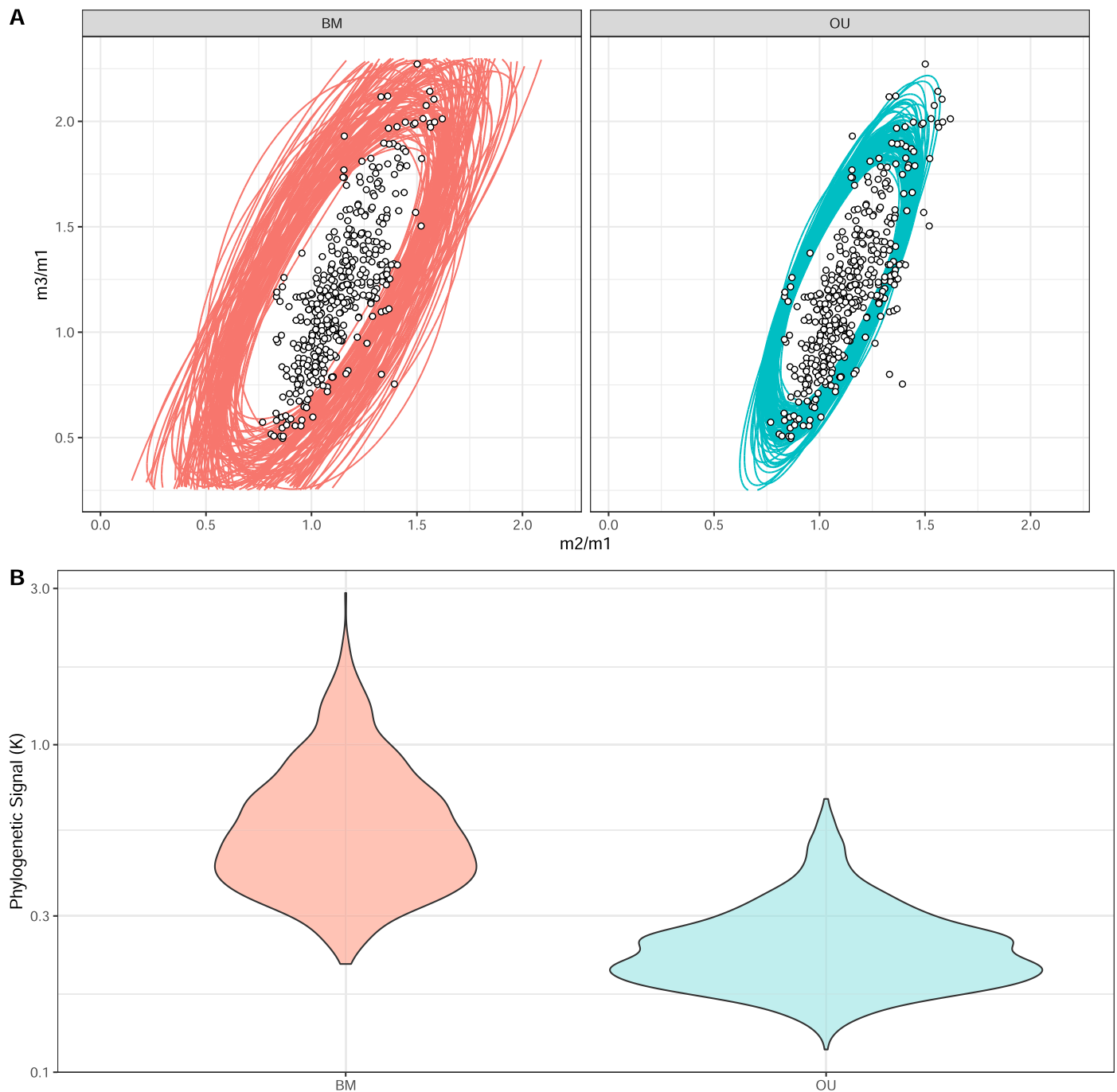
Extended Data Fig. 5 | Disparity and phylogenetic signal under BM and OU models. Simulated disparity (A) and phylogenetic signal (B) assuming the best model (OU) or a brownian motion (BM) model with the same rate parameters as the best OU model. Ellipses represent the covariance matrix of the simulated tip values, and thus do not represent any evolutionary parameter (for example Sigma, H, Omega, stationary variance, etc), but the phenotypic distributions of tips. Dots are observed species averages for comparison.
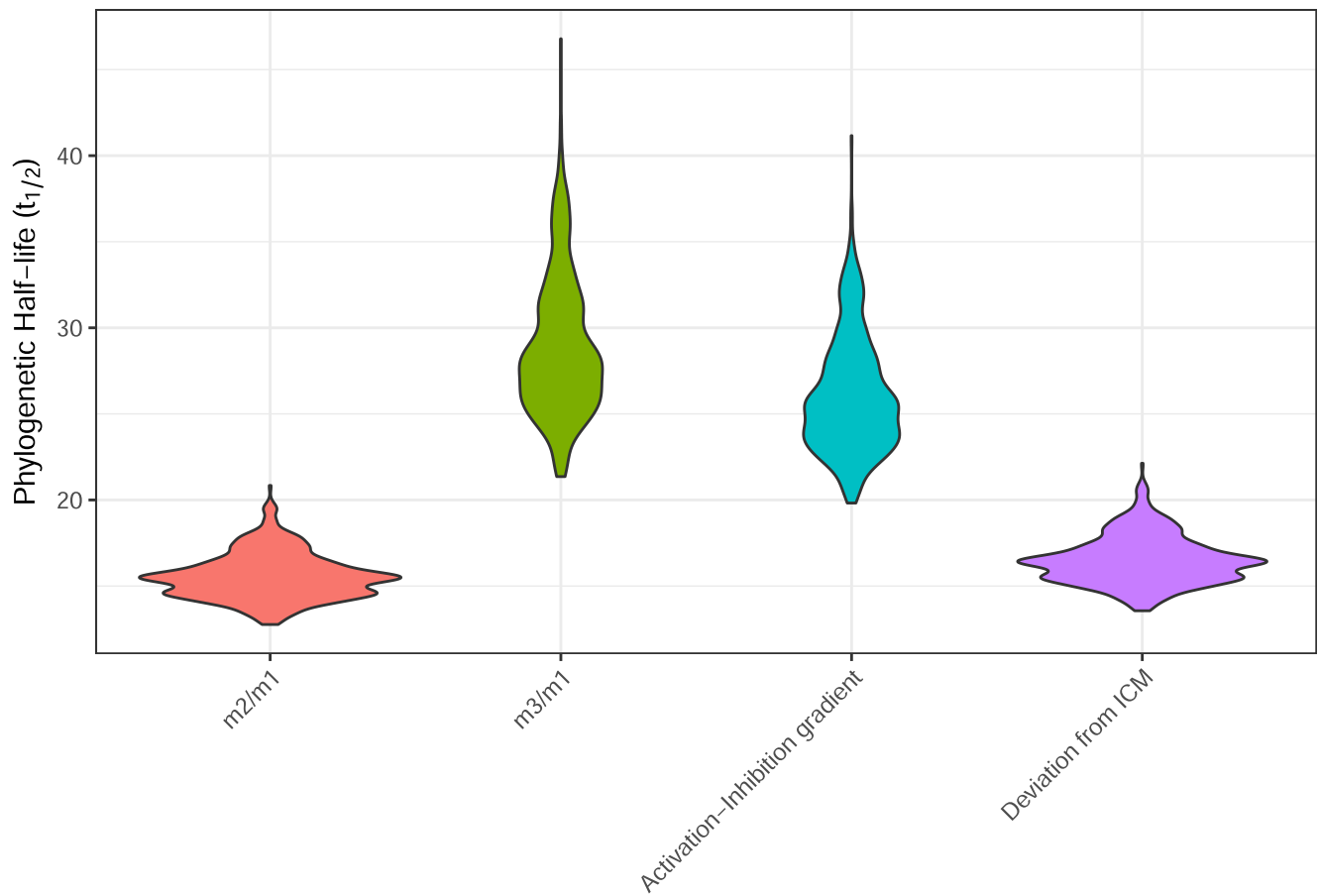
