nature portfolio

https://doi.org/10.1038/s41586-022-04963-z

Supplementary information

Inner ear biomechanics reveals a Late Triassic origin for mammalian endothermy

In the format provided by the authors and unedited

Supplementary Information

Supplementary Methods

Sampling rationale

Our sample captures taxonomic, phylogenetic, morphological, ecological, behavioural, and locomotor diversity, but was constrained by the availability of material. We focused on densely sampling mammals and birds, which we used as independent models for endotherms; we used a large lepidosaurian sample as a model for ectotherms. The remaining extant specimens covered the range of body temperatures in vertebrates, but they are not the focus of this study. Our fossil sample included several mammaliamorphs, where the transition to endothermy is thought to occur. The 2D dataset is less accurate than the 3D datasets but was only used to extend the analysis to lower body temperatures, in order to obtain a broader range of data to assess the relationship between temperature and the TMI. Predictions for the thermal regime of fossil synapsids are based solely on the 3D datasets.

Biomechanics of the Thermo-Motility Index

Introduction

The bilateral inner ears each have three membranous semicircular ducts, which are interconnected and filled with endolymph, a fluid with broadly water-like properties (Fig. 1). Each duct has an inflated part, the ampulla, which contains sensory hair cells atop a transverse ridge, the crista ampullaris. A diaphragm-like cupula fills the remaining ampullar space above the crista, embedding the cilia of the hair cells. When the head moves, the endolymph inside a duct lags behind head rotation because of inertia and presses onto the cupula (Fig. B1). The resulting deflection of the embedded cilia generates action potentials in the afferent ampullary nerve (Highstein et al. 2004), which are integrated with other sensory nerve signals at the level of the brain (Angelaki, & Cullen 2008). This provides the brain with a reference frame of motion (Berthoz 1991) that helps to control essential body functions, like motor coordination (Owen & Lee 1986, Berthoz 1991, Mergner et al. 1997), balance (Berthoz 1991, Highstein et al. 2004), navigation (Fitzpatrick et al. 2006, Valerio, S. & Taube 2016), gait (Allum et al. 1998, Creath, R. et al. 2008), and gaze (Berthoz, & Pozzo 1988, Bronstein 1988, Pozzo et al. 1990).

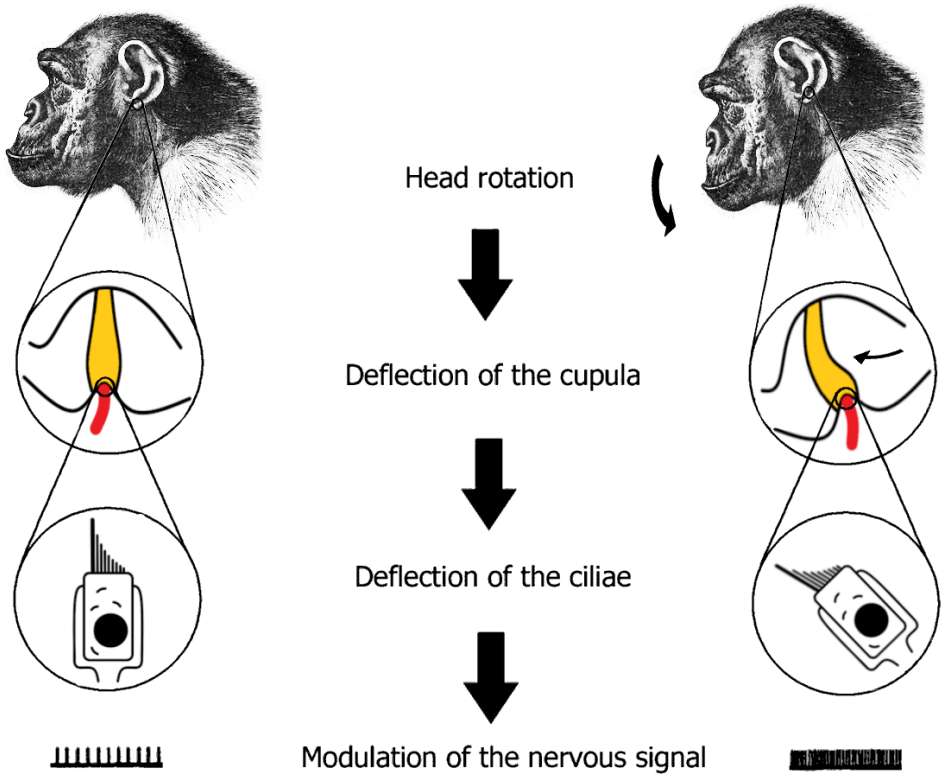


Figure B1: Basic principles of the mechanotransduction of head rotations by the semicircular ducts. The cupula is in yellow, whereas the afferent ampullary nerve bundle is in red.

Endolymph motion in the semicircular ducts is caused by angular accelerations of head motion. However, under natural conditions, the acceleration signal is instantly integrated, such that deflection of the cupula and embedded cilia is proportional to the angular velocity of the head. The semicircular

ducts are thus angular velocity transducers, which require angular accelerations to operate (Rabbitt et al. 2004). Integration of the acceleration signal occurs over a specific frequency bandwidth of head rotations, bracketed by lower and upper corner values. Between these frequencies the relationship between duct response and angular velocity (i.e., the gain or sensitivity) is constant, and the two are in phase with each other. These characteristics help the brain to interpret nervous signals and to reconstruct instantaneous angular velocity and head position (Wilson & Melvill Jones 1979). By contrast, outside the frequency bandwidth the sensitivity to angular velocity depends on the frequency of angular head motion, and the duct response is out of phase, thus degrading the sensory signal presented to the brain. Consequently, the frequency spectrum of naturally occurring angular head motion of an organism is expected to fall well within the frequency bandwidth of its semicircular ducts, where signal transduction is the most effective (Wilson & Melvill Jones 1979, Rabbitt et al. 2004).

A second key property of semicircular duct function concerns the detection range of angular velocities that can be perceived effectively, as limited by noise and saturation thresholds. If the angular velocity of head motion is too low, the signal at the level of the afferent ampullary nerve fibres will not differ sufficiently from noise levels (Wilson & Melvill Jones 1979). Similarly, when the angular velocity of head motion is too high, the resulting nerve signal will no longer be correlated linearly (Rüsch, & Thurm 1989). Complete saturation could even happen for particularly high angular velocities of the head, such that different angular velocities can no longer be differentiated (Rüsch, & Thurm 1989). As with the frequency bandwidth, we expect the distribution of angular velocities during naturally occurring angular head motion to fall within the detection range of its semicircular ducts.

The expected relationship between semicircular duct function and angular head motion for a given axis of rotation can be summarized graphically by plotting the angular velocity against the frequency of angular head motion (Fig. B2). The maximum and minimum frequencies, and maximum and minimum angular velocities that characterize natural angular head motion of an organism jointly form a rectangle with an area representing the full range of naturally occurring head motion (Fig. B2, red rectangle). Likewise, the full range of effective semicircular duct function can be represented by a rectangle combining the detection range, between the noise and saturation angular velocities, with the frequency bandwidth between the lower and upper corner frequencies (Fig B2, blue rectangle). The semicircular ducts are able to monitor angular head motion effectively as long as the head motion rectangle (red) is contained within the semicircular duct function rectangle (blue).

How closely the ranges of angular head motion and semicircular duct function match has consequences for the performance of the system. If frequencies and angular velocities of natural head motion fall easily within the range of effective semicircular duct function the risk of inaccurately detected head motion is low, but the system is underused in that more extreme motion that can could potentially be detected is unlikely to occur. On the other hand, a particularly close match between

motion and function ranges provides the best possible resolution to distinguish between two angular velocities. However, the risk here is that occasional, more extreme motion is not detected accurately.

Considering that semicircular duct function is essential for navigation, gaze stabilisation, and motor coordination, it seems likely that accurately capturing high angular head motion during the most active and intense locomotory behaviours is most important with respect to the selective fitness of an organism. In our analyses, therefore, we do not focus on the full ranges of angular head motion and semicircular duct function, but on their maxima. That is, the highest frequencies and velocities of head motion, and the upper corner frequencies and angular saturation velocities of duct function. The relationship between these maxima is quantified by tuning factors, as represented by arrows in Figure B2. Low tuning factors (long arrows in Fig. B2) indicate that maximum head motion falls well below the upper limit of duct function. High tuning factors (short arrows in Fig, B2) indicate a closer match. In reality, we expect that selection pressures will result in tuning factors that are around an optimal value; a compromise between broad efficacy and resolution.

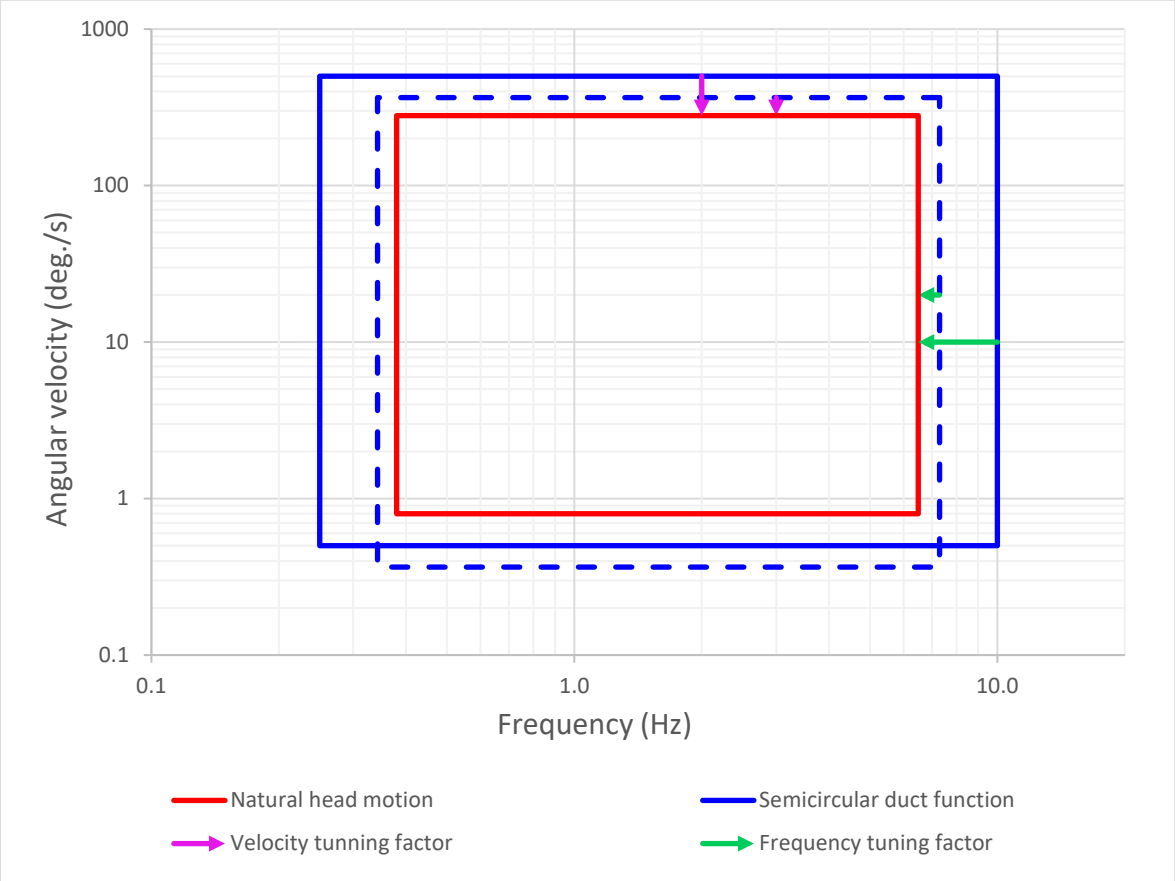


Figure B2: Ranges of semicircular duct function and head motion. Solid and dashed lines respectively correspond to the semicircular duct function of an organism whose body temperature increased from 24°C (average tetrapod ectotherm) to 39°C (average endotherm), assuming its inner ear morphology stayed the same. Tuning factors normally fall between]0;1] and on logarithmic scales they thus fall between]- ∞,0].

Now, as a thought experiment, let us consider an ancestral species with a body temperature of 24°C and a descendent species, similar in size and morphology, but with a body temperature of 39°C. Body temperature directly affects the viscosity of the endolymph, such that it decreases when body temperature increases (Ten Kate & Kuiper 1970, Oman 1981). As a biomechanical consequence the descendent species will show a smaller frequency bandwidth for which semicircular duct function is optimal and the range of detectable angular velocities will be displaced toward lower values, assuming that the duct morphology and all other characteristics of the endolymph remained the same (dashed lines, Fig. B2). If the two species showed similar head motion, these changes will thus result in increased tuning factors (shorter arrows in Fig. B2), with the possibility that the upper limits of angular head motion frequency or velocity fall outside the range of effective semicircular duct function, affecting the selective fitness of the descendant species.

To return to optimal values of the tuning factors (i.e., best trade-off between resolution and efficacy) the descendent species could adopt more sluggish behaviours, reducing its range of natural head motion. However, as increasing body temperatures tend to correlate with increased maximum speed and increased behavioural activity in general (Bennett & Ruben 1979, Garland & Albuquerque 2017, Hirt et al. 2017, Supplementary Data 1, Supplementary Note 2, Models 78-79/88-89), this would not be a plausible response. In fact, the increased body temperature is more likely to have the opposite effect: the need to expand the frequency range of effective semicircular duct function and to move the detection range toward higher angular velocities. To increase fitness and get to an optimal semicircular duct function, it is expected that the descendent species would either change its endolymph composition toward increased viscosity and/or modify aspects of its semicircular duct morphology to compensate for the viscosity decrease. These morphological changes would include reduced radii of curvature and reduced cross-sectional radii of the semicircular duct. The former response is seen in birds (Money 1971) and the latter in mammals (Jones & Spells 1963).

From this thought experiment, it appears that analysing the functional morphology of the semicircular ducts of extinct species may allow us to track changes in body temperature, and the acquisition of endothermy. To do this we developed the Thermo-Motility Index (TMI), which corresponds to a weighted average of semicircular duct function, derived from both the upper corner frequency and angular saturation velocity, as calculated from the duct morphology and endolymph composition, but excluding body temperature (see derivations below). In other words, semicircular duct function (matching behavioural head motion) is the combination of the TMI and body temperature. Thus, after an increase in body temperature, compensatory changes in the TMI are required to keep semicircular duct function constant. Providing information about both body temperature and motility of the head gives the Thermo-Motility Index its name.

In practice, different organisms exhibit different body sizes, body temperatures, and behaviours, all of which will affect their TMI. In order to extract as much information as possible about body temperatures, we will need to adjust the TMI for body size differences (see Supplementary Note 2, Models 26-31), and averaging between phylogenetically related species will help reduce the effects of behavioural variation (see Supplementary Note 2, Model 41). Applications to fossils are affected by an additional issue, which is the lack of information about their membranous labyrinths. Thus, we use statistical relationships between the dimensions of the bony and membranous labyrinths among extant species, along with information about the phylogenetic position of fossils, to obtain the best-possible estimate of fossil TMIs (see Supplementary Note 2, Models 8-19 and 20-25).

Mathematical derivations describing (1) the raw TMI, (2) relationships between dimensions of bony and membranous labyrinths, (3) body size corrections, and (4) the refined TMI are all described in the next sections, together with (5) a comparison of modelled and empirical data.

Mathematical derivations

1. Derivation of the raw Thermo-Motility Index

Description	Explanation
In this study, the upper limit of effective semicircular duct function of an organism $F_{SD\|2}$ (upper functional limit hereafter) is defined as a matrix containing in-plane saturating angular velocities $\dot{\Omega}_2$ (saturating velocities hereafter) and upper corner frequencies ω_2 of the anterior, posterior and lateral semicircular ducts (denoted by indices $_a$, $_p$ and $_l$ hereafter): $$F_{SD\|2} = \begin{bmatrix} \dot{\Omega}_2 \\ \omega_2 \end{bmatrix}$$ $$\dot{\Omega}_2 = \begin{bmatrix} \dot{\Omega}_{2a} \\ \dot{\Omega}_{2p} \\ \dot{\Omega}_{2l} \end{bmatrix} \quad \omega_2 = \begin{bmatrix} \omega_{2a} \\ \omega_{2p} \\ \omega_{2l} \end{bmatrix}$$ Note 1: Information about head rotation is transmitted to the brain through action potentials evoked at the level of ampullary nerve fibres. In this context, the in-plane saturating velocity of a semicircular duct corresponds to the lowest angular velocity of the head leading to saturation of corresponding ampullary nerve fibres, which prevents information to be accurately transmitted (Rüsch, & Thurm 1989). Note 2: Organisms with high in-plane saturation velocities are able to detect higher angular velocities of the head than organisms with low in-plane saturation velocities. Note 3: Saturating velocity is also inversely related to semicircular duct sensitivity (i.e., gain to angular velocities), such that ampullary nerve bundles of sensitive semicircular ducts with high sensitivity will saturate for lower angular velocities than ampullary nerve bundles of less sensitive ones. Note 4: Gain to angular velocity is expressed in degrees of cupula deflection, at the average height of the cilia layer, per degree per second of head rotation.	Here we define the upper limit of effective semicircular duct function, 'upper functional limit' in short, as the set of frequencies and angular velocities above which angular head motion is not optimally detected by the semicircular ducts (i.e., saturating velocities).
The upper functional limit $F_{SD\|2}$ is proportional to endolymph viscosity μ_{T_b}, which is itself affected by body temperature T_b through the formula (modified from Ten Kate and Kuiper 1970): $$\mu_{T_b} = \beta_{\mu_e} 10^{T_R}$$ $$T_R = \frac{247.8}{T_b - 140} \text{ (Equation 1)}$$	The upper functional limit is proportional to endolymph viscosity, which is itself inversely related to body temperature (through a temperature term) and depends on endolymph composition. When body temperature increases, endolymph viscosity and the upper functional limit are both expected to decrease, if endolymph composition is kept constant.

where:

- T_R : Dimensionless ratio (called temperature term hereafter) related to body temperature T_b (in Kelvins).
- β_{μ_e} : Factor linking endolymph viscosity to the temperature term (called physicochemical component of the Thermo-Motility Index hereafter). This factor only depends on the composition and concentration of solutes contained in the endolymph. Based on published evidence (Ten Doesschate 1914, Rauch 1964, Schnieder & Schindler 1964, Steer 1967, Ten Kate and Kuiper 1970, Money 1971, Macdonald & Wells 1991), we use 28.2 µPa.s for low-viscosity endolymph and 47.4 µPa.s for high-viscosity endolymph. For comparison, β_{μ_w} , the same factor for water, equals 24.1 µPa.s.

Note 5: The temperature term is defined between]140:+∞] in Kelvins.

Re-expressing the upper functional limit $F_{SD|2}$ to isolate the temperature term T_R , we obtain:

$$\left(F_{SD|2}\right)_{ij} = 10^{T_R}\left(TMI_{Raw}\right)_{ij} \text{(Equation 2)}$$

$$TMI_{Raw} = \begin{bmatrix} TMI_{\dot{\Omega}_2|Raw} \\ TMI_{\omega_2|Raw} \end{bmatrix} \quad TMI_{\dot{\Omega}_2|Raw} = \begin{bmatrix} TMI_{\dot{\Omega}_{2a}|Raw} \\ TMI_{\dot{\Omega}_{2p}|Raw} \\ TMI_{\dot{\Omega}_{2l}|Raw} \end{bmatrix} \quad TMI_{\omega_2|Raw} = \begin{bmatrix} TMI_{\omega_{2a}|Raw} \\ TMI_{\omega_{2p}|Raw} \\ TMI_{\omega_{2l}|Raw} \end{bmatrix}$$

where the raw Thermo-Motility Index TMI_{Raw} essentially corresponds to what the upper functional limit $F_{SD|2}$ would be if endolymph viscosity was not affected by body temperature (i.e., $\mu_{T_b} = \beta_{\mu_e}$).

Note 6: Indices $_{ij}$ denote element-wise operations. This means that we add/subtract/multiply/divide corresponding elements of matrices, not the matrices themselves.

Note 7: The upper limit of effective semicircular duct function is always higher than the raw TMI.

Note 8: The raw TMI can be seen as the lowest possible upper functional limit, given all non-temperature terms. It explicitly corresponds to what upper functional limit would be when body temperature tends toward infinity.

We show that by separating the temperature term from all other terms of the upper functional limit, we can define a basic version of the Thermo-Motility Index as the upper functional limit, freed from the effects of body temperature. In this context, the upper functional limit becomes directly proportional to this ‘raw’ TMI, through the temperature term. A direct consequence is that the temperature term and the ‘raw’ TMI have to change in inverse proportion to keep the upper functional limit constant.

The ‘raw’ TMI calculated from saturating velocities $TMI_{\dot{\Omega}_2|Raw}$ can be obtained using the formula for semicircular duct sensitivities G_V, excluding the temperature term T_R (modified from David et al. 2016):

$$\left(TMI_{\dot{\Omega}_2|Raw}\right)_{ij} = \frac{\beta_{\mu_e}\left(\lambda_\mu\right)_{ij}\left(L_S\right)_{ij}\left(\theta_2\right)_{ij}}{2\rho\left(\Lambda\right)_{ij}\left(E\right)_{ij}\left(a_S\right)_{ij}^{\ 2}}$$

$$\lambda_\mu = \begin{bmatrix} \lambda_{\mu_a} \\ \lambda_{\mu_p} \\ \lambda_{\mu_l} \end{bmatrix} \quad L_S = \begin{bmatrix} L_{S_a} \\ L_{S_p} \\ L_{S_l} \end{bmatrix} \quad \Lambda = \begin{bmatrix} \Lambda_a \\ \Lambda_p \\ \Lambda_l \end{bmatrix} \quad E = \begin{bmatrix} E_a \\ E_p \\ E_l \end{bmatrix} \quad a_S = \begin{bmatrix} a_{S_a} \\ a_{S_p} \\ a_{S_l} \end{bmatrix} \quad \theta_2 = \begin{bmatrix} \theta_{2a} \\ \theta_{2p} \\ \theta_{2l} \end{bmatrix}$$

where:

- ρ : Density of the endolymph (0.001 g.mm^{-3}, Rabbitt et al. 2004)
- λ_μ : Dimensionless velocity profile factor that depends on the shape of the velocity distribution over semicircular duct cross-section (Rabbitt et al. 2004, David et al. 2016; called wall shape factor hereinafter)
- L_S : Length of the centreline of the slender portion of the semicircular duct (called slender length hereafter)
- Λ : Maximum area enclosed by the centreline of the semicircular duct (called enclosed area hereafter)
- E : Transfer factor linking endolymph volume displacement to the average deflection of cilia bundles (called cilia deflection factor hereafter), considering a Poisson ratio of 0.48 and a Young modulus of 4.26 Pa for the cupula (David et al. 2016).
- a_S : Average cross-sectional area of the slender portion of the semicircular duct (called duct cross-section area hereafter)
- θ_2 : Minimal deflection of cilia bundles that leads to saturation of ampullary nerve fibres of corresponding semicircular ducts (called saturation angle hereafter)

Note 9: The cupula is considered to be clamped at the level of the crista ampullaris and supported elsewhere.

Note 10: The cilia deflection factor is proportional to the thickness of the cupula and inversely proportional to its transverse area.

Note 11: Saturation of ampullary nerve bundles occurs when cilia bundles located below the cupula reach a threshold (i.e., saturation angle). Hence, saturating velocities are directly proportional to saturation angles of cilia bundles. While likely showing interspecific variation, these saturation angles are expected to be small (i.e., ~4°; Rüsch, & Thurm 1989) and to be constrained by the size of tip-links connecting cilia of the same bundle (90–200nm; Sakaguchi et al. 2009, Figure B3).

Here we describe how to calculate the ‘raw’ TMI from saturating velocity. It is proportional to the length and the wall shape factor of the slender portion of the semicircular duct, as well as to the saturation angle of the cilia and endolymph composition. The ‘raw’ TMI calculated from saturating velocity is also inversely proportional to the area enclosed by the semicircular duct, the density of the endolymph, the cilia deflection factor and, in particular, the cross-sectional area of the slender portion of the semicircular duct. Importantly, body temperature is not used to obtain the ‘raw’ TMI calculated from saturating velocity.

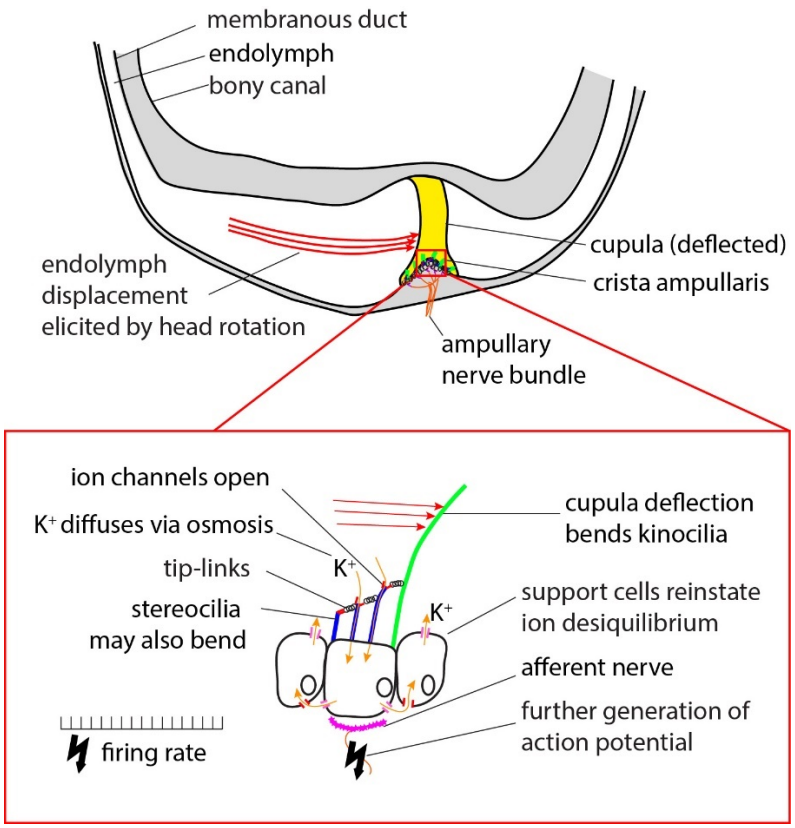


Figure B3: Detailed nanomechanics of the mechanotransduction of head rotations by the semicircular ducts.

Similarly, the 'raw' TMI calculated from upper corner frequencies $TMI_{\omega_2|Raw}$ can be obtained using the formula for upper corner frequencies ω_2 , excluding the temperature term T_R (modified from Rabbitt et al. 2004):

$$\left(TMI_{\omega_2|Raw}\right)_{ij} = \frac{\beta_{\mu_e}\left(\lambda_\mu\right)_{ij}}{2\pi\rho\left(a_S\right)_{ij}}$$

We now describe how to calculate the 'raw' TMI from upper corner frequency. It is proportional to the endolymph composition and to the wall shape factor of the slender portion of the semicircular duct. The 'raw' TMI calculated from upper corner frequency is also inversely proportional to endolymph density and to the cross-sectional area of the slender portion of the semicircular duct. Importantly, body temperature is not used to obtain the 'raw' TMI calculated from upper corner frequency.

Maximum naturally occurring angular head motion $_{\max}M_{Head}$ is expected to fall within angular velocities and ordinary frequencies that allow optimal efficacy of semicircular duct function, which are bracketed at their upper bound by the upper functional limit $F_{SD|2}$. Hence, following equation 2, we can write:

$$10^{T_R}\left(TMI_{Raw}\right)_{ij} \geq \left(_{\max}M_{Head}\right)_{ij}$$

$$10^{T_R}\left(TMI_{Raw}\right)_{ij} = \frac{\left(_{\max}M_{Head}\right)_{ij}}{\left(\tau_{F_{SD|2}}\right)_{ij}}$$

$$\left(TMI_{Raw}\right)_{ij} = \frac{\left(_{\max}M_{Head}\right)_{ij}}{10^{T_R}\left(\tau_{F_{SD|2}}\right)_{ij}} \text{ (Equation 3)}$$

$$_{\max}M_{Head} = \begin{bmatrix} _{\max}\dot{\Omega}_H \\ _{\max}f_H \end{bmatrix} \quad _{\max}\dot{\Omega}_H = \begin{bmatrix} _{\max}\dot{\Omega}_{Ha} \\ _{\max}\dot{\Omega}_{Hp} \\ _{\max}\dot{\Omega}_{Hl} \end{bmatrix} \quad _{\max}f_H = \begin{bmatrix} _{\max}f_{Ha} \\ _{\max}f_{Hp} \\ _{\max}f_{Hl} \end{bmatrix}$$

$$\tau_{F_{SD|2}} = \begin{bmatrix} \tau_{\dot{\Omega}_2} \\ \tau_{\omega_2} \end{bmatrix} \quad \tau_{\dot{\Omega}_2} = \begin{bmatrix} \tau_{\dot{\Omega}_{2a}} \\ \tau_{\dot{\Omega}_{2p}} \\ \tau_{\dot{\Omega}_{2l}} \end{bmatrix} \quad \tau_{\omega_2} = \begin{bmatrix} \tau_{\omega_{2a}} \\ \tau_{\omega_{2p}} \\ \tau_{\omega_{2l}} \end{bmatrix}$$

where

- $\tau_{F_{SD|2}}$: Matrix containing tuning factors, which are constants of proportionality defined between]0;1], indicating how well upper functional limit and maximum angular head motion are attuned to each other (called tuning factor matrix hereafter).
- $_{\max}\dot{\Omega}_H$: Highest angular velocities occurring during natural angular head motion.
- $_{\max}f_H$: Highest ordinary frequencies occurring during natural angular head motion.

Note 12: Highly tunned semicircular duct function is reflected by tuning factors that are close to 1, indicating that most of the frequency bandwidths and detection ranges of the semicircular ducts are overlapping with naturally occurring ordinary frequencies and angular velocities of head motion. In contrast, low tuning coefficients indicate loosely tuned semicircular duct function, optimally working for a large range of angular velocities and ordinary frequencies of head motion, most of which never occur under natural conditions.

Note 13: Tuning factors provide further information about the resolution of semicircular duct function in regard to angular velocity (i.e., values close to 1 indicate that the detection range of the semicircular ducts is put to better use and that angular velocities are better discriminated). However, it should be noted that this observation is relative, because absolute resolution is directly proportional to semicircular duct sensitivity (i.e., for the same tuning coefficients, organisms with high sensitivity will always discriminate smaller angular velocities differences than organisms with low sensitivity).

Maximum angular head motion, defined as the set of highest frequencies and angular velocities of the head that naturally occur in an organism, has to be lower than the upper functional limit for the semicircular ducts to be effective. To quantify how close maximum angular head motion is to the upper functional limit, we define tuning factors as the ratio between these two parameters (or subtraction in logged form). Consequently, we note that the 'raw' TMI is proportional to maximum angular head motion, and inversely proportional to the temperature term and the tuning factors. Note that we do not measure angular head motion in this study, but we need to define it to demonstrate that the TMI intrinsically correlates to body temperatures.

Note 14: Semicircular duct planes were chosen for practical reasons and average angular head motion could similarly be expressed relative to the pitch, roll and yaw axes of the head.

Note 15: Would head motion data be available, probability density functions of angular velocities of the head could be used to define maximal angular velocities, while power spectra of angular positions of the head relative to time could be used to define maximal ordinary frequencies.

If we consider endothermic and ectothermic organisms of similar body size and body plan, we can expect the average body temperature of the endotherms $\overline{T}_{b|Endo}$ to be higher than the average body temperature of the ectotherms $\overline{T}_{b|Ecto}$, such that:

$$\overline{T}_{b|Endo} > \overline{T}_{b|Ecto}$$

and following Equation 1:

$$10^{\overline{T}_{R|Endo}} < 10^{\overline{T}_{R|Ecto}}$$

In addition, based on available evidence (Bennett & Ruben 1979, Garland & Albuquerque 2017, Hirt et al. 2017, Supplementary Data 1, Supplementary Note 2, Models 78-79/88-89), we can also expect endotherms to be faster and more agile than ectotherms, or at least as fast and agile on average. This directly implies that average maximum angular head motion of endotherms $_{\max}\overline{M}_{Head|Endo}$ should be greater than or equal to average maximum angular head motion of ectotherms $_{\max}\overline{M}_{Head|Ecto}$, such that:

$$_{\max}\overline{M}_{Head|Endo} \geq {}_{\max}\overline{M}_{Head|Ecto}$$

Rewriting equation 3 as:

$$\left(\tau_{F_{SD|2}}\right)_{ij}\left(TMI_{Raw}\right)_{ij} = \frac{\left({}_{\max}M_{Head}\right)_{ij}}{10^{T_R}}$$

We can thus state:

$$\frac{\left({}_{\max}\overline{M}_{Head|Endo}\right)_{ij}}{10^{\overline{T}_{R|Endo}}} > \frac{\left({}_{\max}\overline{M}_{Head|Ecto}\right)_{ij}}{10^{\overline{T}_{R|Ecto}}}$$

$$\left\{\left(\overline{\tau}_{F_{SD|2}}\right)_{ij}\left(\overline{TMI}_{Raw}\right)_{ij}\right\}_{Endo} > \left\{\left(\overline{\tau}_{F_{SD|2}}\right)_{ij}\left(\overline{TMI}_{Raw}\right)_{ij}\right\}_{Ecto} \quad \text{(Equation 4)}$$

Equation 4 clearly indicates that a transition from ectothermy to endothermy, along an evolutionary trajectory, should either be marked by (1) an increase in the coefficients of the average tuning factor

Here we describe why the TMI is informative about body temperatures, and about the transition to endothermy.

We note that, because endotherms have higher body temperatures then ectotherms on average, their temperature term is by definition lower.

We further note that, based on available evidence (Bennett & Ruben 1979, Garland & Albuquerque 2017, Hirt et al. 2017, Supplementary Data 1 Supplementary Note 2, Models 78-79/88-89), endotherms are likely to show maximum angular head motion that are higher than, or equal to, ectotherms on average. In any case, it is unlikely that endotherms would show lower maximum angular head motion than ectotherms (i.e., would be less agile).

Here we note that both the ‘raw’ TMI and the tuning factors are proportionally related to a ratio between maximum angular head motion and the temperature term. This ratio is higher in endotherms than ectotherms on average. Although there may be outliers, this implies that the ‘raw’ TMI, the tuning factors, or both of them, are expected to be higher in endotherms than in ectotherms on average.

matrix of endotherm forms $\overline{\tau}_{F_{SD\|2}\|Endo}$, (2) an increase in the average 'raw' TMI of endotherm forms $\overline{TMI}_{Raw\|Endo}$, or (3) a combination of (1) and (2). Although there can be outliers, it should be noted that natural selection likely keeps tuning factors around optimal values, allowing enough discrimination of angular velocities of the head, while at the same time preventing unexpectedly high frequencies, or angular velocities, to escape accurate detection and transduction. Additionally, tuning factors have an upper bound (e.g., 1), preventing them to increase indefinitely. For all these reasons, it is in principle more likely that: $$\overline{\tau}_{F_{SD\|2}\|Endo} \approx \overline{\tau}_{F_{SD\|2}\|Ecto}$$ $$TMI_{Raw\|Endo} > TMI_{Raw\|Ecto}$$ (Equation 5)	However, because the tuning factors are bound between]0,1] and likely optimised around a fixed value, the TMI is probably the only parameter increasing in response to an increase of the ratio between maximum angular head motion and the temperature term, linked to an increase in body temperatures. The TMI is thus expected to reflect increased body temperatures during the evolutionary transition to endothermy.

2. Estimating semicircular duct morphology in fossils

Description	**Explanation**
In practice, fossil species do not preserve the membranous labyrinth. Therefore, to calculate their 'raw' TMI we first need to estimate semicircular duct morphology from preserved bony labyrinth morphology. This can be done by using phylogenetic generalised least square regressions describing relationships between morphometrical parameters of the bony and membranous labyrinths (Supplementary Note 2, Models 8-19). In this context, we redefine the duct cross-sectional area a_S as: $$\left(a_S\right)_{ij} = \left(a_{S\|Fit}\right)_{ij} \cdot \left(a_{S\|Res}\right)_{ij}$$ where $$\left(a_{S\|Fit}\right)_{ij} \propto \left\{\left(r_{S\|Bony}\right)_{ij}, \left(R_{e\|Bony}\right)_{ij}\right\}$$ $$\left(R_{e\|Bony}\right)_{ij} = \frac{\left(M_{e\|Bony}\right)_{ij} + \left(m_{e\|Bony}\right)_{ij}}{2}$$ $$r_{S\|Bony} = \begin{bmatrix} r_{S_a\|Bony} \\ r_{S_p\|Bony} \\ r_{S_l\|Bony} \end{bmatrix} \quad R_{e\|Bony} = \begin{bmatrix} R_{e_a\|Bony} \\ R_{e_p\|Bony} \\ R_{e_l\|Bony} \end{bmatrix} \quad M_{e\|Bony} = \begin{bmatrix} M_{e_a\|Bony} \\ M_{e_p\|Bony} \\ M_{e_l\|Bony} \end{bmatrix} \quad m_{e\|Bony} = \begin{bmatrix} m_{e_a\|Bony} \\ m_{e_p\|Bony} \\ m_{e_l\|Bony} \end{bmatrix}$$	Fossil species do not preserve the membranous labyrinth. Therefore, we cannot directly calculate their 'raw' TMI. Here we describe how we estimated the duct cross-sectional area of fossil specimens, from the radius of curvature and the cross-sectional thickness of their bony semicircular canals, by using phylogenetic generalised least square regressions described in Supplementary Note 2, Models 8-19 (R^2 = 0.81-0.88).

and where: - $a_{S\|Fit}$: Fitted value of the duct cross-sectional area, predicted from the radius of curvature $R_{e\|Bony}$ and the cross-sectional thickness $r_{S\|Bony}$ of the bony semicircular canal, using Models 8-10 from Supplementary Note 2. - $a_{S\|Res}$: Residual variation resulting from the ratio between the actual duct cross-sectional area a_S and the fitted value of the duct cross-sectional area $a_{S\|Fit}$. - $R_{e\|Bony}$: Radius of curvature of the bony semicircular canal, measured from the semi-major axis $M_{e\|Bony}$ and semi-minor axis $m_{e\|Bony}$ of the ellipse best fitting the semicircular canal torus. - $r_{S\|Bony}$: Cross-sectional thickness of the slender portion of the bony semicircular canal.	
Similarly, we redefine the slender length L_S as: $$\left(L_S\right)_{ij} = \left(L_{S\|Fit}\right)_{ij} \cdot \left(L_{S\|Res}\right)_{ij}$$ where $$\left(L_{S\|Fit}\right)_{ij} \propto \left(L_{S\|Bony}\right)_{ij}$$ and where: - $L_{S\|Fit}$: Fitted value of the slender length, predicted from the length of the slender portion of the bony semicircular canal $L_{S\|Bony}$, using Models 14-16 from Supplementary Note 2. - $L_{S\|Res}$: Residual variation resulting from the ratio between the actual slender length L_S and the fitted value of the slender length $L_{S\|Fit}$. - $L_{S\|Bony}$: Length of the slender portion of the bony semicircular canal.	Here we describe how we estimated the slender length of fossil specimens, from the slender length of their bony semicircular canals, by using phylogenetic generalised least square regressions described in Supplementary Note 2, Models 14-16 (R^2 = 0.98-0.99).
We also redefine the enclosed area Λ as:	Here we describe how we estimated the enclosed area of fossil specimens, from the

$$\left(\Lambda\right)_{ij} = \left(\Lambda_{Fit}\right)_{ij} \cdot \left(\Lambda_{Res}\right)_{ij}$$

where

$$\left(\Lambda_{Fit}\right)_{ij} \propto \left\{\left(R_{e|Bony}\right)_{ij}, \left(\sigma_{e|Bony}\right)_{ij}\right\}$$

$$\left(\sigma_{e|Bony}\right)_{ij} = \frac{\left(\Lambda_{e|Bony}\right)_{ij}}{\left(R_{e|Bony}\right)_{ij}^{2}}$$

$$\left(\sigma_{e|Bony}\right)_{ij} = F\left(e_{e|Bony}\right)_{ij}$$

$$\sigma_{e|Bony} = \begin{bmatrix} \sigma_{e_a|Bony} \\ \sigma_{e_p|Bony} \\ \sigma_{e_l|Bony} \end{bmatrix} \quad \Lambda_{e|Bony} = \begin{bmatrix} \Lambda_{e_a|Bony} \\ \Lambda_{e_p|Bony} \\ \Lambda_{e_l|Bony} \end{bmatrix} \quad e_{e|Bony} = \begin{bmatrix} e_{e_a|Bony} \\ e_{e_p|Bony} \\ e_{e_l|Bony} \end{bmatrix}$$

and where:

- Λ_{Fit} : Fitted value of the enclosed area, predicted from the radius of curvature $R_{e|Bony}$ and the shape factor $\sigma_{e|Bony}$ of the bony semicircular canal, using Models 11-13 from Supplementary Note 2.
- Λ_{Res} : Residual variation resulting from the ratio between the actual enclosed area Λ and the fitted value of the enclosed area Λ_{Fit} .
- $\sigma_{e|Bony}$: Ratio between the area enclosed by the ellipse best fitting the bony semicircular canal torus $\Lambda_{e|Bony}$ and the square of the corresponding radius of curvature $R_{e|Bony}$. This parameter is called shape factor hereafter and is directly related to the eccentricity of the ellipse best-fitting the bony semicircular canal torus $e_{e|Bony}$ (Fig. B4).

Note 16: The shape factor increases when the eccentricity decreases (i.e., the semicircular canal torus becomes more circular).

radius of curvature and the shape factor of their bony semicircular canals, by using phylogenetic generalised least square regressions described in Supplementary Note 2, Models 11-13 (R^2 = 0.99-1.00). The shape factor is defined as the ratio between the area enclosed by the ellipse best fitting the semicircular canal torus and the corresponding squared radius of curvature.

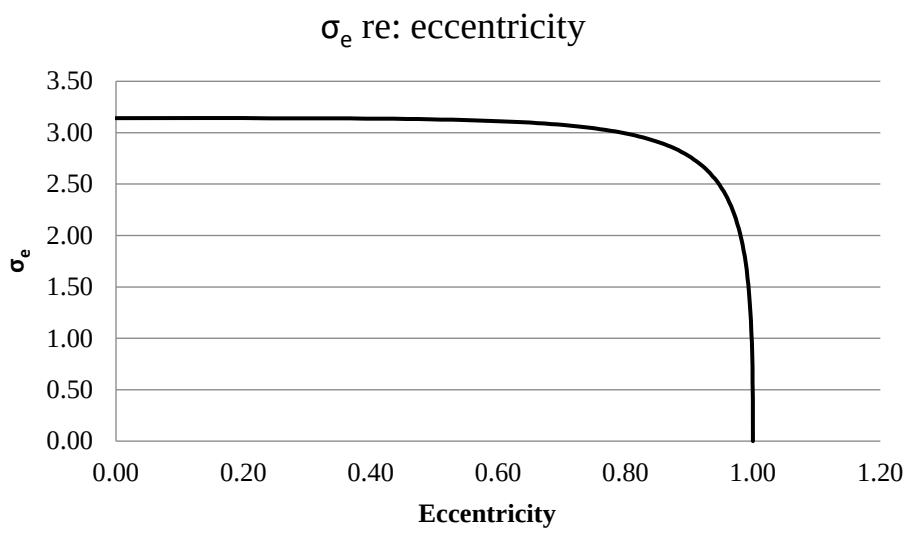


Fig. B4 Relationship between the shape factor σ_e and the eccentricity of the semicircular canal torus.

Finally, we redefine the cilia deflection factor E as:

$$\left(E\right)_{ij} = \left(E_{Fit}\right)_{ij} \cdot \left(E_{Res}\right)_{ij}$$

where

$$\left(E_{Fit}\right)_{ij} \propto \left(R_{e|Bony}\right)_{ij}$$

and where:

- E_{Fit} : Fitted value of the cilia deflection factor, predicted from the radius of curvature $R_{e|Bony}$ of the bony semicircular canal, using Models 17-19 from Supplementary Note 2.
- E_{Res} : Residual variation resulting from the ratio between the actual cilia deflection factor E and the fitted value of the cilia deflection factor E_{Fit} .

Here we describe how we estimated the cilia deflection factor of fossil specimens, from the radius of curvature of their bony semicircular canals, by using phylogenetic generalised least square regressions described in Supplementary Note 2, Models 17-19 (R^2 = 0.87-0.90).

To take into account the uncertainty in the TMI of fossil specimens, resulting from their lack of preserved membranous labyrinths, we re-express the 'raw' TMI, calculated from saturating velocities, as: $$\left(TMI_{\dot{\Omega}_2\|Raw}\right)_{ij} = \left(TMI_{\dot{\Omega}_2\|Fit}\right)_{ij} \cdot \left(TMI_{\dot{\Omega}_2\|Res}\right)_{ij} \cdot \beta_{\mu_e} \cdot \left(\theta_2\right)_{ij}$$ where $$\left(TMI_{\dot{\Omega}_2\|Fit}\right)_{ij} = \frac{\left(L_{S\|Fit}\right)_{ij}}{\left(\Lambda_{Fit}\right)_{ij}\left(E_{Fit}\right)_{ij}\left(a_{S\|Fit}\right)_{ij}^{2}}$$ $$\left(TMI_{\dot{\Omega}_2\|Res}\right)_{ij} = \frac{\left(\lambda_\mu\right)_{ij}\left(L_{S\|Res}\right)_{ij}}{2\rho\left(\Lambda_{Res}\right)_{ij}\left(E_{Res}\right)_{ij}\left(a_{S\|Res}\right)_{ij}^{2}}$$ $$TMI_{\dot{\Omega}_2\|Fit} = \begin{bmatrix} TMI_{\dot{\Omega}_{2a}\|Fit} \\ TMI_{\dot{\Omega}_{2p}\|Fit} \\ TMI_{\dot{\Omega}_{2l}\|Fit} \end{bmatrix} \quad TMI_{\dot{\Omega}_2\|Res} = \begin{bmatrix} TMI_{\dot{\Omega}_{2a}\|Res} \\ TMI_{\dot{\Omega}_{2p}\|Res} \\ TMI_{\dot{\Omega}_{2l}\|Res} \end{bmatrix}$$ and where: - $TMI_{\dot{\Omega}_2\|Fit}$: Fitted duct-morphology component of the 'raw' TMI, calculated from the saturating velocity. - $TMI_{\dot{\Omega}_2\|Res}$: Residual duct-morphology component of the 'raw' TMI, calculated from the saturating velocity. This component also incorporates the wall shape factor λ_μ and the density of the endolymph ρ.	Here we re-expressed the 'raw' TMI, calculated from saturating velocities, as a function of (1) duct morphology that can be estimated from the semicircular canals, (2) residual duct morphology that cannot be estimated from the morphology of the semicircular canals, (3) the physicochemical component of the TMI, and (4) the saturating angles, which are unknown. This was done to better distinguish the components of the TMI, calculated from saturating velocities, that are robustly calculated, from the ones that are potentially more labile.
Similarly, we re-express the 'raw' TMI, calculated from upper corner frequencies, as: $$\left(TMI_{\omega_2\|Raw}\right)_{ij} = \left(TMI_{\omega_2\|Fit}\right)_{ij} \cdot \left(TMI_{\omega_2\|Res}\right)_{ij} \cdot \beta_{\mu_e}$$ where $$\left(TMI_{\omega_2\|Fit}\right)_{ij} = \frac{1}{\left(a_{S\|Fit}\right)_{ij}}$$	Similarly, we re-expressed the 'raw' TMI, calculated from upper corner frequencies, as a function of (1) duct morphology that can be estimated from the semicircular canals, (2) residual duct morphology that cannot be estimated from the morphology of the semicircular canals, and (3) the physicochemical component of the TMI.

$$\left(TMI_{\omega_2|Res}\right)_{ij} = \frac{\left(\lambda_\mu\right)_{ij}}{2\pi\rho\left(a_{S|Res}\right)_{ij}}$$

$$TMI_{\omega_2|Fit} = \begin{bmatrix} TMI_{\omega_{2a}|Fit} \\ TMI_{\omega_{2p}|Fit} \\ TMI_{\omega_{2l}|Fit} \end{bmatrix} \quad TMI_{\omega_2|Res} = \begin{bmatrix} TMI_{\omega_{2a}|Res} \\ TMI_{\omega_{2p}|Res} \\ TMI_{\omega_{2l}|Res} \end{bmatrix}$$

and where:

- $TMI_{\omega_2|Fit}$: Fitted duct-morphology component of the 'raw' TMI, calculated from the upper corner frequency.
- $TMI_{\omega_2|Res}$: Residual duct-morphology component of the 'raw' TMI, calculated from the upper corner frequency. This component also incorporates the wall shape factor λ_μ and the density of the endolymph ρ.

Combining the new expressions for the 'raw' TMI, we obtain:

$$\left(TMI_{Raw}\right)_{ij} = \left(TMI_{Fit}\right)_{ij} \cdot \left(TMI_{Res}\right)_{ij} \cdot \beta_{\mu_e} \cdot \left(\varepsilon_\theta\right)_{ij} \quad \text{(Equation 6)}$$

$$TMI_{Fit} = \begin{bmatrix} TMI_{\dot{\Omega}_2|Fit} \\ TMI_{\omega_2|Fit} \end{bmatrix} \quad TMI_{Res} = \begin{bmatrix} TMI_{\dot{\Omega}_2|Res} \\ TMI_{\omega_2|Res} \end{bmatrix} \quad \varepsilon_\theta = \begin{bmatrix} \theta_2 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

where

- TMI_{Fit} : Fitted duct-morphology component of the 'raw' TMI.
- TMI_{Res} : Residual duct-morphology component of the 'raw' TMI.
- ε_θ : Error factor reflecting uncertainty in the saturating angle and its impact on the saturating velocity components of the 'raw' TMI (called deflection error hereafter).

Here we simply combine the formulae obtained in the two previous steps, to express the 'raw' TMI as a function of (1) duct morphology that can be estimated from the semicircular canals, (2) residual duct morphology that cannot be estimated from the morphology of the semicircular canals, (3) the physicochemical component of the TMI, and (4) an error factor, associated with the saturating angles.

Although the fitted duct-morphology component of the ‘raw’ Thermo-Motility Index TMI_{Fit}, which is obtained from bony labyrinth information by using phylogenetic generalised least square regressions (Supplementary Note 2, Models 8-19), does not capture the full variation of the ‘raw’ Thermo-Motility Index TMI_{Raw}, we can improve our estimations for fossil specimens. Indeed, the distribution of the residual duct-morphology component of the ‘raw’ Thermo-Motility Index TMI_{Res} shows a significant phylogenetic signal in extant tetrapods (Supplementary Note 2, Models 20-25). This property could be used to increase the amount of variation of the ‘raw’ TMI explained by our estimations. We explore this possibility by re-expressing the residual duct-morphology component of the ‘raw’ Thermo-Motility Index TMI_{Res} as:

$$\left(TMI_{Res}\right)_{ij} = \left(TMI_{Res|Pred}\right)_{ij} \cdot \left(\varepsilon_{Res}\right)_{ij} \quad \text{(Equation 7)}$$

where

$$\left(TMI_{Res|Pred}\right)_{ij} \propto \left\{\left(TMI_{Res}\right)_{ij}, \varphi\right\}$$

$$TMI_{Res|Pred} = \begin{bmatrix} TMI_{\dot{\Omega}_2|Res|Pred} \\ TMI_{\omega_2|Res|Pred} \end{bmatrix} \quad TMI_{\dot{\Omega}_2|Res|Pred} = \begin{bmatrix} TMI_{\dot{\Omega}_{2a}|Res|Pred} \\ TMI_{\dot{\Omega}_{2p}|Res|Pred} \\ TMI_{\dot{\Omega}_{2l}|Res|Pred} \end{bmatrix} \quad TMI_{\omega_2|Res|Pred} = \begin{bmatrix} TMI_{\omega_{2a}|Res|Pred} \\ TMI_{\omega_{2p}|Res|Pred} \\ TMI_{\omega_{2l}|Res|Pred} \end{bmatrix}$$

$$\varepsilon_{Res} = \begin{bmatrix} \varepsilon_{\dot{\Omega}_2|Res} \\ \varepsilon_{\omega_2|Res} \end{bmatrix} \quad \varepsilon_{\dot{\Omega}_2|Res} = \begin{bmatrix} \varepsilon_{\dot{\Omega}_{2a}|Res} \\ \varepsilon_{\dot{\Omega}_{2p}|Res} \\ \varepsilon_{\dot{\Omega}_{2l}|Res} \end{bmatrix} \quad \varepsilon_{\omega_2|Res} = \begin{bmatrix} \varepsilon_{\omega_{2a}|Res} \\ \varepsilon_{\omega_{2p}|Res} \\ \varepsilon_{\omega_{2l}|Res} \end{bmatrix}$$

and where:

- $TMI_{Res|Pred}$: Estimated residual duct-morphology component of the ‘raw’ TMI, predicted from known residual duct-morphology components of the ‘raw’ TMI of other specimens, and from their phylogenetic relatedness with the specimen of interest φ.
- ε_{Res}: Error factor reflecting the uncertainty in the estimated residual duct-morphology component of the ‘raw’ Thermo-Motility Index $TMI_{Res|Pred}$ (called residual error hereafter).

Here we develop the idea that, because the residual duct-morphology component of the ‘raw’ TMI shows significant phylogenetic signal (Supplementary Note 2, Models 20-25), known residual duct-morphology components of various species could in principle be used, along with a phylogeny, to estimate the residual duct-morphology component of the TMI of the specimen of interest. Nevertheless, we note that part of the residual duct-morphology component cannot be predicted and represents an error factor of the TMI.

Considering all potential sources of error in the estimation of the 'raw' TMI of fossil specimens, we note that the physicochemical component of the TMI β_{μ_e} can be re-expressed as: $$\beta_{\mu_e} = \beta_{\mu_e\|Pred} . \varepsilon_{\beta_{\mu_e}}$$ (Equation 8) where - $\beta_{\mu_e\|Pred}$: Estimated physicochemical component of the TMI (i.e., low-viscosity endolymph by default, high-viscosity endolymph for birds and pleuronectids) - $\varepsilon_{\beta_{\mu_e}}$: Error factor reflecting the uncertainty in the estimated physicochemical component of the 'raw' TMI $\beta_{\mu_e\|Pred}$ (called the viscosity error hereafter).	Here we note that we do not know the actual physicochemical component of the TMI for fossil specimens, and while we use available knowledge to posit the most parsimonious assumption about its value, this uncertainty leads to another error factor of the TMI.
Combining equations 6, 7, and 8, we re-express the Thermo-Motility Index TMI_{Raw} as: $$\left(TMI_{Raw}\right)_{ij} = \left(TMI_{Raw\|Pred}\right)_{ij} \cdot \left(\varepsilon\right)_{ij}$$ where $$\left(\varepsilon\right)_{ij} = \varepsilon_{\beta_{\mu_e}} \cdot \left(\varepsilon_{Res}\right)_{ij} \cdot \left(\varepsilon_{\theta}\right)_{ij}$$ $$\left(TMI_{Raw\|Pred}\right)_{ij} = \left(TMI_{Fit}\right)_{ij} \cdot \left(TMI_{Res\|Pred}\right)_{ij} \cdot \beta_{\mu_e\|Pred}$$ $$\left(TMI_{Raw\|Pred}\right)_{ij} = \frac{\left(\max M_{Head}\right)_{ij}}{10^{T_R} . \left(\tau_{F_{SD\|2}}\right)_{ij} \cdot \left(\varepsilon\right)_{ij}}$$ (Equation 9) $$TMI_{Raw\|Pred} = \begin{bmatrix} TMI_{\dot{\Omega}_2\|Raw\|Pred} \\ TMI_{\omega_2\|Raw\|Pred} \end{bmatrix} \quad TMI_{\dot{\Omega}_2\|Raw\|Pred} = \begin{bmatrix} TMI_{\dot{\Omega}_{2a}\|Raw\|Pred} \\ TMI_{\dot{\Omega}_{2p}\|Raw\|Pred} \\ TMI_{\dot{\Omega}_{2l}\|Raw\|Pred} \end{bmatrix} \quad TMI_{\omega_2\|Raw\|Pred} = \begin{bmatrix} TMI_{\omega_{2a}\|Raw\|Pred} \\ TMI_{\omega_{2p}\|Raw\|Pred} \\ TMI_{\omega_{2l}\|Raw\|Pred} \end{bmatrix}$$ and where: - $TMI_{Raw\|Pred}$: 'raw' TMI, as estimated from morphometrical parameters of the bony labyrinth of the specimen, and from the known residual duct-morphology TMI_{Res} and physicochemical β_{μ_e} components of the 'raw' TMI of other specimens, and their phylogenetic relatedness φ with the specimen of interest.	Here we describe the components of the 'raw' TMI, as estimated for fossil specimens. In addition to the fitted and residual duct-morphology components, and the physicochemical component, we note an error term which reflects the uncertainty in endolymph composition, in unpredictable variability of semicircular duct morphology and in the saturation angles. The estimated 'raw' TMI is proportional to maximum angular head motion, and inversely proportional to the temperature term (which decreases when temperature increases), to the tuning matrix and to the error term of the TMI.

- ε : Error term reflecting the uncertainty in the estimated 'raw' TMI of the specimen of interest (called error term of the TMI hereafter). Note 17: While not perfect, Model 47 (Supplementary Data 2) suggests that the estimated (size-corrected weighted average) TMI reflects about 75% of the variation of the actual TMI.	

3. Correcting the Thermo-Motility Index for body size

Description	**Explanation**
Like semicircular duct function (Jones & Spells 1963), the 'raw' TMI is affected by body size (Supplementary Note 2, Models 26-31). Therefore, to compare the TMI of different species, we first need to scale them to a common size reference. Size correction of the estimated 'raw' TMI is achieved through the following formulae: $$\left(TMI_{Cor\|Pred}\right)_{ij} = \left(TMI_{Cor\|Pred\|Fit}\right)_{ij} . \left(TMI_{Raw\|Pred\|Res}\right)_{ij}$$ (Equation 10) where $$\left(TMI_{Raw\|Pred\|Res}\right)_{ij} = \frac{\left(TMI_{Raw\|Pred}\right)_{ij}}{\left(TMI_{Raw\|Pred\|Fit}\right)_{ij}}$$ $$\left(TMI_{Raw\|Pred\|Fit}\right)_{ij} \propto M_B$$ $$\left(TMI_{Cor\|Pred\|Fit}\right)_{ij} \propto \left\{M_B = 1g\right\}$$ $$TMI_{Cor\|Pred} = \begin{bmatrix} TMI_{\dot{\Omega}_2\|Cor\|Pred} \\ TMI_{\omega_2\|Cor\|Pred} \end{bmatrix} \quad TMI_{\dot{\Omega}_2\|Cor\|Pred} = \begin{bmatrix} TMI_{\dot{\Omega}_{2a}\|Cor\|Pred} \\ TMI_{\dot{\Omega}_{2p}\|Cor\|Pred} \\ TMI_{\dot{\Omega}_{2l}\|Cor\|Pred} \end{bmatrix} \quad TMI_{\omega_2\|Cor\|Pred} = \begin{bmatrix} TMI_{\omega_{2a}\|Cor\|Pred} \\ TMI_{\omega_{2p}\|Cor\|Pred} \\ TMI_{\omega_{2l}\|Cor\|Pred} \end{bmatrix}$$ and where: - $TMI_{Cor\|Pred}$: Size-corrected TMI, corresponding to the estimated 'raw' TMI of the specimen of interest if its body mass M_B was equal to 1g. - $TMI_{Cor\|Pred\|Fit}$: Fitted value of the size-corrected TMI of the specimen of interest, predicted from a body mass M_B equal to 1g, and using Models 26-31 from Supplementary Note 2.	Because the 'raw' TMI is correlated to body size (Supplementary Note 2, Models 26-31), proper comparisons between different species require to scale it to a standard body mass (e.g., 1g).

- $TMI_{Raw|Pred|Res}$: Residual variation resulting from the ratio between the estimated 'raw' Thermo-Motility Index $TMI_{Raw|Pred}$ and the fitted value of the estimated 'raw' Thermo-Motility Index $TMI_{Raw|Pred|Fit}$, predicted from the actual body mass of the specimen of interest M_B.

Using equations 9 and 10, we can thus write:

$$\left(TMI_{Cor|Pred}\right)_{ij} = \frac{\left(\max M_{Head|Cor}\right)_{ij}}{10^{T_R}.\left(\tau_{F_{SD|2}}\right)_{ij} \cdot \left(\varepsilon\right)_{ij}} \text{ (Equation 11)}$$

where

$$\left(\max M_{Head|Cor}\right)_{ij} = \left(\max M_{Head|Cor|Fit}\right)_{ij} . \left(\max M_{Head|Res}\right)_{ij}$$

$$\left(\max M_{Head|Res}\right)_{ij} = \frac{\left(\max M_{Head}\right)_{ij}}{\left(\max M_{Head|Fit}\right)_{ij}}$$

$$\left(\max M_{Head|Fit}\right)_{ij} \propto M_B$$

$$\left(\max M_{Head|Cor}\right)_{ij} \propto \left\{M_B = 1g\right\}$$

$$\max M_{Head|Cor} = \begin{bmatrix} \max \dot{\Omega}_{H|Cor} \\ \max f_{H|Cor} \end{bmatrix} \quad \max \dot{\Omega}_{H|Cor} = \begin{bmatrix} \max \dot{\Omega}_{Ha|Cor} \\ \max \dot{\Omega}_{Hp|Cor} \\ \max \dot{\Omega}_{Hl|Cor} \end{bmatrix} \quad \max f_H = \begin{bmatrix} \max f_{Ha|Cor} \\ \max f_{Hp|Cor} \\ \max f_{Hl|Cor} \end{bmatrix}$$

and where:

- $\max M_{Head|Cor}$: Size-corrected maximum angular head motion, corresponding to the estimated maximum angular head motion of the specimen of interest if its body mass M_B was equal to 1g.
- $\max M_{Head|Cor|Fit}$: Fitted value of the size-corrected maximum angular head motion of the specimen of interest, predicted from a body mass M_B equal to 1g, and using phylogenetic generalized least square regressions between maximum angular head motion and body mass.

Similarly, we note that because maximum angular head motion is also likely correlated to body size (Jones & Spells 1963), we theoretically need to scale it to a standard body mass (e.g., 1g) to compare different species. Note however that we do not actually measure angular head motion in this study, nor correct it for body size differences, but we need to define it to demonstrate that the TMI intrinsically correlates to body temperatures.

- $_{\max}M_{Head\|Res}$: Residual variation resulting from the ratio between the maximum angular head motion $_{\max}M_{Head}$ and the fitted value of the maximum angular head motion $_{\max}M_{Head\|Fit}$, predicted from the actual body mass of the specimen of interest M_B. Note 18: Note that the temperature term, the tuning factor matrix and the error term of the TMI could all be correlated to body mass in theory. However, we have evidence suggesting this is not the case for the temperature term and the error term of the TMI (Supplementary Note 2, Models 20-25, 81). Concerning the tuning factor matrix, we suspect that as a parameter bounded between]0;1], it is unlikely to linearly correlate to body mass.	

4. Refining the Thermo-Motility Index

Description	**Explanation**
The size-corrected estimated Thermo-Motility Index $TMI_{Cor\|Pred}$ is actually a set composed of six Thermo-Motility Indices, calculated for each of the three semicircular ducts, from the saturating velocity and the upper corner frequency. However, these Thermo-Motility Indices are not equally informative concerning body temperatures (Supplementary Note 2, Models 32-37). To optimise our predictions of body temperatures and classifications of the thermoregulatory regime of fossil specimens, we average these six Thermo-Motility Indices, using Akaike weights (Supplementary Note 2, Section 11), such that: $$TMI = \sum\left(\log_{10}\left(TMI_{Cor\|Pred}\right)_{ij}.\left(W\right)_{ij}\right) \text{ (Equation 12)}$$ $$W = \begin{bmatrix} w_{\dot{\Omega}_{2a}} \\ w_{\dot{\Omega}_{2p}} \\ w_{\dot{\Omega}_{2l}} \\ w_{\omega_{2a}} \\ w_{\omega_{2p}} \\ w_{\omega_{2l}} \end{bmatrix}$$ where: - TMI : Size-corrected, logged, weighted average of the TMI. This corresponds to the TMI we use in our study (just called TMI in the main text and hereafter).	There are actually six different size-corrected estimated Thermo-Motility Indices, and they do not carry the same amount of information about body temperatures (Supplementary Note 2, Models 32-37). Thus, we average these Thermo-Motility Indices, using weights reflecting how informative they are regarding body temperatures (Supplementary Note 2, Section 11), to optimise body temperature predictions and classifications of thermoregulatory regimes of fossil specimens. The weighted average TMI we obtain (which is logged) corresponds to the TMI we use in this study, and that we mention in the main text.

- W : Matrix of Akaike weights reflecting the information about body temperatures contained in each individual Thermo-Motility Indices composing a set (Supplementary Note 2, Section 11).	
Following equations 11 and 12, it appears that the Thermo-Motility Index TMI is related to the size-corrected, logged, weighted average of maximum angular head motion ${}_{\max}M_{Head\|Log\|W}$, to the size-corrected, logged, weighted average of the tuning factor $\tau_{F_{SD\|2}\|Log\|W}$, to the size-corrected, logged, weighted average of the error term $\varepsilon_{Log\|W}$, and to the temperature term T_R, such that: $$TMI = {}_{\max}M_{Head\|Log\|W} - T_R - \tau_{F_{SD\|2}\|Log\|W} - \varepsilon_{Log\|W}$$ (Equation 13) where $${}_{\max}M_{Head\|Log\|W} = \sum\left(\log_{10}\left({}_{\max}M_{Head\|Cor}\right)_{ij}.(W)_{ij}\right)$$ $$\tau_{F_{SD\|2}\|Log\|W} = \sum\left(\log_{10}\left(\tau_{F_{SD\|2}}\right)_{ij}.(W)_{ij}\right)$$ $$\varepsilon_{Log\|W} = \sum\left(\log_{10}\left(\varepsilon\right)_{ij}.(W)_{ij}\right)$$ and where: - ${}_{\max}M_{Head\|Log\|W}$: Size-corrected, logged, weighted average of maximum angular head motion. - $\tau_{F_{SD\|2}\|Log\|W}$: Size-corrected, logged, weighted average of the tuning factor (i.e., bounded between $]-\infty,0]$). - $\varepsilon_{Log\|W}$: Size-corrected, logged, weighted average of the error term of the TMI.	Here we note that the TMI is a function of (1) the temperature term, (2) maximum angular head motion, (3) the tuning factor, and (4) the error term. Parameters 2 to 4 are the weighted average of size-corrected and logged values. Increases in the maximum angular head motion and/or body temperature (reflected by a decrease in the temperature term), and/or decreases in the tuning factor and/or error term all lead to an increase of the TMI.

5. Empirical data versus model

Description	Explanation
If body temperature was only affecting the Thermo-Motility Index TMI through its impact on the viscosity of the endolymph, the TMI would be expected to compensate for the change in the temperature term T_R in an isometric way, to keep the upper functional limit $F_{SD\|2}$ constant. Regressing the Thermo-Motility Index TMI onto the temperature term T_R , and following equation 14, we should thus obtain the following regression equation: $$TMI = (-1) \cdot T_R + C$$ where $$C = {}_{\max} M_{Head\|Log\|W} - \tau_{F_{SD\|2}\|Log\|W} - \varepsilon_{Log\|W}$$ However, an empirical phylogenetic generalised least square regression obtained from real world data (Supplementary Note 2, Model 42) rather suggests that: $$TMI = (-3.7) \cdot T_R + D$$ which directly imply that $$(-2.7) \cdot T_R + D = {}_{\max} M_{Head\|Log\|W} - \tau_{F_{SD\|2}\|Log\|W} - \varepsilon_{Log\|W} \quad \text{(Equation 14)}$$ From equation 14 it clearly appears that either (1) the size-corrected, weighted average, logged maximum angular head motion ${}_{\max} M_{Head\|Log\|W}$, (2) the size-corrected, weighted average, logged tuning factor $\tau_{F_{SD\|2}\|Log\|W}$, (3) the size-corrected, weighted average, logged error term $\varepsilon_{Log\|W}$, or (4) a combination of 1, 2, and 3, is related to body temperature through the temperature term T_R .	Here we discuss the discrepancy we observe between what we expected (i.e., isometric scaling between the TMI and the temperature term, to keep semicircular duct function constant) and what we empirically obtained (i.e., a steeper slope, Supplementary Note 2, Model 42). We note that a steeper slope implies that the temperature term not only correlates to the TMI but also to any of the maximum angular head motion, the tuning factor or the residual term.

If we assume that:

$$_{\max} M_{Head|Log|W} \propto -T_R$$

this would imply that the size-corrected, weighted average, logged maximum angular head motion $_{\max} M_{Head|Log|W}$ increases when the temperature term T_R decreases, and thus when body temperature increases (see equation 1). This is in agreement with the assumption that increased body temperatures lead to increased behavioural activity, and potentially to increased maximum angular head motion, on average (Bennett & Ruben 1979, Garland & Albuquerque 2017, Hirt et al. 2017, Supplementary Data 1, Supplementary Note 2, Models 78-79/88-89).

We interpret the evidence for a relationship between body temperature and behavioural activity (Bennett & Ruben 1979, Garland & Albuquerque 2017, Hirt et al. 2017, Supplementary Data 1, Supplementary Note 2, Models 78-79/88-89) as potentially supporting a correlation between maximum angular head motion and the temperature term.

Alternatively, if we assume that:

$$\tau_{F_{SD|2}|Log|W} \propto T_R$$

this would suggest that the size-corrected, weighted average, logged tuning factor $\tau_{F_{SD|2}|Log|W}$ increases when the temperature term T_R increases, and thus when body temperature decreases. However, as illustrated in Fig. B2, when body temperature increases, the tuning factor will tend to increase (short arrows in Fig. B2), if the semicircular ducts do not compensate. A positive correlation between the tuning factor and the temperature term thus seems unlikely. Additionally, we expect the tuning factor to be optimised around a fixed value, compromising between broad efficacy and resolution of the semicircular ducts, rather than to correlate with body temperature.

We discuss the possibility of a negative correlation between the tuning factor and body temperature, but conclude it is unlikely because, if anything, the contrary would be expected (Fig. B2).

Finally, if we assume that:

$$\varepsilon_{Log|W} \propto T_R$$

this would imply that the size-corrected, weighted average, logged error term $\varepsilon_{Log|W}$ increases when the temperature term T_R increases, and thus when body temperature decreases. As previously mentioned, the error term ε consists of three components, which are the residual error ε_{Res}, the viscosity error $\varepsilon_{\beta_{\mu_e}}$ and the deflection error ε_{θ}.

We then discuss the possibility of a negative correlation between the error term and body temperature.

Empirical evidence demonstrates that the residual error ε_{Res} is not correlated with the temperature term (Supplementary Note 2, Models 82-83).

We note that the residual error factor of the error term is not correlated to the temperature term (Supplementary Note 2, Models 82-83).

High endolymph viscosity has been observed in birds, which are endotherms possessing the highest recorded body temperatures. This suggests that organisms with high body temperatures, but 'normal' semicircular duct morphology, could compensate the temperature increase by evolving a high-viscosity endolymph. Because we use a low-viscosity endolymph by default in our study, if anything, we expect to see an increase of the viscosity error $\varepsilon_{\beta_{\mu_e}}$ when body temperature increase. This is contrary to what is assumed above (i.e., that the error term $\varepsilon_{Log\|W}$ increases when body temperature decreases), suggesting that a correlation between the viscosity error and body temperature is unlikely.	We note that, if anything, the viscosity error factor of the error term would be expected to increase with body temperature, not decrease.
The saturating velocity components of the TMI are the only components affected by the deflection error and cupula biomechanics. In this context, Models 32 and 35 (Supplementary Note 2) show that the temperature term correlates differently with the upper corner frequency and saturating velocity components of the anterior TMI (slopes of -0.30 *vs* -0.18). If not related to the difference in the signals captured by these two Thermo-Motility Indices (i.e., angular velocity and ordinary frequencies), this difference in slopes may suggest that a ratio between (1) the actual saturation velocity of a semicircular duct, and (2) the saturation velocity of the same duct, as derived from its TMI, would decrease with body temperature. This may imply that the temperature term either correlates with the deflection error, or that body temperature affects cupula biomechanics in a way that we did not take into account in this study. If body temperature does affect the deflection error, this would mean that the saturating angle decreases when body temperature increases (to compensate for the potential increase in the TMI, calculated from the saturating velocity). However, as previously explained, this seems unlikely because the saturating angle is probably strongly constrained structurally (see Note 11). On the other hand, we cannot exclude that an increase in body temperature could lead the cupula to deflect more than we anticipated, maybe by reducing its stiffness. In any case, this observation would only apply to the saturating velocity components of the TMI, and, as such, this could only explain the difference in slopes between Models 32 and 35, not the observation that the TMI, as a whole, increases more than expected when body temperature increases.	Then, we note that it is unlikely that the deflection error factor of the error term is correlated to the temperature term, because of structural constraints linked to the microanatomy of the cilia. We cannot however exclude that body temperature affects cupula deflection in ways we did not foresaw, but this cannot explain the steeper slope we observe in itself, because the upper corner frequency component of the Thermo-Motility Index is not affected by cupular biomechanics.
In conclusion, taking all of these considerations into account, it seems that the discrepancy between the theoretical result we expected (i.e., isometric scaling between the TMI and the temperature term) and what we empirically obtained (i.e., a steeper slope) most likely results from a positive correlation between maximum angular head motion and body temperature.	We thus conclude that the steeper slope we empirically observe most likely reflects a positive correlation between maximum angular head motion and body temperature.

References

Allum, J. H., Bloem, B. R., Carpenter, M. G., Hulliger, M. & Hadders-Algra, M. Proprioceptive control of posture: a review of new concepts. Gait Posture 8, 214–242 (1998).

Angelaki, D. E. & Cullen, K. E. Vestibular system: The many facets of a multimodal sense. *Annu. Rev. Neurosci*. **31**, 125–150 (2008).

Bennett, A. F. & Ruben, J. A. Endothermy and activity in vertebrates. Science 206, 649–654 (1979).

Berthoz, A. Reference frames for the perception and control of movement In Brain and space (ed. Paillard, J.) 81–111 (Oxford University Press, USA, 1991).

Berthoz, A. & Pozzo, T. Intermittent head stabilization during postural and locomotory tasks in humans In Posture and Gait: Development, adaptation and modulation (eds Amblard, B., Berthoz, A. & Clarac, F.) 189–198 (Excerpta medica, Amsterdam, 1988).

Bronstein, A. M. Evidence for a vestibular input contributing to dynamic head stabilization in man. Acta Otolaryngol. (Stockh.) 105, 1–6 (1988).

Creath, R. et al. The role of vestibular and somatosensory systems in intersegmental control of upright stance. *J. Vest. Res*. **18**, 39–49 (2008).

David, R. et al. Assessing morphology and function of the semicircular duct system: introducing new in-situ visualization and software toolbox. Sci. Rep. 6, 1-14 (2016).

Fitzpatrick, R. C., Butler, J. E. & Day, B. L. Resolving head rotation for human bipedalism. *Curr. Biol*. **16**, 1509–1514 (2006).

Garland, T. & Albuquerque, R. L. Locomotion, energetics, performance, and behavior: a mammalian perspective on lizards, and vice versa. Integr. Comp. Biol. 57, 252-266 (2017).

Highstein, S. M., Fay, R. R. & Popper, A. N. *The Vestibular System* (Springer, 2004).

Hirt, M. R. et al. A general scaling law reveals why the largest animals are not the fastest. Nat. Ecol. Evol. 1, 1116–1122 (2017).

Jones, G. M., &. & Spells, K. E. A theoretical and comparative study of the functional dependence of the semicircular canal upon its physical dimensions. Proc. R. Soc. Lond. B Biol. Sci. 157, 403–419 (1963).

Macdonald, J. A. & Wells, R. M. G. in Biology of Antarctic fish (eds di Prisco, G., Maresca, B. & Tota B.) 163-178. (Springer, 1991).

Mergner, T., Huber, W. & Becker, W. Vestibular-neck interaction and transformation of sensory coordinates. J. Vestib. Res. 7, 347–367 (1997).

Money, K. E. et al. Physical properties of fluids and structures of vestibular apparatus of the pigeon. Am. J. Physiol. Leg. Content 220, 140–147 (1971).

Oman, C. M. in The Vestibular System: Function and Morphology (ed Gualtierotti, T.) 251–274 (Springer, 1981).

Owen, B. & Lee, D. Establishing a frame of reference for action In Motor development in children: aspect of coordination and control (eds Wade, M. & Whiting, H.) 341–360 (Martinus Nijhoff, Dordrecht, 1986).

Pozzo, T., Berthoz, A. & Lefort, L. Head stabilization during various locomotor tasks in humans. 1. Normal subjects. Exp. Brain Res. 82, 97–106 (1990).

Rabbitt, R. D., Damiano, E. R. & Grant, J. W. in The vestibular system (eds Highstein, S. M. et al) 153–2

Rauch, S. Biochemie das Hörorgans (Thieme Verlag, 1964).

Rüsch, A., & Thurm, U. Cupula displacement, hair bundle deflection, and physiological responses in the transparent semicircular canal of young eel. Pflügers Archiv, 413(5), 533-545.01 (1989).

Sakaguchi, H., Tokita, J., Müller, U., & Kachar, B. (2009). Tip links in hair cells: molecular composition and role in hearing loss. Current opinion in otolaryngology & head and neck surgery, 17(5), 388.

Schnieder, E. A. & Schindler, K. in Biochemie des Hörorgans (ed Rauch, S.) 164-168 (Georg Thieme Verlag, 1964).

Steer, R. W. The influence of angular and linear acceleration and thermal stimulation on the human semicircular canal (Massachusetts Institute of Technology, 1967).

Ten Doesschate, G. Onderzoekingen gedaan in het Fysiologisch Laboratorium der Utrechtse Hogeschool 5e reeks XIV. De eigenschappen van de endolymph van Beenvisschen. (1914).

Ten Kate, J. H. & Kuiper, J. W. The viscosity of the Pike's endolymph. J. Exp. Biol. 53, 495–500 (1970).

Valerio, S. & Taube, J. S. Head Direction Cell Activity Is Absent in Mice without the Horizontal Semicircular Canals. J. Neurosci. 36, 741–754 (2016).

Wilson, V. J. & Melvill Jones, G. (Eds.) Mammalian Vestibular Physiology, Plenum Press, New York and London (1979).

The endolymph viscosity of vertebrates

Phylogenetic distribution of available measurements of endolymph viscosity at a given temperature indicates that a low-viscosity endolymph was the basal condition for Euarchontoglires and Euteleostei (Methods). Therefore, pending on new data, low-viscosity endolymph is currently inferred to be the basal condition for Osteichtyes. Conversely, feral rock pigeons, which possess a high-viscosity endolymph, do not show any peculiarities in terms of locomotor behaviours, body temperature, or labyrinth morphology when compared to other birds. We thus assume that a high-viscosity endolymph is typical for Aves. This raises the question as to when a high-viscosity endolymph originated in diapsids. Data from audition in various lizard species suggest that an endolymph with water-like viscosity would better fit empirical results in these species (Manley 2006, 2014). These results indicate that a high-viscosity endolymph would be unlikely for lepidosaurs, and the acquisition of a high-viscosity endolymph would have occurred between the origin of birds and the divergence of Lepidosauria and Archelosauria. In this context it should be noted that Crocodilia and Testudines are in line with the theoretical and empirical relationships between the TMI and body temperature (Fig. 2), which suggests, *a posteriori*, the retention of low-viscosity endolymph in these taxa.

References

Manley, G. A. Spontaneous otoacoustic emissions from free-standing stereovillar bundles of ten species of lizard with small papillae. *Hearing Res*. **212**, 33-47 (2006).

Manley, G. A. in *Perspectives on auditory research* (eds Popper, A. N. & Fay, R. R.) 321-341 (Springer, 2014).

Step-by-step summary of the Analyses

For fossil specimens and specimens without membranous labyrinth information:

1. Either obtain body mass information directly from the specimen or predict body mass by using Model 7 (Supplementary Note 2) with measurements of condylobasal and condylo-anteroorbital lengths taken from the specimen.
2. Scan the inner ear of the specimen using micro-computed tomography (µCT).
3. Segment the bony semicircular canals using visualisation software (e.g., Avizo).
4. Measure the minor and major axes of each semicircular canal, as illustrated in Extended Data Figure 2 (see Methods).
5. Measure the length of the slender portion of each semicircular canal, as illustrated in Extended Data Figure 2 (see Methods).
6. Measure the average thickness of the slender portion of each semicircular canal, as illustrated in Extended Data Figure 2 (see Methods).
7. Calculate the radius of curvature of each semicircular canal, as explained in Supplementary Methods, Biomechanics section.
8. Calculate eccentricity of the torus of each semicircular canal, and corresponding shape factor, as explained in Supplementary Methods, Biomechanics section.
9. Predict the fitted cross-sectional area of the slender portion of each semicircular duct using Models 8-10 (Supplementary Note 2) with the radius of curvature and the thickness of the slender portion of each semicircular canal obtained in steps 6 and 7.
10. Predict the fitted enclosed area of each semicircular duct using Models 1113 (Supplementary Note 2) with the radius of curvature and the shape factor of the torus of each semicircular canal obtained in steps 7 and 8.
11. Predict the fitted length of the slender portion of each semicircular duct using Models 14-16 (Supplementary Note 2) with the length of the slender portion of each semicircular canal obtained in step 5.
12. Predict the fitted cilia deflection factor of each semicircular duct using Models 17-19 (Supplementary Note 2) with the radius of curvature obtained in step 7.

13. Compute the fitted duct-morphology component of each of the six ‘raw’ Thermo-Motility Indices (i.e., calculated for the three ducts from the upper corner frequency and the saturating velocity), as explained in the Supplementary Methods, Biomechanics of the Thermo-Motility Index, using the fitted enclosed area, the fitted cilia deflection factor, and the fitted cross-sectional area and fitted length of the slender portion, obtained for each semicircular duct in steps 9-12.

14. Predict the residual duct-morphology component of each ‘raw’ Thermo-Motility Index by using ancestral state reconstructions at the node where the specimen branches out from a tree that only contain specimens for which the residual duct-morphology is known (e.g., Supplementary Data 2, 3D Membranous set). Although we provide a set of 47 specimens for which residual duct-morphology components are known, this set can be expanded by subtracting fitted duct-morphology predicted from the bony labyrinths of specimens from the actual duct morphology calculated from respective membranous labyrinth, as explained in the Supplementary Methods, Biomechanics of the Thermo-Motility Index.

15. Calculate the ‘raw’ Thermo-Motility Indices by adding fitted duct-morphology components (obtained in step 13) and residual duct-morphology components (obtained in step 14) with the physicochemical component, as described in Supplementary Methods, Biomechanics of the Thermo-Motility Index. The physicochemical component should be taken as low-viscosity by default but should be changed to high-viscosity for specimens for which there is evidence of a particularly viscous endolymph (e.g., birds).

16. Correct the ‘raw’ Thermo-Motility Indices for body size differences using Models 26-31 (Supplementary Note 2) with body mass information obtained in step 1, as explained in the Supplementary Methods, Biomechanics of the Thermo-Motility Index.

17. Compute the size-corrected weighted average Thermo-Motility Index by using the size-corrected Thermo-Motility Indices obtained in step 16 with Akaike weights provided in section 11 of Supplementary Note 2. Note that we provide the $\log_{10}$ version of the Thermo-Motility Index in this study.

For specimens with membranous labyrinth information:

1. As for step 1 above.
2. Scan the inner ear of the specimen using μCT and contrast agents to visualise the soft tissues (David et al. 2016).
3. Segment the membranous semicircular duct system using visualisation software (e.g., Avizo).
4. Prepare a 3D model of the membranous semicircular duct system as explained in David et al. (2016) and use Ariadne software (www.earbank.org) to directly obtain relevant metrics. Alternatively, measure the enclosed area, the length, and cross-sectional area of the slender portion of each semicircular duct, and the cilia deflection factor for each cupula, via other means.
5. Calculate 'raw' Thermo-Motility Indices by combining duct-morphology components (obtained in step 4) with the physicochemical component, as described in the Supplementary Methods, Biomechanics section. The physicochemical component should be taken as low-viscosity by default but should be changed to high-viscosity for specimens for which there is evidence of a particularly viscous endolymph (e.g., birds).
6. As for step 16 above.
7. As for step 17 above.

References

David, R., Stoessel, A., Berthoz, A., Spoor, F., & Bennequin, D. (2016). Assessing morphology and function of the semicircular duct system: introducing new in-situ visualization and software toolbox. Scientific reports, 6(1), 1-14.

Supplementary Note 1 – Definitions

Movement – all musculoskeletal motion of an animal unrelated to action resulting from physiological responses (e.g., cardiorespiratory muscles, internal organ movements, blinking).

Behavioural activity – a situation in which an animal's movement is occurring. Behavioural activity can be empirically measured in multiple ways, of which we conceptually entertain the following in our paper: (1) *temporal activity*, length of time (within a day, year, reproductive cycle, *etc*.) during which the animal is moving, (2) *spatial activity*, total distance traversed by the animal during a certain period of time (e.g., home range is a frequently noted example in animal ecology); (3) *aerobic activity*, quantification of animal movements resulting mostly from aerobic metabolic pathways over a certain period of time [e.g., maximum aerobic speed, average animal overall body acceleration (OBDA), or average angular frequency, velocity and acceleration of the head]; and (4) *anaerobic activity*, quantification of animal movements resulting mostly from anaerobic metabolic pathways [e.g., maximum speed, maximum OBDA, or maximum angular frequency, velocity and acceleration of the head].

Body temperature **–** active non-shivering animal's temperature measured at its core (e.g., using implants under the skin, not on the skin or in the animal's extremities).

Ambient temperature – temperature of the medium (e.g., air, water, substrate) measured in the immediate vicinity of the animal.

Endotherm – an animal producing heat throughout the entire body via metabolism (not shivering and/or muscular thermogenesis), maintaining nearly constant body temperature largely independently from external conditions, and excluding phases of short-term torpor, aestivation, and hibernation.

Our definition includes monotremes, whose body temperature is relatively stable and significantly higher than ambient temperature when cold (Nicol & Anderson 2006), but resort to behavioural thermoregulation – as in many other mammals (Terrien et al. 2011) – when ambient temperatures are too hot (Brice et al. 2002). Our definition is similar to that of Clarke & Pörtner (2010).

Ectotherm – an animal whose body temperature correlates to ambient temperature of its natural habitat during active phases of life.

Note that this definition includes 'seasonal endothermy' as seen in some tegu lizards (Tattersall et al. 2016) and pythons (Brashears & DeNardo 2013), as well as 'regional endothermy' as seen in tunniforms (Carey 1969, Lawson et al. 2010), opah (Wegner et al. 2015), lamniform sharks (Fritsches et al. 2005), and mobulid

rays (Alexander 1996). This also includes ‘gigantothermy’, which seems to be a by-product of increased body heat generated by muscular thermogenesis through constant flapping as seen in leatherback turtles (Burns et al. 2015, Sato 2014), that is also associated with mechanisms, such as high thermal inertia and counter-current heat exchange structures, to avoid heat loss (e.g., Wallace and Jones 2008). Importantly, the naked mole rat, *Heterocephalus glaber*, and *Xenospermophilus tereticaudus* are here considered to be the only mammalian ectotherms, because their body temperature follows environmental temperature (Wooden and Walsberg 2002, Woodley & Buffenstein 2002) despite their ability to generate their own heat via regulated non-shivering thermogenesis (see definition in IUPS thermal commission 2003).

Note that most small mammals species are ectothermic at birth and during early post-natal development, because their small size and an incompletely developed pelage drastically increase their thermal conductance (Dzal and Milsom 2021).

Thermoregulatory regime – the condition of being an ectotherm or an endotherm.

Low body temperature endotherm – endotherm with body temperature between 31-35°C. This definition is essentially that of Lovegrove’s (2019) ‘basoendotherm’, except that we define a lower bound for the temperature range, as the lowest body temperature ever recorded for an active endotherm. This range of body temperatures excludes ~75% of all sampled mammals and all birds.

Low body temperature ectotherm – ectotherm with body temperature ≤25°C. Following this definition, all amphibians are included as well as ~75% of turtles (not tortoises), ~25% of lepidosaurs, and ~25% crocodilians.

High body temperature ectotherm – ectotherm with body temperature ≥32°C. Following this definition, ~25% of lepidosaurs, and ~25% crocodilians are included.

Importantly, some of these definitions are different from those of the IUPS thermal commission (2003). These are working definitions explicitly stating what we mean when using terms such as ‘endothermy’ or ‘ectothermy’. For instance, our definitions include the notion that the body temperatures of endotherms and ectotherms should be compared during periods of activity, which is not implied in the IUPS thermal commission (2003) definition: “The pattern of thermoregulation in which the body temperature depends on a high (tachymetabolic) and controlled rate of heat production”.

References

Alexander, R.L., 1996. Evidence of brain-warming in the mobulid rays, *Mobula tarapacana* and Manta birostris (Chondrichthyes: Elasmobranchii: Batoidea: Myliobatiformes). Zoological Journal of the Linnean Society, 118(2), pp.151-164.

Burns, T.J., McCafferty, D.J. and Kennedy, M.W., 2015. Core and body surface temperatures of nesting leatherback turtles (Dermochelys coriacea). Journal of thermal biology, 51, pp.15-22.

Brashears JA, DeNardo DF. 2013 Revisiting python thermogenesis: brooding Burmese pythons (*Python bivittatus*) cue on body, not clutch,temperature. J. Herpetol. 47, 440–444. (doi:10.1670/12-050)

Carey, F. G., & Teal, J. M. (1969). Regulation of body temperature by the bluefin *tuna. Comparative Biochemistry and Physiology, 28(1), 205-213*.

Dzal, Y.A. and Milsom, W.K., 2021. Effects of hypoxia on the respiratory and metabolic responses to progressive cooling in newborn rodents that range in heterothermic expression. Experimental Physiology, 106(4), pp.1005-1023.

Fritsches, K.A., Brill, R.W. and Warrant, E.J., 2005. Warm eyes provide superior vision in swordfishes. Current Biology, 15(1), pp.55-58.

Katsufumi Sato; Body temperature stability achieved by the large body *mass of sea turtles*. J Exp Biol *15 October 2014; 217 (20): 3607–3614*

IUPS Thermal Commission, 2003. Glossary of terms for thermal physiology. Journal of Thermal Biology 28, 75–106.

Lawson, G. L., Castleton, M. R., & Block, B. A. 2010. Movements and diving behavior of Atlantic bluefin tuna Thunnus thynnus in relation to water column structure in the northwestern Atlantic. Marine Ecology Progress Series, 400, 245-265.

Lovegrove, B.G., 2019. Fires of Life. Yale University Press.

Nicol, S. and Andersen, N.A., 2006. Body temperature as an indicator of egg-laying in the echidna, Tachyglossus aculeatus. Journal of Thermal Biology, 31(6), pp.483-490.

Tattersall, G.J., Leite, C.A., Sanders, C.E., Cadena, V., Andrade, D.V., Abe, A.S. and Milsom, W.K., 2016. Seasonal reproductive endothermy in tegu lizards. Science Advances, 2(1), p.e1500951.

Terrien, J., Perret, M. and Aujard, F., 2011. Behavioral thermoregulation in mammals: a review. Front Biosci, 16(4), pp.1428-1444.

Wallace, B.P. and Jones, T.T., 2008. What makes marine turtles go: a review of metabolic rates and their consequences. Journal of Experimental Marine Biology and Ecology, 356(1-2), pp.8-24.

Wegner, N.C., Snodgrass, O.E., Dewar, H. and Hyde, J.R., 2015. Whole-body endothermy in a mesopelagic fish, the opah, *Lampris guttatus*. Science, 348(6236), pp.786-789.

427 Wooden, K.M. and Walsberg, G.E., 2002. Effect of environmental temperature on body
428 temperature and metabolic heat production in a heterothermic rodent, *Spermophilus tereticaudus*. Journal
429 of Experimental Biology, 205(14), pp.2099-2105.

430 Woodley, R., & Buffenstein, R. (2002). Thermogenic changes with chronic cold exposure in the
431 naked mole-rat (*Heterocephalus glaber*). Comparative Biochemistry and Physiology Part A: Molecular &
432 Integrative Physiology, 133(3), 827-834.

Supplementary Note 2. Description of the statistical analyses.

Abbreviations

Body mass: **BM**
Condylobasal length: **CBL**
Condylo-anteroorbital length: **CAOL**
3D Bony dataset: **3DB**
3D Membranous dataset: **3DM**
2D Membranous dataset: **2DM**
Anterior canal/duct: **_a**
Posterior canal/duct: **_p**
Lateral canal/duct: **_l**
Bony semicircular canals: **_B**
Radius of curvature: **R**
Shape factor of the torus: **σ** (see Supplementary Methods, Biomechanics)
Eccentricity of the torus: **e**
Enclosed area of the torus: **Λ**
Cross-sectional area of the middle section of the cupula: **aC_ _Md**
Average thickness of the crista ampullaris: **tCr**
Cross-sectional radius of the slender portion: **rS**
Length of the slender portion: **LS**
Cross-sectional area of the slender portion of the membranous semicircular duct: **aS**
Average height of the cilia: **hc**
Average thickness of the cupula: **tC**
Cilia deflection factor: **E** (see Supplementary Methods, Biomechanics)
Temperature term: $\mathbf{T_R}$ (see Supplementary Methods, Biomechanics)
Body temperature: $\mathbf{T_B}$
Upper corner frequency: $\boldsymbol{\omega_2}$ (see Supplementary Methods, Biomechanics)
Saturating angular velocity: $\boldsymbol{\Omega_2}$ (see Supplementary Methods, Biomechanics)
Thermo-Motility Index: **TMI** (see Supplementary Methods, Biomechanics)

Thermoregulatory regime: **RT**

Phylogenetic generalized least square regression: **PGLS-R**

Residual duct-morphology component of the Thermo-Motility Index: **_res**

Fitted duct-morphology component of the Thermo-Motility Index: **_fitted**

Physiochemical factor of the endolymph part of the Thermo-Motility Index: **_endolymph**

Values before size adjustment: **_raw**

Values after size adjustment: **_adjusted**

Standard average of all Thermo-Motility Indices: **_unweighted**

Values of specimens from the 3D Membranous dataset estimated as done for fossils: **_estimated**

Maximum anaerobic speed: **max_speed**

Maximum aerobic speed: **max_speed_aero**

Body size category: **Size**

Stride frequency: $\mathbf{F_{stride}}$

Impact frequency: $\mathbf{F_{impact}}$

Note on assumption testing

Tested assumptions of the following phylogenetic generalised least square regressions (PGLS-R) and phylogenetic ANOVAs are explained in detail in Mundry (2014) and on the blog of Liam Revell http://blog.phytools.org/2013/02/a-comment-on-distribution-of-residuals.html. For all PGLS-R and phylogenetic logistic regressions we assume that evolution follows a Brownian Motion model with branch lengths corrected with a Pagel's λ estimated through maximum likelihood. Assumptions of the phylogenetic logistic regression follow Stoltzfus (2011).

