Behavioural adaptations in egg laying ancestors facilitate evolutionary transitions to live birth

Amanda K. Pettersen[1*], Nathalie Feiner[1], Daniel W.A. Noble[2], Geoffrey M. While[3], Charlie K. Cornwallis[1#] & Tobias Uller[1#]

1. Department of Biology, Lund University, Lund, 22 362, Sweden
2. Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, 2600, Australia
3. School of Natural Sciences, University of Tasmania, Hobart, 7005, Australia

#Joint senior author

*Corresponding author: Amanda K. Pettersen. Email: amanda.pettersen@sims.org.au

Keywords: Viviparity, Oviparity, Thermal plasticity, Behavioural plasticity, Reproductive mode

Abstract

Live birth is a key innovation that has evolved from egg laying over 100 times in reptiles. One significant feature in this transition is the thermal conditions experienced by developing embryos. Adult lizards and snakes often have preferred body temperatures that can be lethal to developing embryos and should prevent egg retention: how has viviparity repeatedly evolved in the face of this pervasive mismatch? Here we resolve this paradox by conducting phylogenetic analyses using data on thermal preference from 224 species. Thermal mismatches between mothers and offspring are widespread but resolved by gravid females behaviourally down-regulating their body temperature towards the thermal optimum of embryos. Importantly, this thermoregulatory behaviour evolved in ancestral egg-laying species before the evolutionary emergence of live birth. Maternal thermoregulatory behaviour therefore bypasses constraints imposed by a slowly evolving thermal physiology and is likely to have been a key requirement for repeated transitions to live birth.

Introduction

The evolution of live bearing is an important life-history transition in vertebrates (1–3). The ecological conditions that favour the transition from egg laying (oviparity) to live birth (viviparity) are relatively well understood, especially in reptiles, with particularly strong support for the adaptive value of viviparity in cool climates (3, 4). By retaining embryos throughout development, mothers can buffer offspring from suboptimal nest temperatures, thereby ensuring higher hatching success and increased offspring viability (4). Live bearing also allows females to maintain offspring at higher temperatures than those of potential nest sites, effectively increasing developmental rates to facilitate early-season hatching (5, 6). This has allowed reptile species to diversify and persist in cool climates globally (7).

Despite its adaptive advantage, it is unclear how viviparity initially evolved. A major challenge to the evolution of viviparity is that adult lizards and snakes typically have preferred body temperatures that exceed the upper lethal limit of embryos (8). For example, the average nest temperature for the Iberian emerald lizard, *Lacerta schreiberi,* is 24°C, rarely exceeding 30°C, while the preferred body temperature of females is 33°C (9). Since embryos are well adapted to the temperatures they typically experience in the nest, they generally have limited capacity to sustain development at temperatures outside a narrow thermal window (10, 11). Prolonged exposure to temperatures beyond this range can result in offspring abnormalities and even death (12). A mismatch between thermal optima of embryos and females should therefore prevent the maternal retention of embryos throughout development, and therefore evolutionary transitions to live birth (8).

Despite this apparent constraint, live birth has evolved over 100 times in squamate reptiles (13, 14). How can the repeated evolution of viviparity be reconciled with the widespread thermal mismatch between embryos and adults? There are two potential scenarios by which viviparity can evolve in the presence of maternal-embryo thermal mismatches.

First, viviparity may only evolve in species where there is no mismatch, that is, adult body temperature (P_{bt}) and embryo thermal optima (T_{opt}) are well aligned. Second, females may behaviourally adjust their body temperature when pregnant to close the gap between adult and embryo thermal optima, even when substantial mismatches exist. Such plasticity may come at a cost to female performance, but temporarily shifting body temperature while gravid to match the thermal optima of embryos could eliminate thermal barriers to the evolution of viviparity.

Elucidating how the mismatch between female and offspring thermal optima have been resolved in viviparous lineages is challenging. Retracing the evolutionary steps that have led to viviparity requires the reconstruction of thermal preferences of adults and embryos at the very origins of these lineages. Here, we collate data on the thermal preferences of female adults (163 species) and embryos (50 species) of squamate reptiles, and use phylogenetic analyses to reconstruct the evolutionary history of live birth (Table S1). First, we estimated the ancestral states of preferred body temperature of females and embryos to quantify the thermal conflict between mothers and embryos at the origins of viviparity. Second, we tested the prediction that transitions to viviparity have been restricted to oviparous lineages with a low degree of maternal-embryo mismatch. This was assessed by calculating if viviparity has originated more frequently from oviparous ancestors where female *Pbt* overlaps with the thermal optima of embryos. Finally, we examined if females adjust their body temperature when gravid to reduce the mismatch between adult and embryo thermal optima, and if so, whether this behaviour evolved prior to the emergence of viviparity, or afterwards in response to such mismatch.