Extended Data Fig. 6 | Phylogenetic half-lifes. Phylogenetic half life for ICM ratios (m2/m1 and m3/m1) and components of the ICM model (Activation-Inhibition gradient and deviations from the ICM). Violin plots represent the distribution of values within the 95% confidence interval for the best model.
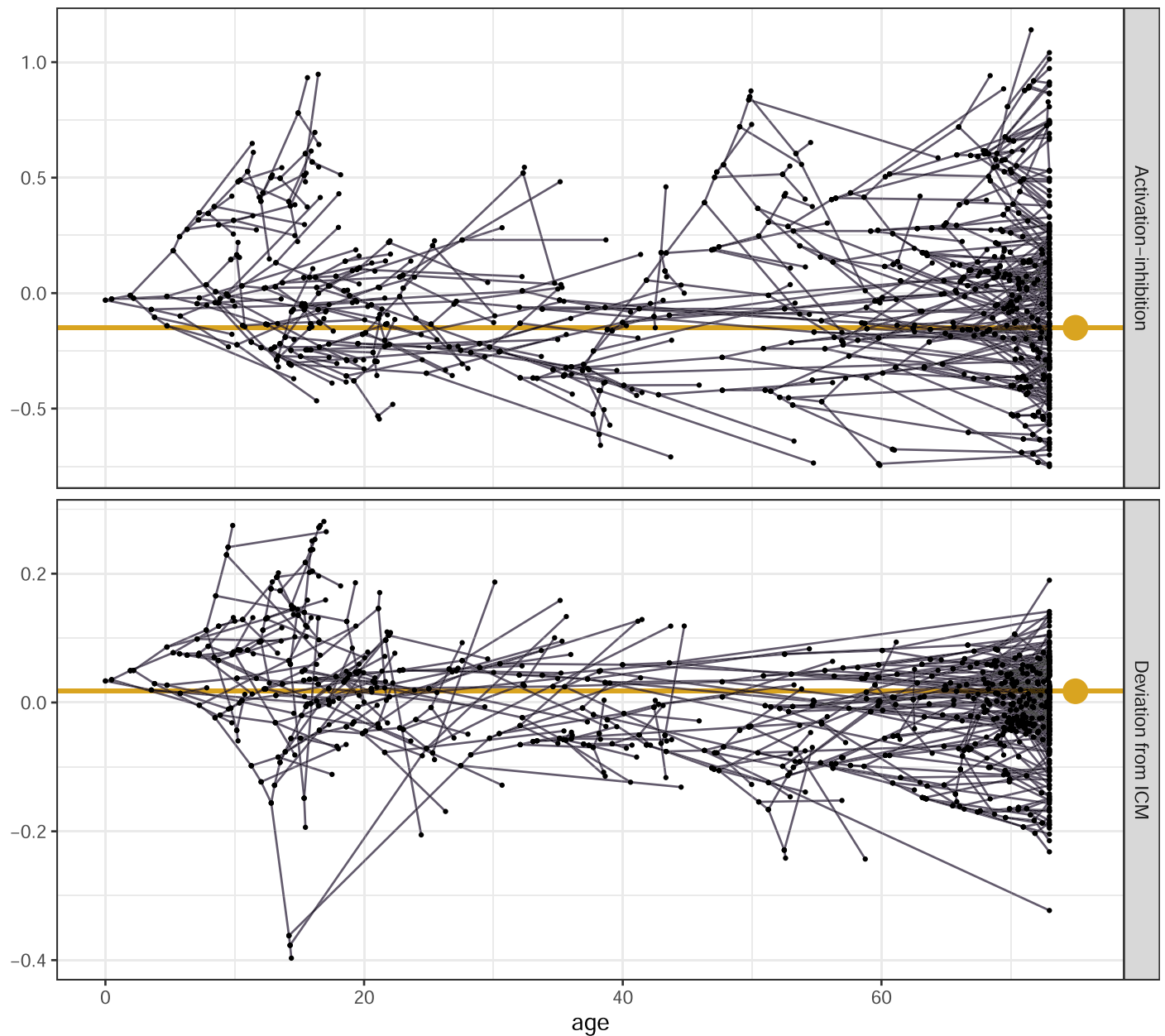
Extended Data Fig. 7 | Traitgrams of the components of the ICM for Primates. Black lines represent the phylogeny mapped to measured trait values (black points) and golden lines and golden dots represents each trait evolutionary optimum.
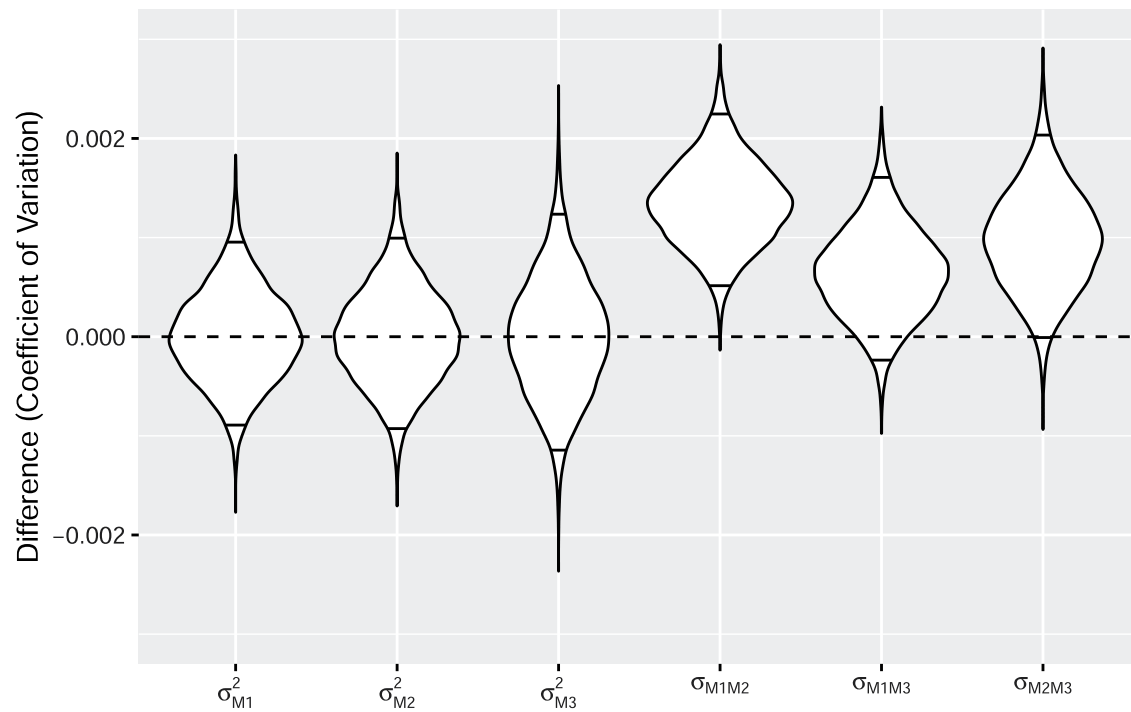
Extended Data Fig. 8 | Comparison between the Monte Carlo and analytical approaches to estimate covariances of areas. Differences between the Monte Carlo sampling approach for generating covariances for areas and the analytical approximation. Values are equal to the difference in coefficient of variation between matrices. Horizontal lines within violins highlights the 95% interval for each matrix cell. The first three entries are each area variance and the latter three are the areas covariances. The subscript indicates which trait (variances) or traits (covariances) are being compared.