1. PGLS-R of the average height of cilia located on the cristae ampullares against average thicknesses of the cristae ampullares and cross-sectional areas of corresponding middle sections of cupulae

Models 1-3 provide very significant regressions ($P << 0.001$), with high coefficients of determination (R^2 = 0.66–0.71). They were tested on specimens for which we have the membranous labyrinth, and were used to predict the average height of cilia (hc) located on the anterior, posterior, and lateral crista ampullaris of *Etroplus maculatus*, *Poecilia sphenops*, *Steatocranus tinanti*, *Rana pipiens*, *Eleutherodactylus limbatus*, *Xenopus laevis*, and *Trachemys scripta*, for which we lacked this information. The cross-sectional area of the middle section of the cupula (aC_Md) and the thickness of the crista ampullaris (tCr) of each membranous semicircular duct were used as predictors. Predicted cilia height are used in Models 4-6 to predict the cilia deflection factor of the same species.

Model 1: $\log_{10}$(hc_a) ~ $\log_{10}$(aC_a_Md)

Branch length transformations:

Pagel's λ [ML]: 0.880

lower bound: 0.000, $P = 3.682810^{-05}$ *

upper bound: 1.000, $P = 0.021109$ *

95.0% CI: (0.391, 0.996)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.796632	0.053918	33.321	$< 2.2\ 10^{-16}$ *
$\log_{10}$(aC_a_Md)	0.386736	0.038028	10.170	$8.928\ 10^{-13}$ *

P values calculated using the exact method.

Adjusted R^2: 0.7092, AICc: -97.89501

F statistic: 103.4 on 1 and 41 DF

n = 43 (3DM set where only specimens with measurements of height of cilia areas were included)

P value: $8.929\ 10^{-13}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98655, P = 0.8882, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S1) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.39634, U = 0.86002, P = 0.2892, P value computed via approximation method).

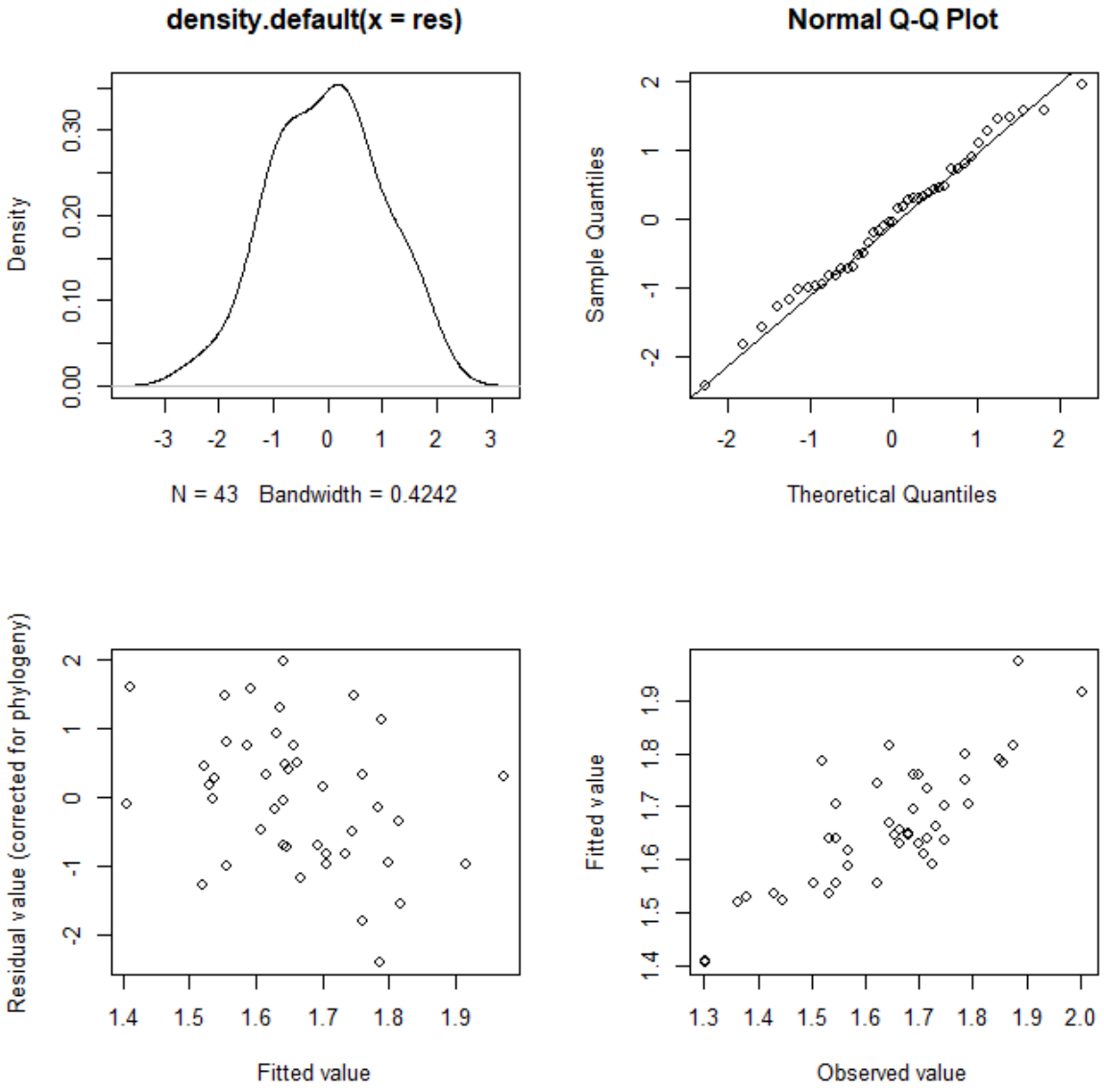


Figure S1 – Diagnostic plots for model 1.

Model 2: $\log_{10}$(hc_p) ~ $\log_{10}$(aC_p_Md)

Branch length transformations:

Pagel's λ [ML]: 1.000

lower bound: 0.000, *P* = 0.00012739 *

upper bound: 1.000, *P* = 1

95.0% CI: (0.224, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.798244	0.067019	26.8318	$< 2.2\ 10^{-16}$ *
$\log_{10}$(aC_p_Md)	0.367619	0.039909	9.2114	$1.545\ 10^{-11}$ *

P values calculated using the exact method.

Adjusted R^2: 0.6663, AICc: -98.1533

F statistic: 84.85 on 1 and 41 DF

n = 43 (3DM set where only specimens with measurements of height of cilia areas were included)

P value: $1.545\ 10^{-11}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.98274, *P* = 0.7549, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S2) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.35323, *U* = 0.86501, *P* = 0.3303, *P* value computed via approximation method).

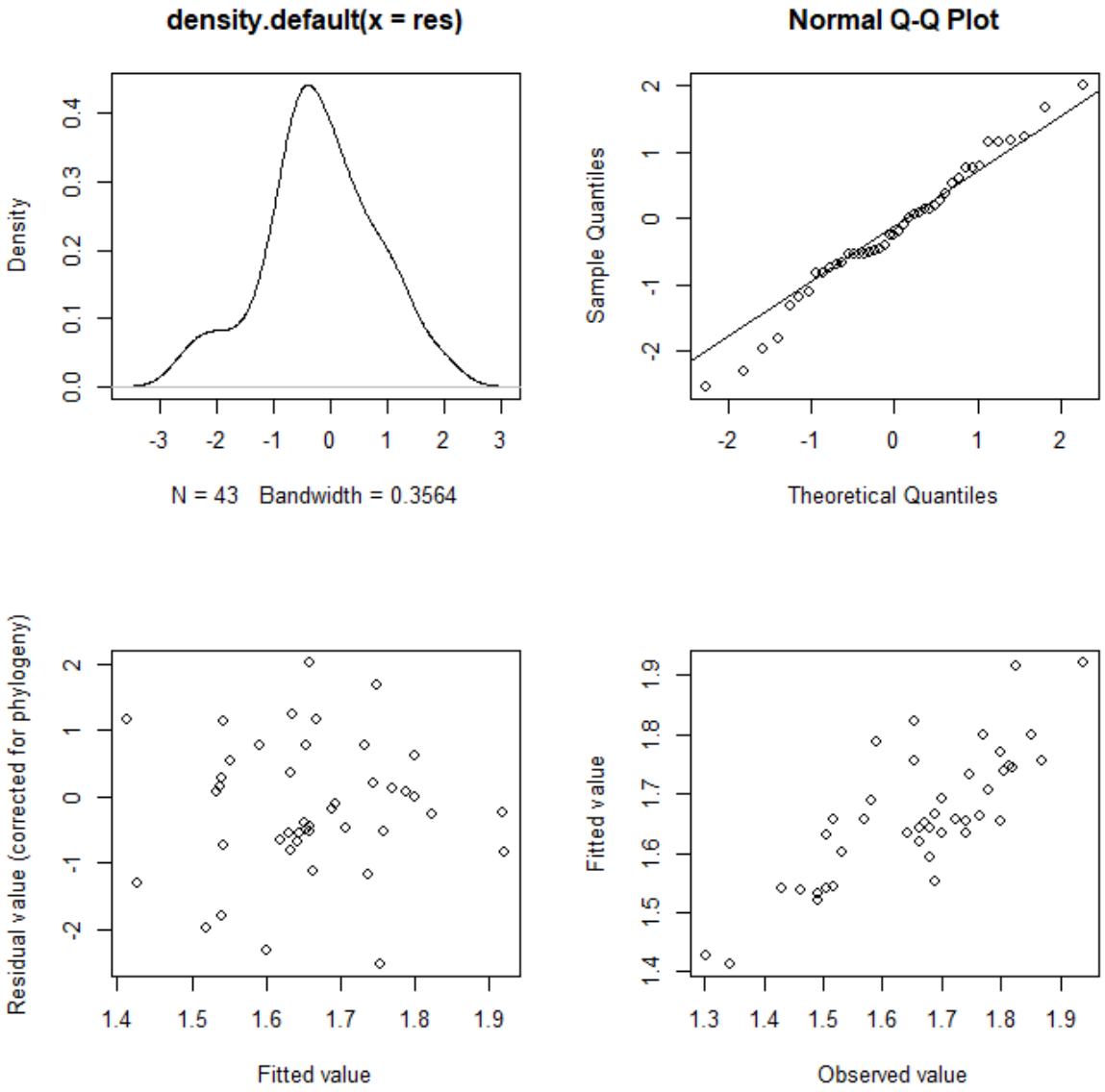


Figure S2 – Diagnostic plots for model 2.

Model 3: $\log_{10}$(hc_l) ~ $\log_{10}$(aC_l_Md) + $\log_{10}$(tCr_l)

Branch length transformations:

Pagel's λ [ML]: 0.903

lower bound: 0.000, $P = 1.9025^{-05}$ *

upper bound: 1.000, $P = 0.0052485$ *

95.0% CI: (0.408, 0.990)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.903657	0.072589	26.2252	$< 2.2\ 10^{-16}$ *
$\log_{10}$(aC_l_Md)	0.140197	0.079271	1.7686	0.084787
$\log_{10}$(tCr_l)	0.369391	0.136356	2.7090	0.009972 *

P values calculated using the exact method.

Adjusted R^2: 0.6619, AICc: -96.83884

F statistic: 41.14 on 2 and 39 DF

n = 42 (3DM set where only specimens with measurements of height of cilia areas were included)

As the initial Q-Q plot and Shapiro-Wilk test (*W* = 0.94825, *P* = 0.05149, *P* value calculated via approximation method) both indicated potential non-normality of the phylogenetic residuals (left skewed), the most influential specimen (*Columba palumbus*) was excluded from this regression.

P value: $2.467\ 10^{-10}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.98125, *P* = 0.7097, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S3) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.36193, *U* = 0.86062, *P* = 0.3125, *P* value computed via approximation method).

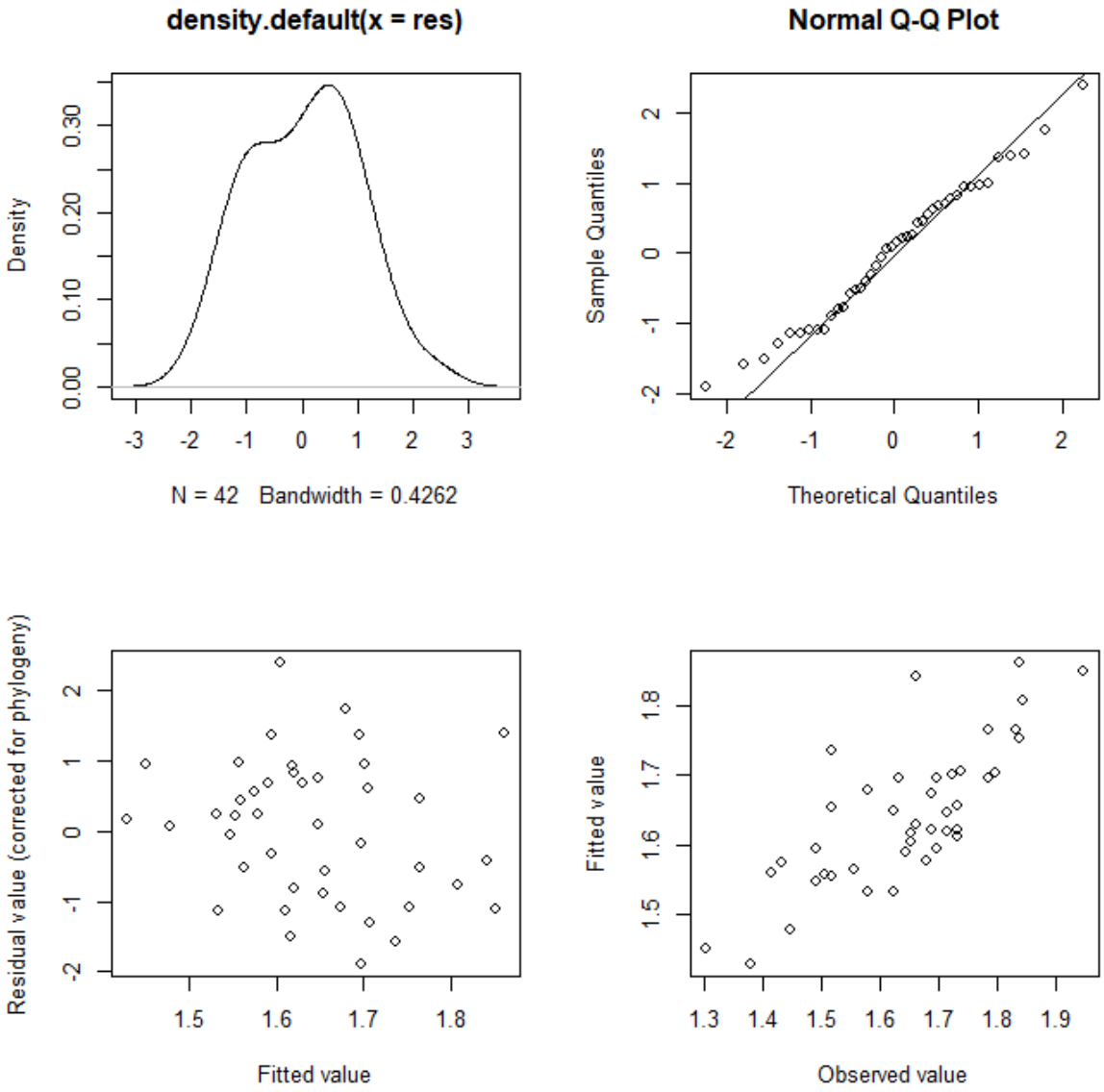


Figure S3 – Diagnostic plots for model 3.

2. PGLS-R of the cilia deflection factors against corresponding average thicknesses of cupulae models and cross-sectional areas of middle sections of cupulae

Models 4-6 provide very significant regressions ($P << 0.001$), with very high coefficients of determination ($R^2 = 0.99–1.00$). They were tested on specimens for which we have the membranous labyrinth, and were used to predict the cilia deflection factors (E, see Supplementary Methods, Biomechanics) of *Etroplus maculatus*, *Poecilia sphenops*, *Steatocranus tinanti*, *Rana pipiens*, *Eleutherodactylus limbatus*, *Xenopus laevis*, and *Trachemys scripta*, for which we lack this information. The cross-sectional area of the middle section of the cupula (aC_Md) and the average thickness of cupula models (tC) of each membranous semicircular duct were used jointly as predictors. Predicted cilia deflection factors are used to calculate 'raw' Thermo-Motility Indices of these species (TMI: see Supplementary Methods, Biomechanics).

Model 4: $\log_{10}(E\_a) \sim \log_{10}(aC\_a\_Md) + \log_{10}(tC\_a)$

Branch length transformations:

Pagel's λ [ML]: 0.975

lower bound: 0.000, $P = 2.4059\ 10^{-13}$ *

upper bound: 1.000, $P = 0.32699$

95.0% CI: (0.775, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.430798	0.039871	-10.8048	$1.976\ 10^{-13}$ *
$\log_{10}$(aC_a_Md)	-1.717530	0.042586	-40.3304	$< 2.2\ 10^{-16}$*
$\log_{10}$(tC_a)	0.421972	0.077029	5.4781	$2.555\ 10^{-06}$*

P values calculated using the exact method.

Adjusted R^2: 0.9928, AICc: -154.2394

F statistic: 2915 on 2 and 40 DF

n = 43 (3DM set where only specimens with measurements of cilia deflection factors were included)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98424, P = 0.8111, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S4) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.59937, U = 0.83529, P = 0.149, P value computed via approximation method).

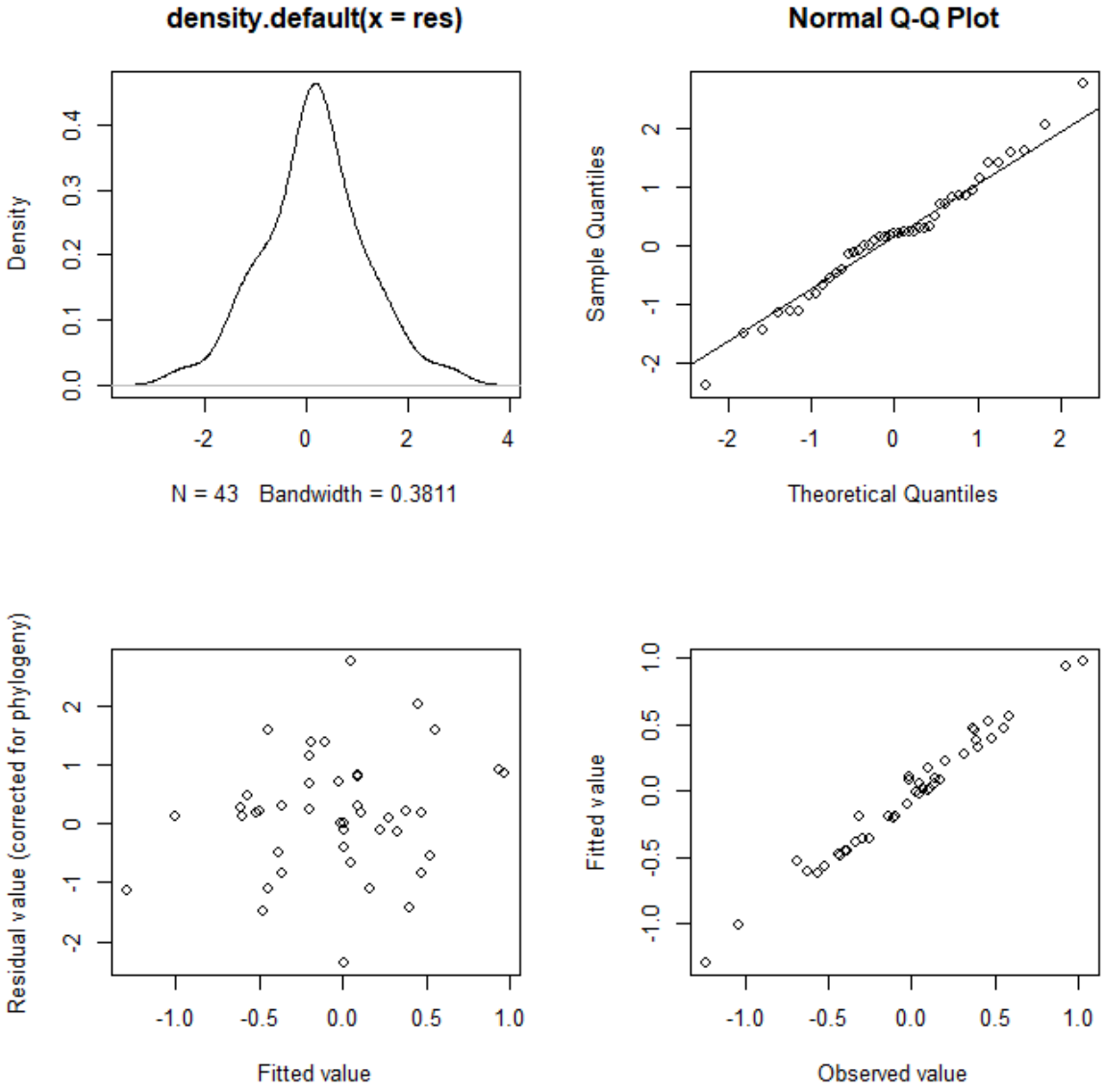


Figure S4 – Diagnostic plots for model 4.

Model 5: $\log_{10}(E\_p) \sim \log_{10}(aC\_p\_Md) + \log_{10}(tC\_p)$

Branch length transformations:

Pagel's λ [ML]: 0.953

lower bound: 0.000, $P = 1.5454\ 10^{-13}$ *

upper bound: 1.000, $P = 0.082336$

95.0% CI: (0.734, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.407967	0.039456	-10.3398	$9.850\ 10^{-13}$*
$\log_{10}$(aC_p_Md)	-1.737006	0.044106	-39.3822	$< 2.2\ 10^{-16}$ *
$\log_{10}$(tC_p)	0.476793	0.082155	5.8036	$9.707\ 10^{-07}$ *

P values calculated using the exact method.

Adjusted R^2: 0.9926, AICc: -148.5563

F statistic: 2739 on 2 and 39 DF

n = 42 (3DM set where only specimens with measurements of cilia deflection factors were included)

As the initial Q-Q plot and Shapiro-Wilk test (*W* = 0.95344, *P* = 0.07969, *P* value calculated via approximation method) both indicated potential non-normality of the phylogenetic residuals, the most influential specimen (*Sus scrofa*) was excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.97724, *P* = 0.5571, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S5) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.04672, *U* = 0.89534, *P* = 0.7691, *P* value computed via approximation method).

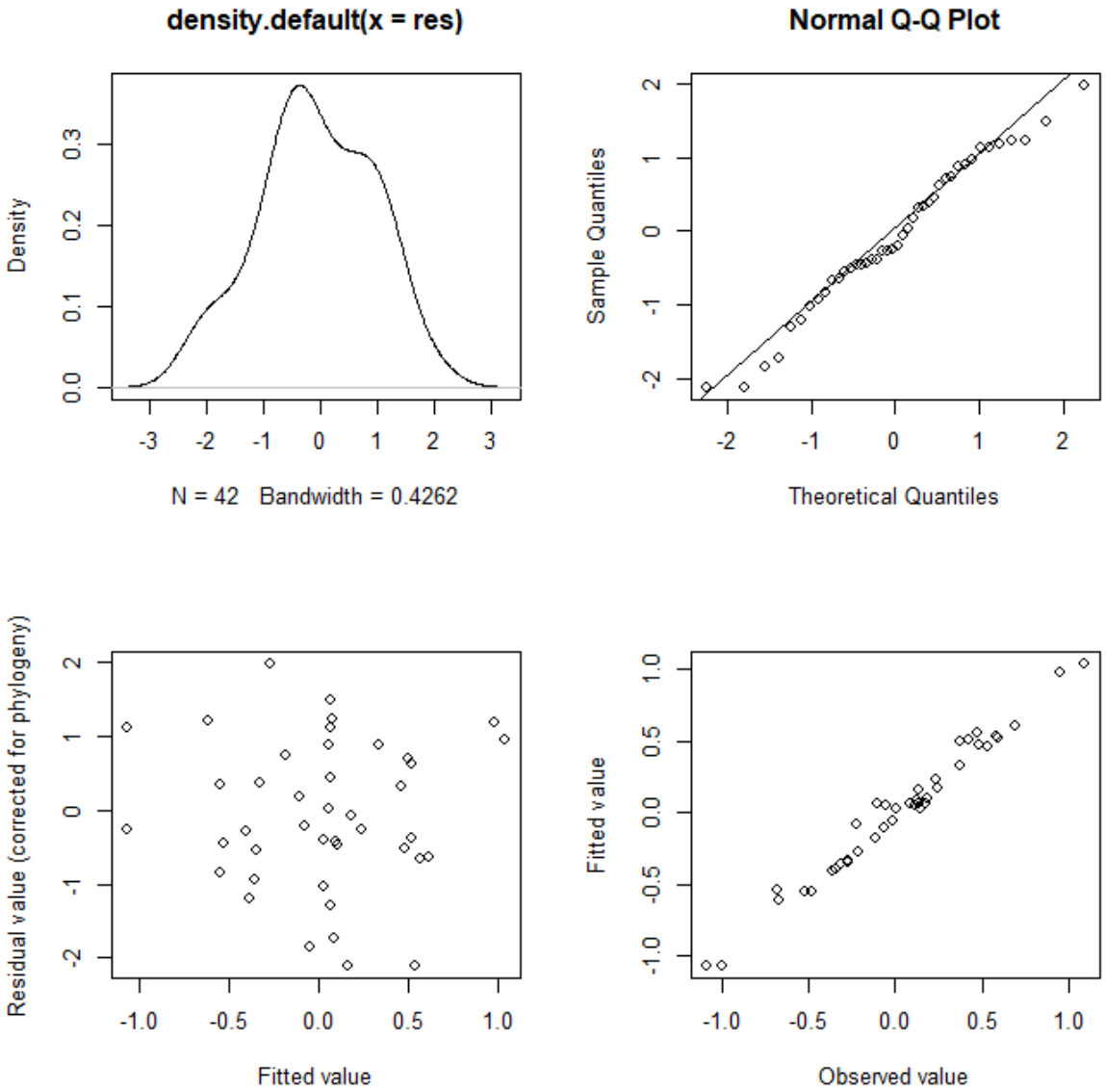


Figure S5 – Diagnostic plots for model 5.

Model 6: $\log_{10}(E\_l) \sim \log_{10}(aC\_l\_Md) + \log_{10}(tC\_l)$

Branch length transformations:

Pagel's λ [ML]: 0.816

lower bound: 0.000, $P = 6.9795\ 10^{-07}$ *

upper bound: 1.000, $P = 0.00028796$ *

95.0% CI: (0.394, 0.971)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.416932	0.024222	-17.2129	$< 2.2\ 10^{-16}$*
$\log_{10}$(aC_l_Md)	-1.654916	0.031048	-53.3011	$< 2.2\ 10^{-16}$*
$\log_{10}$(tC_l)	0.363648	0.057953	6.2749	$1.941\ 10^{-07}$ *

P values calculated using the exact method.

Adjusted R^2: 0.9962, AICc: -183.9703

F statistic: 5547 on 2 and 40 DF

n = 43 (3DM set where only specimens with measurements of cilia deflection factors were included)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.96647$, $P = 0.2378$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S6) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 1.90392$, $U = 0.91164$, $P = 0.91164$, *P* value computed via approximation method).

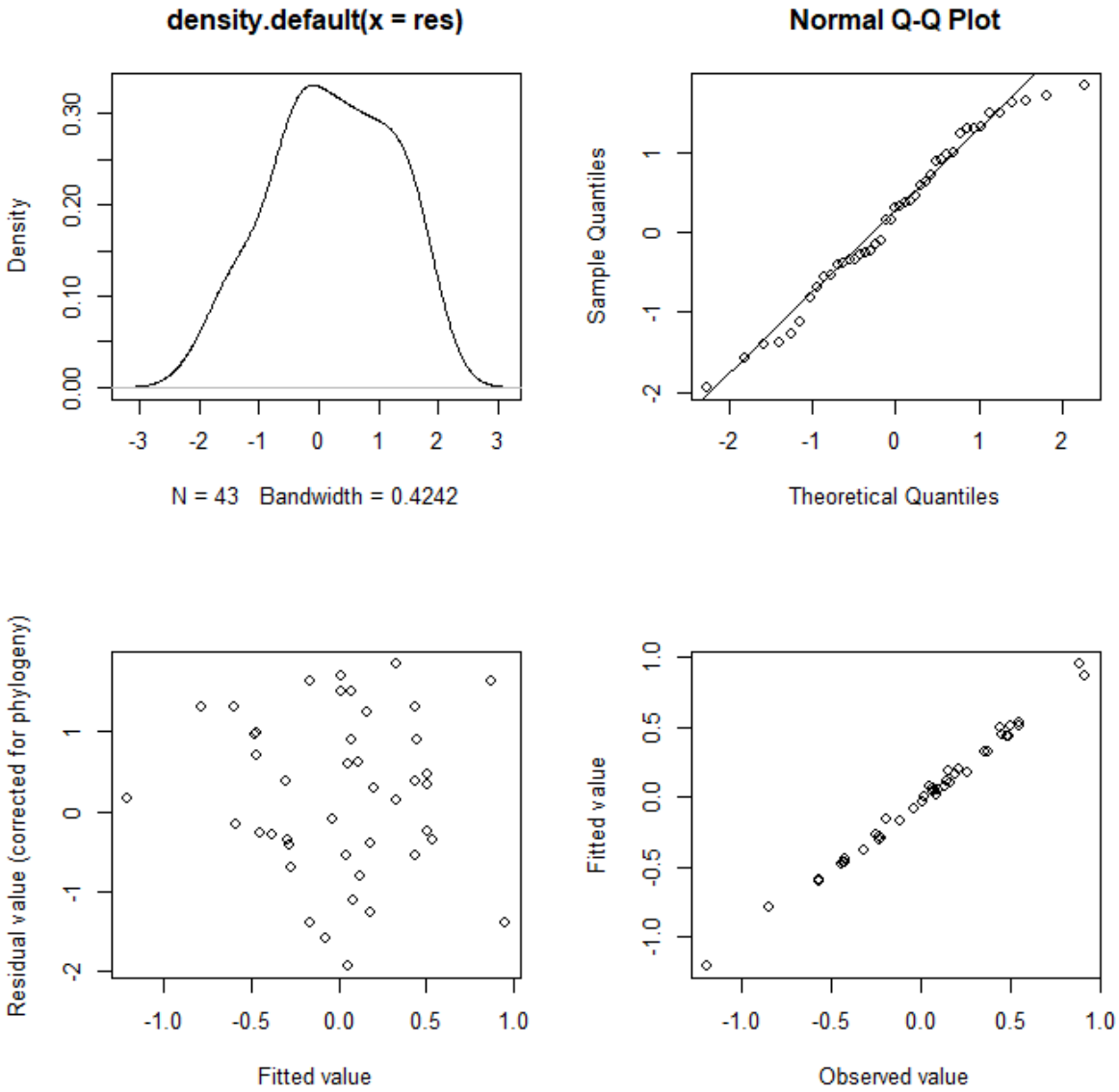


Figure S6 – Diagnostic plots for model 6.

3. PGLS-R of body mass against condylobasal and condylo-anteroorbital lengths

Model 7 provides a very significant regression ($P << 0.001$), with a very high coefficient of determination ($R^2 = 0.89$). It was tested on extant specimens for which we know the body mass (BM), and was used to predict it for fossils and some extant specimens which did not preserve soft tissues. The condylobasal length (CBL) and the condylo-anteroorbital length (CAOL) were jointly used as predictors. While condylobasal length remains a good predictor of body mass when used alone, it can lead to erroneous results in longirostrine taxa (e.g., gharial) and taxa with short snouts (e.g., humans). The condylo-anteroorbital length is not affected by these issues and helps improving predictions. Predicted body masses are used in models 20-25 to size-correct fitted duct-morphology components of Thermo-Motility Indices (TMI_fitted_ω_2 and TMI_fitted_Ω_2, see Supplementary Methods, Biomechanics); in model 38 to assess whether the (size-corrected weighted average) Thermo-Motility Index (TMI, see Supplementary Methods, Biomechanics) is indeed uncorrelated to body mass; and in model 49 to size-correct the radius of curvature of the anterior canal of tetrapod specimens.

Model 7: $\log_{10}(BM) \sim \log_{10}(CBL) + \log_{10}(CAOL)$

Branch length transformations:

Pagel's λ [ML]: 0.869

lower bound: 0.000, $P = < 2.2\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.2\ 10^{-16}$ *

95.0% CI: (0.739, 0.942)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-2.44591	0.22301	-10.9676	$< 2.2\ 10^{-16}$ *
$\log_{10}$(CBL)	1.15316	0.21851	5.2774	$3.036\ 10^{-7}$ *
$\log_{10}$(CAOL)	2.19446	0.24502	8.9561	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.89, AICc: 146.2422

F statistic: 935.9 on 2 and 229 DF

n = 232 (3DB set without specimens lacking body mass information)

A potential outlier (*Phascolarctos cinereus*, G = 3.49303, U = 0.94718, P = 0.0472 *, P value computed via approximation method) was excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.99688, P = 0.9291, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S7) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.95698, U = 0.96198, P = 0.3317, P value computed via approximation method).
- We tested whether phylogeny structured the distribution of the ($\log_{10}$) body mass residuals. We concluded that phylogeny does partially structure the distribution (K = 0.360846, P value $< 1\ 10^{-4}$ *, P value calculated via approximation method) and should be taken into account to improve predictions in fossil species.

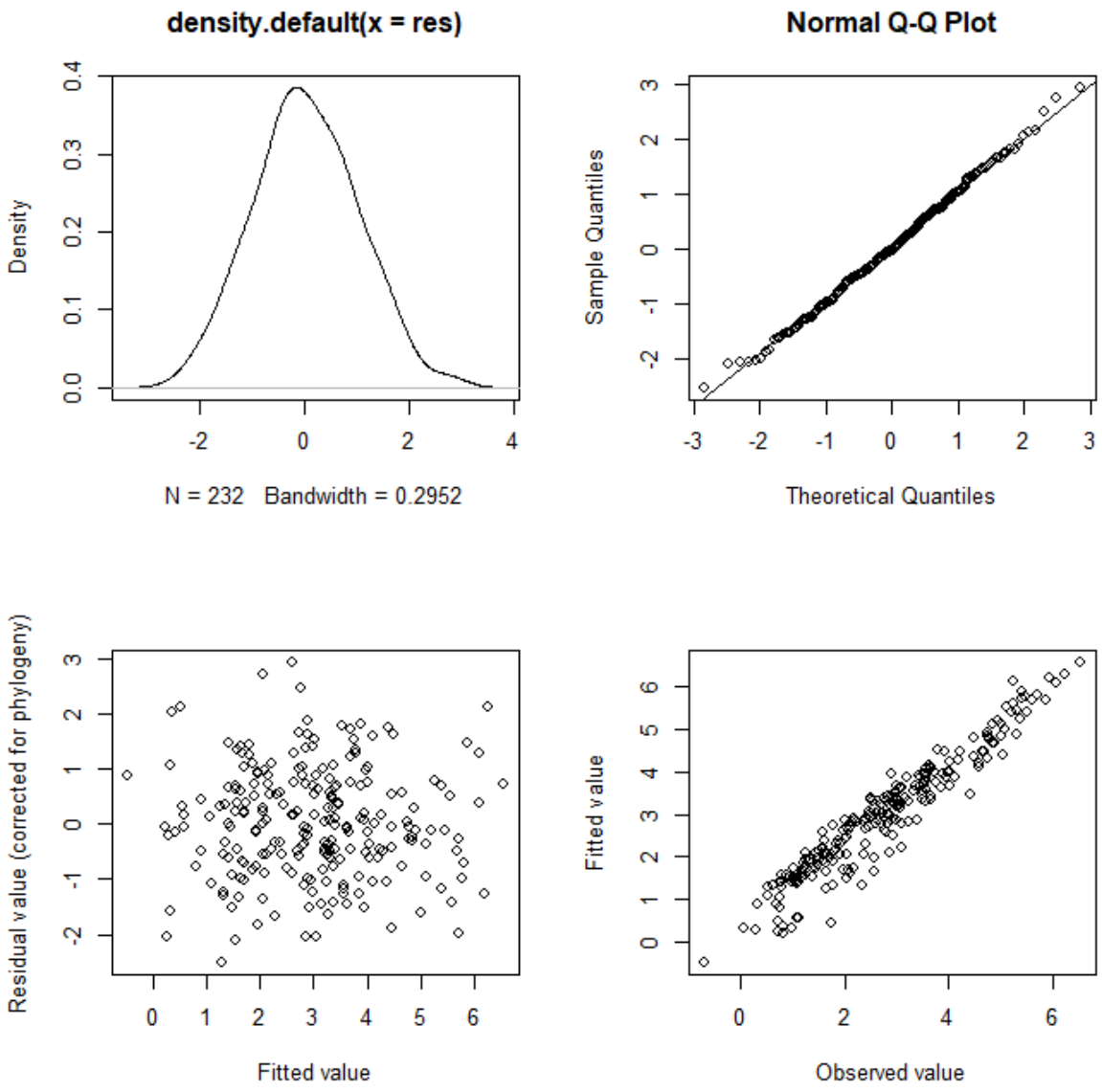


Figure S7 – Diagnostic plots for model 7.

4. PGLS-R of the cross-sectional areas of slender portions of membranous semicircular ducts against cross-sectional radii of slender portions and radii of curvature of corresponding bony semicircular canals

Models 8-10 provide very significant regressions ($P << 0.001$), with very high coefficients of determination ($R^2 = 0.81–0.88$). They were tested on specimens for which we have the bony and membranous labyrinths, and were used to predict fitted values for cross-sectional areas of slender semicircular duct portions (aS, see Supplementary Methods, Biomechanics) of extant and extinct species for which only the bony labyrinth was available (e.g., fossil specimens). The cross-sectional radius of the slender portion (rS_B) and the radius of curvature (R_B) of each bony semicircular canal were jointly used as predictors. Predicted fitted cross-sectional areas of slender semicircular duct portions are used to calculate fitted duct-morphology components of Thermo-Motility Indices (TMI_fitted_ω_2 and TMI_fitted_Ω_2, see Supplementary Methods, Biomechanics).

Model 8: $\log_{10}$(aS_a) ~ $\log_{10}$(rS_a_B) + $\log_{10}$(R_a_B)

Branch length transformations:

Pagel's λ [ML]: 0.462

lower bound: 0.000, p = 0.17214

upper bound: 1.000, p = 0.0034888 *

95.0% CI: (NA, 0.930)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.53842	0.12079	-4.4574	$5.65\ 10^{-05}$ *
$\log_{10}$(rS_a_B)	1.13972	0.12599	9.0463	$1.34\ 10^{-11}$ *
$\log_{10}$(R_a_B)	0.17252	0.12698	1.3587	0.1812

P values calculated using the exact method.

Adjusted R^2: 0.8803, AICc: -77.03506

F statistic: 170.1 on 2 and 44 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

p-value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.9701, P = 0.2672, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S8) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.44283, U = 0.86745, P = 0.2792, P value computed via approximation method).

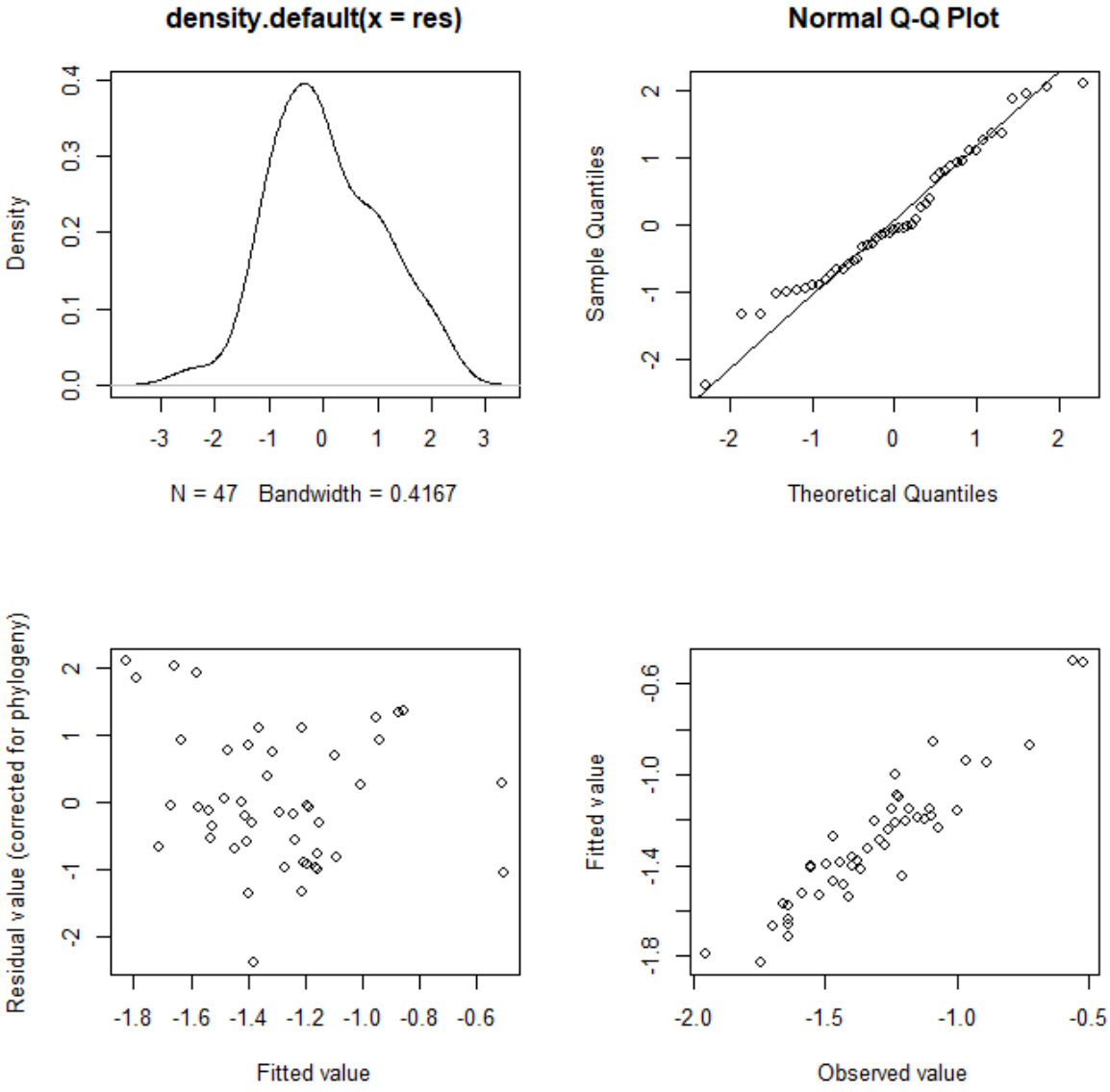


Figure S8 – Diagnostic plots for model 8.

941 Model 9: $\log_{10}(aS\_p) \sim \log_{10}(rS\_p\_B) + \log_{10}(R\_p\_B)$

942 Branch length transformations:

943 Pagel's λ [ML]: 0.560

944 lower bound: 0.000, *P* = 0.34471

945 upper bound: 1.000, *P* = 0.0002221 *

946 95.0% CI: (NA, 0.963)

947 *P* values approximated using a likelihood ratio test against the χ^2 distribution

948 Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.76282	0.12243	-6.2308	$1.69\ 10^{-07}$ *
$\log_{10}$(rS_p_B)	0.88531	0.13778	6.4254	$8.79\ 10^{-08}$ *
$\log_{10}$(R_p_B)	0.41666	0.13982	2.9800	0.004727 *

949 *P* values calculated using the exact method.

950 Adjusted R^2: 0.8712, AICc: -72.87318

951 *F* statistic: 153.2 on 2 and 43 DF

952 *n* = 46 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were
953 included)

954 A potential outlier (*Microcebus*, *G* = 3, *U* = 0.8001, *P* = 0.03827 *, *P* value computed via approximation
955 method) was excluded from this regression.

956 *P* value: $< 2.2\ 10^{-16}$ *

957 *P* value calculated using the exact method.

958 Assumption testing:

959 - We tested whether the distribution of phylogenetic residuals significantly differed from normal
960 using the Shapiro-Wilk test (*W* = 0.98692, *P* = 0.8789, *P* value calculated via approximation
961 method).
962 - Observations are independent from each other thanks to the phylogenetic comparative method.
963 - We visually checked whether the residuals were heteroscedastic (Fig. S9) and concluded they were
964 not.
965 - We looked for outliers using Grubb's test and found none (*G* = 2.76190, *U* = 0.82672, *P* = 0.09221,
966 *P* value computed via approximation method).

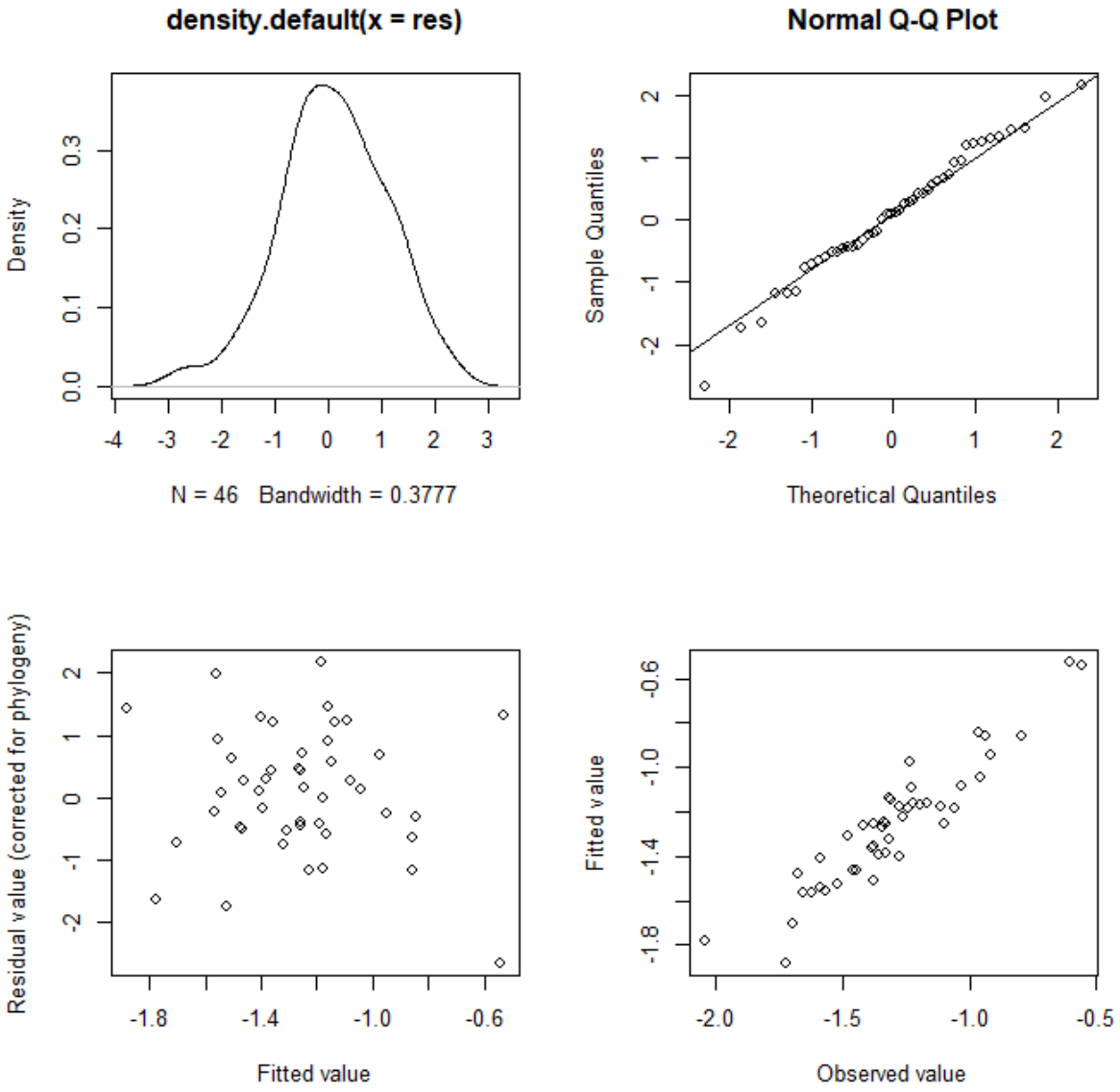


Figure S9 – Diagnostic plots for model 9.

Model 10: $\log_{10}(aS\_l) \sim \log_{10}(rS\_l\_B) + \log_{10}(R\_l\_B)$

Pagel's λ [ML]: 0.000

lower bound: 0.000, *P* = 1

upper bound: 1.000, *P* = 0.00039448*

95.0% CI: (NA, 0.852)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.86196	0.12270	-7.0250	$1.184\ 10^{-8}$ *
rS_l_B	0.83090	0.12797	6.4931	$7.005\ 10^{-8}$ *
R_l_B	0.29901	0.13568	2.2038	0.03294*

P values calculated using the exact method.

Adjusted R^2: 0.8128, AICc: -66.65833

F statistic: 98.72 on 2 and 43 DF

n = 46 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

A potential outlier (*Canis familiaris*, *G* = 3.03780, *U* = 0.79503, *P* = 0.03282 *, *P* value computed via approximation method) was excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.98655, *P* = 0.8882, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S10) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.46991, *U* = 0.86142, *P* = 0.2495, *P* value computed via approximation method).

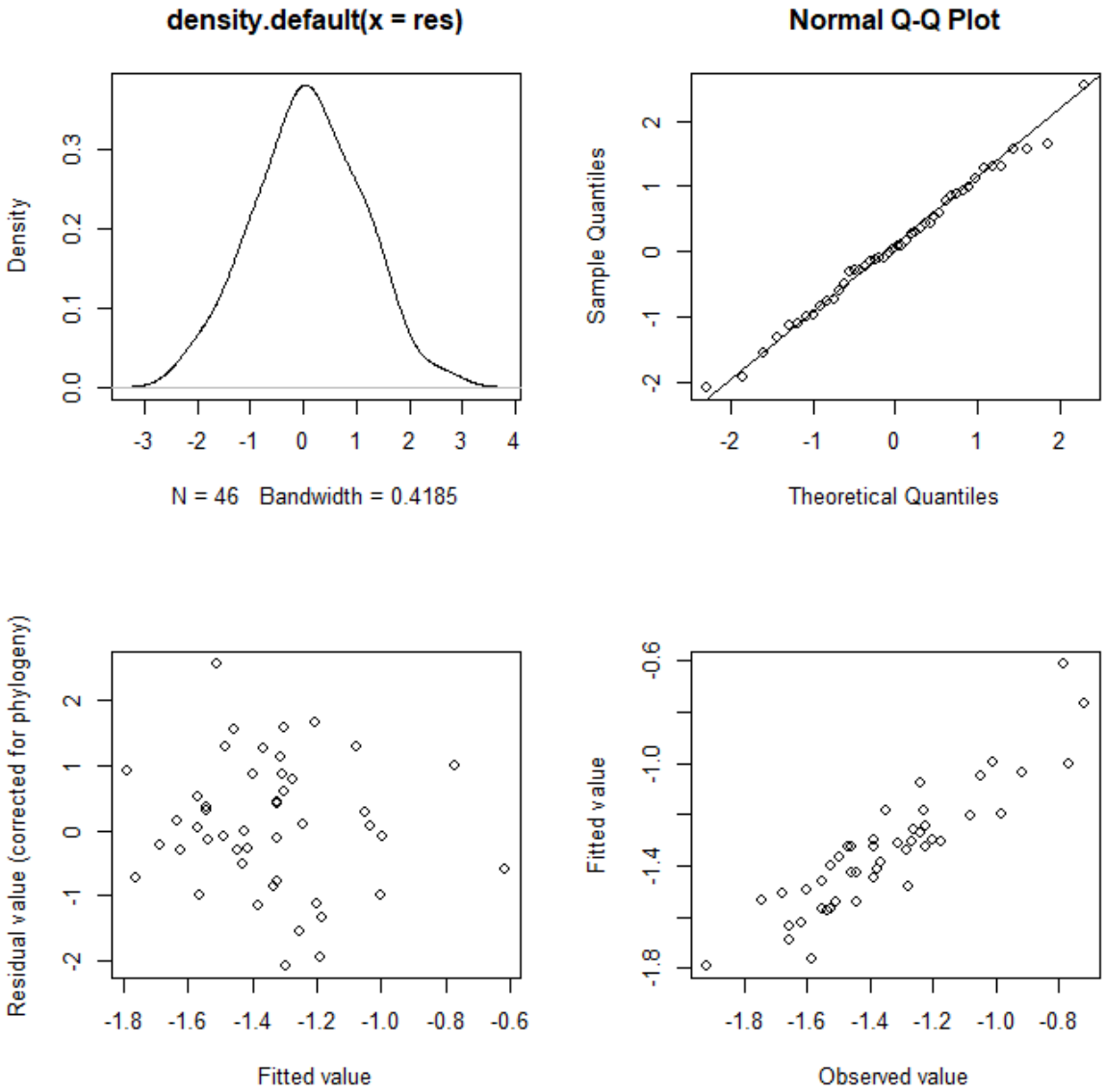


Figure S10 – Diagnostic plots for model 10.

5. PGLS-R of the enclosed areas of membranous semicircular ducts against shape factors and radii of curvature of corresponding bony semicircular canals

Models 11-13 provide very significant regressions (P << 0.001), with very high coefficients of determination (R^2 = 0.99–1.00). They were tested on specimens for which we have the bony and membranous labyrinths, and were used to predict fitted values for enclosed areas of membranous semicircular ducts (Λ, see Supplementary Methods, Biomechanics) of extant and extinct species for which only the bony labyrinth was available (e.g., fossil specimens). The shape factor (SF_B, see Supplementary Methods, Biomechanics) and the radius of curvature (R_B) of each bony semicircular canal were jointly used as predictors. Predicted fitted enclosed areas of membranous semicircular ducts are used to calculate fitted duct-morphology components of Thermo-Motility Indices (TMI_fitted_ω_2 and TMI_fitted_Ω_2, see Supplementary Methods, Biomechanics).

Model 11: $\log_{10}(\Lambda\_a) \sim \sigma\_a\_B + \log_{10}(R\_a\_B)$

Branch length transformations:
Pagel's λ [ML]: 0.225
  lower bound: 0.000, *P* = 0.47752
  upper bound: 1.000, *P* = 0.0020245 *
  95.0% CI: (NA, 0.952)
P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.107967	0.130305	-0.8286	0.4119225
σ_a_B	0.178141	0.044216	4.0289	0.0002242 *
$\log_{10}$(R_a_B)	2.046204	0.027267	75.0419	< 2.2 10^{-16} *

P values calculated using the exact method.

Adjusted R^2: 0.9921, AICc: -160.8835
F statistic: 2835 on 2 and 43 DF
n = 46 (3DM set where only specimens with measurements of radii of curvature of bony canals were included)
A potential outlier (*Tupaia javanica*, *G* = 3.0089, *U* = 0.7989, *P* = 0.03691 *, *P* value computed via approximation method) was excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.98014$, $P = 0.6115$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S11) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.36191$, $U = 0.87328$, $P = 0.3491$, *P* value computed via approximation method).

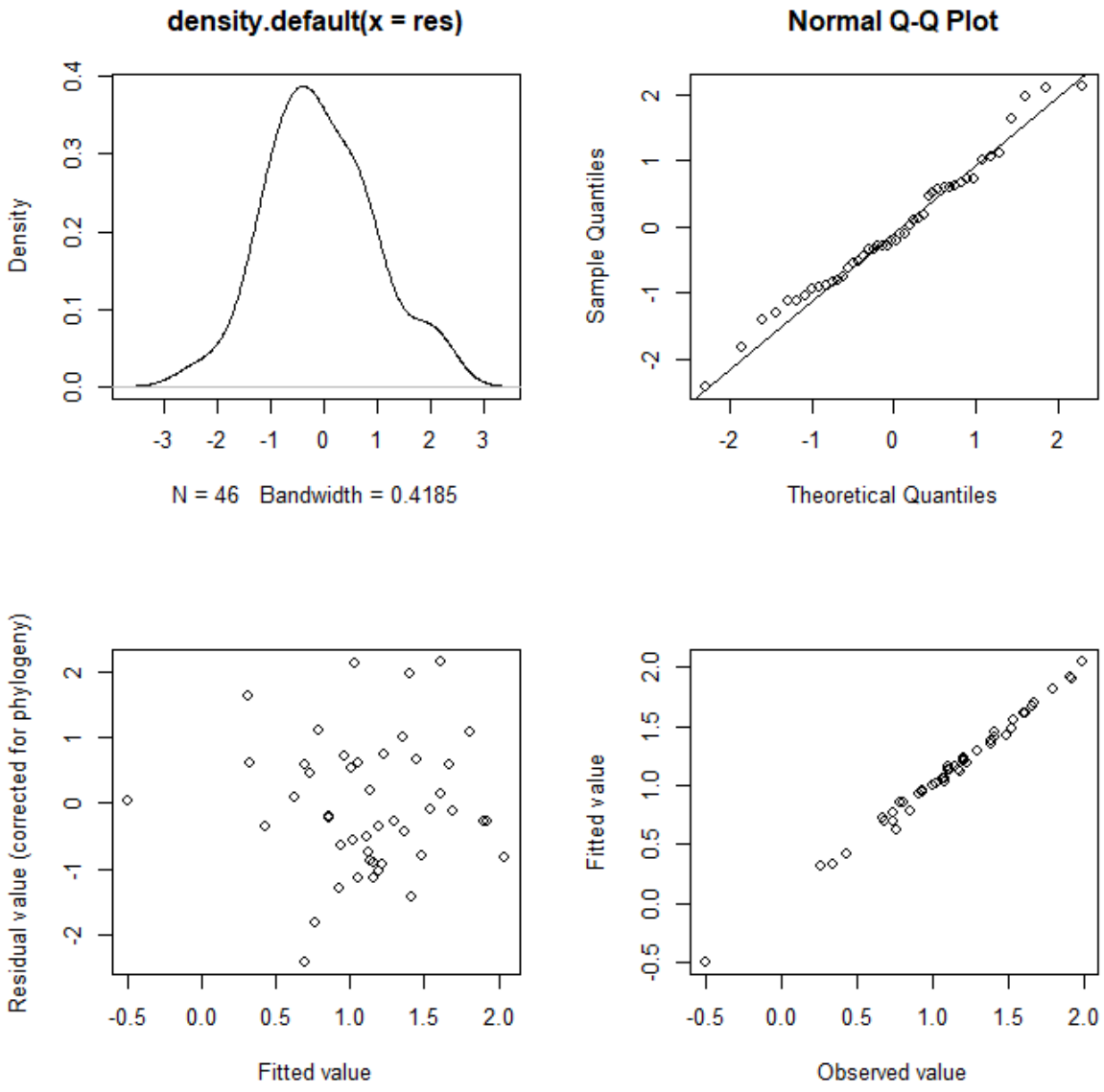


Figure S11 – Diagnostic plots for model 11.

Model 12: $\log_{10}(\Lambda\_p) \sim \sigma\_p\_B + \log_{10}(R\_p\_B)$

Branch length transformations:

Pagel's λ [ML]: 0.900

lower bound: 0.000, $P = 8.4217\ 10^{-06}$ *

upper bound: 1.000, $P = 0.0093308$ *

95.0% CI: (0.621, 0.991)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.179881	0.161946	-1.1107	0.2727128
σ_p_B	0.207927	0.055072	3.7755	0.0004749 *
$\log_{10}$(R_p_B)	2.031402	0.018735	108.4302	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.9962, AICc: -204.0711

F statistic: 6029 on 2 and 44 DF

n = 47 (3DM set where only specimens with measurements of radii of curvature of bony canals were included)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.97346$, $P = 0.3566$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S12) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.33144$, $U = 0.87927$, $P = 0.3926$, *P* value computed via approximation method).

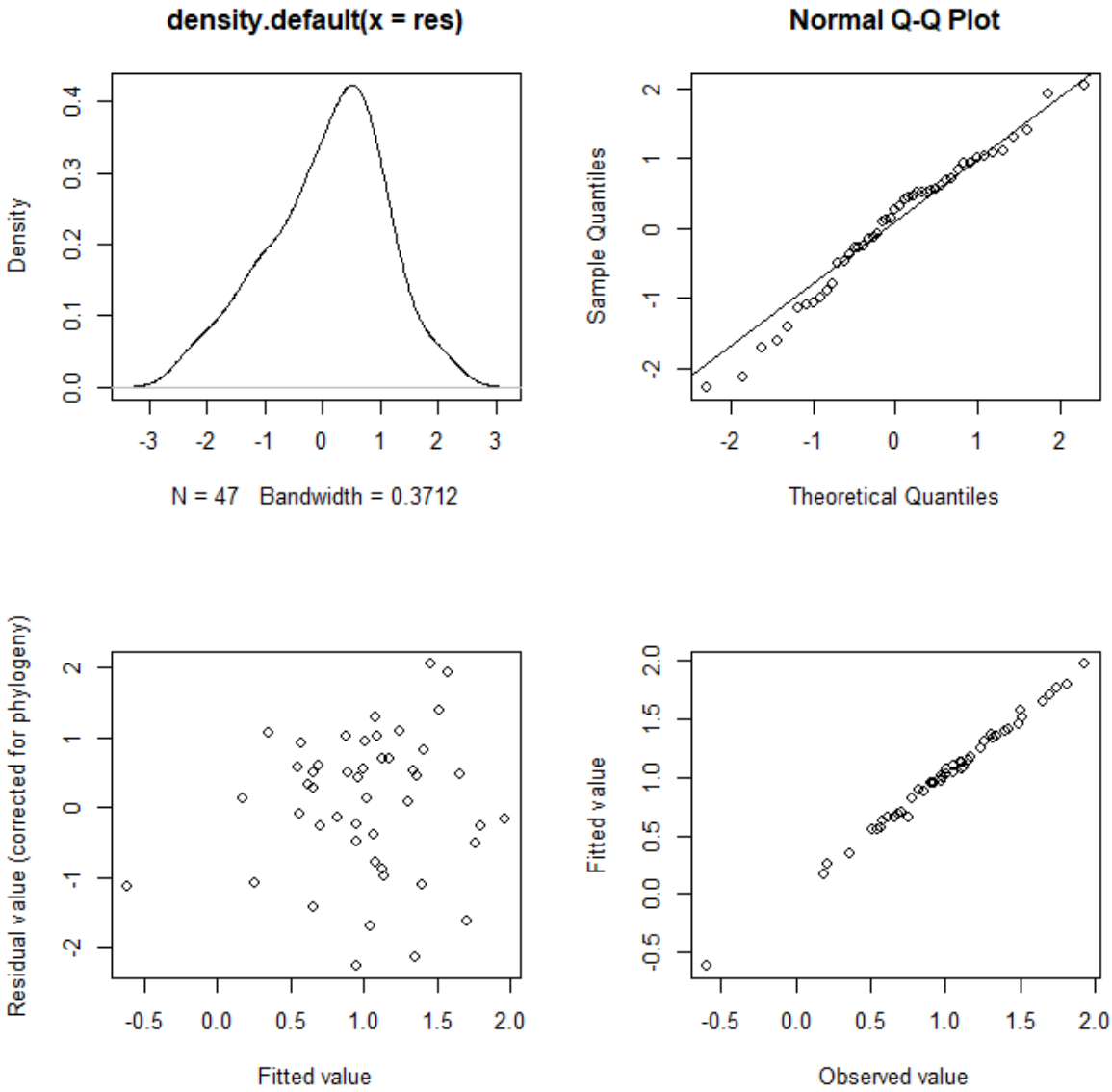


Figure S12 – Diagnostic plots for model 12.

Model 13: $\log_{10}(\Lambda\_l) \sim \sigma\_l\_B + \log_{10}(R\_l\_B)$

Branch length transformations:

Pagel's λ [ML]: 0.906

lower bound: 0.000, $P = 1.6521\ 10^{-08}$ *

upper bound: 1.000, $P = 0.068921$

95.0% CI: (0.663, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.030313	0.310458	0.0976	0.9227
σ_l_B	0.131409	0.100960	1.3016	0.1998
$\log_{10}$(R_l_B)	1.976542	0.023386	84.5190	$< 2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.9937, AICc: -186.4223

F statistic: 3601 on 2 and 44 DF

n = 47 (3DM set where only specimens with measurements of radii of curvature of bony canals were included)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.96817$, $P = 0.2255$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S13) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.2486$, $U = 0.8877$, $P = 0.5006$, *P* value computed via approximation method).

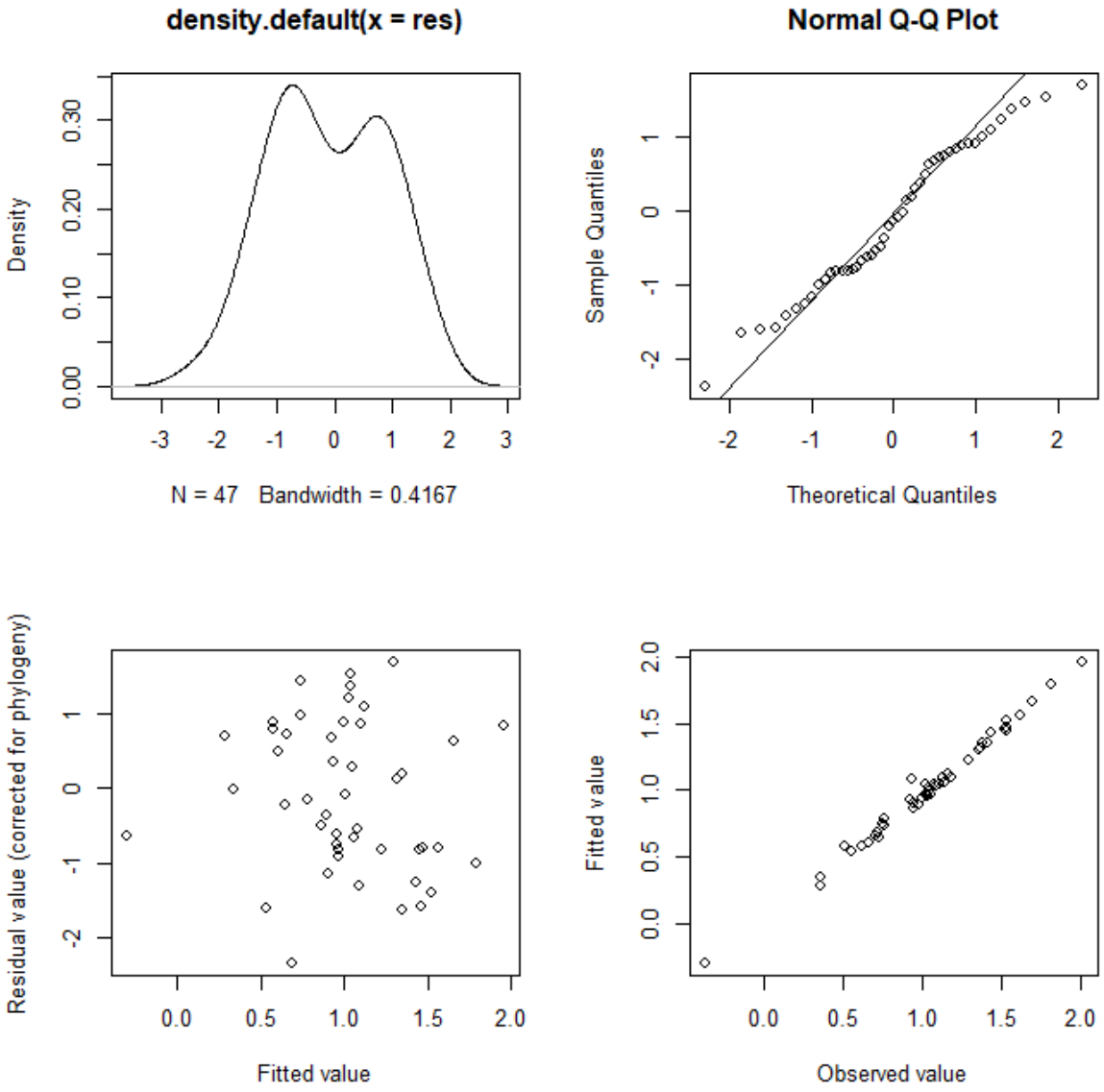


Figure S13 – Diagnostic plots for model 13.

6. PGLS-R of the lengths of slender portions of membranous semicircular ducts against lengths of slender portions of corresponding bony semicircular canals

Models 14-16 provide very significant regressions ($P << 0.001$), with very high coefficients of determination (R^2 = 0.98–0.99). They were tested on specimens for which we have the bony and membranous labyrinths, and were used to predict fitted values for lengths of slender semicircular duct portions (LS, see Supplementary Methods, Biomechanics) of extant and extinct species for which only the bony labyrinth was available (e.g., fossil specimens). The length of the slender portion of each bony semicircular canal (LS_B) was used as predictor. Predicted fitted lengths of slender semicircular duct portions are used to calculate fitted duct-morphology components of Thermo-Motility Indices (TMI_fitted_ω_2 and TMI_fitted_Ω_2, see Supplementary Methods, Biomechanics).

Model 14: $\log_{10}$(LS_a) ~ $\log_{10}$(LS_a_B)

Branch length transformations:

Pagel's λ [ML]: 0.534

lower bound: 0.000, P = 0.20545

upper bound: 1.000, P = 0.00031243 *

95.0% CI: (NA, 0.936)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.0027586	0.0181044	0.1524	0.8796
$\log_{10}$(LS_a_B)	0.9790946	0.0171502	57.0894	$< 2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.9861, AICc: -202.9758

F statistic: 3259 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of lengths of slender portions of bony canals were included)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98205, P = 0.6783, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S14) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.0550, U = 0.9062, P = 0.8526, P value computed via approximation method).

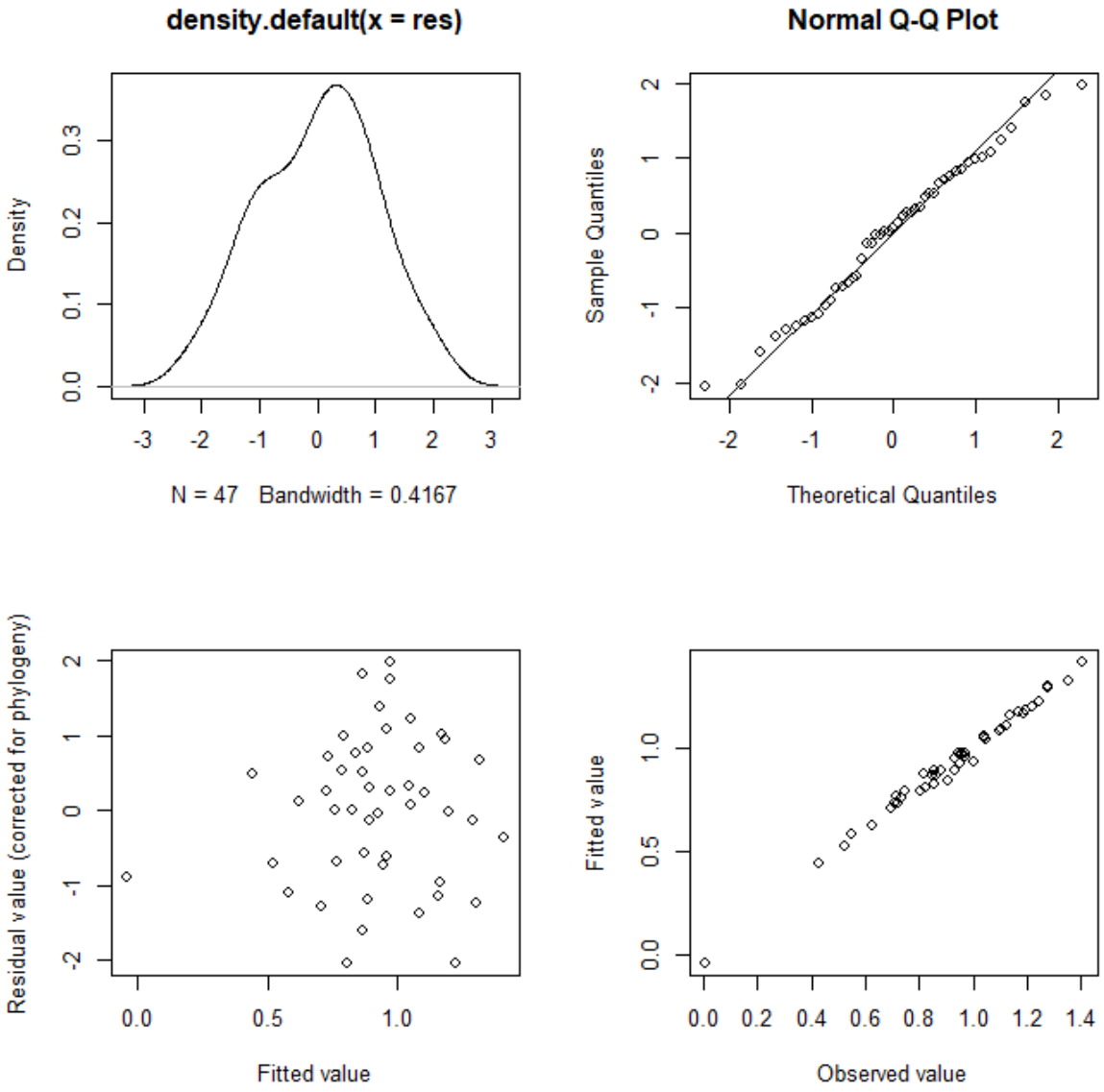


Figure S14 – Diagnostic plots for model 14.

Model 15: $\log_{10}$(LS_p) ~ $\log_{10}$(LS_p_B)

Branch length transformations:

Pagel's λ [ML]: 0.590

lower bound: 0.000, P = 0.032186 *

upper bound: 1.000, P = 2.1286 10^{-06} *

95.0% CI: (0.030, 0.918)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.0039861	0.0206368	0.1932	0.8477
$\log_{10}$(LS_p_B)	0.9694468	0.0198064	48.9460	< 2 10^{-16} *

P values calculated using the exact method.

Adjusted R^2: 0.9812, AICc: -190.6329

F statistic: 2396 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of lengths of slender portions of bony canals were included)

P value: < 2.2 10^{-16} *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.97768, P = 0.5, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S15) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.1101, U = 0.9011, P = 0.7362, P value computed via approximation method).

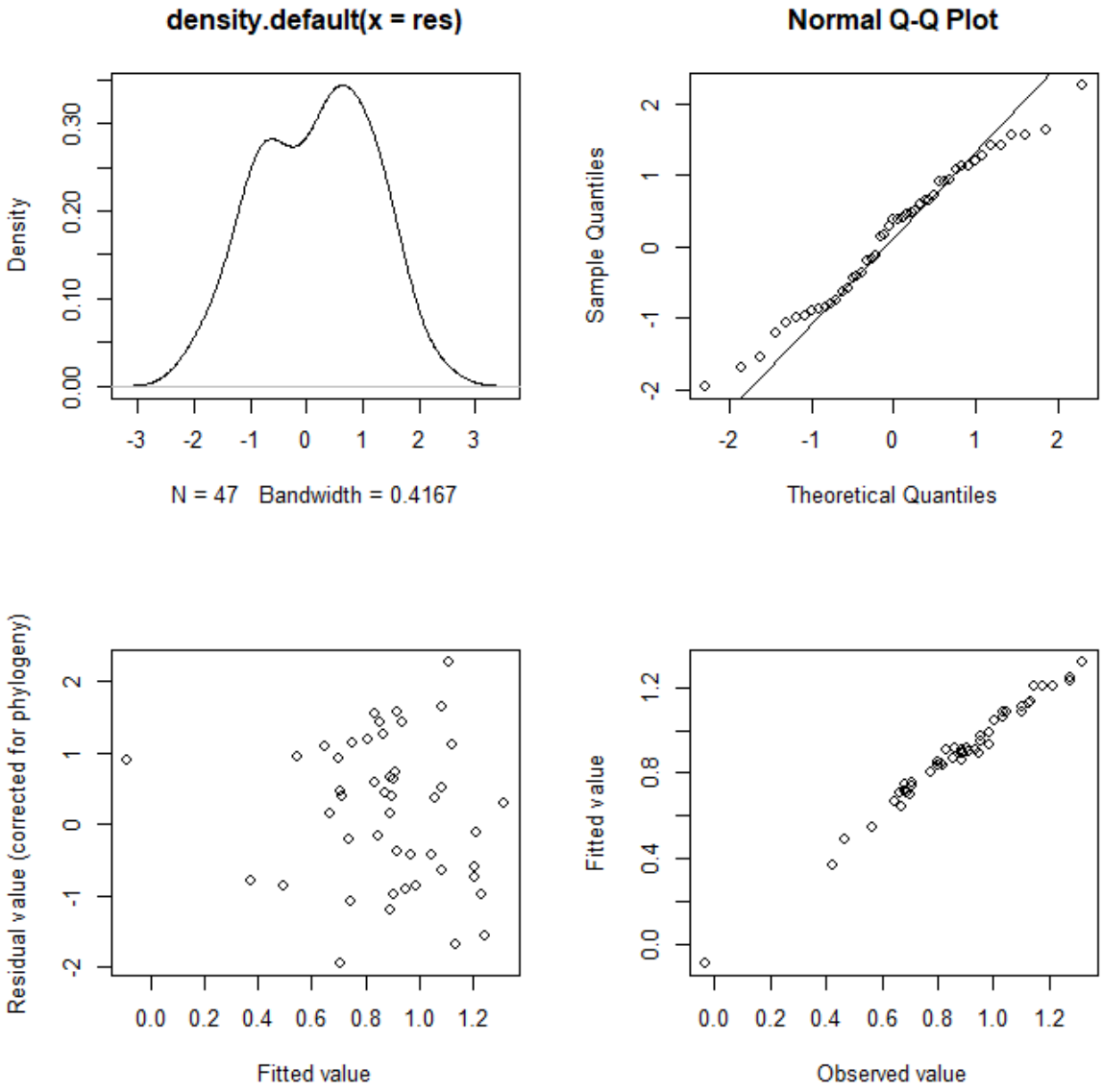


Figure S15 – Diagnostic plots for model 15.

Model 16: $\log_{10}$(LS_l) ~ $\log_{10}$(LS_l_B)

Branch length transformations:

Pagel's λ [ML]: 0.474

lower bound: 0.000, *P* = 0.038058 *

upper bound: 1.000, *P* = 0.0051219 *

95.0% CI: (0.011, 0.930)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.065183	0.024303	-2.6821	0.01026 *
$\log_{10}$(LS_l_B)	1.014299	0.023568	43.0379	< 2 10^{-16} *

P values calculated using the exact method.

Adjusted R^2: 0.9763, AICc: -176.7123

F statistic: 1852 on 1 and 44 DF

n = 46 (3DM set where only specimens with measurements of lengths of slender portions of bony canals were included)

A potential outlier (*Papio hamadryas*, *G* = 3.47582, *U* = 0.73165, *P* = 0.004509 *, *P* value computed via approximation method) was excluded from this regression.

P value: < 2.2 10^{-16} *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.97132, *P* = 0.3099, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S16) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.36856, *U* = 0.87256, *P* = 0.3421, *P* value computed via approximation method).

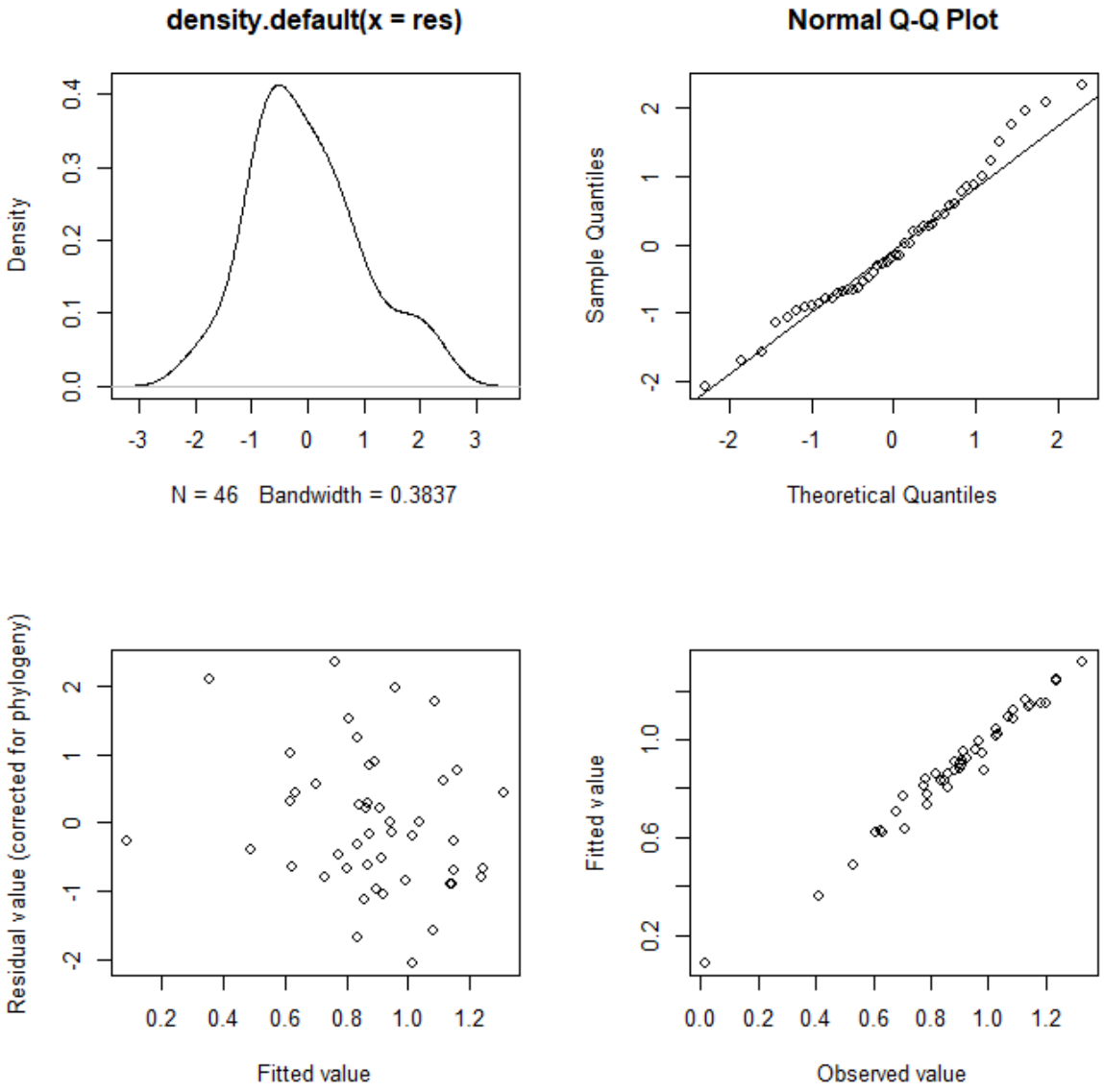


Figure S16 – Diagnostic plots for model 16.

7. PGLS-R of the cilia deflection factors against radii of curvature of corresponding bony semicircular canals

Models 17-19 provide very significant regressions ($P << 0.001$), with very high coefficients of determination (R^2 = 0.87–0.90). They were tested on specimens for which we have the bony and membranous labyrinths, and were used to predict cilia deflection factors (E, see Supplementary Methods, Biomechanics) of extant and extinct species for which only the bony labyrinth was available (e.g., fossil specimens). The radius of curvature of each bony semicircular canal (R_B) was used as predictor. These models show that cilia deflection factors decrease when radii of curvature of the semicircular canals increase. This reflects compensations occurring at the level of the cupula to stabilise the mechanical sensitivity of the semicircular ducts (X_S, see David et al. 2016) under increased endolymph volumic displacements resulting from increased body size. Predicted fitted cilia deflection factors are used to calculate fitted duct-morphology components of Thermo-Motility Indices (TMI_fitted_ω_2 and TMI_fitted_Ω_2, see Supplementary Methods, Biomechanics).

Model 17: $\log_{10}$(E_a) ~ $\log_{10}$(R_a_B)

Branch length transformations:

Pagel's λ [ML]: 0.000

lower bound: 0.000, P = 1

upper bound: 1.000, P = 0.0015941 *

95.0% CI: (NA, 0.349)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	2.059095	0.055306	37.231	$< 2.2\ 10^{-16}$ *
$\log_{10}$(R_a_B)	-2.265650	0.133665	-16.950	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.8721, AICc: -30.34293

F statistic: 287.3 on 1 and 41 DF

n = 43 (3DM set where only specimens with measurements of cilia deflection factors were included)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98891, P = 0.948, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S17) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.26890, U = 0.87451, P = 0.4251, P value computed via approximation method).

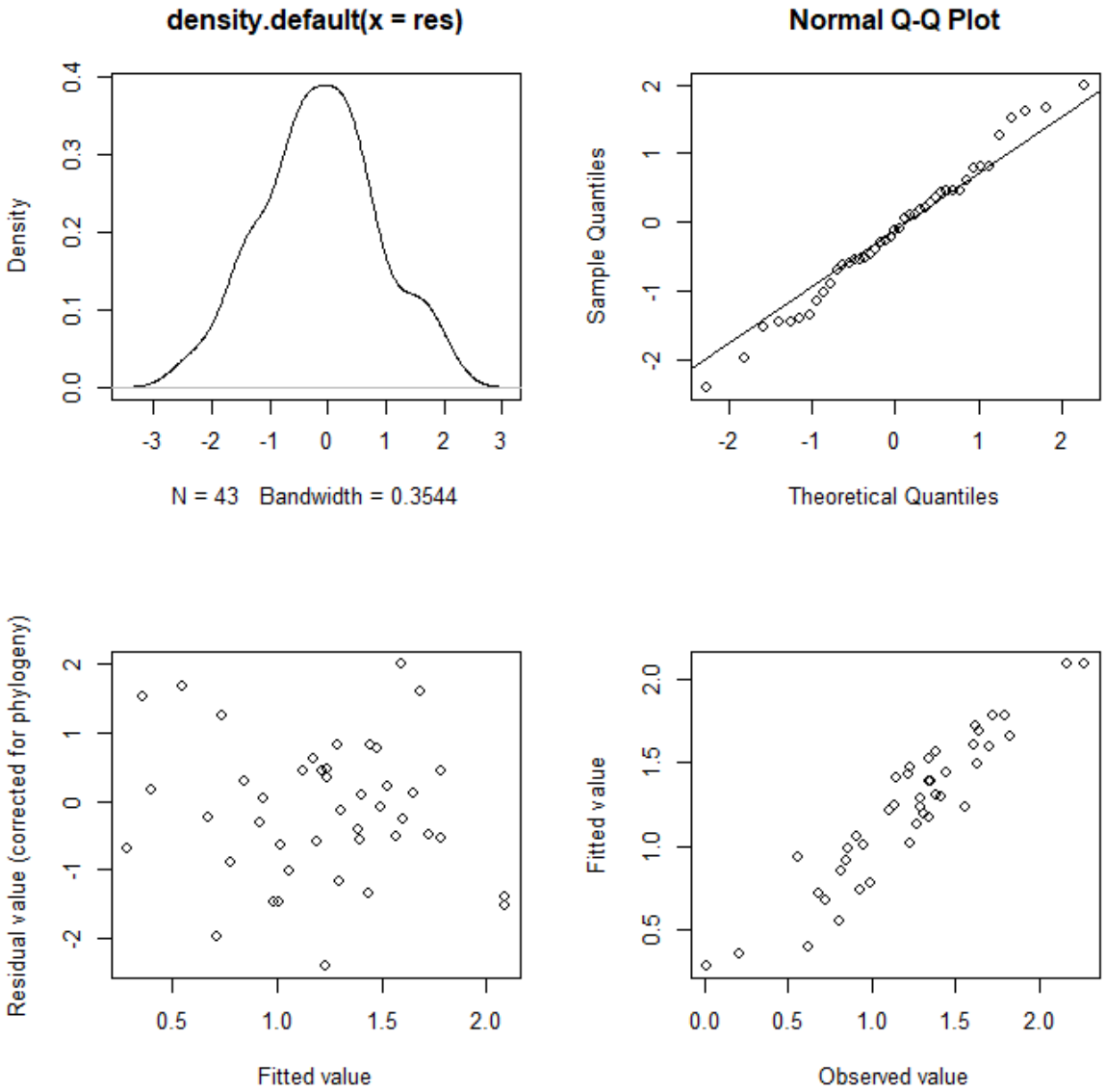


Figure S17 – Diagnostic plots for model 17.

Model 18: $\log_{10}$(E_p) ~ $\log_{10}$(R_p_B)

Branch length transformations:

Pagel's λ [ML]: 0.479

lower bound: 0.000, *P* = 0.0015795 *

upper bound: 1.000, *P* = 0.05985

95.0% CI: (0.097, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.94717	0.07375	26.402	< 2.2 10^{-16} *
$\log_{10}$(R_p_B)	-2.24141	0.11701	-19.156	< 2.2 10^{-16} *

P values calculated using the exact method.

Adjusted R^2: 0.897, AICc: -37.94298

F statistic: 366.9 on 1 and 41 DF

n = 43 (3DM set where only specimens with measurements of cilia deflection factors were included)

P value: < 2.2 10^{-16} *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.97323, *P* = 0.4061, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S18) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.1805, *U* = 0.8841, *P* = 0.5481, *P* value computed via approximation method).

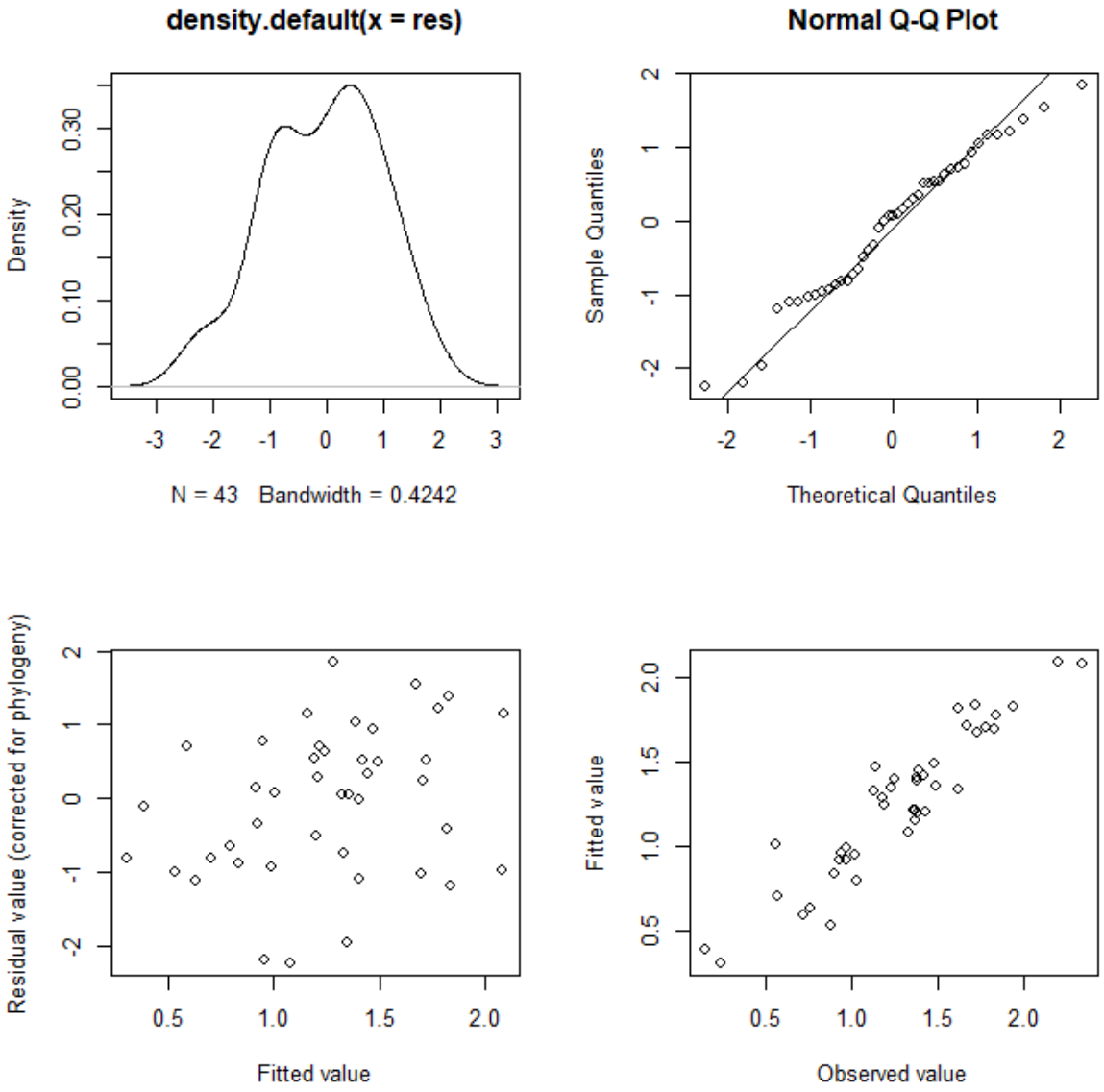


Figure S18 – Diagnostic plots for model 18.

Model 19: $\log_{10}$(E_l) ~ $\log_{10}$(R_l_B)

Branch length transformations:

Pagel's λ [ML]: 0.000

lower bound: 0.000, *P* = 1

upper bound: 1.000, *P* = 0.0014198 *

95.0% CI: (NA, 0.479)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.971029	0.042041	46.883	$< 2.2\ 10^{-16}$ *
$\log_{10}$(R_l_B)	-2.204359	0.115526	-19.081	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.8963, AICc: -44.22581

F statistic: 364.1on 1 and 41 DF

n = 43 (3DM set where only specimens with measurements of cilia deflection factors were included)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.97245, *P* = 0.3826, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S19) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.00035, *U* = 0.90246, *P* = 0.8913, *P* value computed via approximation method).

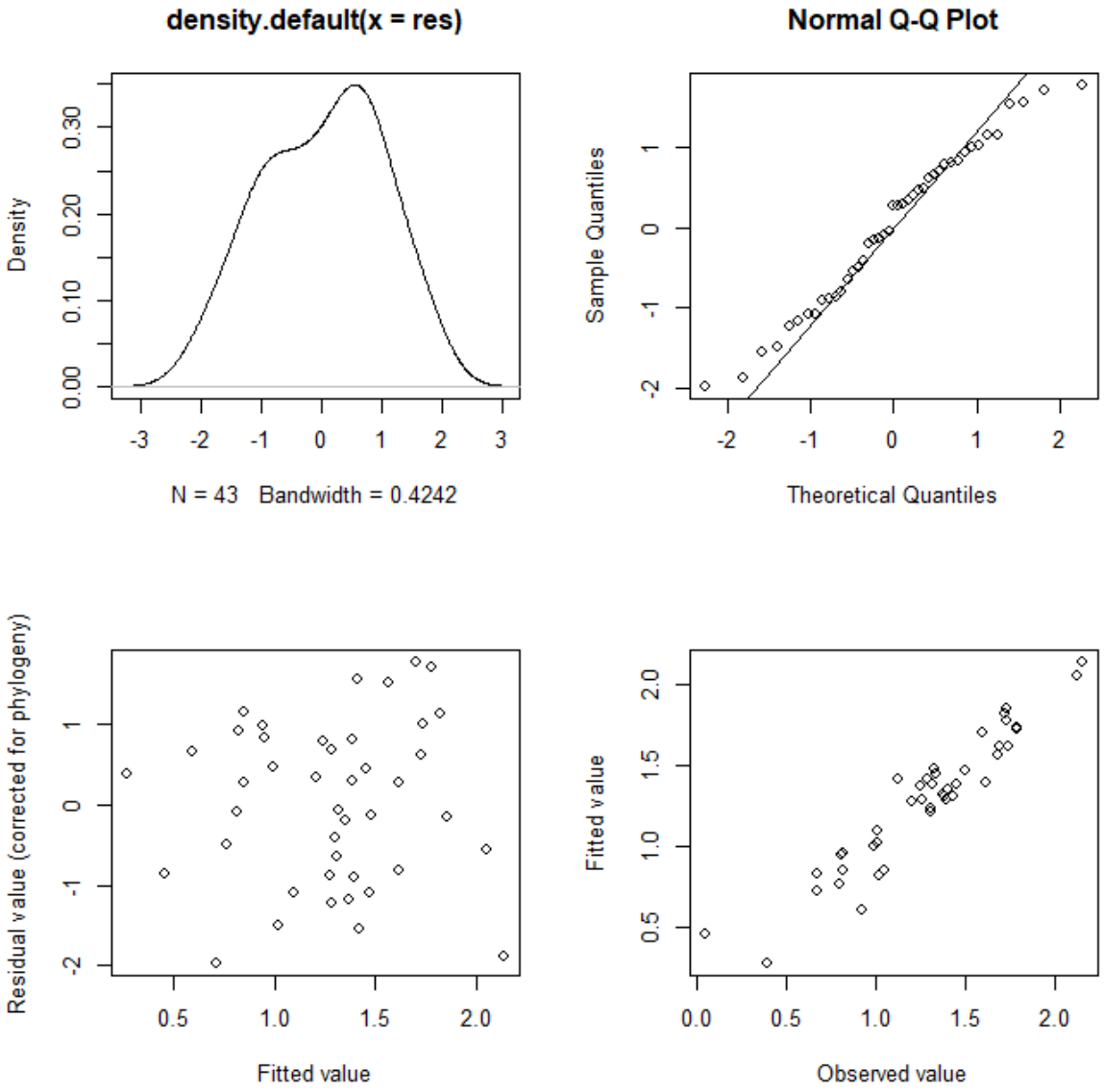


Figure S19 – Diagnostic plots for model 19.

8. PGLS-R of the residual duct-morphology components of the 'raw' Thermo-Motility Indices against the cubic root of body mass, and phylogenetic signal

Models 20-25 were tested on specimens for which we have the bony and membranous labyrinths, to check whether residual duct-morphology components of Thermo-Motility Indices ($TMI\_res\_\omega_2$ and $TMI\_res\_\Omega_2$, see Supplementary Methods, Biomechanics) were (1) correlated to the cubic root of body mass ($\sqrt[3]{BM}$) and (2) reflected phylogeny. We concluded that residual duct-morphology components of Thermo-Motility Indices were not correlated to body mass but carried phylogenetic information. Taking this into account, we predict residual duct-morphology components of Thermo-Motility Indices of extant and extinct species lacking membranous labyrinth information, by using ancestral state reconstruction at nodes where they branch out (Fig. S20-26). Estimated residual duct-morphology components of Thermo-Motility Indices are used, along with fitted duct-morphology and physiochemical components, to calculate 'raw' Thermo-Motility Indices (see Supplementary Methods, Biomechanics).

Model 20: $\log_{10}(TMI\_res\_\omega_2\_a) \sim \log_{10}(\sqrt[3]{BM})$

Adjusted R^2: -0.008166, AICc: -75.79865

F statistic: 0.6274 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: 0.4325, Corrected *P* value: 0.4325

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether phylogeny structured the distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the anterior semicircular duct (Fig. S20). We concluded that phylogeny does partially structure the distribution ($K = 0.391967$, $P = 0.0088$ *, *P* value calculated via approximation method) and should be taken into account to improve predictions of fossil specimens.
- We tested whether the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the anterior semicircular duct, was correlated to ($\log_{10}$) body size. We found no significant correlations.

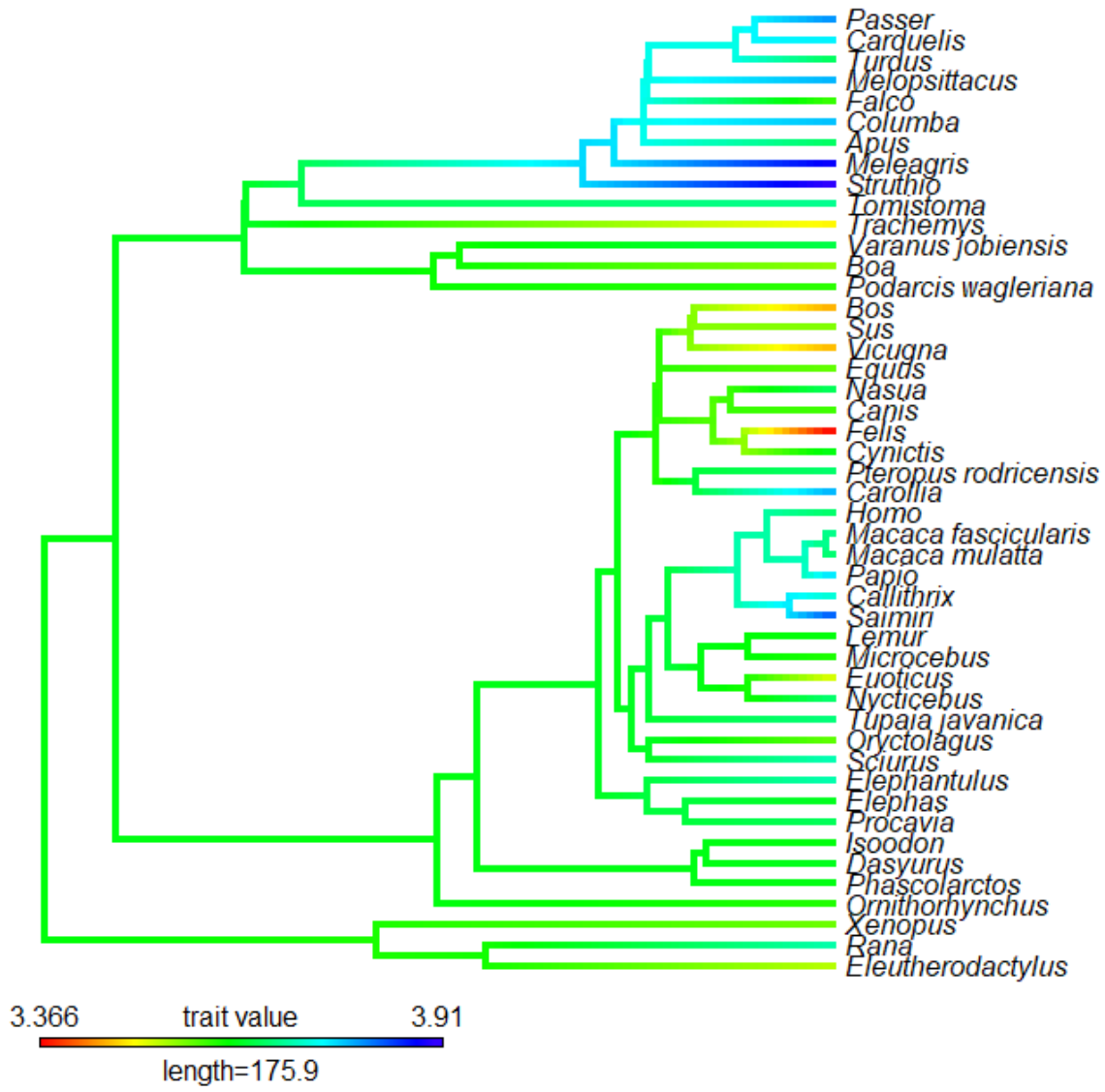


Figure S20 – Phylogenetic distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the anterior semicircular duct.

Model 21: $\log_{10}(\text{TMI\_res\_}\omega_2\text{\_p}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Adjusted R^2: -0.00403, AICc: -72.34938

F statistic: 0.8153 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: 0.3714, Corrected *P* value: 0.4325

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether phylogeny structured the distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the posterior semicircular duct (Fig. S21). We concluded that phylogeny does partially structure the distribution (*K* = 0.402279, *P* = 0.0046 *, *P* value calculated via approximation method) and should be taken into account to improve predictions of fossil specimens.
- We tested whether the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the posterior semicircular duct, was correlated to ($\log_{10}$) body size. We found no significant correlations.

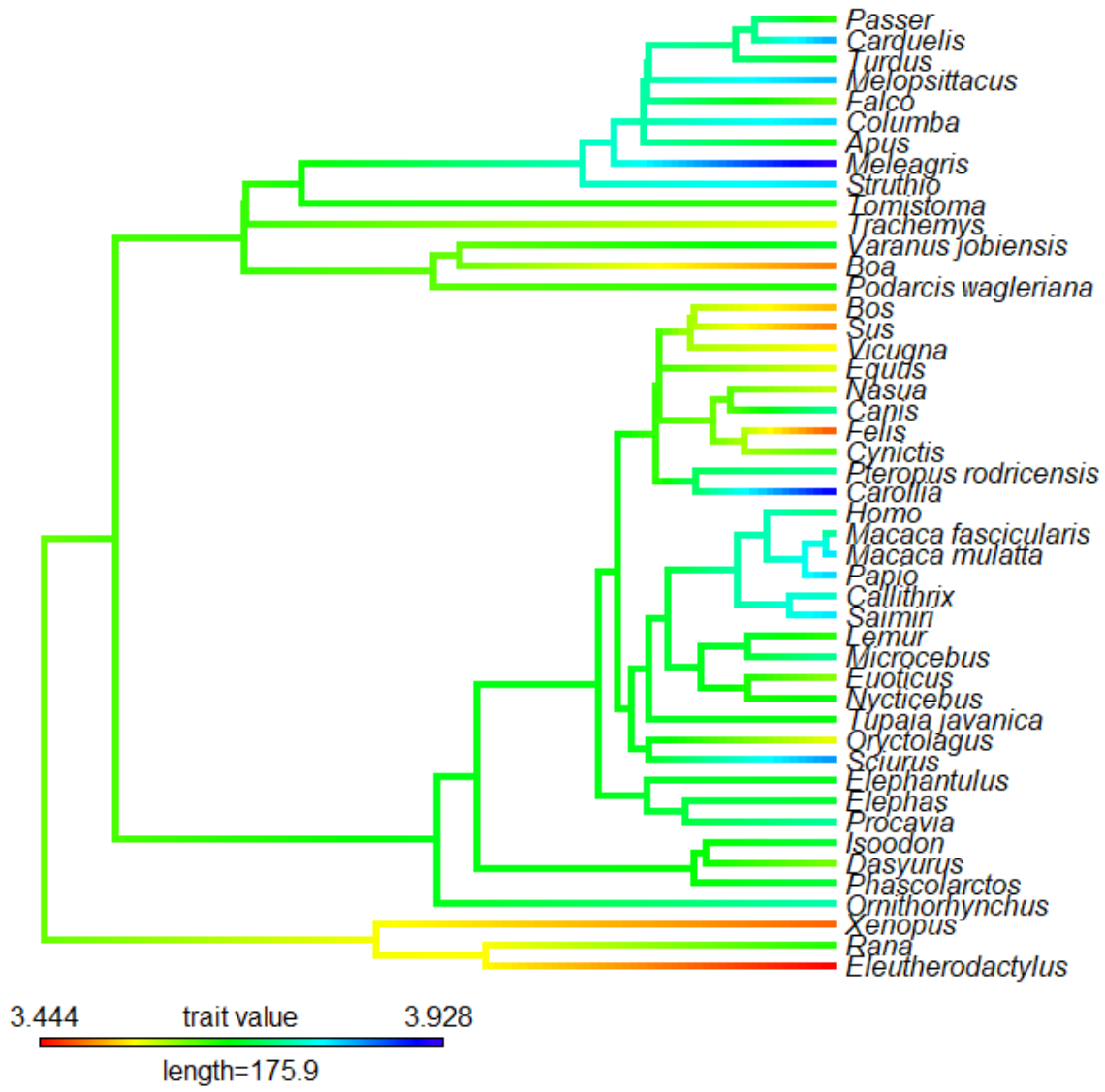


Figure S21 – Phylogenetic distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the posterior semicircular duct.

Model 22: $\log_{10}(\text{TMI\_res\_}\omega_2\text{\_l}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Adjusted R^2: 0.01547, AICc: -68.35134

F statistic: 1.723 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: 0.196, Corrected *P* value: 0.39200

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether phylogeny structured the distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the lateral semicircular duct (Fig. S22). We concluded that phylogeny does partially structure the distribution (*K* = 0.35715, *P* = 0.0175 *, *P* value calculated via approximation method) and should be taken into account to improve predictions of fossil specimens.
- We tested whether the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the lateral semicircular duct, was correlated to ($\log_{10}$) body size. We found no significant correlations.

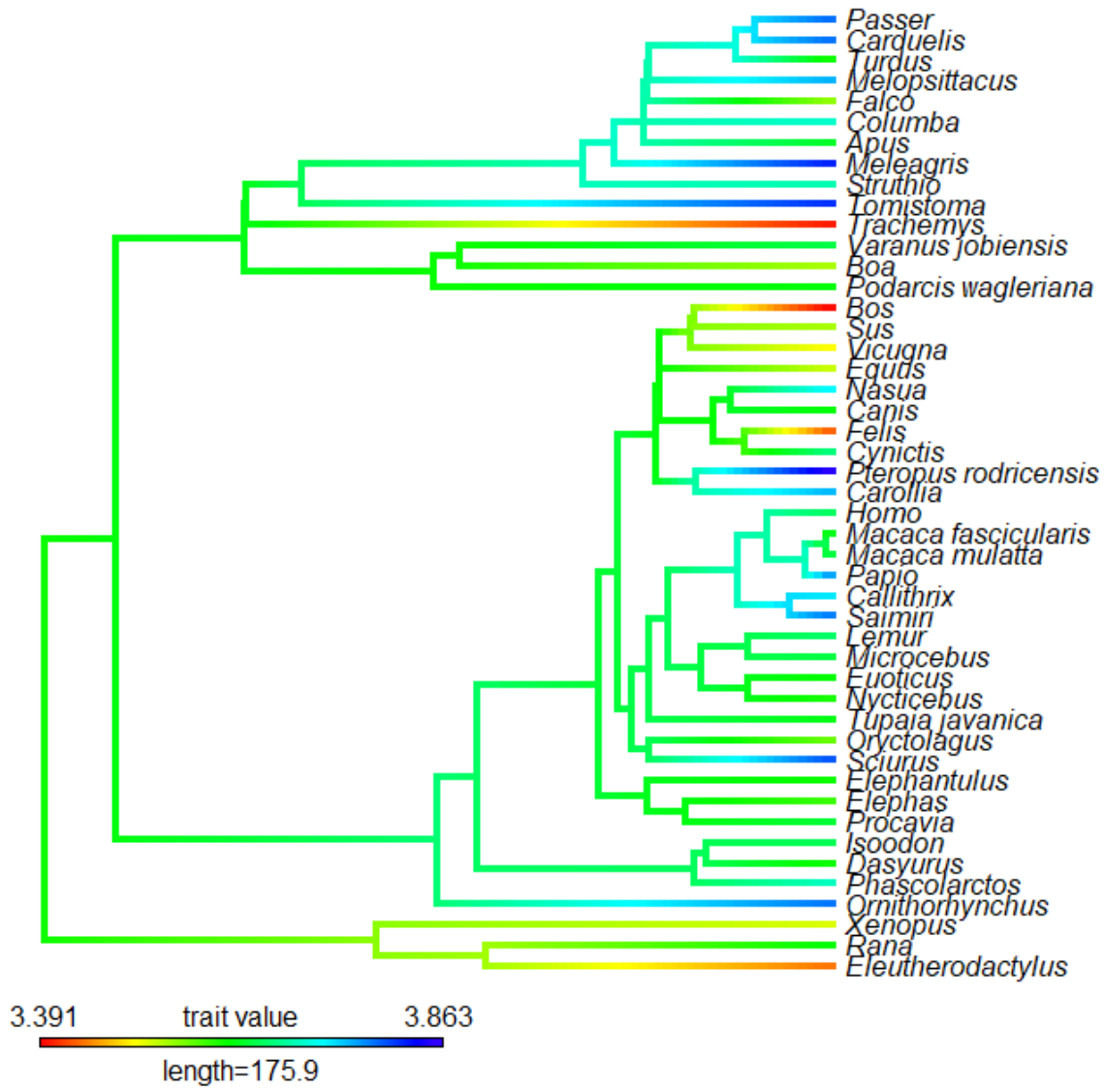


Figure S22 – Phylogenetic distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the upper corner frequency of the lateral semicircular duct.

Model 23: $\log_{10}(\text{TMI\_res\_}\Omega_2\text{\_a}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Adjusted R^2: 0.004553, AICc: -15.1846

F statistic: 1.21 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: 0.2771, Corrected *P* value: 0.41565

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether phylogeny structured the distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the anterior semicircular duct (Fig. S23). We concluded that phylogeny does partially structure the distribution (K = 0.515755, P = 1 10^{-04} *, *P* value calculated via approximation method) and should be taken into account to improve predictions of fossil specimens.
- We tested whether the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the anterior semicircular duct, was correlated to ($\log_{10}$) body size. We found no significant correlations.

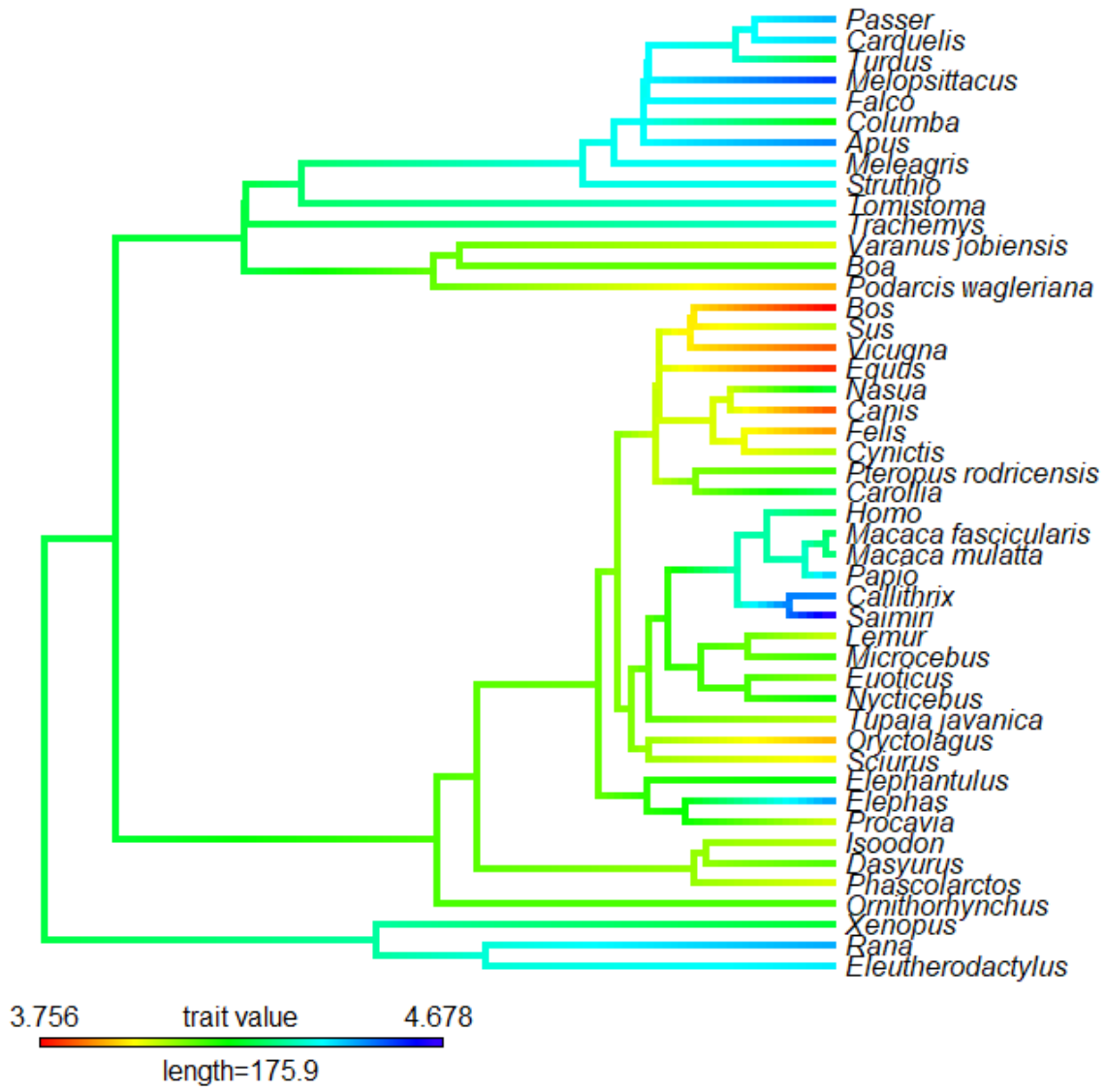


Figure S23 – Phylogenetic distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the anterior semicircular duct.

Model 24: $\log_{10}(\text{TMI\_res\_}\Omega_2\text{\_p}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Adjusted R^2: 0.02897, AICc: -24.41756

F statistic: 2.372 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: 0.1305, Corrected *P* value: 0.39150

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether phylogeny structured the distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the posterior semicircular duct (Fig. S24). We concluded that phylogeny does partially structure the distribution (K = 0.468593, P = 5 10^{-04} *, *P* value calculated via approximation method) and should be taken into account to improve predictions of fossil specimens.
- We tested whether the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the posterior semicircular duct, was correlated to ($\log_{10}$) body size. We found no significant correlations.

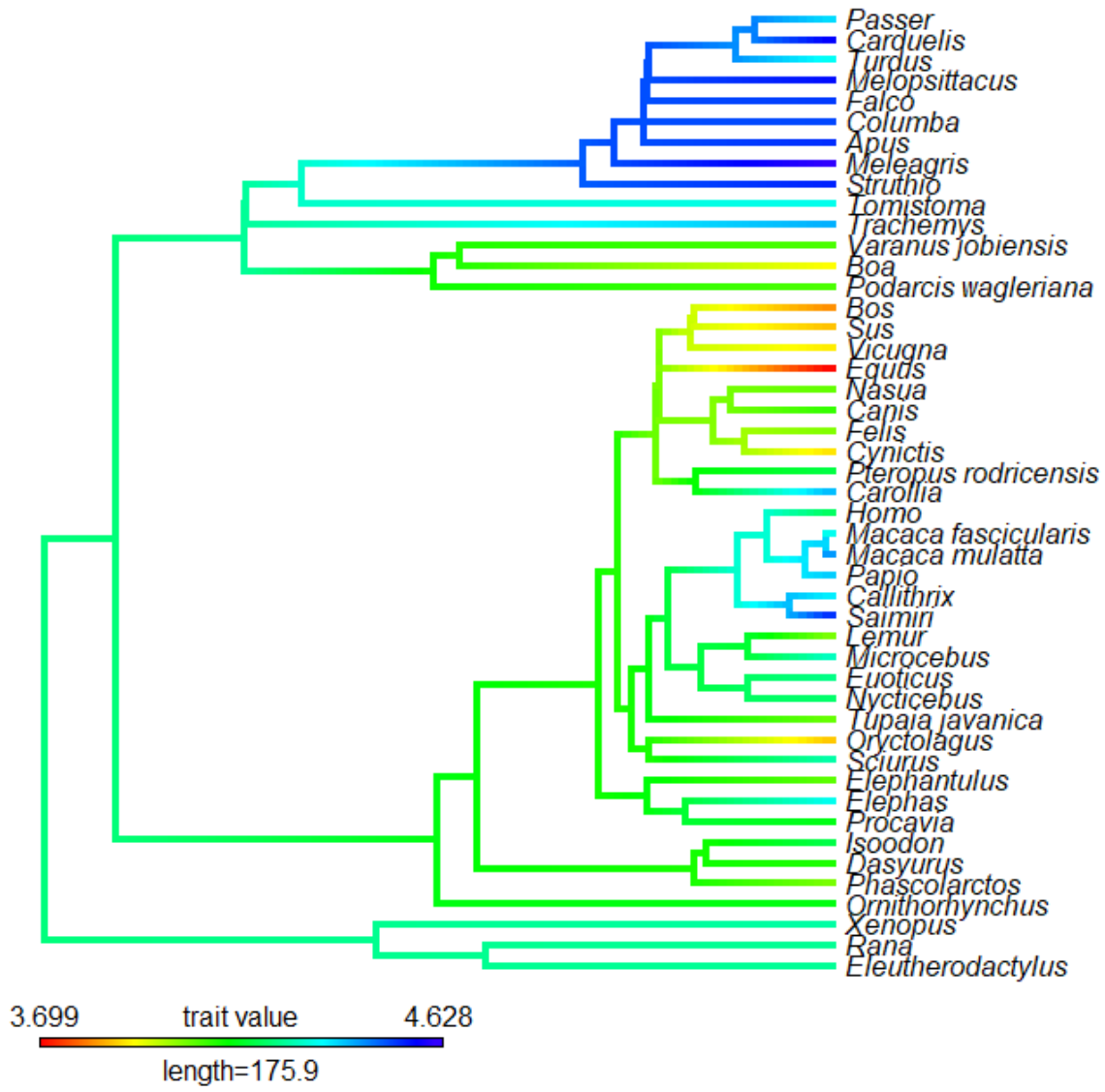


Figure S24 – Phylogenetic distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the posterior semicircular duct.

Model 25: $\log_{10}(\text{TMI\_res\_}\Omega_2\text{\_l}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Adjusted R^2: 0.03812, AICc: -15.05711

F statistic: 2.823 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: 0.09985, Corrected *P* value: 0.39150

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether phylogeny structured the distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct (Fig. S25). We concluded that phylogeny does partially structure the distribution (K = 0.372178, P = 0.0149 *, P value calculated via approximation method) and should be taken into account to improve predictions of fossil specimens.
- We tested whether the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct, was correlated to ($\log_{10}$) body size. We found no significant correlations.

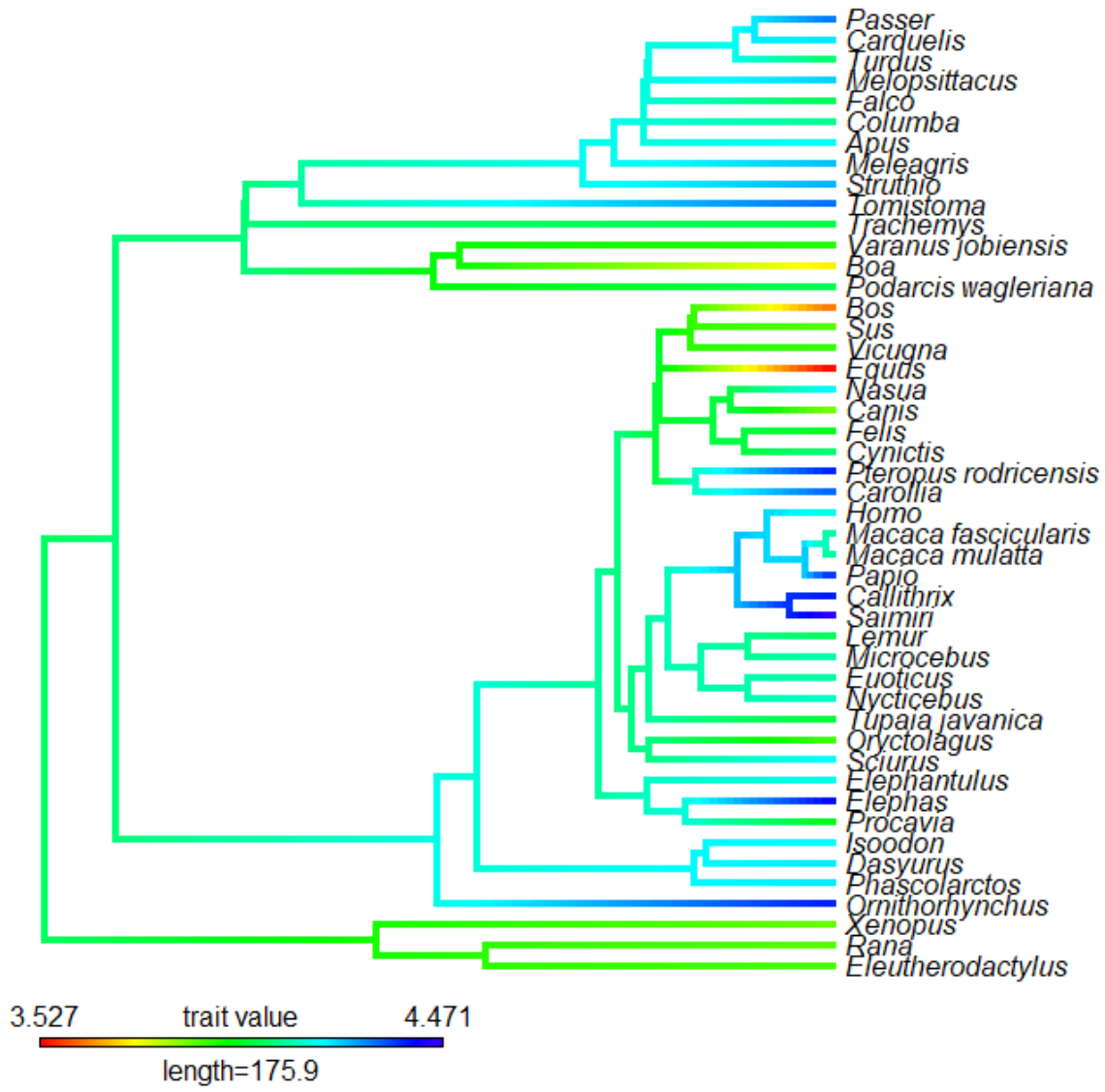


Figure S25 – Phylogenetic distribution of the ($\log_{10}$) residual duct-morphology component of the Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct.

9. PGLS-R of the 'raw' Thermo-Motility Indices against the cubic root of body mass

Models 26-31 provide very significant regressions (P << 0.001), with coefficients of determination R^2 ranging from 0.46–0.69. They were tested on specimens for which we have the bony labyrinth, and were used to size-correct 'raw' Thermo-Motility Indices (TMI_raw_ω_2 and TMI_raw_Ω_2, see Supplementary Methods, Biomechanics). In practice, this was done by combining fitted Thermo-Motility Indices obtained for a body mass (BM) of 1g with residuals obtained for each species, for each model. Size-corrected Thermo-Motility Indices (TMI_ω_2 and TMI_Ω_2, see Supplementary Methods, Biomechanics) are provided in Supplementary Data 2, and are used in Models 32-37 to assess which one best predict the temperature term (T_R, see Supplementary Methods, Biomechanics).

Model 26: $\log_{10}(\text{TMI\_raw\_}\omega_2\text{\_a}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Branch length transformations:

Pagel's λ [ML]: 0.777

lower bound: 0.000, $P = < 2.2\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.2\ 10^{-16}$ *

95.0% CI: (0.662, 0.866)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.759262	0.071015	10.691	$< 2.2\ 10^{-16}$ *
$\log_{10}(\sqrt[3]{\text{BM}})$	-0.616676	0.025371	-24.307	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.6651, AICc: -300.1331

F statistic: 590.8 on 1 and 296 DF

n = 298 (3DB set)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.99641, P = 0.7402, P value calculated via approximation method).

- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S26) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 3.3669, U = 0.9617, P = 0.1015, P value computed via approximation method).

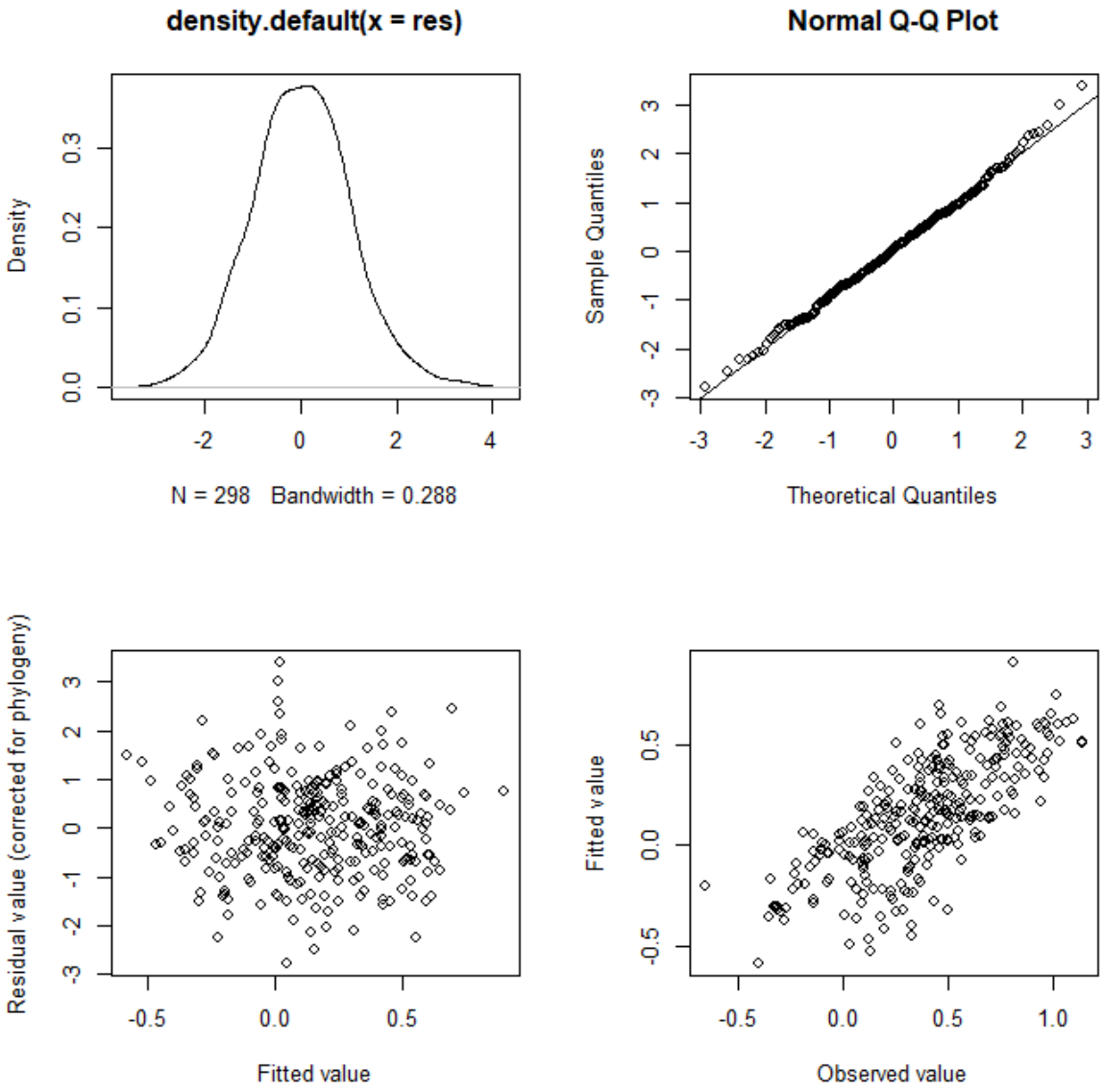


Figure S26 – Diagnostic plots for model 26.

Model 27: $\log_{10}(\text{TMI\_raw\_}\omega_2\text{\_p}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Branch length transformations:

Pagel's λ [ML]: 0.789

lower bound: 0.000, $P = < 2.2\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.2\ 10^{-16}$ *

95.0% CI: (0.675, 0.875)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.815601	0.069059	11.810	$< 2.2\ 10^{-16}$ *
$\log_{10}(\sqrt[3]{\text{BM}})$	-0.629477	0.024364	-25.837	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.6925, AICc: -323.5753

F statistic: 667.5 on 1 and 295 DF

n = 297 (3DB set)

As the initial Shapiro-Wilk test (*W* = 0.98909, *P* = 0.02464 *, *P* value calculated via approximation method) indicated potential non-normality of the phylogenetic residuals, the most influential specimen (*Microgomphodon oligocynus*) was excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.99369, *P* = 0.2503, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S27) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 3.3830, *U* = 0.9612, *P* = 0.09519, *P* value computed via approximation method).

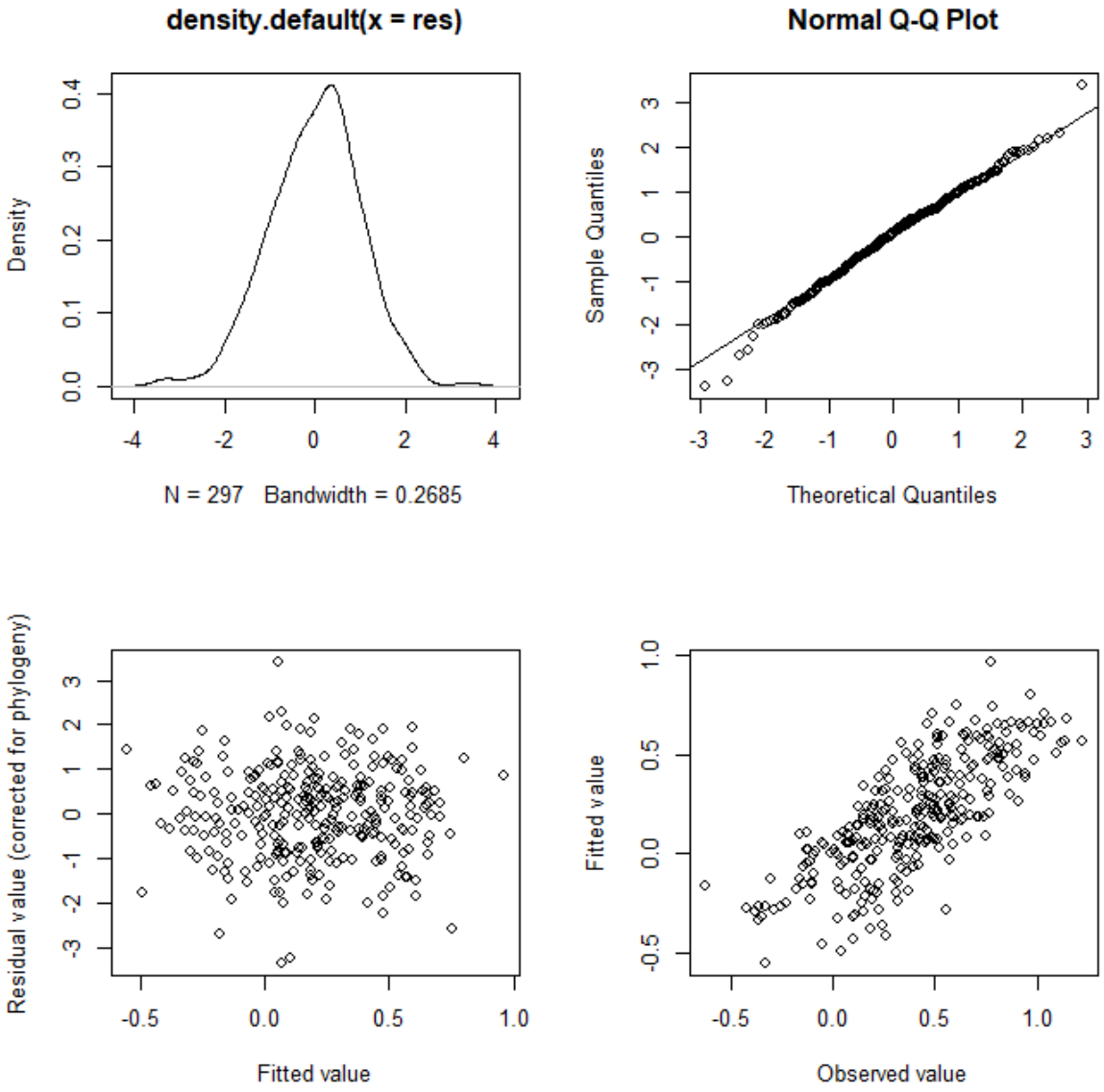


Figure S27 – Diagnostic plots for model 27.

Model 28: $\log_{10}(\text{TMI\_raw\_}\omega_2\text{\_l}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Branch length transformations:

Pagel's λ [ML]: 0.801

lower bound: 0.000, $P = < 2.2\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.2\ 10^{-16}$ *

95.0% CI: (0.693, 0.881)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.767119	0.062027	12.367	$< 2.2\ 10^{-16}$ *
$\log_{10}(\sqrt[3]{\text{BM}})$	-0.552946	0.021664	-25.524	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.6873, AICc: -396.6525

F statistic: 651.5 on 1 and 295 DF

n = 297 (3DB set)

As the initial Shapiro-Wilk test (*W* = 0.99163, *P* = 0.08939, *P* value calculated via approximation method) indicated potential non-normality of the phylogenetic residuals, the most influential specimen (*Niassodon mfumukasi*) was excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.99235, *P* = 0.1303, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S28) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 3.19109, *U* = 0.96548, *P* = 0.1928, *P* value computed via approximation method).

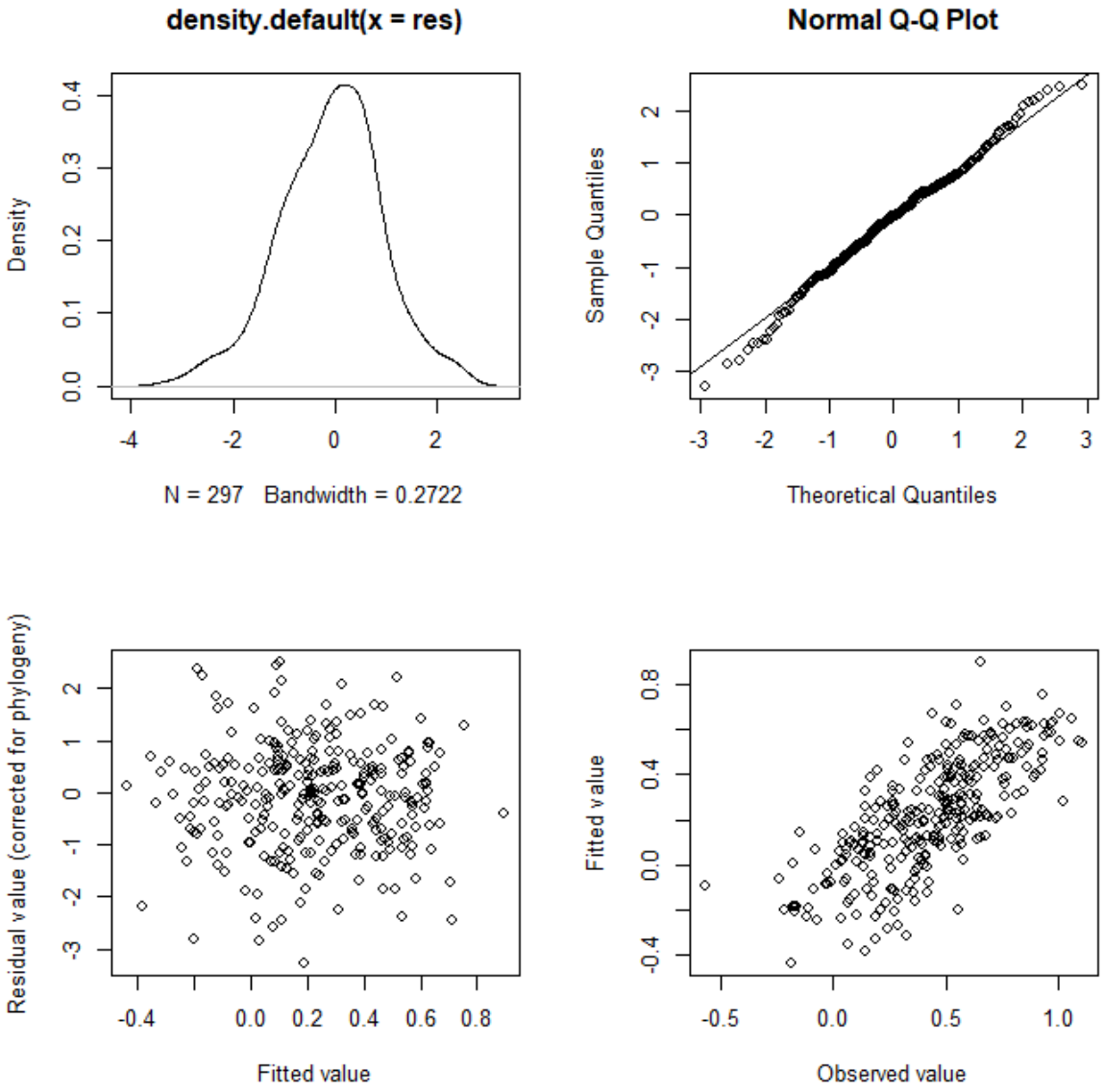


Figure S28 – Diagnostic plots for model 28.

Model 29: $\log_{10}(\text{TMI\_raw\_}\Omega_2\text{\_a}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Branch length transformations:

Pagel's λ [ML]: 0.669

lower bound: 0.000, $P = < 2.2\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.2\ 10^{-16}$ *

95.0% CI: (0.525, 0.790)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.925786	0.094700	9.776	$< 2.2\ 10^{-16}$ *
$\log_{10}(\sqrt[3]{\text{BM}})$	-0.610012	0.038062	-16.027	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.4645, AICc: -50.01016

F statistic: 256.9 on 1 and 294 DF

n = 296 (3DB set)

Potential outliers (*Kawingasaurus fossilis*, *G* = 3.70721, *U* = 0.95357, *P* = 0.02654 *, *Microgomphodon oligocynus*, *G* = 4.25867, *U* = 0.93852, *P* = 0.002289 *, *P* values computed via approximation method) were excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.99723, *P* = 0.8974, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S29) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 3.13643, *U* = 0.96654, *P* = 0.2331, *P* value computed via approximation method).

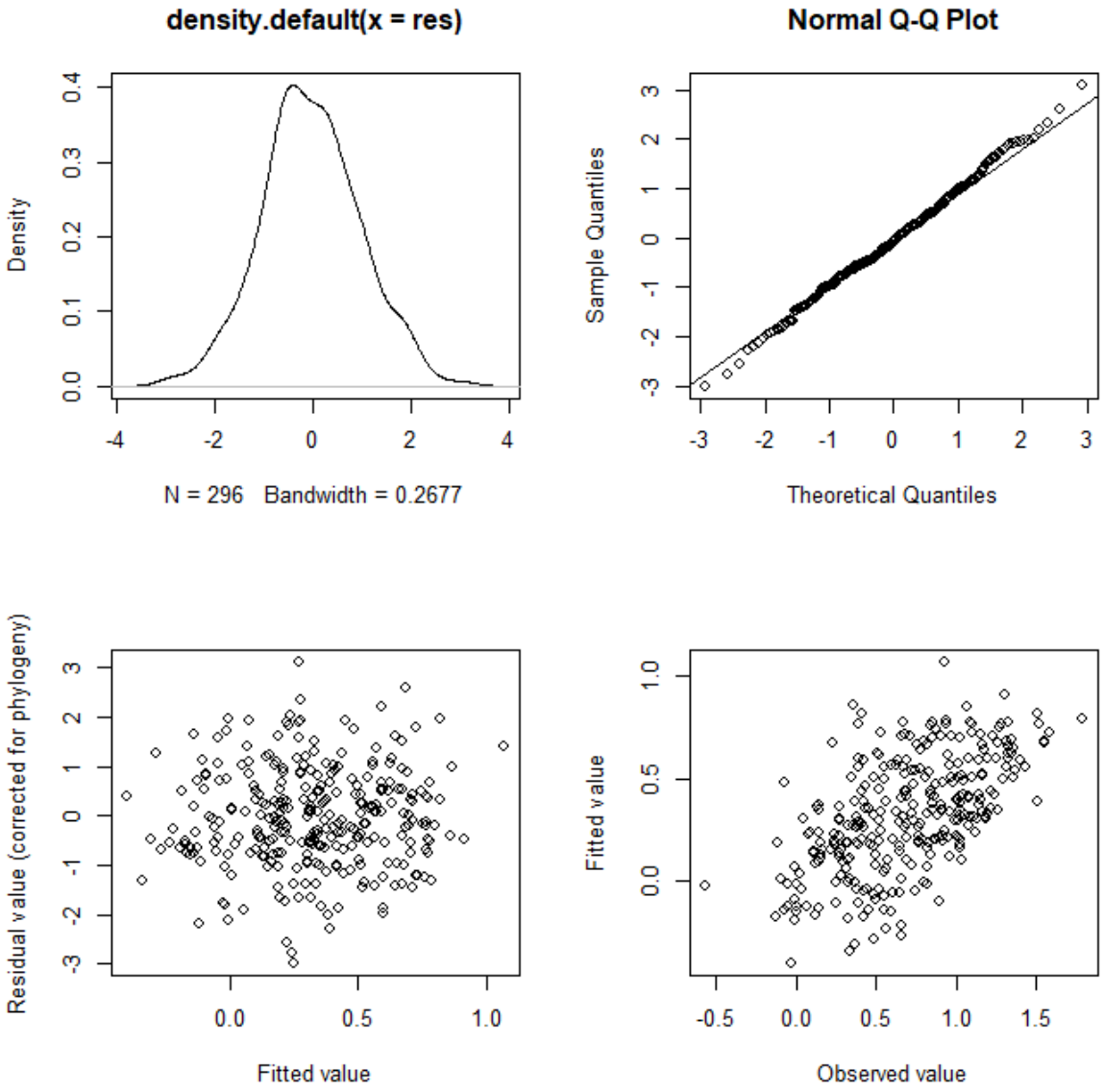


Figure S29 – Diagnostic plots for model 29.

Model 30: $\log_{10}(\text{TMI\_raw\_}\Omega_2\text{\_p}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Branch length transformations:

Pagel's λ [ML]: 0.651

lower bound: 0.000, $P = < 2.2\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.2\ 10^{-16}$ *

95.0% CI: (0.510, 0.773)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.089078	0.085466	12.743	$< 2.2\ 10^{-16}$ *
$\log_{10}(\sqrt[3]{\text{BM}})$	-0.653329	0.034562	-18.903	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.5454, AICc: -100.1469

F statistic: 357.3 on 1 and 296 DF

n = 298 (3DB set)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.99244$, $P = 0.1349$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S30) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 3.5047$, $U = 0.9585$, $P = 0.05987$, *P* value computed via approximation method).

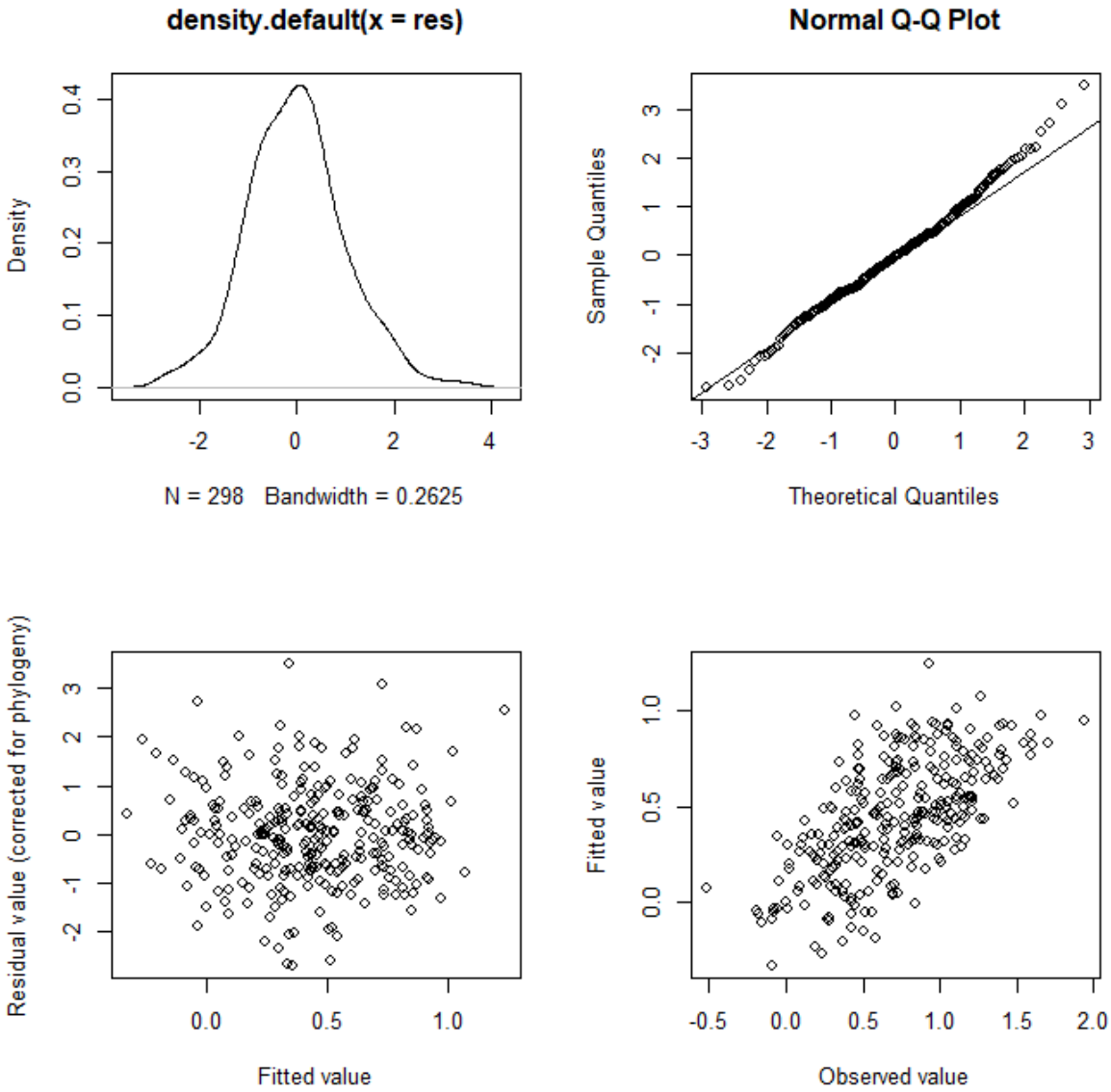


Figure S30 – Diagnostic plots for model 30.

Model 31: $\log_{10}(\text{TMI\_raw\_}\Omega_2\text{\_l}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Branch length transformations:

Pagel's λ [ML]: 0.670

lower bound: 0.000, $P = < 2.2\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.2\ 10^{-16}$ *

95.0% CI: (0.540, 0.780)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.947763	0.075953	12.478	$< 2.2\ 10^{-16}$ *
$\log_{10}(\sqrt[3]{\text{BM}})$	-0.488995	0.030173	-16.206	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.4684, AICc: -183.5676

F statistic: 262.6 on 1 and 296 DF

n = 298 (3DB set)

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.99353$, $P = 0.23$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S31) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 3.08601$, $U = 0.96783$, $P = 0.28$, *P* value computed via approximation method).

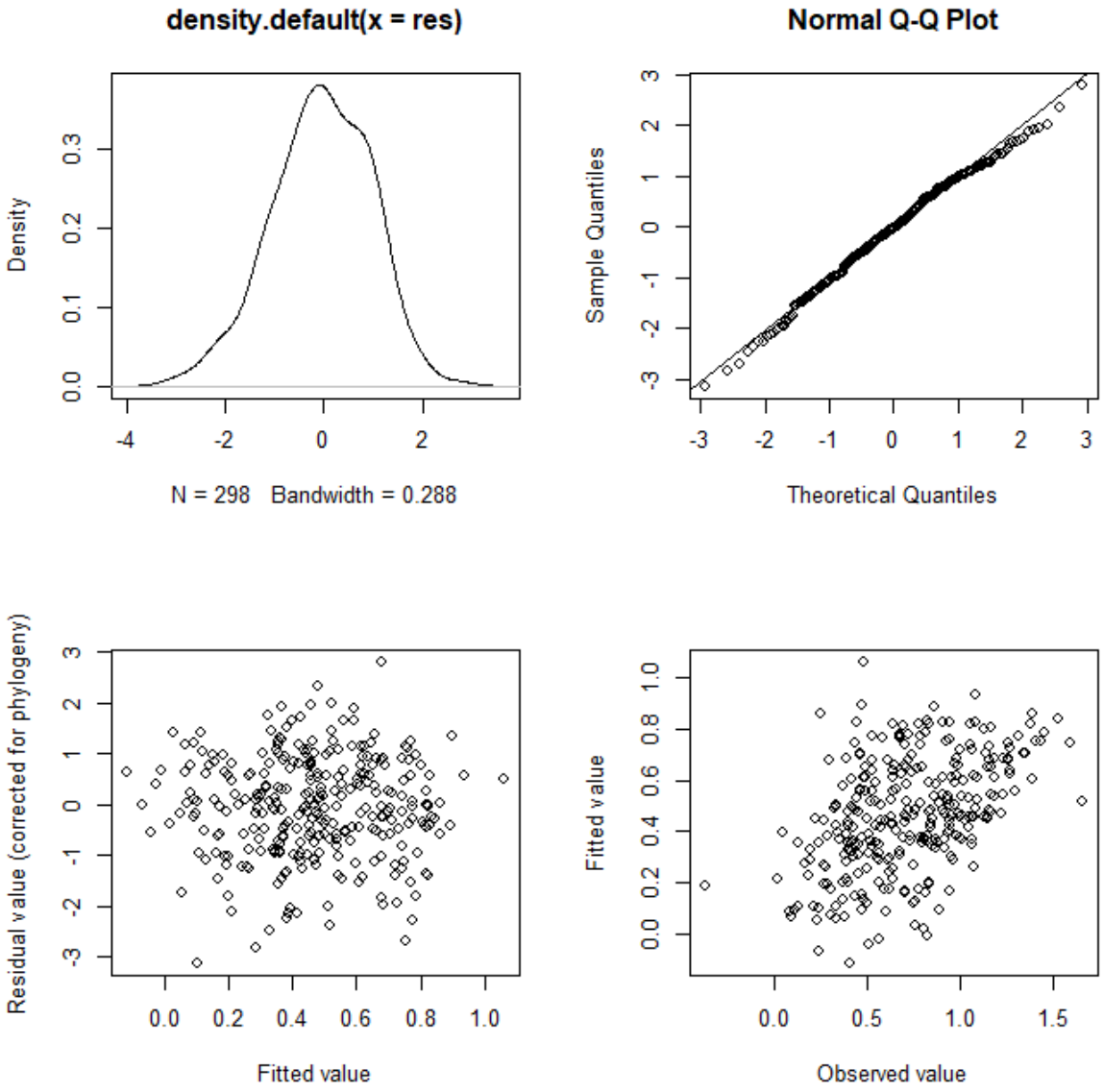


Figure S31 – Diagnostic plots for model 31.

10. PGLS-R of the temperature term against size-corrected Thermo-Motility Indices

Models 32-37 provide significant regressions ($P < 0.01$), with coefficients of determination (R^2) ranging between 0.23–0.83. They were tested on specimens for which we know the body temperature, and were used to assess which size-corrected Thermo-Motility Indices (TMI_raw_ω_2 and TMI_raw_Ω_2, see Supplementary Methods, Biomechanics) best predict the temperature term (T_R, see Supplementary Methods, Biomechanics), which decreases when body temperature increases. Because Thermo-Motility Indices are not equally informative about the temperature term (e.g., AICc of Model 35 << AICc of Model 37), Akaike weights obtained from these models (Section 11) were applied to size-corrected Thermo-Motility Indices to compute the (size-corrected weighted average) Thermo-Motility Index (see Supplementary Methods, Biomechanics).

In these models we used average values for the clades Acanthopterygii, Anguimorpha, Anura, Atlantogenata, Batoidea, Caudata, Crocodylia, Elopomorpha, Euarchontoglires, Galloanserae, Gekkota, Gymnophiona, Holocephali, Iguania, Lacertoidea, Laurasiatheria, Marsupialia, Monotremata, Neoaves, Otocephala, Palaeognathae, Paracanthopterygii, Protacanthopterygii, Rhynchocephalia, Scincomorpha, Selachii, Serpentes, and Testudines. As Batoidea was found to be a potential outlier for Model 37 (G = 2.74272, U = 0.71107, P = 0.04451 *, P value computed via approximation method) we decided to exclude its most heterogeneous species (*Dipturus batis*, *Urolophus paucimaculatus*, *Gymnura micrura*) when computing the average for Models 32-37. These groups were chosen to best balance taxonomic sampling with robust averaging of Thermo-Motility Indices.

Model 32: $T_R \sim \log_{10}$(TMI ω_2_a)

Branch length transformations:

Pagel's λ [ML]: 0.000

lower bound: 0.000, P = 1

upper bound: 1.000, P = 0.096853

95.0% CI: (NA, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.800050	0.021777	82.660	$< 2.2\ 10^{-16}$ *
$\log_{10}$(TMI ω_2_a)	-0.295580	0.025628	-11.533	$1.004\ 10^{-11}$ *

P values calculated using the exact method.

Adjusted R^2: 0.8302, AICc: -92.98236

F statistic: 133 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: $1.004\ 10^{-11}$ *, Corrected *P* value: $6.0240\ 10^{-11}$ *

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.96445, *P* = 0.4419, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S32) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.23110, *U* = 0.80881, *P* = 0.2797, *P* value computed via approximation method).

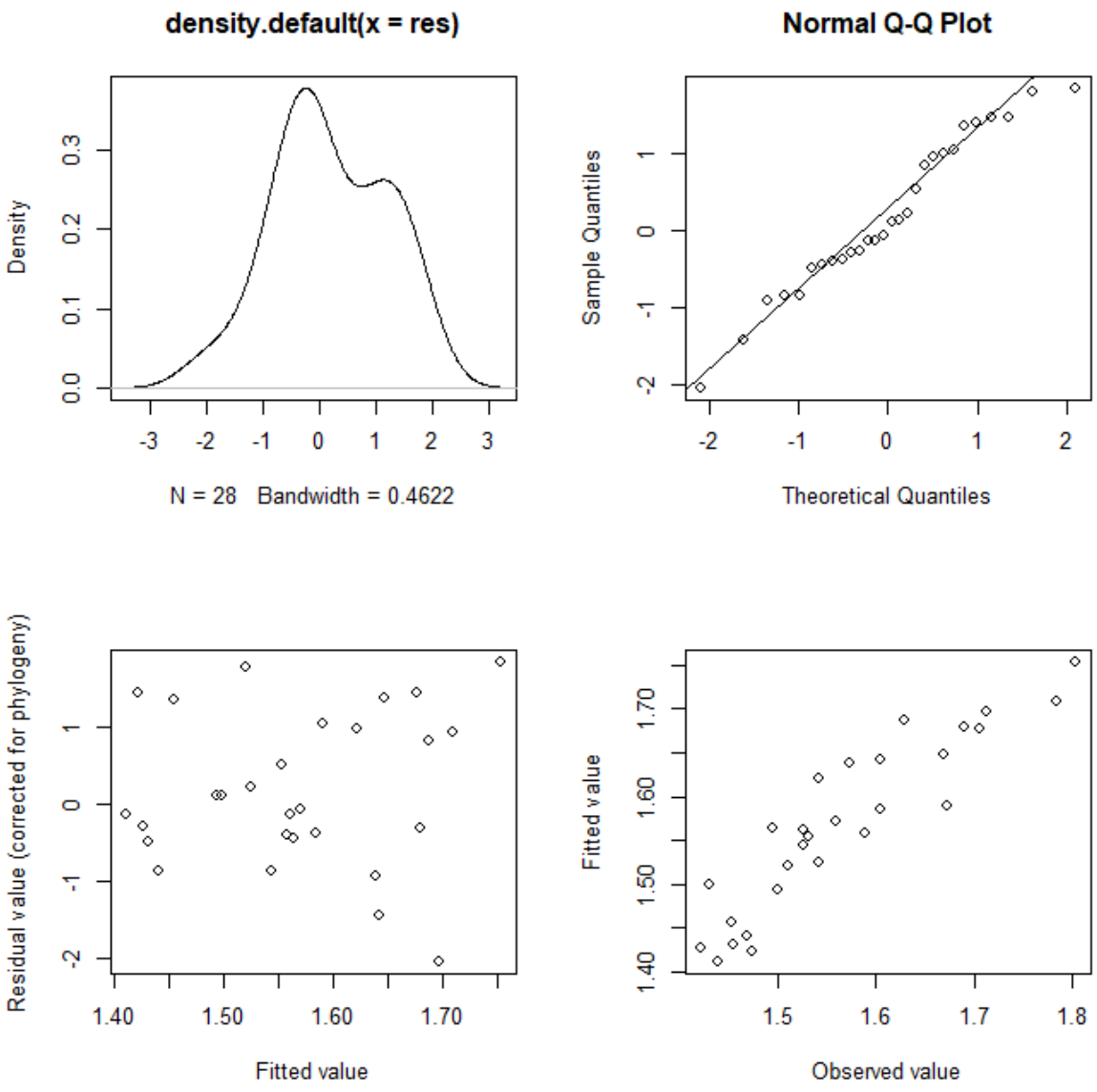


Figure S32 – Diagnostic plots for model 32.

Model 33: $T_R \sim \log_{10}(\text{TMI } \omega_2\_p)$

Branch length transformations:

Pagel's λ [ML]: 0.000

lower bound: 0.000, $P = 1$

upper bound: 1.000, $P = 0.20061$

95.0% CI: (NA, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.818030	0.026833	67.7544	$< 2.2\ 10^{-16}$ *
$\log_{10}$(TMI ω_2_p)	-0.299361	0.030092	-9.9483	$2.361\ 10^{-10}$ *

P values calculated using the exact method.

Adjusted R^2: 0.7839, AICc: -86.23574

F statistic: 98.97 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: 2.361 10-^{10} *, Corrected *P* value: $7.0830\ 10^{-10}$ *

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.989$, $P = 0.9884$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S33) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.36035$, $U = 0.78602$, $P = 0.1847$, *P* value computed via approximation method).

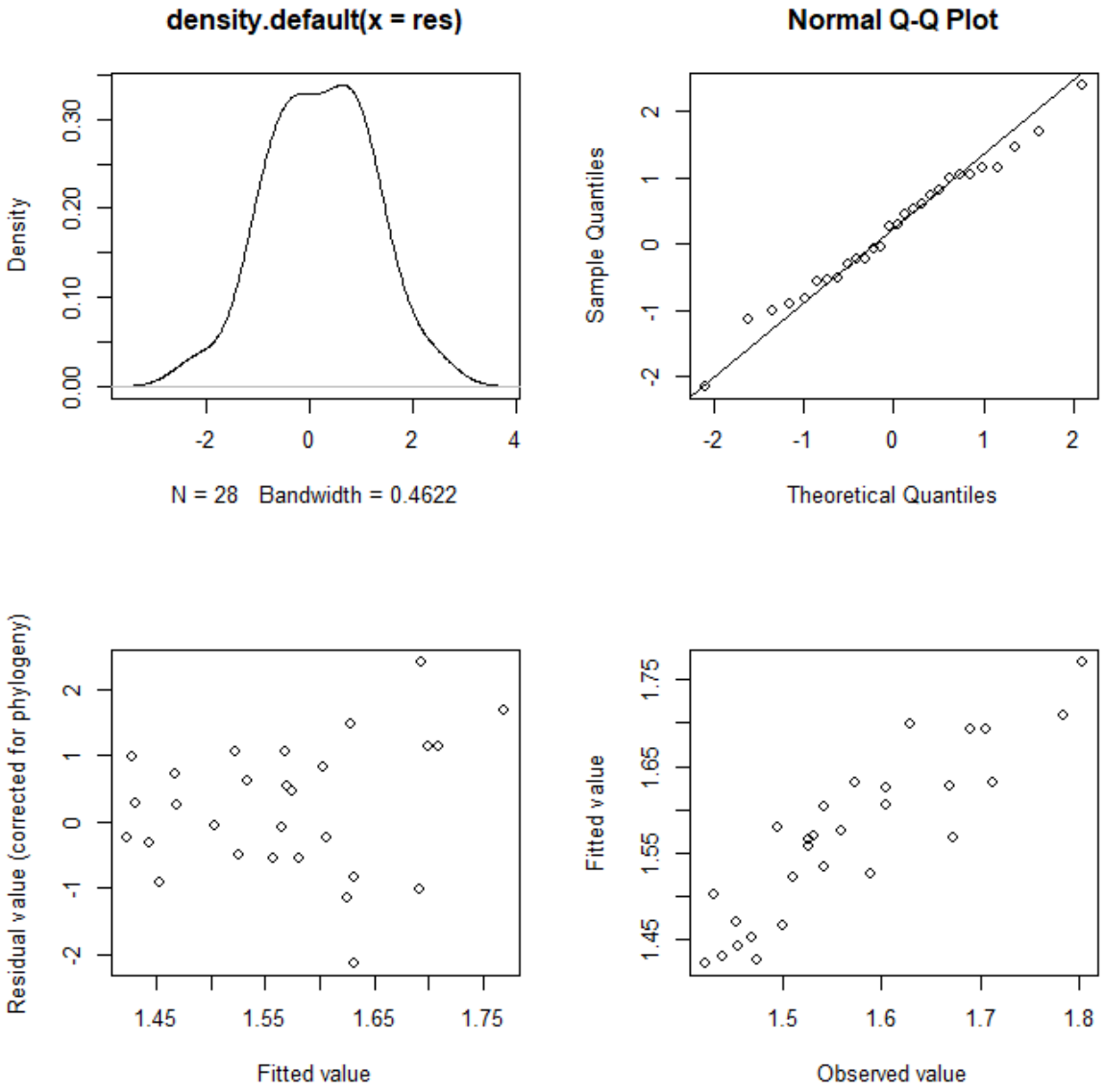


Figure S33 – Diagnostic plots for model 33.

Model 34: $T_R \sim \log_{10}(\text{TMI } \omega_2\_l)$

Branch length transformations:

Pagel's λ [ML]: 1.000

lower bound: 0.000, $P = 0.027098$ *

upper bound: 1.000, $P = 1$

95.0% CI: (0.184, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.791167	0.048498	36.9327	$< 2.2\ 10^{-16}$ *
$\log_{10}$(TMI ω_2_l)	-0.242691	0.056981	-4.2591	0.0002372 *

P values calculated using the exact method.

Adjusted R^2: 0.3883, AICc: -82.31672

F statistic: 18.14 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: 0.0002372 *, Corrected *P* value: $2.8464\ 10^{-04}$ *

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.9635$, $P = 0.4207$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S34) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 1.88712$, $U = 0.86322$, $P = 0.7383$, *P* value computed via approximation method).

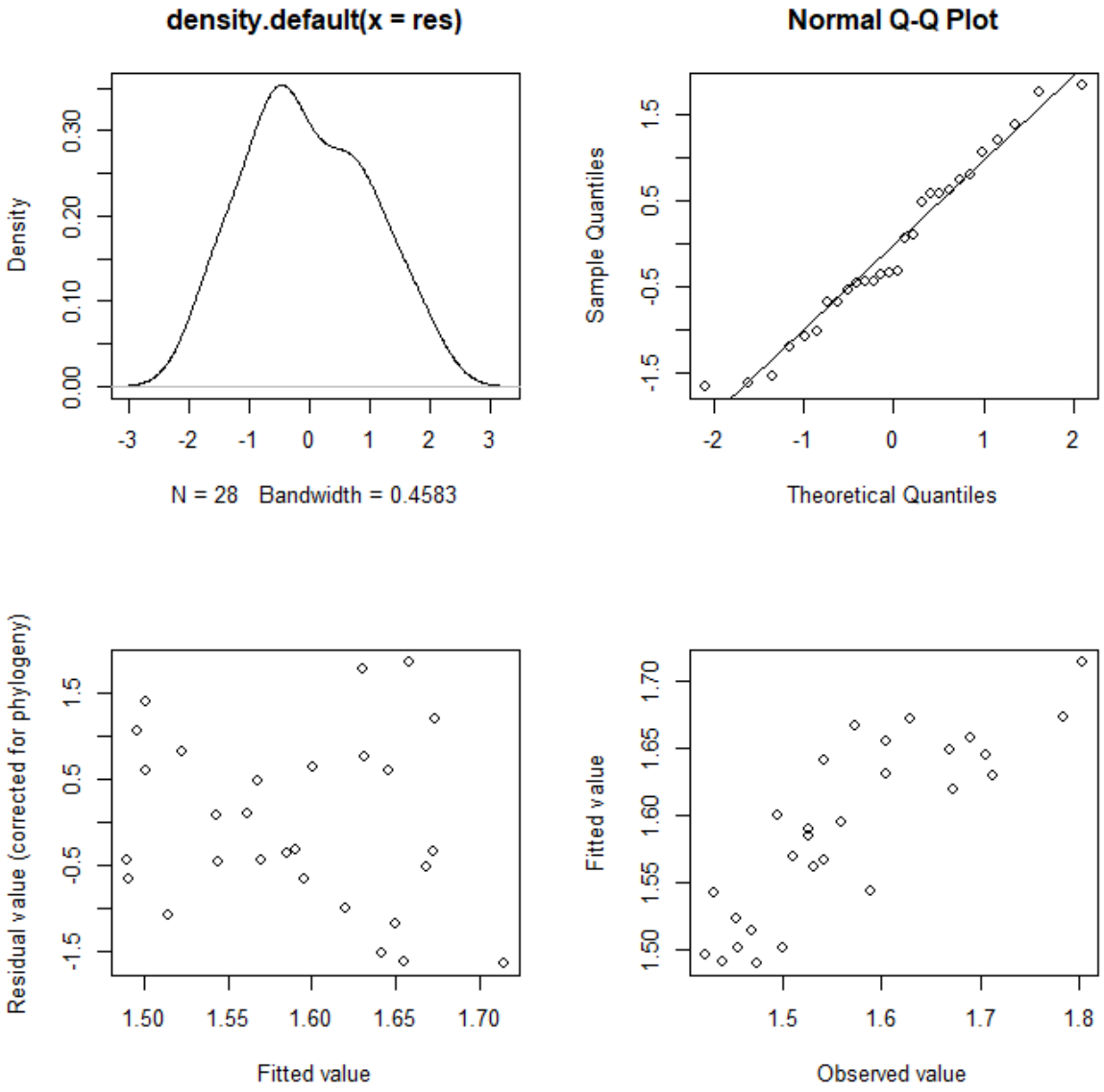


Figure S34 – Diagnostic plots for model 34.

Model 35: $T_R \sim \log_{10}(\text{TMI } \Omega_2\_a)$

Branch length transformations:

Pagel's λ [ML]: 1.000

lower bound: 0.000, *P* = 0.26028

upper bound: 1.000, *P* = 1

95.0% CI: (NA, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.750983	0.029958	58.4484	$< 2.2\ 10^{-16}$ *
$\log_{10}(\text{TMI } \Omega_2\_a)$	-0.175222	0.026801	-6.5378	$6.239\ 10^{-07}$ *

P values calculated using the exact method.

Adjusted R^2: 0.6072, AICc: -94.72086

F statistic: 42.74 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: $6.239\ 10^{-07}$ *, Corrected *P* value: $1.2478\ 10^{-06}$ *

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.96108, *P* = 0.3697, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S35) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.33872, *U* = 0.78992, *P* = 0.1984, *P* value computed via approximation method).

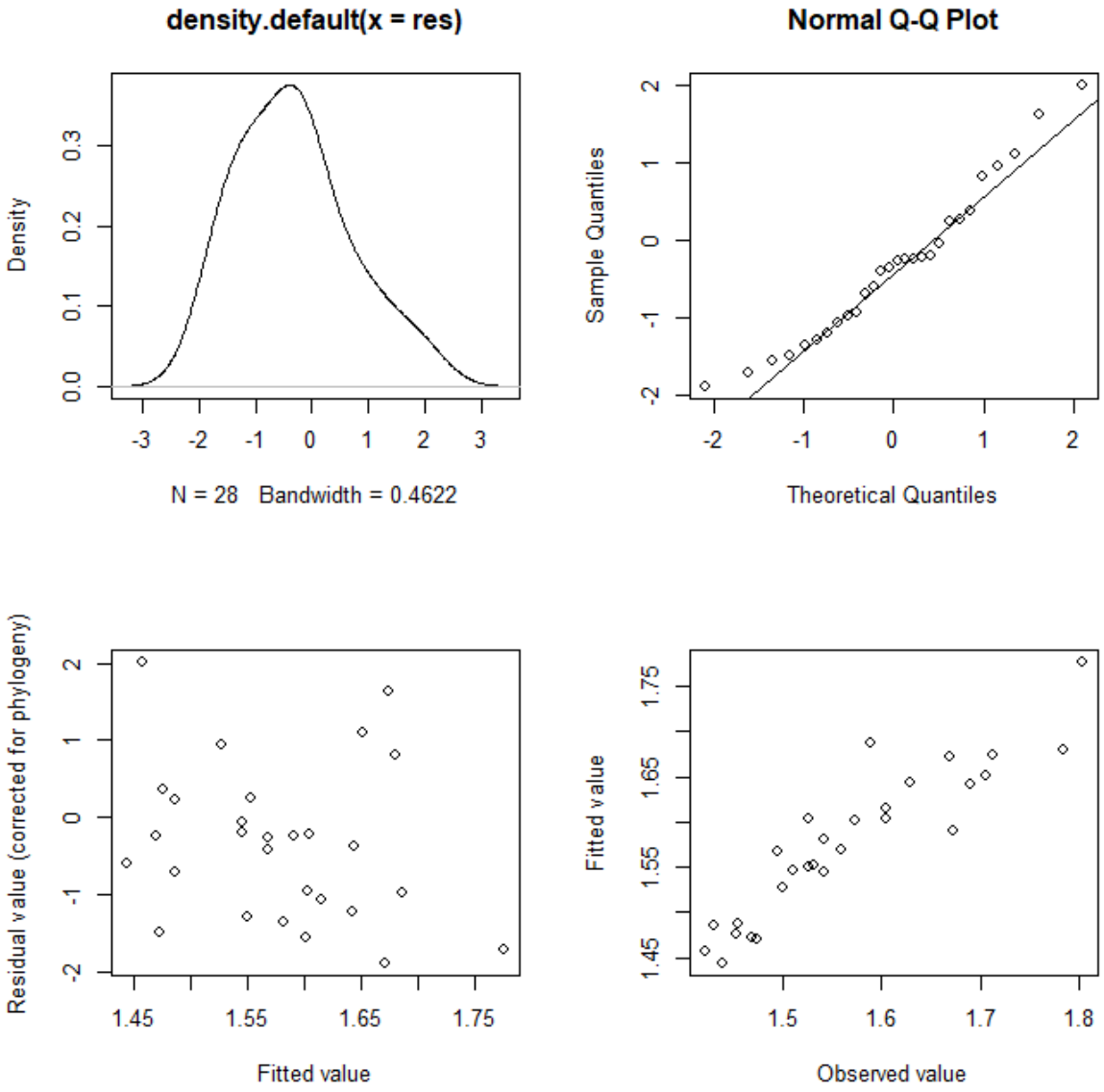


Figure S35 – Diagnostic plots for model 35.

Model 36: $T_R \sim \log_{10}(\text{TMI } \Omega_2\_p)$

Branch length transformations:

Pagel's λ [ML]: 1.000

lower bound: 0.000, $P = 0.0036249$ *

upper bound: 1.000, $P = 1$

95.0% CI: (0.346, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.801330	0.047216	38.1505	$< 2.2\ 10^{-16}$ *
$\log_{10}(\text{TMI } \Omega_2\_p)$	-0.166416	0.036016	-4.6206	$9.152\ 10^{-05}$ *

P values calculated using the exact method.

Adjusted R^2: 0.4298, AICc: -84.28196

F statistic: 21.35 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: $9.152\ 10^{-05}$ *, Corrected *P* value: $1.3728\ 10^{-04}$ *

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.95246$, $P = 0.2283$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S36) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 1.9516$, $U = 0.8537$, $P = 0.6238$, *P* value computed via approximation method).

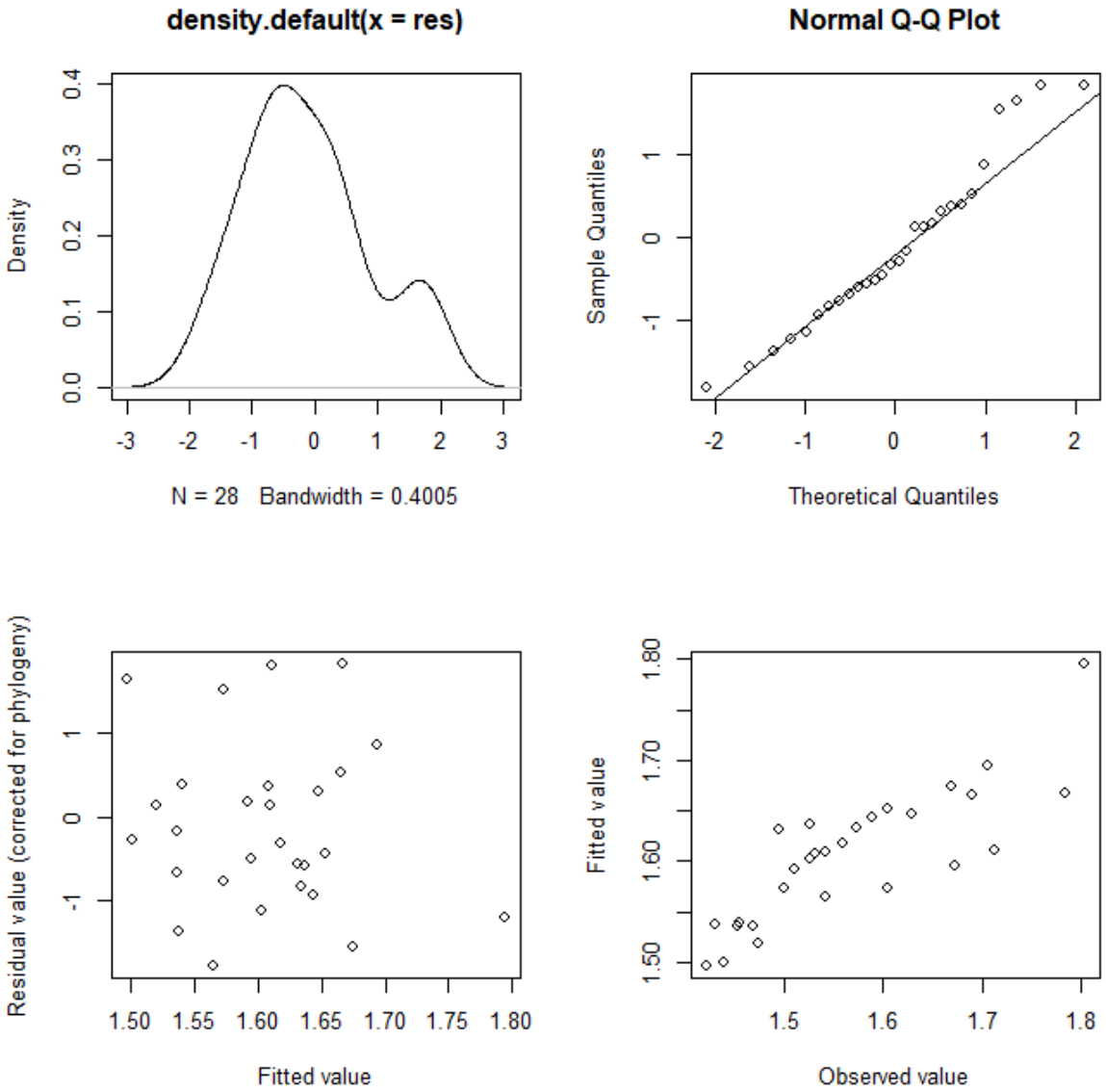


Figure S36 – Diagnostic plots for model 36.

Model 37: $T_R \sim \log_{10}(\text{TMI } \Omega_2\_l)$

Branch length transformations:

Pagel's λ [ML]: 1.000

lower bound: 0.000, $P = 9.5886\ 10^{-07}$ *

upper bound: 1.000, $P = 1$

95.0% CI: (0.763, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.773806	0.059037	30.0457	$< 2.2\ 10^{-16}$ *
$\log_{10}(\text{TMI } \Omega_2\_l)$	-0.145480	0.048201	-3.0182	0.005631 *

P values calculated using the exact method.

Adjusted R^2: 0.231, AICc: -75.9076

F statistic: 9.11 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: 0.005631 *, Corrected *P* value: $5.6310\ 10^{-03}$ *

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.95307$, $P = 0.2364$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S37) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.66783$, $U = 0.72663$, $P = 0.06032$, *P* value computed via approximation method).

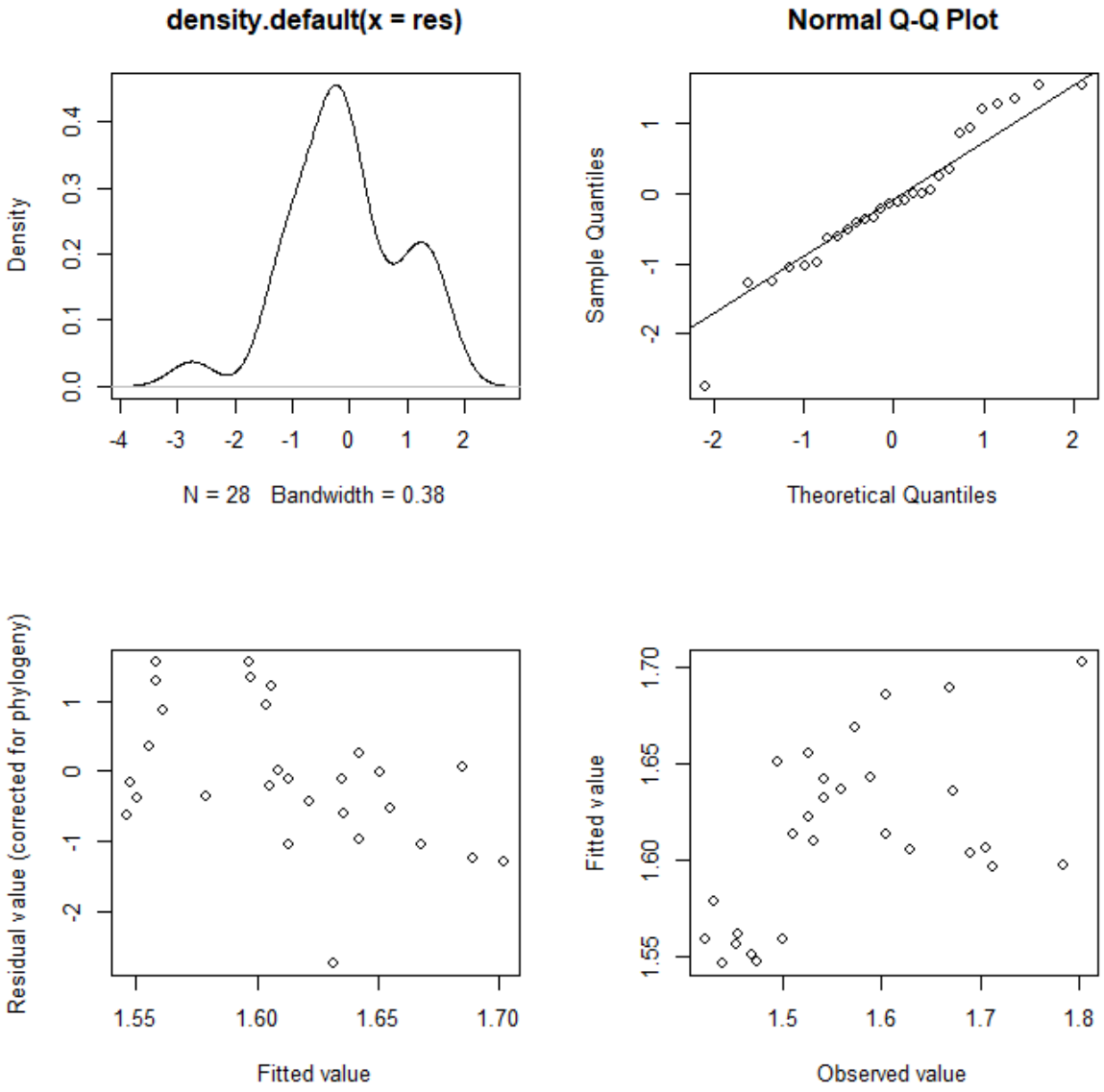


Figure S37 – Diagnostic plots for model 37.

11. Akaike weights used for the computation of the (size-corrected) weighted average Thermo-Motility Index

A total of six size-corrected Thermo-Motility Indices were computed (see Models 20-25), and are respectively related to the upper corner frequency (ω_2) and saturating velocity (Ω_2, which is inversely related to the gain to angular velocity: see Supplementary Methods, Biomechanics) of each of the three semicircular ducts. According to the corrected version of the Akaike Information Criterion (AICc), the size-corrected Thermo-Motility Index that best explains the temperature term is calculated from the saturating velocity of the anterior semicircular duct (TMI_Ω_2_a, AICc = -94.7 for clades). All other models are significantly less likely to inform about the temperature term (relative likelihood < 0.05), except for the size-corrected Thermo-Motility Index calculated from the upper corner frequency of the anterior semicircular duct (TMI_ω_2_a, AICc = - 93.0 for clades). Interestingly, it is the anterior semicircular canal that best explains the temperature term, likely because: (1) lateral canals correlate the most to agility of organisms (Cox & Jeffery 2010), blurring the relationship with body temperature, and (2) the anterior semicircular canal is the most easily measured. The lateral and posterior canals are prone to measurement errors resulting from the occasional presence of a 'secondary common crus' or from fusions between the bony canals and the vestibule. Akaike weights are applied to Thermo-Motility Indices of the anterior semicircular duct (calculated from the upper corner frequency and the saturating velocity) to calculate the size-corrected weighted average Thermo-Motility Index (TMI, see Supplementary Methods, Biomechanics), which is referred to as the Thermo-Motility Index in the main text and is used in subsequent analyses. Size-corrected weighted average Thermo-Motility Indices are provided in Supplementary Data 2, and illustrated in Fig. 4, along with ancestral state estimations computed using Model 66.

Model	AICc	Relative likelihood	Akaike Weights
32: TMI_ω_2_a	-92.98236	0.42	0.30
33: TMI_ω_2_p	-86.23574	0.01	-
34: TMI_ω_2_l	-82.31672	0.00	-
35: TMI_Ω_2_a	-94.72086	1.00	0.70
36: TMI_Ω_2_p	-84.28196	0.01	-
37: TMI_Ω_2_l	-75.9076	0.00	-

12. PGLS-R of the (size-corrected weighted average) Thermo-Motility Index against the cubic root of body mass

Model 38 was tested on specimens for which we know the body mass or predicted it, to check whether the (size-corrected weighted average) Thermo-Motility Index (TMI, see Supplementary Information, Biomechanics) was correlated to the cubic root of body mass ($\sqrt[3]{BM}$). As expected, we found no significant correlation because the (weighted average) Thermo-Motility Index has been size-corrected.

Model 38: $\log_{10}(TMI) \sim \log_{10}(\sqrt[3]{BM})$

Adjusted R^2: -0.002455, AICc: -93.50254
F statistic: 0.1674 on 1 and 339 DF
n = 341 (3DB+3DM+2DM sets)
P value: 0.6827

13. PGLS-R of the residual duct-morphology component of the (size-corrected weighted average) Thermo-Motility Index against the cubic root of body mass, and phylogenetic signal

Model 38 was tested on specimens for which we have the bony and membranous labyrinths, to check whether the residual duct-morphology component of the (size-corrected weighted average) Thermo-Motility Index (TMI_res, see Supplementary Methods, Biomechanics) was (1) correlated to the cubic root of body mass ($\sqrt[3]{BM}$) and (2) reflected phylogeny. We concluded that the residual duct-morphology component of the Thermo-Motility Index was not correlated to body mass but carried phylogenetic information. Therefore, we estimated ancestral states of the residual duct-morphology component of the (size-corrected weighted average) Thermo-Motility Index, at all nodes of a tree depicting phylogenetic relationships between tetrapod species for which we have membranous labyrinths (Fig. S38). We then used these ancestral states estimations to check whether the residual duct-morphology component of the (size-corrected weighted average) Thermo-Motility Index changed between the nodes of Amniota and Mammalia, and concluded that it decreased slightly (-2%).

Model 39: $\log_{10}(TMI\_res) \sim \log_{10}(\sqrt[3]{BM})$

Adjusted R^2: 0.006728, AICc: -36.81854
F statistic: 1.312 on 1 and 45 DF
n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)
P value: 0.2582

P values have been calculated using the exact method.

Predicted values at different nodes:

Clade	Most likely value	95% confidence interval
Amniota	4.06	[3.83;4.29]
Mammalia	3.98	[3.79;4.17]
Diapsida	4.07	[3.87;4.26]
Aves	4.21	[4.06;4.36]

Assumption testing:

- We tested whether phylogeny structured the distribution of the (log_{10}) residual duct-morphology component of the Thermo-Motility Index (Fig. S38). We concluded that phylogeny does partially structure the distribution (K = 0.492937, P = 4 10^{-04} *, P value calculated via approximation method) and should be taken into account to improve predictions of fossil specimens.
- We checked whether the residual duct-morphology component of the Thermo-Motility Index changed between the nodes of Amniota and Mammalia and concluded that it remained relatively stable along this lineage (2% differences).
- We tested whether the (log_{10}) residual duct-morphology component of the Thermo-Motility Index was correlated to the (log_{10}) cubic root of body mass. We found no significant correlations.

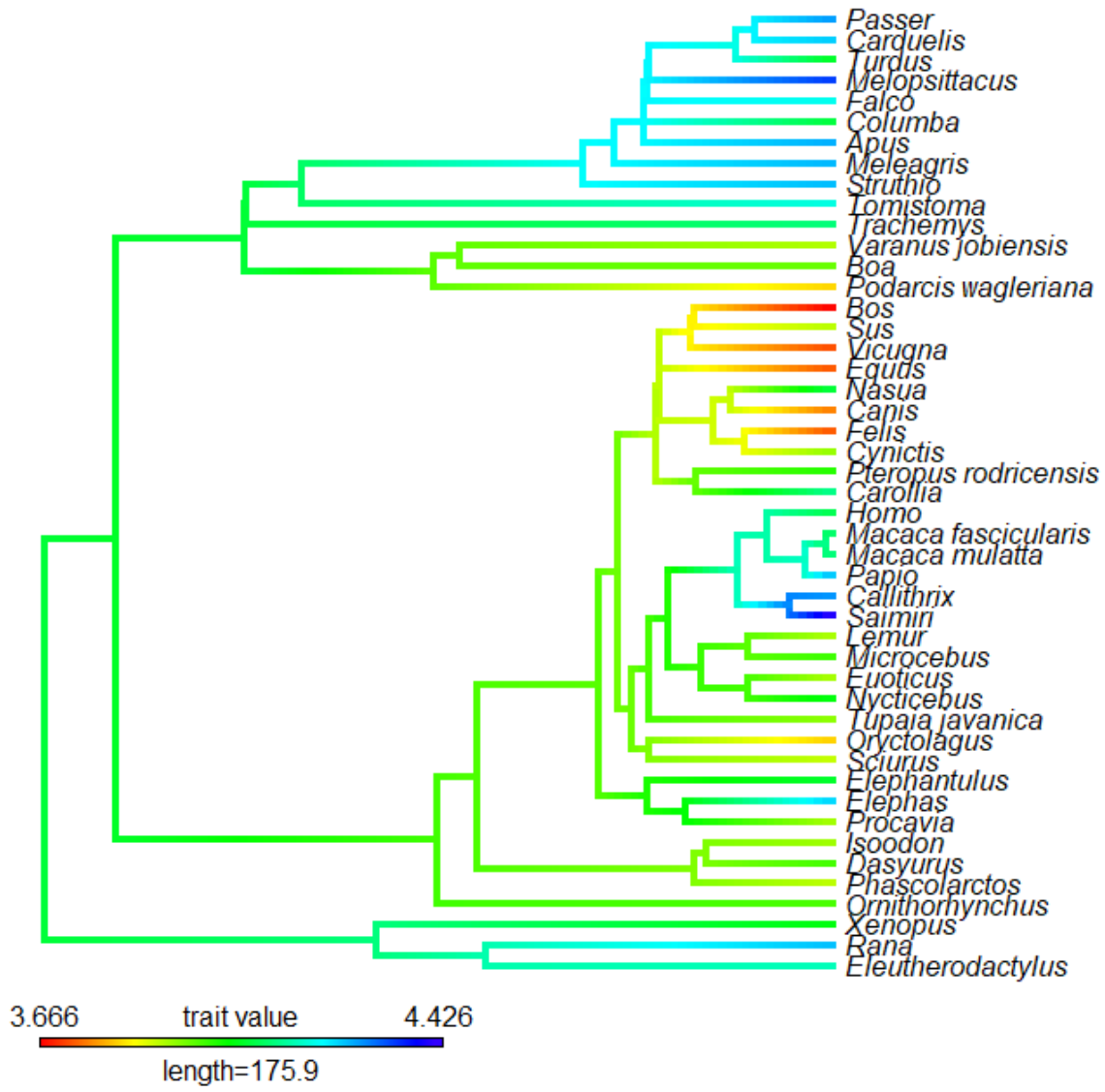


Figure S38 – Phylogenetic distribution of the (log_{10}) residual duct-morphology component of the (size-corrected weighted average) Thermo-Motility Index.