Corresponding author(s): Fabio A. Machado

Last updated by author(s): May 2, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection
Data analysis	All analyses were run under the R programming environment (version 4.2.2). Custom codes are available at the corresponding author's github (https://github.com/MachadoFA)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data compiled and collected for this study will be made available in figshare (df.csv file).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	The study is an evolutionary model fitting investigation using phenotypic averages and errors. A total sample size of 6558 specimens distributed among 480 species.
Research sample	Species spanning both species of Primates, divided between 232 extant and 248 extinct species. Species where chosen to maximize taxonomic coverage.
Sampling strategy	Samples focused on species with known phylogenetic affinities, and where then sampled as comprehensively as possible to access intraspecific sampling errors.
Data collection	Data was primarely gathered from the literature when available, extracted from published or newly taken photographs, or measured directly using callipers. Data compilation was done by VD. Calliper measurements where taken by CSM. Photos where taken by AP. Measurements on photos where made by AS. All data curation was done by FAM.
Timing and spatial scale	N/A
Data exclusions	No data was excluded from the analysis.
Reproducibility	Reproducibility was accessed by using sampling errors and running codes on multiple machines, multiple times.
Randomization	N/A
Blinding	N/A

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☐ ☒	Palaeontology and archaeology
☒ ☐	Animals and other organisms
☒ ☐	Clinical data
☒ ☐	Dual use research of concern
☒ ☐	Plants

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging

Palaeontology and Archaeology

Specimen provenance
Specimens originate from all over the world.

Specimen deposition
All paleontological data was obtained from the literature, either directly from measurement tables or through photographs. Their deposition information is on our list of references used in data gathering.

Dating methods
N/A

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight
Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.

Note that full information on the approval of the study protocol must also be provided in the manuscript.