14. PGLS-R of the temperature term against the (size-corrected weighted average) Thermo-Motility Index, using species values or group averages

Models 40-43 were tested on specimens for which we know the body temperature, to assess correlations between the (size-corrected weighted average) Thermo-Motility Index (TMI, see Supplementary Information, Biomechanics) and a temperature term, which decreases when body temperature increases (T_R, see Supplementary Information, Biomechanics). Note that because relationships between size-corrected Thermo-Motility Indices and the temperature term were used to compute Akaike weights, to obtain the (size-corrected weighted average) Thermo-Motility Index (Models 32-37, Section 11), it is inappropriate to provide p-values for Models 40-42 (Mundry 2011, Mundry, personal communication). As theoretically expected, models 40b and 42 clearly demonstrate the existence of a temperature signal in the Thermo-Motility Index, whether analysing species or clades (AICc Model 40b = -25.3 << AICc Null Model for species = -13.5; AICc Model 42 = -20.9 << AICc Null Model for clades = 6.7). However, models 40 and 41 demonstrate the existence of a TMI signal in body temperature, but much weaker for species than for clades (AICc Model 40 = -846.3 < AICc Null Model for species = -845.7; AICc Model 41 = -98.4 << AICc Null Model for clades = -70.7). Simulations were used to better understand the discrepancies in the relationships between the Thermo-Motility Index and body temperature at the species level, including the asymmetry in signal strength and in the slope values of Models 40 and 40b. Results are discussed in the Simulation section, below. Models 40 and 41 show that the (size-corrected weighted average) Thermo-Motility Index is a very good predictor of the temperature term (and by extension of body temperature) when clades are considered (TMI_{Clades}, Model 41, $R^2 = 0.86$), but not when individual species are considered ($TMI_{Species}$, Model 40, $R^2 = 0.01$). Models 40b and 42 show that the slope between the (size-corrected weighted average) Thermo-Motility Index and the temperature term is steeper than expected ([-1.47; -3.7] instead of -1, see Supplementary Methods, Biomechanics), which suggests that angular head motion is probably positively correlated to body temperature and further drives the increase of the (size-corrected weighted average) Thermo-Motility Index when body temperature increases. Model 43 shows that while being significantly correlated to the temperature term ($P << 0.001$) and showing a high coefficient of determination ($R^2 = 0.61$), the standard average (i.e., using equal weights instead of Akaike weights) of the six size-corrected Thermo-Motility Indices ($TMI\_unweighted_{Clades}$) is less informative concerning the temperature term than the (size-corrected weighted average) Thermo-Motility Index (AICc Model 41 = -98.4 << AICc Model 43 = -91.2), confirming that using a weighted average of size-corrected Thermo-Motility Indices is beneficial for increasing the accuracy of body temperature predictions. Model 41 is used in Section 24 to compute probability distribution of predicted body temperatures for fossil synapsids. The

relationship between the (size-corrected weighted average) Thermo-Motility Index and body temperature is illustrated in Figure 2.

Models 40 and 40b are based on all relevant specimens, while Models 41-43 are based on average values for the clades Acanthopterygii, Anguimorpha, Anura, Atlantogenata, Batoidea, Caudata, Crocodylia, Elopomorpha, Euarchontoglires, Galloanserae, Gekkota, Gymnophiona, Holocephali, Iguania, Lacertoidea, Laurasiatheria, Marsupialia, Monotremata, Neoaves, Otocephala, Palaeognathae, Paracanthopterygii, Protacanthopterygii, Rhynchocephalia, Scincomorpha, Selachii, Serpentes, and Testudines.

Simulation:

We used simulations to better understand why the slope from Model 40 ($T_R \sim \log_{10}(TMI_{Species})$) is not the inverse of the slope from Model 40b ($\log_{10}(TMI_{Species}) \sim T_R$), and why signal strength from the independent variable was very high in one case (Model 40b) and, unexpectedly, very weak in its reciprocal (Model 40). Simulations allowed us to evaluate and bypass factors typical of real-world samples including sampling uniformity, outliers, and standard deviation. Simulations were performed in R 4.0.3 with the packages phytools 4.0.5, phangorn 2.5.5, geiger 2.0.7, and caper 4.0.4. The scripts are provided in the RDATA zip file.

The simulation was set to mimic our dataset as much as possible. Firstly, we simulated a phylogenetic tree of 1000 extant taxa with a root dated at 473 Ma. Then we isolated a clade of 504 species that we labelled as terrestrial. Species not belonging to this clade (496) were labelled as marine. Among the terrestrial clade we isolated two clades of 109 and 150 species that we labelled as endotherms. Marine and terrestrial species which did not belong to these clades were labelled as ectotherms. The proportion of species was set to resemble what can be observed in extant vertebrates. We then created four pruned trees (i.e., subtrees) corresponding to the two clades of endotherms, the clade of terrestrial species minus the endotherms and the clade of remaining species (i.e., marine).

We simulated the evolution of body temperature on each of these pruned trees using Brownian motion and an evolutionary rate σ^2 of 0.144, which corresponds to the value measured from our dataset when we enforce Brownian evolution. For simulations with the marine pruned tree, we used an ancestral body temperature of 18.5°C (predicted as ancestral for vertebrates in our actual dataset) and bounds of [-2; 30°C] to mimic conditions in extant seas and oceans in which fishes live. For simulations with the ectotherm terrestrial pruned tree (mimicking reptiles), we used an ancestral body temperature of 27°C (similar to the average value for terrestrial ectotherms of our dataset) and bounds of [15; 42°C] to reflect what can be observed in extant terrestrial ectothermic vertebrates. For simulations with the endotherm subtrees, we wanted to mimic

conditions observed in extant mammals and birds. Hence, we used ancestral body temperatures of 37°C and 42°C, and bounds of [31; 39°C] and [39; 45°C] mimicking mammals and birds, respectively.

For each species, the simulated body temperature was transformed into the corresponding temperature term (T_R, see Supplementary Information, Biomechanics), which was used with regression coefficients from Model 42 to compute the fitted TMI. Then, a residual of the TMI was randomly sampled from a distribution of mean = -0.008 and standard deviation = 0.258, and added to the corresponding fitted TMI to obtain the (size-corrected weighted average) Thermo-Motility Index of the simulated species. The mean and standard deviation of the distribution from which residuals have been sampled were obtained from our original dataset.

PGLS regressions were then fitted on simulated data (for species and clades), to compare with results obtained from Models 40-42. All regressions provided significant results, but the p-value of the model $T_R \sim \log_{10}(TMI_{Species})$ was higher than the other three regressions ($3.745\ 10^{-13} > 2.2\ 10^{-16}$). Furthermore, the R^2 of the model $T_R \sim \log_{10}(TMI_{Species})$ was much lower than the R^2 of its reciprocal model $\log_{10}(TMI_{Species}) \sim T_R$ ($0.05 << 0.64$). R^2 were equal for clade models (0.98) and higher than for species models. Finally, all models captured the correct slope (or its inverse when applicable) except the model $T_R \sim \log_{10}(TMI_{Species})$, which provided a much lower slope than the one used to simulate the data (-0.01 *vs* -0.27).

As described in the Biomechanics section, the TMI is theoretically expected to adapt to changes in body temperature. What our simulations demonstrate is that when you enforce that link, (1) the signal is better retrieved for clades than species (higher R^2) and (2) the coefficients of the relationship are better estimated with the TMI than the temperature term as the dependent variable. Interestingly, R^2 for model $T_R \sim \log_{10}(TMI_{Species})$ increases to 0.46 when the standard deviation used for the distribution of the residual TMI is divided by 4. The estimated slope also gets closer to slope used in the simulation (-0.12 *vs* 0.27). This strongly suggests that variation in the TMI that is independent from body temperature is blurring the relationship between the two parameters.

In conclusion, our simulations demonstrate that although Model 40 may suggest that the link between the TMI and body temperature is weak, Model 42 is the most suitable model to assess the relationship between the two parameters. It is important to note that, from a practical point of view, Models 40 and 41 are the most relevant because they inform us about the predictive power of using the TMI to estimate body temperature. Note that if a bidirectional relationship was theoretically predicted between TMI and T_R, instead of TMI adapting to T_R, we should use a phylogenetic reduced major axis to assess the correlation between the two variables (Smith 2009). Such a regression provides significant results with our dataset when testing against a slope of -1 (p-value $<< 0.001$, $R^2 = 0.02$, slope = -6.1, Pagel's $\lambda = 0.94$).

Model 40: $T_R \sim \log_{10}(TMI_{Species})$

Adjusted R^2: 0.007264, AICc: -846.2728

n = 230 (3DB+3DM+2DM sets where only specimens with measurements of body temperature were included)

Model 40b: $\log_{10}(TMI_{Species}) \sim T_R$

Branch length transformations:

Pagel's λ [ML]: 0.716

95.0% CI: (0.510, 0.853)

Coefficients:

	Estimate	Std. Error
(Intercept)	3.06505	0.56566
T_R	-1.46839	0.33970

Adjusted R^2: 0.07, AICc: -25.33475

n = 230 (3DB+3DM+2DM sets where only specimens with measurements of body temperature were included)

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.99693, P = 0.9365, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S38b) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.80320, U = 0.96554, P = 0.5442, P value computed via approximation method).

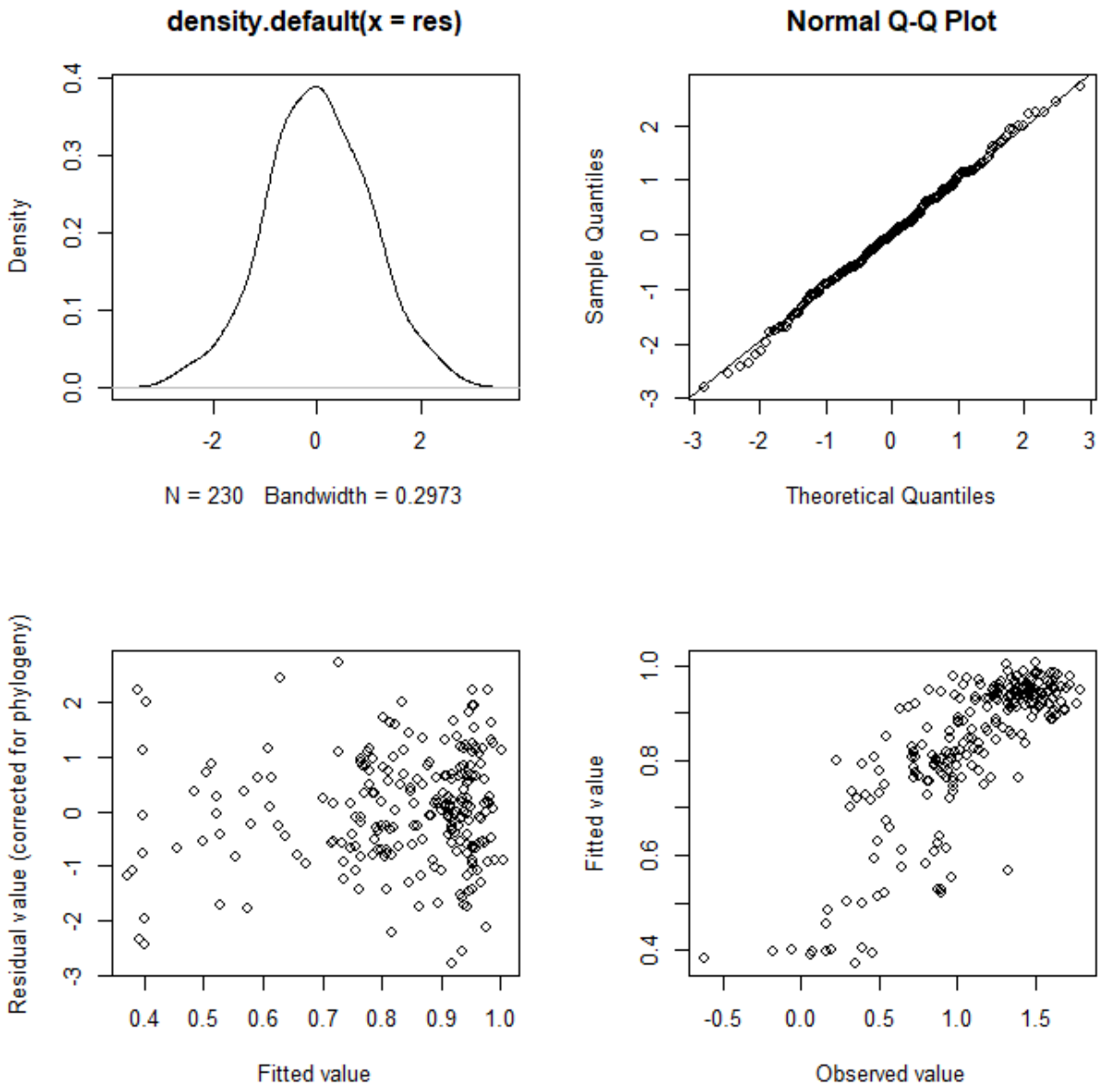


Figure S38b – Diagnostic plots for model 40b.

Model 41: $T_R \sim \log_{10}(TMI_{Clades})$

Branch length transformations:

Pagel's λ [ML]: 0.000

95.0% CI: (NA, NA)

Coefficients:

	Estimate	Std. Error
(Intercept)	1.785333	0.018336
$\log_{10}(TMI_{Clades})$	-0.233125	0.018047

Adjusted R^2: 0.86, AICc: -98.42916

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.97978, P = 0.845, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S39) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.32337, U = 0.79267, P = 0.2086, P value computed via approximation method).

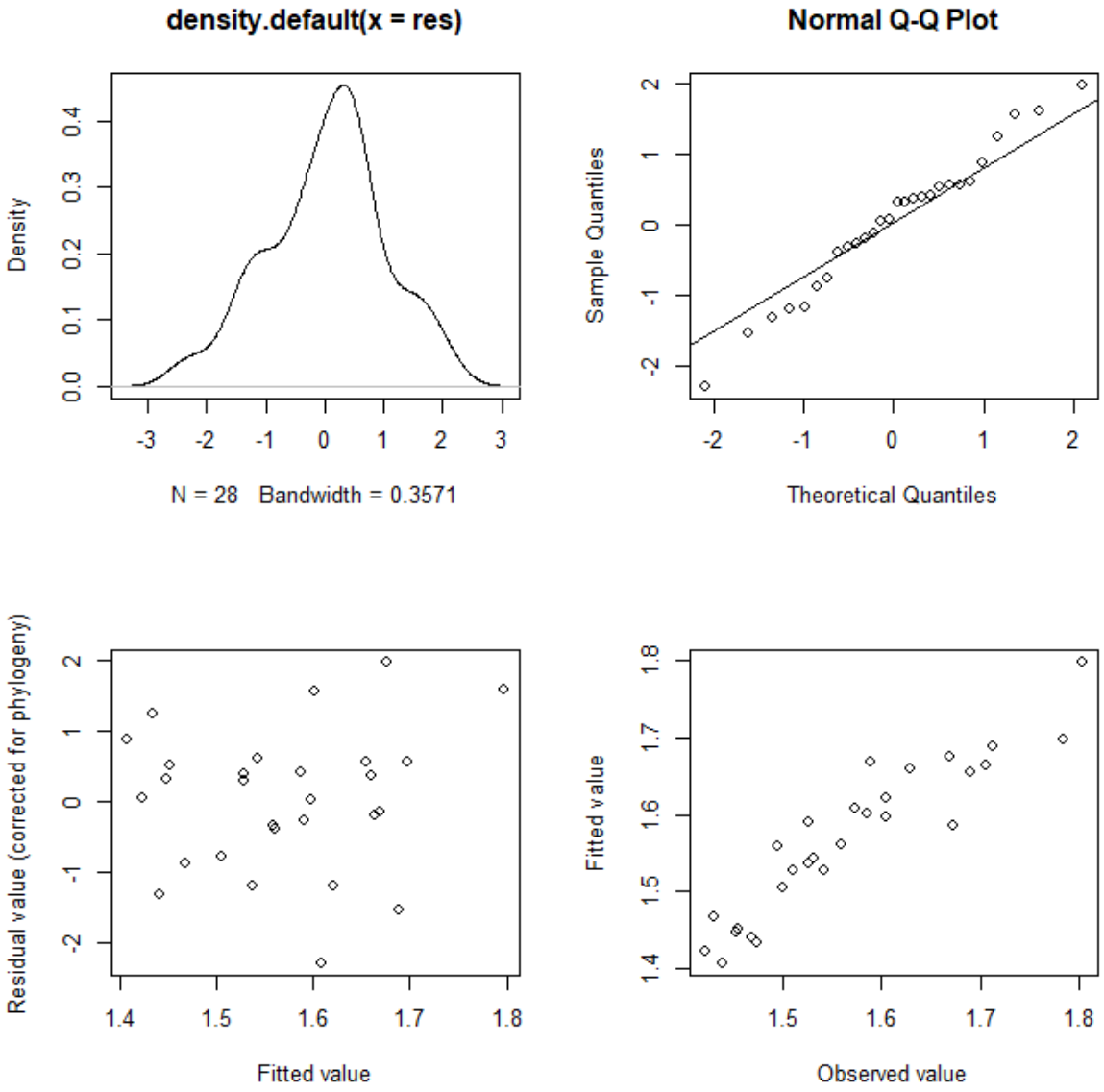


Figure S39 – Diagnostic plots for model 41.

Model 42: $\log_{10}(TMI_{Clades}) \sim T_R$

Branch length transformations:

Pagel's λ [ML]: 0.000

95.0% CI: (NA, NA)

Coefficients:

	Estimate	Std. Error
(Intercept)	6.75065	0.45194
T_R	-3.71127	0.28731

Adjusted R^2: 0.86, AICc: -20.93762

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.96297, P = 0.4091, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S40) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.24008, U = 0.80727, P = 0.272, P value computed via approximation method).

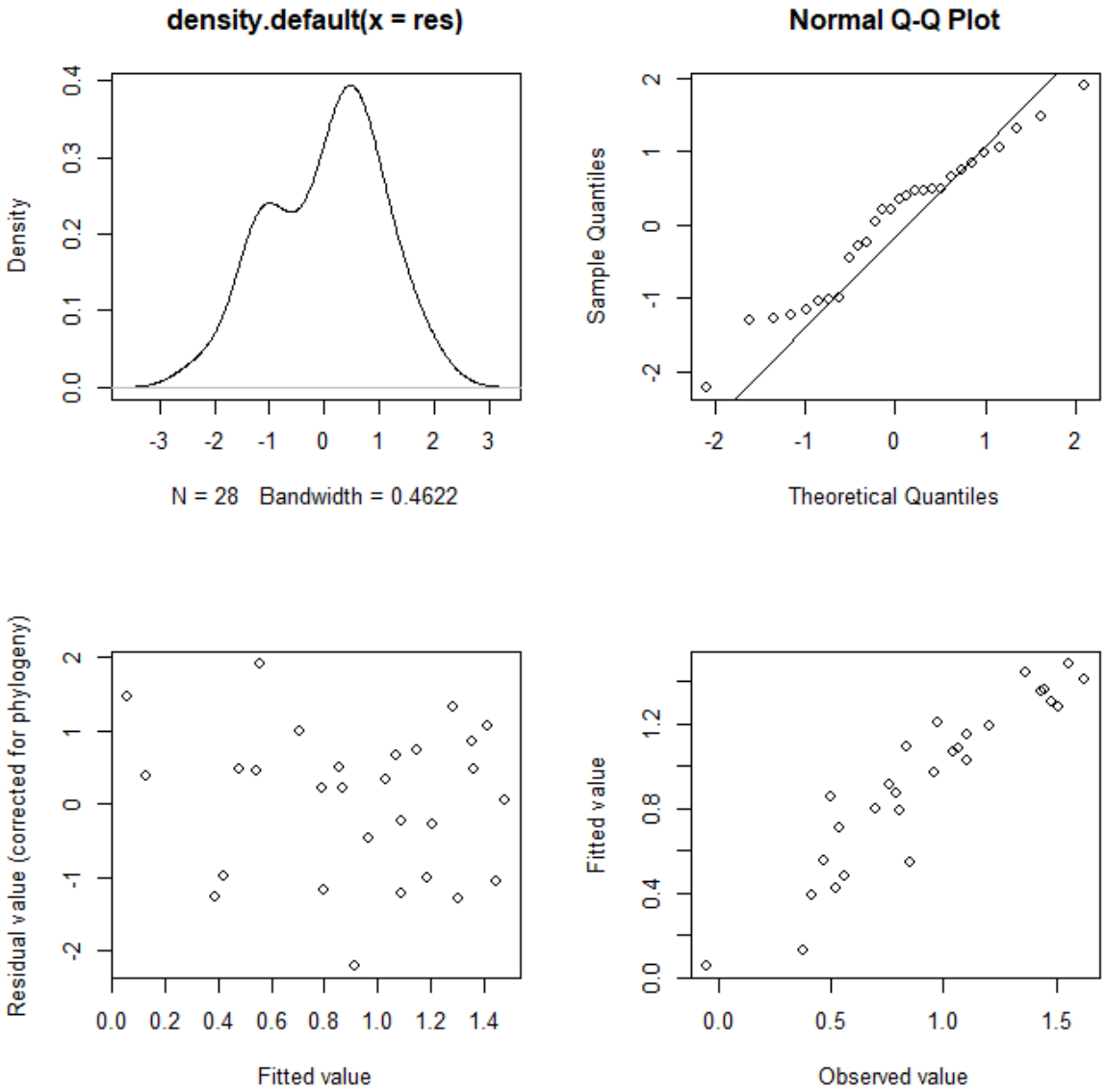


Figure S40 – Diagnostic plots for model 42.

Model 43: $T_R \sim \log_{10}(TMI\_unweighted_{Clades})$

Branch length transformations:

Pagel's λ [ML]: 0.782

lower bound: 0.000, $P = 0.082106$

upper bound: 1.000, $P = 0.64897$

95.0% CI: (NA, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.825289	0.036987	49.3495	$< 2.2\ 10^{-16}$ *
$\log_{10}(TMI\_unweighted_{Clades})$	-0.250376	0.038060	-6.5784	$5.633\ 10^{-07}$ *

P values calculated using the exact method.

Adjusted R^2: 0.6102, AICc: -91.15107

F statistic: 43.28 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: $5.633\ 10^{-07}$ *, Corrected *P* value: $5.633\ 10^{-07}$ *

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.97846$, $P = 0.812$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S41) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.18631$, $U = 0.81641$, $P = 0.3208$, *P* value computed via approximation method).

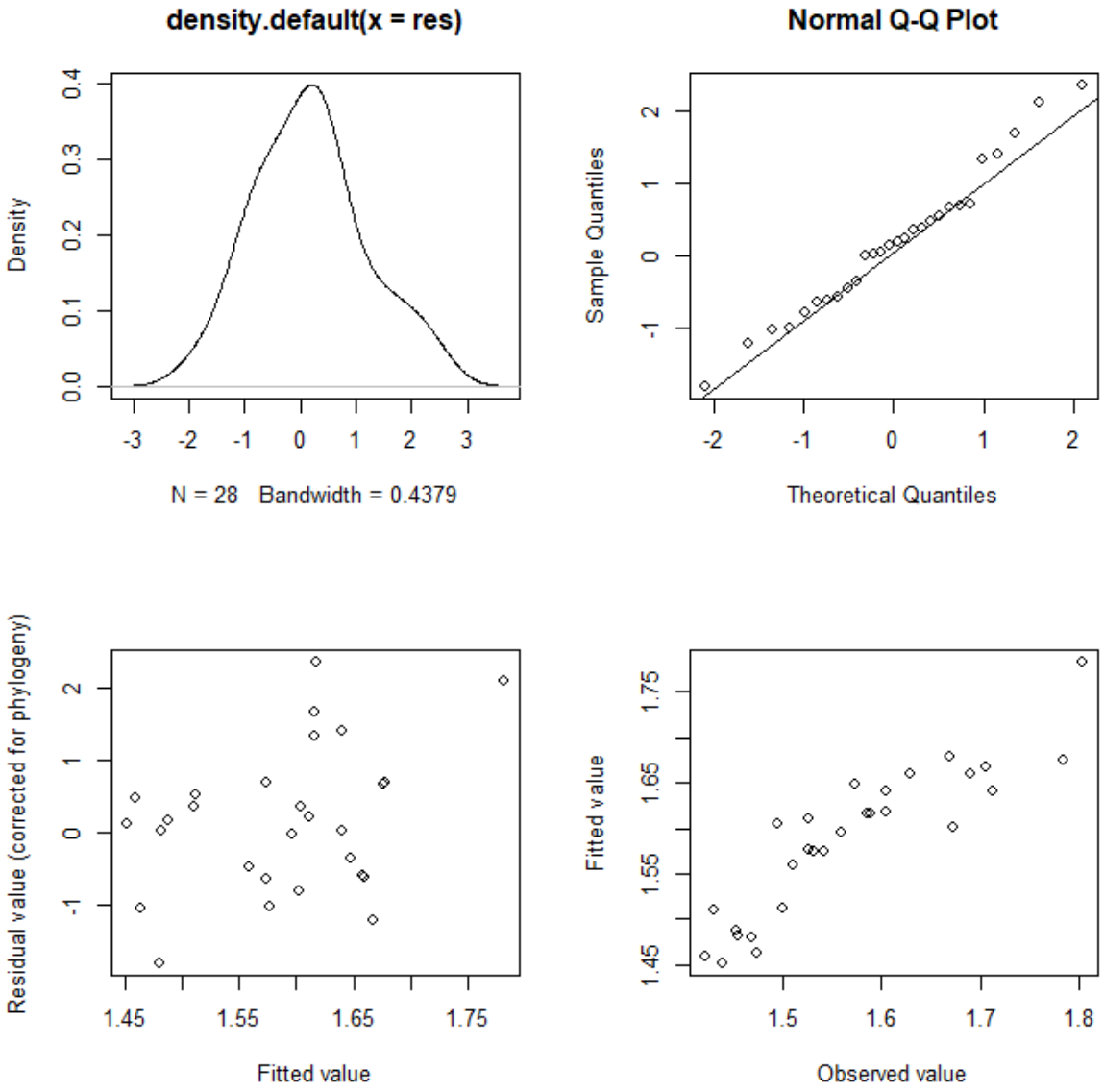


Figure S41 – Diagnostic plots for model 43.

15. Distribution of the (size-corrected weighted average) Thermo-Motility Index and comparison with the standard average of all size-corrected Thermo-Motility Indices

We checked whether the distribution of the standard average of all size-corrected Thermo-Motility Indices differed from the distribution of the (size-corrected weighted average) Thermo-Motility Index that we use in the main text and subsequent analyses (Figure S42). We concluded that while the (size-corrected weighted average) Thermo-Motility Index allows for finer predictions of body temperatures (see Section 14), its overall distribution is similar to that of the standard average of all size-corrected Thermo-Motility Indices. This suggest that only using information from the anterior semicircular canal in the (size-corrected weighted average) Thermo-Motility Index is not biasing the results and, as a corollary, this confirms that the main conclusions of our study are not a by-product of our averaging procedure.

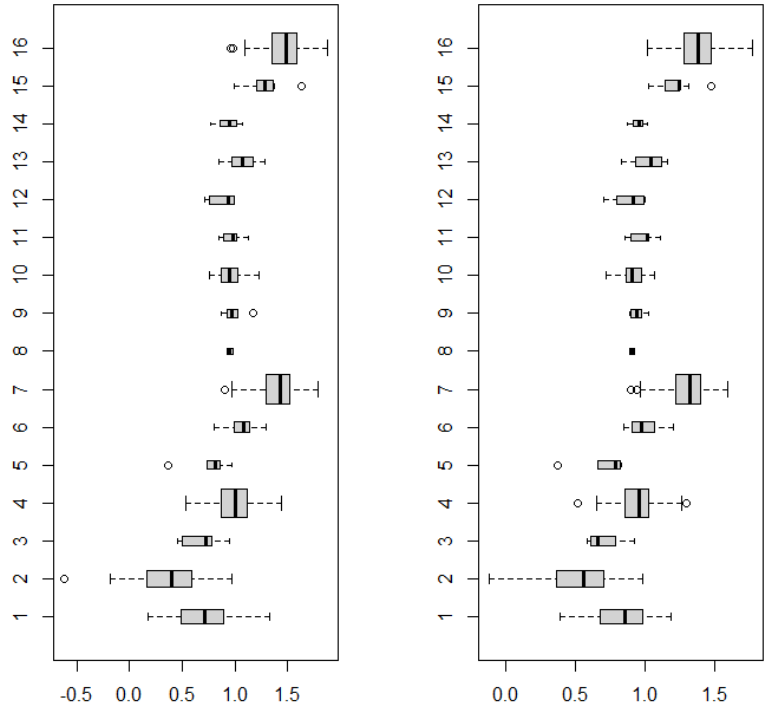


Figure S42 – Distribution of the (size-corrected weighted average) Thermo-Motility Index (on the left) and the average of all size-corrected Thermo-Motility Indices (on the right). Boxplot centre, median; box boundaries, first and third quartiles; whiskers, 1.5 × IQR from boundaries. Individual dots represent outliers. 1. Chondrichtyans (n =19), 2. Actinopterygians (n =24), 3. Amphibians (n =7), 4. Lepidosaurs (n = 67), 5. Turtles (n = 6), 6. Crocodilians (n = 9), 7. Birds (n = 72), 8. Non-therapsid synapsids (n = 2), 9. Non-neotherapsid therapsids (n = 5), 10. Anomodonts (n = 16), 11. Gorgonopsians (n = 5), 12. Therocephalians (n = 6), 13. Non-probainognathian cynodonts (n = 9), 14. Non-mammaliamorph probainognathians (n = 3), 15. Non-mammal mammaliamorphs (n = 9), 16. Mammals (n = 77).

16. PGLS-R of the components of the (size-corrected weighted average) Thermo-Motility Index against the (size-corrected weighted average) Thermo-Motility Index

Models 44-46 were used to assess which component of the (size-corrected weighted average) Thermo-Motility Index (i.e., fitted duct-morphology component TMI_fitted, residual duct-morphology component TMI_res, physicochemical component TMI_endolymph, see Supplementary Methods, Biomechanics) best explained the (size-corrected weighted average) Thermo-Motility Index (TMI: see Supplementary Methods, Biomechanics). These models were tested on specimens for which we have the bony and membranous labyrinths, therefore reducing the uncertainty in their computed (size-corrected weighted average) Thermo-Motility Index to uncontrolled variation of the physiochemical component only (i.e., actual deviations from low-viscosity or high-viscosity endolymph, see Supplementary Methods, Biomechanics). Models 44 and 46 provide significant regressions ($P < 0.05$) with coefficients of determination (R^2) of 0.46 and 0.13, respectively, whereas Model 45 is not significant. These results are reassuring in that they demonstrate that the fitted duct-morphology, which is the only component that can be directly calculated from measurements available in fossil specimens, is the most influential component of the (size-corrected weighted average) Thermo-Motility Index.

Model 44: $\log_{10}(\text{TMI}) \sim \log_{10}(\text{TMI\_fitted})$

Branch length transformations:

Pagel's λ [ML]: 0.833

lower bound: 0.000, $P = 6.7942\ 10^{-07}$ *

upper bound: 1.000, $P = 0.0019365$ *

95.0% CI: (0.551, 0.970)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.20517	0.16326	1.2567	0.2153
$\log_{10}$(TMI_fitted)	0.57417	0.09083	6.3214	1.04 10-07 *

P values calculated using the exact method.

Adjusted R^2: 0.4586, AICc: -49.67936

F statistic: 39.96 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: $1.04\ 10^{-07}$ *, Corrected P value: $3.1200\ 10^{-07}$*

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98297, P = 0.7172, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S43) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.30250, U = 0.88224, P = 0.4278, P value computed via approximation method).

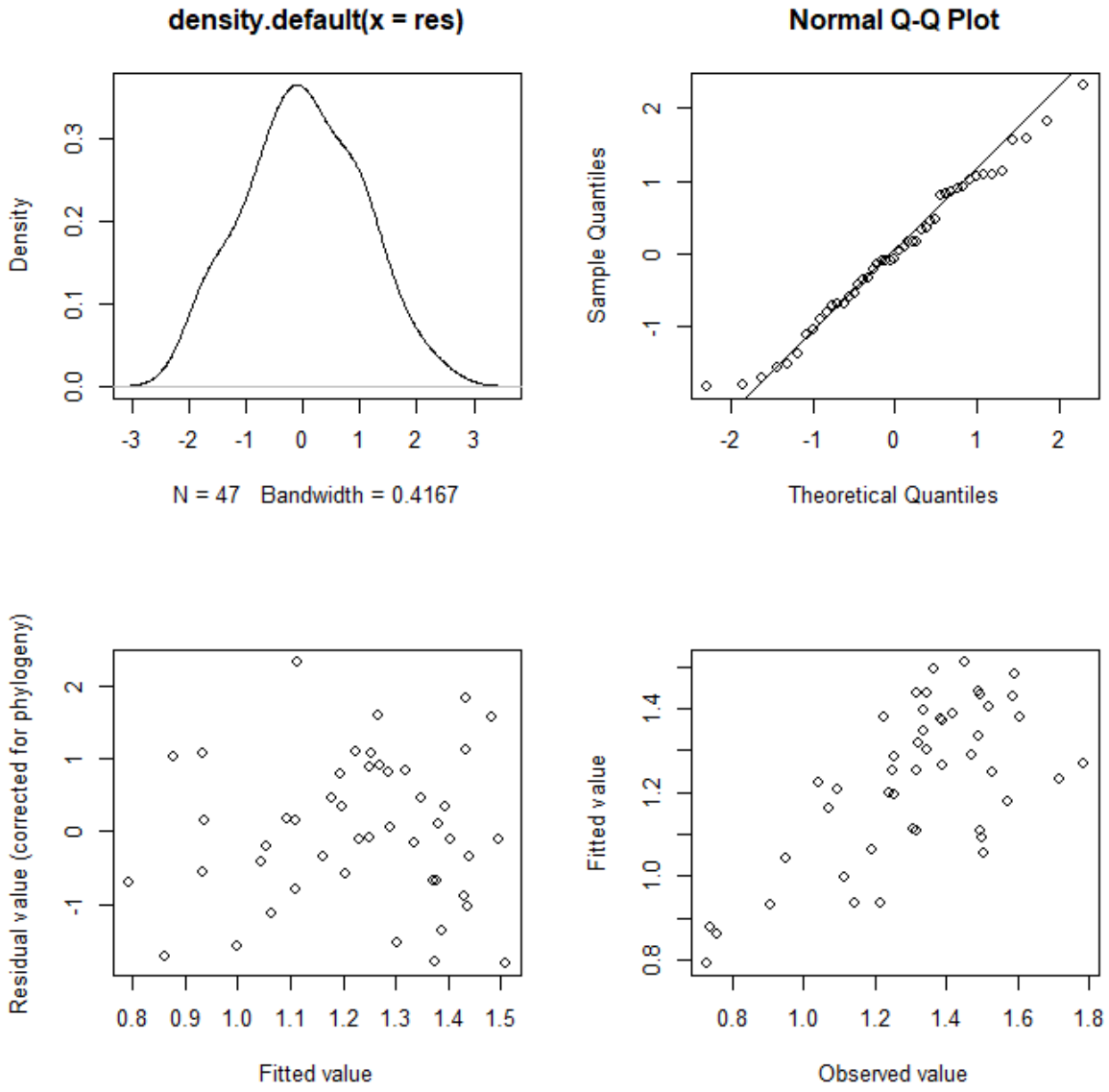


Figure S43 – Diagnostic plots for model 44.

Model 45: $\log_{10}(TMI) \sim \log_{10}(TMI\_res)$

Adjusted R^2: -0.0008264, AICc: -20.80875

F statistic: 0.962 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: 0.3319, Corrected *P* value: 0.3319

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Model 46: $\log_{10}$(TMI) ~ $\log_{10}$(TMI_endolymph)

Branch length transformations:

Pagel's λ [ML]: 0.735

lower bound: 0.000, $P = 3.596\ 10^{-07}$ *

upper bound: 1.000, $P = 1.6361\ 10^{-06}$ *

95.0% CI: (0.418, 0.925)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	10.49489	3.40658	3.0808	0.003516 *
$\log_{10}$(TMI_endolymph)	2.07928	0.75384	2.7583	0.008370 *

P values calculated using the exact method.

Adjusted R^2: 0.1256, AICc: -26.7818

F statistic: 7.608 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: 0.00837 *, Corrected *P* value: 1.2555 10-02 *

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.97421$, $P = 0.3793$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S44) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.4136$, $U = 0.8706$, $P = 0.3058$, *P* value computed via approximation method).

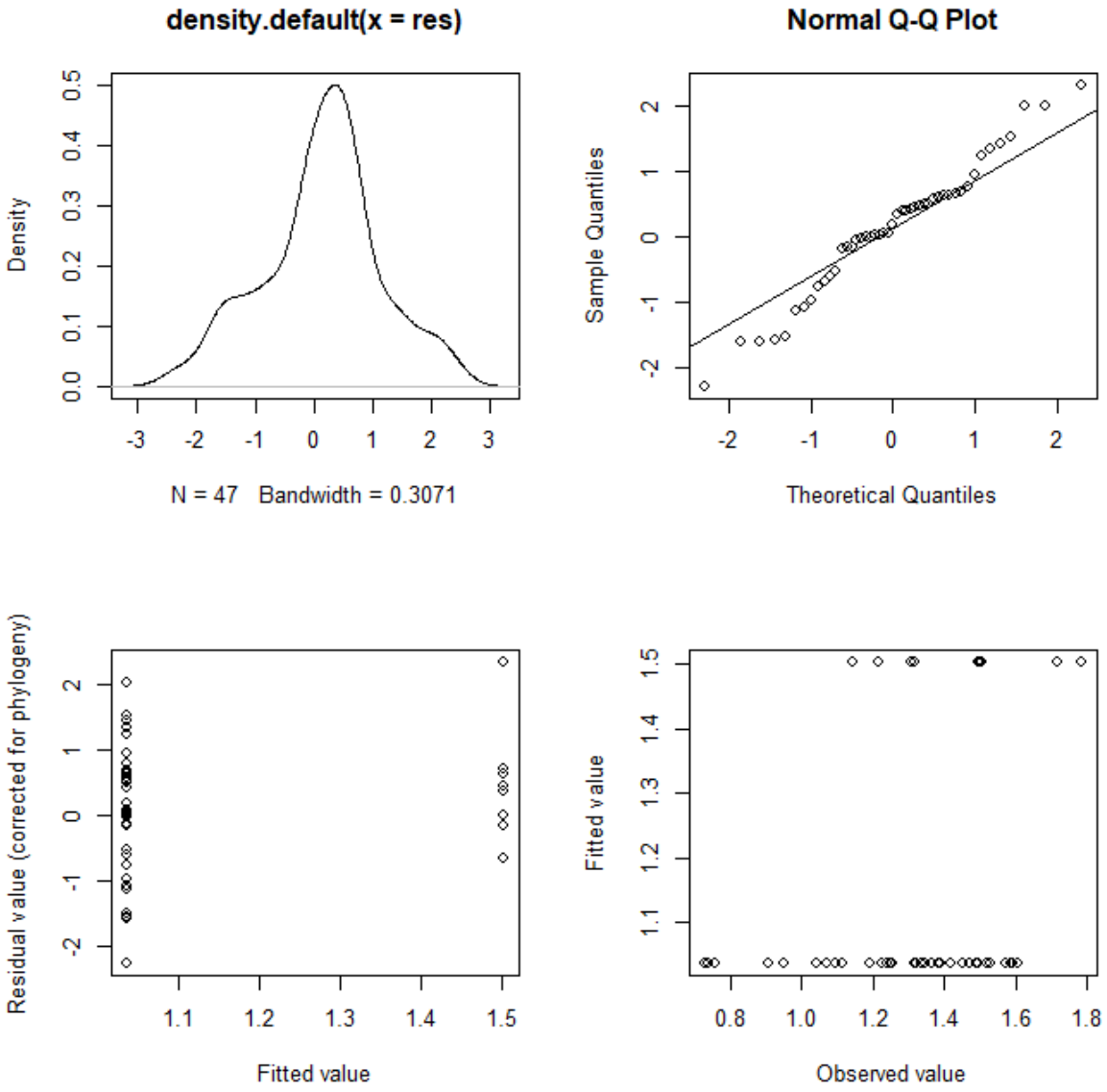


Figure S44 – Diagnostic plots for model 46.

17. PGLS-R of the estimated (size-corrected weighted average) Thermo-Motility Index against the actual (size-corrected weighted average) Thermo-Motility Index

Model 47 provides a very significant regression ($P << 0.001$), with a very high coefficient of determination ($R^2 = 0.75$). It was tested on specimens for which we have the bony and membranous labyrinths, and was used to assess if the (size-corrected weighted average) Thermo-Motility Index, calculated the same way as for fossil specimens (TMI_estimated), is a good estimator of the (size-corrected weighted average) Thermo-Motility Index, as computed from complete information of membranous semicircular duct morphometry (TMI). Results suggest that our framework for estimating the (size-corrected weighted average) Thermo-Motility Index of fossil species provides relatively accurate estimations of the 'true' (size-corrected weighted average) Thermo-Motility Index ($R^2 = 0.75$). They also demonstrate that using estimated residual duct-morphology components along with fitted duct-morphology components is significantly more accurate than just using the latter (AICc Model 47 = -64.1 << AICc Model 43 = -49.7).

Model 47: $\log_{10}(TMI) \sim \log_{10}(TMI\_estimated)$

Branch length transformations:

Pagel's λ [ML]: 0.000

lower bound: 0.000, $P = 1$

upper bound: 1.000, $P = 3.0899\ 10^{-08}$ *

95.0% CI: (NA, 0.487)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.271674	0.089915	3.0215	0.004139
$\log_{10}$(TMI_estimated)	0.791128	0.066800	11.8432	$1.998\ 10^{-15}$ *

P values calculated using the exact method.

Adjusted R^2: 0.7517, AICc: -64.10206

F statistic: 140.3 on 1 and 45 DF

n = 47 (3DM set where only specimens with measurements of cross-sectional radii of bony canals were included)

P value: $2.007\ 10^{-15}$ *

P values have been calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.97239, P = 0.3256, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S45) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.66134, U = 0.84268, P = 0.1358, P value computed via approximation method).

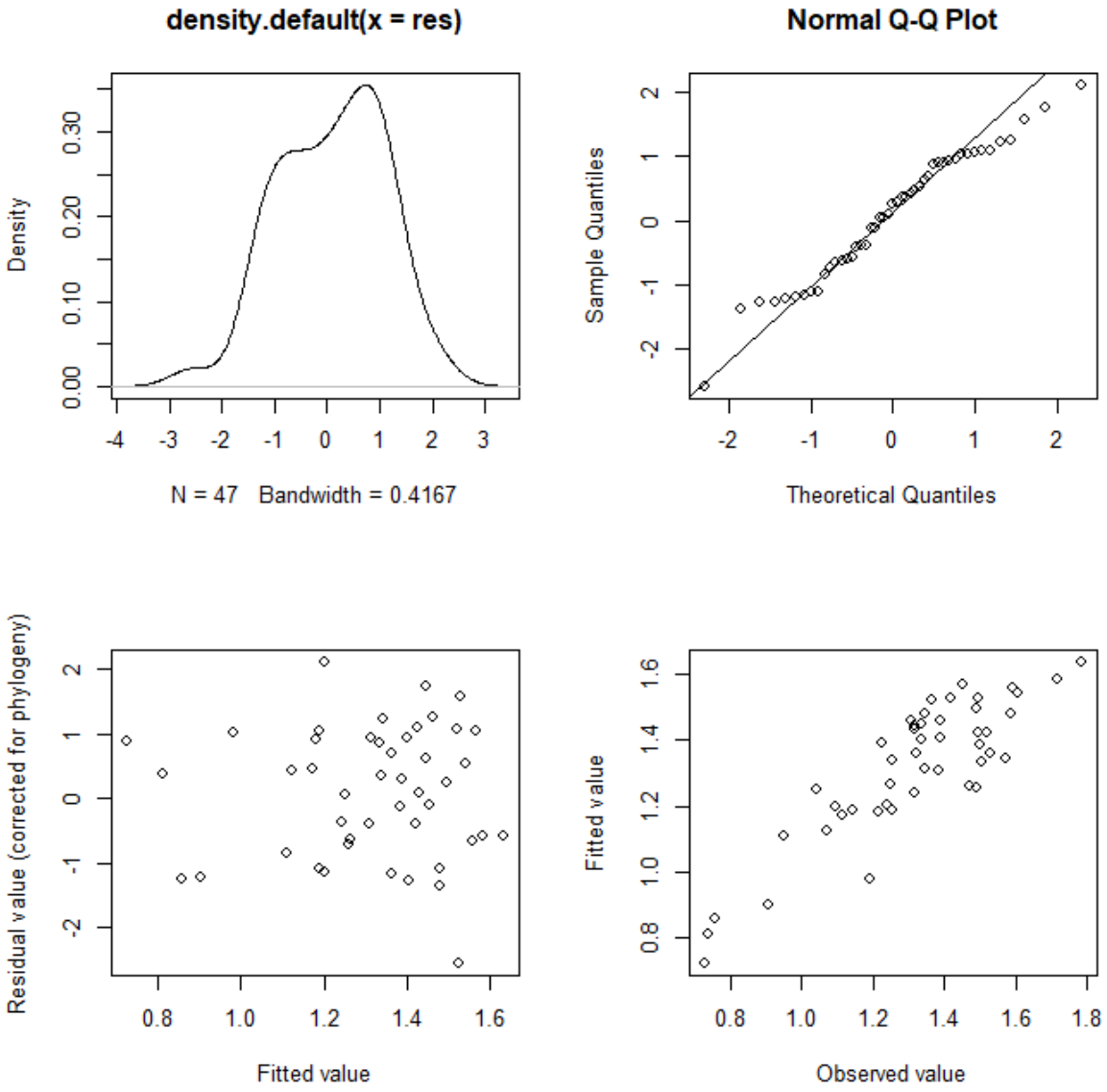


Figure S45 – Diagnostic plots for model 47.

18. PGLS ANOVA for comparing the (size-corrected weighted average) Thermo-Motility Index of endotherms *versus* ectotherms

Model 48 is very significant ($P << 0.001$), with a very large effect size (Glass's $\Delta = 1.77$), but a low coefficient of determination ($R^2 = 0.06$). It was tested on specimens for which we know the body temperature, and was used to check if the (size-corrected weighted average) Thermo-Motility Index (TMI, see Supplementary Methods, Biomechanics) of endotherms (RT = endotherm) is significantly higher than that of ectotherms (RT = ectotherm), as we theoretically posited. Results confirm that the (size-corrected weighted average) Thermo-Motility Index of endotherms is indeed significantly higher than that of ectotherms.

Model 48: $\log_{10}(TMI) \sim RT$

Branch length transformations:

Pagel's λ [ML]: 0.770

lower bound: 0.000, $P = < 2.22\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.22\ 10^{-16}$ *

95.0% CI: (0.615, 0.877)

P values approximated using a likelihood ratio test against the χ^2 distribution

Response: TMI

	Df	Sum Sq	Mean Sq	*F* statistic	Pr(>|t|)
RT	1	0.004177	0.0041761	18.318	$2.605\ 10^{-5}$*
Residuals	268	0.061098	0.0002280		

P values calculated using the exact method.

Adjusted R^2: 0.06049, AICc: -57.45117

n = 270 (3DB+3DM+2DM sets where only specimens with known thermoregulatory regime were included) Potential outliers (*Babyrousa babyrussa*, $G = 3.86279$, $U = 0.94574$, $P = 0.01261$ *, *Basiliscus basiliscus barbouri*, $G = 4.30570$, $U = 0.93234$, $P = 0.001658$ *, *Gambelia wislizenii*, $G = 4.04807$, $U = 0.93998$, $P = 0.005509$ *, *Morunasaurus annularis*, $G = 4.21144$, $U = 0.93479$, $P = 0.002576$ *, *Leiocephalus barahonensis*, $G = 4.15954$, $U = 0.93616$, $P = 0.00327$ *, *Liolaemus chilensis*, $G = 4.14800$, $U = 0.93628$, $P = 0.003434$ *, *Carollia perspicillata*, $G = 3.77358$, $U = 0.94706$, $P = 0.01798$ *, *P* values computed via approximation method) were excluded from this regression.

Effect size calculated using the Glass's $\Delta = 1.77$, 95% CI = [1.48, 2.06], because variances between the factors are still affected by phylogeny and are not similar according to the Bartlett's K-squared test = 57.012, df = 1, $P = 4.331\ 10^{-14}$ * (P value computed via exact method).

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.99641$, $P = 0.8013$, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S46) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.96164$, $U = 0.96727$, $P = 0.3847$, P value computed via approximation method).

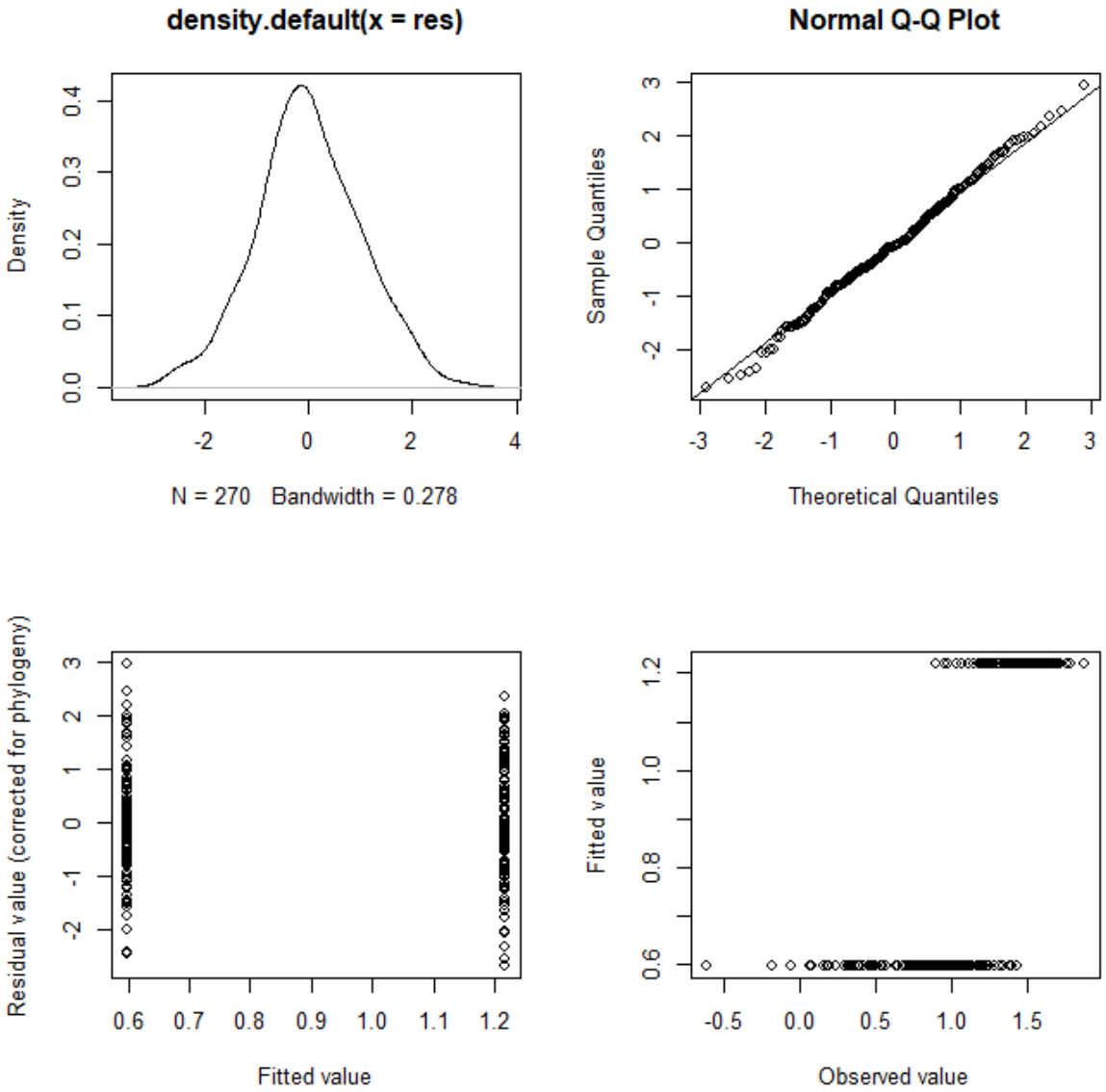


Figure S46 – Diagnostic plots for model 48.

19. PGLS-R of morphological parameters against body size or semicircular canal size

Models 49, 51, and 52 provide very significant regressions (P << 0.001), with very high coefficients of determination (R^2 = 0.72-0.94). They were tested on specimens for which we have the bony labyrinth, and were used to size-correct radii of curvature (R_a_B), and cross-sectional radii (rS_a_B) and lengths (LS_a_B) of slender portions of anterior semicircular canals. In practice, this was done by combining fitted radii of curvature obtained for a body mass (BM) of 1g, and fitted cross-sectional radii and lengths of slender portions obtained for a radius of curvature of 1mm, with residuals obtained for each species, for each model. Size-corrected radii of curvature (R_a_B_adjusted), and cross-sectional radii (rS_a_B_adjusted) and lengths (LS_a_B_adjusted) of slender portions are used in Models 53-56, along with non-corrected measurements of torus eccentricity (e_a_B: see Model 50), to assess which morphological parameters of the anterior semicircular canals best reflect the (size-corrected weighted average) Thermo-Motility Index (TMI: see Supplementary Methods, Biomechanics). Size-corrected morphological parameters are illustrated in Extended Data Figure 3.

Model 49: $\log_{10}(R\_a\_B) \sim \log_{10}(\sqrt[3]{BM})$

Branch length transformations:

Pagel's λ [ML]: 0.812

lower bound: 0.000, $P = < 2.22\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.22\ 10^{-16}$ *

95.0% CI: (0.715, 0.884)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.190636	0.055064	-3.4621	0.0006165 *
$\log_{10}(\sqrt[3]{BM})$	0.515617	0.018631	27.6756	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.7237, AICc: -483.4633

F statistic: 765.7 on 1 and 291 DF

n = 293 (3DB set)

Potential outliers (*Didelphis virginiana*, *G* = 3.96152, *U* = 0.94698, *P* = 0.008963 *, *Phascolarctos cinereus*, *G* = 3.92215, *U* = 0.94768, *P* = 0.01057 *, *Dasyurus viverrinus*, *G* = 3.82315, *U* = 0.94994, *P* = 0.01606, *P* values computed via approximation method) were excluded from this regression. As Shapiro-

Wilk tests (W = 0.99128, P = 0.07609, W = 0.9902, P = 0.04528 *, P value calculated via approximation method) indicated potential non-normality of the phylogenetic residuals after removing outliers, the most influential specimens (*Varanus salvator*, *Varanus giganteus*) were excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.99334, P = 0.2201, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S47) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.88777, U = 0.97134, P = 0.5358, P value computed via approximation method).

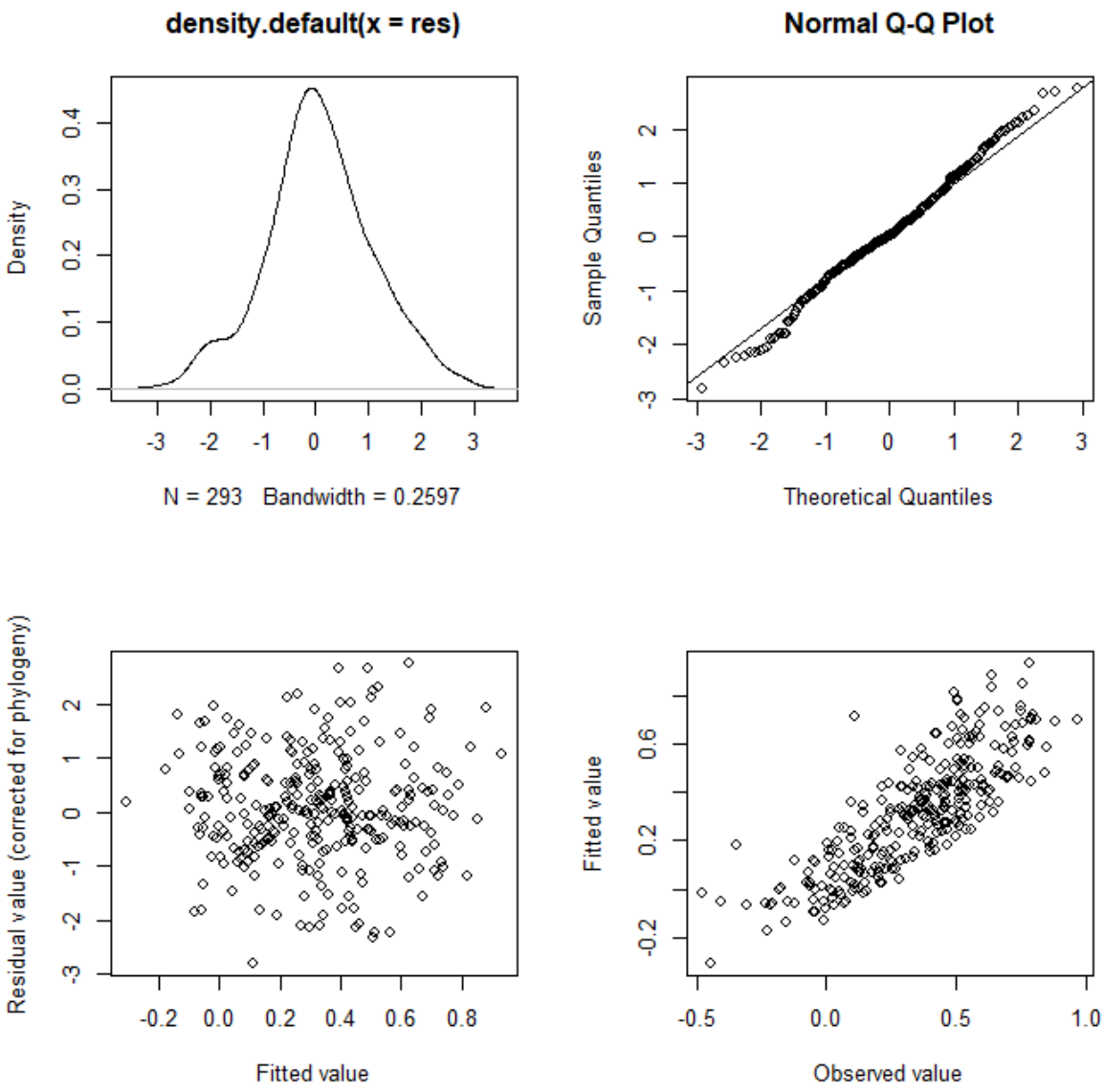


Figure S47 – Diagnostic plots for model 49.

Model 50: logit(e_a_B) ~ $\log_{10}$(R_a_B)

Adjusted R^2: -0.001316, AICc: 586.5895

F statistic: 0.6096 on 1 and 296 DF

n = 298 (3DB set)

P value: 0.4356, Corrected *P* value: 0.4356

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Model 51: $\log_{10}(\text{LS\_a\_B}) \sim \log_{10}(\sqrt[3]{\text{BM}})$

Branch length transformations:

Pagel's λ [ML]: 0.791

lower bound: 0.000, $P = < 2.22\ 10^{-16}$ *

upper bound: 1.000, $P = 2.259\ 10^{-07}$ *

95.0% CI: (0.604, 0.917)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.555582	0.023481	23.660	$< 2.2\ 10^{-16}$ *
$\log_{10}$(R_a_B)	1.011606	0.016849	60.039	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.9442, AICc: -680.8537

F statistic: 3605 on 1 and 212 DF

n = 214 (3DB set)

Potential outliers (*Dicynodontoides cf. recurvidens*, *G* = 4.0968, *U* = 0.9433, *P* = 0.004884 *, *Kawingasaurus fossilis*, *G* = 3.8289, *U* = 0.9503, *P* = 0.01586 *, *Cynosaurus suppostus*, *G* = 4.01074, *U* = 0.94529, *P* = 0.007145 *, *Pristerodon mackayi*, *G* = 4.64462, *U* = 0.92637, *P* = 0.0003317 *, *Scalenodon angustifrons*, *G* = 4.05347, *U* = 0.94373, *P* = 0.005847 *, *Endothiodon tolani*, *G* = 4.54454, *U* = 0.92903, *P* = 0.0005502 *, *Diictodon feliceps*, *G* = 4.11049, *U* = 0.94174, *P* = 0.004472 *, *Herpetoskylax hopsoni*, *G* = 4.05083, *U* = 0.94322, *P* = 0.005842 *, *Lemurosaurus pricei*, *G* = 4.00182, *U* = 0.94439, *P* = 0.007249 *, *Abajudon kaayai*, *G* = 4.44197, *U* = 0.93125, *P* = 0.0009048 *, *Niassodon mfumukasi*, *G* = 4.52295, *U* = 0.92847, *P* = 0.0005994 *, *Endothiodon bathystoma*, *G* = 4.40398, *U* = 0.93195, *P* = 0.001082 *, *Eodicynodon oosthuizeni*, *G* = 4.37082, *U* = 0.93273, *P* = 0.001267 *, *Patranomodon nyaphulii*, *G* = 4.37023, *U* = 0.93251, *P* = 0.001265 *, *Moschops capensis*, *G* = 4.47506, *U* = 0.92899, *P* = 0.0007486 *, *Leucocephalus wewersi*, *G* = 4.43298, *U* = 0.93007, *P* = 0.0009193 *, *Coracias garrulus*, *G* = 3.98989, *U* = 0.94315, *P* = 0.007384 *, *Tyto alba*, *G* = 3.9599, *U* = 0.9438, *P* = 0.0084 *, *Athene cunicularia*, *G* = 3.9702, *U* = 0.9433, *P* = 0.007988 *, *Aquila chrysaetos*, *G* = 3.98087, *U* = 0.94279, *P* = 0.007585 *, *Buteo buteo*, *G* = 3.9418, *U* = 0.9437, *P* = 0.008977 *, *Circus cyaneus*, *G* = 3.95822, *U* = 0.94303, *P* = 0.008312 *, *Pandion haliaetus*, *G* = 4.05254, *U* = 0.94006, *P* = 0.005423 *, *Chiniquodon theotonicus*, *G* = 4.30826, *U* = 0.93201, *P* = 0.00163 *, *Sagittarius serpentarius*, *G* = 3.93022, *U* = 0.94321, *P* = 0.009279 *, *Vultur*

gryphus, *G* = 3.91461, *U* = 0.94345, *P* = 0.009891 *, *Opisthocomus hoazin*, *G* = 3.88079, *U* = 0.94422, *P* = 0.01141 *, *Fregata magnificens*, *G* = 3.87996, *U* = 0.94404, *P* = 0.0114 *, *Phalacrocorax harrisi*, *G* = 4.01729, *U* = 0.93978, *P* = 0.006185 *, *Ardea cinerea*, *G* = 3.9658, *U* = 0.9411, *P* = 0.00775 *, *Phaethon lepturus*, *G* = 3.90807, *U* = 0.94258, *P* = 0.009949 *, *Ciconia nigra*, *G* = 3.92580, *U* = 0.94184, *P* = 0.009162 *, *Lumkuia fuzzi*, *G* = 4.03499, *U* = 0.93833, *P* = 0.005603 *, *Oudenodon bainii*, *G* = 4.14195, *U* = 0.93477, *P* = 0.003414 *, *Corvus corax*, *G* = 4.96476, *U* = 0.90592, *P* = 4.903 10^{-05} *, *Dicrurus paradiseus*, *G* = 4.93653, *U* = 0.90663, *P* = 5.715 10^{-05} *, *Ara macao*, *G* = 4.26384, *U* = 0.93008, *P* = 0.001891 *, *Aulacephalodon sp.*, *G* = 4.68013, *U* = 0.91543, *P* = 0.0002296 *, *Psittacus erithacus*, *G* = 4.8000, *U* = 0.9107, *P* = 0.0001197 *, *Melopsittacus undulatus*, *G* = 4.49945, *U* = 0.92123, *P* = 0.0005807 *, *Falco subbuteo*, *G* = 4.82739, *U* = 0.90897, *P* = 0.0001017 *, *Pteroglossus bailloni*, *G* = 4.74201, *U* = 0.91182, *P* = 0.000161 *, *Ramphastos dicolorus*, *G* = 4.74879, *U* = 0.91122, *P* = 0.0001543 *, *Dendrocopos medius*, *G* = 4.7504, *U* = 0.9108, *P* = 0.0001519 *, *Scylacocephalus sp.*, *G* = 4.16409, *U* = 0.93119, *P* = 0.002912 *, BP/1/155, *G* = 4.83220, *U* = 0.90697, *P* = 9.598 10^{-05} *, *Alcedo atthis*, *G* = 4.69410, *U* = 0.91186, *P* = 0.0002019 *, *Ciconia ciconia*, *G* = 4.7288, *U* = 0.9102, *P* = 0.0001666 *, *Balaeniceps rex*, *G* = 4.78394, *U* = 0.90772, *P* = 0.0001227 *, *Fulmarus glacialis*, *G* = 4.68533, *U* = 0.91113, *P* = 0.0002077 *, *Diomedea exulans*, *G* = 4.6871, *U* = 0.9107, *P* = 0.0002045 *, *Pelagodroma marina*, *G* = 4.72291, *U* = 0.90896, *P* = 0.0001677 *, *Spheniscus demersus*, *G* = 4.72022, *U* = 0.90869, *P* = 0.000169 *, *Gavia immer*, *G* = 4.7205, *U* = 0.9083, *P* = 0.0001677 *, *Larus argentatus*, *G* = 4.71608, *U* = 0.90809, *P* = 0.0001706 *, *Creagrus furcatus*, *G* = 4.66988, *U* = 0.90951, *P* = 0.0002172 *, *Gelochelidon nilotica*, *G* = 4.6477, *U* = 0.9100, *P* = 0.0002428 *, *Alca torda*, *G* = 4.62091, *U* = 0.91066, *P* = 0.000278 *, GPIT/RE/7124, *G* = 4.34641, *U* = 0.92063, *P* = 0.001121 *, *Apus apus*, *G* = 3.80821, *U* = 0.93881, *P* = 0.01331 *, *Scolopax rusticola*, *G* = 4.4125, *U* = 0.9175, *P* = 0.0007972 *, *Actitis hypoleucos*, *G* = 4.49846, *U* = 0.91389, *P* = 0.0005123 *, *Rana pipiens*, *G* = 5.11207, *U* = 0.88832, *P* = 1.713 10^{-05} *, *Dimetrodon sp.*, *G* = 4.00065, *U* = 0.93131, *P* = 0.00558 *, *Lissotriton vulgaris*, *G* = 4.82418, *U* = 0.89969, *P* = 8.823 10^{-05} *, *Grus grus*, *G* = 4.22685, *U* = 0.92266, *P* = 0.001923 *, *Cuculus canorus*, *G* = 4.2112, *U* = 0.9229, *P* = 0.002061 *, *Kembawacela kitchingi*, *G* = 4.19275, *U* = 0.92324, *P* = 0.002237 *, *Sangusaurus parringtonii*, *G* = 4.81714, *U* = 0.89823, *P* = 8.914 10^{-05} *, *Proteus anguinus*, *G* = 4.48810, *U* = 0.91127, *P* = 0.0005131 *, *Choerosaurus dejageri*, *G* = 4.20886, *U* = 0.92162, *P* = 0.002034 *, *Crocodylus moreletii*, *G* = 4.0805, *U* = 0.9260, *P* = 0.003699 *, *Xenopus laevis*, *G* = 4.91305, *U* = 0.89224, *P* = 5.042 10^{-05} *, *Pseudotherium argentinus*, *G* = 3.67316, *U* = 0.93895, *P* = 0.02161 *, *Riograndia guaibensis*, *G* = 3.66896, *U* = 0.93881, *P* = 0.02187 *, *Luangwa drysdalli*, *G* = 3.62704, *U* = 0.93993, *P* = 0.02586 *, *Pelodytes caucasicus*, *G* = 3.38364, *U* = 0.94724, *P* = 0.06717, *Hipposaurus boonstrai*, *G* = 3.89558, *U* = 0.92974, *P* = 0.008097 *, *Lystrosaurus declivis*, *G* = 3.56097, *U* = 0.94102, *P* = 0.03309 *, *Eothyris parkeri*, *G* = 3.73931, *U* = 0.93466, *P* = 0.01574 *, *Lystrosaurus murrayi*, *G* = 3.5811, *U* = 0.9398, *P* =

0.03017 *, *P* values computed via approximation method) were excluded from this regression. As Shapiro-Wilk tests (W = 0.98168, P = 0.005162 *, W = 0.97631, P = 0.0008157 *, W = 0.98894, P = 0.08838, *P* value calculated via approximation method) indicated potential non-normality of the phylogenetic residuals after removing outliers, the most influential specimens (*Thrinaxodon liorhinus*, *Tritylodon longaevus*, *Ictidosuchoides* sp.) were excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *, Corrected *P* value: $3.3\ 10^{-16}$

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.99241, P = 0.3361, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S48) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 3.21278, U = 0.95131, P = 0.1239, *P* value computed via approximation method).

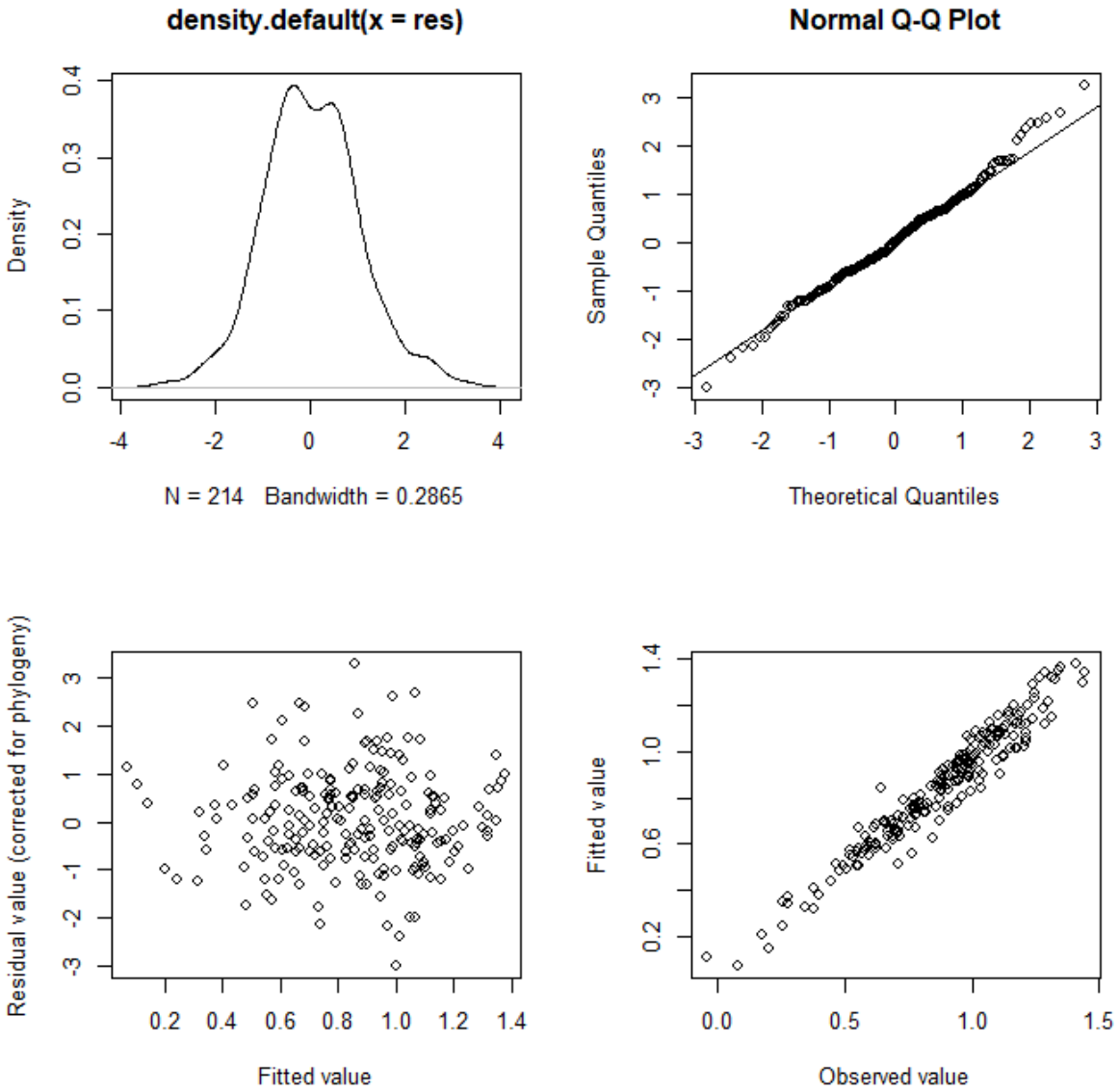


Figure S48 – Diagnostic plots for model 51.

Model 52: $\log_{10}(rS\_a\_B) \sim \log_{10}(R\_a\_B)$

Branch length transformations:

Pagel's λ [ML]: 0.740

lower bound: 0.000, $P = < 2.22\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.22\ 10^{-16}$ *

95.0% CI: (0.594, 0.846)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.765618	0.045063	-16.99	$< 2.2\ 10^{-16}$ *
$\log_{10}$(R_a_B)	0.798263	0.029369	27.18	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.7151, AICc: -505.3266

F statistic: 738.8 on 1 and 293 DF

n = 295 (3DB set)

A potential outlier (*Salvator merianae*, *G* = 3.49603, *U* = 0.95871, *P* = 0.06193, *P* values computed via approximation method) was excluded from this regression. As Shapiro-Wilk tests (*W* = 0.99143, *P* = 0.0822, *W* = 0.98947, *P* = 0.03099 *, *P* value calculated via approximation method) indicated potential non-normality of the phylogenetic residuals after removing outliers, the most influential specimens (*Tupaia glis*, *Tupinambis teguixin*) were excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *, Corrected *P* value: $3.3\ 10^{-16}$

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.99232, *P* = 0.1312, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S49) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 3.3367, *U* = 0.9620, *P* = 0.1124, *P* value computed via approximation method).

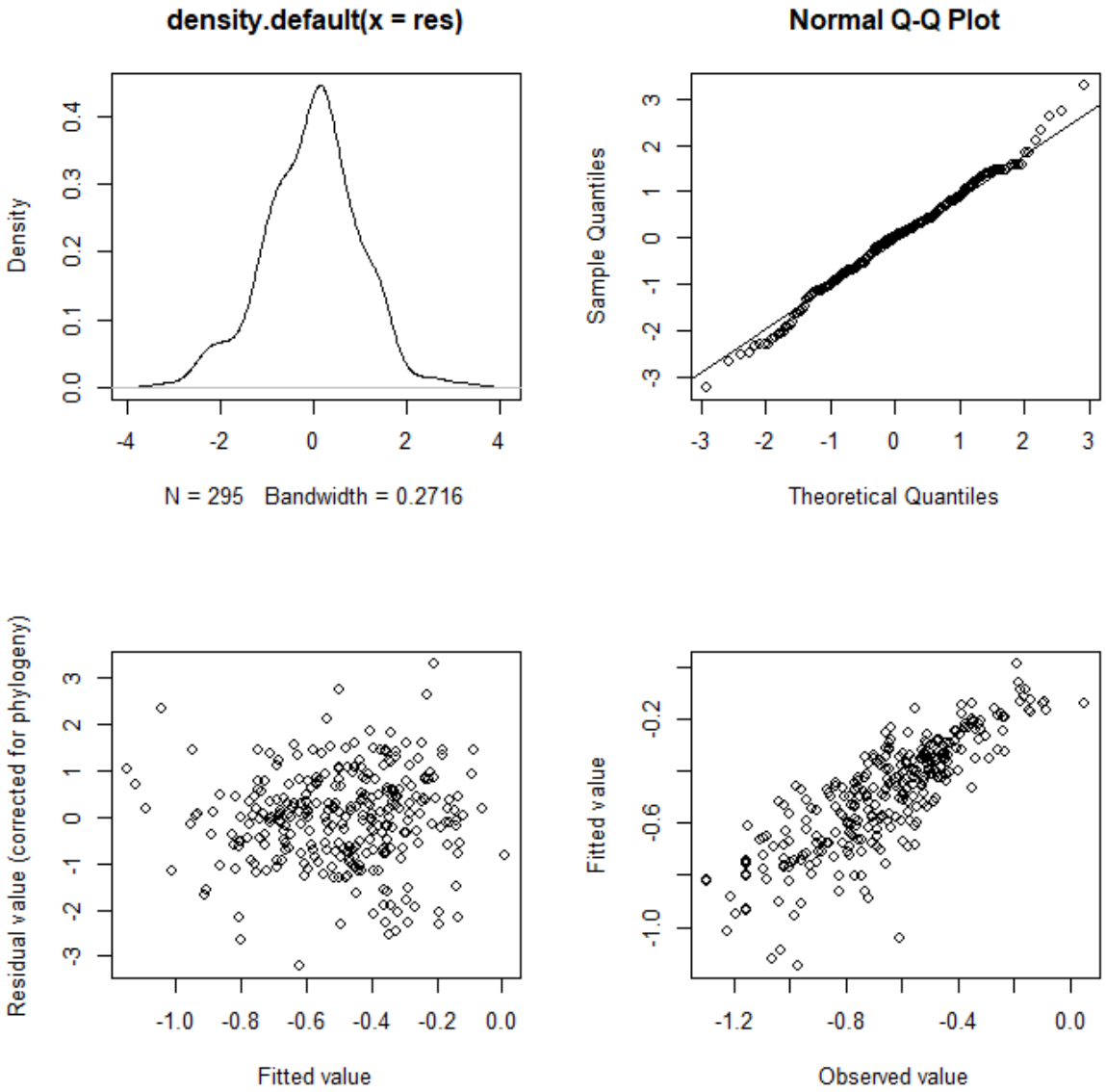


Figure S49 – Diagnostic plots for model 52.

20. PGLS-R of the (size-corrected weighted average) Thermo-Motility Index against size-adjusted morphological parameters of the semicircular canals

Models 53 and 56 provide very significant regressions (P << 0.001), with coefficients of determination (R^2) ranging between 0.08–0.63. They were tested on specimens for which we have the bony labyrinth, and were used to assess which morphological parameters of the anterior semicircular canals best reflect the (size-corrected weighted average) Thermo-Motility Index (TMI, see Supplementary Methods, Biomechanics). Results suggest that the size-corrected radius of curvature (R_a_B_adjusted) and the size-corrected cross-sectional radius of the slender portion (rS_a_B_adjusted) are both negatively correlated to the (size-corrected weighted average) Thermo-Motility Index, but that the cross-sectional radius of the slender portion of the anterior semicircular canal is the best predictor of the (size-corrected weighted average) Thermo-Motility Index. This finding means that an anterior semicircular canal with cross-sections of the slender portion appearing thinner than expected for the size of the radius of curvature, as in most mammals, are generally indicative of an elevated (size-corrected weighted average) Thermo-Motility Index.

Model 53: $\log_{10}(\text{TMI}) \sim \log_{10}(\text{R\_a\_B\_adjusted})$

Branch length transformations:

Pagel's λ [ML]: 0.705

lower bound: 0.000, $P = < 2.22\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.22\ 10^{-16}$ *

95.0% CI: (0.567, 0.815)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.778163	0.077695	10.016	$< 2.2\ 10^{-16}$ *
$\log_{10}$(R_a_B_adjusted)	-0.516719	0.101238	-5.104	$5.958\ 10^{-07}$

P values calculated using the exact method.

Adjusted R^2: 0.07779, AICc: -157.1054

F statistic: 26.05 on 1 and 296 DF

n = 298 (3DB set)

P value: $5.958\ 10^{-07}$ *, Corrected *P* value: $1.1916\ 10^{-06}$

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.9962, P = 0.6956, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S50) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.86548, U = 0.97226, P = 0.5865, P value computed via approximation method).

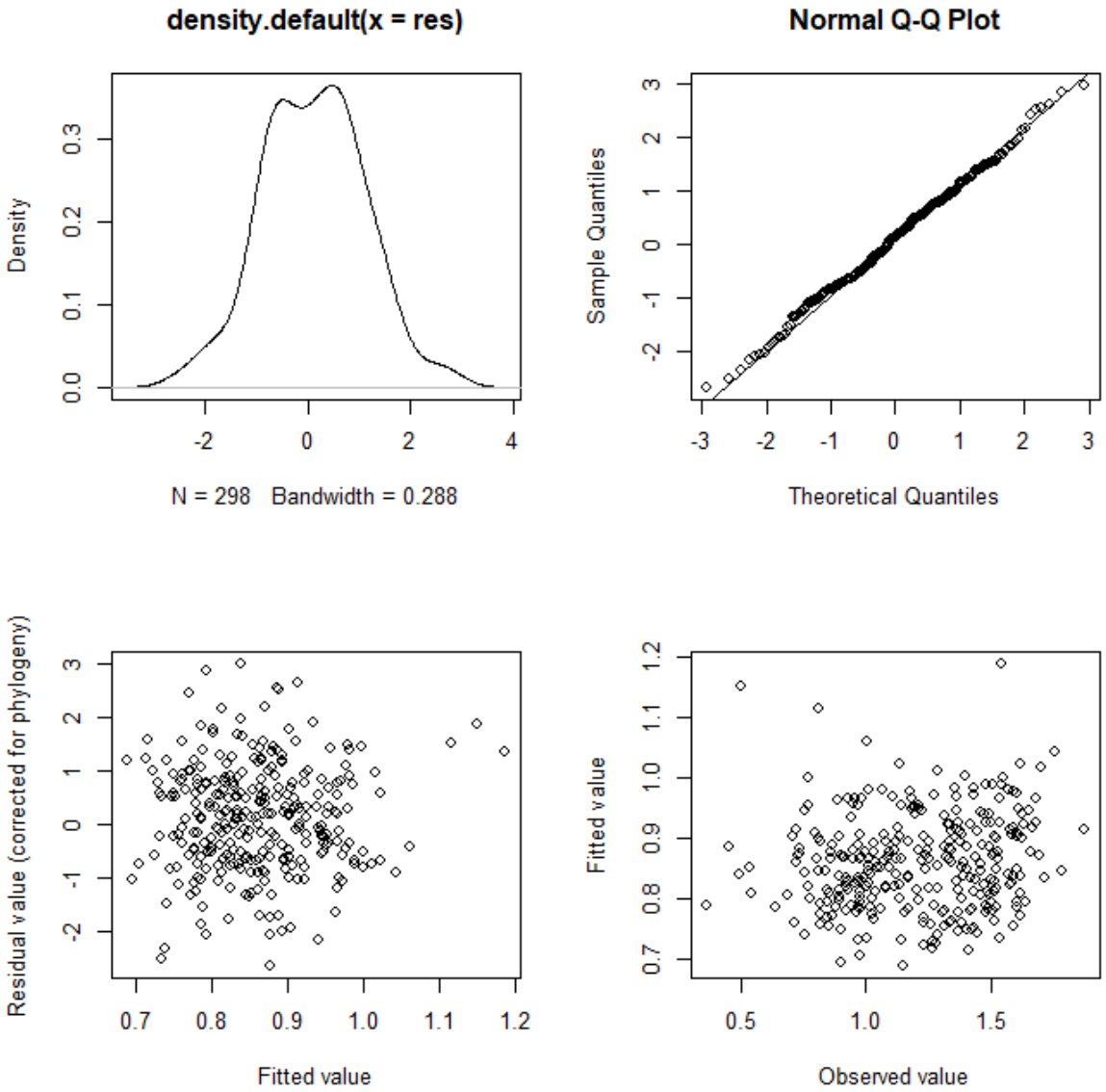


Figure S50 – Diagnostic plots for model 53.

Model 54: $\log_{10}$(TMI) ~ logit(e_a_B)

Adjusted R^2: -0.003092, AICc: -132.096

F statistic: 0.08457 on 1 and 296 DF

n = 298 (3DB set)

P value: 0.7714, Corrected *P* value: 0.9818

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Model 55: $\log_{10}$(TMI) ~ $\log_{10}$(LS_a_B_adjusted)

Adjusted R^2: -0.003377, AICc: -132.0139

F statistic: 0.0005224 on 1 and 296 DF

n = 298 (3DB set)

P value: 0.9818, Corrected *P* value: 0.9818

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Model 56: $\log_{10}$(TMI) ~ $\log_{10}$(rS_a_B_adjusted)

Branch length transformations:

Pagel's λ [ML]: 0.795

lower bound: 0.000, $P = < 2.22\ 10^{-16}$ *

upper bound: 1.000, $P = < 2.22\ 10^{-16}$ *

95.0% CI: (0.691, 0.872)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.307508	0.078202	-3.9322	0.0001055 *
$\log_{10}$(rS_a_B_adjusted)	-1.522221	0.068004	-22.3842	$< 2.2\ 10^{-16}$ *

P values calculated using the exact method.

Adjusted R^2: 0.6337, AICc: -414.1055

F statistic: 501.1 on 1 and 288 DF

n = 290 (3DB set)

Potential outliers (*Galesaurus planiceps*, *G* = 3.66871, *U* = 0.95453, *P* = 0.03109 *, *Caecilia tentaculata*, *G* = 4.47840, *U* = 0.93155, *P* = 0.0007759 *, *Scylacocephalus sp.*, *G* = 3.78404, *U* = 0.95096, *P* = 0.01896

*, *Dixeya nasuta*, G = 4.16353, U = 0.94022, P = 0.003501 *, *Eleutherodactylus limbatus*, G = 4.75985, U = 0.92161, P = 0.0001768 *, P values computed via approximation method) were excluded from this regression. As Shapiro-Wilk tests (W = 0.98571, P = 0.00479 *, W = 0.99005, P = 0.04134 *, W = 0.98277, P = 0.001349 *, P value calculated via approximation method) indicated potential non-normality of the phylogenetic residuals after removing outliers, the most influential specimens (*Pelodytes caucasicus*, *Eothyris parkeri*, BP/1/155) were excluded from this regression.

P value: $< 2.2\ 10^{-16}$ *, Corrected P value: $8.8\ 10^{-16}$

P values have been calculated using the exact method and corrected for the False Discovery Rate.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.99434, P = 0.3559, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S51) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 3.33316, U = 0.96142, P = 0.1118, P value computed via approximation method).

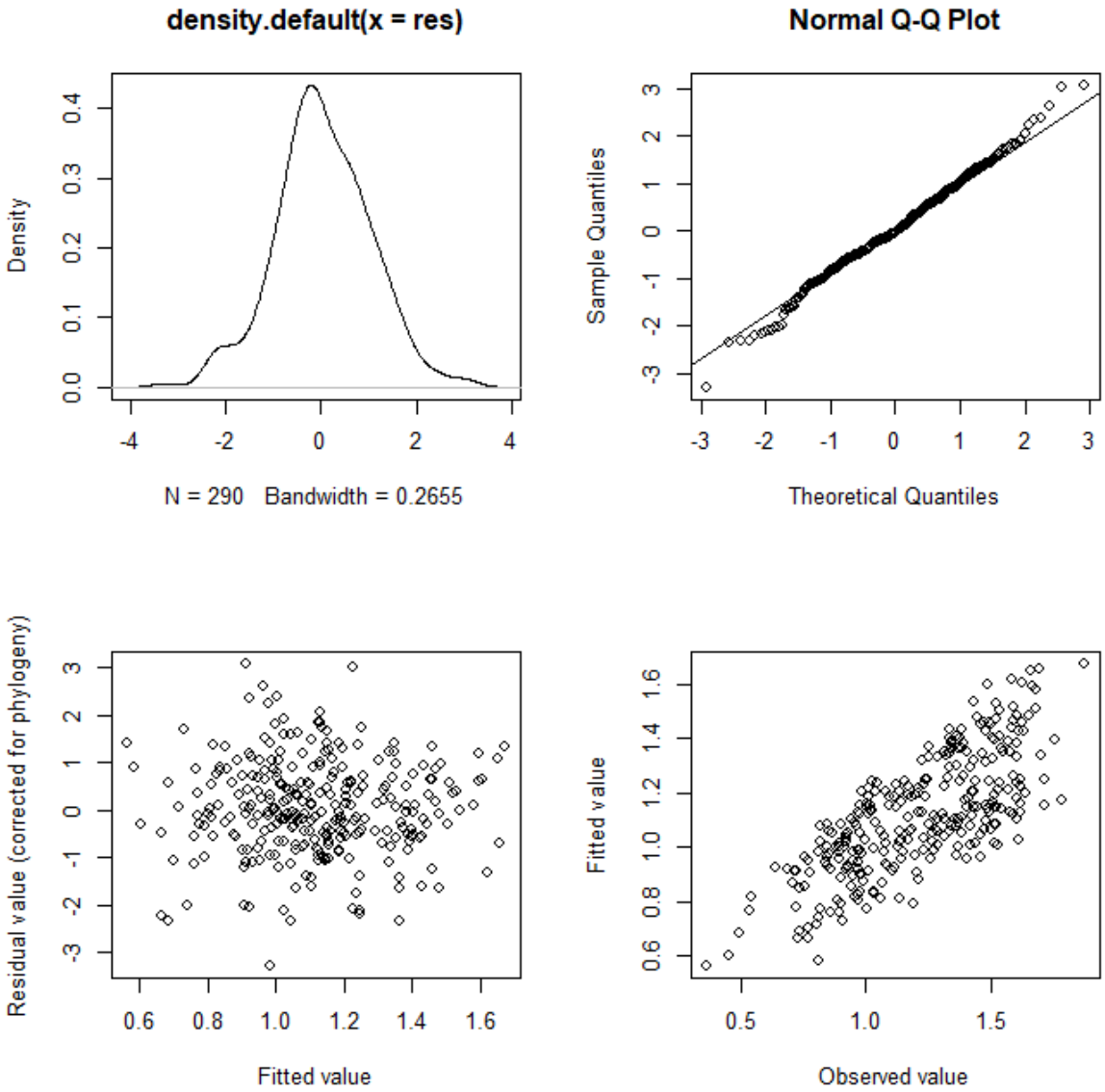


Figure S51 – Diagnostic plots for model 56.

21. Phylogenetic logistic regression of the endothermic status against the (size-corrected weighted average) Thermo-Motility Index

Model 57 provides a significant logistic regression (P <0.05) with a very low cross-validated error rate of 5%, when regarding probabilities P > 0.70 for species and P > 0.53 for clades as indicating endothermy, P < 0.30 for species and P < 0.47 for clades as implying ectothermy, and intermediate probabilities as expressing uncertainty. Model 57 was tested on amniote specimens for which we know the thermoregulatory regime, and was used to calculate probabilities that fossil species, and groups, have thermoregulatory regimes comparable to that of extant endotherms. Computed probabilities are provided in Extended Data Tables 1 and 3 and were used, along with Model 66, to compute the ancestral state estimations provided in Extended Data Table 2 and illustrated in Figure 3 and Extended Data Figure 6.

Model 57: RT ~ $\log_{10}$(TMI)

Random effect (phylogenetic signal σ^2):

σ^2	Pr
9.709	$5.697\ 10^{-09}$ *

P value calculated using an approximate likelihood test based on the conditional restricted maximum likelihood.

Coefficients:

	Value	Std. Error	Z score	*P* value
(Intercept)	-8.0243	4.2004	-1.9104	0.05608
$\log_{10}$(TMI)	6.8740	3.3382	2.0592	0.03948 *

P values calculated using approximate penalized quasi-likelihood tests.

n = 227 (3DB set where only amniote specimens with known thermoregulatory regime were included).

Effect size for species data was determined via leave-p-out cross-validation, with p = 76 specimens randomly sampled and 100 iterations. Specimens with probabilities above P = 0.70 or below P = 0.30 were predicted as endothermic and ectothermic, respectively. In total, 74% of specimens were assigned the correct thermoregulatory regime, whereas 5% were assigned the wrong one. The thermoregulatory regime of 20% of specimens was labelled as uncertain.

Effect size for clade data was determined via leave-p-out cross-validation, with p = all specimens from six randomly sampled clades and 100 iterations. Clades with probabilities above P = 0.53 or below P = 0.47

were predicted as endothermic and ectothermic, respectively. In total, 88% of clades were assigned the correct thermoregulatory regime, whereas 4% were assigned the wrong one. The thermoregulatory regime of 9% of clades was labelled as uncertain.

Assumption testing:

- The outcome of the logistic regression is a binary variable (Endothermic regime: YES or NO)
- Observations are independent from each other thanks to the phylogenetic comparative method.
- The logit of the outcome is linearly associated with the predictor variable (Fig. S52), as expected.
- We looked for outliers using Grubb's test and found none (G = 1.46342, U = 0.99048, P = 1, P value computed via approximation method).

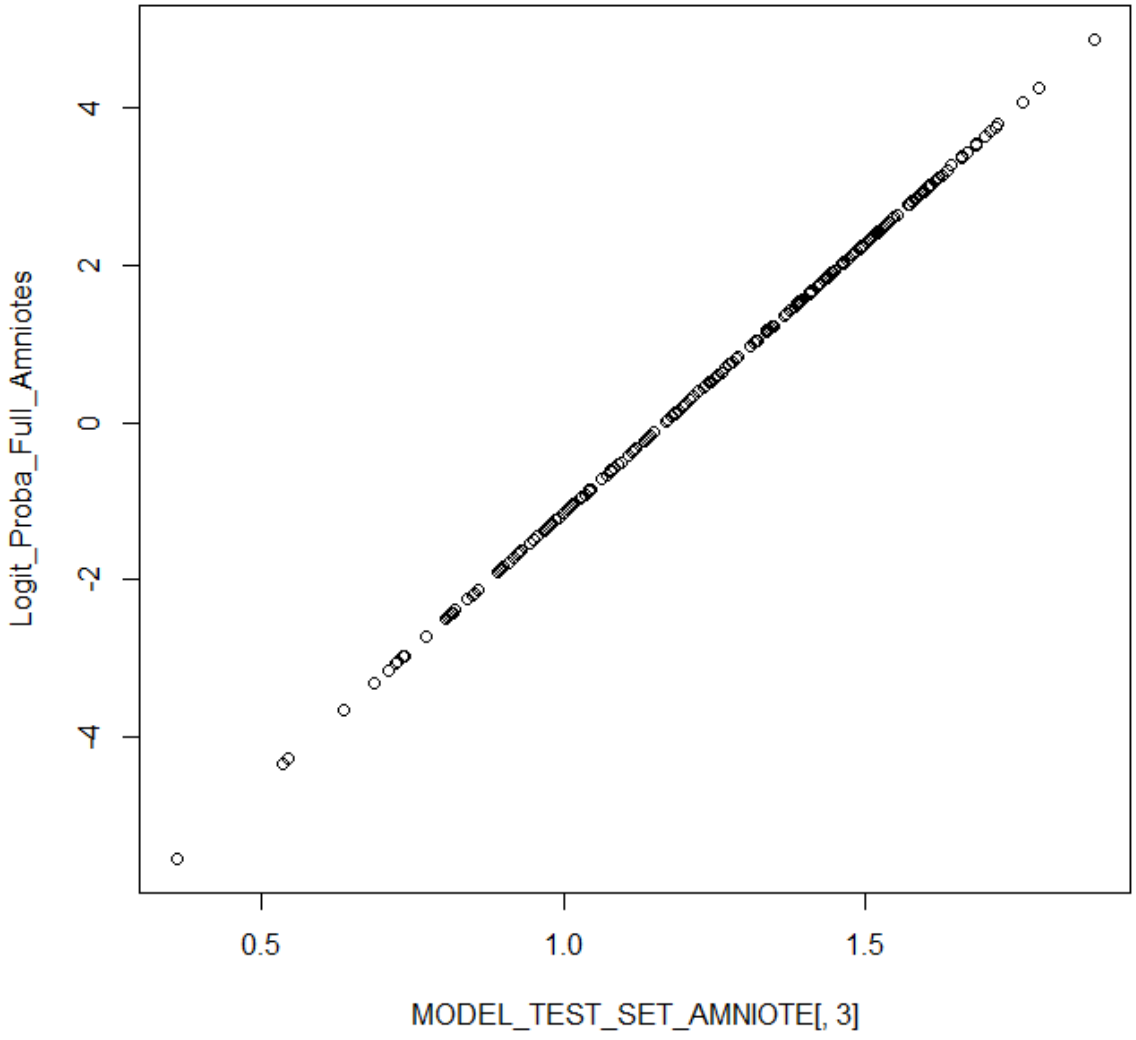


Figure S52 – Diagnostic plot for model 57.

22. Determination of the model best explaining the evolution of the (size-corrected weighted average) Thermo-Motility Index of fossil synapsids

Models 58-64 correspond to various evolutionary models that were fitted on the (size-corrected weighted average) Thermo-Motility Index of amniotes, using average values for Theria, Monotremata, Aves, Crocodylia, Lepidosauria, and Testudines, and a phylogenetic tree focusing on fossils, with *Diadectes* as the outgroup. We used the corrected version of the Akaike information criterion (AICc) to select which model best explained the evolution of the Thermo-Motility Index (see below).

Model	AICc	Relative likelihood
58: Brownian motion	-19.3	0.00
59: Ornstein–Uhlenbeck process	-27.7	0.00
60: Early burst	-17.2	0.00
61: Continuous trend	-17.1	0.00
62: Pagel's λ	-43.8	1.00
63: Pagel's δ	-19.6	0.01
64: Pagel's κ	-34.6	0.00

We found that Brownian motion along the proposed tree, with branch lengths scaled using a Pagel's λ of 0.50, was the model best explaining the evolution of the (size-corrected weighted average) Thermo-Motility Index of fossil synapsids (Model 62). This result suggests that the variation observed in the (size-corrected weighted average) Thermo-Motility Index is less explained by the temporal structure of the tree than what would be expected according to a pure Brownian motion process. Using Model 62, we run further analyses to see if either (1) a change in the rate of evolution of the (size-corrected weighted average) Thermo-Motility Index after the Permo-Triassic boundary (Model 65), or (2) shifts in the rate of evolution of the (size-corrected weighted average) Thermo-Motility Index (Model 66), better explained its evolution than Model 62 (see below).

Model	AICc	Relative likelihood
62: Pagel's λ	-46.0	0.00
65: Pagel's λ with change of rate at the Permo-Triassic boundary	-42.0	0.00
66: Pagel's λ with a major shift along the branch directly leading to Mammaliamorpha	-66.8	1.00

The evolutionary model assuming that the Permo-Triassic crisis had an effect onto the rate of evolution of the (size-corrected weighted average) Thermo-Motility Index (Model 65) was less supported than Model 62. Concerning shifts in the rate of evolution of the (size-corrected weighted average) Thermo-Motility Index, all potential locations, and up to five shifts were tested. Results suggest that Model 66, which corresponds to a major shift in the rate of evolution along the branch directly leading to Mammaliamorpha, best explains the evolution of the (size-corrected weighted average) Thermo-Motility Index. The tree reflecting the best evolutionary model (Fig. S53) was used to estimate ancestral states of the (size-corrected weighted average) Thermo-Motility Index, and probabilities of having a thermoregulatory regime comparable to those of extant endotherms. Results are illustrated in Figures 3, 4, Extended Data Figure 6 and Extended Data Tables 2. More details are provided below for each evolutionary model we tested.

Model 58: Brownian motion

Brownian variance: 0.00236667
Root state: 0.9387194
AICc: -19.26999

Model 59: Ornstein–Uhlenbeck process

Alpha [ML]: 0.05590538, 95% CI = [0.02440807, 0.1268502]
Brownian variance: 0.005399526
Root state: 1.135667
AICc: -27.68446

Model 60: Early burst

ACDC [ML]: -0.002175848, 95% CI = [-0.01355329, 0.009106489]
Brownian variance: 0.002834733
Root state: 0.938343
AICc: -17.22236

Model 61: Continuous trend

Slope [ML]: $-6.153115\ 10^{-05}$, Standard Error = 0.0007338026
Brownian variance: 0.002366582
Root state: 0.9423194, Standard Error = 0.1790141
AICc: -17.07657

Model 62: Pagel's λ

Pagel's λ [ML]: 0.5032402, 95% CI = [0.160908, 0.8148193]
Brownian variance: 0.0004623377
Root state: 0.9461507
AICc: -43.76172

Model 63: Pagel's δ

Pagel's δ [ML]: 1.852006, 95% CI = [0.807271, 3.021368]
Brownian variance: $3.132764\ 10^{-05}$
Root state: 0.9468019
AICc: -19.55448

Model 64: Pagel's κ

Pagel's κ [ML]: 0.3267571, 95% CI = [0.05934661, 0.6220893]
Brownian variance: 0.007261385
Root state: 0.9253141
AICc: -34.56887

Model 65: Pagel's λ with change of rate at the Permo-Triassic boundary

Brownian variance: $5.097349\ 10^{-04}$
Root state: $9.446408\ 10^{-01}$
Time boundary: 251.9
Rate before boundary: $8.071330\ 10^{-01}$
Rate after boundary: 1.272500
AICc: $-4.202544\ 10^{1}$

Model 66: Pagel's λ with a major shift along the branch directly leading to Mammaliamorpha

We tested from 0 to 5 shifts in the rate of evolution, with no *a priori* assumptions about where the shifts were expected to occur. The best model we retrieved support one major shift along the branch going toward Mammaliamorpha and a minor shift in *Diadectes*. The latter minor shift seems artefactual because it always occurred at the level of the outgroup, irrespective of the phylogenetic configuration we tested.

Number of shifts: 2
Position of first shift: *Diadectes*

Rate of the first shift: 1 10^{-08}, 95% CI = [NA, 1 10^{-08}]

Position of the second shift: along the branch directly leading to Mammaliamorpha

Rate of the second shift: 511.144082164479, 95% CI = [42.0328986152098, NA]

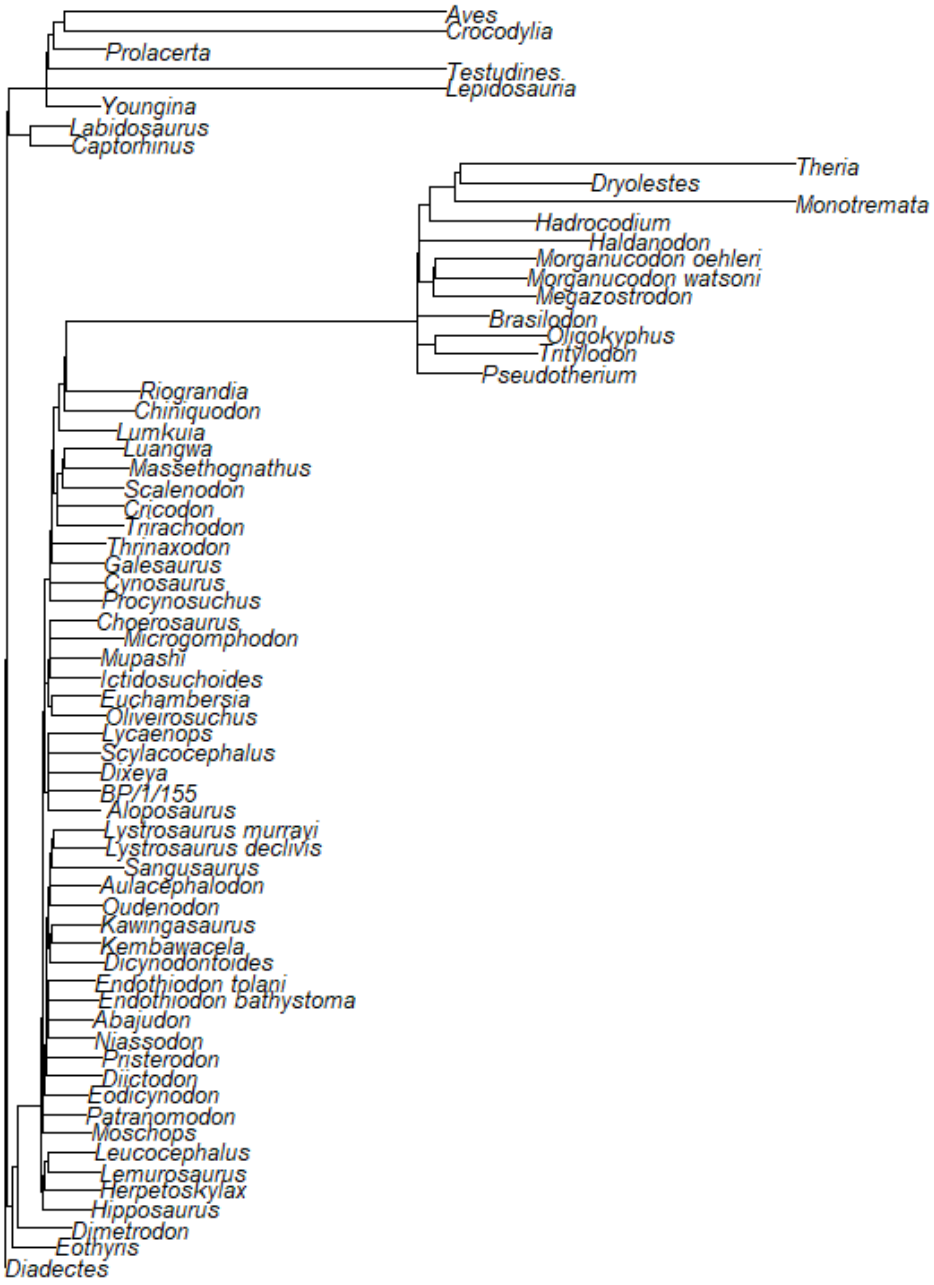


Figure S53 – Tree reflecting evolutionary model 66.

23. Determination of phylogenetic clusters best explaining the distribution of the (size-corrected weighted average) Thermo-Motility Index of fossil synapsids

We used phylogenetic ANOVAs to find clusters (grades and clades) best summarising the distribution of the (size-corrected weighted average) Thermo-Motility Index of fossil synapsids along the mammalian lineage. In practice we tested all possible phylogenetic clusters (i.e., giving the same group label to all members of a clade or grade), excluding clusters branching out the mammalian lineage, and considering 1–5 clusters. For each number of clusters, we provide the corrected version of the Akaike Information Criterion and the description of the best model that was obtained (see below). Model 67, which best summarises the distribution of the (size-corrected weighted average) Thermo-Motility Index of fossil synapsids along the mammalian lineage, shows a coefficient of determination $R^2 = 0.38$. Model 67 contains four clusters, which correspond to non-eucynodont synapsids (minus *Thrinaxodon*), non-probainognathain eucynodonts (plus *Thrinaxodon*), non-mammaliamorph probainognathians, and non-mammal mammaliamorphs. We used these clusters to describe the evolution of the body temperature of fossil synapsids (Section 24), illustrated in Extended Data Figure 5.

Model	AICc	Relative likelihood
One cluster	-48.6	0.00
- *Non-mammalian synapsids*		
Two clusters	-61.6	0.28
- *Non-mammaliamorph synapsids*		
- *Non-mammal mammaliamorphs*		
Three clusters	-63.2	0.63
- *Non-eucynodont synapsids (minus* Thrinaxodon*)*		
- *Non-mammaliamorph eucynodonts (plus* Thrinaxodon*)*		
- *Non-mammal mammaliamorphs*		
Four clusters	-64.1	1.00
- *Non-eucynodont synapsids (minus* Thrinaxodon*)*		
- *Non-probainognathians eucynodonts (plus* Thrinaxodon*)*		
- *Non-mammaliamorph probainognathians*		
- *Non-mammal mammaliamorphs*		

Five clusters	-62.9	0.53
- *Non-eutheriodont synapsids*		
- *Non-eucynodont eutheriodonts (minus* Thrinaxodon*)*		
- *Non-probainognathian eucynodonts (plus* Thrinaxodon*)*		
- *Non-mammaliamorph probainognathians*		
- *Non-mammalian mammaliamorphs*		

Model 67: Four clusters

Branch length transformations:

Pagel's λ [ML]: 0.000

95.0% CI: (NA, 0.252)

Response: TMI

	Df	Sum Sq	Mean Sq
Clusters	3	0.0088226	0.00294086
Residuals	51	0.0126753	0.00024854

Value of the Thermo-Motility-Index:

Cluster	Most likely value	95% confidence interval
Non-eucynodont synapsids (minus *Thrinaxodon*)	0.96	[0.92,0.99]
Non-probainognathian eucynodonts (plus *Thrinaxodon*)	1.11	[0.96,1.26]
Non-mammaliamorph probainognathians	0.94	[0.73,1.14]
Non-mammalian mammaliamorphs	1.29	[1.13,1.44]

Adjusted R^2: 0.3757, AICc: -64.1243

n = 55 (3DB sets non-mammalian fossil synapsids were included)

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98195, P = 0.5752, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S54) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.58133, U = 0.87432, P = 0.2177, P value computed via approximation method).

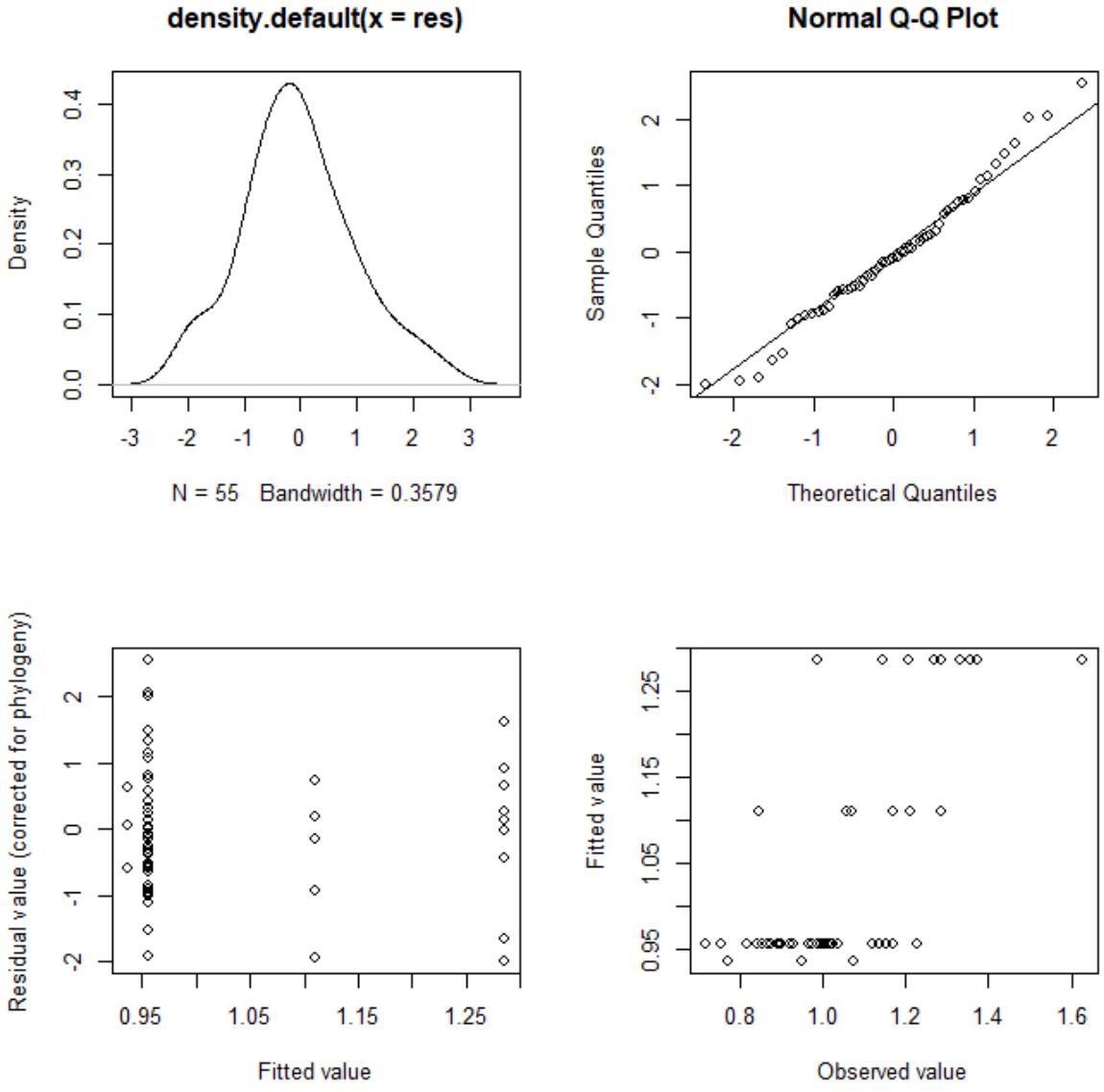


Figure S54 – Diagnostic plot for model 67.

24. Prediction of body temperature of fossil synapsids and related statistics

To predict the most likely body temperatures (T_B) of fossil synapsids, with 95% confidence intervals, we had to combine fitted temperature terms (i.e., T_R a ratio inversely related to body temperature, see Supplementary Methods, Biomechanics), predicted from the (size-corrected weighted average) Thermo-Motility Index by using Model 41, with residual values of the temperature term, simulated with Brownian motion by using extant species values and phylogenetic relationships as guides. Residuals of the temperature term are a measure of the mismatch between the metric we use (i.e., the Thermo-Motility Index) and the actual body temperature of the organism. This mismatch exists because the (size-corrected weighted average) Thermo-Motility Index carries behavioural information in addition to temperature information. Furthermore, the value of the (size-corrected weighted average) Thermo-Motility Index we compute in this study may slightly differ from the real value, because we do not have information on (1) the specific endolymph chemistry and (2) the residual variation relative to bony/membranous correlations, for each species we study. It is important to realise that these uncontrolled parameters (e.g., behaviours, endolymph chemistry, residual variation relative to bony/membranous correlations) are actually taken into account in our computations of probability distributions of predicted body temperatures. These distributions can thus be used directly to assess the likelihood of different assumptions (e.g., bird-like endolymph viscosity in cynognathians, sluggish behaviours in anomodonts) and their effect on the evolution of body temperature in fossil synapsids. Predicted body temperatures are provided in Table 1 and Supplementary Data 4. Computed probability distributions of predicted body temperatures are illustrated in Extended Data Figure 5.

The computation of probability distributions of predicted body temperatures was achieved in several steps that we describe below for clarity:

- Step 1: We used the (size-corrected weighted average) Thermo-Motility Index of fossil specimens to predict their fitted temperature term, using Model 41.

- Step 2: For all extant species with a high-viscosity endolymph chemistry (e.g., birds), we adjusted the (size-corrected weighted average) Thermo-Motility Index to that of a low-viscosity endolymph chemistry. This was done to increase the residual variation of the temperature term in Step 4, hence taking into account the possibility that high-viscosity endolymph could have evolved in clades branching out along the mammalian lineage.

- Step 3: We used the (size-corrected weighted average) Thermo-Motility Index of extant specimens for which we know the body temperature, to predict their fitted temperature term, using Model 41. We used

species in this step, rather than clades, because we wanted to further increase the residual variation of the temperature, to account for the highest possible uncertainty in body temperature predictions.

- Step 4: We subtracted the fitted temperature term of extant specimens from their actual temperature term to compute the residual distribution of the temperature term (Fig. S55).

- Step 5: We tested whether phylogeny structured the distribution of the residuals of the temperature term of extant species (Fig. S56). We concluded that phylogeny strongly structured the distribution (K = 0.943612, P = < 0.0001 *, P value calculated via approximation method) and should be taken into account to improve predictions of fossil specimens.

- Step 6: We fitted several models of evolution on the residual temperature term, using a phylogenetic tree that focused only on extant species for which we know the body temperature. We used the corrected version of the Akaike information criterion (AICc) to select the model that best explained the evolution of the residual temperature term (Model 71, see Table below, and further below for details on evolutionary models).

Model	AICc	Relative likelihood
68: Brownian motion	-576	0.00
69: Ornstein–Uhlenbeck process	-629	0.04
70: Early burst	-604	0.00
71: Pagel's λ	-636	1.00
72: Pagel's δ	-617	0.00
73: Pagel's κ	-598	0.00

- Step 7: We used the residual temperature term of extant species for which we know the body temperature, along with the phylogenetic tree corresponding to the best evolutionary model (Model 71), to estimate ancestral states of residual temperature terms and their 95% confidence interval for the node *Eothyris*-Mammalia (i.e., the 'root', -0.0087, 95% C.I. [-0.0461,0.0286]) and all other crownward nodes along the mammalian lineage, up to the node *Hadrocodium*+Mammalia.

Step 8 to 13 were replicated 15,000 times.

- Step 8: We randomly sampled a value from the predicted probability distribution of the residual temperature term computed in Step 7 for the 'root' (Fig. S57).

- Step 9: Using the value sampled in Step 8 as a starting point, we randomly simulated Brownian evolution of the residual temperature term along the tree corresponding to the best evolutionary model (Model 71), obtained in Step 6, by using the evolutionary rate obtained in the same model (Fig. S57).

- Step 10: We assessed whether values obtained during the simulation for nodes along the mammalian lineage fell within the 95% confidence interval predicted for these same nodes in Step 7. If at least one simulated value at any of these nodes fell outside the 95% confidence interval, we rejected the simulation. This is because the evolutionary history of stem-mammals has already happened and we know the outcome. Therefore, this step allows us to prevent the simulation to produce meaningless values that would significantly deviate from what is observed in extant mammals and what was most likely the condition at their base.

- Step 11: For simulations that successfully passed the test in Step 10, we combined the fitted temperature term of each fossil synapsid species to their simulated residual temperature, to obtain a predicted temperature term.

- Step 12: For simulations that successfully passed the test in Step 10, we computed the mean of predicted temperature terms for each relevant fossil synapsid groups (e.g., non-mammalian mammaliamorphs, cynognathians). Computing the means of fossil groups during each iteration of the simulation, and not just averaging the final distributions of each relevant specimens, allows us to acknowledge the impact of shared evolutionary history when computing the probability distributions of group means.

- Step 13: For simulations that successfully passed the test in Step 10, we computed the difference between mean predicted temperature terms for each relevant fossil synapsid groups (e.g., non-mammalian mammaliamorphs *vs* cynognathians). As for Step 12, computing the differences in the mean predicted temperature terms between fossil groups during each iteration of the simulation, and not just calculating the difference between the final distributions of each fossil synapsid groups, allows us to acknowledge the impact of shared evolutionary history when computing probability distributions of group mean differences.

In total we obtained 9647 valid simulations out of 15,000

- Step 14: We computed the mean and standard deviation of the 9647 predicted temperature terms, for each fossil synapsid species.

- Step 15: We computed the mean and standard deviation of the 9647 mean predicted temperature terms calculated in Step 12, for each fossil synapsid group. Corresponding values are provided in Table 1.

- Step 16: We computed the mean and standard deviation of the 9647 differences in the mean predicted temperature terms between fossil groups calculated in Step 13, for each fossil synapsid group combination (Fig. S59).

- Step 17: We computed probability distributions of predicted body temperatures for each fossil synapsid groups, using their respective probability distributions of predicted temperature term (see Supplementary Methods, Biomechanics). Calculated probability distributions of predicted body temperatures are illustrated in Extended Data Figure 5.

- Step 18: Using calculated means and standard deviations of predicted temperature terms, and assuming a normal distribution, we computed the probability that the body temperature of fossil synapsid groups was higher or equal to 31°C. Results are provided in Table 1.

- Step 19: Using calculated means and standard deviations of differences in the mean predicted temperature terms between fossil groups, and assuming a normal distribution, we computed the probability that the body temperature of fossil synapsid groups was higher or equal to that of non-mammal mammaliamorphs. Results are provided in Table 1.

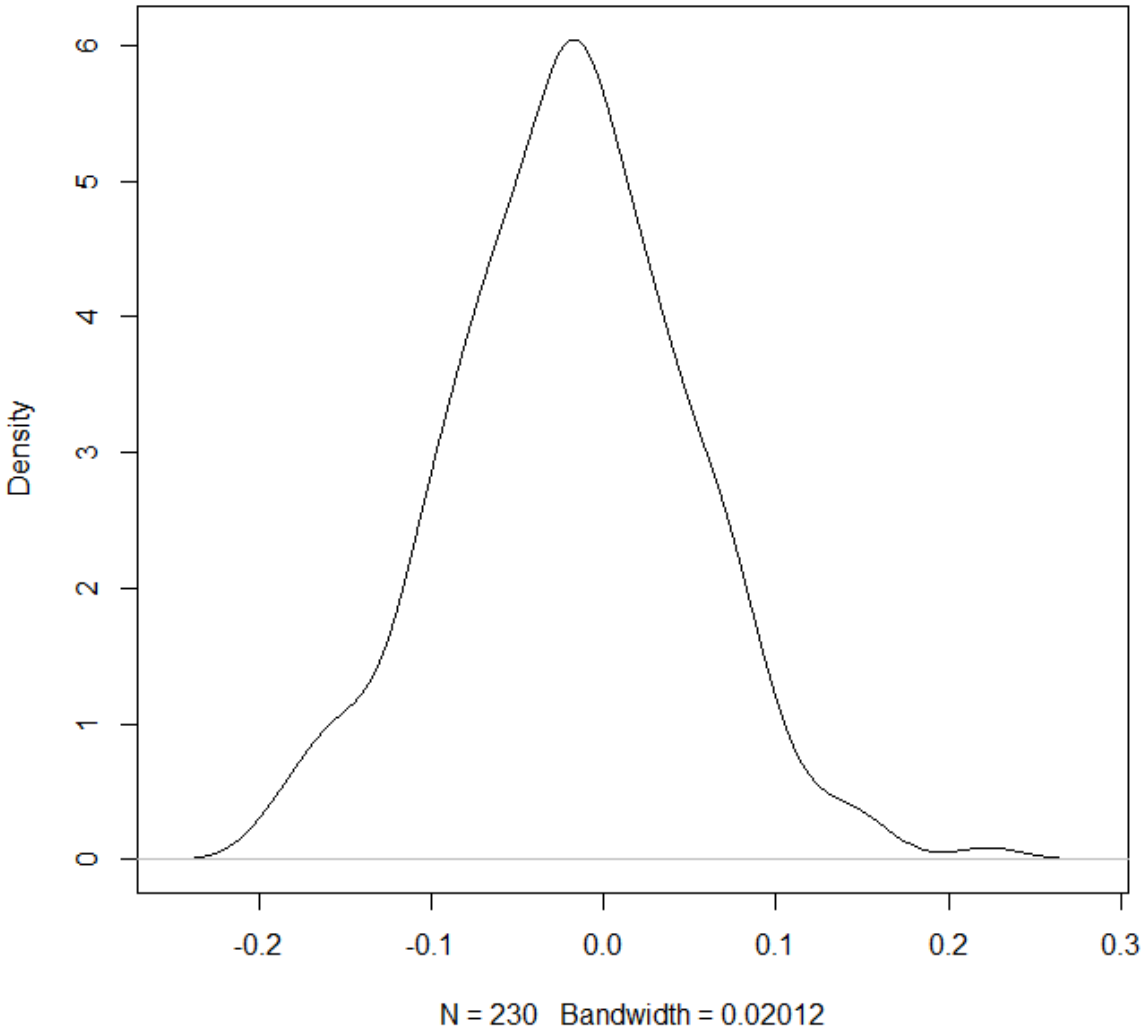


Figure S55 – Distribution of residuals of the temperature term of extant species.

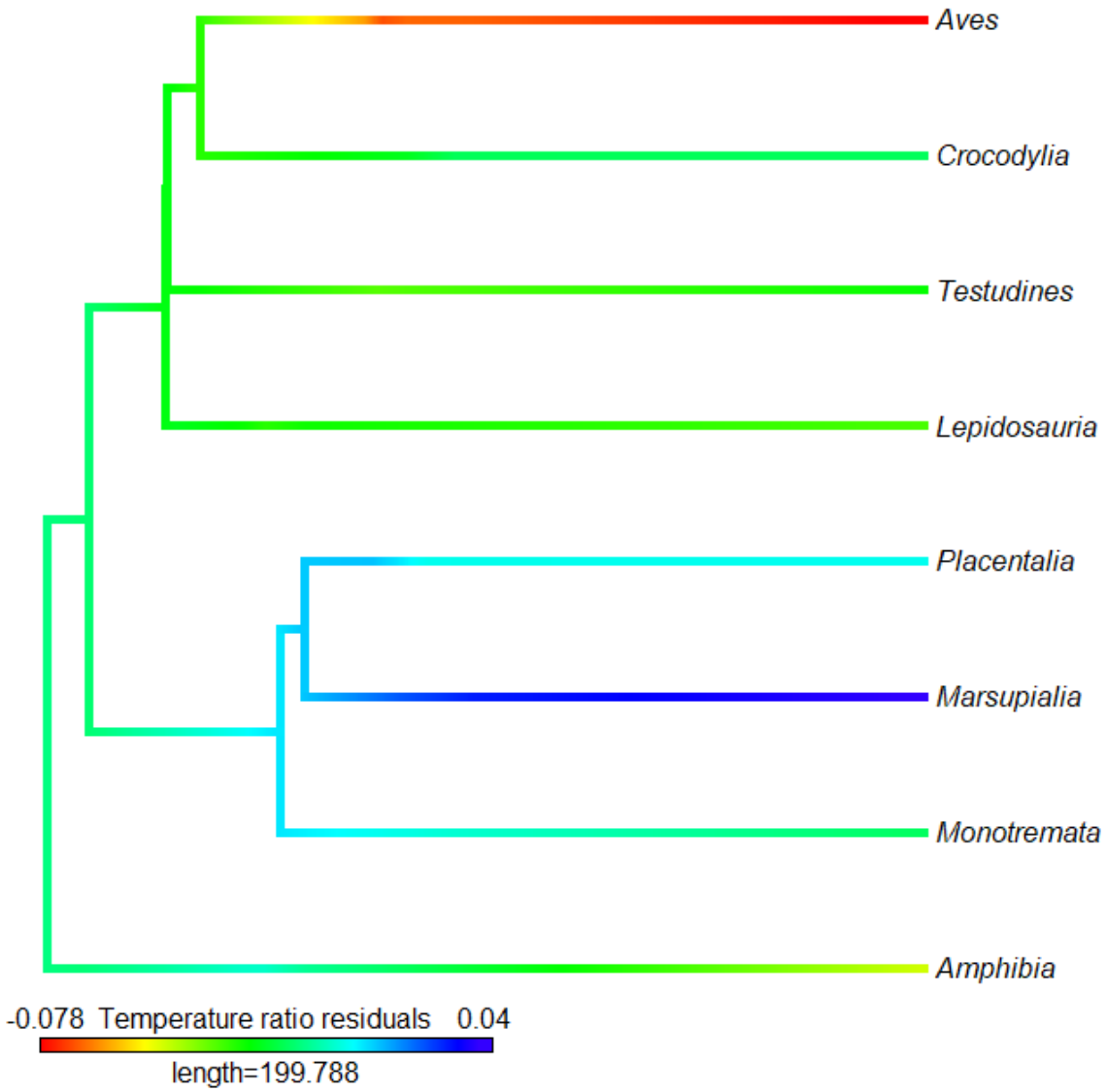


Figure S56 – Phylogenetic distribution of the residuals of the temperature term.

Model 68: Brownian motion

Brownian variance: $4.96622\ 10^{-05}$

Root state: 0.002008276

AICc: -576.1901

Model 69: Ornstein–Uhlenbeck process

Alpha [ML]: 0.01067376, 95% CI = [0.007442972, 0.01458885]

Brownian variance: $9.720138\ 10^{-05}$

Root state: -0.009617323

AICc: -629.2932

Model 70: Early burst

ACDC [ML]: 0.006333472, 95% CI = [0.005374112, 0.006333472]

Brownian variance: $3.001478\ 10^{-06}$

Root state: -0.001039781

AICc: -604.3833

Model 71: Pagel's λ

Pagel's λ [ML]: 0.610036, 95% CI = [0.3339972, 0.808227]

Brownian variance: $1.290695\ 10^{-05}$

Root state: 0.0002273189

AICc: -635.9712

Model 72: Pagel's δ

Pagel's δ [ML]: 5, 95% CI = [4.386971, 5]

Brownian variance: $2.631947\ 10^{-16}$

Root state: -0.005187285

AICc: -616.7988

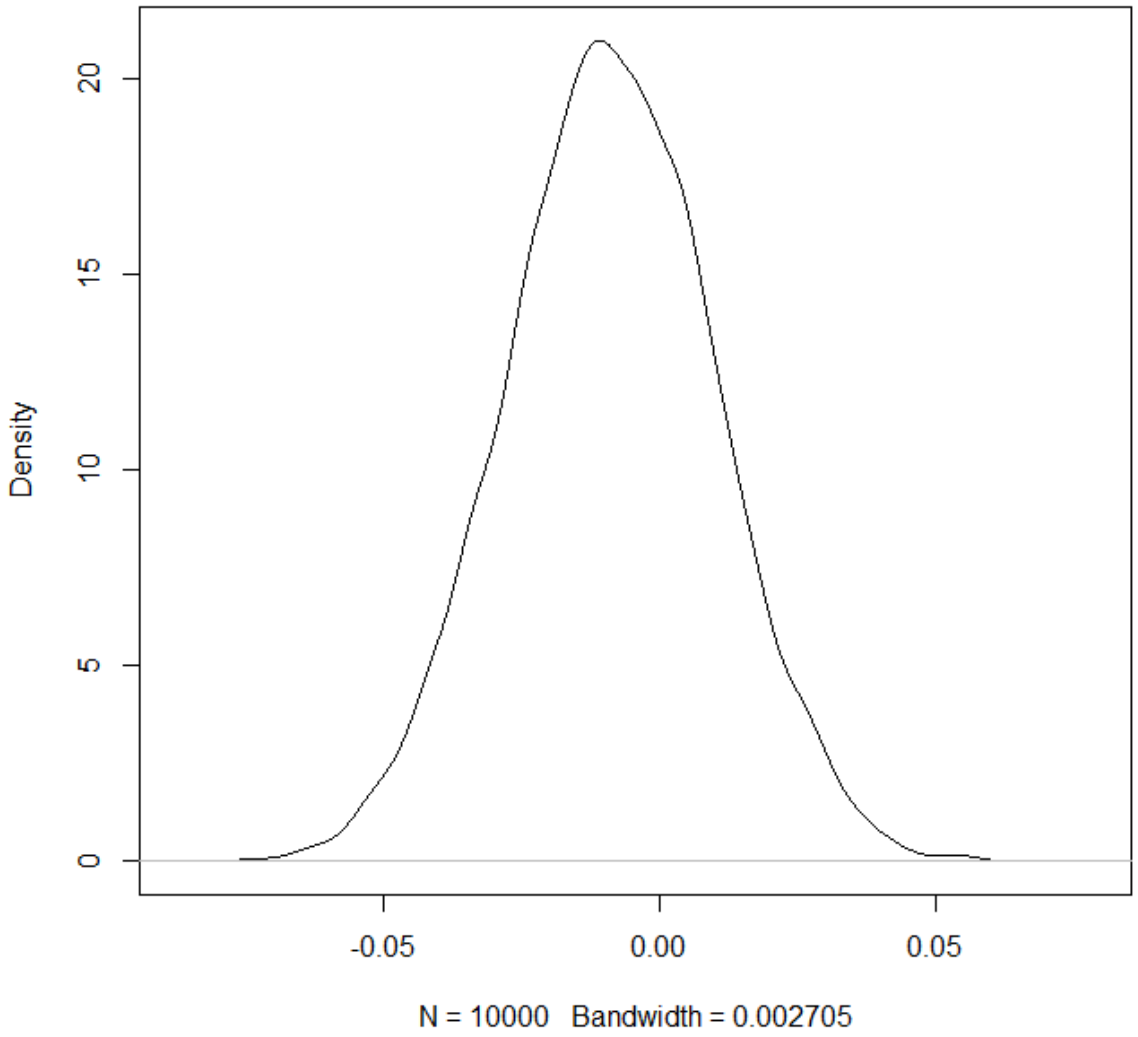
Model 73: Pagel's κ

Pagel's κ [ML]: 0.4225069, 95% CI = [0.1957719, 0.6581099]

Brownian variance: 0.0003877604

Root state: 0.00719144

AICc: -597.8352



3070

3071 Figure S57 – Probability density function of the value of the temperature residual at the node *Eothyris*-

3072 Mammalia.

3073

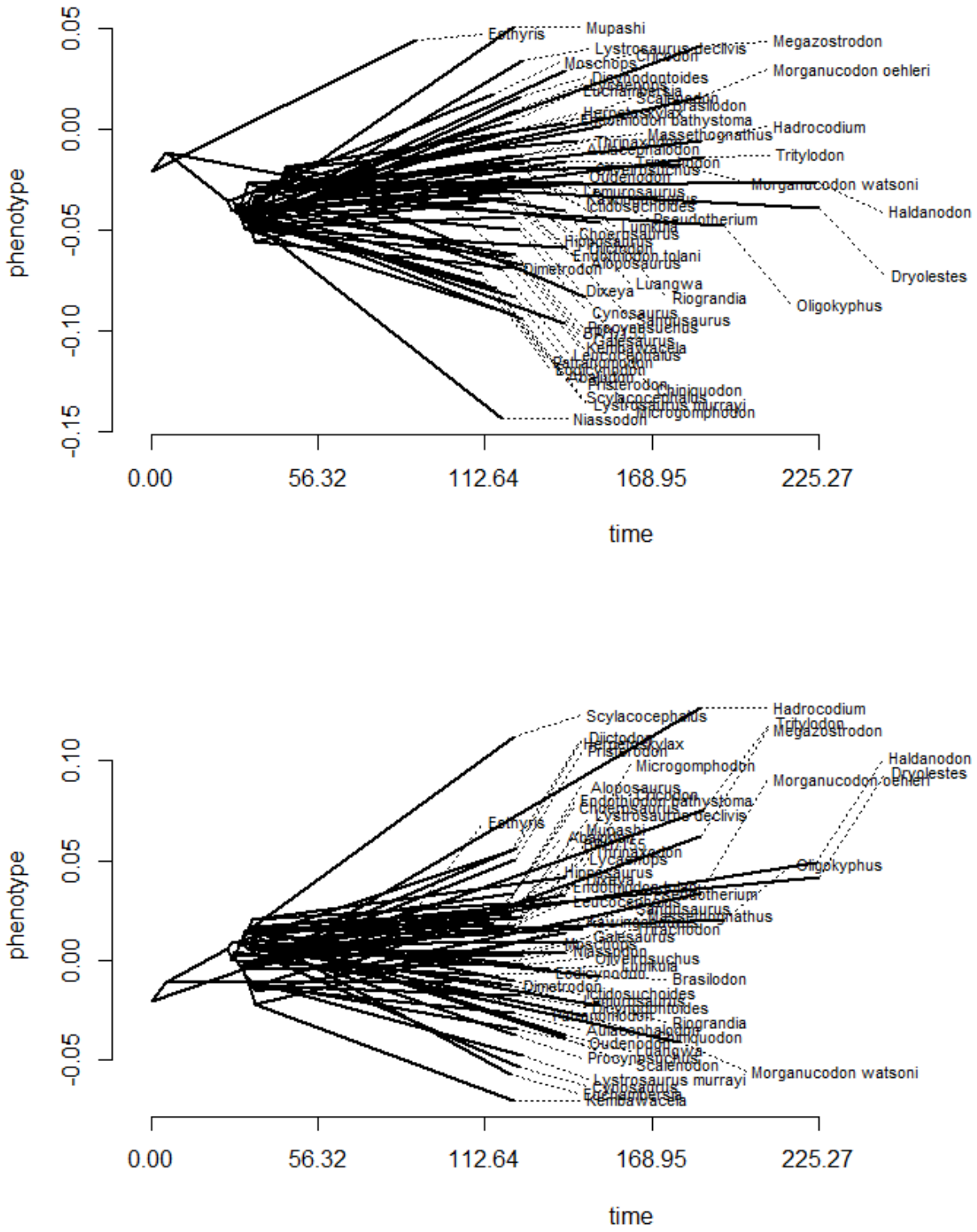


Figure S58 – Example of two simulations of the evolution of the residual of the temperature term in fossil synapsids under Model 71. The simulations have been obtained in R using seeds 27 and 72 respectively.

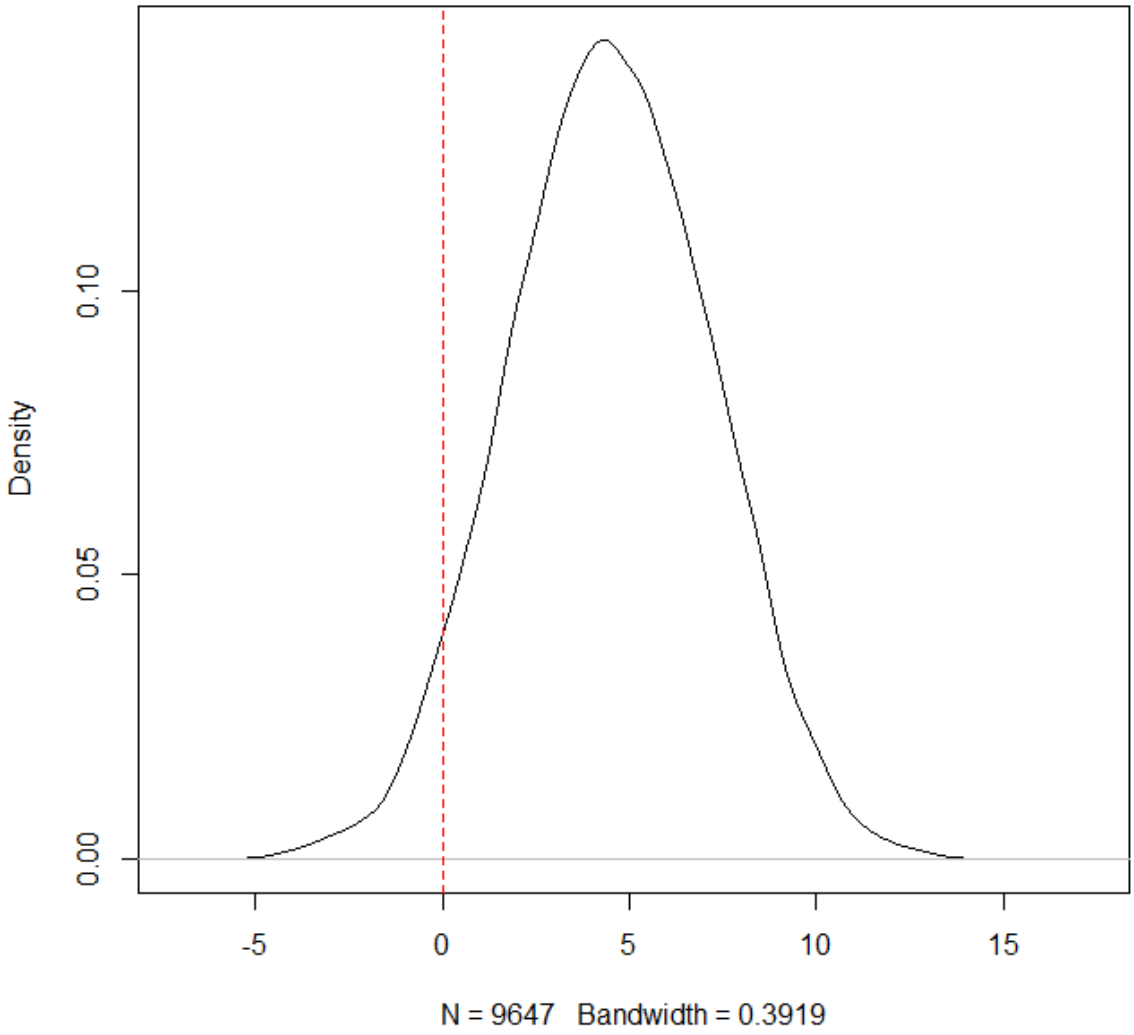


Figure S59 – Probability density function of the differences in the mean predicted temperature terms between non-mammal mammaliamorphs and non-probainognathian eucynodonts (plus *Thrinaxodon*). Note that while the body temperature distributions of the two group show some overlap in Extended Data Figure 5, the likelihood that non-probainognathian eucynodonts (plus *Thrinaxodon*) had higher body temperatures than non-mammalian mammaliamorphs is actually very low (this corresponds to the area of the distribution that is on the left of the red dashed line) when shared history is taken into account. This means that simulated body temperature values on the high end of the distribution of non-probainognathian eucynodonts (plus *Thrinaxodon*) are generally accompanied by simulated body temperature values on the high end of the distribution of non-mammal mammaliamorphs (see Extended Data Figure 5), because of shared ancestry.

25. PGLS-R of the Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct against the (size-corrected weighted average) Thermo-Motility Index

Model 74 provides a very significant regression ($P << 0.001$), with a coefficient of determination (R^2) of 0.57. It was tested on clades for which we know the average body temperature, and was used to assess the relationship between the size-corrected Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct (TMI_Ω_2_l, Supplementary Methods, Biomechanics) and the (size-corrected weighted average) Thermo-Motility Index (TMI, Supplementary Methods, Biomechanics). The resulting regression line is illustrated on Extended Data Figure 7. The intercept was modified to better highlight the separation between extant endothermic and ectothermic amniotes, as well as the intermediate position of fossil synapsids. Because the lateral semicircular canal seems the most related to behavioural agility (Cox & Jeffery 2010), and because its size-corrected Thermo-Motility index, calculated from the saturating velocity, is the least good predictor of the temperature term (see Section 11), the relationship between the size-corrected Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct, and the (size-corrected weighted average) Thermo-Motility Index is expected to provide insight onto a possible uncoupling between body temperature and behavioural activity. In this context, Extended Data Figure 7 shows that while non-mammal mammaliamorphs are within extant endotherms in regard to the (size-corrected weighted average) Thermo-Motility Index, they are intermediate between extant endotherms and ectotherms in regard to the size-corrected Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct. We interpret this observation as a potential indication that non-mammalian mammaliamorphs reached mammal-like body temperatures first, and that mammal-like locomotor activity developed at a later stage. In addition, it is interesting to note that for a given (size-corrected weighted average) Thermo-Motility Index, fossil synapsids generally plot higher than extant ectotherms for the size-corrected Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct. We interpret this observation as hinting at increased behavioural activity in fossil synapsids when compared to diapsid ectotherms with similar body temperatures. Nevertheless, note that these observations could also be explained by significant differences between residual duct-morphology components of the size-corrected Thermo-Motility Index of the lateral semicircular duct, calculated from the saturating velocity, and of the (size-corrected weighted average) Thermo-Motility Index. Furthermore, while a change in endolymph viscosity would isometrically affect both Thermo-Motility Indices, it could also potentially explain some of these observations, because the relationship between the size-corrected Thermo-Motility Index, calculated from the saturating velocity of the lateral semicircular duct, and the (size-corrected weighted average) Thermo-Motility Index is not isometric.

Model 74: log10(TMI_Ω_2_l) ~ log10(TMI)

Branch length transformations:

Pagel's λ [ML]: 1.000

lower bound: 0.000, $P = 6.7778\ 10^{-06}$ *

upper bound: 1.000, $P = 1$

95.0% CI: (0.814, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.56641	0.11294	5.0151	$3.229\ 10^{-05}$ *
$\log_{10}$(TMI)	0.68503	0.11248	6.0902	$1.952\ 10^{-06}$ *

P values calculated using the exact method.

Adjusted R^2: 0.572, AICc: -22.1491

F statistic: 37.09 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: $1.952\ 10^{-06}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.97171$, $P = 0.6271$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S60) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.51873$, $U = 0.75633$, $P = 0.1064$, *P* value computed via approximation method).

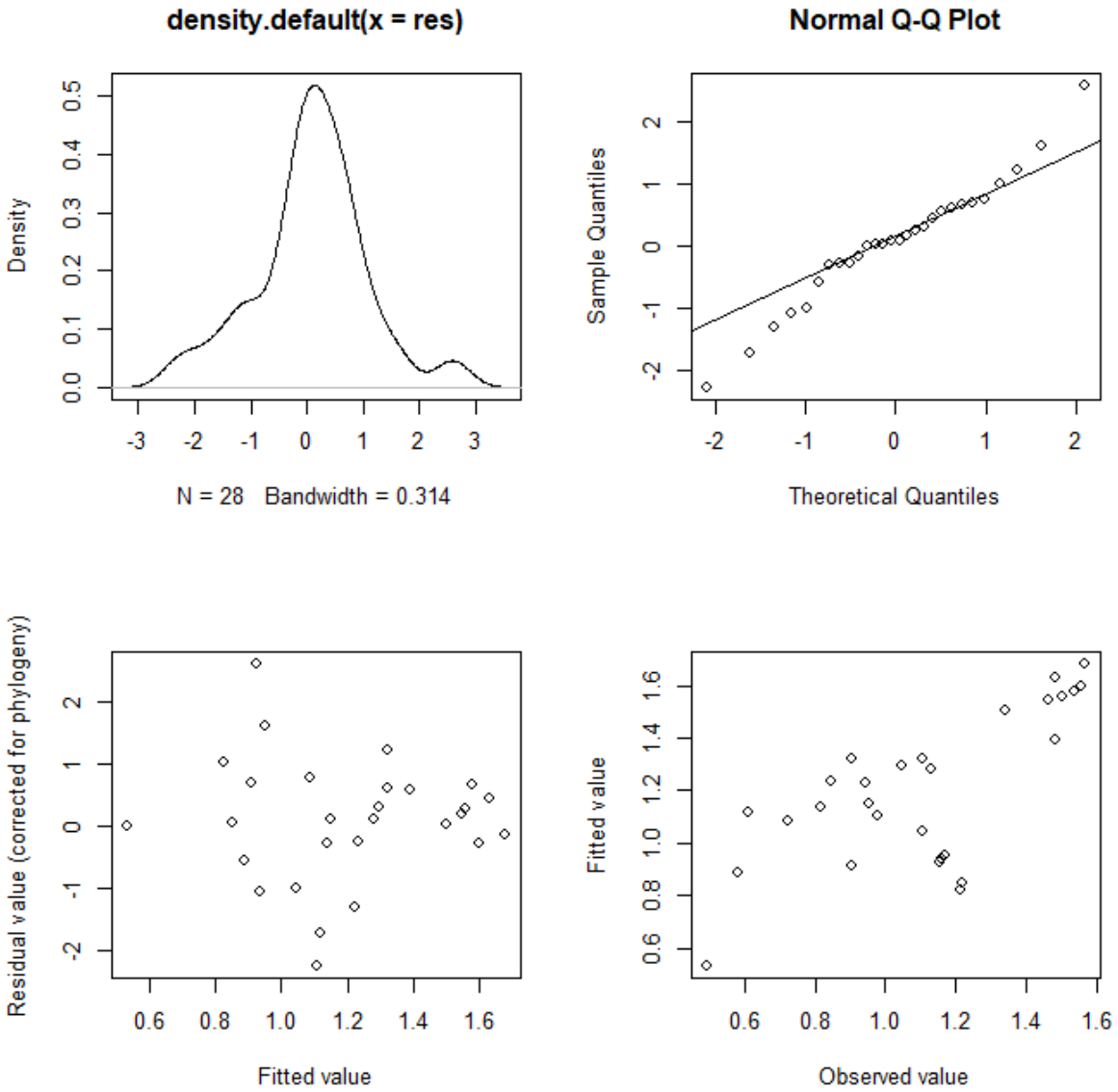


Figure S60 – Diagnostic plots for model 74.

26. Assessing the likelihood that the real group mean of the (size-corrected weighted average) Thermo-Motility Index of non-mammaliamorph probainognathians could affect the existence of a shift to endothermy

In this study, we concluded that although there was a gradual stochastic increase in body temperatures from most basal fossil synapsids (e.g., *Eothyris*) to cynognathians (see Table 1, Extended Data Figure 5), endothermy was abruptly acquired along the branch directly leading to Mammaliamorpha. This is reflected by a major shift in the evolutionary rate of the (size-corrected weighted average) Thermo-Motility Index at this location (see Fig. 4). Arguably, these observations are partially resulting from the particularly low (size-corrected) Thermo-Motility Indices we obtained for non-mammaliamorph probainognathians sampled in this study, whose mean is similar to what is found in non-eucynodont synapsids (see Table 1). Therefore, because our non-mammaliamorph probainognathian sample is small (n = 3), we wanted to assess how likely it would be for their true mean (size-corrected weighted average) Thermo-Motility Index to be higher than the one we calculated, and how this would affect our conclusions. This was achieved in several steps that we describe below for clarity:

- Step 1: We computed the standard deviation of the (size-corrected weighted average) Thermo-Motility Index of anomodonts (n = 16, stdv = 0.125), mammals (n = 77, stdv = 0.177), birds (n = 72, stdv = 0.179), and lepidosaurs (n = 67, stdv = 0.201). These groups were chosen because Anomodontia is the largest fossil group sampled, whereas mammals, birds, and lepidosaurs are the largest extant groups sampled in this study.

- Step 2: We calculated the mean standard deviation of the (size-corrected weighted average) Thermo-Motility Index, between the standard deviation of anomodonts, and the mean of the standard deviations of mammals, birds, and lizards. We used this mean standard deviation (stdv = 0.155) in the next steps, as an indication of what the standard deviation of non-mammaliamorph probainognathians would likely be, if we would have had a much larger sample. We did not average all standard deviations directly, because it would have biased our estimate of the true standard deviation of non-mammaliamorph probainognathians toward higher values. Indeed, because extant groups have a longer evolutionary history than non-mammaliamorph probainognathians, the variation in their (size-corrected weighted average) Thermo-Motility Index is also expected to be higher, simply because of Brownian motion.

- Step 3: We derived a theoretical distribution of the (size-corrected weighted average) Thermo-Motility Index for a very large non-mammaliamorph probainognathian sample, assuming a normal distribution whose mean is similar to that of cynognathians, and whose standard deviation equals the one calculated in Step 2 (Fig. S60).

-Step 4: We randomly sampled three specimens from the theoretical distribution obtained in Step 3, and we calculated the corresponding mean (size-corrected weighted average) Thermo-Motility Index. Step 4 was replicated 10,000 times, resulting in 10,000 different means.

- Step 5: We computed the mean and standard deviation of the 10,000 mean (size-corrected weighted average) Thermo-Motility Indices obtained in Step 4. Assuming normality, the resulting distribution (Fig. S61) illustrates the likelihood of obtaining a given mean (size-corrected weighted average) Thermo-Motility Index, when randomly sampling three specimens from a population with a mean similar to that of cynognathians and a standard deviation comparable to what is observed in the larger groups of this study (e.g., mammals).

- Step 6: Using the probability density function calculated in Step 5, we computed the probability, when sampling three specimens from a population of non-mammaliamorph probainognathians whose true mean is similar to that of cynognathians and whose standard deviation is comparable to what is observed in the larger groups of this study (e.g., mammals), of randomly obtaining a mean (size-corrected weighted average) Thermo-Motility Index that is lower than, or equal to, the one we calculated in our study.

If the true mean of non-mammaliamorph probainognathians is actually similar to that of cynognathians, we calculated that there was only a probability of 4.2% to obtain, by chance alone, a mean (size-corrected weighted average) Thermo-Motility Index that is lower or equal to the one we obtained in our study. This result strongly suggests that the true mean of non-mammaliamorph probainognathians is lower than the mean of cynognathians. It further supports our interpretation concerning the gradual stochastic increase observed in body temperatures, from most basal fossil synapsids (e.g., *Eothyris*) to cynognathians, which likely stopped, or even reversed, in non-mammaliamorph probainognathians.

To go further, and assess the potential effect our small non-mammaliamorph probainognathian sample could have had on our conclusions concerning the existence of an abrupt shift at the root of Mammaliamorpha, we searched for the highest possible mean (size-corrected weighted average) Thermo-Motility Index that non-mammaliamorph probainognathian could have had without affecting the existence of the shift. We found that adding 0.086 to the (size-corrected weighted average) Thermo-Motility Index of *Lumkuia fuzzi*, *Chiniquodon theotonicus*, and *Riograndia guaibensis*, resulting in a mean of 1.018 for non-mammaliamorph probainognathians, was the threshold above which the existence of a shift at the root of Mammaliamorpha was not the most likely evolutionary model anymore (Model 75; Fig. S62). We thus assessed how likely it would be, if the true mean (size-corrected weighted average) Thermo-Motility Index of non-mammaliamorph probainognathians was higher than this threshold of 1.018, to obtain by chance

alone the mean we calculated in our study. This was achieved in several steps that we describe below for clarity:

- Steps 1b and 2b are the same as Steps 1 and 2.

- Steps 3b-6b correspond to Steps 3-6 when replacing the cynognathian-like mean we used for the theoretical true mean of the (size-corrected weighted average) Thermo-Motility Index of non-mammaliamorph probainognathians, with the threshold of 1.018 above which the shift in evolutionary rate at the root of Mammaliamorpha is lost (Figs. S63-64).

If the true mean of non-mammaliamorph probainognathians is actually higher than 1.018, we calculated that there was only a probability of 16.9% to obtain, by chance alone, a mean (size-corrected weighted average) Thermo-Motility Index that is lower or equal to the one we obtained in our study, and to wrongly conclude on the existence of an abrupt shift to endothermy at the root of Mammaliamorpha. This result suggests that the true mean of non-mammaliamorph probainognathians is likely lower than the threshold of 1.018, and that our interpretations concerning the existence of an abrupt shift to endothermy at the root of Mammaliamorpha is likely not a by-product of our small non-mammaliamorph probainognathian sample but an actual palaeobiological event.

As a final test, we checked whether a mean (size-corrected weighted average) Thermo-Motility Index, similar to that of non-mammal mammaliamorphs, could have been already achieved in cynognathians, and found with a larger sample. Like the two previous analyses, this was achieved in several steps that we describe below for clarity:

- Steps 1c and 2c are the same as Steps 1 and 2.

- Steps 3c-6c correspond to Steps 3-6 when replacing the cynognathian-like mean we used for the theoretical true mean of the (size-corrected weighted average) Thermo-Motility Index of non-mammaliamorph probainognathians, with a non-mammal mammaliamorph-like mean we use for the theoretical true mean of the (size-corrected weighted average) Thermo-Motility Index of cynognathians (Figs. S65-66). Five specimens were randomly sampled in Step 4c (to mirror the size of our cynognathian sample), instead of three in Step 4.

If the true mean of cynognathians is actually similar to that of non-mammal mammaliamorphs, we calculated that there was only a probability of 0.2% to obtain, by chance alone, a mean (size-corrected weighted average) Thermo-Motility Index that is lower or equal to the one we obtained in our study. This result strongly suggests that the true mean of cynognathians is significantly lower than the mean of non-

mammal mammaliamorphs, again supporting the existence of a shift to endothermy at a later time, in Mammaliamorpha.

Model 75: Pagel's λ with a major shift along the branch directly leading to Mammaliamorpha (using the threshold non-mammaliamorph probainognathian mean)

We tested from 0–2 shifts in the rate of evolution, with no *a priori* assumptions about where the shifts were expected to occur. The best model we retrieved support one major shift along the branch going toward Mammaliamorpha and a minor shift in *Diadectes*. This minor shift seems artefactual because it always occurred at the level of the outgroup, irrespective of the phylogenetic configuration we tested.

Number of shifts: 2

Position of first shift: *Diadectes*

Rate of the first shift: $1\ 10^{-08}$, 95% CI = [NA, $1\ 10^{-08}$]

Position of the second shift: along the branch directly leading to Mammaliamorpha

Rate of the second shift: 330.6, 95% CI = [19.5, NA]

Model	AICc	Relative likelihood
Pagel's λ	-49.30	0.00
75: Pagel's λ with a major shift along the branch directly leading to Mammaliamorpha	-67.14	1.00

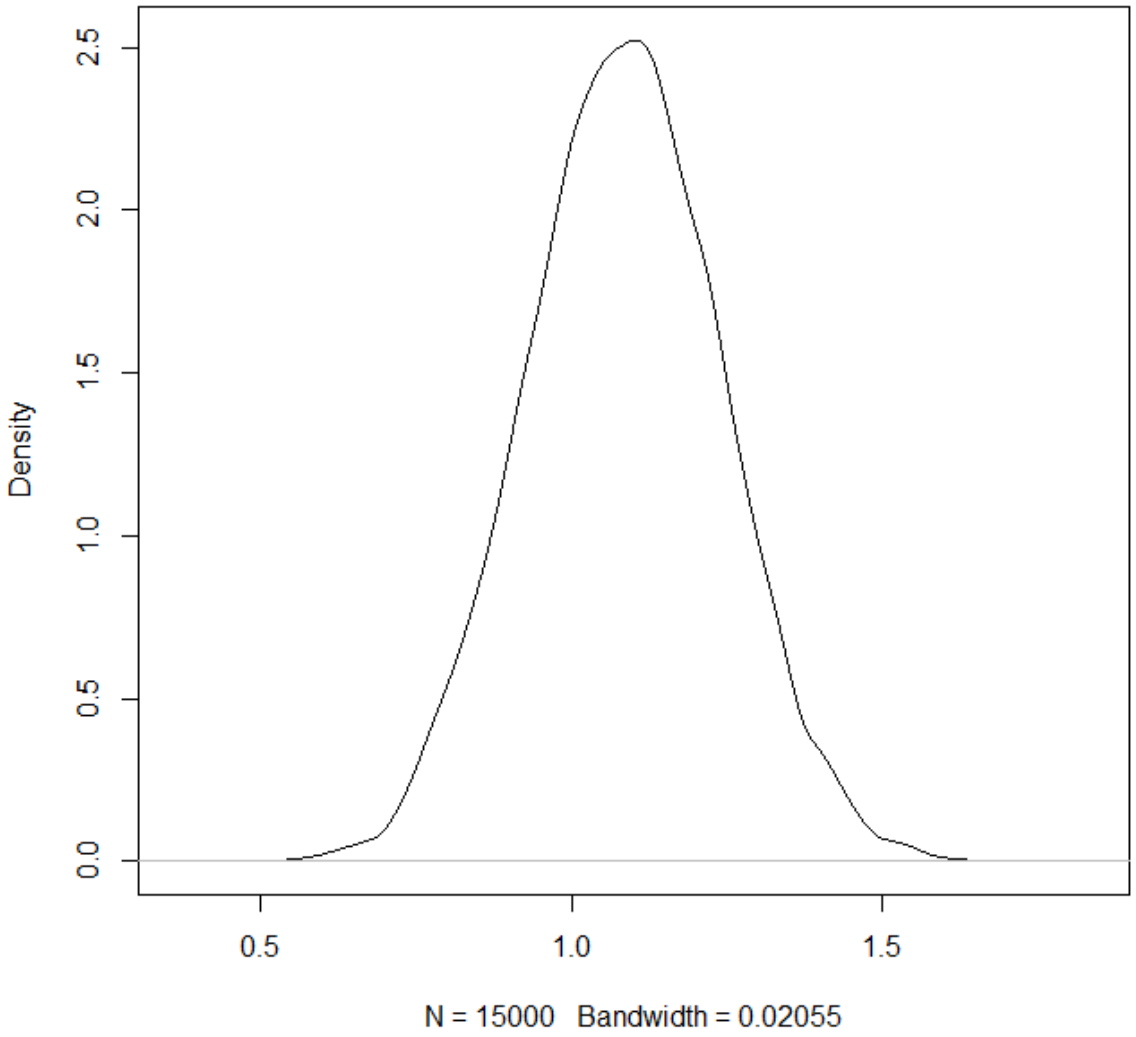


Figure S60 – Theoretical distribution of a very large non-mammaliamorph probainognathian sample (n = 15,000 specimens), assuming a normal distribution whose mean is similar to that of cynognathians, and whose standard deviation is comparable to what is observed in the larger groups of this study (e.g., mammals).

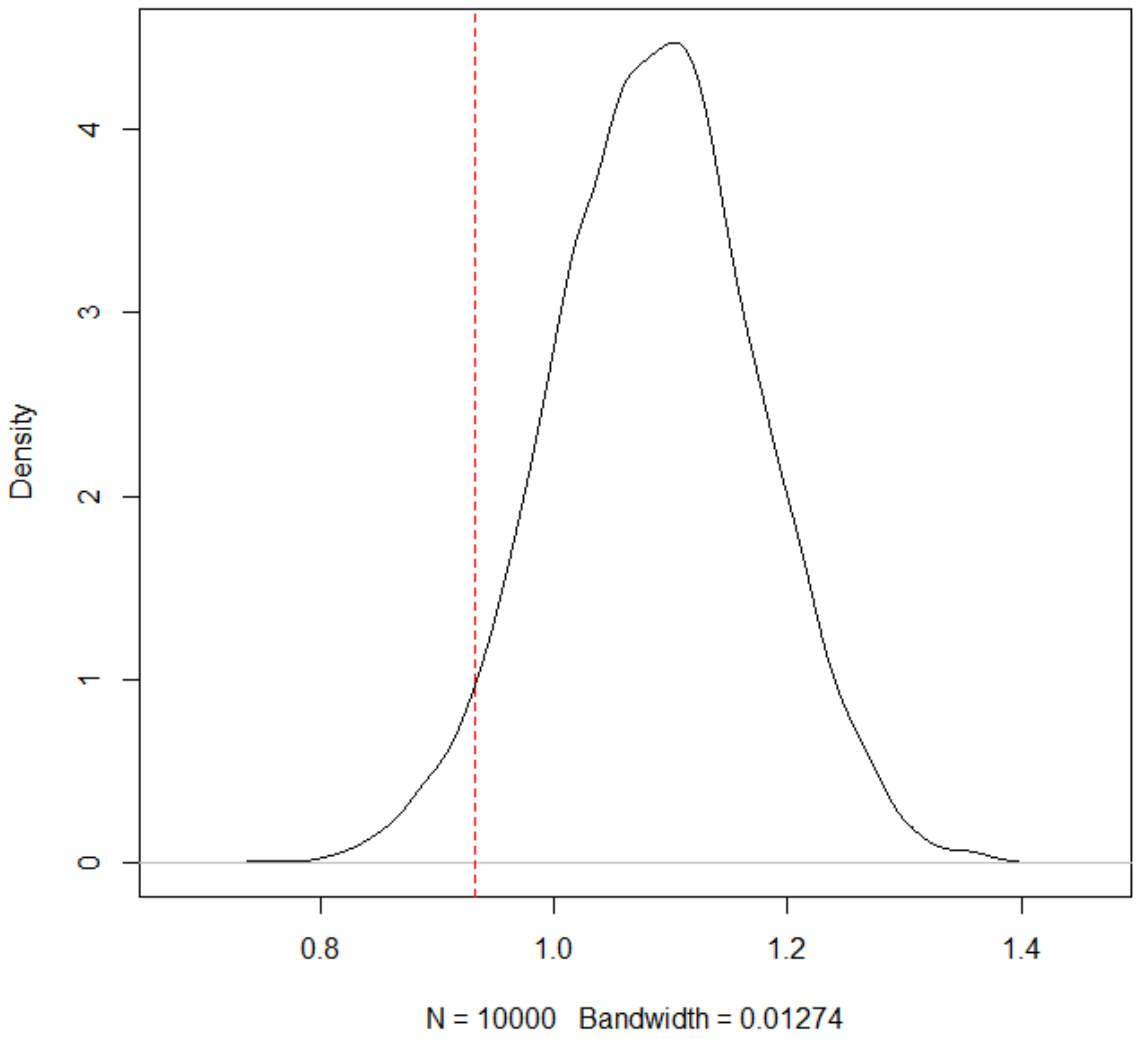


Figure S61 – Probability density function of the mean (size-corrected weighted average) Thermo-Motility Index, computed by using three randomly sampled non-mammaliamorph probainognathians from a population (i.e., Fig. S60) whose mean is similar to that of cynognathians, and whose standard deviation is comparable to what is observed in the larger groups of this study (e.g., mammals). The red dashed line represents the mean of non-mammaliamorph probainognathians we computed in our study. The area of the distribution located on the left of the red dashed line corresponds to the probability (P = 4.2%) of getting the mean we computed in our study, if the true mean of non-mammaliamorph probainognathians was similar to that of cynognathians. This very low probability suggests that the true mean of non-mammaliamorph probainognathians was likely lower than that of cynognathians. Indeed, if the true mean of non-mammaliamorph probainognathians was closer to the one we computed, both the area on the left of the red dashed line and the probability would increase.

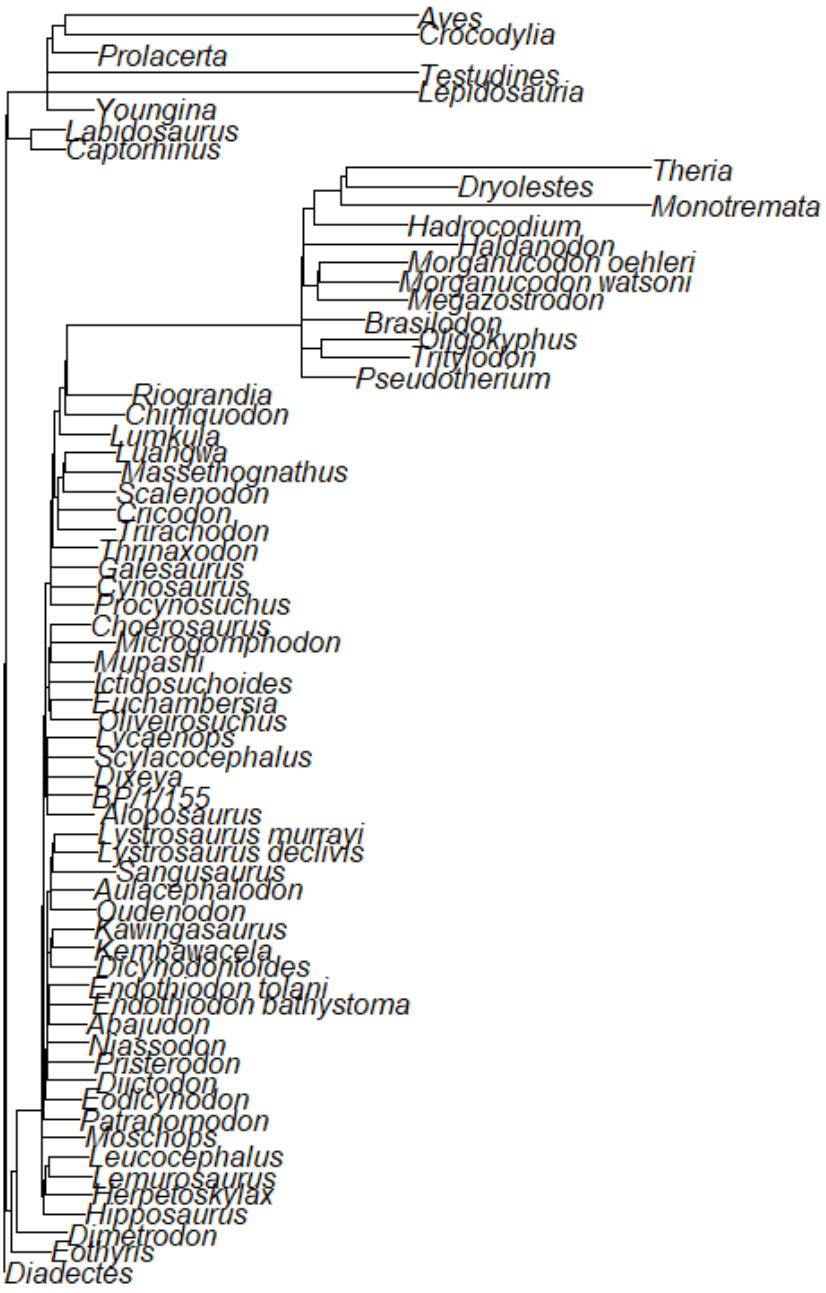


Figure S62 – Tree reflecting evolutionary model 75.

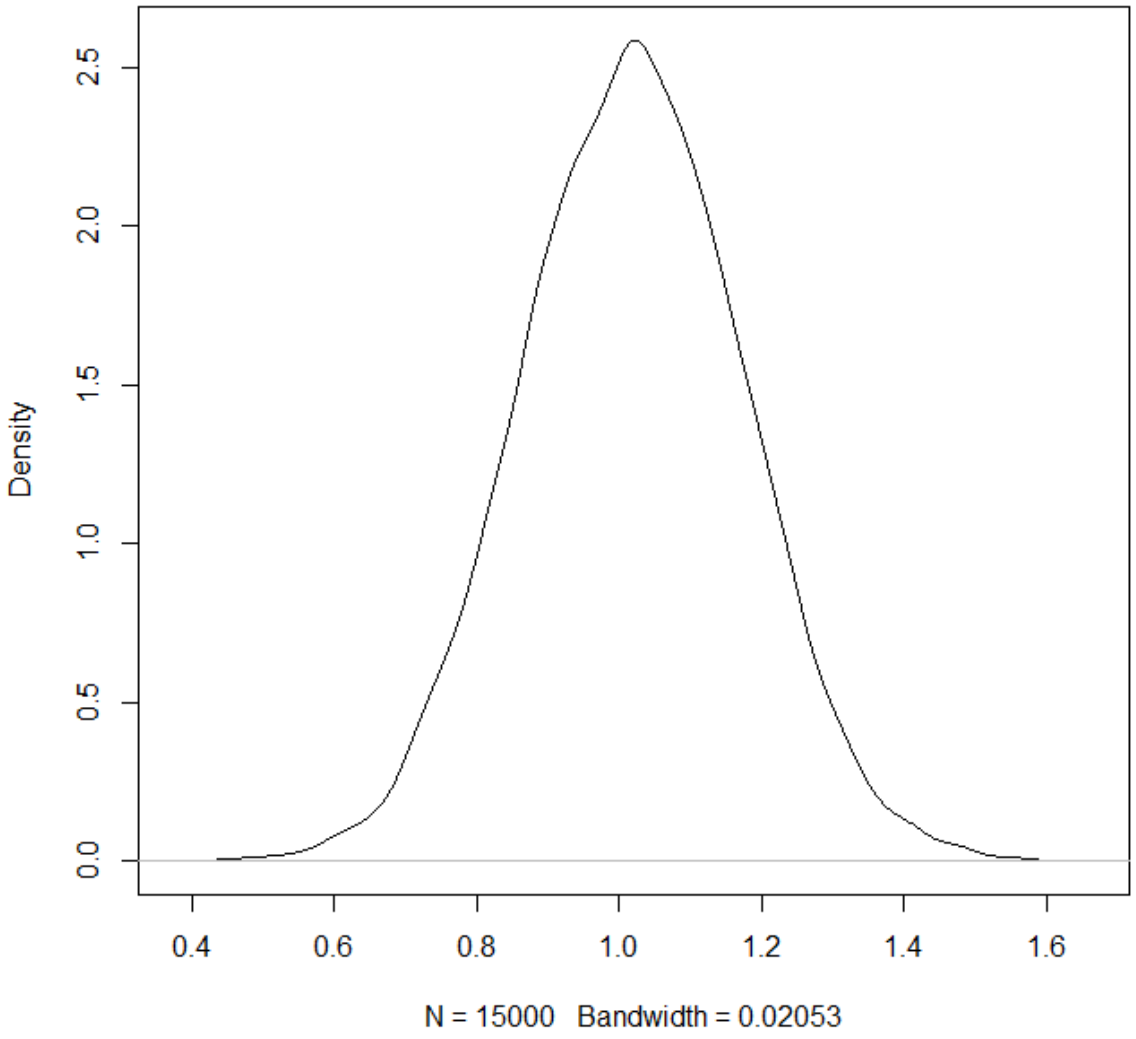


Figure S63 – Theoretical distribution of a very large non-mammaliamorph probainognathian sample (n = 15,000 specimens), assuming a normal distribution whose mean equals 1.018 (the threshold above which a shift in evolutionary rate on the branch directly leading to Mammaliamorpha is not the most informative model), and whose standard deviation is comparable to what is observed in the larger groups of this study (e.g., mammals).

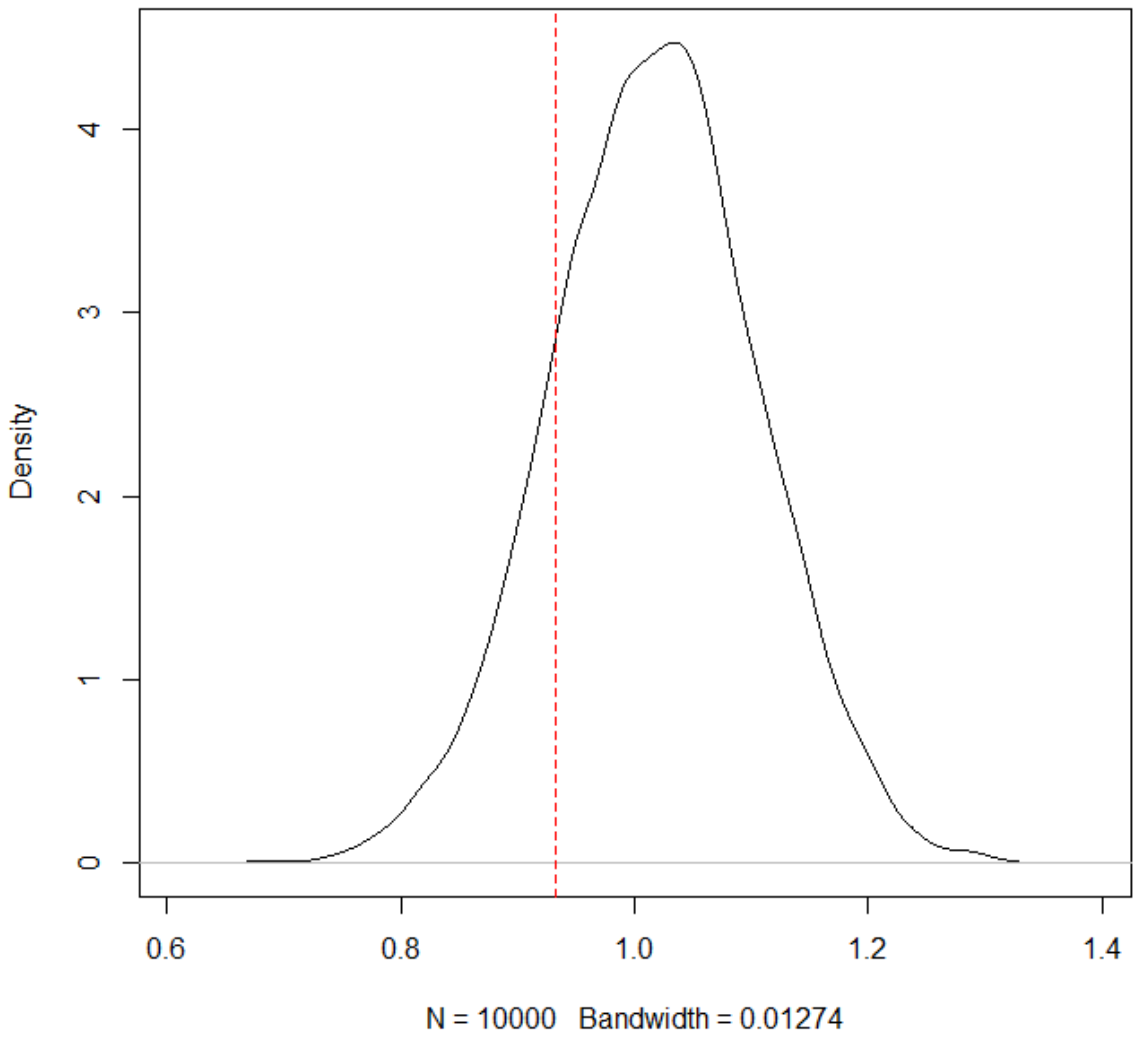


Figure S64 – Probability density function of the mean (size-corrected weighted average) Thermo-Motility Index, computed by using three randomly sampled non-mammaliamorph probainognathians from a population (i.e., Fig. S63) whose mean equals 1.018 (the threshold above which a shift in evolutionary rate on the branch directly leading to Mammaliamorpha is not the most informative model), and whose standard deviation is comparable to what is observed in the larger groups of this study (e.g., mammals). The red dashed line represents the mean of non-mammaliamorph probainognathians we computed in our study. The area of the distribution located on the left of the red dashed line corresponds to the probability (P = 16.9%) of getting the mean we computed in our study, if the true mean of non-mammaliamorph probainognathians was similar to that of cynognathians. This low probability suggests that the true mean of non-mammaliamorph probainognathians was likely lower than that threshold, supporting our conclusions on the abruptness of the shift to endothermy. Indeed, if the true mean of non-mammaliamorph probainognathians was closer to the one we computed, both the area on the left of the red dashed line and the probability would increase.

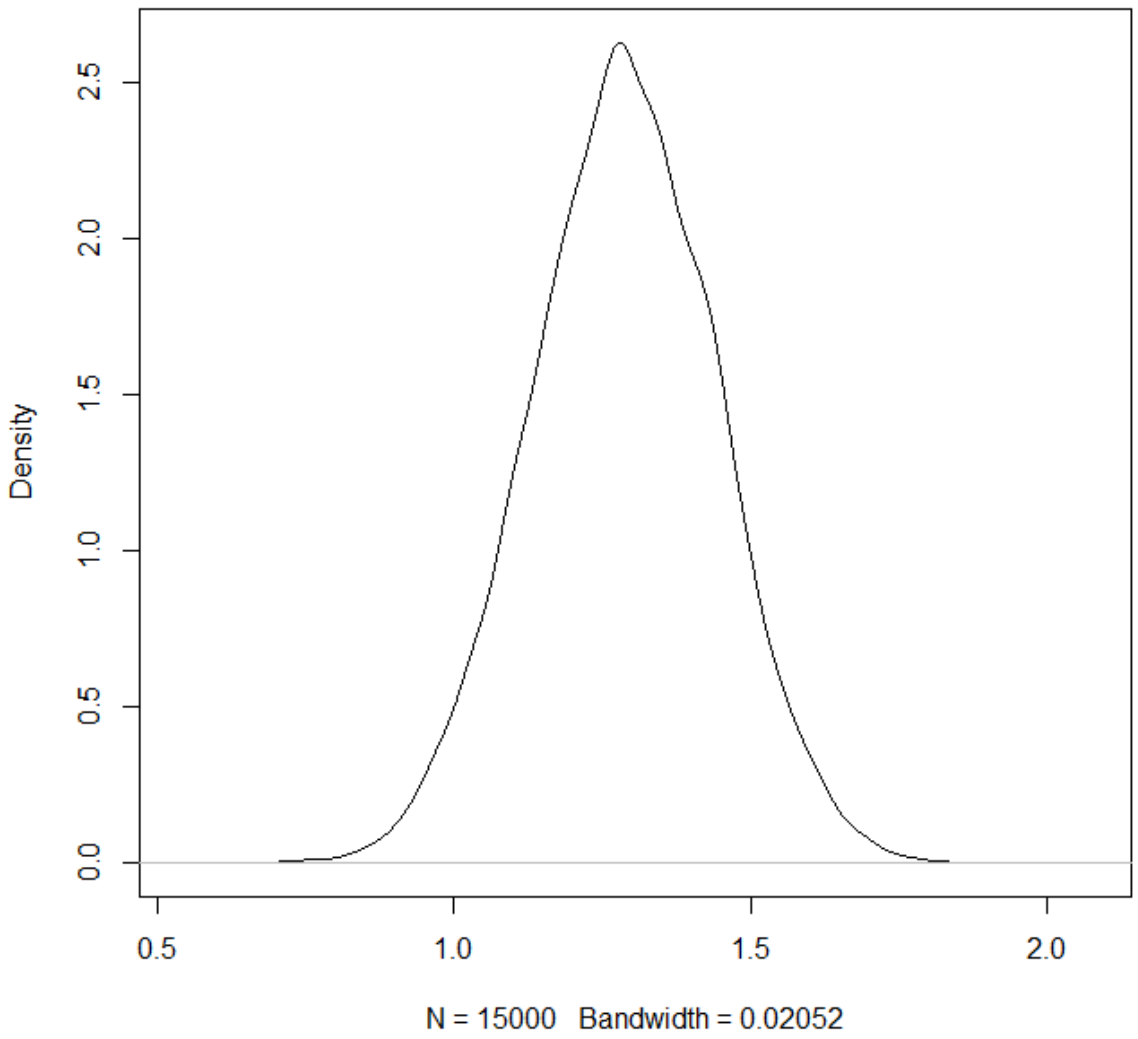


Figure S65 – Theoretical distribution of a very large cynognathian sample (n = 15,000 specimens), assuming a normal distribution whose mean is similar to that of non-mammal mammaliamorphs, and whose standard deviation is comparable to what is observed in the larger groups of this study (e.g., mammals).

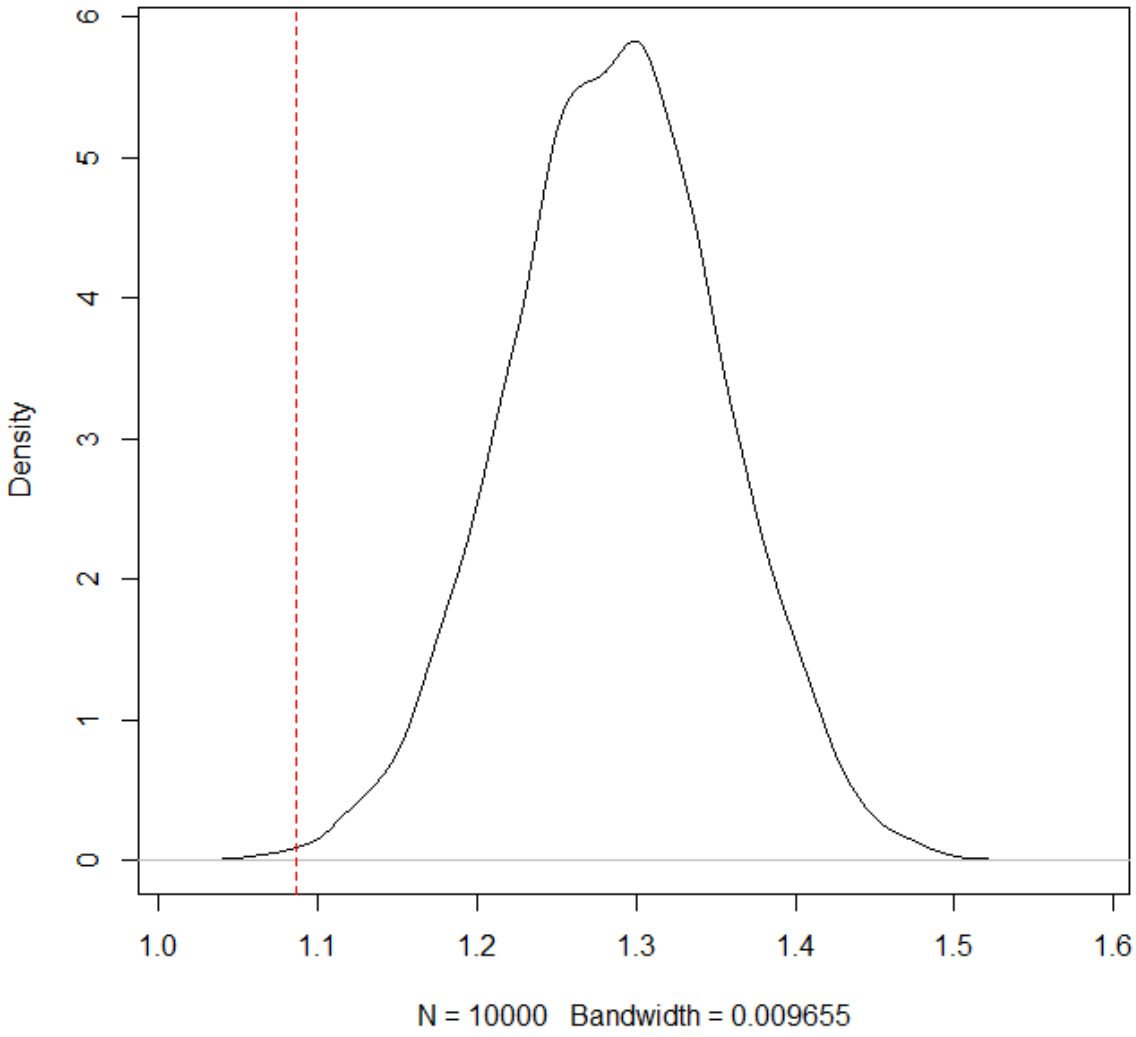


Figure S66 – Probability density function of the mean (size-corrected weighted average) Thermo-Motility Index, computed by using five randomly sampled cynognathians from a population (i.e., Fig. S65) whose mean is similar to that of non-mammal mammaliamorphs, and whose standard deviation is comparable to what is observed in the larger groups of this study (e.g., mammals). The red dashed line represents the mean of cynognathians we computed in our study. The area of the distribution located on the left of the red dashed line corresponds to the probability (P = 0.2%) of getting the mean we computed in our study, if the true mean of cynognathians was similar to that of non-mammal mammaliamorphs. This extremely low probability suggests that the true mean of cynognathians was likely lower than that of non-mammal mammaliamorphs. Indeed, if the true mean of cynognathians was closer to the one we computed, both the area on the left of the red dashed line and the probability would increase.

27. PGLS-R of maximum anaerobic and anaerobic speeds against body mass

Models 76 and 77 provide very significant regressions (P << 0.001), with coefficients of determination R^2 ranging from 0.44-0.45. They were tested on specimens for which we know the body mass, and were used to size-correct maximum aerobic (max_speed_aero) and anaerobic (max_speed) speeds. Body mass (BM) was used as a predictor for maximum aerobic and anaerobic speeds, with different slopes and intercepts depending on the size category (i.e., < 50 kg = small, > 50 kg = large), following observations made in Hirt et al. (2017) that maximum speed seems to peak at about 50 kg before decreasing. In practice, size corrections were done by combining fitted maximum speeds obtained for a body mass (BM) of 1g with residuals obtained for each species, for each model. The size-corrected maximum speed (max_speed_adjusted) is used in Model 78 to assess whether endotherms show higher maximum speeds than ectotherms, in Model 79 to test if maximum speed correlates with average body temperature for clades, and in Model 80 to check the relationship between maximum aerobic and anaerobic speeds.

Model 76: $\log_{10}$(max_speed) ~ $\log_{10}$(BM)*Size

Branch length transformations:

Pagel's λ [ML]: 0.999

lower bound: 0.000, $P = < 2.2\ 10^{-16}$ *

upper bound: 1.000, $P = 0.1703$

95.0% CI: (0.987, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.427579	0.316030	4.5172	$9.470\ 10^{-06}$*
$\log_{10}$(BM)	-0.019651	0.029224	-0.6724	0.5019
Size = Small	-0.614899	0.065513	-9.3859	$< 2.2\ 10^{-16}$ *
$\log_{10}$(BM): Size = Small	0.211014	0.030775	6.8566	$5.014\ 10^{-11}$ *

P values calculated using the exact method.

Adjusted R^2: 0.4428, AICc: -48.05779

F statistic: 71.45 on 3 and 263 DF

n = 267 (Supplementary Data 1, maximum anaerobic speed set)

Potential outliers (*Spermophilus citellus*, $G = 6.68677$, $U = 0.89504$, $P = 1.384\ 10^{-09}$ *, *Terrapene ornata*, $G = 5.24349$, $U = 0.93531$, $P = 2.117\ 10^{-05}$ *, *Hemidactylus turcicus*, $G = 5.25761$, $U = 0.93481$, $P = 1.944\ 10^{-05}$ *, *Nymphicus hollandicus*, $G = 4.49572$, $U = 0.95222$, $P = 0.00115$ *, *Litoria caerulea*, $G = 4.16122$, $U = 0.95897$, $P = 0.005599$ *, *Litoria dahlii*, $G = 5.4902$, $U = 0.9284$, $P = 4.826\ 10^{-06}$ *, *Microhyla fissipes*, $G = 4.72754$, $U = 0.94679$, $P = 0.0003529$ *, *Chiromantis doriae*, $G = 4.36739$, $U = 0.95448$, $P = 0.002119$ *, *Eulamprus kosciuskoi*, $G = 4.39220$, $U = 0.95385$, $P = 0.001876$ *, *Trachylepis striata*, $G = 4.07942$, $U = 0.96009$, $P = 0.007987$ *, *Trachylepis variegata*, $G = 3.90376$, $U = 0.96337$, $P = 0.01718$ *, *Ctenotus regius*, $G = 4.38884$, $U = 0.95359$, $P = 0.00189$ *, *Eulamprus tympanum*, $G = 4.3732$, $U = 0.9538$, $P = 0.002031$ *, *Somateria mollissima*, $G = 4.87500$, $U = 0.94246$, $P = 0.0001585$ *, *Anas acuta*, $G = 4.78643$, $U = 0.94439$, $P = 0.0002529$ *, *Petrosaurus mearnsi*, $G = 3.73649$, $U = 0.96603$, $P = 0.03418$ *, *Sceloporus merriami*, $G = 3.78083$, $U = 0.96514$, $P = 0.02839$ *, *Uta stansburiana*, $G = 3.71300$, $U = 0.96629$, $P = 0.03743$ *, *Anser caerulescens*, $G = 3.93457$, $U = 0.96206$, $P = 0.01472$ *, *Anas platyrhynchos*, $G = 3.9791$, $U = 0.9611$, $P = 0.0121$ *, *Phrynosoma modestum*, $G = 4.36581$, $U = 0.95305$, $P = 0.002054$ *, *Phrynosoma cornutum*, $G = 4.3023$, $U = 0.9543$, $P = 0.00277$ *, *Ctenotus taeniolatus*, $G = 4.48379$, $U = 0.95024$, $P = 0.001151$ *, *Phrynosoma platyrhinos*, $G = 4.09798$, $U = 0.95833$, $P = 0.007061$ *, *Falco peregrinus*, $G = 4.39775$, $U = 0.95189$, $P = 0.00174$ *, *Uma scoparia*, $G = 4.28036$, $U = 0.95431$, $P = 0.003035$ *, *Eulamprus quoyii*, $G = 4.35370$, $U = 0.95261$, $P = 0.002137$ *, *Cophosaurus texanus*, $G = 4.4712$, $U = 0.9499$, $P = 0.001205$ *, *Uma rufopunctata*, $G = 4.30038$, $U = 0.95353$, $P = 0.002737$ *, *Anolis pulchellus*, $G = 3.7381$, $U = 0.9648$, $P = 0.03266$ *, *Carlia fusca*, $G = 4.44265$, $U = 0.95016$, $P = 0.001373$ *, *Nanorana yunnanensis*, $G = 4.34785$, $U = 0.95214$, $P = 0.002164$ *, *Diceros bicornis*, $G = 3.67836$, $U = 0.96566$, $P = 0.04133$ *, *Phylloscopus trochilus*, $G = 4.54347$, $U = 0.94747$, $P = 0.0008266$ *, *Urocitellus undulatus*, $G = 4.27487$, $U = 0.95338$, $P = 0.003029$ *, *Xenopus laevis*, $G = 4.31290$, $U = 0.95243$, $P = 0.002524$ *, *Urosaurus nigricaudus*, $G = 4.23331$, $U = 0.95405$, $P = 0.003656$ *, *Chiasmocleis bassleri*, $G = 4.30575$, $U = 0.95234$, $P = 0.002594$ *, *Urosaurus ornatus*, $G = 4.19888$, $U = 0.95456$, $P = 0.004262$ *, *Callisaurus draconoides*, $G = 4.23733$, $U = 0.95361$, $P = 0.003555$ *, *Leptodactylus rhodomystax*, $G = 4.76534$, $U = 0.94117$, $P = 0.0002588$ *, *Allobates femoralis*, $G = 4.74784$, $U = 0.94145$, $P = 0.0002827$ *, *Ameerega trivittata*, $G = 4.79549$, $U = 0.94011$, $P = 0.0002192$ *, *Duttaphrynus melanostictus*, $G = 4.74117$, $U = 0.94131$, $P = 0.0002907$ *, *Rhinella margaritifera*, $G = 4.69763$, $U = 0.94223$, $P = 0.0003634$ *, *Dasyuroides byrnei*, $G = 3.63523$, $U = 0.96532$, $P = 0.04738$ *, *Isoodon obesulus*, $G = 4.3577$, $U = 0.9499$, $P = 0.001961$ *, *Antechinus flavipes*, $G = 4.37242$, $U = 0.94942$, $P = 0.001821$ *, *Bathytoshia lata*, $G = 4.1445$, $U = 0.9542$, $P = 0.00525$ *, *Anguilla anguilla*, $G = 4.25189$, $U = 0.95166$, $P = 0.003185$ *, *Cyclorana australis*, $G = 4.48245$, $U = 0.94613$, $P = 0.001047$ *, *Dendropsophus triangulum*, $G = 4.44938$, $U = 0.94678$, $P = 0.001228$ *, *Sminthopsis macroura*, $G = 3.75239$, $U = 0.96205$, $P = 0.02853$ *, *Cnemidophorus sexlineatus*, $G = 4.41395$, $U = 0.94734$, $P = 0.00145$

*, *Aspidoscelis inornata*, G = 4.46751, U = 0.94591, P = 0.001111 *, *Antechinomys laniger*, G = 4.6479, U = 0.9413, P = 0.0004469 *, *Antechinus stuartii*, G = 4.52295, U = 0.94426, P = 0.0008383 *, *Squatina californica*, G = 4.18472, U = 0.95215, P = 0.004239 *, *Gallotia atlantica*, G = 4.34024, U = 0.94839, P = 0.002034 *, *Gallotia stehlini*, G = 4.32292, U = 0.94866, P = 0.002202 *, *Hypsiboas punctatus*, G = 4.35507, U = 0.94775, P = 0.001881 *, *Gallotia simonyi*, G = 4.32542, U = 0.94832, P = 0.002161 *, *Aspidoscelis tigris*, G = 5.01222, U = 0.93041, P = $6.186\ 10^{-05}$ *, *Crocodylus niloticus*, G = 4.29703, U = 0.94871, P = 0.002456 *, *Cyclorana longipes*, G = 4.5908, U = 0.9413, P = 0.0005788 *, *Litoria inermis*, G = 4.14993, U = 0.95189, P = 0.004842 *, *Litoria pallida*, G = 4.23380, U = 0.94979, P = 0.003274 *, *Litoria nasuta*, G = 4.12265, U = 0.95226, P = 0.005447 *, *Dermochelys coriacea*, G = 4.13972, U = 0.95173, P = 0.005022 *, *Trichechus manatus*, G = 4.05346, U = 0.95359, P = 0.007392 *, *Glyphoglossus yunnanensis*, G = 4.50944, U = 0.94239, P = 0.0008524 *, *Rhacophorus dugritei*, G = 4.0547, U = 0.9533, P = 0.007303 *, *Eremitalpa namibensis*, G = 4.45871, U = 0.94336, P = 0.001088 *, *Plestiodon skiltonianus*, G = 4.76441, U = 0.93514, P = 0.0002273 *, *Bradypus tridactylus*, G = 4.17796, U = 0.94999, P = 0.004126 *, *Pseudemoia entrecasteauxii*, G = 4.37763, U = 0.94493, P = 0.0016 *, *Cercartetus concinnus*, G = 3.87778, U = 0.95667, P = 0.01558 *, *Osteocephalus planiceps*, G = 4.64144, U = 0.93774, P = 0.0004251 *, *Gorilla gorilla*, G = 4.26226, U = 0.94734, P = 0.002745 *, *Hyla annectans*, G = 4.57366, U = 0.93919, P = 0.0005964 *, *Presbytis*, G = 3.96889, U = 0.95408, P = 0.01036 *, *Varanus giganteus*, G = 4.96089, U = 0.92804, P = $7.595\ 10^{-05}$ *, *Dendropsophus rhodopeplus*, G = 4.61148, U = 0.93764, P = 0.0004861 *, *Homo sapiens*, G = 4.37315, U = 0.94375, P = 0.001587 *, *Dendropsophus sarayacuensis*, G = 4.58525, U = 0.93798, P = 0.0005513 *, *Sylvilagus floridanus*, G = 4.24486, U = 0.94669, P = 0.002904 *, *Macroscelides proboscideus*, G = 3.9115, U = 0.9546, P = 0.01303 *, *Oryctolagus cuniculus*, G = 4.34084, U = 0.94392, P = 0.001827 *, *Casuarius casuarius*, G = 4.66712, U = 0.93498, P = 0.0003562 *, *Elephas maximus*, G = 4.70540, U = 0.93371, P = 0.0002905 *, *Loxodonta africana*, G = 4.62319, U = 0.93581, P = 0.0004432 *, *Liopholis whitii*, G = 4.87499, U = 0.92842, P = 0.0001165 *, *Trachylepis occidentalis*, G = 4.9054, U = 0.9273, P = $9.822\ 10^{-05}$ *, *Litoria bicolor*, G = 4.4573, U = 0.9398, P = 0.001011 *, *Monodelphis brevicaudata*, G = 4.08309, U = 0.94933, P = 0.005922 *, *Varanus mertensi*, G = 4.83913, U = 0.92861, P = 0.000139 *, *Lepus europaeus*, G = 4.49799, U = 0.93813, P = 0.0008158 *, *Varanus storri*, G = 4.27003, U = 0.94407, P = 0.002465 *, *Lepus americanus*, G = 3.93036, U = 0.95247, P = 0.0115 *, *Numenius phaeopus*, G = 4.47453, U = 0.93821, P = 0.0009059 *, *Sceloporus occidentalis*, G = 3.90120, U = 0.95288, P = 0.01295 *, *Morus bassanus*, G = 4.45231, U = 0.93844, P = 0.001003 *, *Diomedea exulans*, G = 4.43562, U = 0.93871, P = 0.001085 *, *Alburnus alburnus*, G = 3.94650, U = 0.95133, P = 0.01052 *, *Lepus arcticus*, G = 4.31357, U = 0.94167, P = 0.001948 *, *Oncorhynchus mykiss*, G = 4.11868, U = 0.94666, P = 0.004832 *, *Hirundapus caudacutus*, G = 4.07378, U = 0.94765, P = 0.005903 *, *Geococcyx californianus*, G = 4.04646, U = 0.94818, P =

0.00665 *, *Gazella subgutturosa*, G = 3.67480, U = 0.95713, P = 0.03257 *, *Lepus townsendii*, G = 3.80972, U = 0.95378, P = 0.01853 *, *Notechis scutatus*, G = 3.6040, U = 0.9585, P = 0.04306 *, *Apus apus*, G = 4.84594, U = 0.92473, P = 0.0001245 *, *Ambystoma tigrinum nebulosum*, G = 4.0981, U = 0.9460, P = 0.005162 *, *Ambystoma maculatum*, G = 4.12345, U = 0.94515, P = 0.00458 *, *Squalius cephalus*, G = 3.94044, U = 0.94975, P = 0.01036 *, *Tamiasciurus hudsonicus*, G = 4.30980, U = 0.93969, P = 0.001896 *, *Laticauda colubrina*, G = 5.24154, U = 0.91051, P = $1.282\ 10^{-05}$ *, *Hamptophryne boliviana*, G = 4.1799, U = 0.9429, P = 0.003472 *, *Leptodactylus leptodactyloides*, G = 4.34278, U = 0.93817, P = 0.001597 *, *Mesocricetus brandti*, G = 4.06231, U = 0.94572, P = 0.005905 *, *Rhinella proboscidea*, G = 4.35260, U = 0.93727, P = 0.001503 *, *Leuciscus leuciscus*, G = 3.96274, U = 0.94783, P = 0.009098 *, *Sciurus niger*, G = 4.30747, U = 0.93815, P = 0.001853 *, *Sciurus vulgaris*, G = 4.32939, U = 0.93731, P = 0.00166 *, *Scardinius erythrophthalmus*, G = 3.95135, U = 0.94761, P = 0.009451 *, *Oreobates quixensis*, G = 4.39980, U = 0.93482, P = 0.001167 *, *Pedioplanis lineoocellata*, G = 4.9326, U = 0.9178, P = $7.137\ 10^{-05}$ *, *Tamias amoenus*, G = 4.16231, U = 0.94127, P = 0.003598 *, *Phacochoerus aethiopicus*, G = 3.83312, U = 0.95002, P = 0.01552 *, *Nucras tessellata*, G = 4.87070, U = 0.91903, P = $9.895\ 10^{-05}$ *, *Heliobolus lugubris*, G = 4.81422, U = 0.92063, P = 0.000134 *, *Zalophus californianus*, G = 3.61344, U = 0.95513, P = 0.03814 *, *Pedioplanis namaquensis*, G = 4.77703, U = 0.92131, P = 0.0001621 *, *Esox lucius*, G = 3.9027, U = 0.9473, P = 0.01126 *, *Takydromus septentrionalis*, G = 5.25641, U = 0.90406, P = $1.052\ 10^{-05}$ *, *Iberolacerta monticola*, G = 4.7132, U = 0.9226, P = 0.0002244 *, *Zootoca vivipara*, G = 4.6408, U = 0.9247, P = 0.0003268 *, *Tamias minimus*, G = 3.95774, U = 0.94504, P = 0.008703 *, *Archaeolacerta bedriagae*, G = 4.6457, U = 0.9240, P = 0.0003153 *, *Lacerta agilis*, G = 5.25043, U = 0.90259, P = $1.057\ 10^{-05}$ *, *Lacerta viridis*, G = 5.28423, U = 0.90098, P = $8.57\ 10^{-06}$ *, *Acanthodactylus erythrurus*, G = 5.28738, U = 0.90051, P = $8.355\ 10^{-06}$ *, *Lacerta schreiberi*, G = 5.22271, U = 0.90258, P = $1.224\ 10^{-05}$ *, *Podarcis tiliguerta*, G = 4.95393, U = 0.91204, P = $5.773\ 10^{-05}$ *, *Ammospermophilus leucurus*, G = 4.28507, U = 0.93395, P = 0.001867 *, *Tamias striatus*, G = 4.0599, U = 0.9405, P = 0.005317 *, *Sciurus carolinensis*, G = 4.18532, U = 0.93653, P = 0.002968 *, *Lepus californicus*, G = 4.41687, U = 0.92906, P = 0.0009668 *, *Gallotia caesaris*, G = 4.32955, U = 0.93159, P = 0.001477 *, *Psammodromus algirus*, G = 4.38666, U = 0.92926, P = 0.001106 *, *Rutilus rutilus*, G = 3.76740, U = 0.94763, P = 0.01862 *, *Marmota monax*, G = 4.30786, U = 0.93127, P = 0.001609 *, *Xerospermophilus tereticaudus*, G = 4.44395, U = 0.92631, P = 0.0008152 *, *Acanthodactylus scutellatus*, G = 5.75286, U = 0.87605, P = $3.865\ 10^{-07}$ *, *Acanthodactylus pardalis*, G = 5.92530, U = 0.86801, P = $1.174\ 10^{-07}$ *, P values computed via approximation method) were excluded from this regression. As Shapiro-Wilk tests (W = 0.98708, P = 0.001776 *, W = 0.99034, P = 0.01366 *, W = 0.99241, P = 0.0518 *, W = 0.98756, P = 0.01005 *, W = 0.98391, P = 0.003469 *, W = 0.98523, P = 0.006788 *, P value calculated via approximation method) indicated potential non-normality of the phylogenetic residuals after

removing outliers, the most influential specimens (*Myrmecobius fasciatus*, *Calypte anna*, *Trachylepis spilogaster*, *Scincella lateralis*, *Anolis gundlachi*, *Lepus alleni*) were excluded from this regression.

Note that size-corrected maximum anaerobic speeds obtained with Model 76 do strongly correlate to size-corrected maximum anaerobic speed obtained with a model derived from the complete dataset (n = 428, including potential outliers), with a coefficient of determination (R^2) of 0.80 (*P* value: $< 2.2\ 10^{-16}$ *).

P value: $< 2.2\ 10^{-16}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.99163$, $P = 0.1334$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S67) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 3.38438$, $U = 0.95678$, $P = 0.08403$, *P* value computed via approximation method).

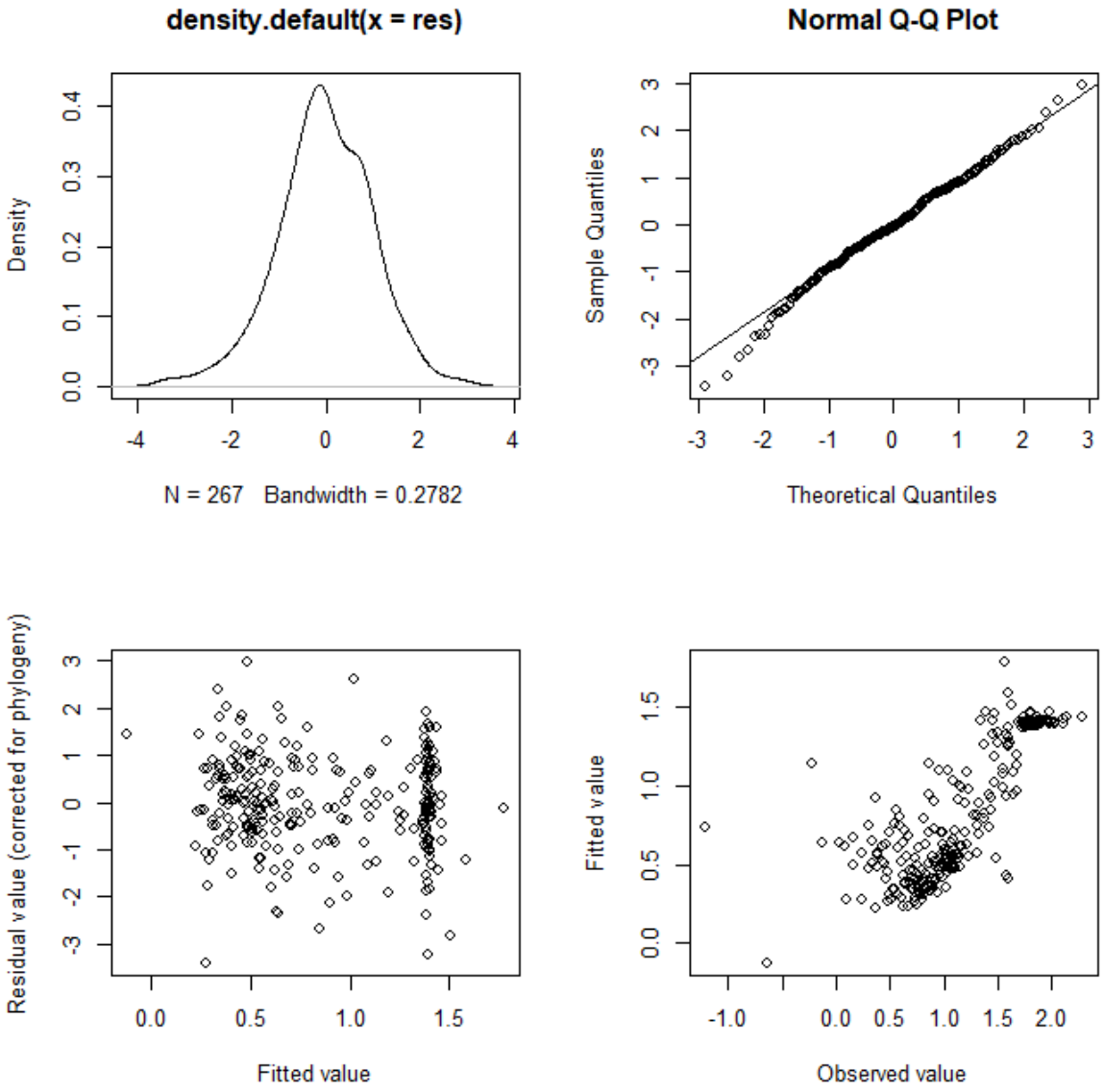


Figure S67 – Diagnostic plots for model 76.

Model 77: $\log_{10}$(max_speed_aero) ~ $\log_{10}$(BM)*Size

Branch length transformations:

Pagel's λ [ML]: 0.871

lower bound: 0.000, P = 6.7499 10^{-07} *

upper bound: 1.000, P = 0.045237

95.0% CI: (0.546, 0.999)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.612678	0.646281	0.9480	0.3505
$\log_{10}$(BM)	0.191314	0.287480	0.6655	0.5107
Size = Small	-0.018297	0.611517	-0.0299	0.9763
$\log_{10}$(BM): Size = Small	0.085100	0.291805	0.2916	0.7725

P values calculated using the exact method.

Adjusted R^2: 0.4456, AICc: 12.7454

F statistic: 10.11 on 3 and 31 DF

n = 35 (Supplementary Data 1, maximum aerobic speed set)

P value: 8.427 10^{-05} *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.99217, P = 0.9961, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S68) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.14447, U = 0.86076, P = 0.477, P value computed via approximation method).

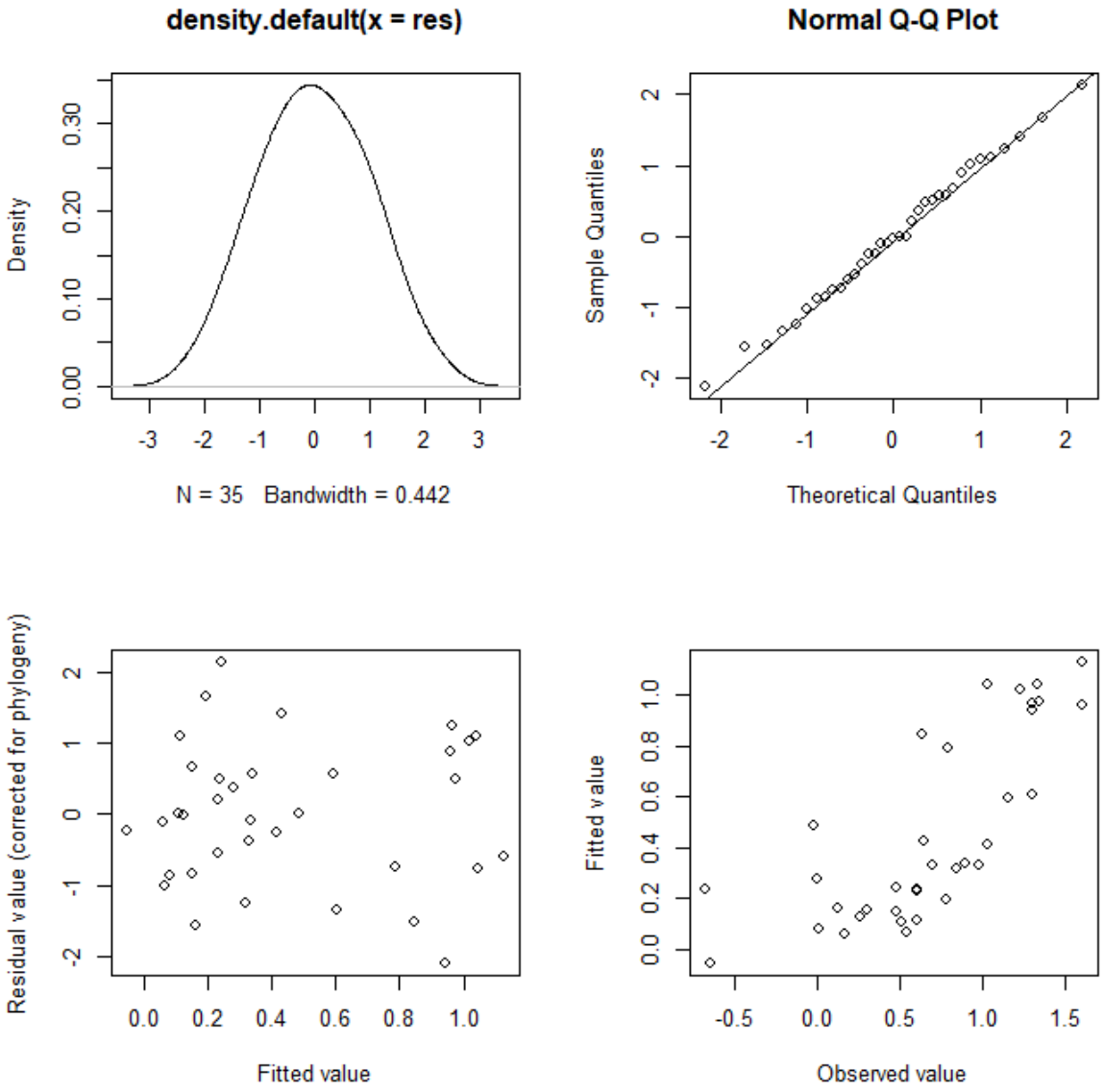


Figure S68 – Diagnostic plots for model 77.

28. PGLS ANOVA for comparing the size-corrected maximum anaerobic speed of endotherms *versus* ectotherms

Model 78 is significant ($P < 0.01$), with a very large effect size (Glass's $\Delta = -1.50$), but a low coefficient of determination ($R^2 = 0.05$). It was tested on specimens for which we know the thermoregulatory regime, and was used to check if the size-corrected maximum anaerobic speed (max_speed_adjusted) of endotherms (RT = endotherm) is significantly higher than that of ectotherms (RT = ectotherm). Results confirm that the size-corrected maximum anaerobic speed of endotherms is significantly higher than that of ectotherms.

Model 78: $\log_{10}$(max_speed_adjusted) ~ RT

Branch length transformations:

Pagel's λ [ML]: 0.970

lower bound: 0.000, $P = < 2.22\ 10^{-16}$ *

upper bound: 1.000, $P = 6.3699\ 10^{-06}$ *

95.0% CI: (0.937, 0.988)

P values approximated using a likelihood ratio test against the χ^2 distribution

Response: TMI

	Df	Sum Sq	Mean Sq	*F* statistic	Pr(>|t|)
RT	1	0.007475	0.0074752	9.079	0.002994 *
Residuals	165	0.135853	0.0008234		

P values calculated using the exact method.

Adjusted R^2: 0.04641, AICc: 4.011184

n = 167 (Supplementary Data 1, maximum anaerobic speed set size-corrected using Model 76)

Potential outliers (*Spermophilus citellus*, *G* = 5.57717, *U* = 0.92698, *P* = 2.885 10-06 *, Ter*rapene ornata*, *G* = 4.68599, *U* = 0.94833, *P* = 0.0004441 *, *Hemidactylus turcicus*, *G* = 4.1948, *U* = 0.9585, *P* = 0.004828 *, *Equus quagga*, *G* = 4.32132, *U* = 0.95585, *P* = 0.002669 *, *Callisaurus draconoides*, *G* = 4.46608, *U* = 0.95274, *P* = 0.001326 *, *Dendropsophus sarayacuensis*, *G* = 4.52041, *U* = 0.95146, *P* = 0.001012 *, *Holbrookia elegans*, *G* = 4.68296, *U* = 0.94779, *P* = 0.0004442 *, *Cophosaurus texanus*, *G* = 4.63467, *U* = 0.94873, *P* = 0.0005668 *, *Cyclorana australis*, *G* = 4.69900, *U* = 0.94718, *P* = 0.0004064 *, *Eulamprus kosciuskoi*, *G* = 4.83147, *U* = 0.94402, *P* = 0.0002025 *, *Osteocephalus planiceps*, *G* = 4.47256, *U* = 0.95191, *P* = 0.001262 *, *Urosaurus ornatus*, *G* = 4.46033, *U* = 0.95206, *P* = 0.001336 *, *Hyla annectans*, *G* = 4.5632, *U* = 0.9497, *P* = 0.0008004 *, *Uta stansburiana*, *G* = 4.44674, *U* = 0.95212,

P = 0.001419 *, *Urosaurus graciosus*, *G* = 4.46959, *U* = 0.95151, *P* = 0.001265 *, *Acinonyx jubatus*, *G* = 4.62884, *U* = 0.94787, *P* = 0.0005695 *, *Somateria mollissima*, *G* = 5.33909, *U* = 0.93047, *P* = 1.146 10-05 *, Tri*gla lyra*, *G* = 4.55734, *U* = 0.94922, *P* = 0.0008116 *, *Felis catus*, *G* = 4.32304, *U* = 0.95419, *P* = 0.002534 *, *Sphyraena barracuda*, *G* = 4.58373, *U* = 0.94838, *P* = 0.0007065 *, *Uma scoparia*, *G* = 4.18450, *U* = 0.95687, *P* = 0.004808 *, *Hippotragus niger*, *G* = 4.88313, *U* = 0.94112, *P* = 0.0001477 *, *Sminthopsis macroura*, *G* = 5.02577, *U* = 0.93748, *P* = 6.753 10-05 *, Hel*icolenus dactylopterus*, *G* = 4.58196, *U* = 0.94791, *P* = 0.0007038 *, *Antilope cervicapra*, *G* = 4.15364, *U* = 0.95708, *P* = 0.005472 *, *Eudorcas thomsonii*, *G* = 4.48875, *U* = 0.94975, *P* = 0.001113 *, *Ambystoma tigrinum nebulosum*, *G* = 4.08980, *U* = 0.95818, *P* = 0.007261 *, *Nanger granti*, *G* = 4.98433, *U* = 0.93774, *P* = 8.341 10-05 *, G*azella dorcas*, *G* = 4.06862, *U* = 0.95841, *P* = 0.007936 *, *Petrosaurus mearnsi*, *G* = 4.82537, *U* = 0.94135, *P* = 0.0001957 *, *Sceloporus merriami*, *G* = 4.5465, *U* = 0.9478, *P* = 0.0008222 *, *Urosaurus nigricaudus*, *G* = 4.89735, *U* = 0.93928, *P* = 0.0001322 *, *Casuarius casuarius*, *G* = 3.9946, *U* = 0.9595, *P* = 0.01089 *, *Uma rufopunctata*, *G* = 4.88310, *U* = 0.93933, *P* = 0.0001417 *, *Dromaius novaehollandiae*, *G* = 4.02471, *U* = 0.95868, *P* = 0.009477 *, *Sceloporus graciosus*, *G* = 4.54616, *U* = 0.94714, *P* = 0.0008102 *, *Phrynosoma mcallii*, *G* = 4.88652, *U* = 0.93877, *P* = 0.0001377 *, *Phrynosoma platyrhinos*, *G* = 4.94437, *U* = 0.93716, *P* = 0.0001002 *, *Phrynosoma modestum*, *G* = 4.85222, *U* = 0.93932, *P* = 0.0001644 *, *Phrynosoma cornutum*, *G* = 4.91861, *U* = 0.93749, *P* = 0.0001145 *, *Leptodactylus leptodactyloides*, *G* = 4.53378, *U* = 0.94675, *P* = 0.0008478 *, *Sceloporus jarrovii*, *G* = 4.75523, *U* = 0.94127, *P* = 0.0002719 *, *Sceloporus magister*, *G* = 4.64087, *U* = 0.94391, *P* = 0.0004901 *, *Anolis pulchellus*, *G* = 4.63163, *U* = 0.94399, *P* = 0.0005121 *, *Crocodylus johnstoni*, *G* = 3.78210, *U* = 0.96255, *P* = 0.02607 *, *Sceloporus clarkii*, *G* = 4.60543, *U* = 0.94433, *P* = 0.0005811 *, *Leptodactylus rhodomystax*, *G* = 4.69249, *U* = 0.94205, *P* = 0.0003706 *, *Anolis krugi*, *G* = 4.51034, *U* = 0.94632, *P* = 0.0009306 *, *Anolis gundlachi*, *G* = 4.53624, *U* = 0.94556, *P* = 0.0008151 *, *Anolis poncensis*, *G* = 4.4235, *U* = 0.9481, *P* = 0.001417 *, *Heliobolus lugubris*, *G* = 4.5806, *U* = 0.9442, *P* = 0.0006478 *, *Anolis cristatellus*, *G* = 4.32226, *U* = 0.95018, *P* = 0.002292 *, *Anolis stratulus*, *G* = 4.56698, *U* = 0.94423, *P* = 0.0006889 *, *Sceloporus occidentalis*, *G* = 4.64138, *U* = 0.94225, *P* = 0.0004704 *, *Anolis evermanni*, *G* = 4.66716, *U* = 0.94145, *P* = 0.0004105 *, *Anolis lemurinus*, *G* = 4.35545, *U* = 0.94887, *P* = 0.00193 *, *Anolis limifrons*, *G* = 4.3224, *U* = 0.9495, *P* = 0.002253 *, *Anolis humilis*, *G* = 4.27102, *U* = 0.95057, *P* = 0.002864 *, *Sceloporus virgatus*, *G* = 4.43246, *U* = 0.94661, *P* = 0.001316 *, *Coleonyx variegatus*, *G* = 4.38356, *U* = 0.94764, *P* = 0.001663 *, *Anolis lineatopus*, *G* = 4.49011, *U* = 0.94492, *P* = 0.0009836 *, *Micropterus salmoides*, *G* = 4.47144, *U* = 0.94522, *P* = 0.001075 *, *Lepidodactylus lugubris*, *G* = 4.38877, *U* = 0.94708, *P* = 0.001605 *, *Gekko gecko*, *G* = 4.43123, *U* = 0.94591, *P* = 0.001301 *, *Ctenotus uber*, *G* = 4.41553, *U* = 0.94614, *P* = 0.001399 *, *Sceloporus cowlesi*, *G* = 4.67839, *U* = 0.93937, *P* = 0.0003721 *, *Sceloporus woodi*, *G* = 4.63351, *U* = 0.94036, *P* = 0.0004671 *, *Gonatodes concinnatus*, *G*

= 4.38586, *U* = 0.94642, *P* = 0.001599 *, *Ornithorhynchus anatinus*, *G* = 4.50497, *U* = 0.94331, *P* = 0.0008877 *, *Gambelia wislizenii*, *G* = 4.06378, *U* = 0.95374, *P* = 0.007129 *, *Sceloporus undulatus*, *G* = 4.56230, *U* = 0.94153, *P* = 0.0006611 *, *Tachyglossus aculeatus*, *G* = 4.52103, *U* = 0.94242, *P* = 0.0008104 *, *Monodelphis brevicaudata*, *G* = 4.68150, *U* = 0.93809, *P* = 0.0003565 *, *Anolis carolinensis*, *G* = 4.52185, *U* = 0.94208, *P* = 0.0008011 *, *Didelphis marsupialis*, *G* = 4.53531, *U* = 0.94157, *P* = 0.0007461 *, *Cercartetus concinnus*, *G* = 4.68149, *U* = 0.93756, *P* = 0.0003524 *, *Macropus eugenii*, *G* = 4.31155, *U* = 0.94689, *P* = 0.002213 *, *Bettongia penicillata*, *G* = 4.32873, *U* = 0.94631, *P* = 0.002031 *, *Anolis sagrei*, *G* = 4.52595, *U* = 0.94114, *P* = 0.0007703 *, *Crotaphytus collaris*, *G* = 4.3246, *U* = 0.9461, *P* = 0.002057 *, *Anolis frenatus*, *G* = 4.65589, *U* = 0.93735, *P* = 0.0003946 *, *Podarcis siculus*, *G* = 4.58619, *U* = 0.93903, *P* = 0.0005618 *, *Potorous tridactylus*, *G* = 4.47302, *U* = 0.94184, *P* = 0.0009876 *, *Trichechus manatus*, *G* = 4.59582, *U* = 0.93842, *P* = 0.0005307 *, *Amblyrhynchus cristatus*, *G* = 4.49370, *U* = 0.94096, *P* = 0.0008844 *, *Dipsosaurus dorsalis*, *G* = 4.4587, *U* = 0.9417, *P* = 0.001048 *, *Axis axis*, *G* = 5.13544, *U* = 0.92243, *P* = 2.812 10-05 *, Tap*irus terrestris*, *G* = 4.72768, *U* = 0.93407, *P* = 0.0002639 *, *Trapelus savignii*, *G* = 4.58680, *U* = 0.93776, *P* = 0.0005448 *, *Stellagama stellio*, *G* = 4.72542, *U* = 0.93374, *P* = 0.0002649 *, *Macropus rufus*, *G* = 4.77553, *U* = 0.93213, *P* = 0.0002024 *, *Isoodon obesulus*, *G* = 4.29510, *U* = 0.94493, *P* = 0.002265 *, *Myrmecobius fasciatus*, *G* = 4.47404, *U* = 0.94007, *P* = 0.0009453 *, *Ctenophorus nuchalis*, *G* = 4.53933, *U* = 0.93812, *P* = 0.0006792 *, *Dasyuroides byrnei*, *G* = 4.71065, *U* = 0.93316, *P* = 0.0002803 *, *Leiolepis belliana*, *G* = 4.41110, *U* = 0.94122, *P* = 0.001274 *, *Antechinus flavipes*, *G* = 4.46699, *U* = 0.93953, *P* = 0.0009636 *, *Varanus storri*, *G* = 4.43661, *U* = 0.94017, *P* = 0.001115 *, *Varanus acanthurus*, *G* = 4.66943, *U* = 0.93353, *P* = 0.0003418 *, *Balaenoptera musculus*, *G* = 5.60546, *U* = 0.90391, *P* = 1.528 10-06 *, Var*anus caudolineatus*, *G* = 4.56478, *U* = 0.93608, *P* = 0.0005803 *, *Varanus gilleni*, *G* = 4.33273, *U* = 0.94224, *P* = 0.00182 *, *Varanus eremius*, *G* = 4.35926, *U* = 0.94135, *P* = 0.001595 *, *Cyclorana longipes*, *G* = 4.36209, *U* = 0.94109, *P* = 0.001567 *, *Oryctolagus cuniculus*, *G* = 4.53129, *U* = 0.93623, *P* = 0.0006761 *, *Litoria caerulea*, *G* = 4.37907, *U* = 0.94026, *P* = 0.001431 *, *Sylvilagus floridanus*, *G* = 4.53592, *U* = 0.93571, *P* = 0.000655 *, *Varanus glauerti*, *G* = 4.49226, *U* = 0.93674, *P* = 0.0008122 *, *Varanus tristis*, *G* = 4.6261, *U* = 0.9327, *P* = 0.0004099 *, *Dendropsophus triangulum*, *G* = 4.49346, *U* = 0.93631, *P* = 0.0008006 *, *Litoria dahlii*, *G* = 4.35689, *U* = 0.93993, *P* = 0.001562 *, *Varanus scalaris*, *G* = 4.44104, *U* = 0.93739, *P* = 0.00103 *, *Varanus mitchelli*, *G* = 4.59536, *U* = 0.93275, *P* = 0.0004716 *, *Litoria bicolor*, *G* = 4.41199, *U* = 0.93761, *P* = 0.001174 *, *Varanus panoptes*, *G* = 4.48960, *U* = 0.93519, *P* = 0.0007954 *, *Dendropsophus rhodopeplus*, *G* = 4.44261, *U* = 0.93633, *P* = 0.001001 *, *Varanus pilbarensis*, *G* = 4.53668, *U* = 0.93339, *P* = 0.0006221 *, *Varanus brevicauda*, *G* = 4.30322, *U* = 0.93988, *P* = 0.001956 *, *Varanus gouldii*, *G* = 4.50976, *U* = 0.93354, *P* = 0.0007033 *, *Carlia fusca*, *G* = 4.47406, *U* = 0.93437, *P* = 0.0008374 *, *Chiromantis doriae*, *G* = 4.3866, *U* = 0.9367, *P* = 0.001285 *, *Rhacophorus*

dugritei, *G* = 5.43555, *U* = 0.90249, *P* = 3.863 10-06 *, Var*anus kingorum*, *G* = 4.2955, *U* = 0.9389, *P* = 0.001979 *, *Varanus rosenbergi*, *G* = 4.61661, *U* = 0.92919, *P* = 0.0003986 *, *Varanus giganteus*, *G* = 4.31454, *U* = 0.93795, *P* = 0.001791 *, *Varanus mertensi*, *G* = 4.36811, *U* = 0.93619, *P* = 0.001375 *, *Varanus varius*, *G* = 4.41336, *U* = 0.93464, *P* = 0.001097 *, *Taeniura lymma*, *G* = 4.35247, *U* = 0.93622, *P* = 0.001471 *, *Bathytoshia lata*, *G* = 4.34599, *U* = 0.93619, *P* = 0.001511 *, *Anguilla anguilla*, *G* = 4.33538, *U* = 0.93629, *P* = 0.001584 *, *Dipodomys microps*, *G* = 4.57731, *U* = 0.92874, *P* = 0.0004721 *, *Alosa fallax*, *G* = 4.11735, *U* = 0.94214, *P* = 0.00439 *, *Cyprinus carpio*, *G* = 4.10254, *U* = 0.94236, *P* = 0.004678 *, *Carassius auratus*, *G* = 4.36831, *U* = 0.93443, *P* = 0.001325 *, *Duttaphrynus melanostictus*, *G* = 4.63527, *U* = 0.92591, *P* = 0.0003432 *, *Rhinella margaritifera*, *G* = 3.95335, *U* = 0.94592, *P* = 0.009025 *, *Platichthys flesus*, *G* = 3.78294, *U* = 0.95031, *P* = 0.01874 *, *Gallotia atlantica*, *G* = 4.67069, *U* = 0.92399, *P* = 0.0002809 *, *Gallotia stehlini*, *G* = 4.95294, *U* = 0.91423, *P* = 6.038 10-05 *, Psa*mmodromus algirus*, *G* = 5.02160, *U* = 0.91152, *P* = 4.074 10-05 *, Psa*mmodromus hispanicus*, *G* = 5.15377, *U* = 0.90648, *P* = 1.887 10-05 *, Ple*uronectes platessa*, *G* = 4.34181, *U* = 0.93339, *P* = 0.001453 *, *Leuciscus leuciscus*, *G* = 5.26670, *U* = 0.90164, *P* = 9.525 10-06 *, Alb*urnus alburnus*, *G* = 4.13378, *U* = 0.93919, *P* = 0.00386 *, *Hypsiboas punctatus*, *G* = 4.54815, *U* = 0.92612, *P* = 0.0005117 *, *Hypsiboas hobbsi*, *G* = 3.87707, *U* = 0.94612, *P* = 0.01207 *, *Hypsiboas lanciformis*, *G* = 3.96205, *U* = 0.94353, *P* = 0.008282 *, *Zalophus californianus*, *G* = 3.92568, *U* = 0.94437, *P* = 0.009679 *, *Scinax ruber*, *G* = 4.05972, *U* = 0.94029, *P* = 0.005298 *, *Sphaenorhynchus lacteus*, *G* = 3.94836, *U* = 0.94331, *P* = 0.008683 *, *Litoria rothii*, *G* = 3.92095, *U* = 0.94389, *P* = 0.009751 *, *Scardinius erythrophthalmus*, *G* = 5.17676, *U* = 0.90184, *P* = 1.539 10-05 *, Eso*x lucius*, *G* = 5.16488, *U* = 0.90193, *P* = 1.641 10-05 *, Onc*orhynchus mykiss*, *G* = 4.23357, *U* = 0.93386, *P* = 0.002308 *, *Salmo salar*, *G* = 4.29285, *U* = 0.93175, *P* = 0.00173 *, *Lycaon pictus*, *G* = 4.79227, *U* = 0.91463, *P* = 0.0001332 *, *Podarcis tiliguerta*, *G* = 4.44338, *U* = 0.92633, *P* = 0.0008176 *, *Lepus alleni*, *G* = 4.29989, *U* = 0.93075, *P* = 0.001647 *, *Squalius cephalus*, *G* = 4.87030, *U* = 0.91083, *P* = 8.531 10-05 *, Smin*thopsis crassicaudata*, *G* = 4.11872, *U* = 0.93599, *P* = 0.00384 *, *Antechinomys laniger*, *G* = 4.10894, *U* = 0.93605, *P* = 0.003997 *, *Salmo trutta*, *G* = 4.22583, *U* = 0.93184, *P* = 0.00229 *, *Pygoscelis papua*, *G* = 4.14824, *U* = 0.93407, *P* = 0.003283 *, *Istiompax indica*, *G* = 4.75376, *U* = 0.91308, *P* = 0.0001557 *, *Pygoscelis adeliae*, *G* = 4.13405, *U* = 0.93401, *P* = 0.003472 *, *Mugil cephalus*, *G* = 4.37659, *U* = 0.92576, *P* = 0.001081 *, *Necturus maculosa*, *G* = 4.60303, *U* = 0.91756, *P* = 0.0003405 *, *Podarcis muralis*, *G* = 4.36855, *U* = 0.92545, *P* = 0.001113 *, *Tamiasciurus hudsonicus*, *G* = 4.89274, *U* = 0.90612, *P* = 7.034 10-05 *, Amb*ystoma maculatum*, *G* = 4.48282, *U* = 0.92088, *P* = 0.0006213 *, *Nanorana yunnanensis*, *G* = 4.53699, *U* = 0.91864, *P* = 0.0004685 *, *Chiasmocleis bassleri*, *G* = 4.05995, *U* = 0.93459, *P* = 0.004714 *, *Hamptophryne boliviana*, *G* = 4.03841, *U* = 0.93503, *P* = 0.005174 *, *Glyphoglossus yunnanensis*, *G* = 4.14590, *U* = 0.93125, *P* = 0.003137 *, *Microhyla fissipes*, *G* = 4.03593, *U* = 0.93458, *P* = 0.005179 *, *Rhacophorus rhodopus*, *G* =

4.69413, *U* = 0.91115, *P* = 0.0001994 *, *Panthera pardus*, *G* = 4.3923, *U* = 0.9219, *P* = 0.000941 *, *Galeocerdo cuvier*, *G* = 4.26285, *U* = 0.92613, *P* = 0.001763 *, *Uperoleia lithomoda*, *G* = 5.11462, *U* = 0.89323, *P* = 1.835 10-05 *, Isu*rus oxyrinchus*, *G* = 4.17364, *U* = 0.92861, *P* = 0.002667 *, *Carcharodon carcharias*, *G* = 4.16416, *U* = 0.92864, *P* = 0.002774 *, *Merlangius merlangus*, *G* = 4.06895, *U* = 0.93159, *P* = 0.004294 *, *Eremiascincus fasciolatus*, *G* = 3.98620, *U* = 0.93407, *P* = 0.006219 *, *Loxodonta africana*, *G* = 3.70120, *U* = 0.94292, *P* = 0.02124 *, *Podarcis hispanicus*, *G* = 3.62086, *U* = 0.94514, *P* = 0.02943 *, *Melanogrammus aeglefinus*, *G* = 4.12558, *U* = 0.92849, *P* = 0.003233 *, *Ursus americanus*, *G* = 3.63902, *U* = 0.94413, *P* = 0.02706 *, *Canis familiaris*, *G* = 3.88844, *U* = 0.93593, *P* = 0.009347 *, *Canis lupus*, *G* = 4.00880, *U* = 0.93162, *P* = 0.005437 *, *Laticauda laticaudata*, *G* = 4.54215, *U* = 0.91183, *P* = 0.0004064 *, *Notechis scutatus*, *G* = 5.59038, *U* = 0.86587, *P* = 8.438 10-07 *, Den*droaspis polylepis*, *G* = 4.01413, *U* = 0.93055, *P* = 0.00522 *, *Podarcis pityusensis*, *G* = 4.04450, *U* = 0.92919, *P* = 0.00452 *, *Podarcis lilfordi*, *G* = 4.0582, *U* = 0.9284, *P* = 0.00422 *, *Podarcis bocagei*, *G* = 4.08474, *U* = 0.92714, *P* = 0.003713 *, *Laticauda frontalis*, *G* = 3.54367, *U* = 0.94492, *P* = 0.03803 *, *Pollachius virens*, *G* = 3.99034, *U* = 0.92986, *P* = 0.005652 *, *Laticauda colubrina*, *G* = 4.39599, *U* = 0.91449, *P* = 0.0008134 *, *Lacerta schreiberi*, *G* = 4.80316, *U* = 0.89747, *P* = 9.422 10-05 *, Lac*erta viridis*, *G* = 4.69523, *U* = 0.90159, *P* = 0.0001689 *, *Lacerta agilis*, *G* = 4.43362, *U* = 0.91185, *P* = 0.0006598 *, *Ursus maritimus*, *G* = 4.61075, *U* = 0.90424, *P* = 0.0002617 *, *Trisopterus luscus*, *G* = 4.02409, *U* = 0.92673, *P* = 0.004682 *, *Ursus arctos*, *G* = 4.44367, *U* = 0.91025, *P* = 0.0006145 *, *Hemiergis decresiensis*, *G* = 4.32598, *U* = 0.91455, *P* = 0.001101 *, *Hemiergis peronii*, *G* = 4.03475, *U* = 0.92533, *P* = 0.00438 *, *Canis latrans*, *G* = 4.41798, *U* = 0.91005, *P* = 0.0006859 *, *Scincella lateralis*, *G* = 4.25344, *U* = 0.91624, *P* = 0.001541 *, *Eulamprus quoyii*, *G* = 4.11134, *U* = 0.92138, *P* = 0.003015 *, *Trachylepis spilogaster*, *G* = 3.8314, *U* = 0.9314, *P* = 0.0106 *, *Argyrosomus regius*, *G* = 3.66543, *U* = 0.93692, *P* = 0.02134 *, *Gadus morhua*, *G* = 3.86376, *U* = 0.92958, *P* = 0.009095 *, *Marmota monax*, *G* = 3.54760, *U* = 0.94035, *P* = 0.03415 *, *Perca fluviatilis*, *G* = 3.58836, *U* = 0.93869, *P* = 0.0288 *, *Cottus gobio*, *G* = 3.74745, *U* = 0.93281, *P* = 0.01477 *, *Nasua nasua*, *G* = 4.51681, *U* = 0.90192, *P* = 0.0003871 *, *Nasua narica*, *G* = 3.96516, *U* = 0.92405, *P* = 0.005614 *, *Pantherophis guttatus*, *G* = 4.71409, *U* = 0.89213, *P* = 0.0001328 *, *Boiga irregularis*, *G* = 4.78644, *U* = 0.88825, *P* = 8.825 10-05 *, Tha*mnophis elegans*, *G* = 4.50626, *U* = 0.90046, *P* = 0.0003968 *, *Coluber constrictor*, *G* = 4.47978, *U* = 0.90114, *P* = 0.0004518 *, *Crotalus cerastes*, *G* = 3.60271, *U* = 0.93575, *P* = 0.02591 *, *Boa constrictor*, *G* = 4.88742, *U* = 0.88116, *P* = 4.812 10-05 *, Elg*aria kingii*, *G* = 4.65250, *U* = 0.89177, *P* = 0.0001767 *, *Gallotia caesaris*, *G* = 4.5798, *U* = 0.8946, *P* = 0.0002589 *, *Enhydra lutris*, *G* = 4.75600, *U* = 0.88576, *P* = 9.843 10-05 *, Mel*es meles*, *G* = 4.70384, *U* = 0.88769, *P* = 0.0001302 *, *Elephas maximus*, *G* = 4.08831, *U* = 0.91473, *P* = 0.002958 *, *Procyon lotor*, *G* = 4.38865, *U* = 0.90123, *P* = 0.0006774 *, *Zootoca vivipara*, *G* = 4.56073, *U* = 0.89279, *P* = 0.0002749 *, *Iberolacerta monticola*, *G* = 4.31700, *U* = 0.90344, *P* =

0.0009566 *, *Antechinus stuartii*, G = 4.62334, U = 0.88867, P = 0.0001933 *, *Ictidomys tridecemlineatus*, G = 4.68014, U = 0.88532, P = 0.0001406 *, *Xerospermophilus tereticaudus*, G = 4.35141, U = 0.90035, P = 0.000786 *, *Archaeolacerta bedriagae*, G = 5.24274, U = 0.85457, P = 5.008 10-06 *, Ped*ioplanis namaquensis*, G = 5.20892, U = 0.85568, P = 6.123 10-06 *, Tak*ydromus septentrionalis*, G = 5.02017, U = 0.86523, P = 1.925 10-05 *, Emy*docephalus annulatus*, G = 3.86171, U = 0.91983, P = 0.007744 *, *Acanthocybium solandri*, G = 4.17416, U = 0.90582, P = 0.001811 *, *Vicugna vicugna*, G = 4.69788, U = 0.88006, P = 0.0001195 *, *Pedioplanis lineoocellata*, G = 5.24445, U = 0.84971, P = 4.609 10-06 *, Mac*roscelides proboscideus*, G = 5.07336, U = 0.85858, P = 1.32 10-05 *, Bra*dypus tridactylus*, G = 4.09329, U = 0.90743, P = 0.002583 *, *Trachylepis variegata*, G = 3.95350, U = 0.91317, P = 0.004906 *, *Trachylepis occidentalis*, G = 3.96232, U = 0.91229, P = 0.004677 *, *Trachylepis striata*, G = 3.91219, U = 0.91402, P = 0.005827 *, *Eulamprus tympanum*, G = 3.94396, U = 0.91212, P = 0.005007 *, *Scomber scombrus*, G = 3.64743, U = 0.92441, P = 0.01815 *, *Ctenotus taeniolatus*, G = 3.74618, U = 0.91981, P = 0.01185 *, *Xiphias gladius*, G = 5.09542, U = 0.85079, P = 1.047 10-05 *, Kaj*ikia audax*, G = 4.3445, U = 0.8909, P = 0.0007036 *, *Makaira nigricans*, G = 4.65863, U = 0.87382, P = 0.0001316 *, *Istiophorus albicans*, G = 5.19950, U = 0.84191, P = 5.248 10-06 *, Amo*lops tuberodepressus*, G = 5.11322, U = 0.84621, P = 8.91 10-06 *, Bab*ina pleuraden*, G = 5.0868, U = 0.8469, P = 1.036 10-05 *, Ran*a temporaria*, G = 5.02746, U = 0.84956, P = 1.473 10-05 *, Odo*rrana grahami*, G = 5.08411, U = 0.84523, P = 1.026 10-05 *, Het*eromys desmarestianus*, G = 3.60058, U = 0.92191, P = 0.0205 *, P values computed via approximation method) were excluded from this regression. As Shapiro-Wilk tests (W = 0.96678, P = 1.261 10^{-06} *, W = 0.96351, P = 5.237 10^{-07} *, W = 0.98741, P = 0.02037 *, P value calculated via approximation method) indicated potential non-normality of the phylogenetic residuals after removing outliers, the most influential specimens (*Peromyscus leucopus, Pseudemoia entrecasteauxii, Xenopus laevis*) were excluded from this regression.

Note that we obtain a P = 0.01584 *, when using the complete dataset (n = 428), including potential outliers.

Effect size calculated using the Glass's Δ = -1.50, 95% CI = [-2.16, -0.83], because variances between the factors are still affected by phylogeny and are not similar according to the Bartlett's K-squared test = 48.637, df = 1, P = 3.08 10^{-12} * (P value computed via exact method).

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98858, P = 0.1954, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.

- We visually checked whether the residuals were heteroscedastic (Fig. S69) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 3.37363, U = 0.93102, P = 0.05061, P value computed via approximation method).

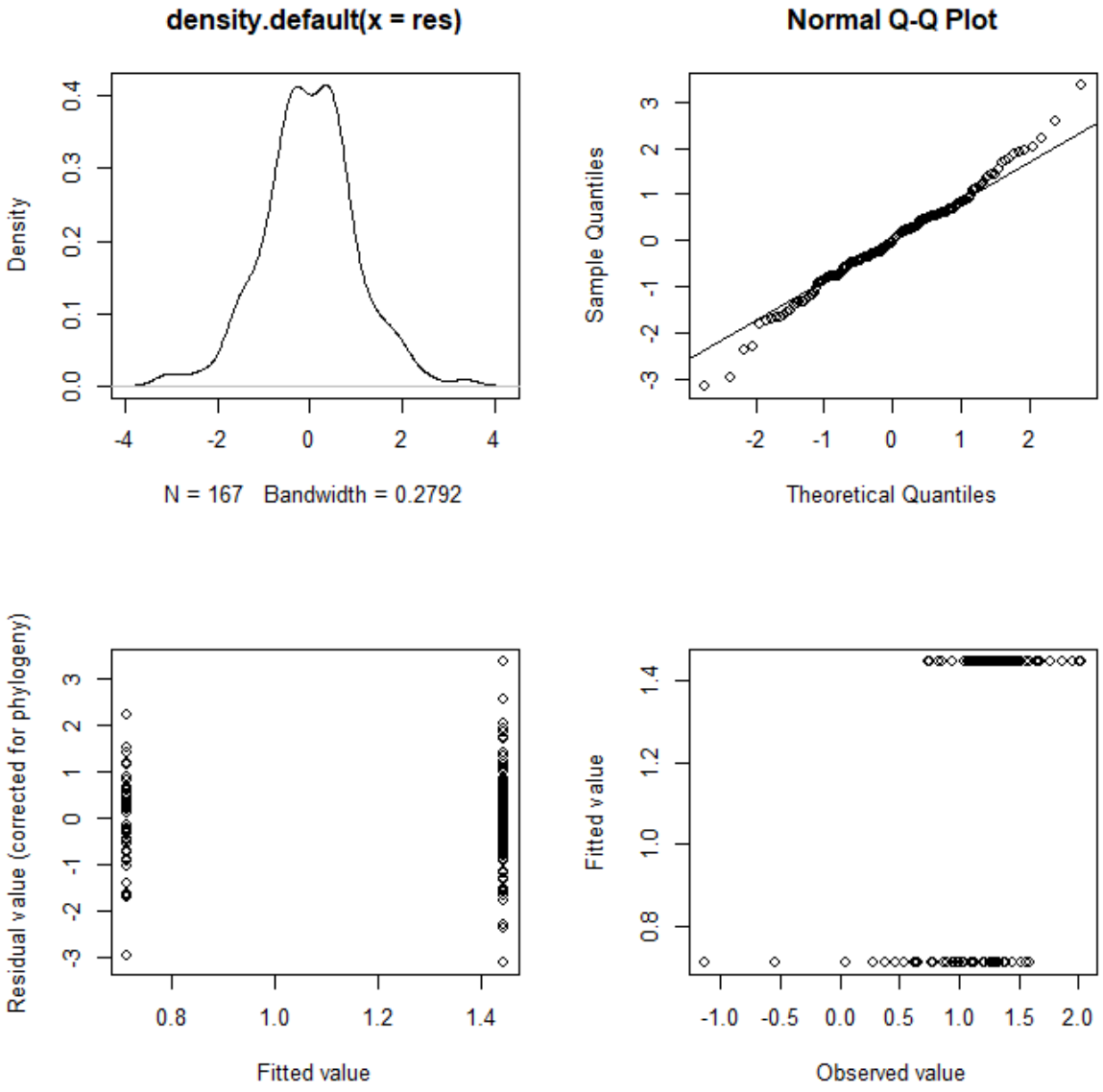


Figure S69 – Diagnostic plots for model 78.

29. PGLS-R of size-corrected maximum anaerobic speed against body temperature

Model 79 provides a significant regression ($P < 0.05$), with a coefficient of determination R^2 equal to 0.20. It was tested on specimens for which we know both the max anaerobic speed and the body temperature, and was used to assess if there was a relationship between size-corrected anaerobic speed (max_speed_adjusted) and body temperature (T_B). Results confirm that the size-corrected maximum anaerobic speed is positively correlated to the body temperature on average for clades.

In this model we used average values for the clades Acanthopterygii, Anguimorpha, Anura, Atlantogenata, Batoidea, Caudata, Crocodylia, Elopomorpha, Euarchontoglires, Galloanserae, Gekkota, Iguania, Lacertoidea, Laurasiatheria, Marsupialia, Monotremata, Neoaves, Otocephala, Palaeognathae, Paracanthopterygii, Protacanthopterygii, Scincomorpha, Selachii, Serpentes, and Testudines. These groups were chosen to best balance taxonomic sampling with robust averaging of size-corrected maximum anaerobic speeds.

This model was also assessed on terrestrial species only, considering only quadrupedal organisms, and excluding arboreal and fossorial behaviours. Results again confirm that the size-corrected maximum anaerobic speed is positively correlated to the body temperature (see Supplementary Data 1).

Model 79: $\log_{10}(\text{max\_speed\_adjusted}) \sim T_B$

Branch length transformations:

Pagel's λ [ML]: 0.001

lower bound: 0.000, $P = 1$

upper bound: 1.000, $P = 0.13624$

95.0% CI: (NA, NA)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	0.4814167	0.1945027	2.4751	0.02112 *

T_B	0.0172926	0.0066247	2.6103	0.01564 *

P values calculated using the exact method.

Adjusted R^2: 0.195, AICc: 17.92757

F statistic: 6.814 on 1 and 23 DF

n = 25 (Supplementary Data 1, maximum anaerobic speed set for clades)

P value: 0.01564 *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.98205, *P* = 0.9226, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S70) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.35450, *U* = 0.75939, *P* = 0.1599, *P* value computed via approximation method).

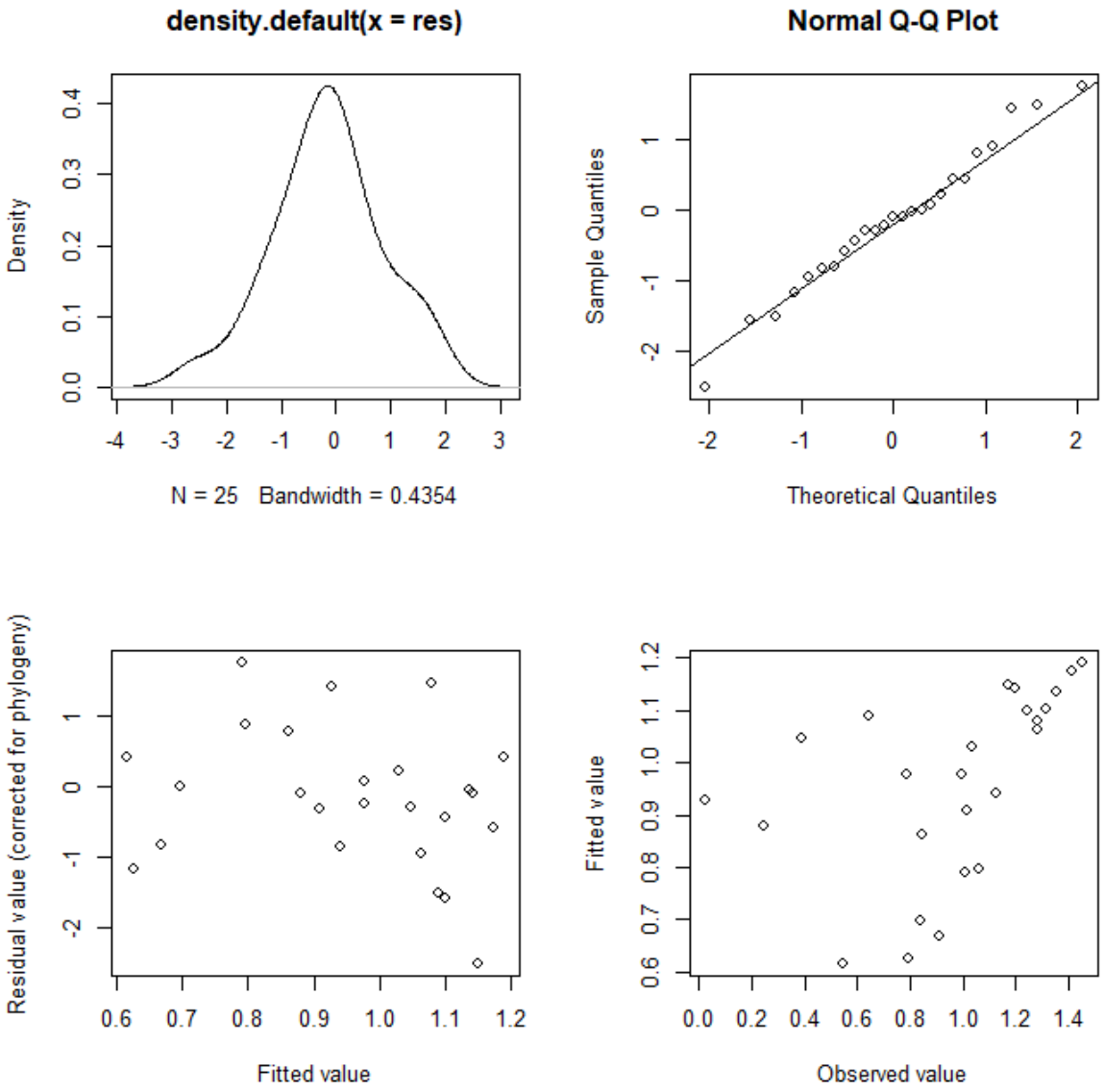


Figure S70 – Diagnostic plots for model 79.

30. PGLS-R of size-corrected maximum aerobic speed against size-corrected maximum anaerobic speed

Model 80 provides a significant regression ($P < 0.01$), with a coefficient of determination R^2 equal to 0.22. It was tested on specimens for which we know both the max aerobic and anaerobic speeds, and was used to assess if there was a relationship between size-corrected aerobic speed (max_speed_aero_adjusted) and size corrected anaerobic speed (max_speed_adjusted). Results confirm that the size-corrected maximum aerobic speed is positively correlated to the size-corrected maximum anaerobic speed.

Model 80: $\log_{10}$(max_speed_aero_adjusted) ~ $\log_{10}$(max_speed_adjusted)

Branch length transformations:

Pagel's λ [ML]: 0.840

lower bound: 0.000, $P = 3.3773\ 10^{-06}$ *

upper bound: 1.000, $P = 0.022651$ *

95.0% CI: (0.466, 0.989)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	-0.62935	0.40855	-1.5405	0.132985
$\log_{10}$(max_speed_adjusted)	0.65211	0.19744	3.3028	0.002309 *

P values calculated using the exact method.

Adjusted R^2: 0.2257, AICc: -2.13814

F statistic: 10.91 on 1 and 33 DF

n = 35 (Supplementary Data 1, maximum aerobic speed set)

P value: 0.002309 *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.9804$, $P = 0.772$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.

- We visually checked whether the residuals were heteroscedastic (Fig. S71) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.19946, U = 0.85353, P = 0.4071, P value computed via approximation method).

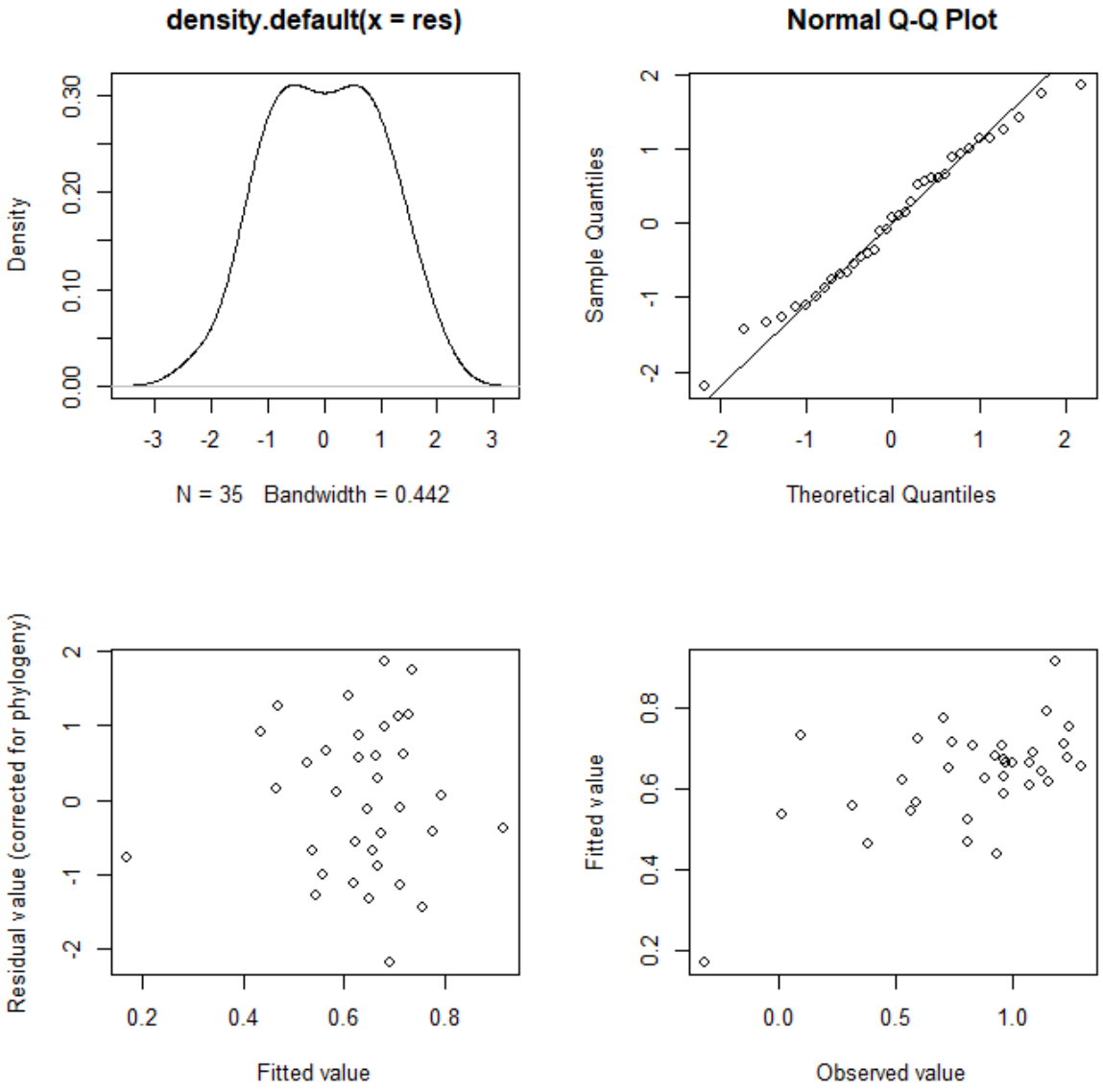


Figure S71 – Diagnostic plots for model 80.

31. PGLS-R of the temperature term against the cubic root of body mass

Model 81 provides a non-significant regression ($P > 0.05$). It was tested on specimens for which we know the body temperature, and was used to assess if there was a relationship between the temperature term (T_R, see Supplementary Methods, Biomechanics) and the cubic root of body mass ($\sqrt[3]{BM}$). Results suggest that there is no correlation between the temperature term and the cubic root of body mass.

In this model we used average values for the clades Acanthopterygii, Anguimorpha, Anura, Atlantogenata, Batoidea, Caudata, Crocodylia, Elopomorpha, Euarchontoglires, Galloanserae, Gekkota, Gymnophiona, Holocephali, Iguania, Lacertoidea, Laurasiatheria, Marsupialia, Monotremata, Neoaves, Otocephala, Palaeognathae, Paracanthopterygii, Protacanthopterygii, Rhynchocephalia, Scincomorpha, Selachii, Serpentes, and Testudines. These groups were chosen to best balance taxonomic sampling with robust averaging of the temperature term.

Model 81: $T_R \sim \log_{10}(\sqrt[3]{BM})$

Adjusted R^2: -0.02161, AICc: -68.87748

F statistic: 0.4289 on 1 and 26 DF

n = 28 (3DB+3DM+2DM set for clades sets where only specimens with measurements of body temperature were included)

P value: 0.5183

P values have been calculated using the exact method.

32. PGLS-R of the residual error of the anterior Thermo-Motility Index against the temperature term

Models 82 and 83 provide non-significant regressions ($P > 0.05$). They were tested on specimens for which we know the body temperature and have the membranous labyrinth, and were used to assess if there were relationships between the temperature term (T_R, see Supplementary Methods, Biomechanics) and the residual error term (ε_{Res}, see Supplementary Methods, Biomechanics). Results suggest that there is no correlation between the temperature term and the residual error term.

Model 82: log10(TMI_res_ω2_a_error) ~ TR

Adjusted R^2: -0.01078, AICc: -76.87477
F statistic: 0.5307 on 1 and 43 DF
n = 45 (3DM set where only specimens with bony labyrinths and measurements of body temperature were included)
P value: 0.4703
P values have been calculated using the exact method.

Model 83: log10(TMI_res_ω2_a_error) ~ TR

Adjusted R^2: 0.03184, AICc: -20.60383
F statistic: 2.447 on 1 and 43 DF
n = 45 (3DM set where only specimens with bony labyrinths and measurements of body temperature were included)
P value: 0.1251
P values have been calculated using the exact method.

33. PGLS-R of stride and impact frequencies against the cubic root of body mass

Models 84 and 85 provide very significant regressions (P << 0.001), with coefficients of determination R^2 ranging from 0.37–0.47. They were tested on specimens for which we know the body mass, and were used to size-correct stride (F_{stride}) and impact (F_{impact}) frequencies. In practice, size corrections were done by combining fitted frequencies obtained for a body mass (BM) of 1g with residuals obtained for each species, for each model. The size-corrected stride (F_{stride}_adjusted) and impact (F_{impact}_adjusted) frequencies are used in Models 86 and 87 to assess whether behaviourally induced head motion can explain variation in the Thermo-Motility Index unexplained by body temperature, and in Models 88 and 89 to check whether stride and/or impact frequencies correlate to body temperature.

Model 84: $\log_{10}(F_{stride}) \sim \log_{10}(\sqrt[3]{BM})$

Branch length transformations:

Pagel's λ [ML]: 0.656

lower bound: 0.000, P = 0.0087238 *

upper bound: 1.000, P = 2.3415 10^{-7} *

95.0% CI: (0.181, 0.891)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.15842	0.14523	7.9763	2.074 10^{-12} *
$\log_{10}(\sqrt[3]{BM})$	-0.65803	0.06723	-9.7877	2.2 10^{-16} *

P values calculated using the exact method.

Adjusted R^2: 0.4745, AICc: 11.65302

F statistic: 95.8 on 1 and 104 DF

n = 106 (3DB + 2DM sets where only specimens with measurements of body temperature and stride frequency)

P value: < 2.2 10^{-16} *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98773, P = 0.4457, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S72) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.27756, U = 0.95013, P = 1, P value computed via approximation method).

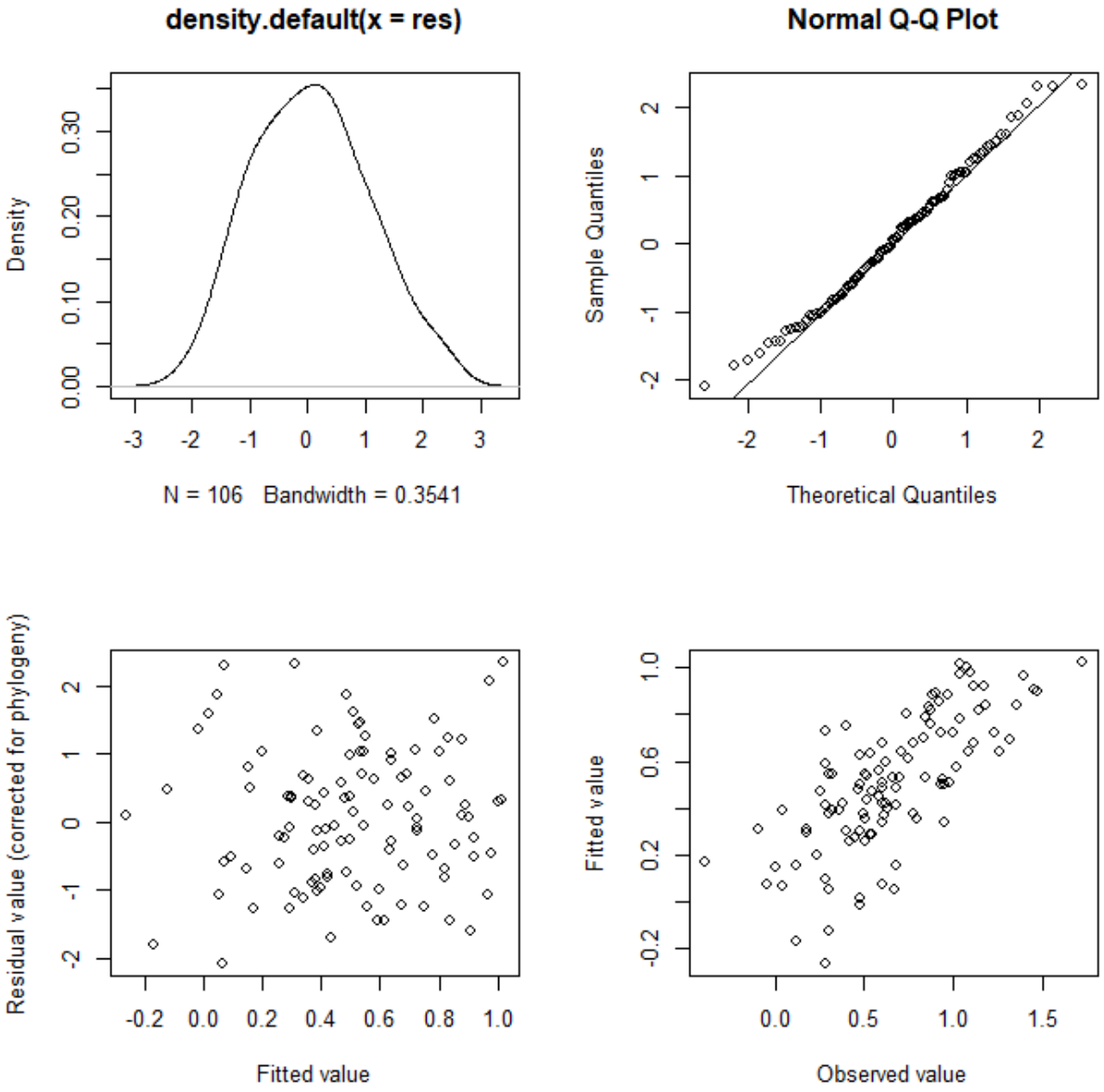


Figure S72 – Diagnostic plots for model 84.

Model 85: $\log_{10}(F_{impact}) \sim \log_{10}(\sqrt[3]{BM})$

Branch length transformations:

Pagel's λ [ML]: 0.646

lower bound: 0.000, $P = 1.5866\ 10^{-5}$ *

upper bound: 1.000, $P = 5.2436\ 10^{-7}$ *

95.0% CI: (0.268, 0.885)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.22531	0.16243	7.5437	$1.798\ 10^{-11}$ *
$\log_{10}(\sqrt[3]{BM})$	-0.59543	0.07590	-7.8449	$4.008\ 10^{-12}$ *

P values calculated using the exact method.

Adjusted R^2: 0.3657, AICc: 37.57292

F statistic: 61.54 on 1 and 104 DF

n = 106 (3DB + 2DM sets where only specimens with measurements of body temperature and stride frequency)

P value: $4.008\ 10^{-12}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.98877$, $P = 0.5239$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S73) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.41147$, $U = 0.94409$, $P = 0.7778$, *P* value computed via approximation method).

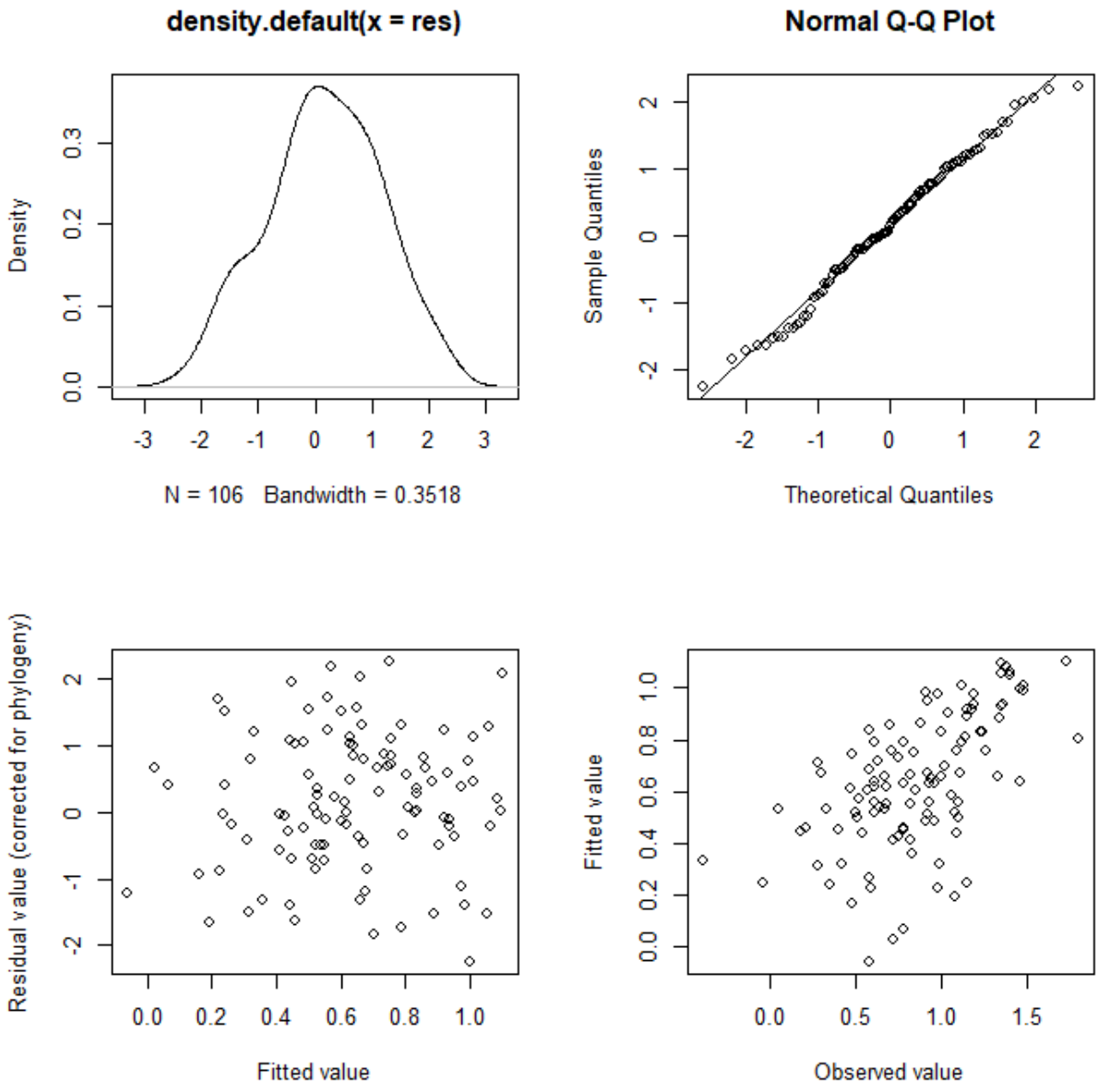


Figure S73 – Diagnostic plots for model 85.

34. PGLS-R of the anterior Thermo-Motility Index against the temperature term and stride or impact frequencies

Models 86 and 87 provide very significant regressions ($P << 0.001$), with coefficients of determination (R^2) ranging from 0.28-0.38. They were tested on specimens for which we know both the stride frequency and the body temperature, and were used to assess if the Thermo-Motility Index (TMI) was related to both the size-corrected impact ($F_{impact_}$adjusted) or stride (F_{stride} adjusted) frequency, and to the temperature term (T_R, see Supplementary Methods, Biomechanics). Results confirm that the Thermo-Motility index is best explained by the combination of impact frequency and body temperature rather than by any of these variables taken individually (AICc Model 86 = -46.1 << AICc $F_{impact_}$adjusted = -35.1 << AICc T_R = -17.7). This further confirms that, at species level, behaviourally induced angular head motion (reflected here by impact/stride frequencies) drives more TMI variation than body temperature, which blurs the relationship between the Thermo-Motility Index and body temperature at this level.

Note that coefficients of determination (R^2) obtained in these regressions are likely underestimates, because stride frequencies reported in the literature were obtained from heterogenous studies, which did not always aim to explore very high speeds or extreme locomotory behaviours. As a result, for a given body mass, the lowest stride frequencies contained in our dataset likely underestimate the maximum frequency of corresponding organisms, more than the highest stride frequencies. This may have resulted in an overestimated spread of maximum frequency values for a given body mass.

Model 86: $\log_{10}(TMI) \sim \log_{10}(F_{impact\_}\text{adjusted}) + T_R$

Branch length transformations:

Pagel's λ [ML]: 0.741

lower bound: 0.000, p = 0.0013816 *

upper bound: 1.000, p = $3.0652\ 10^{-08}$ *

95.0% CI: (0.281, 0.933)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	3.929224	0.933216	4.2104	$5.504\ 10^{-05}$ *
log10($F_{impact_}$adjusted)	0.399593	0.067599	5.9112	$4.538\ 10^{-8}$ *
T_R	-2.251254	0.562631	-4.0013	0.0001196 *

P values calculated using the exact method.

Adjusted R^2: 0.3834, AICc: -46.13615

F statistic: 33.33 on 2 and 102 DF

n = 105 (3DB + 2DM sets where only specimens with measurements of body temperature and stride frequency)

As the initial Q-Q plot and Shapiro-Wilk test (*W* = 0.97893, *P* = 0.09006, *P* value calculated via approximation method) both indicated potential non-normality of the phylogenetic residuals (left skewed), the most influential specimen (*Teratoscincus przewalskii*) was excluded from this regression.

P value: $7.268\ 10^{-12}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.99476, *P* = 0.962, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S74) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.52331, *U* = 0.93819, *P* = 0.5537, *P* value computed via approximation method).

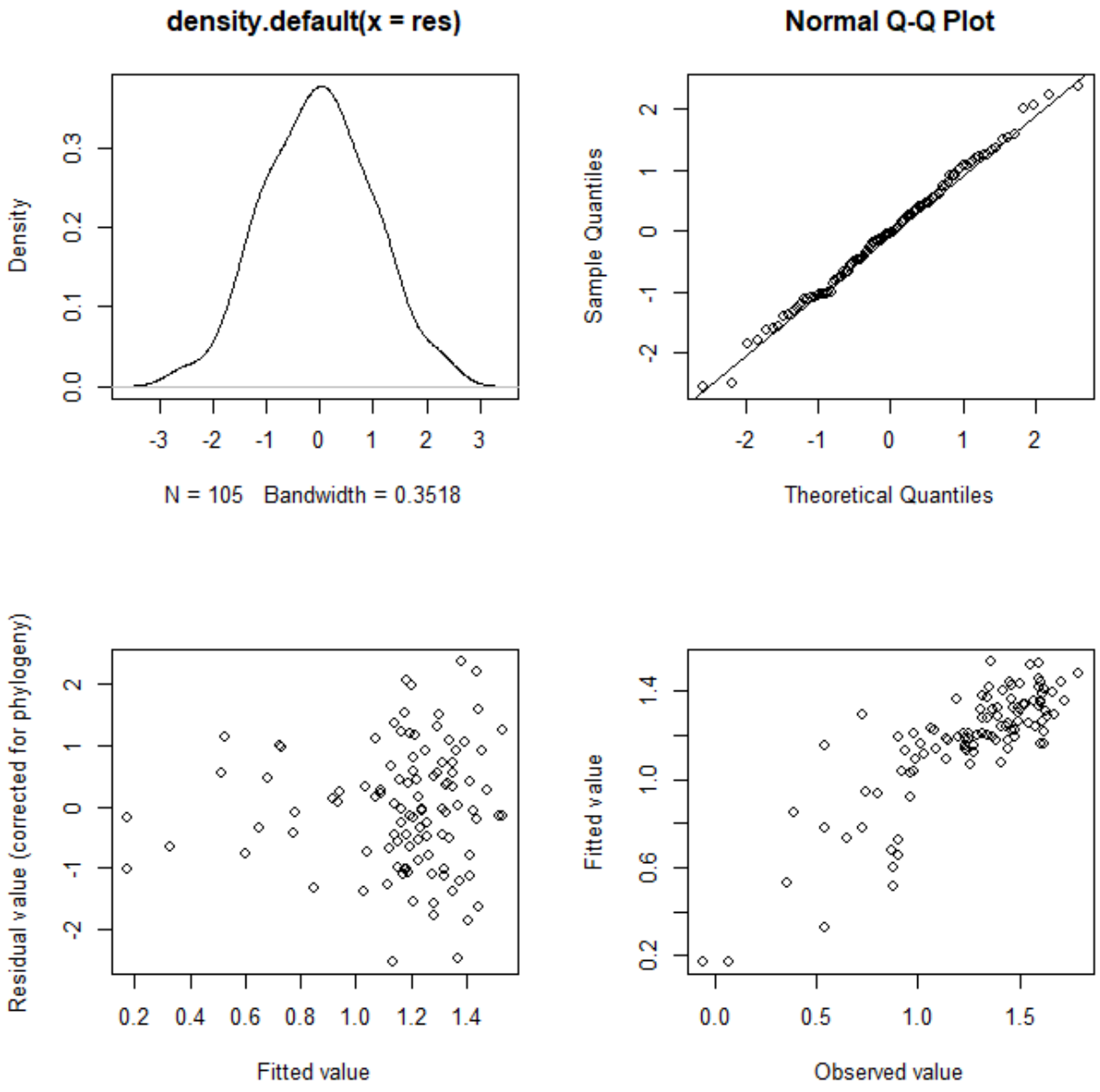


Figure S74 – Diagnostic plots for model 86.

Model 87: $\log_{10}(\text{TMI}) \sim \log_{10}(F_{\text{stride}}\ \text{adjusted}) + T_R$

Branch length transformations:

Pagel's λ [ML]: 0.791

lower bound: 0.000, p = 0.00080069 *

upper bound: 1.000, p = $4.3765\ 10^{-07}$ *

95.0% CI: (0.312, 0.946)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	3.865505	0.954644	4.0492	$9.973\ 10^{-05}$ *
$\log10(F_{\text{stride\_}}\text{adjusted})$	0.360094	0.080327	4.4828	$1.916\ 10^{-5}$ *
T_R	-2.167598	0.574163	-3.7752	0.0002676 *

P values calculated using the exact method.

Adjusted R^2: 0.2807, AICc: -33.15496

F statistic: 21.49 on 2 and 103 DF

n = 106 (3DB + 2DM sets where only specimens with measurements of body temperature and stride frequency)

P value: $1.586\ 10^{-8}$ *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (*W* = 0.99045, *P* = 0.6629, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S75) and concluded they were not.
- We looked for outliers using Grubb's test and found none (*G* = 2.53468, *U* = 0.93823, *P* = 0.5407, *P* value computed via approximation method).

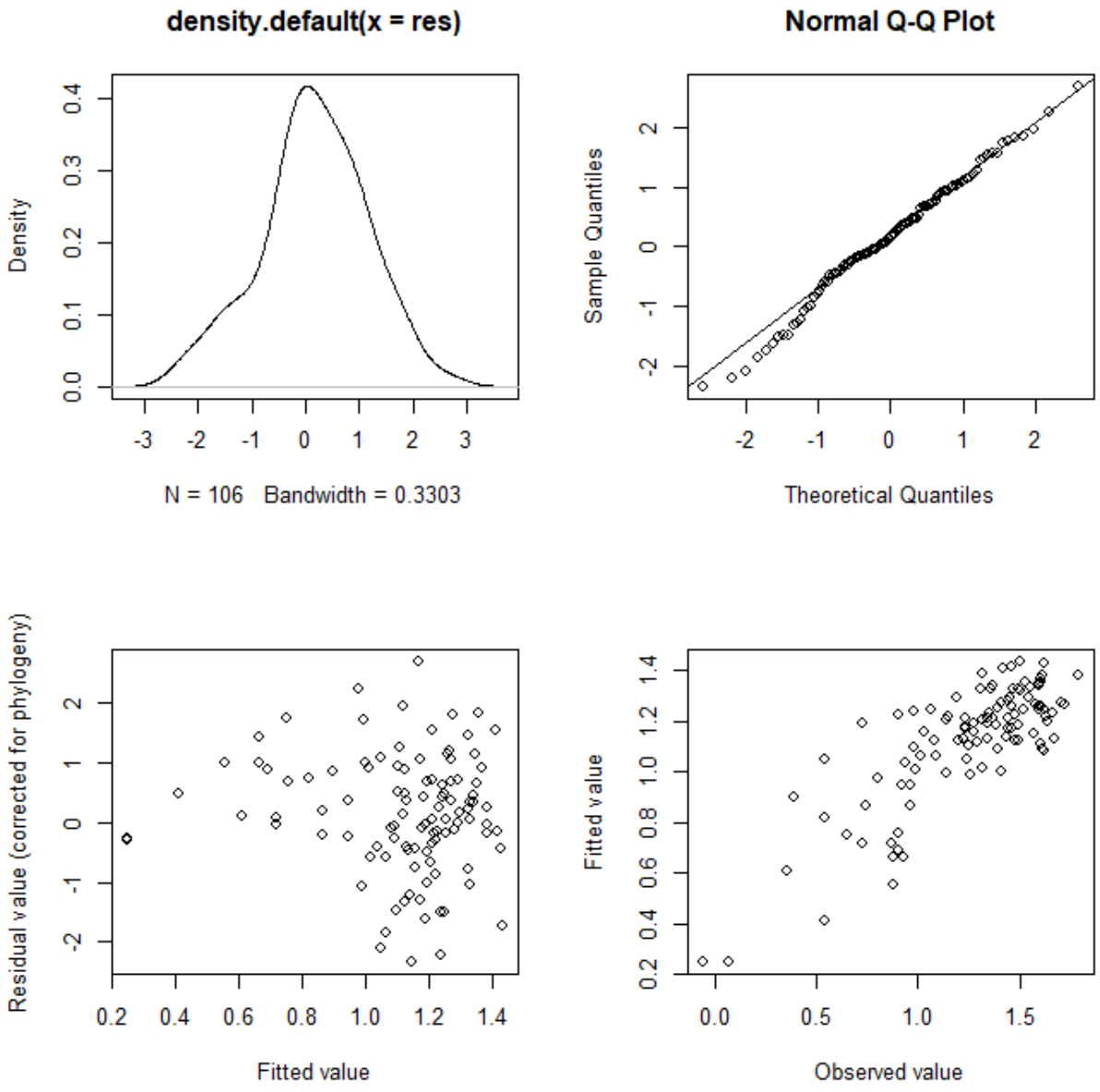


Figure S75 – Diagnostic plots for model 87.

35. PGLS-R of the stride or impact frequencies against the temperature term

Models 88 and 89 provide significant regressions ($P < 0.05$), with coefficients of determination R^2 ranging from 0.04–0.06. They were tested on specimens for which we know both the stride frequency and the body temperature, and were used assess if there was a relationship between size corrected impact ($F_{impact_}$adjusted) or stride ($F_{stride_}$adjusted) frequency and the temperature term (T_R, see Supplementary Methods, Biomechanics). Results confirm that the size-corrected impact/stride frequency (reflecting behaviourally induced angular head motion) is positively correlated to body temperature.

Model 88: $\log_{10}(F_{stride\_}\text{adjusted}) \sim T_R$

Branch length transformations:

Pagel's λ [ML]: 0.606

lower bound: 0.000, $P = 0.042564$ *

upper bound: 1.000, $P = 2.4274\ 10^{-8}$ *

95.0% CI: (0.097, 0.869)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	1.15842	0.14523	7.9763	0.000615 *
T_R	-0.65803	0.06723	-9.7877	0.021880 *

P values calculated using the exact method.

Adjusted R^2: 0.04075, AICc: 5.540257

F statistic: 5.418 on 1 and 103 DF

n = 105 (3DB + 2DM sets where only specimens with measurements of body temperature and stride frequency)

As the initial Q-Q plot and Shapiro-Wilk test ($W = 0.96735$, $P = 0.01029$, *P* value calculated via approximation method) both indicated potential non-normality of the phylogenetic residuals (left skewed), the most influential specimen (*Phoca sp.*) was excluded from this regression.

P value: 0.02188 *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test (W = 0.98449, P = 0.2608, P value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S76) and concluded they were not.
- We looked for outliers using Grubb's test and found none (G = 2.77622, U = 0.92518, P = 0.2494, P value computed via approximation method).

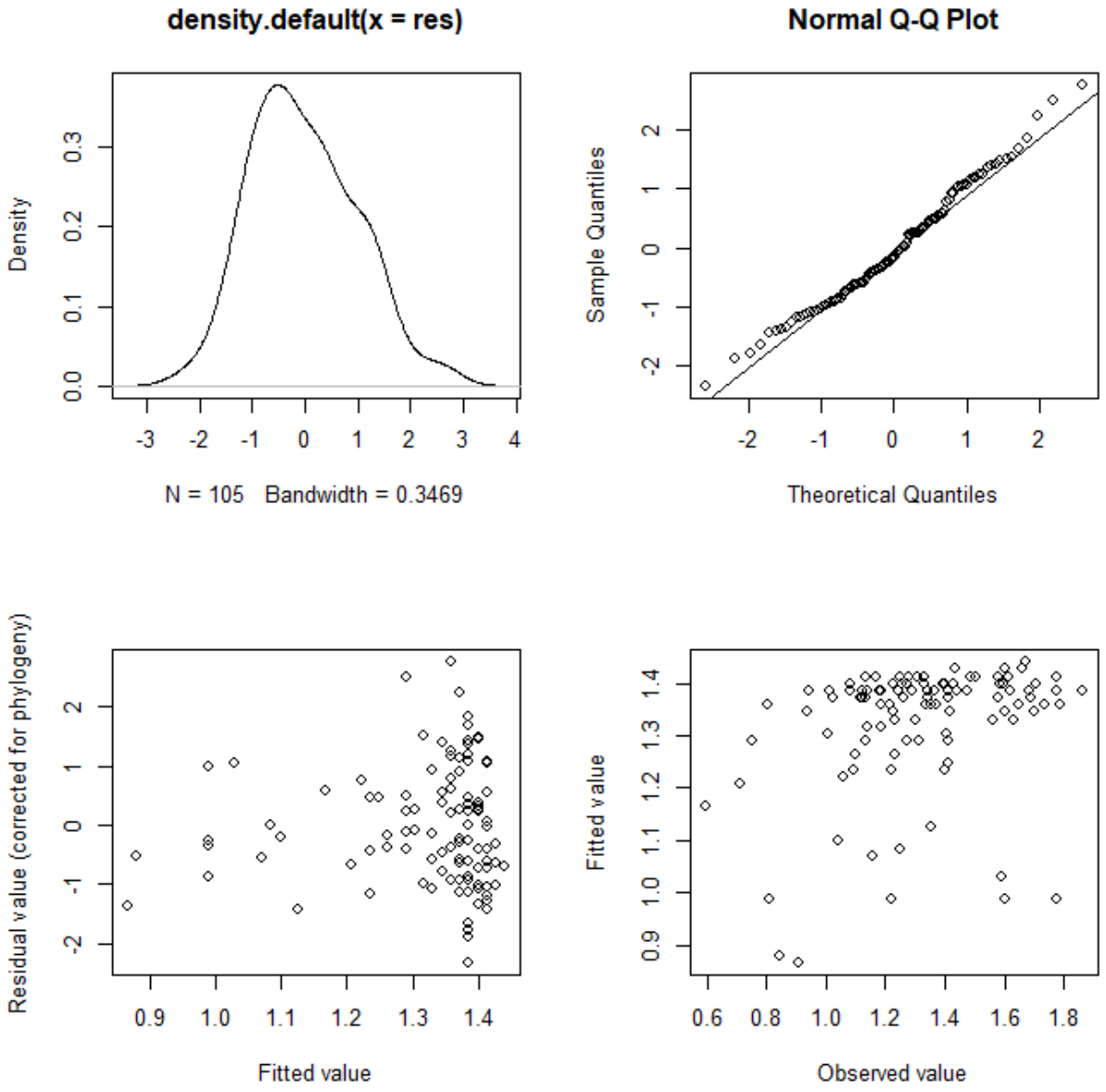


Figure S76 – Diagnostic plots for model 88.

Model 89: $\log_{10}(F_{impact}$ adjusted) ~ T_R

Branch length transformations:

Pagel's λ [ML]: 0.546

lower bound: 0.000, $P = 6.604\ 10^{-5}$ *

upper bound: 1.000, $P = 3.9523\ 10^{-8}$ *

95.0% CI: (0.165, 0.841)

P values approximated using a likelihood ratio test against the χ^2 distribution

Coefficients:

	Estimate	Std. Error	*t* value	Pr(>|t|)
(Intercept)	4.00756	1.02106	3.9249	0.0001561 *
T_R	-1.73105	0.63392	-2.7307	0.0074242 *

P values calculated using the exact method.

Adjusted R^2: 0.05793, AICc: 30.53795

F statistic: 7.457 on 1 and 104 DF

n = 106 (3DB + 2DM sets where only specimens with measurements of body temperature and stride frequency)

P value: 0.007424 *

P value calculated using the exact method.

Assumption testing:

- We tested whether the distribution of phylogenetic residuals significantly differed from normal using the Shapiro-Wilk test ($W = 0.98676$, $P = 0.3794$, *P* value calculated via approximation method).
- Observations are independent from each other thanks to the phylogenetic comparative method.
- We visually checked whether the residuals were heteroscedastic (Fig. S77) and concluded they were not.
- We looked for outliers using Grubb's test and found none ($G = 2.52624$, $U = 0.93864$, $P = 0.5546$, *P* value computed via approximation method).

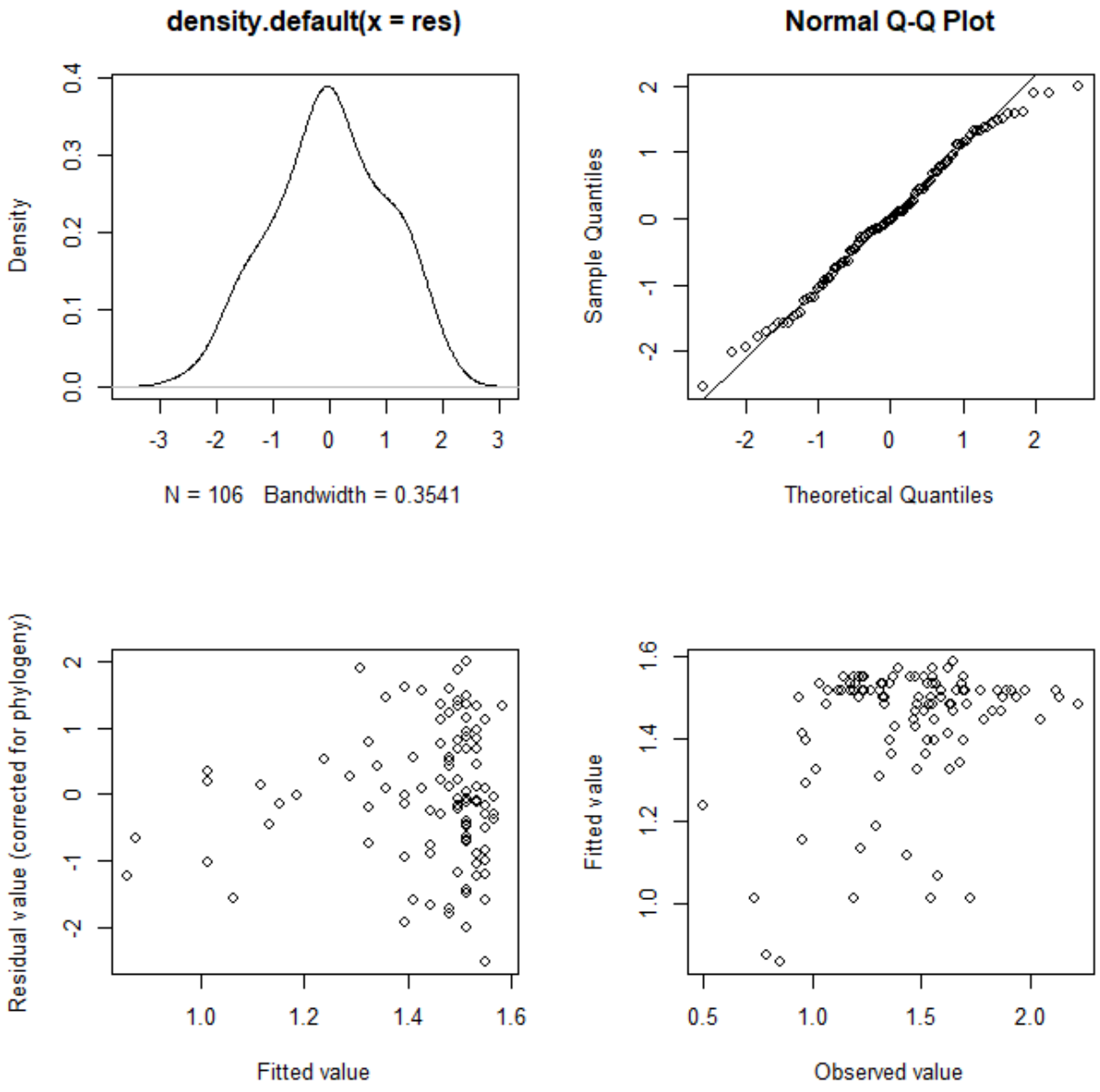


Figure S77 – Diagnostic plots for model 89.

References

Cox, P. G., & Jeffery, N. (2010). Semicircular canals and agility: the influence of size and shape measures. Journal of Anatomy, 216(1), 37-47.

David, R., Stoessel, A., Berthoz, A., Spoor, F., & Bennequin, D., 2016. Assessing morphology and function of the semicircular duct system: introducing new in-situ visualization and software toolbox. Scientific reports, 6(1), 1-14.

Hirt, M. R., Jetz, W., Rall, B. C., & Brose, U. (2017). A general scaling law reveals why the largest animals are not the fastest. Nature Ecology & Evolution, 1(8), 1116-1122.

Mundry, R. (2011). Issues in information theory-based statistical inference—a commentary from a frequentist's perspective. Behavioral Ecology and Sociobiology, 65(1), 57-68.

Mundry, R. in *Modern phylogenetic comparative methods and their application in evolutionary biology* (eds. Garamszegi, L. Z.) 131–153 (Springer, 2014).

Smith, R.J., 2009. Use and misuse of the reduced major axis for line-fitting. American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists, 140(3), pp.476-486.

Stoltzfus, J.C., 2011. Logistic regression: a brief primer. Academic Emergency Medicine, 18(10), pp.1099-1104.

Supplementary Note 3 – *Dimetrodon* photogrammetry

The *Dimetrodon* specimen used in this study (FMNH PR 4976) is a wax endocast model that was reconstructed from a serially-sectioned skull. The inclusion of the specimen was paramount because, despite our efforts to scan multiple non-therapsid synapsids (Supplementary Data 3), only one retrieved usable data. The FMNH collections records do not include information on the original specimen that was sectioned to make the endocast model. However, comments in Olson (1938; also see Romer & Price, 1940) and notes from A. S. Romer preserved in the Museum of Comparative Zoology archives suggest that the endocast was reconstructed from serial sections of MCZ 9560. For measurement purposes, a photogrammetric model of FMNH PR 4976 was created based on 50 individual photographs of the specimen. The model was reconstructed using Colmap 3.5 (https://colmap.github.io/index.html). The endocast is comprised of about 30 slices with an average thickness of 2.147 mm, resulting in a total length of 64.42 mm. We estimate that the condylobasal length of MCZ 9560 was about 260 mm, so the model clearly has been scaled up from the original size of the bony labyrinth. In his description of his serial sectioning technique, Olson (1944; p.9) states: "In cases of skulls under 40mm in length, it has proved very difficult to obtain evenly spaced sections since they should not be over .15mm apart. Intervals of .3 mm to .5 mm are satisfactory for skulls between 50 and 65 mm, and intervals of .75 mm for those between 65 and 100 mm. The best results have been obtained with skulls of over 100 mm by using intervals of .6 to .8mm in the otic region of the brain cases and intervals of 1.0 to 1.5mm for the more anterior parts of the skulls." He also lists slice thickness ranging from 0.37–1.1 mm (average 0.744 mm) for the specimens he describes in that work. Based on those values, we estimate that the endocast has been scaled up by a factor of between 2.68x (0.8mm slice thickness) to 3.58x (0.6 mm slice thickness), as suggested by the original publications (Olson 1938, 1944, Romer et al. 1940) and notes of A. S. Romer. Hence, measurements taken on the wax model were corrected by the average scaling factor. Using either factor instead of the average does not affect the conclusions of the study.

References

Olson, E. C. The occipital, otic, basicranial and pterygoid regions of the Gorgonopsia. J. Morphol. 62,141–175 (1938).

Olson, E. C. Origin of mammals based upon cranial morphology of the therapsid suborders. Geol. Soc. Am., 55 (1944).

Romer, A. S., Price, L. W. & Price, L. I. Review of the Pelycosauria. Geol. Soc. Am. 28(1940).

Supplementary Note 4 – Divergence times and last occurrence datum

Divergence Times

Common ancestor of *Diadectes* and Amniota: The relationships of diadectomorphs are controversial. They have often been recovered as the sister group of Amniota (e.g., Gauthier et al. 1988; Laurin & Reisz 1995, 1997; Lee & Spencer 1997; Kissel & Reisz 2004a; Reisz 2007; Kissel 2010; Liu & Bever 2015; Laurin & Piñeiro 2017; Ford & Benson 2019), but other analyses have recovered them nested within Amniota, either as the sister taxon of Sauropsida (Ruta et al. 2003; Ruta & Coates 2007; Marjanović & Laurin 2019) or Synapsida (Sumida et al. 1992; Berman et al. 1992; Berman 2000, 2013). Here, we used the 'traditional' hypothesis that diadectomorphs are the sister group of amniotes, although we acknowledge that this is an area of continuing research. Despite this stemward position, the known fossil record of diadectomorphs begins after the first appearances of Synapsida and Sauropsida. The oldest records of diadectomorphs are occurrences of *Desmatodon* in the Kasimovian (Late Pennsylvanian) Sangre de Cristo Formation of Colorado and the Red Knob Formation of Pennsylvania (Reisz 2007; Kissel 2010). The nested position of *Desmatodon* within Diadectomorpha and the stratigraphic ranges of other taxa in the clade imply that much of the earliest history of the clade is not recorded (Reisz 2007; Kissel 2010). Therefore, it is necessary to turn to other taxa to calibrate the time of origin of Amniota.

Benton et al. (2015) used the Visean (Middle Mississippian) aïstopod *Lethiscus stocki* to calibrate the divergence of crown Tetrapoda, reflecting the hypothesis that aïstopods are members of Lepospondyli and thus part of total group Amniota (e.g., Ruta et al. 2003; Ruta & Coates 2007). However, reconsideration of the phylogenetic relationships of aïstopods has raised the possibility that they are stem tetrapods, not members of the crown group (Pardo et al. 2017). If that hypothesis is correct, the stem lissamphibian *Balanerpeton* and the stem amniote *Westlothiana* from the upper Visean East Kirkton locality of Scotland (Pardo et al. 2017), which has been dated to 332.9–330.9 Mya (Benton et al. 2015), are the next relevant calibration points. Therefore, the divergence between diadectomorphs and crown amniotes must have occurred between the upper Visean and the Bashkirian given the first occurrences of synapsids and sauropsids in the Joggins Formation of Nova Scotia, which has been calibrated at 318 Mya (Benton et al. 2015; also see below). We use the midpoint of this age range (324.5 Mya) to calibrate the divergence of Diadectomorpha and Amniota.

Synapsida + Sauropsida: The divergence between synapsids and sauropsids occurred by the Bashkirian Stage of the Early Pennsylvanian (late Carboniferous Period). This divergence is calibrated by the presence of the sauropsid *Hylonomus lyelli* and the putative synapsids *Protoclepsydrops haplous* and *Asaphestera*

platyris, which occur in the same stratigraphic horizon in the Joggins Formation of Nova Scotia (Carroll 1964; Mann et al. 2020). *Protoclepsydrops* is known from limited, fragmentary material, and its identity as a synapsid has been questioned (Reisz 1972, 1986; Mann & Paterson 2019). *Asaphestera* was long considered a 'microsaur', but had recently been re-identified as a synapsid, possibly a caseasaur (Mann et al. 2020). The sauropsid identity of *Hylonomus* is better supported (Müller & Reisz 2006), so its presence in the fossil record implies that synapsids and sauropsids must have diverged by this time, even if *Protoclepsydrops* and/or *Asaphestera* are not synapsids. Age estimates for the Joggins Formation range from about 319–310 million years ago (Reisz & Müller 2004, van Tuinen & Hadly 2004), and Benton et al. (2015) recommended of 318 million years ago as a minimum age for this divergence. We use a slightly older date of 320 Mya to accommodate the earliest divergences among sauropsids (see divergence between Captorhinidae and Diapsida, below).

Caseasauria + Eupelycosauria: *Eocasea martinsi* and *Datheosaurus macrourus* are the oldest known relatively certain caseasaurs, and both are likely Gzhelian in age (Reisz & Fröbisch 2014; Spindler et al. 2016). However, the oldest taxa relevant for calibrating this node are the ophiacodontids *Archaeothyris florensis* and *Echinerpeton intermedium*, and the varanopid *Dendromaia unamakiensis* (assuming varnopids are synapsids; see discussion in Ford & Benson 2019) from the slightly older Morien Group of Nova Scotia (Reisz 1972; Maddin et al. 2019; Mann & Paterson 2019), which is likely Moscovian–Kasimovian in age (see discussion in van Tuinen & Hadley 2004). The presence of ophiacodontids at this time implies the caseasaurs had also diverged by this point, regardless of the recent debate on the phylogenetic relationships of early synapsids (e.g., compare the results of Benson 2012 and Brocklehurst et al. 2016). If *Asaphestera* is indeed a caseasaur (Mann et al. 2020), it would provide direct confirmation of this hypothesis. The numerical age of the divergence should be between 309.2 Mya (the older end of the range given for the age of *Archaeothyris* in van Tuinen & Hadley 2004) and 318 Mya (Benton et al.'s 2015 hard minimum for the divergence of Synapsida). 313.6 Mya is the midpoint of this range.

Sphenacodontidae + Therapsida: Sphenacodontidae is the sister taxon of Therapsida and first appears in the Kasimovian Sangre de Cristo Formation in Colorado (Sumida & Berman 1993). Potentially older sphenacodontid material has been reported from Nova Scotia (Reisz 1972), but given the uncertainty in its identification we use the Sangre de Cristo record as the oldest sphenacodontid. The oldest known therapsid material is substantially younger: the early Permian *Tetraceratops insignis* and the middle Permian *Raranimus dashankouensis* (Liu et al. 2009; Amson & Laurin 2011; although see Spindler 2014, 2020). The Kasimovian ranges in age from 307.0–303.7 Mya, and we use the midpoint of this range (305.35 Mya) as the calibration point for this divergence.

Biarmosuchia + Eutherapsida: Biarmosuchians are generally considered to be the most stemward major therapsid clade (e.g., Sidor & Hopson 1998; Liu et al. 2009, 2010), although *Raranimus dashankouensis* and *Tetraceratops insignis* likely fall even further down the therapsid stem (Liu et al. 2009; Amson & Laurin 2011; Brink et al. 2015) (and *Tetraceratops* may not be a therapsid at all; see Conrad & Sidor 2001; Spindler 2020). Sphenacodontidae, the sister taxon of Therapsida, first appears in the Late Pennsylvanian (Sumida & Berman 1993), but there are no known Pennsylvanian therapsids (although see Spindler 2014). Therefore, while divergences between therapsid lineages could date back to the Pennsylvanian, there is no direct evidence for this. The putative early Permian therapsid *Tetraceratops insignis* (Amson & Laurin 2011; although see Spindler 2020) does not help to resolve this issue because it does not appear to fall within any of the major therapsid clades. The divergence between biarmosuchians and other therapsids is calibrated by the oldest record of Dinocephalia. Specifically, anteosaurid dinocephalians are present in the Golyusherma Assemblage of Russia, which is considered to be early Roadian (early middle Permian) in age (Golubev 2015), making these fossils the oldest known unequivocal therapsid occurrence. The Roadian is 272.3 to 268.8 Mya, and we calibrate this divergence at 271 Mya, in part to accommodate other divergences whose timing is based on the first appearance of dinocephalians (see below).

***Hipposaurus* + Burnetiamorpha:** The biarmosuchians *Luecocephalus*, *Herpetoskylax*, and *Lemurosaurus* occur in the Wuchiapingian *Tropidostoma-Gorgonops* subzone (*Endothiodon* Assemblage Zone) and *Cistecephalus* Assemblage Zone of the Karoo Basin, South Africa (e.g., Sidor & Welman 2003; Sidor & Rubidge 2006; Day et al. 2016, 2018a; Day & Smith 2020; Smith 2020), but *Hipposaurus* is older, occurring in the Capitanian upper Abrahamskraal Formation (*Diictodon-Styracocephalus* subzone, *Tapinocephalus* Assemblage Zone) (e.g., Day et al. 2018; Day & Rubidge 2020). The divergence time of *Hipposaurus* is bracketed by the presence of the more stemward biarmosuchian *Biarmosuchus* in the likely Wordian Ocher Subassemblage of European Russia (e.g., Sennikov & Golubev 2017) and the presence of the more derived bullacephalids *Pachydectes* and *Bullacephalus* in the lunderlying *Eosimops-Glanosuchus* subzone of the *Tapinocephalus* Assemblage Zone (Day & Rubidge 2020). Lanci et al. (2013) provided an age range of 268.5–264.6 Mya for the lower Abrahamskraal Formation, and we use the lower end of this range (268.5 Mya) as the calibration point for this node.

***Herpetoskylax* + Burnetiamorpha:** The biarmosuchian *Leucocephalus* occurs in the Wuchiapingian *Tropidostoma-Gorgonops* subzone of the *Endothiodon* Assemblage Zone of the Karoo Basin, whereas *Herpetoskylax* and *Lemurosaurus* both occur in the slightly younger *Cistecephalus* Assemblage Zone of the Karoo Basin, South Africa (e.g., Sidor & Welman 2003; Sidor & Rubidge 2006; Day et al. 2016, 2018a; Day & Smith 2020; Smith 2020), but they are members of two distinct lineages of biarmosuchians that diverged earlier in the Permian. The age of this divergence depends strongly on whether the lower

Tapinocephalus Assemblage Zone taxa *Pachydectes* and *Bullacephalus* (i.e., the Bullacephalidae) are members of the nested clade Burnetiamorpha or if they fall near the base of the biarmosuchian phylogeny (Day et al. 2016). Phylogenetic hypotheses supporting bullacephalids as burnetiamorphs are more common in the biarmosuchian literature despite their poorer overall fit to stratigraphy (e.g., Sidor & Smith 2007; Kruger et al. 2015; Day et al. 2016, 2018a; Kammerer 2016a), and we use this topology here. Lanci et al. (2013) provide dates ranging from 268.5–264.6 Mya for the lower Abrahamskraal Formation (i.e., the lower *Tapinocephalus* Assemblage Zone) so we use 265 Mya (approximate age of the Wordian-Capitanian boundary) for this calibration point.

***Lemurosaurus* + *Leucocephalus*:** As noted in the previous entry, *Leucocephalus* occurs in strata assigned to the *Tropidostoma-Gorgonops* Subzone (*Endothiodon* Assemblage Zone) of the Karoo Basin, whereas *Lemurosaurus* occurs in strata assigned to the *Cistecephalus* Assemblage Zone (Sidor & Welman 2003; Day et al. 2018a; Day & Smith 2020; Smith 2020). *Lemurosaurus* is typically reconstructed as the most stemward burnetiamorph (e.g., Sidor & Welman 2003; Smith et al. 2006; Kruger et al. 2015; Day et al. 2016; Kammerer 2016a), so the minimum age of this divergence is calibrated by the oldest burnetiamorph records. In addition to *Lemurosaurus*, the oldest burnetiamorphs are *Proburnetia* and *Niuksenitia*, both of which occur in the Ilinskoe faunal assemblage of Russia (e.g., Sennikov & Golubev 2017) whose early Wuchiapingian age likely falls within the *Endothiodon* Assemblage Zone of South Africa, potentially within the *Tropidostoma-Gorgonops* Subzone (Sennikov & Golubev 2017; Schneider et al. 2019; Day & Smith 2020). The *Endothiodon* Assemblage Zone is between 260.26–256.8 Mya, with the *Tropidostoma-Gorgonops* Subzone being about 258 – 256.8 Mya (Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020), and we calibrate this divergence at 260 Mya to allow time for the multiple burnetiamorph lineages to diverge before their first appearances in Russia and South Africa.

Dinocephalia + Neotherapsida: As with the divergence between biarmosuchians and Eutherapsida, this divergence is calibrated by the first appearance of dinocephalians in the Roadian Golyusherma Assemblage of Russia (Golubev 2015). We calibrate this divergence at 270 Mya, placing it in the Roadian, but slightly after the divergence of Biarmosuchians and eutherapsids.

Anomodontia + Theriodontia: The exact time of divergence between Anomodontia and Theriodontia (i.e., Gorgonopsia + Therocephalia + Cynodontia) is somewhat uncertain. In South Africa, anomodonts, therocephalians, and very fragmentary gorgonopsian remains are known from the *Eodicynodon* Assemblage zone in the Abrahamskraal Formation (e.g., Abdala et al. 2008; Day et al. 2018; Rubidge & Day 2020). Dates from near the base of this formation range from 268.5–264.6 Mya (Lanci et al. 2013). The portion of the formation corresponding to the overlying *Tapinocephalus* Assemblage Zone are in the range of 262–

261 Mya (Day et al. 2015), and the underlying Ecca Group has been dated to approximately 290–265 Mya (e.g., Tohver et al. 2015; Rubidge et al. 2016; Belica et al. 2017), but the latter strata are well below any therapsid occurrences. Together, this implies that the oldest South African therapsids (including early anomodonts) are in the Wordian to early Capitanian range. The non-dicynodont anomodont *Otsheria* is known from the Ocher Assemblage of Russia, which also probably falls in the Wordian (Golubev 2015). The most stemward anomodont, *Biseridens* is known from the Dashankou locality in China, where it co-occurs with dinocephalians but not theriodonts (Liu et al. 2010). This locality is often portrayed as older than either the *Eodicynodon* zone or the Ocher Assemblage (e.g., Liu et al. 2009, 2010) based on the primitive gestalt of *Biseridens* and the primitive therapsid *Raranimus*, but its age is poorly constrained and Rubidge & Day (2020) suggested a broad correlation between it and the *Eodicynodon* zone (also see Olroyd & Sidor 2017). Based on this information, the divergence of anomodonts and theriodonts is likely older than 262 Mya and probably older than 264–265 Mya, but younger than 272 Mya. We use 268.5 Mya as the date for this divergence (upper end of the Lanci et al. 2013 range of ages for the lower Abrahamskraal Formation).

***Patranomodon* + Dicynodontia**: *Eodicynodon oosthuizeni*, the most stemward dicynodont, occurs in the eponymous *Eodicynodon* Assemblage Zone (lower Abrahamskraal Formation, Karoo Basin, South Africa), where it co-occurs with the non-dicynodont anomodont *Patranomodon nyaphulii* (e.g., Day et al. 2018; Rubidge & Day 2020). Radiometric dates and biostratigraphic correlations do not exist with sufficient precision to differentiate the first appearance of *Eodicynodon* from the divergence of Anomodontia, so we place the origin of Dicynodontia one million years after the origin of Anomodontia, 267.5 Mya.

Common ancestor of *Eodicynodon* and *Lystrosaurus*: As noted above, the currently-available fossil record lacks sufficient resolution to provide clear ages for divergences at the base of Dicynodontia. Therefore, we calibrate this divergence point one million years after the origin of Dicynodontia, 266.5 Mya.

Common ancestor of *Diictodon* and *Lystrosaurus*: *Diictodon feliceps* is a member of the dicynodont clade Pylaecephalidae. *Eosimops newtoni* and *Robertia broomiana* are the oldest members of this clade, with stratigraphic ranges extending down into the Koornplaats Member of the Abrahamskraal Formation (*Eosimops-Glanosuchus* Subzone of the *Tapinocephalus* Assemblage Zone; Day 2013; Day et al., 2018b; Day & Rubidge 2020). However, its sister lineage (including *Lystrosaurus*) has a potentially older first appearance in *Brachyprosopus broomi*. The latter species also is present in the Abrahamskraal Formation at least as far down as the Leeuvlei Member. A first appearance in the Leeuvlei member would put the first appearance of *Brachyprosopus* the lowermost *Tapinocephalus* Assemblage Zone, near the boundary with the *Eodicynodon* Assemblage Zone (Angielczyk et al. 2016; Day et al. 2020). Together, these data indicate

that this divergence falls between the 261–262 Mya radiometric dates in the upper *Tapinocephalus* zone (Day et al. 2015) and the 268–264 Mya range for the dates from the base of the Abrahamskraal Formation (Lanci et al. 2013). We place this divergence at 264 Mya.

Common ancestor of *Pristerodon* and *Lystrosaurus*: As with the previous calibration, *Brachyprosopus broomi* is likely the oldest relevant taxon whose age is relatively well-constrained. We calibrate this node at 263 Mya, subtracting one million years from the previous calibration.

Endothiodontia + Therochelonia: The oldest members of these lineages, *Abajudon*, *Rastodon*, and *Emydops* co-occur with dinocephalian therapsids in the Karoo (South Africa), Ruhuhu (Tanzania), Mid-Zambezi (Zambia), and Paraná (Brazil) basins (Boos et al. 2016; Day et al. 2018; Olroyd et al. 2018; Day & Rubidge 2020). *Abajudon* and *Rastodon* are assumed to be Capitanian in age because of their co-occurrence with dinocephalians, but their ages are otherwise poorly constrained. The records of *Emydops* in the uppermost Abrahamskraal Formation (upper *Tapinocephalus* Assemblage Zone) in the Karoo Basin are close to the dates of 261.24–260.26 Mya for the upper *Diictodon-Styracocephalus* Subzone of the *Tapinocephalus* Assemblage Zone (Rubidge et al. 2013; Day et al. 2015; Day & Rubidge 2020). We calibrate this node at 262.5 Mya to reflect that the divergence must have occurred after the divergence in the previous calibration but before the ca. 261 Mya first appearance of *Emydops* in the fossil record.

Endothiodontia: *Abajudon kaayai*, which co-occurs with tapinocephalid dinocephalians in the Ruhuhu Formation of Tanzania and the lower Madumabisa Mudstone Formation of the Mid-Zambezi Basin of Zambia is likely the oldest known endothiodont, although radiometric dates are not available for either formation (Angielczyk et al. 2014a; Olroyd et al. 2018). However, the widespread genus *Endothiodon* is not known to co-occur definitively with dinocephalians (it is uncertain whether the *Endothiodon* specimen from the Brazilian Rio do Rasto Formation comes from the same stratigraphic level as the dinocephalians reported from that formation; Boos et al. 2013. 2015; Day & Smith 2020), nor does the recently described genus *Niassodon* (e.g., Ray 2000; Boos et al. 2013; Castanhinha et al. 2013; Cox & Angielczyk 2018; Day et al. 2018; Macungo et al. 2020) implying they are at least slightly younger. We calibrate the base of Endothiodontia at 262 Mya to reflect that divergences within the clade must have post-dated the divergence between Endothiodontia and Therochelonia (see previous entry) and must pre-date the first appearance of *Endothiodon bathystoma* in the lower Poortjie Member of the Teekloof Formation, which is between approximately 261–259 Mya (Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020), and the likely similarly-aged first appearance of *E. tolani* in the Ruhuhu Formation of Tanzania (Angielczyk et al. 2014; Cox & Angielczyk 2015) and the K5 Formation of Mozambique (Macungo et al. 2020).

***Abajudon* +*Endothiodon*:** The presumably Capitanian first appearance of *Abajudon* discussed in the previous entry is the oldest relevant calibration point for this divergence as well. We calibrate this node at 261.5 Mya to reflect the fact that it must post-date the emergence of Endothiodontia and pre-date the first appearance of *E. bathystoma* in the lower Poortjie Member of the Teekloof Formation (Day et al. 2018).

***Endothiodon tolani* + *Endothiodon bathystoma*:** *Endothiodon tolani* was described recently from the middle fossiliferous horizon of the Ruhuhu Formation (*R2* of Olroyd & Sidor 2017), Ruhuhu Basin, Tanzania (Cox & Angielczyk 2015). The age of this horizon is not well constrained, but it has been hypothesized to fall near the Guadalupian-Lopingian boundary based on biostratigraphic comparisons (Angielczyk et al. 2014a; Cox & Angielczyk 2015; Olroyd & Sidor 2017). Recently, Macungo et al. (2020) reported *E. tolani* specimens from the K5 Formation of the Metangula Graben of Mozambique, and unerupted tusks visible in CT-scans of the maxillae of NHCC LB648 allow us to recognize that specimen as the first occurrence of *E. tolani* in the Madumabisa Mudstone Formation of the Mid-Zambezi Basin, Zambia. The ages of these units are not well constrained, but there are reasons to believe that fall within the Lopingian (Castanhinha et al. 2013; Barbolini et al. 2016). *Endothiodon bathystoma* has a cosmopolitan distribution in Gondwana (e.g., Boos et al. 2013; Cox & Angielczyk 2015), but its best-dated records are in the Karoo Basin of South Africa. There, *E. bathystoma* first appears in in the lower Poortjie Member of the Teekloof Formation (Day et al. 2018), which is between approximately 261–259 Mya (Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020). Together, these observations suggest that *E. tolani* and *E. bathystoma* diverged no later than the latest Capitanian, and we calibrate this divergence at 261 Mya.

Emydopoidea + Bidentalia: The records of *Emydops* and *Rastodon* noted in the entry for Endothiodontia + Therochelonia are the oldest relevant calibration points for this node. We calibrated the divergence at 262 Mya based on the fact that it must post-date the divergence of Endothiodontia and Therochelonia, but pre-date the first appearance of *Dicynodontoides* in the Capitanian-Wuchiapingian *Lycosuchus-Eunotosaurus* subzone of the *Endothiodon* Assemblage Zone of the Karoo Basin (Angielczyk et al. 2009; Day & Smith 2020).

***Dicynodontoides* + Cistecephalidae:** The first appearance of *Dicynodontoides* is in the upper Poortjie Member of the Teekloof Formation (*Lycosuchus-Eunotosaurus* subzone of the *Endothiodon* Assemblage Zone) of the Karoo Basin (Angielczyk et al. 2009; Day et al. 2018; Day & Smith 2020), which is between 260.26–259.26 Mya (Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020). The oldest well-documented cistecephalid is likely *Cistecephalus microrhinus*, which first appears in the upper *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone of the Karoo Basin (Smith & Keyser 1995; Day & Smith 2020). This assemblage zone is bracketed by radiometric dates of 259.26 Mya and 256.25 Mya

(Day et al. 2015). The undescribed Mid-Zambezi Basin cistecephalid (see Angielczyk et al. 2019) may be older, but its age is very poorly constrained. We calibrate this node at 259 Mya based on the age of the *Lycosuchus-Eunotosaurus* Subzone.

***Kawingasaurus* + *Kembawacela*:** Our dataset includes two cistecephalid species, *Kawingasaurus fossilis* from the Usili Formation (Ruhuhu Basin, Tanzania; Cox 1972) and *Kembawacela kitchingi* from the upper Madumabisa Mudstone Formation (Luangwa Basin, Zambia; Angielczyk et al. 2019). Angielczyk et al. (2014b) correlated the Usili and upper Madumabisa Mudstone formations with the Wuchiapingian *Cistecephalus* Assemblage Zone of the Karoo Basin, but more recent work suggests that they may encompass parts of the upper *Cistecephalus* Assemblage Zone and the lower *Daptocephalus* Assemblage Zone (Angielczyk & Kammerer 2017; Angielczyk 2019; Smith 2020; Viglietti 2020; also see Sidor et al. 2010; Kammerer 2019). Therefore, the divergence between *Kawingasaurus* and *Kembawacela* must have occurred sometime between the divergence of Cistecephalidae (presumably near the Capitanian-Wuchiapingian boundary; see previous entry) and Wuchiapingian-Changhsingian boundary (i.e., near the *Cistecephalus-Daptocephalus* Assemblage one boundary; see dates in Rubidge et al. 2013; Day et al. 2015). We calibrate this node at 256 Mya, based on Rubidge et al.'s (2013) date of 256.25 Mya near the base of the Oudeberg Member of the Balfour Formation, which corresponds in part to the *Cistecephalus* Assemblage Zone.

Common ancestor of *Oudenodon* and *Lystrosaurus*: The common ancestor of *Oudenodon* and *Lystrosaurus* must have existed after the divergence of the stemward, presumably Capitanian bidentalian *Rastodon* and early cryptodonts and geikiids such as *Tropidostoma*, *Australobarbarus*, and *Bulbasaurus*, all of which are early Wuchiapingian in age (approximately the *Endothiodon* Assemblage Zone of the Karoo Basin or its likely equivalents; e.g., Kurkin, 2011; Benton et al., 2012; Kammerer & Smith 2017; Sennikov & Golubev, 2017; Kammerer & Masyutin, 2018a; Day et al. 2018; Day & Smith 2020). Radiometric dates of 259.26 Mya and 256.25 Mya bracket the *Endothiodon* Assemblage Zone in the Karoo Basin (Rubidge et al. 2013; Day et al. 2015), and we calibrate this node at 258 Mya, reflecting the presence of *Tropidostoma* and *Bulbasaurus* in the *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone (Day et al. 2018).

Common ancestor of *Aulacephalodon* and *Lystrosaurus*: The early Wuchiapingian (*Tropidostoma_Gorgonops* Subzone of the *Endothiodon* Assemblage Zone) geikiid *Bulbasaurus* is the oldest taxon that is relevant to this calibration (Kammerer & Smith 2017). We calibrate this node at 257.5 Mya to reflect the fact that *Bulbasaurus* is present in the *Tropidostoma-Gorgonops* Subzone (Day et al.

2018; Day & Smith 2020) but must have diverged after the common ancestor of *Oudenodon* and *Lystrosaurus* (see previous entry).

Common ancestor of *Lystrosaurus* and *Sangusaurus*: The fossils that are most relevant for calibrating this divergence depend strongly on which taxa are most closely related to *Lystrosaurus* and where Lystrosauridae falls relative to other dicynodontoids in dicynodont phylogeny (*Sangusaurus* dates from the Anisian–Carnian and is well nested within Triassic kannemeyeriiform dicynodonts; Angielczyk et al. 2018; Peecook et al. 2018a). Unfortunately, these are both areas where recent phylogenetic analyses of the group have differed (e.g., compare Angielczyk & Kammerer 2017; Kammerer 2018, 2019; Olroyd et al. 2018; Kammerer et al. 2019; Olivier et al. 2019; Liu 2020). For this analysis, we assumed that the topology of Kammerer (2019) is correct, but our results would not vary dramatically if other recent phylogenies were used as a reference. In this context, *Dicynodon lacerticeps*, *Dicynodon angielczyki*, *Daptocephalus huenei*, *Dinanomodon gilli*, and *Peramodon amalitzkii* all are known from faunal assemblages that are thought to span the Changhsignian-Wuchiapingian boundary (i.e., the *Cistecephalus* and *Daptocephalus* Assemblage zones of South Africa, the Usili Formation of Tanzania, the upper Madumabisa Mudstone Formation of Zambia, and the Sokolki Subassemblage of Russia; see e.g., Viglietti et al. 2016; Sennikov & Golubev 2017; Angielczyk & Kammerer 2017; Angielczyk 2019; Smith 2020; Viglietti 2020; dates from Rubidge et al. 2013). Permian occurrences of *Lystrosaurus* in China and South Africa seem to be restricted to the Changhsingian (i.e., the Guodikeng Formation and the *Lystrosaurus maccaigi-Moschorhinus* Subzone of the *Daptocephalus* Assemblage Zone; dates from Yang et al. 2010; Rubidge et. al. 2013; Gastaldo et al. 2015; also see Liu 2018; Viglietti 2020) and thus are slightly younger. Based on the radiometric dates for the Karoo Basin in Rubidge et al. (2013) and Gastaldo et al. (2015), the age of this node likely falls in the 256–253 Mya range and we calibrate it at 255 Mya.

***Lystrosaurus declivis* + *Lystrosauruus murrayi*:** *Lystrosaurus declivis* and *L. murrayi* are best known from the Karoo Basin of South Africa, where they occur in strata assigned to the earliest Triassic *Lystrosaurus declivis* Assemblage Zone (e.g., Botha & Smith 2006, 2007; Smith & Botha-Brink 2014; Botha & Smith 2020; although see Gastaldo et al. 2020 for an alternative age assessment). Phylogenetic relationships within the genus *Lystrosaurus* have been unstable in recent phylogenetic analyses (e.g., compare Angielczyk & Kammerer 2017; Kammerer 2018, 2019; Olroyd et al. 2018; Kammerer et al. 2019; Olivier et al. 2019; Liu et al. 2020). In some cases, they are reconstructed as sister taxa, meaning their divergence could be as young as earliest Triassic, whereas in other cases species with first occurrences in the late Permian fall between them, implying a late Permian divergence. Here, we place this divergence in the late Permian (equivalent to *Lystrosaurus maccaigi-Moschorhinus* subzone of the *Daptocephalus* Assemblage Zone) and calibrate it at 253 Mya based on the radiometric date of Gastaldo et al. (2015).

Gorgonopsia + Eutheriodontia: The oldest well-characterized gorgonopsian is *Eriphostoma microdon*, which first appears in the Capitanian *Diictodon-Styracocephalus* Subzone of the *Tapinocephalus* Assemblage Zone of the Karoo Basin, South Africa (Kammerer 2014; Kammerer et al. 2015; Day et al. 2018; Day & Rubidge 2020), although fragmentary material has also been reported from the Wordian *Eodicynodon* Assemblage Zone (Abdala & Rubidge 2008; Rubidge & Day 2020). More reliable Wordian to Wordian-Capitanian records exist for Eutheriodontia, in the form of therocephalians in Russia and South Africa (e.g., Abdala & Rubidge 2008; Huttenlocker & Smith 2017), with *Glanosuchus macrops* and *Ictidosaurus angusticeps* from the lower Abrahamskraal Formation (*Eodicynodon* Assemblage Zone) comprising the oldest occurrences of the clade (Huttenlocker & Smith 2017; Day et al. 2018; Rubidge & Day 2020). Dates from near the base of the Abrahamskraal Formation range from 268.5–264.6 Mya (Lanci et al. 2013). We calibrate this node at 265.5 Mya, one million years after our calibration for the divergence of Theriodontia from Anomodontia.

***'Aloposaurus' + Scylacocephalus + Lycaenops + Dixeya* + BP/1/155:** Our dataset includes five African gorgonopsians, but only one (*Lycaenops;* Wuchiapingian–Changhsingian) has been formally included in recent phylogenetic analyses of the clade. An important result of this work is the new hypothesis that African gorgonopsians form a single clade that is distinct from Russian members of the group (Bendel et al. 2018; Kammerer & Masyutin 2018a). *Eriphostoma microdon* is the oldest and most stemward well-characterized African gorgonopsian, occurring in the upper Abrahamskraal and lower Teekloof formations in the Karoo Basin (*Tapinocephalus–Endothiodon* assemblage zones; Kammerer 2014; Kammerer et al. 2015; Bendel et al. 2018; Kammerer & Masyutin 2018a; Day & Rubidge 2020; Day & Smith 2020). As such it provides a useful estimate of the maximum time of divergence for the African specimens in our dataset. The upper *Tapinocephalus* Assemblage Zone has been dated to the late Capitanian (ca. 261 Mya; Rubidge et al. 2013; Day et al. 2015; Day & Rubidge 2020), and the first appearance of *Eriphostoma* is likely slightly older than the dated strata. We calibrate this node at 262 Mya to reflect its position between the divergence of Gorgonopsia and Eutheriodontia (see previous entry) and the 261.24 Mya date for the upper *Tapinocephalus* Zone (Rubidge et al. 2013).

Therocephalia + Cynodontia: The oldest cynodont species are *Charassognathus gracilis* and *Abdalodon diastematicus*, both of which are known from the Hoedemaker Member of the Teekloof Formation in the Karoo Basin (lower Wuchiapingian *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone; Botha et al. 2007; Kammerer 2016b; age from Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020). Therocephalia has a deeper fossil record, with *Glanosuchus macrops* and *Ictidosaurus angusticeps* from the lower Abrahamskraal Formation (*Eodicynodon* Assemblage Zone) comprising the oldest occurrences of the clade (Huttenlocker & Smith 2017; Day et al. 2018; Rubidge & Day 2020). Lanci et al.

(2013) dated the lower Abrahamskraal Formation to 268.5–264.6 Mya, and we calibrate this node at 265 Mya to reflect these ages and also the fact that this divergence must post-date the divergence of Gorgonopsia from Eutheriodontia (see above). Our calibration is notably younger than statistical estimates of the root age for Cynodontia presented by Lukic-Walther et al. (2019), which clustered near the Cisuralian-Guadalupian boundary. However, our calibration is consistent with the known therapsid fossil record.

Akidnognathidae + Baurioidea: The therocephalians in our dataset comprise members of two major clades, Akidnognathidae (*Euchambersia*, *Olivierosuchus*) and Baurioidea (*Ictidosuchoides, Mupashi, Choerosaurus, Microgomphodon*) (e.g., Huttenlocker & Smith 2017; Liu & Abdala 2017; Kammerer & Masyutin 2018b). Of these two clades, baurioids appear earlier in the fossil record, near the Capitanian-Wuchiapingian boundary (e.g., cf. *Ictidosuchoides* in the South African *Lycosuchus-Eunotosaurus* Subzone of the *Endothiodon* Assemblage Zone, *Karenites* in the Russian Kotelnich Subassemblage; Huttenlocker & Smith 2017; Day & Smith 2020). The base of the *Lycosuchus-Eunotosaurus* Subzone is constrained by a radiometric date of 260.26 Mya at the base of the Poortjie Member of the Teekloof Formation (Day & Smith 2020), and we calibrate this node at 260.26 Mya.

***Olivierosuchus + Euchambersia*:** Although it is relatively well-nested within Akidnognathidae, *Euchambersia mirabilis* is one of the first members of the clade to appear in the fossil record (Wuchiapingian *Cistecephalus* Assemblage Zone, Karoo Basin; Benoit et al. 2017a; Huttenlocker & Smith 2017; age from Rubidge et al. 2013). Radiometric dates and revisions to Karoo litho- and biostratigraphy suggest that the base of the *Cistecephalus* Assemblage zone is older the 256 Mya and that the boundary between the *Cistecephalus* and *Daptocephalus* assemblage zones is slightly older than 255 Mya (Rubidge et al. 2013; Viglietti et al. 2016, 2017, 2018a; Smith 2020; Viglietti 2020). Therefore, we calibrate this node at 257 Mya.

Common ancestor of *Ictidosuchoides* and *Microgomphodon*: *Ictidosuchoides longiceps* is one of the oldest baurioids, with well characterized records extending back to the Wuchiapingian *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone of South Africa and a potential occurrence in the Capitanian-Wuchiapingian *Lycosuchus-Eunotosaurus* Subzone of the *Endothiodon* Assemblage Zone (Huttenlocker & Smith 2017; Day et al. 2018; Day & Smith 2020). *Karenites ornamentatus* from the Kotelnich Subassemblage in Russia is likely of similar age (Huttenlocker & Smith 2017; Sennikov & Golubev 2017; Kammerer & Masyutin 2018b). The *Lycosuchus-Eunotosaurus* Subzone is bracketed by radiometric dates of 260.26 Mya and 259.25 Mya (Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020), and we calibrate this node at 259.5 Mya.

Common ancestor of *Mupashi* and *Microgomphodon*: *Mupashi migrator*, from the upper Madumabisa Mudstone Formation of Zambia, is commonly reconstructed as the sister taxon of *Karenites ornamentatus* from Russia (Huttenlocker & Sidor 2016; Huttenlocker & Smith 2017; Liu & Abdala 2017; Kammerer & Masyutin 2018b). *Karenites* occurs in the Kotelnich Subassemblage, which is generally considered to be close to the Capitanian-Wuchiapingian boundary in age (e.g., Kurkin 2011; Benton et al. 2012; Kammerer & Masyutin 2018; Sennikov & Golubev 2018), although radiometric dates do not yet exist for the assemblage. We calibrate the age of this node as 259 Mya, making it slightly younger than the age of the common ancestor of *Ictidosuchoides* and *Microgomphodon*.

Common ancestor of *Choerosaurus* and *Microgomphodon*: *Choerosaurus dejageri* occurs in the Wuchiapingian *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone of the Karoo Basin (Benoit et al. 2016a; Huttenlocker & Smith 2017; Smith & Day 2020), making it one of the oldest baurioids. By contrast *Microgomphodon oligocynus* dates to the Middle–?Late Triassic *Langbergia-Garjainia* Subzone of the *Cynognathus* Assemblage Zone (Abdala et al. 2014; Huttenlocker & Smith 2017; Hancox et al. 2020). The *Tropidostoma-Gorgonops* Subzone is thought to be between 258 Mya and 256.8 Mya in age (Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020). We calibrate this node at 258 Mya, one million years after the age of the common ancestor of *Mupashi* and *Microgomphodon*.

Common ancestor of *Procynosuchus* and mammals: There is uncertainty about the exact phylogenetic placement of *Procynosuchus* (e.g., compare Botha et al. 2007; Kammerer 2016b; Van den Brandt & Abdala 2018; Abdala et al. 2019; Huttenlocker & Sidor 2020; supertree results of Lukic-Walther et al. 2019), which impacts on which taxa are most useful for calibrating this node. If *Procynosuchus* (or a clade comprised of *Procynosuchus* and *Dvinia*) is the most stemward known cynodont lineage, then the presence of *Charassognathus* and *Abdalodon* in the early Wuchiapingian *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone of South Africa (Day & Smith 2020) implies that *Procynosuchus* must have diverged by that time. Alternatively, if *Procynosuchus* is in a more nested position than *Charassognathus* and/or *Abdalodon* (or Charassognathidae *sensu* Huttenlocker & Sidor 2020), then its divergence time might be closer to its actual first appearance near the boundary of the *Cistecephalus* and *Daptocephalus* assemblage zones (e.g., Huttenlocker et al. 2011; Kammerer 2016b, Viglietti et al. 2016; Smith 2020; Viglietti 2020). For the purposes of this analysis, we assumed that *Procynosuchus* falls stemward of *Charassognathus* and *Abdalodon*, and we calibrate this node at 259 Mya (i.e., near the boundary between the boundary between the *Lycosuchus-Eunotosaurus* and *Tropidostoma-Gorgonops* subzones of the *Endothiodon* Assemblage Zone in the Karoo Basin; dates from Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020).

Epicynodontia: *Cynosaurus suppostus* occurs in the upper Wuchiapingian-Changhsingian *Cistecephalus* and *Daptocephalus* assemblage zones of South Africa (Van den Brandt & Abdala 2018; Smith 2020; Viglietti 2020) and provides a calibration point for this node. The base of the *Cistecephalus* Assemblage zone dates to 256.6 Mya (Rubidge et al. 2013; Viglietti et al. 2016, 2017, 2018a; Smith 2020), so we calibrate this node at 258 Mya, in part to accommodate additional divergences that must have occurred in the latest Permian (see next two entries)

Common ancestor of *Galesaurus* and mammals: *Galesaurus planiceps* is a well characterized cynodont known from the Early Triassic *Lystrosaurus declivis* Assemblage Zone of South Africa (e.g., Jasinoski & Abdala 2017a, b; Butler et al. 2018; Pusch et al. 2019; Botha & Smith 2020). However, the presence of the phylogenetically more deeply-nested cynodonts *Nanictosaurus* and *Vetusodon* in the Changhsingian *Daptocephalus* Assemblage Zone of the Karoo Basin (e.g., Abdala et al. 2019; Viglietti 2020) implies that the lineage including *Galesaurus* must have diverged in the Permian as well. The base of the *Daptocephalus* Assemblage Zone is approximately 255 Mya (Rubidge et al. 2013; Viglietti 2020), and we use 257 Mya as a calibration point for this node to accommodate additional divergences that must have occurred in the latest Permian (see next entry).

Common ancestor of *Thrinaxodon* and mammals: *Thrinaxodon liorhinus* is unquestionably the most thoroughly studied non-mammalian cynodont, and it has a well-documented stratigraphic range in the Early Triassic Palingkloof Member of the Balfour Formation and the Katberg Formation in the Karoo Basin (e.g., Botha & Smith 2006, 2020) as well as the lower Fremouw Formation in Antarctica (e.g., Kitching et al. 1972; Colbert & Kitching 1977; Hammer 1990; Peecook et al. 2018b). As in the case of *Galesaurus* (see previous entry), the occurrence of the phylogenetically more-nested cynodonts *Nanictosaurus* and *Vetusodon* in the *Daptocephalus* Assemblage Zone (e.g., Abdala et al. 2019; Viglietti 2020) implies that the lineage including *Thrinaxodon* must have diverged in the Permian as well. We calibrate this node at 256 Mya, reflecting the fact that the base of the *Daptocephalus* Assemblage Zone is slightly younger than this age (Rubidge et al. 2013; Viglietti et al. 2016, 2017, 2018a; Viglietti 2020).

Eucynodontia: A consistent feature of analyses of cynodont phylogeny is the division of 'higher' cynodonts into two major clades, the extinct Cynognathia and Probainognathia, whose extant representatives are mammals (e.g., see supertree in Lukic-Walther et al. 2019), which are united within Eucynodontia. The strata hosting the oldest occurrences of eucynodonts, such as the lower Burgersdorp Formation of the Karoo Basin (*Laingbergia-Garjainia* Subzone of the *Cynognathus* Assemblage Zone), have typically been thought to be Olenekian in age (e.g., Rubidge 2005; Lucas 2010; Schneider et al. 2019; Hancox et al. 2020 although see further discussion in the following entry). However, depending on the identity of its sister taxon, the

eucynodont lineage must have diverged by the Early Triassic (e.g., if *Playtcraniellus* is the sister taxon; Abdala 2007) or the latest Permian (e.g., if *Vetusodon* is the sister taxon; Abdala et al. 2019). For this analysis, we assumed the latter hypothesis is correct, and used Gastaldo et al.'s (2015) radiometric date of 253.48 Mya for the *Lystrosaurus maccaigi-Moschorhinus* subzone of the *Daptocephalus* Assemblage Zone to calibrate this node.

Common ancestor of Trirachodontidae and *Massetognathus*: The most stemward cynognathian lineage in our dataset is Trirachodontidae. In addition to *Trirachodon* itself, Trirachodontidae also includes the genera *Langbergia* and *Cricodon* (Sidor & Hopson 2018), and possibly *Beishanodon* and *Sinognathus* from China (Liu & Abdala 2014). We treat Trirachodontidae as a clade in our phylogeny, but recently Hendrickx et al. (2020) proposed that the trirachodontids may instead be a paraphyletic assemblage on the stem leading to Traversodontidae. *Langbergia,* from the lower Burgersdorp Formation (*Langbergia-Garjainia* Subzone of the *Cynognathus* Assemblage Zone; Hancox et al. 2020) of the Karoo Basin is the stratigraphically lowest-occurring trirachodontid. Traditionally. one record of *Trirachodon berryi* was suggested to have originated in this subzone (Abdala et al. 2006), but Hancox et al.'s (2020) redefined *Cynognathus* zone biostratigraphy would place this specimen in the overlying *Trirachodon-Kannemeyeria* Subzone. *Langbergia* also occurs stratigraphically below other stemward cynognathians such as *Cynognathus* and *Diademodon* (Neveling 2004), making it the most relevant record for calibrating this divergence. As noted in the previous entry, the *Laingbergia-Garjainia* Subzone has generally been regarded as Olenekian in age, although it has not been radiometrically dated directly. Recent radiometric dates of strata that are biostratigraphically-correlated with the *Cynognathus* Assemblage Zone have raised the possibility that it parts of it might be substantially younger than previously thought (e.g., Ottone et al. 2014; Marsicano et al. 2016; also see discussion in Martinelli et al. 2017a; Peecook et al. 2018a; Schneider et al. 2019). Detrital zircon crystals from the lower part of the underlying Katberg Formation with a minimum age of 250 ± 5 Mya (Viglietti et al. 2018b) also are consistent with a younger age for the *Cynognathus* Assemblage Zone. However, the wide error range on the latter date and the existence of dates for other biostratigraphically-correlated strata more in accordance with the traditional hypothesis (Liu et al. 2018) suggest that further work on the problem is needed. For the purposes of this paper, we have maintained the assumption that subzone A of the *Cynognathus* Assemblage Zone is Olenekian (Hancox et al. 2020). We calibrate this node at 248 Mya, in the late Olenekian, to accommodate the possibility of an Olenekian age for at least part of the underlying Katberg Formation. This is slightly younger than the statistically estimated root age of 249.34 Mya for Cynognathia presented by Lukic-Walther et al. (2019).

Common ancestor of *Trirachodon* and *Cricodon*: *Trirachodon berryi* is an index fossil for the *Trirachodon-Kannemeyeria* Subzone of the of the *Cynognathus* Assemblage Zone of the South Africa

Karoo Basin (Hancox et al. 2020). The two species of *Cricodon* recognized by Hopson & Sidor (2018) include occurrences in the *Trirachodon-Kannemeyeria* Subzone of the *Cynognathus* Assemblage Zone (*C. kannemeyeri*) as well as the overlying *Cricodon-Ufudocyclops* Subzone, the Ntawere Formation of Zambia, and the Lifua Member of the Manda Beds of Tanzania (all *C. metabolus*; Hopson & Sidor 2018; Peecook et al. 2018a; Hancox et al. 2020). As noted in the previous entry, the age of the *Cynognathus* Assemblage Zone and many biostratigraphically-correlated assemblages has been the subject of recent debate. We calibrate this divergence at 247.5 Mya (latest Olenekian), slightly older than the traditionally-assumed early Anisian age for the *Trirachodon-Kannemeyeria* Subzone (e.g., Hancox et al. 2020, although see discussion above).

Common ancestor of *Scalenodon* and *Massetognathus*: Traditionally, the Lifua Member of the Manda Beds, which hosts the cynodont *Scalenodon angustifrons* (e.g., Liu & Abdala 2014), has been considered to be Anisian–early Ladinian in age (e.g., Lucas 1998, 2010; Rubidge 2005), although no radiometric dates are available for the Lifua Member or nearby biostratigraphically-correlated strata in southern Africa (e.g., Ntawere Formation of Zambia, Burgersdorp Formation of South Africa). Recently published radiometric dates from South American strata (Philipp et al. 2013, 2018; Ottone et al. 2014; Marsicano et al. 2016; Langer et al. 2018) and the increasing number of cynodonts (including *Scalenodon* itself) and other taxa shared between basins in South America and southern Africa (e.g., Abdala et al. 2013; Martinelli et al. 2017a; Melo et al. 2017; Peecook et al. 2018a) have raised the strong possibility that *Scalenodon* is no older than late Ladinian and could be early Carnian in age. A Ladinian–early Carnian age for *Massetognathus* is relatively certain (e.g., Marsicano et al. 2016; Schmitt et al. 2019). We calibrate this node at 239 Mya, reflecting the fact that the *Dinodontosaurus* Assemblage Zone of Brazil (Santa Maria Supersequence) is likely older than 236–237 Mya and may fall within the late Ladinian (e.g., Martinelli et al. 2017a; Melo et al. 2017; Philipp et al. 2018; Schmitt et al. 2019), as well as the need to accommodate the divergence between *Massetognathus* and *Luangwa*, both of which also are present in the *Dinodontosaurus* Assemblage Zone of Brazil (see next entry).

Common ancestor of *Luangwa* and *Massetognathus*: As its name suggests, *Luangwa* was first discovered in the upper portion of the Triassic Ntawere Formation of the Luangwa Basin, Zambia (Brink 1963), but it was subsequently found in Brazil, Namibia, and Tanzania (Abdala & Sa-Teixeira 2004; Abdala & Smith 2009; Martinelli et al. 2017a; Peecook et al. 2018a). The upper Ntawere Formation was traditionally considered to be Anisian in age (Rubidge 2005), but the presence of *Luangwa* in Brazilian strata radiometrically dated to near the Ladinian–Carnian boundary strongly suggest that the Ntawere Formation is younger than previously appreciated (Peecook et al. 2018a; also see previous entry). We calibrate this divergence at 238 Mya, reflecting the fact that the *Dinodontosaurus* Assemblage zone of Brazil (Santa

Maria Supersequence) is likely older than 236–237 Mya (e.g., Martinelli et al. 2017a; Melo et al. 2017; Philipp et al. 2018; Schmitt et al. 2019).

Common ancestor of *Lumkuia* and mammals: There has been uncertainty over the last two decades as to the identity of the most stemward member of Probainognathia, the cynodont clade that includes mammals as its extant representatives (e.g., compare Hopson & Kitching 2001; Liu & Olsen 2010; Martínez et al. 2013a; Ruta et al. 2013; Martinelli et al. 2017a, b, c; Stefanello et al. 2018; Lukic-Walther et al. 2019; Wallace et al. 2019). Here, we follow the hypothesis that *Lumkuia fuzzi* is the most stemward probainognathian, which appears to be an emerging consensus among recent analyses (also see Benoit et al. 2019). The only known specimen of *Lumkuia* was collected from the *Trirachodon-Kannemeyeria* Subzone of the *Cynognathus* Assemblage Zone (Burgersdorp Formation, Karoo Basin) (Hopson & Kitching 2001; Hancox et al. 2020). As noted in several entries above, the *Cynognathus* Assemblage zone was traditionally regarded as Anisian (Middle Triassic) in age based on biostratigraphical correlations, but new radiometric dates from South America have raised the possibility of an age as young as Ladinian–Carnian (Middle–Late Triassic). Therefore, the divergence between *Lumkuia* and the lineage including other probainognathians could be as early as Olenekian (based on the presence of cynognathians in the record at this time and assuming a 'traditional' age for the *Cynognathus* zone; see entry for the common ancestor of Trirachodontidae and *Massetognathus* above), or as late as Ladinian (assuming an age no older than Ladinian for most of the 'classic' African and South American cynodont-bearing basins) (e.g., see discussion in Martinelli et al. 2017a). We calibrate this node at 244 Mya (i.e., the midpoint of the Olenekian–Ladinian age range). The difference between this calibration and that for the divergence between *Trirachodon* and *Massetognathus* implies the existence of a probainognathian ghost lineage following the divergence of Cynognathia and Probainognathia. Our calibration is also notably younger than the statistically estimated root age of 251.9 Mya for Probainognathia presented by Lukic-Walther et al. (2019).

Common ancestor of *Chiniquodon* and mammals: Following the taxonomic revision of Abdala & Giannini (2002), *Chiniquodon* is one of the stratigraphically and geographically widest-ranging early probainognathians (see reviews in Martinelli et al. 2017a; Abdala & Gaetano 2018). The oldest records of the genus are from the *Dinodontosaurus* Assemblage Zone (Santa Maria Supersequence) of Brazil and the Chañares Formation of Argentina (e.g., Abdala & Giannini 2002; Abdala & Gaetano 2018; Schmitt et al. 2019), which are late Ladian–early Carnian in age (Marsicano et al. 2016; Philipp et al. 2018). We calibrate this node at 238 Mya, reflecting the stratigraphic position of the *Dinodontosaurus* Assemblage Zone (Pinherios-Chiniquá sequence) below Philipp et al.'s (2018) radiometric date of 237 Mya for the base of the overlying *Santacruzodon* Assemblage Zone (Santa Cruz sequence).

Common ancestor of *Riograndia* and mammals: The ictidosaur (*sensu* Martinelli & Rougier 2007) *Riograndia guaibensis* is known from the *Riograndia* Assemblage Zone (Candelária Sequence, Santa Maria Supersequence) of the Paraná Basin, Brazil (e.g., Soares et al. 2011). Rocks from this assemblage zone were recently dated at 225.42 Mya (Norian) by Langer et al. (2018). This is almost certainly an underestimate of the age of this divergence, and a more likely estimate of the age is calibrated by the presence of the more stemward prozostrodonts *Prozostrodon*, *Therioherpeton*, and *Alemoatherium* in the underlying *Hyperodapedon* Assemblage Zone (e.g., Martinelli et al. 2016, 2017b; Pacheco et al. 2017), which is thought to be as old as 231.4 Mya (Carnian) based on biostratigraphical correlation with the radiometrically-dated Ischigualasto Formation of Argentina (Martínez et al. 2013b). We calibrate this divergence at 234 Mya, in part to accommodate the presumably Carnian divergence of tritylodontids, brasilodontids, and mammaliaforms (see next entry).

Common ancestor of Tritylodontidae + *Pseudotherium* and mammals:

Triylodontids are known primarily from the Jurassic and Cretaceous (e.g., Abdala & Gaetano 2018), but *Oligokyphus* has been reported from the Late Triassic (Rhaetian) of Nova Scotia and Germany (Fedak et al. 2015). However, there are older Triassic fossils that are relevant to calibrating this divergence, but the question of which specific taxa are most important depends strongly on the phylogenetic relationships assumed for tritylodontids. Two main phylogenetic positions have been proposed for tritylodontids. One posits that they are derived cynognathians (e.g., Hopson & Kitching 2001; Sues & Jenkins 2006; Bonaparte & Crompton 2017; Sidor & Hopson 2018), whereas the other places them within Probainognathia, close to the base of Mammaliaformes (e.g., Abdala 2007; Liu & Olsen 2010; Ruta et al. 2013; Martinelli et al. 2016; Lukic-Walther et al. 2019; Wallace et al. 2019). Among phylogenies that place tritylodontids within Probainognathia, there also is debate concerning the branching order of tritylodontids, trithelodontids, and brasilodontids relative to Mammaliaformes (e.g., compare Abdala 2007; Liu & Olsen 2010; Ruta et al. 2013; Martinelli et al. 2016, 2017b; Wallace et al. 2019; also see summary supertree of Lukic-Walther et al. 2019). Resolving these questions is beyond the scope of this analysis. Here, we assumed the branching order: (Trithelodontidae (Tritylodontidae (Brasilodontidae, Mammaliaformes))), consistent with the analyses of Liu & Olsen (2010), Martinelli et al. (2016), and Wallace et al. (2019).

Brasilodon and other taxa of potential relevance to the age of this correlation (e.g., *Botucaratherium*) co-occur with *Riograndia* in the Norian *Riograndia* Assemblage Zone of Brazil (e.g., Martinelli et al. 2016, 2017b). *Adelobasileus* from the Tecovas Formation of Texas (Lucas & Luo 1993), which is likely more closely related to mammals than *Brasilodon*, was shown as having an older occurrence by Martinelli et al. (2017b; also see e.g., Abdala & Gaetano 2018) that potentially implied a Carnian age

for this divergence. However, recent work suggests a younger Norian age for *Adelobasileus* (Sarigül 2017). *Tikitherium* and *Gondwanadon* from the Tiki Formation of India (Datta & Das 1996; Datta 2005) have been suggested to be members of Mammaliaformes (e.g., Luo & Martin 2007; Debuysschere et al. 2015; see reviews in Abdala & Gaetano 2018; Martin 2018), and the Tiki Formation typically is biostratigraphically correlated with the Carnian *Hyperodapedon* Assemblage Zone and the Ischigualasto Formation (e.g., Ray et al. 2016; Bhat et al. 2018). If this correlation is correct, it would indicate that divergences between trithelodontids, tritylodontids, and mammaliforms all occurred no later than the late Carnian. The recent discovery of *Pseudotherium* in the Ischigualasto Formation (radiometrically dated to 231.4– 225.9 Mya; Martínez et al. 2013b), which may be the sister taxon of Tritylodontidae (Wallace et al. 2019; although see below), provides additional support for a Carnian divergence among these lineages. Therefore, we calibrate this divergence at 233 Mya, in part to accommodate the divergences between *Brasilodon* and mammals and between *Pseudotherium* and tritylodontids (see below).

Common ancestor of *Pseudotherium* and Tritylodontidae: Tritylodontids are primarily a post-Triassic mammaliamorph radiation, although their oldest records extend into the Late Triassic (see previous and following entries). The recently described species *Pseudotherium argentinus* from the Carnian Ischigualasto Formation of Argentina may be the sister taxon of Tritylodontidae, although branch support for this hypothesis is weak (Wallace et al. 2019) and ongoing research suggests an alternative phylogenetic placement (A.G. Martinelli, unpublished data). For simplicity, we assumed that *Pseudotherium* is a stem tritylodont because this is the primary hypothesis that has been presented in the literature at this time. The Ischigualasto Formation has been dated radiometrically to 231.4–225.9 Mya (Martínez et al. 2013b), so we calibrate this divergence at 232 Mya.

Common ancestor of *Tritylodon* and *Oligokyphus*: *Oligokyphus* is generally considered to be the most stemward tritylodontid (see review in Velazco et al. 2017) and is the only member of the clade that is well documented from the Late Triassic (Fedak et al. 2015). Putative tritylodontid postcrania reported from the Los Colorados Formation of Argentina, which is Nornian in age (Kent et al. 2014), would constitute an even older record for the group, but a recent re-assessment found that they could only be identified as an indeterminate non-mammaliaform cynodont (Martinelli & Soares 2016; Gaetano et al. 2017). If the possible stem-tritylodontid status of *Pseudotherium* is assumed to be correct (see discussion in Wallace et al. 2019 and above) it provides an even older, potentially Carnian upper bound for divergences among tritylodontids. The McCoy Brook Formation record of *Oligokyphus* is approximately 201.45 Mya (Fedak et al. 2015), providing a minimum age for this divergence; other material of *Oligokyphus* is of likely Early Jurassic in age or has less precise age constraints (e.g. Clemens & Martins 2014; Whiteside et al. 2016). We calibrate this divergence at 204.9 Mya, the midpoint of the Rhaetian Stage.

Common ancestor of *Brasilodon* and mammals: *Brasilodon quadrangularis* occurs in strata assigned to the *Riograndia* Assemblage Zone of the Candelária Sequence (Santa Maria Supersequence) in the Paraná Basin, Brazil (e.g., Bonaparte et al. 2003; Martinelli et al. 2016, 2017b; Guignard et al. 2019). Rocks from this assemblage zone were recently dated at 225.42 Mya (Norian) by Langer et al. (2018). As discussed above (see entry for divergence between Tritylodontidae + *Pseudotherium* and mammals), however, divergences between trithelodontids, tritylodontids, and mammaliforms all occurred no later than the late Carnian. Therefore, we calibrate this divergence at 232 Mya, slightly older than the radiometric date for the lower Ischigualasto Formation (Martínez et al. 2013b).

Common ancestor of Morganucodonta and mammals: Morganucodonts are best known from the latest Triassic and Early Jurassic, but Norian records of the clade also exist (Debuysschere et al. 2015). If *Gondwanadon* is a morganucodont (Datta & Das 1996; Kielan Jaworowska et al. 2004; Debuysschere et al. 2015) it would push the first appearance of the clade into the Carnian, based on the biostratigraphical correlation of the Tiki Formation with the Carnian *Hyperodapedon* Assemblage Zone and the Ischigualasto Formation (e.g., Ray et al. 2016; Bhat et al. 2018). The presence of the putative docodont or docodont relative *Tikitherium* in the Tiki Formation (Datta 2005; Luo & Martin 2007; Luo et al. 2015; Panciroli et al. 2019) also implies that morganucodonts must have diverged by the Carnian because Docodonta is consistently recovered crownward of Morganucodonta in phylogenetic analyses (e.g., see reviews in Abdala & Gaetano 2018; Martin 2018). We calibrate this divergence at 230.5 Mya, slightly younger than the radiometric date of 231.4 Mya for the Ischigualasto Formation, but still within the Carnian.

Common ancestor of *Morganucodon watsoni and Morganucodon oehleri:* Despite early occurrences of morganucodontans as early as the late early Rhaetian from the Howell quarry in Wales (Whiteside et al. 2016), material that can be confidently attributed to *Morganucodon watsoni* is only present at the Triassic–Jurassic transition in the St. Brides Island palaeocommunity (Whiteside et al. 2016). *Morganucodon oehleri* is known from the Zhangjia'ao Member (*sensu* Fang et al. 2000) of the Lufeng Formation of China (Luo & Wu 1994), which is considered to be Sinemurian in age based on biostratigraphical comparisons (e.g., Luo & Wu 1994; Sullivan et al. 2013). However, as noted above, there is considerable uncertainty about phylogenetic relationships within Morganucodonta, which complicates identifying the correct calibration point for this divergence. Therefore, we use the age of the oldest species of *Morganucodon*, *Morganucodon peyeri* from the Hallau locality in Switzerland, which is dated as late Norian–early Rhaetian (Whiteside et al. 2017). We calibrate this divergence at 207 Mya, the latest possible occurrence of *M. peyeri*.

Common ancestor of *Haldanodon* and mammals: *Haldanodon exspectatus* is a well-studied docodont known from the Alcobaça Formation of the Guimarota coal mine of Portugal (e.g., Lillegraven & Krusat

1991; Martin 2005, 2018; Ruf et al. 2013), which is considered Kimmeridgian in age (Schudack 2000a, b). *Haldanodon* is well-nested within Docodonta (Meng et al. 2015; Panciroli et al. 2019), and the oldest members of the clade date back to the Bathonian (Panciroli et al. 2019). However, the presence of the putative docodont or docodont relative *Tikitherium* in the Carnian Tiki Formation (Datta 2005; Luo & Martin 2007; Luo et al. 2015; Panciroli et al. 2019) implies a long ghost lineage for Docodonta. Therefore, we calibrate this node at 230 Mya (Carnian), slightly younger than the divergence of Morganucodonta (see above).

Common ancestor of *Morganucodon* and *Megazostrodon*: Species of *Morganucodon* and *Megazostrodon* span the Triassic–Jurassic boundary, but the oldest records of both genera are in the Rhaetian (Debuysschere et al. 2015), providing a minimum age for their divergence. Potential records of *Hallutherium* and *Brachyzostrodon* in the Norian of Greenland and Poland (Jenkins et al. 1994; Świło et al. 2014; Debuysschere et al. 2015), and the presence of *Gondwanadon* in the Carnian of India (Datta & Das 1996), could imply an even older divergence time for *Morganucodon* and *Megazostrodon*, depending on the pattern of phylogenetic relationships within Morganucodonta. However, most mammaliaform phylogenetic analyses include few morganucodonts, leaving relationships within the clade (and even the monophyly of the clade) in question (e.g., Rougier et al. 2007; Gaetano & Rougier 2012; Zhou et al. 2013; Close et al. 2015; Luo et al. 2015; Meng et al. 2015; Huttenlocker et al. 2018; Panciroli et al. 2019; also see Abdala & Gaetano 2018). Given this uncertainty, we calibrate this node at the base of the Rhaetian Stage, 208.5 Mya.

Common ancestor of *Hadrocodium* and mammals: *Hadrocodium wui* is known from the Early Jurassic (Sinemurian) Lufeng Formation of China (Luo et al. 2001). The maximum age of this divergence is calibrated by the Carnian occurrence of the potential docodont *Tikitherium*, given that docodonts fall stemward of *Hadrocodium* in most recent mammaliaform phylogenies (e.g., Zhou et al. 2013; Close et al. 2015; Luo et al., 2015; Meng et al. 2015; Huttenlocker et al. 2018). The minimum age of this divergence depends strongly on whether the clade Haramiyida falls crownward or stemward of *Hadrocodium* (e.g., compare results of Zhou et al. 2013; Close et al. 2015; Luo et al., 2015; Meng et al. 2015; Huttenlocker et al. 2018). If the former option is correct, Norian–Rhaetian occurrences of haramiyidans such as *Haramiyavia* and *Thomasia* (Hahn 1973; Jenkins et al. 1997; Clemmensen et al. 2016) would imply that the lineage including *Hadrocodium* must have diverged by the Norian, whereas the latter would suggest a minimum divergence time closer to the Early Jurassic (although note that the analysis of Close et al. 2015 still implies a Triassic age for this node even when *Hadrocodium* is in a more crownward position). Resolving this phylogenetic problem is beyond the scope of this analysis. We calibrate this node at 214 Mya (late Norian), which is consistent with the case where haramiyidans are crownward of *Hadrocodium*

and close to the divergence time estimate of Close et al.'s (2015) for a topology in which *Hadrocodium* occupied a more crownward position.

Common ancestor of *Dryolestes* and Theria: *Dryolestes* and its closest relatives (Dryolestidae) are cladotherian mammals known from the Jurassic and Cretaceous periods. *Dryolestes* itself is best known from the Late Jurassic (Kimmeridgian–Tithonian) of Europe and North America, but other records of Dryolestidae extend back into the Middle Jurassic (see reviews in Kielan-Jaworowska et al. 2004; Martin 2018), with the Bathonian *Anthracolestes sergeii* representing the oldest well-characterized member of the clade (Averianov et al. 2014). Additional evidence for a divergence of dryolestids by the Bathonian comes from *Amphitherium*, a member of Zatheria (the clade including therian mammals) (e.g., Averianov et al. 2013; Close et al. 2015; Luo et al. 2015; Huttenlocker et al. 2018), which is known from the lower Bathonian Tayton Limestone Formation of England (Butler & Clemens 2001). We calibrate this node at the base of the Bathonian (168.3 Mya).

Common ancestor of Captorhinidae and Diapsida: Captorhinidae is a clade of important Permian eureptiles, but only one species is known from the Carboniferous, the Gzhelian *Euconcordia cunninghami* from the Hamilton Quarry of Kansas (Müller & Reisz 2005; Reisz et al. 2016). However, several older (Bashkirian to Moscovian) taxa, including *Hylonomus*, *Brouffia*, and *Paleothyris* likely either fall on the captorhinid stem or diapsid stem, but have been difficult to place phylogenetically (Müller & Reisz 2006). *Hylonomus* is the oldest of these taxa and occurs in the Joggins Formation of Nova Scotia, whose estimated age range is about 319–310 Mya (Reisz & Müller 2004; van Tuinen & Hadly 2004; Benton et al. 2015). We calibrate this node at 319 Mya, which is slightly older than Benton et al.'s (2015) recommendation of 318 Mya for the age of the synapsid-sauropsid divergence.

***Captorhinus* and *Labidosaurus*:** *Labidosaurus* is a derived captorhinid that occurs in the Kungurian Arroyo Formation of Texas (*sensu* Lucas 2006; equivalent to the lower Clear Fork Group of Hentz 1988) (Dodick & Modesto 1995), whereas the first occurrence of the genus *Captorhinus* is represented by records of *Captorhinus laticeps* from the Artinskian Petrolia Formation of Texas (Heaton 1979; taxonomy following Dodick & Modesto 1995; stratigraphy based on Hentz 1988; Lucas 2006; Schneider et al. 2019). Therefore, we calibrate this divergence at 286 Mya, in the mid-Artinskian.

Common ancestor of *Youngina* and Sauria: *Youngina capensis* is a well-studied stemward diapsid, but recent phylogenetic analyses have differed on whether it represents a distinct lineage or a member of a larger subclade (Younginiformes), as well as the membership of that subclade when present (e.g., Bickelmann et al. 2009; Reisz et al. 2011; Ezcurra et al. 2014; Turner et al. 2017; Simões et al. 2018). For

simplicity, we treated *Youngina* as an individual lineage although the inclusion of other putative younginiforms would not result in a dramatic age increase. The earliest occurrence of *Youngina* is in strata of the Hoedemaker Member (Teekloof Formation; Karoo Basin) that are assigned to the *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone (Smith & Evans 1996; Day & Smith 2020). This assemblage zone is bracketed by radiometric dates of 259.26 Mya and 256.25 Mya (Day et al. 2015; Day & Smith 2020). However, this occurrence postdates the more crownward divergence of Testudines and Archosauromorpha, which occurred no later than the late Capitanian. Therefore, we calibrate this node at 263 Mya to accommodate the divergence of Testudines and Archosauromorpha.

Common ancestor of Testudines and Archosauromorpha: Although our sample does not include any fossil turtles, the age of this node has implications for other calibrations in the diapsid portion of the tree that require some discussion. The phylogenetic position of turtles within Sauropsida has been controversial (see review in Schoch & Sues 2019). A consensus is emerging that turtles are members of Diapsida, but there is still debate over where they fall relative to lepidosauromorphs, archosaurs, and extinct lineages such as sauropterygians (e.g., Lyson et al. 2010; Field et al. 2014; Bever et al. 2015; Crawford et al. 2015; Irisarri et al. 2017; Schoch & Sues 2015, 2018; Li et al. 2018; Simões et al. 2018). Here, we follow the hypothesis that turtles are more closely related to archosaurs than lepidosaurs.

The Permian archosauromorph fossil record is extremely sparse and of limited use for calibrating this node (e.g., Ezcurra 2016; also next entry). Very little is known of the Permian portion of the turtle lineage but an increasing amount of data suggests that *Eunotosaurus africanus* from the Karoo Basin of South Africa is the oldest known stem turtle (e.g., Lyson et al. 2010, 2013, 2014, 2016; Bever et al. 2015, 2016). Day (2013) and Day et al. (2013) reported that *Eunotosaurus* occurs in the upper portion of the Abhrahamskraal Formation and the Poortjie Member of the Teekloof Formation in that Karoo Basin of South Africa (equivalent to the *Eosimops-Glanosuchus* and *Diictodon-Styracocephalus* subzones of the *Tapinocepahlus* Assemblage Zone and the *Lycosuchus-Eunotosaurus* Subzone of the *Endothiodon* Assemblage Zones; Day & Rubidge 2020; Day & Smith 2020). Occurrences of *Eunotosaurus* in the upper Abrahamskraal formation are close to the dates of 261.24–260.26 Mya (Capitanian) for the upper *Tapinocephalus* Assemblage Zone (Rubidge et al. 2013; Day et al. 2015). We calibrate this node at 262 Mya (but see discussion in Marjanović 2019).

Common ancestor of *Prolacerta* and Archosauriformes: *Prolacerta broomi* and its sister taxon *Kadimakara australiensis* fall near the base of Archosauriformes within Archosauromorpha, and they occur in Induan (Early Triassic) strata in Antarctica, Australia, and South Africa (e.g., Ezcurra 2016; Spiekman 2018). However, the minimum age of this node is calibrated by the presence of the archosauriform

Archosaurus rossicus in the upper Changhsingian Vyazniki Assemblage of Russia (e.g., Sennikov & Golubev 2006, 2017). The Russian taxon *Eorasaurus olsoni* could push this divergence farther back into the Permian (Sennikov 1997; Ezcurra et al. 2014; Bernardi et al. 2015; Ezcurra 2016), but its status as an archosauriform (and even a sauropsid) has been challenged (Peecook et al. 2018c). Sennikov & Golubev (2006) considered the Vyazniki Assemblage to be transitional between more typical late Permian tetrapod assemblages and those of the Early Triassic, implying an age close to the Permo–Triassic boundary. We calibrate this node at 254 Mya, near the base of the Changhsingian Stage, although the very limited Permian archosauromorph fossil record raises the possibility that this is an underestimate.

Last Appearance Datum

Eothyris: *Eothyris parkeyi* is known from a single specimen that was collected in the Petrolia Formation of Texas (Reisz et al. 2009). The Petrolia Formation has produced a tetrapod assemblage assigned to the Seymouran Land Vertebrate Faunachron (e.g., Lucas 2006, 2018), which is considered to be middle–late Artinskian in age (Schneider et al. 2019), so we calibrate this last occurrence at 287 Mya.

***Dimetrodon*:** *Dimetrodon* is a very abundant and stratigraphically long-ranging genus known from North America and western Europe (e.g., Romer & Price 1940; Reisz 1986; Berman et al. 2001). *Dimetrodon angelensis* from the San Angelo Formation of Texas (Olson 1962) is the youngest species of *Dimetrodon*, although it is known from limited, fragmentary material. The San Angelo Formation tetrapod assemblage is assigned to the Littlecrotonian Land Vertebrate Faunauchron (Lucas 2006, 2018), and the San Angelo formation is considered to be late Kungurian in age based on fusilinids that occur in the formation (Schneider et al. 2019). Therefore, we calibrate the last occurrence of *Dimetrodon* at 275 Mya.

***Hipposaurus*:** *Hipposaurus boonstrai* primarily occurs in the *Diictodon-Styracocephalus* Subzone of the *Tapinocephalus* Assemblage Zone (Abrahamskraal Formation, Karoo Basin, South Africa; Day & Rubidge 2020). The stratigraphically highest specimens of the species range at least as high as the Karelskraal Member of the Abrahamskraal Formation, and one specimen may reach the lower Poortjie Member of the overlying Teekloof Formation, although its exact stratigraphic position is uncertain (Day 2013). The uppermost occurrences of *Hipposaurus* are close to the dates of 261.24–260.26 Mya for the upper *Tapinocephalus* Assemblage Zone (Rubidge et al. 2013; Day et al. 2015). Therefore, we calibrate this last occurrence at 261 Mya.

***Herpetoskylax*:** *Herpetoskylax hopsoni* occurs in strata of the Teekloof Formation that are assigned to the *Cistecephalus* Assemblage Zone (Karoo Basin, South Africa) (Sidor & Rubidge 2006). *Cistecephalus* Assemblage Zone strata are approximately bracketed by radiometric dates of 256.25 Mya and 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020). We calibrate this last occurrence at 255 Mya, near the end of *Cistecephalus* zone times.

***Leucocephalus*:** *Leucocephalus wewersi* is known from a single specimen collected in the Hoedemaker Member of the Middleton Formation (Karoo Basin, South Africa), which is assigned to the *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone (Day et al. 2018a; Day & Smith 2020). The *Tropidostoma-Gorgonops* Subzone is bracketed by radiometric dates of 259.26 Mya and 256.25 Mya (Rubidge et al. 2013; Day et al. 2015) and Day & Smith (2020) considered it to begin about 258 Mya, so we calibrate its last occurrence at 258 Mya.

***Lemurosaurus*:** *Lemurosaurus pricei* is known from rocks of the Balfour and Middleton formations (Karoo Basin, South Africa) that are assigned to the *Cistecephalus* Assemblage Zone (Sidor & Welman 2003; Sidor & Rubidge 2006; Smith 2020). *Cistecephalus* Assemblage Zone strata are approximately bracketed by radiometric dates of 256.25 Mya and 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020). We calibrate this last occurrence at 255 Mya, near the end of *Cistecephalus* zone times.

***Moschops*:** *Moschops* is one of the more abundant tapinocephalid taxa known from the Karoo Basin of South Africa, and its stratigraphic range extends into the Karelskraal Member of the Abrahamskraal Formation (*Diictodon-Styracocephalus* Subzone of the *Tapinocephalus* Assemblage Zone) (Day 2013; Day & Rubidge 2020). Radiometric dates of 261.24 Mya to 260.26 Mya have been reported from the upper *Tapinocephalus* Assemblage Zone (Rubidge et al. 2013; Day et al. 2015), and we calibrate this last occurrence at 261 Mya.

***Patranomodon*:** *Patranomodon nyaphulii* is known from a single specimen collected in rocks of the Abrahamskraal Formation (Karoo Basin, South Africa) assigned to the *Eodicynodon* Assemblage Zone (Rubidge & Hopson 1996; Rubidge & Day 2020). Dates from near the base of this formation range from 268.5–264.6 Mya (Lanci et al. 2013), and the upper portion of the *Tapinocephalus* Assemblage Zone has been dated to about 262–261 Mya (Day et al. 2015). We calibrate this last occurrence at 265 Mya.

***Eodicynodon*:** *Eodicynodon oosthuizeni* occurs in the *Eodicynodon* Assemblage Zone of the Abrahamskraal Formation (Karoo Basin, South Africa), with its highest occurrence falling in the Koornplaats Member (e.g., Day 2013; Day et al. 2018). The last occurrence of *E. oosthuizeni* is stratigraphically higher than the occurrence of *Patranomodon* (Day 2013; Day et al. 2018), so we calibrate this last occurrence at 264 Mya.

***Diictodon*:** Although *Diictodon feliceps* is also known from China and Zambia (Angielczyk & Sullivan 2008), its full temporal range is best documented in the Karoo Basin of South Africa. There, *D. feliceps* ranges through most of the *Daptocephalus* Assemblage Zone (Viglietti et al. 2016; Viglietti 2020). This places it close to the level of Gastaldo et al.'s (2015) radiometric date of 253.48 Mya for the *Lystrosaurus maccaigi-Moschorhinus* Subzone, and we calibrate this last occurrence at 253 Mya.

***Pristerodon*:** Like *Diictodon*, the temporal range of *Pristerodon mackayi* is best documented in the Karoo Basin, where its last occurrence is in the upper portion of the *Dicynodon-Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone (Viglietti 2020). Therefore, we calibrate this last occurrence at 253.5 Mya, slightly older than Gastaldo et al.'s (2015) radiometric date of 253.48 Mya.

***Niassodon*:** The single known specimen of *Niassodon mfumukasi* was collected in the K5 Formation of the Metangula Graben (Mozambique), which was with the *Cistecephalus* Assemblage Zone of the Karoo Basin by Castanhinha et al. (2013). However, a detrital zircon-based maximum depositional age of 258.85 Mya for the K5c horizon reported by Araújo et al. (2020) implies an age close to the boundary of the South African *Lycosuchus-Eunotosaurus* and *Tropidostoma-Gorgonops* subzones of the *Endothiodon* Assemblage Zone (Day & Smith 2020). We use an age of 258 Mya to calibrate this last occurrence.

Abajudon: *Abajudon kaayai* has been reported from the Ruhuhu Formation (Ruhuhu Basin, Tanzania) and the lower Madumabisa Mudstone Formation (Mid-Zambezi Basin, Zambia) (Angielczyk et al. 2014a). In each area, it co-occurs with tapinocephalid dinocephalians (Simon et al. 2010, 2014), leading to a broad correlation with the *Tapinocephalus* Assemblage Zone of the South African Karoo Basin (Olroyd & Sidor 2017; Day & Rubidge 2020). Radiometric dates of 261.24–260.26 Mya have been reported from the upper *Tapinocephalus* Assemblage Zone (Rubidge et al. 2013; Day et al. 2015; Day & Rubidge 2020) and we calibrate this last occurrence at 260 Mya, in part to accommodate our calibrations for divergence times within Endothiodontia (see above).

***Endothiodon tolani*:** *Endothiodon tolani* was first described from the Ruhuhu Formation of Tanzania, where its range appears to be limited to an interval near the Guadalupian-Lopingian boundary (Angielczyk et al. 2014a; Cox & Angielczyk 2015; Olroyd & Sidor 2017). New records from the Metangula Graben of Mozambique (Macungo et al. 2020) and the mid-Zambezi Basin of Zambia (this paper) do not have precisely-constrained ages but could be as young as late Wuchiapingian (~*Cistecephalus* Assemblage Zone of the South African Karoo Basin; Castanhinha et al. 2013; Barbolini et al. 2016; Araújo et al. 2020). Here, we calibrate the last occurrence of *E. tolani* at 258 Mya, reflecting its co-occurrence with *Niassodon* in the K5 Formation (see above), although we acknowledge that it may eventually be shown to have a younger last occurrence as the ages of the Mozambican and Zambian records become more certain.

***Endothiodon bathystoma*:** *Endothiodon bathystoma* is stratigraphically and geographically wide-ranging (e.g., Cox 1964; Ray 2000; Boos et al. 2013; Angielczyk et al. 2014b; Cox & Angielczyk 2015; Macungo et al. 2020), but its temporal range is best constrained in the Karoo Basin of South Africa. In the Karoo, *E. bathystoma* ranges into the *Cistecephalus* Assemblage Zone (Smith et al. 2012; Smith 2020), which is approximately bracketed by radiometric dates of 256.25 and 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith et al. 2020). We calibrate this last occurrence at 256 Mya, reflecting the fact that *Endothiodon* does not appear to reach the upper *Cistecephalus* Zone in the Karoo (Viglietti et al. 2016; Smith 2020). Occurrences of *Endothiodon* in the upper Madumabisa Mudstone Formation (Zambia) and the Usili Formation (Tanzania) may imply a slightly later last occurrence because these formations may sample a

time interval near the *Cistecephalus*/*Daptocephalus* assemblage zone boundary (Angielczyk & Kammerer 2017; Angielczyk 2019; Viglietti 2020), but the ages of those formations have not been corroborated with radiometric dates.

***Dicynodontoides*:** The genus *Dicynodontoides* is known from India, South Africa, Tanzania, and Zambia (e.g., Ray & Bandyopadhyay 2003; Angielczyk et al. 2009, 2014b), but its temporal range is best constrained in the South African Karoo Basin. There, its last occurrence is very close to the traditionally-recognized Permo–Triassic boundary (Angielczyk et al. 2009; Smith & Botha-Brink 2014; Viglietti et al. 2016; Viglietti 2020). Therefore, we calibrate this last occurrence at 252 Mya.

***Kawingasaurus*:** *Kawingasaurus fossilis* is known only from the Usili Formation of the Ruhuhu Basin, Tanzania (Cox 1972). Angielczyk et al. (2014b) correlated the Usili Formation with the Wuchiapingian *Cistecephalus* Assemblage Zone of the Karoo Basin, but more recent work suggests that it may encompass parts of the upper *Cistecephalus* Assemblage Zone and the lower *Daptocephalus* Assemblage Zone (Angielczyk & Kammerer 2017; Angielczyk 2019; also see Sidor et al. 2010; Kammerer 2019; Viglietti 2020). Radiometric dates from the South African Karoo Basin suggest the *Cistecephalus*-*Daptocephalus* zone boundary is approximately 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020; Viglietti 2020), and Gastaldo et al. (2015) reported a date of 253.48 Mya for the *Lystrosaurus maccaigi*-*Moschorhinus* Subzone of the *Daptocephalus* Assemblage Zone. We calibrate this last occurrence at 254 Mya, placing it in the *Dicynodon*-*Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone.

***Kembawacela*:** *Kembawacela kitchingi* is a newly-described cistecephalid dicynodont known only from the upper Madumabisa Mudstone Formation of the Luangwa Basin (Zambia) (Angielczyk et al. 2019). Angielczyk et al. (2014b) correlated the upper Madumabisa Mudstone Formation with the Wuchiapingian *Cistecephalus* Assemblage Zone of the Karoo Basin, but more recent work suggests that it may encompass parts of the upper *Cistecephalus* Assemblage Zone and the lower *Daptocephalus* Assemblage Zone (Angielczyk & Kammerer 2017; Angielczyk 2019; also see Sidor et al. 2010; Smith 2020; Viglietti 2020). Radiometric dates from the South African Karoo Basin suggest the *Cistecephalus*-*Daptocephalus* zone boundary is approximately 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020; Viglietti 2020), and Gastaldo et al. (2015) reported a date of 253.48 Mya for the *Lystrosaurus maccaigi*-*Moschorhinus* Subzone of the *Daptocephalus* Assemblage Zone. We calibrate this last occurrence at 254 Mya, placing it in *Dicynodon*-*Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone.

***Oudenodon*:** *Oudenodon bainii* is a widespread dicynodont in southern Africa (e.g., Botha & Angielczyk 2007; Sidor et al. 2010; Castanhinha et al. 2013; Angielczyk et al. 2014b), and like many such taxa its

temporal range is best known in the South African Karoo Basin. In the Karoo, *Oudenodon* ranges into the *Lystrosaurus maccaigi-Moschorhinus* Subzone of the *Daptocephalus* Assemblage Zone, but disappears some distance below the Permo–Triassic boundary (Smith & Botha-Brink 2014; Viglietti et al. 2016; Viglietti 2020). We calibrate this last occurrence at 253 Mya, taking into account the radiometric date of Gastaldo et al. (2015) for the *Lystrosaurus maccaigi-Moschorhinus* Subzone and the fact that the last occurrence of *Oudenodon* is stratigraphically close to that of *Diictodon* (Viglietti 2020).

***Aulacephalodon*:** The specimen of *Aulacephalodon* used in this analysis is an undescribed specimen that originated in the upper Madumabisa Mudstone Formation of the Luangwa Basin, Zambia. Angielczyk et al. (2014b) correlated the upper Madumabisa Mudstone Formation with the Wuchiapingian *Cistecephalus* Assemblage Zone of the Karoo Basin, but more recent work suggests that it may encompass parts of the upper *Cistecephalus* Assemblage Zone and the lower *Daptocephalus* Assemblage Zone (Angielczyk & Kammerer 2017; Angielczyk 2019; also see Sidor et al. 2010; Smith 2020; Viglietti 2020). The stratigraphic range of *Aulacephalodon* is better constrained in the Karoo Basin itself, where it disappears in the lower portion of the *Lystrosaurus maccaigi-Moschorhinus* Subzone of the *Daptocephalus* Assemblage Zone. Radiometric dates from the Karoo Basin suggest the *Cistecephalus-Daptocephalus* zone boundary is approximately 255.2 Mya (Rubidge et al. 2013; Day et al. 2015), and Gastaldo et al. (2015) reported a date of 253.48 Mya for the *Lystrosaurus maccaigi-Moschorhinus* Subzone. We calibrate this last occurrence at 254 Mya, placing it in lower *Daptocephalus* Assemblage Zone times (i.e., the *Dicynodon-Theriognathus* Subzone), reflecting its potentially slightly greater age than the youngest specimens known from the Karoo.

***Lystrosaurus murrayi*:** The genus *Lystrosaurus* is famous for its Pangaean geographic range (e.g., Fröbisch 2009). These occurrences have long been thought to fall in the Early Triassic, but the exact temporal range encompassed by *Lystrosaurus*-bearing strata is somewhat uncertain, particularly in regard to whether any of the strata were deposited in the Olenekian (e.g., Rubidge 2005; Lucas 2010; Schneider et al. 2019). In the South African Karoo Basin, the last occurrence of *L. murrayi* is in the Swartberg Member of the Katberg Formation, slightly below that last occurrence of *L. declivis* in the upper Katberg Formation (both *Lystrosaurus declivis* Assemblage Zone; Botha & Smith 2006, 2007, 2020). We calibrate the last occurrence of *L. murrayi* at 251.2 Mya, the boundary between the Induan and the Olenekian.

***Lystrosaurus declivis*:** *Lystrosaurus declivis* has a slightly longer stratigraphic range in the South African Karoo Basin than *L. murrayi* (Botha & Smith 2006, 2007, 2020). Therefore, we calibrate its last occurrence at 251 Mya, in the early Olenekian.

Sangusaurus: The genus *Sangusaurus* occurs in the Lifua Member of the Manda Beds (Ruhuhu Basin, Tanzania) and the upper Ntawere Formation (Luangwa Basin, Zambia) (Angielczyk et al. 2018). Traditionally, these strata have been considered to be Anisian–early Ladinian in age (e.g., Lucas 1998, 2010; Rubidge 2005; Hancox et al. 2020), but recent dating of biostratigraphically-correlated strata in South America suggests a younger, Ladinian–Carnian age may be more likely (e.g., Peecook et al. 2018a). We calibrate this last occurrence at 237 Mya (latest Ladinian), although we acknowledge the uncertainty that accompanies this estimate (e.g., Schneider et al. 2019).

'Aloposaurus': This specimen (GPIT/RE/7124) has been variously assigned to the genera *Aloposaurus*, *Aelurosaurus*, and Gorgonopsia *incertae sedis* (see review in Araújo et al. 2017). It was collected in *Cistecephalus* Assemblage Zone strata in the South African Karoo Basin (Araújo et al. 2017), establishing a maximum age for the specimen regardless of its taxonomic assignment. However, based on their compositions in the revisions of Sigogneau-Russell (1989) and Gebauer (2007), the genera *Aloposaurus* and *Aelurosaurus* both have stratigraphic ranges that extend into the *Daptocephalus* Assemblage Zone. We calibrate this last occurrence at 254 Mya, in the *Dicynodon-Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone, although we acknowledge that it is somewhat uncertain until a firm taxonomic assignment exists for this specimen.

***Lycaenops*:** The taxonomy of the genus *Lycaenops* was most recently reviewed by Gebauer (2007), who recognized five named species and a possible unnamed sixth species. Some more recent works (e.g., Kammerer 2016c) considered certain species historically referred to the genus but did not focus on formally revising the genus. Specimens of *Lycaenops* range into the *Lystrosaurus maccaigi-Moschorhinus* Subzone of the *Daptocephalus* Assemblage Zone (e.g., supplementary data of Viglietti et al. 2016; Viglietti 2020). We calibrate this last occurrence at 253 Mya, reflecting Gastaldo et al.'s (2015) date of 253.48 Mya for the *Lystrosaurus maccaigi-Moschorhinus* Subzone.

***Scylacocephalus*:** The specimen in question (BP/1/216) was identified as *Scylacocephalus watermeyri* in Benoit et al. (2017b). However, the genus *Scylacocephalus* was considered a synonym of *Aelurosaurus* or *Aloposaurus* in the revisions of Sigogneau-Russell (1989) and Gebauer (2007), respectively. Based on their compositions in the revisions of Sigogneau-Russell (1989) and Gebauer (2007), the genera *Aloposaurus* and *Aelurosaurus* both have stratigraphic ranges that extend into the *Daptocephalus* Assemblage Zone. We calibrate this last occurrence at 254 Mya, in the *Dicynodon-Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone, although we acknowledge that it is somewhat uncertain pending a revision of *Scylacocephalus* in the taxonomic framework for Gorgonopsia that is currently emerging (e.g., Kammerer 2016c, 2017; Kammerer& Masyutin 2018a).

***Dixeya*:** GPIT/RE/7119 was originally described as *Dixeya nasuta* by von Huene (1950). Sigogneau-Russell (1989) and Gebauer (2007) suggested that it may instead represent *Arctognathus*, but Kammerer (2015) noted that the specimen differed from more certain *Arctognathus* material and preferred to treat *Dixeya* as a valid taxon. Araújo et al. (2017) recently described the endocranial anatomy of this specimen but noted that its systematic placement was uncertain. GPIT/RE/7119 was collected in the Usili Formation of the Ruhuhu Basin, Tanzania. Based on biostratigraphy, the fossil assemblage of the Usili Formation has been considered to correlate with the *Cistecephalus* Assemblage Zone of the South African Karoo Basin (Sidor et al. 2010; Angielczyk et al. 2014a, b), although recent discoveries in the correlative upper Madumabisa Mudstone Formation of the Zambian Luangwa Basin (Angielczyk & Kammerer 2017; Angielczyk 2019; also see Kammerer 2019; Smith 2020; Viglietti 2020) raise the possibility that the Usili Formation may span the *Cistecephalus–Daptocephalus* assemblage zone boundary. No radiometric dates exist for the Usili Formation, but radiometric dates from the Karoo Basin suggest the *Cistecephalus-Daptocephalus* zone boundary is approximately 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020; Viglietti 2020), and Gastaldo et al. (2015) reported a date of 253.48 Mya for the *Lystrosaurus maccaigi-Moschorhinus* Subzone of the *Daptocephalus* Assemblage Zone. We calibrate this last occurrence at 254 Mya, which would make it equivalent to the *Dicynodon-Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone.

BP/1/155: BP/1/155 is a gorgonopsian in the collections of the Evolutionary Studies Institute (University of the Witwatersrand). Benoit et al. (2016b, 2017c, d) included it in their studies of therapsid endocranial anatomy and identified it as Gorgonopsia indet. The specimen was collected in strata assigned to the *Cistecephalus* Assemblage Zone in the Karoo Basin of South Africa (Benoit et al. 2017c). *Cistecephalus* Assemblage Zone strata are approximately bracketed by radiometric dates of 256.25 Mya and 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020). We calibrate this last occurrence at 255 Mya, near the end of *Cistecephalus* zone times.

Euchambersia: *Euchambersia mirabilis* is known from only two specimens that were collected in strata assigned to the *Cistecephalus* Assemblage Zone in the South African Karoo Basin (e.g., Benoit 2016). The *Cistecephalus* zone is approximately bracketed by radiometric dates of 256.25 and 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020), and we calibrate this last occurrence at 255 Mya.

***Olivierosuchus*:** *Olivierosuchus parringtoni* is known from strata assigned to the Early Triassic *Lystrosaurus declivis* Assemblage Zone in the Karoo Basin, South Africa (e.g., Botha & Smith 2006, 2020; Huttenlocker & Smith 2017). The *Lystrosaurus declivis* Assemblage Zone is generally considered to be Induan to potentially early Olenekian in age (e.g., Rubidge 2005; Lucas 2010; Schneider et al. 2019; Botha

& Smith 2020). We calibrate the last occurrence of *Olivierosuchus* at 251 Mya, in the earliest Olenekian, to accommodate this uncertainty.

***Ictidosuchoides*:** *Ictidosuchoides longiceps* has a long stratigraphic range in the Karoo Basin of South Africa, with its latest occurrences falling in the lower *Daptocephalus* Assemblage Zone (Viglietti et al. 2016; Huttenlocker & Smith 2017). Radiometric dates from the South African Karoo Basin suggest the *Cistecephalus-Daptocephalus* zone boundary is approximately 255.2 Mya (Rubidge et al. 2013; Day et al. 2015), and Gastaldo et al. (2015) reported a date of 253.48 Mya for the upper *Daptocephalus* Assemblage Zone. We calibrate this last occurrence at 254 Mya, placing it in lower *Daptocephalus* Assemblage Zone times.

***Microgomphodon*:** *Microgomphodon oligocynus* is known from the Burgersdorp Formation of the South African Karoo Basin and the upper Omingonde Formation of the Namibian Otjiwarongo Basin, with the latter record representing its youngest occurrence (Abdala et al. 2014; Hancox et al. 2020). The upper Omingonde Formation is biostratigraphically correlated with the *Dinodontosaurus* Assemblage Zone of the Santa Maria Supersequence, Brazil (e.g., Abdala et al. 2013; Martinelli et al. 2017a; Melo et al. 2017; Peecook et al. 2018a). The *Dinodontosaurus* Assemblage Zone is likely older than 236–237 Mya (e.g., Martinelli et al. 2017a; Melo et al. 2017; Philipp et al. 2018; Schmitt et al. 2019), so we calibrate the last occurrence of *Microgomphodon* at 237 Mya.

***Choerosaurus*:** *Choerosaurus dejageri* occurs in the Wuchiapingian *Tropidostoma-Gorgonops* Subzone of the *Endothiodon* Assemblage Zone of the Karoo Basin (Benoit et al. 2016a; Huttenlocker & Smith 2017). The *Tropidostoma-Gorgonops* Subzone is thought to be between 258 Mya and 256.8 Mya in age (Rubidge et al. 2013; Day et al. 2015; Day & Smith 2020). We calibrate this last occurrence at 256.5 Mya, near what is presumably the end of *Tropidostoma-Gorgonops* Subzone time.

***Mupashi*:** *Mupashi migrator* is known from a single specimen collected in the upper Madumabisa Mudstone Formation of the Luangwa Basin, Zambia (Huttenlocker & Sidor 2016). Angielczyk et al. (2014b) correlated the upper Madumabisa Mudstone Formation with the Wuchiapingian *Cistecephalus* Assemblage Zone of the Karoo Basin, but more recent work suggests that it may encompass parts of the upper *Cistecephalus* Assemblage Zone and the lower *Daptocephalus* Assemblage Zone (Angielczyk & Kammerer 2017; Angielczyk 2019; also see Sidor et al. 2010; Smith 2020; Viglietti 2020). Radiometric dates from the South African Karoo Basin suggest the *Cistecephalus-Daptocephalus* zone boundary is approximately 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020, Viglietti 2020), and Gastaldo et al. (2015) reported a date of 253.48 Mya for the *Lystrosaurus maccaigi-Moschorhinus* subzone of the

Daptocephalus Assemblage Zone. We calibrate this last occurrence at 254 Mya, placing it in the *Dicynodon-Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone.

***Procynosuchus*:** *Procynosuchus delaharpeae* is a geographically wide-ranging Permian cynodont (e.g., see review in Kammerer 2016b), but its temporal range is best constrained in the Karoo Basin of South Africa. There, it ranges into the *Dicynodon-Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone (Viglietti 2020). Gastaldo et al. (2015) reported a date of 253.48 Mya for the overlying *Lystrosaurus maccaigi-Moschorhinus* Subzone, so we calibrate the last occurrence of *Procynosuchus* at 253.5 Mya.

***Cynosaurus*:** *Cynosaurus suppostus* occurs in the upper Wuchiapingian-Changhsingian *Cistecephalus* and *Daptocephalus* assemblage zones of South Africa (Van den Brandt & Abdala 2018; Smith 2020; Viglietti 2020), and Viglietti (2020) recorded its stratigraphic range as nearly reaching the Permo–Triassic boundary. Therefore, we calibrate its last occurrence at 252 Mya.

***Galesaurus*:** *Galesaurus planiceps* has a short stratigraphic range in the *Lystrosaurus declivis* Assemblage Zone of the Karoo Basin (South Africa) (Botha & Smith 2006, 2020). The *Lystrosaurus declivis* Assemblage Zone is generally considered to be Induan to potentially early Olenekian in age (e.g., Rubidge 2005; Lucas 2010; Schneider et al. 2019; Botha & Smith 2020). We calibrate its last occurrence at 251.2 Mya, at the top of the Induan Stage.

***Thrinaxodon*:** *Thrinaxodon liorhinus* has a well-documented stratigraphic range that extends through the *Lystrosaurus* Assemblage Zone in the South African Karoo Basin (e.g., Botha & Smith 2006, 2020; Smith & Botha-Brink 2014), and it is also found in the biostratigraphically-correlated lower Fremouw Formation of Antarctica (e.g., Kitching et al. 1972; Colbert & Kitching 1977; Hammer 1990; Peecook et al. 2018b). The exact temporal range encompassed by *Lystrosaurus*-bearing strata is somewhat uncertain, particularly in regard to whether any of the strata were deposited in the Olenekian (e.g., Rubidge 2005; Lucas 2010; Schneider et al. 2019). We calibrate the last occurrence of *Thrinaxodon* at 251 Mya, in the earliest Olenekian, to accommodate this uncertainty.

***Trirachodon*:** In their recent taxonomic revision, Hopson & Sidor (2018) recognized one valid species of *Trirachodon*, *T. berryi*. *Trirachodon berryi* is best known from the *Trirachodon-Kannemeyeria* Subzone of the *Cynognathus* Assemblage Zone (Karoo Basin, South Africa; Abdala et al. 2006; Hopson & Sidor 2018; Hancox et al. 2020). Additional material from the upper Omingonde Formation of Namibia has been referred to *Trirachodon* (Keyser 1973; Smith & Swart 2002), but much of this material has been referred to other non-trirachodontid cynodonts subsequently (Abdala et al. 2006; Abdala & Smith 2009). A remaining Namibian specimen may represent *T. berryi*, but this identification is not completely certain

(Abdala et al. 2006). Here, we treat the Namibian record as a valid occurrence of *Trirachodon*. The upper Omingonde Formation is biostratigraphically correlated with the *Dinodontosaurus* Assemblage Zone of the Santa Maria Supersequence of Brazil (e.g., Abdala et al. 2013; Martinelli et al. 2017a; Melo et al. 2017; Peecook et al. 2018a). The *Dinodontosaurus* Assemblage zone of Brazil (Santa Maria Supersequence) is likely older than 236–237 Mya (e.g., Martinelli et al. 2017a; Melo et al. 2017; Philipp et al. 2018; Schmitt et al. 2019), so we calibrate the last occurrence of *Trirachodon* at 237 Mya.

***Cricodon*:** Species of *Cricodon* have been reported from the *Trirachodon-Kannemeyeria* and *Cricodon-Ufudocyclops* subzones of the *Cynoganthus* Assemblage Zone (Karoo Basin, South Africa), the Lifua Member of the Manda Beds (Ruhuhu Basin, Tanzania), and the upper Ntawere Formation (Luangwa Basin, Zambia) (e.g., Abdala et al. 2005; Hopson & Sidor 2018; Peecook et al. 2018a; Hancox et al. 2020). Traditionally, these strata have been considered to be Anisian–early Ladinian in age (e.g., Lucas 1998, 2010; Rubidge 2005; Hancox et al. 2020), but recent dating of biostratigraphically-correlated strata in South America suggests a younger, Ladinian–Carnian age may be more likely (e.g., Peecook et al. 2018a). We calibrate this last occurrence at 237 Mya (latest Ladinian), although we acknowledge the uncertainty that accompanies this estimate.

***Scalenodon*:** The genus *Scalenodon* is represented by two species, *S. angustifrons* from the Tanzanian Lifua Member of the Manda Beds and *S. ribeiroae* from the *Dinodontosaurus* Assemblage Zone of the Santa Maria Supersequence, Brazil (e.g., Liu & Abdala 2014; Melo et al. 2017). The *Dinodontosaurus* Assemblage zone of Brazil (Santa Maria Supersequence) is likely older than 236–237 Mya (e.g., Martinelli et al. 2017a; Melo et al. 2017; Philipp et al. 2018; Schmitt et al. 2019), so we calibrate the last occurrence of *Scalenodon* at 237 Mya.

***Luangwa*:** The genus *Luangwa* is known from the upper Ntawere Formation of Zambia, the upper Omingonde Formation of Namibia, the Lifua Member of the Manda Beds of Tanzania, and the *Dinodontosaurus* Assemblage Zone of the Santa Maria Supersequence, Brazil (Brink 1963; Abdala & Sa-Teixeira 2004; Abdala & Smith 2009; Peecook et al. 2018a). All of these units have been biostratigraphically correlated with each other (e.g., Martinelli et al. 2017a; Peecook et al. 2018a) and are likely in the range of 236–237 Mya based on the estimated age of the *Dinodontosaurus* Assemblage Zone (e.g., Martinelli et al. 2017a; Melo et al. 2017; Philipp et al. 2018; Schmitt et al. 2019). We calibrate the last occurrence of *Luangwa* at 237 Mya.

***Massetognathus*:** The genus *Massetognathus* is known from the *Massetognathus-Chanaresuchus* Assemblage Zone of the Chañares Formation (Argentina) and the *Dinodontosaurus* and *Santacruzodon*

assemblage zones of the Santa Maria Supersequence (Brazil) (e.g., Schmitt et al. 2019). The maximum depositional age of the *Santacruzodon* Assemblage Zone is approximately 237 Mya (Philipp et al. 2018); radiometric dates for the Chañares Formation range from approximately 236–233 Mya (Marsicano et al. 2016; Ezcurra et al. 2017). Based on the date for the upper Chañares Formation, we calibrate the last occurrence of *Massetognathus* at 233 Mya.

***Lumkuia*:** *Lumkuia fuzzi* is known from a single specimen collected in the *Trirachodon-Kannemeyeria* Subzone of the *Cynognathus* Assemblage Zone (Karoo Basin, South Africa) (Hopson & Kitching 2001; Hancox et al. 2020). The *Trirachodon-Kannemeyeria* Subzone has generally been regarded as Anisian in age (e.g., Lucas 1998, 2010; Rubidge 2005; Hancox et al. 2020), although it has not been radiometrically dated directly. Recent radiometric dates of strata that are biostratigraphically-correlated with the *Cynognathus* Assemblage Zone have raised the possibility that it parts of it might be substantially younger than previously thought (e.g., Ottone et al. 2014; Marsicano et al. 2016; see discussion in Peecook et al. 2018a). Detrital zircon crystals from the lower part of the underlying Katberg Formation with a minimum age of 250 ± 5 Mya (Viglietti et al. 2018b) also are consistent with a younger age for the *Cynognathus* Assemblage zone. However, the wide error range on the latter date and the existence of dates for other biostratigraphically-correlated strata more in accordance with the traditional hypothesis (Liu et al. 2018) suggest that further work on the problem is needed (also see Schneider et al. 2019; Hancox et al. 2020). We calibrate the last occurrence of *Lumkuia* at 242 Mya (Anisian–Ladinian boundary), although we acknowledge that that its exact age will remain uncertain until the questions surrounding the ages of many Middle–Late Triassic faunal assemblages are more firmly resolved.

***Chiniquodon*:** The genus *Chiniquodon* is a geographically-widespread and temporally long-ranging probainognathian cynodont (e.g., Abdala & Gaetano 2018). The latest occurrence of *Chiniquodon* is in the Carnian–Norian Ischigualasto Formation of Argentina (e.g., Martínez & Forster 1996; Abdala & Giannini 2002; Martínez et al. 2013b). Radiometric dates of 231.4 and 225.9 Mya are available for the basal and upper Ischigualasto Formation, respectively (see review in Martínez et al. 2013b). Specimens of *Chiniquodon* are found in the lower Ischigualasto Formation (*Scaphonyx-Exaeretodon-Herrerasaurus* biozone; Martínez et al. 2013b), suggesting they are closer to 231 Mya in age than to 225 Mya. Based on these data, we calibrate the last occurrence of *Chiniquodon* at 230 Mya (late Carnian).

***Riograndia*:** *Riograndia guaibensis* is known from the *Riograndia* Assemblage Zone (Candelária Sequence, Santa Maria Supersequence) of the Paraná Basin, Brazil (e.g., Soares et al. 2011). Rocks from this assemblage zone were recently dated at 225.42 Mya (Norian) by Langer et al. (2018). We calibrate the last occurrence of *Riograndia* at 225 Mya.

***Brasilodon*:** *Brasilodon quadrangularis* occurs in strata assigned to the *Riograndia* Assemblage Zone of the Candelária Sequence (Santa Maria Supersequence) in the Paraná Basin, Brazil (e.g., Bonaparte et al. 2003; Martinelli et al. 2016, 2017b; Guignard et al. 2019). Rocks from this assemblage zone were recently dated at 225.42 Mya (Norian) by Langer et al. (2018). We calibrate the last occurrence of *Brasilodon* at 225 Mya.

***Pseudotherium*:** The only known specimen of *Pseudotherium argentinus* was collected in the upper La Peña Member of the Ischigualasto Formation (Ischigualasto-Villa Unión Basin, Argentina), which corresponds to the lower *Scaphonyx-Exaeretodon-Herrerasaurus* biozone (Wallace et al. 2019). This is close to a radiometrically-dated horizon that has produced an age of 231.4 Mya (Carnian; Martínez et al. 2013b). Therefore, we calibrate the last occurrence of *Pseudotherium* at 231 Mya.

***Tritylodon*:** *Tritylodon longaevus* is best known from the upper Elliot Formation of South Africa and Lesotho, but its range may extend into the lower portion of the overlying Clarens Formation if *Tritylodontoideus maximus* is a junior synonym of *Tritylodon longaevus* (e.g., for the possible synonymy see Hopson & Kitching 1972; Gaetano et al. 2017; for stratigraphic range information see Kitching & Raath 1984; Knoll 2005; Sciscio et al. 2017; Bordy et al. 2020; Viglietti et al. 2020). For the purposes of this analysis, we assumed that the proposed synonymy is correct. Recent radiometric dates suggest that the Clarens Formation ranges in age from 190.5–186.7 Mya, with a preferred maximum depositional age of 187.5 Mya (Pliensbachian; Rademan 2018; Bordy et al. 2020; Viglietti et al. 2020). We calibrate the last occurrence of *Tritylodon* at 190 Mya, near the older end of this range, reflecting the occurrence of the '*Tritylodontoideus maximus*' specimen low in the Clarens Formation (Kitching & Raath 1984; Bordy et al. 2020; Viglietti et al. 2020). A minor increase in age would be needed if the range of *Tritylodon* is restricted to occurrences in the Elliot Formation: preferred maximum depositional ages for rock samples from the upper Eliliot Formation Range from 191.1–202.33 Mya (Bordy et al. 2020).

***Oligokyphus*:** The genus *Oligokyphus* is known from North America (Sues 1985; Fedak et al. 2015), Europe (Hennig 1922; Kühne 1956), and China (Luo & Sun 1993). The youngest occurrences of the genus are from the Pliensbachian Kayenta Formation (Sues 1985; Marsh 2014) and Pliensbachian fissure-fill deposits in the United Kingdom (Kühne 1956; Whiteside et al. 2016). Marsh (2014) reported a date of 183.7 Mya (late Pliensbachian) for the Kayenta Formation, and we calibrate the last occurrence of *Oligokyphus* at 183 Mya.

***Morganucodon watsoni*:** The stratigraphic range of *Morganucodon watsoni* extends at least until the early Sinemurian from the St. Brides palaeo-island tetrapod community preserved in fissure fills in the United

Kingdom, but it could reach the Pleinsbachian based on the associated marine invertebrate fauna (Whiteside et al. 2016). We conservatively calibrate this last occurrence at 198 Mya (early Sinemurian).

***Morganucodon oehleri*:** *Morganucodon oehleri* is known from the Zhangjia'ao Member (*sensu* Fang et al. 2000) of the Lufeng Formation of China (Luo & Wu 1994), which is considered to be Sinemurian in age based on biostratigraphical comparisons (e.g., Luo & Wu 1994; Sullivan et al. 2013). We calibrate the last occurrence of *M. oehleri* at 191 Mya (latest Sinemurian).

***Megazostrodon*:** Although also known from the Rhaetian of Europe (Debuysschere et al. 2015), the youngest records of the genus *Megazostrodon* are known from the Early Jurassic of South Africa and Lesotho. *Megazostrodon rudnerae* occurs in the upper Elliot Formation (e.g., Kitching & Raath 1984; Scisio et al. 2017; Bordy et al. 2020; Viglietti et al. 2020), with some records coming from just below the contact with the overlying Clarens Formation (Gow 1986; Bordy et al. 2020). Recent radiometric dates suggest that the upper Elliot Formation spans the latest Rhaetian or Hettangian to the Sinemurian and that the Clarens Formation is Pliensbachian in age (Scisio et al. 2017; Rademan 2018; Bordy et al. 2020). Therefore, we calibrate the last occurrence of *Megazostrodon* at 191 Mya (latest Sinemurian).

Haldanodon: *Haldanodon exspectatus* is known from the Alcobaça Formation of the Guimarota coal mine of Portugal (e.g., Lillegraven & Krusat 1991; Martin 2005, 2018; Ruf et al. 2013), which is considered Kimmeridgian in age (Schudack 2000a, b). We calibrate the last occurrence of *Haldanodon* at 152 Mya, at the end of the Kimmeridgian.

***Hadrocodium*:** *Hadrocodium wui* is known from the Zhangjia'ao Member (*sensu* Fang et al. 2000) of the Lufeng Formation of China (Luo et al. 2001), which is considered to be Sinemurian in age based on biostratigraphical comparisons (e.g., Luo & Wu 1994; Sullivan et al. 2013). We calibrate the last occurrence of *Hadrocodium* at 191 Mya (latest Sinemurian).

***Dryolestes*:** The genus *Dryolestes* is known from the Late Jurassic (Kimmeridgian–Tithonian) of Europe and North America (e.g., Martin 2018). The youngest records of *Dryolestes* are from high in the Brushy Basin Member of the Morrison Formation of western North America (e.g., Turner & Peterson 1999; Foster 2003). Radiometric dates suggest that these occurrences are early Tithonian in age (e.g., Trujillo et al. 2014; Trujillo & Kowallis 2015). Based on these dates, we calibrate the last occurrence of *Dryolestes* at 151 Mya (earliest Tithonian).

***Diadectes*:** The stratigraphically highest occurrences of the genus *Diadectes* are in the Vale Formation (*sensu* Lucas 2006; equivalent to the middle Clear Fork Group of Hentz 1988) of Texas (Kissel 2010). The

Vale Formation is assigned to the Redtankian Land Vertebrate Faunachron (e.g., Lucas 2006, 2018), which is considered to be mid-Kungurian (mid-upper Leonardian) in age (Schneider et al. 2019). We calibrate the last occurrence of *Diadectes* at 276 Mya, in the early part of the late Kungurian.

***Captorhinus*:** The genus *Captorhinus* has a comparably long stratigraphic range in the early Permian of North America, but its latest reliable occurrences are in the Vale Formation (*sensu* Lucas 2006; equivalent to the middle Clear Fork Group of Hentz 1988) of Texas (e.g., LeBlanc et al. 2015). The Vale Formation is assigned to the Redtankian Land Vertebrate Faunachron (e.g., Lucas 2006, 2018), which is considered to be mid-Kungurian (mid-upper Leonardian) in age (Schneider et al. 2019). We calibrate the last occurrence of *Captorhinus* at 276 Mya, in the early part of the late Kungurian.

***Labidosaurus*:** *Labidosaurus hamatus* occurs in the Arroyo Formation (*sensu* Lucas 2006; equivalent to the lower Clear Fork Group of Hentz 1988) of Texas (Modesto et al. 2007). The Arroyo Formation is assigned to the Redtankian Land Vertebrate Faunachron (e.g., Lucas 2006, 2018), which is considered to be mid-Kungurian (mid-upper Leonardian) in age (Schneider et al. 2019). We calibrate the last occurrence of *Labidosaurus* at 277 Mya, in the early part of the late Kungurian.

***Youngina*:** The stratigraphic range of *Youngina capensis* extends into strata assigned to the *Dicynodon-Theriognathus* Subzone of the *Daptocephalus* Assemblage Zone in the South African Karoo Basin (Viglietti et al. 2016; Viglietti 2020). We calibrate its last occurrence at 254 Mya, taking into account that the *Cistecephalus-Daptocephalus* zone boundary is approximately 255.2 Mya (Rubidge et al. 2013; Day et al. 2015; Smith 2020; Viglietti 2020), and Gastaldo et al.'s (2015) date of 253.48 Mya for the *Lystrosaurus maccaigi-Moschorhinus* Subzone of the *Daptocephalus* Assemblage Zone.

***Prolacerta*:** In the Karoo Basin, *Prolacerta broomi* occurs primarily in strata of the Katberg Formation that are assigned to the *Lystrosaurus declivis* Assemblage Zone (e.g., Botha & Smith 2006, 2020). It is also found in the biostratigraphically-correlated lower Fremouw Formation of Antarctica (e.g., Spiekman 2018). The exact temporal range encompassed by *Lystrosaurus*-bearing strata is somewhat uncertain, particularly in regard to whether any of the strata were deposited in the Olenekian (e.g., Rubidge 2005; Lucas 2010; Schneider et al. 2019). We calibrate the last occurrence of *Prolacerta* at 251 Mya, in earliest Olenekian, to accommodate this uncertainty.

References

Abdala, F. Redescription of *Platycraniellus elegans* (Therapsida, Cynodontia) from the Lower Triassic of South Africa, and the cladistic relationships of eutheriodonts. Palaeontology **50**,591–618 (2007).

Abdala, F. & Giannini N. P. Chiniquodontid cynodonts: systematic and morphometric considerations. Palaeontology **45**,1151–1170 (2002).

Abdala, F. & Gaetano L. C. in *The Late Triassic World. Topics in Geobiology, vol. 46* (ed. Tanner, L) 407–445 (Springer, 2018).

Abdala, F. & Sá-Teixeira, A. M. A traversodontid cynodont of African affinity in the South America Triassic. Palaeontol. Afr. **40**, 11–22 (2004).

Abdala, F. & Smith, R. M. H. A Middle Triassic cynodont fauna from Namibia and its implications for the biogeography of Gondwana. Journal of Vertebrate Paleontology **29**, 837–851(2009).

Abdala, F., Hancox, P. J. & Neveling, J. Cynodonts from the uppermost Burgersdorp Formation, South Africa, and their bearing on the biostratigraphy and correlation of the Triassic *Cynognathus* Assemblage Zone. *J. Vertebr. Paleontol*. **25**,192–199 (2005).

Abdala, F., Neveling, J., & Welman, J. A new trirachodontid cynodont from the lower levels of the Burgersdorp Formation (Lower Triassic) of the Beaufort Group, South Africa and the cladistic relationships of Gondwanan gomphodonts. *Zool. J. Linn. Soc.* **147**, 383–413 (2006).

Abdala, F., Rubidge, B.S. & van den Heever, J. The oldest therocephalians (Therapsida, Eutheriodontia) and the early diversification of Therapsida. *Palaeontology* **51**, 1011–1024 (2008).

Abdala, F. Marsicano, C. A., Smith, R. M. H. & Swart, R. Strengthening western Gondwanan correlations: a Brazilian dicynodont (Synapsida, Anomodontia) in the Middle Triassic of Namibia. *Gondwana Res.* **23**, 1151–1162 (2013).

Abdala, F., Jashashvili, T., Rubidge, B. S. & van den Heever, J. in *Early Evolutionary History of the Synapsida* (eds. Kammerer, C. F., Angielczyk, K. D. & Fröbisch, J.) 209–231 (Springer, 2014).

Abdala, F., Gaetano, L. C., Smith, R. M. H., & Rubidge, B. S. A new large cynodont from the late Permian (Lopinigan) of the South African Karoo Basin and its phylogenetic significance. *Zool. J. Linn. Soc*. **186**, 983–1005 (2019).

Amson, E. & Laurin, M. On the affinities of *Tetraceratops insignis*, an Early Permian synapsid. *Acta Palaeontol. Pol.* 56:301–312 (2011).

Angielczyk, K. D. First occurrence of the dicynodont *Digalodon* (Therapsida, Anomodontia) from the Lopingian upper Madumabisa Mudstone Formation, Luangwa Basin, Zambia. Palaeontologia Africana 53:219–225 (2019).

Angielczyk, K. D. & Kammerer, C. F. The cranial morphology, phylogenetic position and biogeography of the upper Permian dicynodont *Compsodon helmoedi* van Hoepen (Therapsida, Anomodontia). *Papers Palaeontol*. 3: 513–545 (2017).

Angielczyk, K. D. & Sullivan, C.: *Diictodon feliceps* (Owen, 1876), a dicynodont (Therapsida, Anomodontia) species with a Pangaean distribution. *J. Vertebr. Paleontol*. **28**, 788–802 (2008).

Angielczyk, K. D. et al. Taxonomic revision and new observations on the postcranial skeleton, biogeography, and biostratigraphy of the dicynodont genus *Dicynodontoides*, the senior subjective synonym of *Kingoria* (Therapsida, Anomodontia). *J. Vertebr. Paleontol*. 29: 1174–1187 (2009).

Angielczyk., K. D. et al. New dicynodonts (Therapsida, Anomodontia) and updated tetrapod stratigraphy of the Permian Ruhuhu Formation (Songea Group, Ruhuhu Basin) of southern Tanzania. *J. Vertebr. Paleontol.* **34**, 1408–1426 (2014a).

Angielczyk, K.D. et al. (2014b): Permian and Triassic dicynodont (Therapsida: Anomodontia) faunas of the Luangwa Basin, Zambia: taxonomic update and implications for dicynodont biogeography and biostratigraphy. In: Kammerer C.F., Angielczyk K.D. & Fröbisch J. (eds.). Early Evolutionary History of the Synapsida. Springer, Dordrecht: 93–138.

Angielczyk, K. D., Rubidge, B. S., Day, M. O. & Lin, F A reevaluation of Brachyprosopus broomi and Chelydontops altidentalis, dicynodonts (Therapsida, Anomodontia) from the middle Permian Tapinocephalus Assemblage Zone of the Karoo Basin, South Africa. Journal of Vertebrate Paleontology 36: e1078342 (2016).

Angielczyk, K. D., Hancox, P. J. & Nabavizadeh, A. A redescription of the Triassic kannemeyeriiform dicynodont Sangusaurus (Therapsida, Anomodontia), with an analysis of its feeding system. In: Sidor, C.A. & Nesbitt S. J. (eds.). Vertebrate and Climatic Evolution in the Triassic Rift Basins of Tanzania and Zambia. Society of Vertebrate Paleontology Memoir 17: 189–227 (2018).

5349 Angielczyk, K. D., Benoit, J., & Rubidge, B. S. A new tusked cistecephalid dicynodont (Therapsida,
5350 Anomodontia) from the upper Permian upper Madumabisa Mudstone Formation, Luangwa Basin, Zambia.
5351 Papers in Palaeontology (2019).

5352 Araújo, R. et al. Aspects of gorgonopsian paleobiology and evolution: insights from the basicranium,
5353 occiput, osseous labyrinth, vasculature, and neuroanatomy. PeerJ 5:e3119 (2017).

5354 Araújo, R. et al. Biostratigraphic refinement of tetrapod-bearing beds from the Metangula Graben (Niassa
5355 Province, Mozambique): new radiometric dating and the first Lower Triassic tetrapod fossils from
5356 Mozambique. Palaeontologia Africana, 54: 56–68 (2020).

5357 Averianov, A. O., Martin T., & Lopatin, A. V. A new phylogeny for basal Trechnotheria and Cladotheria
5358 and affinities of South American endemic Late Cretaceous mammals. Naturwissenschaften 100:311–326
5359 (2013).

5360 Averianov, A. O., Martin, T., and Lopatin, A. The oldest dryolestid mammal from the Middle Jurassic of
5361 Siberia. Journal of Vertebrate Paleontology 34:924–931 (2014).

5362 Barbolini, N. et al. Resolving the age of Madumabisa fossil vertebrates: palynological evidence from the
5363 Mid-Zambezi Basin of Zambia. Palaeogeography, Palaeoclimatology, Palaeoecology 457:117–128 (2016).

5364 Belica, M. E. et al. Refining the chronostratigraphy of the Karoo Basin, South Africa: magnetostratigraphic
5365 constraints support an early Permian age for the Ecca Group. Geophysical Journal International 211:1354–
5366 1374 (2017).

5367 Bendel, E.-M et al. Cranial anatomy of the gorgonopsian *Cynariops robustus* based on CT-reconstruction.
5368 PloS ONE 13(11): e0207367 (2018).

5369 Benoit, J. A review of the “venomous therocephalian” hypothesis and how multiple re-portrayals of
5370 Euchambersia have influenced its success and vice versa. Bulletin de la Société Géologique de France
5371 187:217–224 (2017).

5372 Benoit, J. et al. Cranial bosses of *Choerosaurus dejageri* (Therapsida, Therocephalia): earliest evidence of
5373 cranial display structures in eutheriodonts. PLoS One 11 (8):e0161457 (2016a).

5374 Benoit, J., Manger, P. R. & Rubidge, B. S. Palaeoneurological clues to the evolution of defining mammalian
5375 soft tissue traits. Scientific Reports 6:25604 (2016b).

Benoit, J. et al. The bony labyrinth of late Permian Biarmosuchia: palaeobiology and diversity in non-mammalian Therapsida. Palaeontologia Africana 52: 58–77 (2017b).

Benoit, J. et al. Reappraisal of the envenoming capacity of *Euchambersia mirabilis* (Therapsida, Therocephalia) using µCT-scanning techniques. PLoS One 12(2):e0172047 (2017a).

Benoit, J. et al. Endocranial casts of pre-mammalian therapsids reveal an unexpected neurological diversity at the deep evolutionary root of mammals. Brain, Behavior, and Evolution **90**, 311–333 (2017c).

Benoit, J. et al. The evolution of the maxillary canal in Probainognathia (Cynodontia, Synapsida): reassessment of the homology of the infraorbital foramen in mammalian ancestors. Journal of Mammalian Evolution. (2019).

Benson, R. B. J. Interrelationships of basal synapsids: cranial and postcranial morphological partitions suggest different topologies. Journal of Systematic Palaeontology **10**, 601–624 (2012).

Benton, M. J. et al. Preservation of exceptional vertebrate assemblages in Middle Permian fluviolacustrine mudstones of Kotel'nich, Russia: stratigraphy, sedimentology, and taphonomy. Palaeogeography, Palaeoclimatology, Palaeoecology 319–320:58–83 (2012).

Benton, M. J., et al. Constraints on the timescale of animal evolutionary history. Palaeontologia Electronica 18.1.1FC: 1–106 (2015).

Berman, D. S. Origin and early evolution of the amniote occiput: Journal of Paleontology 74:938–956 (2000).

Berman, D. S. Diadectomorphs, amniotes or not? In: Lucas, S. G. (eds.). The Carboniferous–Permian Transition. New Mexico Museum of Natural History and Science Bulletin 60:22–35 (2013).

Berman, D. S. et al. A new species of *Dimetrodon* (Synapsida: Sphenacodontidae) from the Lower Permian of Germany records first occurrence of genus outside of North America. Canadian Journal of Earth Sciences 38: 803–812 (2001).

Berman, D. S, Sumida, S. S. & Lombard, R. E. Reinterpretation of the temporal and occipital regions in Diadectes and the relationships of diadectomorphs: Journal of Paleontology 66:481–499 (1992).

Bernardi, M. et al. The origin and early radiation of archosauriforms: integrating the skeletal and footprint record. PloS One 10(6):e0128449 (2015).

Bever, G. S. et al. Evolutionary origin of the turtle skull. Nature 525:239–242 (2015).

Bever, G. S. et al. The amniote temporal roof and the diapsid origin of the turtle skull. Zoology 119:471–473 (2016).

Bhat, M. S., Ray, S. & Datta, P. M. A new hybodont shark (Chondrichthyes, Elasmobranchii) from the Upper Triassic Tiki Formation of India with remarks on its dental histology and biostratigraphy. Journal of Paleontology 92:221–239 (2018).

Bickelmann, C., Müller, J., & Reisz, R. R. The enigmatic diapsid *Acerosodontosaurus piveteaui* (Reptilia: Neodiapsida) from the Upper Permian of Madagascar and the paraphyly of "younginiform" reptiles. Canadian Journal of Earth Sciences 46:651–661 (2009).

Bonaparte, J. F. Los tetrápodos del sector superior de la Formación Los Colorados, La Rioja, Argentina (Triásico Superior). Opera Lilloana 22:87–102 (1971).

Bonaparte, J. F. & Crompton, A. W. Origin and relationships of the Ictidosauria to non-mammalian cynodonts and mammals. Historical Biology 30:174–182 (2017).

Bonaparte, J. F. et al. The sister group of mammals: small cynodonts from the Late Triassic of Southern Brazil. Revista Brasileira de Paleontologia 5:5–27 (2003).

Boos, A. D. S. et al. On the presence of the late Permian dicynodont Endothiodon in Brazil. Palaeontology 56:837–848 (2013).

Boos, A. D. S. et al. A tapinocephalid dinocephalian (Synapsida, Therapsida) from the Rio do Rasto Formation Paraná Basin, Brazil): taxonomic, onotgenetic and biostratigraphic considerations. Journal of South American Earth Sciences 63:375–384 (2015).

Boos, A. D. S. et al. A new dicynodont (Therapsida, Anomodontia) from the Permian of southern Brazil and its implications for bidentalian origins. PloS One 11(5): e0155000 (2016).

Bordy, E. M. et al. A chronostratigraphic framework for the upper Storberg Group: implications for the Triassic-Jurassic boundary in southern Africa. Earth-Science Reviews 203:103120 (2020).

Botha, J. & Angielczyk, K. D. An integrative approach to distinguishing the Late Permian dicynodonts Oudenodon bainii and *Tropidostoma microtrema* (Therapsida: Anomodontia). Palaeontology 50: 1175–1209 (2007).

5430 Botha J. & Smith, R. M. H. Rapid vertebrate recovery in the Karoo Basin of South Africa following the
5431 end-Permian extinction. Journal of African Earth Sciences 45:502–514 (2006).

5432 Botha, J. & Smith, R. M. H. *Lystrosaurus* species composition across the Permo-Triassic Boundary in the
5433 Karoo Basin of South Africa. Lethaia 40, 125–137 (2007).

5434 Botha, J., Abdala, F. & Smith, R. The oldest cynodont: new clues on the origin and early diversification of
5435 the Cynodontia. Zoological Journal of the Linnean Society 149:477–492 (2007).

5436 Botha, J., & Smith R. M. H. Biostratigraphy of the *Lystrosaurus declivis* Assemblage Zone (Beaufort
5437 Group, Karoo Supergroup), South Africa. South African Journal of Geology 123:207–216 (2020).

5438 Brocklehurst, N. et al. A re-description of '*Mycterosaurus*' *smithae*, an early Permian eothyridid, and its
5439 impact on the phylogeny of pelycosaurian-grade synapsids. PLoS One 11(6): e0156810 (2016b).

5440 Butler, E., Abdala, F., & Botha-Brink, J. Postcranial morphology of the Early Triassic epicynodont
5441 *Galesaurus planiceps* (Owen) from the Karoo Basin, South Africa. Papers in Palaeontology **5**, 1–32 (2018).

5442 Butler, P. M. & Clemens, W. A. Dental morphology of the Jurassic holotherian mammal *Amphitherium*,
5443 with a discussion of the evolution of mammalian post-canine dental formulae. Palaeontology **44**, 1–20
5444 (2001).

5445 Carroll, R. L. The earliest reptiles. Journal of the Linnean Society, Zoology 45, 61–83 (1964).

5446 Castanhinha, R. et al. Bringing dicynodonts back to life: paleobiology and anatomy of a new emydopoid
5447 genus from the Upper Permian of Mozambique. PLoS One 8 (12):e80974 (2013).

5448 Clemens, W. A. & Martin, T. Review of the non-tritylodontid synapsids from bone beds in the Rhaetic
5449 Sandstone, southern Germany. Paläontologische Zeitschrift 88:461–479 (2014).

5450 Clemmensen, L. B. et al. The vertebrate-bearing Late Triassic Fleming Fjord Formation of central
5451 Greenland revisited: stratigraphy, palaeoclimate and new palaeontological data. In: Kear, B.P., Lindgren,
5452 J., Hurum, J.H., Milàn, J., & Vajda, V. (eds.). Mesozoic biotas of Scandinavia and its Arctic Territories.
5453 Geological Society of London, Special Publication 434:31–47 (2016).

5454 Close, R. A. et al. Evidence for a mid-Jurassic adaptive radiation in mammals. Current Biology **25**, 2137–
5455 2142 (2015).

Colbert, E. H. & Kitching, J. W. Triassic cynodont reptiles from Antarctica. American Museum Novitates 2611:1–30 (1977).

Conrad, J. & Sidor, C. A. Re-evaluation of *Tetraceratops insignis* (Synapsida: Sphenacodontia). Journal of Vertebrate Paleontology 31(abstract volume), 42A (2001).

Cox, C. B. On the palate, dentition and classification of the fossil reptile Endothiodon and related genera. American Museum Novitates 2171: 1–25 (1964).

Cox, C. B. A new digging dicynodont from the upper Permian of Tanzania. In: Joysey, K.A. & Kemp, T.S. (eds.). Studies in Vertebrate Evolution. Oliver and Boyd, Edinburgh: 173–189 (1972).

Cox, C. B. & Angielczyk, K. D. A new endothiodont dicynodont (Therapsida, Anomodontia) from the Permian Ruhuhu Formation (Songea Group) of Tanzania and its feeding system. Journal of Vertebrate Paleontology 35:e935388 (2015).

Crawford, N. G. et al. A phylogenomic analysis of turtles. Molecular Phylogenetics and Evolution 83:250–257 (2014).

Datta, P. M. Earliest mammal with transversely expanded upper molar from the Late Triassic (Carnian) Tiki Formation, South Rewa Gondwana Basin, India. Journal of Vertebrate Paleontology 25:200–207 (2005).

Datta, P. M. & Das, D. P. Discovery of the oldest fossil mammal from India. Indian Minerals 50:217–222 (1996).

David, R. et al. Assessing morphology and function of the semicircular duct system: introducing new in-situ visualization and software toolbox. *Sci. Rep*. **6**, 32772 (2016).

Day, M. O. Middle Permian continental biodiversity changes as reflected in the Beaufort Group of South Africa: a bio- and lithostratigraphic review of the Eodicynodon, Tapinocephalus and Pristerognathus Assemblage Zones. Unpublished PhD dissertation, University of the Witwatersrand, Johannesburg (2013).

Day, M. et al. Biostratigraphic correlation in the Karoo: the case of the Middle Permian parareptile *Eunotosaurus*. South African Journal of Science **109** (2013).

Day, M. O. et al. When and how did the terrestrial mid-Permian mass extinction occur? Evidence from the tetrapod record of the Karoo Basin, South Africa. Proceedings of the Royal Society B 282: 20150834 (2015).

Day, M. O., Rubidge, B. S. & Abdala, F. A new mid-Permian burnetiamorph therapsid from the main Karoo Basin of South Africa and a phylogenetic review of Burnetiamorpha. Acta Palaeontologica Polonica 61:701–719 (2016).

Day, M. O. et al. A new species of burnetiid (Therapsida, Burnetiamorpha) from the early Wuchiapingian of South Africa and implications for the evolutionary ecology of the family Burnetiidae. Papers in Palaeontology **4**, 453–475 (2018a).

Day, M. O. et al. Evolutionary rates of mid-Permian tetrapods from South Africa and the role of temporal resolution in turnover reconstruction. Paleobiology **44**, 347–367 (2018b).

Day, M. O. & Rubidge, B. S. Biostratigraphy of the *Tapinocephalus* Assemblage Zone (Beaufort Group, Karoo Supergroup), South Africa. South African Journal of Geology **123**, 149–164 (2020).

Day, M. O. & Smith, R. M. H. Biostratigraphy of the *Endothiodon* Assemblage Zone (Beaufort Group, Karoo Supergroup), South Africa. South African Journal of Geology **123**, 165–180 (2020).

Debuysschere, M., Gheerbrant, E. & Allain, R. Earliest known European mammals: a review of the Morganucodonta from Saint-Nicolas-de-Port (Upper Triassic, France). Journal of Systematic Palaeontology 13:825–855 (2015).

Dodick, J. T. & Modesto, S. P. The cranial anatomy of the captorhinid reptile Labidosaurikos meachami from the lower Permian of Oklahoma. Palaeontology 38:687–711 (1995).

Ezcurra, M. D. The phylogenetic relationships of basal archosauromorphs, with an emphasis on the systematics of proterosuchian archosauriforms. PeerJ 4:e1778 (2016).

Ezcurra, M. D. et al. Deep faunistic turnovers preceded the rise of dinosaurs in southwestern Pangaea. Nature Ecology and Evolution 1:1477–1483 (2017).

Ezcurra, M. D., Scheyer, T. M. & Butler, R. J. The origin and early evolution of Sauria: Reassessing the Permian saurian fossil record and the timing of the crocodile-lizard divergence. PloS One **9**, e89165 (2014).

Fang, X. et al. Division of Lower, Middle, and Upper Jurassic subdivision in the Lufeng Area, Yunnan. Proceedings of the Third National Stratigraphical Conference of China. Beijing, Geological Publishing House: 208–214 (2000).

5509 Fedak, T. J., Sues, H.-D., & Olsen, P. E. First record of the tritylodontid cynodont Oligokyphus and
5510 cynodont postcranial bones from the McCoy Brook Formation, of Nova Scotia, Canada. Canadian Journal
5511 of Earth Sciences 52:244–249 (2015).

5512 Field, D. J. et al. Toward consilience in reptile phylogeny: miRNAs support an archosaur, not lepidosaur,
5513 affinity for turtles. Evolution & Development 16:189–196 (2014).

5514 Ford, D. P. & Benson, R. B. J. A redescription of *Orovenator mayorum* (Sauropsida, Diapsida) using high-
5515 resolution μCT, and the consequences for early amniote phylogeny. Papers in Palaeontology 5:197–239
5516 (2019).

5517 Foster, J. R. Paleoecological analysis of the vertebrate fauna of the Morrison Formation (Upper Jurassic),
5518 Rocky Mountain Region, U.S.A. New Mexico Museum of Natural History and Science Bulletin 23:1–95
5519 (2003).

5520 Fröbisch, J. Composition and similarity of global anomodont-bearing tetrapod faunas. Earth-Science
5521 Reviews 95: 119–157 (2009).

5522 Garland, T. The relation between maximal running speed and body mass in terrestrial mammals. Journal of
5523 Zoology, 199(2):157-170 (1983).

5524 Gastaldo, R. A. et al. The base of the Lystrosaurus Assemblage Zone, Karoo Basin, predates the end-
5525 Permian extinction. Nature Communications 11:1428 (2020).

5526 Gauthier, J., Kluge, A. J. & Rowe, T. Amniote phylogeny and the importance of fossils. Cladistics 4: 105–
5527 209 (1988).

5528 Gaetano, L. C. & Rougier, G. W. First amphilestid from South America. A molariform from the Jurassic
5529 Cañadon Asfalto Formation, Patagonia, Argentina. Journal of Mammalian Evolution 19:235–248 (2012).

5530 Gaetano, L. C., Abdala F. & Govender R. The postcranial skeleton of the Lower Jurassic Tritylodon
5531 longaevus from southern Africa. Ameghiniana 54:1–35 (2017).

5532 Gastaldo, R. A. et al. Is the vertebrate-defined Permian-Triassic boundary in the Karoo Basin, South Africa,
5533 the terrestrial expression of the end-Permian marine event? Geology 43:939–942 (2015).

Gebauer, E. V. I. Phylogeny and evolution of the Gorgonopsia with a special reference to the skull and skeleton of GPIT/RE/7113 (‘*Aelurognathus*’ *parringtoni*). Unpublished PhD Dissertation, Eberhard-Karls-Universität Tübingen, Tübingen (2007).

Golubev, V. K. Dinocephalian stage in the history of the Permian tetrapod fauna of Eastern Europe. Paleontological Journal 49:1346–1352 (2015).

Gow, C. E. A new skull of Megazostrodon (Mammalia, Triconodonta) from the Elliot Formation (Lower Jurassic) of southern Africa. Palaeontologia Africana 20:13–23 (1986).

Guignard, M. L., Martinelli, A. G., & Soares, M. B. The postcranial anatomy of Brasilodon quadrangularis and the acquisition of mammaliaform traits among non-mammaliaform cynodonts. PloS ONE 14(5):e0216672 (2019).

Hahn, G. Neue Zähne von Haramiyiden aus der deutschen Ober-Trias und ihre Beziehungen zu den Multituberculaten. Palaeontographica, Abteilung A 142:1–15 (1973).

Hammer, W. R. *Thrinaxodon* from Graphite Peak, central Transantarctic Mountains, Antarctica. Antarctic Journal of the United States 25:37–38 (1990).

Hancox, P. J., Neveling, J., & Rubidge, B. S. Biostratigraphy of the Cynognathus Assemblage Zone (Beaufort Group, Karoo Supergroup), South Africa. South African Journal of Geology 123:217–238 (2020).

Heaton, M. J. Cranial anatomy of primitive captorhinid reptiles from the Late Pennsylvanian and Early Permian Oklahoma and Texas. Oklahoma Geological Survey Bulletin 127:1–84 (1979).

Henrickx et al. A new traversodontid cynodont with a peculiar postcanine dentition from the Middle/Late Triassic of Namibia and dental evolution in basal gomphodonts. Journal of Systematic Palaeontology 18(20):1669–1706 (2020).

Hennig, E. Die Säugerzähne des württembergischen Rhät-Lias-Bonebeds. Neues Jahrbuch für Mineralogie, Geologie und Paläontologie, Beilage-Band 46:181–267 (1922).

Hentz, T. F. Lithostratigraphy and paleoenvironments of Upper Paleozoic continental red beds, north-central Texas: Bowie (new) and Wichita (revised) groups. Texas Bureau of Economic Geology Report of Investigations 170:1–55 (1988).

5560 Hopson, J. A. & Kitching, J. W. A revised classification of cynodonts (Reptilia: Therapsida) Palaeontologia
5561 Africana 14:71–85 (1972).

5562 Hopson, J. A. & Kitching, J. W. A probainognathian cynodont from South Africa and the phylogeny of
5563 nonmammalian cynodonts. Bulletin of the Museum of Comparative Zoology 156:3–35 (2001).

5564 Huene, F. von. Die Theriodontier des ostafrikanischen Ruhuhu-Gebietes in der Tübinger Sammlung. Neues
5565 Jahrbuch für Geologie und Paläontologie 92:47–136 (1950).

5566 Huttenlocker, A. K. & Sidor, C. A. The first karenitid (Therapsida, Therocephalia) from the upper Permian
5567 of Gondwana and the biogeography of Permo-Triassic therocephalians. Journal of Vertebrate Paleontology
5568 36:e1111897 (2016).

5569 Huttenlocker, A. K. & Smith, R. M. H. New whaitsioids (Therapsida: Therocephalia) from the Teekloof
5570 Formation of South Africa and therocephalian diversity during the end-Guadalupian extinction. PeerJ
5571 5:e3868 (2017).

5572 Huttenlocker, A. K., Sidor, C. A. & Smith, R. M. H. A new specimen of Promoschorhynchus (Therapsida:
5573 Therocephalia: Akidnognathidae) from the Lower Triassic of South Africa and its implications for
5574 theriodont survivorship across the Permo-Triassic boundary. Journal of Vertebrate Paleontology 31:405–
5575 421 (2011b).

5576 Huttenlocker, A. K. et al. Late-surviving stem mammal links the lowermost Cretaceous of North America
5577 and Gondwana. Nature 558:108–112 (2018).

5578 Huttenlocker, A. K. & Sidor, C. A. A basal nonmammaliaform cynodont from the Permian of Zambia and
5579 the origins of mammalian endocranial and postcranial anatomy. Journal of Vertebrate Paleontology. DOI:
5580 10.1080/02724634.2020.1827413 (2020).

5581 Irisarri, I. et al. Phylotranscriptomic consolidation of the jawed vertebrate timetree. Nature Ecology &
5582 Evolution 1:1370–1378 (2017).

5583 Jasinoski, S. C. & Abdala, F. Cranial ontogeny of the Early Triassic basal cynodont Galesaurus planiceps.
5584 The Anatomical Record 300:353–381 (2017a).

5585 Jasinoski, S. C. & Abdala, F. Aggregations and parental care in the Early Triassic basal cynodonts
5586 Galesaurus planiceps and Thrinaxodon liorhinus. PeerJ 5:e2875 (2017b).

5587 Jenkins, F. A. Jr., Crompton, A. W. & Downs, W. R. Mesozoic mammals from Arizona: New evidence on
5588 mammalian evolution. Science 222:1233–1235 (1983).

5589 Jenkins, F. A. Jr. et al. Late Triassic continental vertebrates and depositional environments of the Fleming
5590 Fjord Formation, Jameson Land, East Greenland. Meddelelser om Grønland 32:3–25 (1994).

5591 Jenkins, F. A. Jr. et al. Haramiyids and Triassic mammalian evolution. Nature 385:715–718 (1997).

5592 Kammerer, C. F. A redescription of *Eriphostoma microdon* Broom, 1911 (Therapsida, Gorgonopsia) from
5593 the Tapinocephalus Assemblage Zone of South Africa and a review of Middle Permian gorgonopsians. In:
5594 Kammerer, C.F., Angielczyk, K.D. & Fröbisch, J. (eds.). Early Evolutionary History of the Synapsida.,
5595 171–184 (Springer, 2014).

5596 Kammerer, C. F. Cranial osteology of Arctognathus curvimola, a short-snouted gorgonopsian from the late
5597 Permian of South Africa. Papers in Palaeontology 1:41–58 (2015).

5598 Kammerer, C. F. Two unrecognized burnetiamorph specimens from historical Karoo collections.
5599 Palaeontologia Africana 50: 64–75 (2016a).

5600 Kammerer, C. F. A new taxon of cynodont from the Tropidostoma Assemblage Zone (upper Permian) of
5601 South Africa, and the early evolution of Cynodontia. Papers in Palaeontology 2:387–397 (2016b).

5602 Kammerer, C. F. Systematics of the Rubidgeinae (Therapsida: Gorgonopsia). PeerJ 4: e1608 (2016c).

5603 Kammerer, C. F. Anatomy and relationships of the South African gorgonopsian Arctops (Therapsida,
5604 Theriodontia). Papers in Palaeontology 3: 583–611 (2017).

5605 Kammerer, C. F. The first skeletal evidence of a dicynodont from the lower Elliot Formation of South
5606 Africa. Palaeontologia Africana 52: 102–128 (2018).

5607 Kammerer, C. F. Revision of the Tanzanian dicynodont *Dicynodon huenei* (Therapsida: Anomodontia) from
5608 the Permian Usili Formation. PeerJ 7:e7420 (2019).

5609 Kammerer, C. F. & Masyutin, V. Gorgonopsian therapsids (*Nochnitsa* gen. nov. and *Viatkogorgon*) from the
5610 Permian Kotelnich locality of Russia. PeerJ 6:e4954 (2018a).

5611 Kammerer, C. F. & Smith, R. M. H. An early geikiid dicynodont from the *Tropidostoma* Assemblage Zone
5612 (late Permian) of South Africa. PeerJ 5:e2913 (2017).

5613 Kammerer, C. F., Smith, R. M. H., Day, M. O. & Rubidge, B. S. New information on the morphology and
5614 stratigraphic range of the mid-Permian gorgonopsian *Eriphostoma microdon* Broom, 1911. Papers in
5615 Palaeontology 1:201–221 (2015).

5616 Kammerer, C. F. et al. A new kannemeyeriiform dicynodont (*Ufudocylops mukanelai*, gen. et sp. nov.) from
5617 Subzone C of the Cynognathus Assemblage Zone, Triassic of South Africa, with implications for
5618 biostratigraphic correlation with other African Triassic faunas. Journal of Vertebrate Paleontology. DOI:
5619 10.1080/02724634.2019.1596921 (2019).

5620 Kent, D. V. et al. Age constraints on the dispersal of dinosaurs in the Late Triassic from magnetochronology
5621 of the Los Colorados Formation (Argentina). Proceedings of the National Academy of Sciences 111:7958–
5622 7963 (2014).

5623 Keyser, A. W. A new Triassic vertebrate fauna from South West Africa. Palaeontologia Africana 16:1–15
5624 (1973).

5625 Kielan-Jaworowska, Z., Cifelli, R. L. & Luo, Z.-X. Mammals from the Age of Dinosaurs. Columbia
5626 University Press, New York (2004).

5627 Kissel, R. Morphology, Phylogeny, and Evolution of Diadectidae (Cotylosauria: Diadectomorpha).
5628 Unpublished Ph.D. dissertation, University of Toronto (2010).

5629 Kissel, R. A. & Reisz, R. R. *Ambedua pusillus*, new genus, new species, a small diadectid (Tetrapoda:
5630 Diadectomorpha) from the Lower Permian of Ohio, with a consideration of diadectomorph phylogeny.
5631 Annals of Carnegie Museum 73:197–212 (2004b).

5632 Kissel, R. A. & Reisz, R. R. Synapsid fauna of the Upper Pennsylvanian Rock Lake Shale near Garnett,
5633 Kansas and the diversity pattern of early amniotes. In: Arratia, G., Wilson, M.V.H. & Cloutier, R. (eds.).
5634 Recent Advances in the Origin and Early Radiation of Vertebrates. Verlag Dr. Friedrich Pfeil, Munich: 409–
5635 428 (2004b).

5636 Kitching, J. W. & Raath, M. A. Fossils from the Elliot and Clarens formations (Karoo Sequence) of the
5637 northeastern Cape, Orange Free State and Lesotho, and a suggested biozonation based on tetrapods.
5638 Palaeontologia Africana 25:111–125 (1984).

5639 Kitching J. W. et al. *Lystrosaurus* Zone (Triassic) fauna from Antarctica. Science 175:524–527 (1972).

Knoll, F. The tetrapod fauna of the Upper Elliot and Clarens formations in the main Karoo Basin (South Africa and Lesotho). Bulletin de la Société géologique de France 176: 81–91 (2005).

Kruger, A. et al. *Lende chiweta*, a new therapsid from Malawi, and its influence on burnetiamorph phylogeny and biogeography. Journal of Vertebrate Paleontology 35:e1008698 (2015).

Kühne, W. G. The Liassic therapsid *Oligokyphus*. Trustees of the British Museum, London (1956).

Kumar, S. et al. TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. Mol Biol Evol (2017).

Kurkin, A. A. Permian anomodonts: paleobiogeography and distribution of the group. Paleontological Journal 45:432–444 (2011).

Lanci, L. et al. Upper Permian magnetic stratigraphy of the lower Beaufort Group, Karoo Basin. Earth and Planetary Science Letters 375:123–134 (2013).

Langer, M. C., Ramezani, J. & Da Rosa, Á. A. S. U-Pb age constraints on dinosaur rise from south Brazil. Gondwana Research 57:133–140 (2018).

Laurin, M. & Piñeiro, G. H. A reassessment of the taxonomic position of mesosaurs, and a surprising phylogeny of early amniotes. Frontiers in Earth Science 5:88 (2017).

Laurin, M. & Reisz, R. R. A reevaluation of early amniote phylogeny. Zoological Journal of the Linnean Society 113: 165–223 (1995).

Laurin, M. & Reisz, R. R. A new perspective on tetrapod phylogeny. In: Sumida, S.S. & Martin, K.L.M. (eds.). Amniote Origins: Completing the Transition to Land. Academic Press, San Diego: 9–59 (1997).

LeBlanc, A. R. H. et al. Multiple tooth-rowed captorhinids from the early Permian fissures fills of the Bally Mountain Locality of Oklahoma. Vertebrate Anatomy, Morphology, Palaeontology 1:35–49 (2015).

Lee, M. S. Y. & Spencer P. S. Crown-clades, key characters, and taxonomic stability: when is an amniote not an amniote? In: Sumida, S.S. & Martin, K.L.M. (eds.). Amniote Origins: Completing the Transition to Land. 61–84 (Academic Press, 1997).

Li, C., et al. A Triassic stem turtle with an edentulous beak. Nature 560:476–479 (2018).

Lillegraven, J. A. & Krusat, G. Cranio-mandibular anatomy of *Haldanodon exspectatus* (Docodonta; Mammalia) from the Late Jurassic of Portugal and its implications to the evolution of mammalian characters. Contributions to Geology, University of Wyoming 28:39–138 (1991).

Liu, J. New progress on the correlation of Chinese terrestrial Permo-Triassic strata. Vertebrata PalAsiatica 56:327–342 (2018).

Liu, J. & Abdala, F. Phylogeny and taxonomy of the Traversodontidae. In: Kammerer, C.F., Angielczyk, K.D. & Fröbisch, J. (eds.). Early Evolutionary History of the Synapsida. 255–279 (Springer, 2014).

Liu, J. & Abdala, F. The tetrapod fauna of the upper Permian Naobaogou Formation of China: 1. *Shiguaignathus wangi* gen. et sp. nov., the first akidnognathid therocephalian from China. PeerJ 5:e4150 (2017).

Liu, J. & Bever, G. S. The last diadectomorph sheds light on Late Palaeozoic tetrapod biogeography. Biology Letters 11:20150100 (2015).

Liu, J. & Olsen, P. E. The phylogenetic relationships of Eucynodontia (Amniota: Synapsida). Journal of Mammalian Evolution 17:151–176 (2010).

Liu, J., Rubidge, B. S. & Li, J. New basal synapsid supports Laurasian origin for therapsids. Acta Palaeontologica Polonica 54:393–400 (2009).

Liu, J., Rubidge, B. S. & Li, J. A new specimen of Biseridens qilianicus indicates its phylogenetic position as the most basal anomodont. Proceedings of the Royal Society B 277:285–292 (2010).

Liu, J. et al. High precision temporal calibration of Middle Triassic vertebrate biostratigraphy: U=Pb zircon constraints for the Sinokannemeyeria Fauna and Yonghesuchus. Vertebrata Palasiatica 56:16–24 (2018).

Liu, J. *Taoheodon baizhijuni*, gen. et sp. nov. (Anomodontia, Dicynodontoidea), from the upper Permian Sunjiagou Formation of China and its implications. J. Vertebr. Paleontol. (2020).

Lucas, S. G. Global tetrapod biostratigraphy and biochronology. Palaeogeography, Palaeoclimatology, Palaeoecology 143:347–384 (1998).

Lucas, S. G. Global Permian tetrapod biostratigraphy and biochronology. In: Lucas, S.G., Cassinis, G. & Schneider, J.W. (eds.). Non-marine Permian Biostratigraphy and Biochronology. Geological Society of London, Special Publication 265: 65–93 (2006).

Lucas, S. G. The Triassic timescale based on nonmarine tetrapod biostratigraphy and biogeography. In: Lucas, S.G. (ed.). The Triassic Timescale. Geological Society of London, Special Publication 334:447–500 (2010).

Lucas, S. G. Permian tetrapod biochronology, correlation and evolutionary events. In: Lucas, S.G. & Shen, S.Z. (eds.). The Permian Timescale. Geological Society of London, Special Publication 450: 405–444 (2018).

Lucas, S. G. & Luo, Z.-X. *Adelobasileus* from the Upper Triassic of west Texas: the oldest mammal. Journal of Vertebrate Paleontology 13:309–334 (1993).

Lukic-Walther, M. et al. Diversity patterns of nonmammalian cynodonts (Synapsida, Therapsida) and the impact of taxonomic practice and research history on diversity estimates. Paleobiology 45:56–69 (2019).

Luo, Z. & Sun A. *Oligokyphus* (Cynodontia: Tritylodontidae) from the lower Lufeng Formaton (lower Jurassic) of Yunnan, China. Journal of Vertebrate Paleontology 13:477–482 (1995).

Luo, Z.-X. & Martin, T. Analysis of molar structure and phylogeny of docodont genera. Bulletin of Carnegie Museum of Natural History 39:27–47 (2007).

Luo, Z.-X. & Wu, X.-C. The small tetrapods of the Lower Lufeng Formation, Yunnan, China. In: Fraser, N.C. & Sues, H.-D. (eds.). In the Shadow of the Dinosaurs: Early Mesozoic Tetrapods. Cambridge University Press, Cambridge: 251–270 (1994).

Luo, Z.-X., Crompton, A. W., & Sun, A.-L. A new mammaliaform from the Early Jurassic and evolution of mammalian characteristics. Science 292:1535–1540 (2001).

Luo, Z.-X., et al. Mandibular and dental characteristics of Late Triassic mammaliaform Haramiyavia and their ramifications for basal mammal evolution. Proceedings of the National Academy of Sciences 112:E7101–7109 (2015).

Lyson. T. R. et al. Transitional fossils and the origin of turtles. Biology Letters 6:830–833 (2010).

Lyson, T. R. et al. Evolutionary origin of the turtle shell. Current Biology 23:1113–1119 (2013).

Lyson, T. R. et al. Origin of the unique ventilatory apparatus of turtles. Nature Communications 5:5211 (2014).

5718 Lyson, T. R. et al. Fossorial origin of the turtle shell. Current Biology 26:1887–1894 (2016).

5719 Macungo, Z. et al. *Endothiodon* (Therapsida, Anomodontia) specimens from the middle/late Permian of the
5720 Metangula Graben (Niassa Province, Mozambique) increase complexity to the taxonomy of the genus.
5721 Journal of African Earth Sciences (2020).

5722 Maddin, H. C. et al. Varanopid from the Carboniferous of Nova Scotia reveals evidence of parental care in
5723 amniotes. Nature Ecology & Evolution 4:50–56 (2019).

5724 Maddison, W. P. & Maddison, D. R. Mesquite: a modular system for evolutionary analysis. Version 3.51
5725 http://www.mesquiteproject.org. (2018).

5726 Mann, A. & Paterson, R. S. Cranial osteology and systematics of the enigmatic early 'sail-backed' synapsid
5727 Echinerpeton intermedium Reisz, 1972, and a review of the earliest pelycosaurs. Journal of Systematic
5728 Palaeontology (2019).

5729 Mann, A., et al. Reassessment of historic ‘microsaurs’ from Joggins, Nova Scotia, reveals hidden diversity
5730 in the earliest amniote ecosystem. Papers in Palaeontology (2020)

5731 Marjanović, D. Recalibrating the transcriptomic timetree of jawed vertebrates. BioRxiv preprint. (2019)

5732 Marjanović, D. & Laurin, M. Phylogeny of Paleozoic limbed vertebrates reassessed through revision and
5733 expansion of the largest published relevant data matrix. PeerJ 6:e5565 (2019).

5734 Marsh, A. D. Preliminary U-Pb detrital zircon dates from the Kayenta Formation of Arizona: Paleo Bios
5735 32(1, Suppl.): 10 (2014).

5736 Marsicano, C. A. et al. (2016) The precise temporal calibration of dinosaur origins. Proceedings of the
5737 National Academy of Sciences 113:509–513.

5738 Martin, T. Postcranial anatomy of *Haldanodon exspectatus* (Mammalia, Docodonta) from the Late Jurassic
5739 (Kimmeridgian) of Portugal and its bearing for mammalian evolution. Zool. J. Linn. Soc. 145:219–248
5740 (2005).

5741 Martin, T. Mesozoic mammals – early mammalian diversity and ecomorphological adaptations. In: Zachos,
5742 F.E., & Asher, R. (eds.) Handbook of Zoology: Mammalia: Mammalian Evolution, Diversity, and
5743 Systematics. De Gruyter, Berlin: 199–299 (2018).

5744 Martinelli, A. G. & Rougier, G. M. On *Chaliminia musteloides* (Eucynodontia Tritheledontidae) from the
5745 Late Triassic of Argentina, and the phylogeny of Ictidosauria. Journal of Vertebrate Paleontology 27:442–
5746 460 (2007).

5747 Martinelli, A. G., Soares, M. B. & Schwanke, C. Two new cynodonts (Therapsida) from the Middle-early
5748 Late Triassic of Brazil and comments on South American probainognathians. PloS ONE 11:e0162945
5749 (2016).

5750 Martinelli, A. G. & Soares, M. B. Evolution of South American cynodonts. Contribuciones del Museo
5751 Argentino de Ciencias Naturales “Bernardino Rivadavia” 6: 183-197 (2016).

5752 Martinelli, A. G. et al. The African cynodont *Aleodon* (Cynodontia, Probainognathia) in the Triassic of
5753 southern Brazil and its biostratigraphic significance. PloS ONE 12(6):e0177948 (2017a).

5754 Martinelli, A. G. et al. A new cynodont from the Santa Maria Formation, south Brazil, improves Late
5755 Triassic probainognathian diversity. Papers in Palaeontology 3:301–423 (2017b).

5756 Martinelli, A. G. et al. The Triassic eucynodont *Candelariodon barberenai* revisited and the early diversity
5757 of stem prozostrodontians. Acta Palaeontologia Polonica 62:527–542 (2017c).

5758 Martínez, R. N. & Forster, C. A. The skull of *Probelesodon sanjuanensis*, sp. nov., from the Late Triassic
5759 Ischigualasto Formation of Argentina. Journal of Vertebrate Paleontology 16:285–291 (1996).

5760 Martínez, R. N., Fernandez, E., & Alcober, O. A. A new non-mammaliaform eucynodont from the Carnian-
5761 Norian Ischigualasto Formation, northwestern Argentina. Revista Brasileira de Paleontologia 16:61–76
5762 (2013a).

5763 Martínez, R. N., et al. Vertebrate succession in the Ischigualasto Formation. In: Sereno, P.C. (ed.). Basal
5764 sauropodomorphs and the vertebrate fossils record of the Ischigualasto Formation (Late Triassic: Carnian-
5765 Norian) of Argentina. Society of Vertebrate Paleontology Memoir 12:10–30 (2013b).

5766 Melo, T. P., Martinelli, A. G., & Soares, M. B. A new gomphodont cynodont (Traversodontidae) from the
5767 Middle-Late Triassic *Dinodontosaurus* Assemblage Zone of the Santa Maria Supersequence, Brazil.
5768 Palaeontology 60:571–582 (2017).

5769 Meng, Q.-J. et al. An arboreal docodont from the Jurassic and mammaliaform ecological diversification.
5770 Science 347:764–768 (2015).

Modesto, S. P. et al. The skull and palaeoecological significance of *Labidosaurus hamatus*, a captorhinid reptile from the Lower Permian of Texas. Zoological Journal of the Linnean Society 149:237–262 (2007).

Müller, J. & Reisz, R. R. An early captorhinid reptile (Amniota, Eureptilia) from the Upper Carboniferous of Hamilton, Kansas. Journal of Vertebrate Paleontology 25:561–568 (2005).

Müller, J. & Reisz, R. R. The phylogeny of early eureptiles: comparing parsimony and Bayesian approaches in the investigation of a basal fossil clade. Systematic Biology 55:503–511 (2006).

Neveling, J. Stratigraphic and sedimentological investigation of the contact between the Lystrosaurus and the *Cynognathus* assemblage zones (Beaufort Group: Karoo Supergroup). Council for Geoscience Bulletin 137:1–165 (2004).

Olivier, C. et al. New dicynodonts (Therapsida, Anomodontia) from near the Permo-Triassic boundary of Laos: implications of dicynodont survivorship across the Permo-Triassic mass extinction and the paleobiogeography of Southeast Asian blocks. Journal of Vertebrate Paleontology (2019)

Olroyd, S. L. & Sidor, C. A. A review of the Guadalupian (middle Permian) global tetrapod fossil record. Earth-Science Reviews 171:538–597 (2017).

Olroyd, S. L., Sidor, C. A. & Angielczyk, K. D. New materials of the enigmatic dicynodont *Abajudon kaayai* (Therapsida, Anomodontia) from the lower Madumabisa Mudstone Formation, middle Permian of Zambia. Journal of Vertebrate Paleontology (2018).

Olson, E. C. Late Permian terrestrial vertebrates, U.S.A. and U.S.S.R. Transactions of the American Philosophical Society New Series 52: 3–224 (1962).

Ottone, E. G. et al. A new Late Triassic age from the Puesto Viejo Group (San Rafael depocenter, Argentina): SHRIMP U-Pb zircon dating and biostratigraphic correlations across southern Gondwana. Journal of South American Earth Sciences 56:186–199 (2014).

Pacheco, C. P. et al. *Prozostrodon brasiliensis*, a probainognathian cynodont from the Late Triassic of Brazil: second record and improvements on its dental anatomy. Historical Biology 30:475–485 (2017).

Panciroli, E., Benson, R. B. J, & Luo, Z.-X. The mandible and dentition of *Borealestes serendipitus* (Docodonta) from the Middle Jurassic of Skye, Scotland. Journal of Vertebrate Paleontology (2019).

Pardo, J. D. et al. Hidden morphological diversity among early tetrapods. Nature 546:642–645 (2017).

Peecook, B. R. et al. Updated geology and vertebrate paleontology of the Triassic Ntawere Formation of northeastern Zambia, with special emphasis on the archosauromorphs. In: Sidor, C.A. & Nesbitt S. J. (eds.). Vertebrate and Climatic Evolution in the Triassic Rift Basins of Tanzania and Zambia. Society of Vertebrate Paleontology Memoir 17:8–38 (2018a).

Peecook, B. R., Smith, R. M. H. & Sidor, C. A. A novel archosauromorph from Antarctica and an updated review of a high-latitude vertebrate assemblage in the wake of the end-Permian mass extinction. Journal of Vertebrate Paleontology 28(6):e1536664 (2018b).

Peecook, B. R. Vertebral osteology of *Hipposaurus boonstrai* (Therapsida, Biarmosuchia) from the middle Permian of South Africa, with implications for the evolution of Archosauromorpha. Journal of Vertebrate Paleontology, Program and Abstracts, 2018, 195–196 (2018c).

Philipp, R. P. et al. Proveniência por U-Pb LA-ICP-MS em zircão detrítico e idade de deposição da Formação Santa Maria, Triássico da Bacia do Paraná, RS: evidências da estruturação do Arco do Rio Grande; pp. 154–157 in VIII Symposium International on Tectonics–XIV Simpósio Nacional de Estudos Tectônicos 2013, 19–23 May 2013, Chapada dos Guimarães, Mato Grosso (2013).

Philipp, R. P. et al. Middle Triassic SW Gondwana paleogeography and sedimentary dispersal revealed by integration of stratigraphy and U-Pb zircon analysis: The Santa Cruz Sequence, Paraná Basin, Brazil. Journal of South American Earth Sciences 88:216–237 (2019).

Prum, R. O. et al. A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. Nature 526:569-573 (2015).

Pusch, L. C., Kammerer, C. F., & Fröbisch, J. Cranial anatomy of the early cynodont *Galesaurus planiceps* and the origin of mammalian endocranial characters. Journal of Anatomy 234:592–621(2019).

Rabbitt, R. D., Damiano, E. R. & Grant, J. W. Biomechanics of the semicircular canals and otolith organs. In The vestibular system (pp. 153-201). Springer, New York, NY (2004).

Rademan, Z. Radiometric dating and stratigraphic reassessment of the Elliot and Clarens formations; near Maphutseng and Moyeni, Kingdom of Lesotho, southern Africa. Unpublished Master's thesis, University of Stellenbosch (2018).

Ray, S. Endothiodont dicynodonts from the Late Permian Kundaram Formation, India. Palaeontology 43:375–404 (2000).

Ray, S. & Bandyopadhyay, S. Late Permian vertebrate community of the Pranhita-Godavari Valley, India. Journal of Asian Earth Sciences 21:643–654 (2003).

Ray, S. et al. Vertebrate fauna from the Late Triassic Tiki Formation of India: new finds and their biostratigraphic implications. The Palaeobotanist 65:47–59 (2016).

Reisz, R. R. Pelycosaurian reptiles from the Middle Pennsylvanian of North America. Bulletin of the Museum of Comparative Zoology 144:27–62 (1972).

Reisz, R. R. Pelycosauria. Handbuch der Paläoherpetologie 17A: Gustav Fischer Verlag, Stuttgart (1986).

Reisz, R. R. The cranial anatomy of basal diadectomorphs and the origin of amniotes. In: Anderson, J. S. & Sues, H.-D. (eds.). Major Transitions in Vertebrate Evolution. Indiana University Press, Bloomington:228–252 (2007).

Reisz, R. R. & Fröbisch, J. The oldest caseid synapsid from the late Pennsylvanian of Kansas, and the evolution of herbivory in terrestrial vertebrates. PLoS One 9: e94518 (2014).

Reisz, R. R. & Müller, J. Molecular timescales and the fossil record: a paleontological perspective. Trends in Genetics 20:237–241 (2004).

Reisz, R. R., Godfrey, S. J. & Scott, D. *Eothyris* and *Oedaleops*: do these Early Permian synapsids from Texas and New Mexico form a clade? Journal of Vertebrate Paleontology 29: 39–47 (2009).

Reisz, R. R., Haridy, Y. & Müller, J. *Euconcordia* nom. nov., a replacement name for the captorhinid eureptile Concordia Müller and Reisz, 2005 (non Kingsley, 1880), with new data on its dentition. Vertebrate Anatomy Morphology Paleontology 3:1–6 (2016).

Reisz, R. R., Modesto, S. P. & Scott, D. M. A new Early Permian reptile and its significance in early diapsid evolution. Proceedings of the Royal Society B 278:3731–3737 (2011).

Romer, A. S. & Price, L. I. Review of the Pelycosauria. Geological Society of America Special Papers 28: 1–538 (1940).

Rougier, G. W. et al. New Jurassic mammals from Patagonia, Argentina: a reappraisal of australosphenidan morphology and interrelationships. American Museum Novitates 3566:1–54 (2007).

5851 Rubidge, B. S. Re-uniting lost continents—fossil reptiles form the ancient Karoo and their wanderlust.
5852 South African Journal of Geology 108:135–172 (2005).

5853 Rubidge, B. S. & Hopson, J. A. A primitive anomodont therapsid from the base of the Beaufort Group
5854 (Upper Permian) of South Africa. Zoological Journal of the Linnean Society 117: 115–139 (1996).

5855 Rubidge, B. S. et al. High-precision temporal calibration of Late Permian vertebrate biostratigraphy: U-Pb
5856 zircon constraints from the Karoo Supergroup, South Africa. Geology 41: 363–366 (2013).

5857 Rubidge, B. S. & Day, M. O. Biostratigraphy of the Eodicynodon Assemblage Zone (Beaufort Group,
5858 Karoo Supergroup), South Africa. South African Journal of Geology 123:141–148 (2020).

5859 Ruf, I., Luo, Z.-X., & Martin, T. Reinvestigation of the basicranium of *Haldanodon exspectatus*
5860 (Mammaliaformes, Docodonta). Journal of Vertebrate Paleontology 33:382–400 (2013).

5861 Ruta, M. & Coates, M. I. Dates, nodes, and character conflict: addressing the lissamphibian origin problem.
5862 Journal of Systematic Palaeontology 5:69–122 (2007).

5863 Ruta, M. et al. The radiation of cynodonts and the ground plan of mammalian morphological diversity.
5864 Proceedings of the Royal Society Series B 280: 20131865 (2013).

5865 Ruta, M., Coates, M. I. & Quicke, D. L. J. Early tetrapod relationships revisited. Biological Reviews
5866 78:215–345 (2003).

5867 Sarigül, V. New archosauromorph fragments from the Dockum Group of Texas and assessment of the
5868 earliest dinosaurs in North America. Historical Biology 30:1059–1079 (2017).

5869 Schmitt, M. R. et al. On the occurrence of the traversodontid *Massetognathus ochagaviae* (Synapsida,
5870 Cynodontia) in the early Late Triassic *Santacruzodon* Assemblage Zone (Santa Maria Supersequence,
5871 southern Brazil): Taxonomic and biostratigraphic implications. Journal of South American Earth Sciences
5872 93:36–50 (2019).

5873 Schneider, J. W. et al. Late Paleozoic–early Mesozoic continental biostratigraphy — links to the Standard
5874 Global Chronostratigraphic Scale. Palaeoworld. (2019).

5875 Schoch, R. R. & Sues, H.-D. A Middle Triassic stem-turtle and the evolution of the turtle body plan. Nature
5876 523:584–587 (2015).

Schoch, R. R. & Sues, H.-D. Osteology of the Middle Triassic stem-turtle Pappochelys rosinae and the early evolution of the turtle skeleton. Journal of Systematic Palaeontology 16:927–965 (2018).

Schoch, R. R. & Sues, H.-D. The origin of the turtle body plan: evidence from fossils and embryos. Palaeontology (2019).

Schudack, M. Geological seeting and dating of the Guimarota beds. Pp. 21–26 in T. Martin and B. Krebs (eds.), Guimarota: a Jurassic Ecocsystem. (Verlag Dr. Friedrich Pfeil, 2000a).

Schudack, M. Ostracods and charophytes from the Guimarota beds. Pp. 33–36 in T. Martin and B. Krebs (eds.), Guimarota: a Jurassic Ecocsystem (Verlag Dr. Friedrich Pfeil, 2000b).

Sciscio, L. et al. Magnetostratigraphy across the Triassic-Jurassic boundary in the main Karoo Basin. Gondwana Res.177–192 (2017).

Sennikov, A. G. An enigmatic reptile from the Upper Permian of the Volga River Basin. Paleontological Journal 31:94–101 (1997).

Sennikov, A. G. & Golubev, V. K. Vyazniki Biotic Assemblage of the terminal Permian. Paleontological Journal 40 (suppl.4):S475–S481 (2006).

Sennikov, A. G. & Golubev, V. K. Sequence of Permian tetrapod faunas of Eastern Europe and the Permian-Triassic ecological crisis. Paleontological Journal **51**, 600–611 (2017).

Sidor, C. A. & Hopson, J. A. Ghost lineages and “mammalness”: assessing the temporal pattern of character acquisition in the Synapsida. Paleobiology 24:254–273 (1998).

Sidor, C. A. & Hopson, J. A. *Cricodon metabolus* (Cynodontia: Gomphodtontia) from the Triassic Ntawere Formation of northeastern Zambia: patterns of tooth replacement and a systematic review of the Trirachodontidae. In: Sidor, C.A. & Nesbitt S. J. (eds.). Vertebrate and Climatic Evolution in the Triassic Rift Basins of Tanzania and Zambia. Society of Vertebrate Paleontology Memoir 17:39–64 (2018).

Sidor, C. A. & Rubidge, B. S. *Herpetoskylax hopsoni*, a new biarmosuchian (Therapsida: Biarmosuchia) from the Beaufort Group of South Africa. In: Carrano, M.T., Gaudin, T.J., Blob, R.W. & Wible, J.R. (eds.). Amniote Paleobiology: Perspectives on the Evolution of Mammals, Birds, and Reptiles. University of Chicago Press, Chicago: 76–113 (2006).

5903 Sidor, C. A. & Smith, R. M. H. A second burnetiamorph therapsid from the Permian Teekloof Formation of
5904 South Africa and its associated fauna. Journal of Vertebrate Paleontology 27: 420–430 (2007).

5905 Sidor, C. A. & Welman, J. A second specimen of *Lemurosaurus pricei* (Therapsida: Burnetiamorpha).
5906 Journal of Vertebrate Paleontology 23: 631–642 (2003).

5907 Sidor, C. A. et al. Tapinocephalids (Therapsida, Dinocephalia) from the Permian Madumabisa Mudstone
5908 Formation (Lower Karoo, Mid-Zambezi Basin) of southern Zambia. Journal of Vertebrate Paleontology 34:
5909 980–986 (2014).

5910 Sidor, C. A. et al. Tetrapod fauna of the lowermost Usili Formation (Songea Group, Ruhuhu Basin) of
5911 southern Tanzania, with a new burnetiid record. Journal of Vertebrate Paleontology 30: 696–703 (2010).

5912 Sigogneau-Russell, D. Theriodontia I. Handbuch der Paläoherpetologie 17B. Gustav Fischer Verlag,
5913 Stuttgart (1989).

5914 Simões, T. R. et al. The origin of squamates revealed by a Middle Triassic lizard from the Italian Alps.
5915 Nature 557:706–709 (2018).

5916 Simon, R. V. et al. First record of a tapinocephalid (Therapsida: Dinocephalia) from the Ruhuhu Formation
5917 (Songea Group) of southern Tanzania. Journal of Vertebrate Paleontology 30: 1289–1293 (2010).

5918 Smith, R. M. H. & Botha-Brink, J. Anatomy of a mass extinction: sedimentological and taphonomic
5919 evidence for drought-induced die-offs at the Permo-Triassic boundary in the main Karoo Basin, South
5920 Africa. Palaeogeography, Palaeoclimatology, Palaeoecology 15: 99–118 (2014).

5921 Smith, R. M. H. & Evans, S. E. New material of Youngina: evidence of juvenile aggregation in Permian
5922 diapsid reptiles. Palaeontology 39:289–303 (1996).

5923 Smith, R. M. H. & Keyser, A. W. Biostratigraphy of the Tropidostoma Assemblage Zone. In: Rubidge, B.S.
5924 (ed.). Biostratigraphy of the Beaufort Group (Karoo Supergroup). South African Committee for
5925 Stratigraphy Biostratigraphic Series 1:18–22 (1995).

5926 Smith, R. M. H. & Swart, R. Changing fluvial environments and vertebrate taphonomy in response to
5927 climatic drying in a mid-Triassic rift valley fill: the Omingonde Formation (Karoo Supergroup) of central
5928 Namibia. Palaios 17:249–267 (2002).

Smith, R. M. H., Rubidge, B. S. & Sidor, C. A. A new burnetiid (Therapsida: Biarmosuchia) from the upper Permian of South Africa and its biogeographic implications. Journal of Vertebrate Paleontology 26:331–343 (2006).

Smith, R. M. H., Rubidge, B. S. & van der Walt, M. Therapsid biodiversity patterns and environments of the Karoo Basin, South Africa. In: Chinsamy-Turan, A. (ed.). Forerunners of Mammals: Radiation, Histology, Biology. Indiana University Press, Bloomington and Indianapolis: 223–246 (2012).

Smith, R. M. H. Biostratigraphy of the Cistecephalus Assemblage Zone (Beaufort Group, Karoo Supergroup), South Africa. South African Journal of Geology 123:181–190 (2020).

Soares, M. B., Schultz, C. L., & Horn, B. L. D. New information on Riograndia guaibensis Bonaparte, Ferigolo & Ribeiro, 2001 (Eucynodontia, Trithelodontidae) from the Late Triassic of southern Brazil: anatomical and biostratigraphic implications. Anais da Academia Brasileira de Ciências 83:329–354 (2011).

Spiekman, S. N. F. A new specimen of Prolacerta broomi from the lower Fremouw Formation (Early Triassic) of Antarctica, its biogeographical implications and a taxonomic revision. Scientific Reports 8:17996 (2018).

Spindler, F. Reviewing the question of the oldest therapsid. – Paläontologie, Stratigraphie, Fazies (22), Freiberger Forschungshefte, C 548: 1–7 (2014).

Spindler, F., Falconnet, J. & Fröbisch, J. Callibrachion and Datheosaurus, two historical and previously mistaken basal caseasaurian synapsids from Europe. Acta Palaeontologica Polonica 61:597–616 (2016).

Spindler, F. The skull of *Tetraceratops insignis* (Synapsida, Sphenacodontia. Palaeovertebrata 43(1)-e1 (2020).

Stefanello, M. et al. Skull anatomy and phylogenetic assessment of a large specimen of Ecteniniidae (Eucynodontia; Probainognathia) from the Upper Triassic of southern Brazil. Zootaxa 4457:351–378 (2018).

Sues, H.-D. First record of the tritylodontid Oligokyphus (Synapsida) from the lower Jurassic of western North America. Journal of Vertebrate Paleontology 5:328–335 (1985).

Sues, H.-D. & Jenkins, F. A. The postcranial skeleton of Kayentatherium wellesi from the Lower Jurassic Kayenta Formation of Arizona and the phylogenetic significance of postcranial features. In: Carrano, M.

T., Gaudin, T.J., Blob, R.W. & Wible, J.R. (eds.). Amniote Paleobiology: Perspectives on the Evolution of Mammals, Birds, and Reptiles. University of Chicago Press, Chicago: 114–152 (2006).

Suh, A. The phylogenomic forest of bird trees contains a hard polytomy at the root of Neoaves. Zoologica Scripta 45:50-62 (2016).

Sullivan, C. et al. Pelvic morphology of a tritylodontid (Synapsida: Eucynodontia) from the Lower Jurassic of China, and some functional and phylogenetic implications. Comptes Rendus Palevol 12:505–518 (2013).

Sumida, S. S. & Berman, D. S. The pelycosaurian (Amniota: Synapsida) assemblage from the Late Pennsylvanian Sangre de Cristo formation of central Colorado. Annals of Carnegie Museum 62:293–310 (1993).

Sumida, S. S., Lombard, R. E., & Berman, D. S. Morphology of the atlas-axis complex of the late Paleozoic tetrapod suborders Diadectomorpha and Seymouriamorpha: Philosophical Transactions of the Royal Society of London B 336:259–273 (1992).

Świło, M. Niedźwiedzki, G. & Sulej, T. Mammal-like tooth from the Upper Triassic of Poland. Acta Palaeontologia Polonica 59:815–820 (2014).

Tattersall, G. J. et al. Seasonal reproductive endothermy in tegu lizards. Science advances, 2(1):e1500951 (2016).

Tohver, E. et al. Magnetostratigraphic constraints on the age of the lower Beaufort Group, western Karoo Basin, South Africa, and a critical analysis of existing U-Pb geochronological data. Geochemistry, Geophysics, Geosystems 3649–3665(2015).

Trujillo, K. C. & Kowallis, B. J. Recalibrated legacy 40Ar/39Ar ages for the Upper Jurassic Morrison Formation, Western Interior, U.S.A. Geology of the Intermountain West 2:1–8 (2015).

Trujillo, K. C. et al. A U/Pb age for the Mygatt-Moore Quarry, Upper Jurassic Morrison Formation, Mesa County, Colorado. Volumina Jurassica 12:107–114 (2014).

Turner, A. H., Pritchard, A. C., & Matzke, N. J. Empirical and Bayesian approaches to fossil-only divergence times: a study across three reptile clades. PloS One 12(2):e0169885 (2017).

5981 Turner, C. E. & Peterson, F. Biostratigraphy of dinosaurs in the Upper Jurassic Morrison Formation of the
5982 Western Interior, U.S.A. In Gillette, D.D. (ed.). Vertebrate Paleontology in Utah. Utah Geological Survey
5983 Miscellaneous Publication 99-1:77–114 (1999).

5984 van Tuinen, M. & Hadly, E. A. Error in estimation of rate and time inferred from the early amniote fossil
5985 record and avian molecular clocks. Journal of Molecular Evolution 59:267–276 (2004).

5986 Van den Brandt, M. J. & Abdala, F. Cranial morphology and phylogenetic analysis of Cynosaurus suppostus
5987 (Therapsida, Cynodontia) from the upper Permian of the Karoo Basin, South Africa. Palaeontologia
5988 Africana 52:201–221 (2018).

5989 Velazco, P. M., Buczek, A. J., & Novacek, M. J. Two new tritylodontids (Synapsida, Cynodontia,
5990 Mammaliamorpha from the Upper Jurassic, southwestern Mongolia. American Museum Novitates 3874:1–
5991 35 (2017).

5992 Viglietti, P.A. et al. The *Daptocephalus* Assemblage Zone (Lopingian), South Africa: A proposed
5993 biostratigraphy based on a new compilation of stratigraphic ranges. Journal of African Earth Sciences
5994 113:153–164 (2016).

5995 Viglietti, P. A., Rubidge, B. S., & Smith, R. M. H. Revised lithostratigraphy of the upper Permian Balfour
5996 and Teekloof formations of the main Karoo Basin, South Africa. South African Journal of Geology 120:45–
5997 60 (2017).

5998 Viglietti, P. A., Smith, R. M. H., & Rubidge, B. S. Changing palaeoenvironments and tetrapod populations
5999 in the *Daptocephalus* Assemblage Zone (Karoo Basin, South Africa) indicate early onset of the Permo-
6000 Triassic mass extinction. Journal of African Earth Sciences 138:102–111 (2018a).

6001 Viglietti, P. A. et al. U-Pb detrital zircon dates and provenance data from the Beaufort Group (Karoo
6002 Supergroup) reflect sedimentary recycling and air-fall tuff deposition in the Permo-Triassic Karoo foreland
6003 Basin. Journal of African Earth Sciences 143:59–66 (2018b).

6004 Viglietti, P. A. Biostratigraphy of the *Daptocephalus* Assemblage Zone (Beaufort Group, Karoo
6005 Supergroup), South Africa. South African Journal of Geology 123:191–206 (2020).

6006 Viglietti, P. A. et al. Biostratigraphy of the *Massospondylus* Assemblage Zone (Stormberg Group, Karoo
6007 Supergroup), South Africa. South African Journal of Geology 123:149–262 (2020).

Wallace, R.V. S., Martínez, R. & Rowe, T. First record of a basal mammaliamorph from the early Late Triassic Ischigualasto Formation of Argentina. PloS ONE 14(8):e0218791 (2019).

Whiteside, D. I. et al. The Late Triassic and Early Jurassic fissure faunas from Bristol and South Wales: stratigraphy and setting. Palaeontologia Polonica 67:257–287 (2016).

Yang, W. et al. Depositional environments and cyclo- and chronostratigraphy of uppermost Carboniferous–Triassic fluvial–lacustrine deposits, southern Bogda Mountains, NW China — a terrestrial paleoclimatic record of mid-latitude Pangea. Global and Planetary Change 73:15–113 (2010).

Zhou, C.-F. et al. A Jurassic mammaliaform and the earliest mammalian evolutionary adaptations. Nature 500:163–167 (2013).