Results and Discussion

Wide-spread mismatch in thermal optima of oviparous mothers and their embryos

Across oviparous lizards and snakes, there were many species with a significant disparity in preferred body temperature (P_{bt}) and the temperature that maximizes hatching success (T_{opt}) (Multi-response BPMM (MR-BPMM): posterior mode (PM) of difference = 5.60, 95% Credible Interval (CI): -0.60, 10.18, *pMCMC* = 0.06. Figure 1 & 2; Table S3). The female P_{bt} of oviparous species is on average 4 °C higher than T_{opt} (raw data mean ± SD P_{bt} 102 species: 31.40 ± 4.31 ºC; T_{opt} 50 species: 27.25 ± 2.42 ºC). Incubation experiments have shown that such a difference can significantly reduce hatching success, illustrating that the retention of offspring throughout development may jeopardise embryo survival (15, 16); reviewed in (17). P_{bt} and T_{opt} of oviparous species were both estimated to have high phylogenetic inertia (Figure 1; MR-BPMM: P_{bt} phylo H^2: 0.99, CI: 0.91, 1.0. T_{opt} phylo H^2: 0.93, CI: 0.72, 0.98. Table S3) and female P_{bt} and offspring T_{opt} only show weak evidence of coevolving (Figure 2; phylogenetic correlation (MR-BPMM): PM = 0.57, CI: -0.09, 0.97, *pMCMC* = 0.08. Table S4), suggesting that this mismatch may be difficult to resolve.

Transitions to live birth occur despite mismatches between females and embryos

We found evidence that preferred body temperature was lower in viviparous, compared with oviparous, lineages (MR-BPMM: PM = -0.20, CI: -0.43, -0.02, *pMCMC* = 0.02; Table S6). This is consistent with the hypothesis that viviparity is an adaptation to cool climates (2, 3, 18). However, reconstructions of ancestral preferred body temperatures of adults indicated that in the majority of lineages, lower preferred body temperatures evolved after, not before, transitions to live bearing. This was evident from the lack of difference in P_{bt} between ancestors of egg-laying and live-bearing lineages (MR-BPMM, PM = 0.04, CI: -3.96, 4.27, *pMCMC* = 0.42; Table S7). Additionally, live-bearing lineages with egg-laying ancestors

showed higher P_{bt} than those with live-bearing ancestors (MR-BPMM, PM = 6.78, CI: -0.64, 8.74, *pMCMC* = 0.06; Table S7).

Ancestral reconstructions of *Pbt* and *Topt* of oviparous species showed that transitions from egg laying to live bearing were no more likely to occur when adult and embryo thermal optima were similar (% of lineages with thermal mismatches: ancestors of live bearing 84%, ancestors of egg laying species 85%. Table S5). Only 5 of the estimated 31 evolutionary origins of live birth had adult and embryo thermal optima that overlapped, which was similar to the proportion of lineages where live bearing did not occur (Table S5). In the remaining 26 egg-laying ancestors of live-bearing species, the thermal optimum of embryos was significantly lower than the preferred body temperature of adults (there were no cases of T_{opt} being significantly higher than P_{bt}; Table S5). In summary, the degree of similarity between adult P_{bt} and embryo T_{opt} does not predict transitions towards viviparity, and the lower P_{bt} observed for viviparous taxa appears to have evolved after transitions to live birth, not before.

Female behavioural plasticity resolves the constraints imposed by thermal physiology

Next, we tested the hypothesis that the thermal mismatch between females and embryos during the evolution of viviparity is resolved by females adjusting their body temperature when gravid to meet the thermal requirements of their embryos (19). We collected data on the preferred body temperature of gravid ($P_{bt\text{-}g}$) and non-gravid ($P_{bt\text{-}ng}$) adult females from 52 species of lizards and snakes were collected from published field and experimental studies (live bearing n = 32, egg laying n = 20. Table S1). Using these data, an effect size of female behavioural plasticity, Hedges' *g* was calculated (positive values = higher P_{bt} when gravid, negative values = lower P_{bt} when gravid).

Across those 52 species, there was a negative phylogenetic correlation between P_{bt} and the degree to which females reduced their body temperature when gravid (MR-BPMM:

PM = -0.29, CI: -0.62, -0.02, *pMCMC* = 0.02. Figure 4; Table S8). Females with high *Pbt* have therefore evolved to behaviourally reduce their body temperature when gravid to a greater extent than females with low *Pbt*. This relationship was not restricted to viviparous lineages, which is expected if female behavioural plasticity is an adaptation that evolves in response to selection for embryo retention. Instead, comparable phylogenetic correlations between Hedges' *g* and *Pbt* were evident in both oviparous and viviparous taxa, showing that behavioural adjustment of temperature when gravid occurs to a similar extent in both egg laying and live bearing species (Egg laying: PM = -0.61, CI: -0.92, -0.08, *pMCMC* = 0.02. Live bearing: PM = -0.88, CI: -0.96, -0.30, *pMCMC* = 0.005. Table S9).

To further investigate the role of female plasticity in the evolution of live bearing, we used phylogenetic models to reconstruct values of Hedges' *g* for the ancestors of egg-laying and live-bearing species. Consistent with the results from our correlational analyses, we found that, in the ancestors of both egg laying and live bearing species, thermal mismatches between adults and embryos were resolved by females adjusting their body temperature when gravid (Figure 4A; MR-BPMM: Egg-laying Hedges' *g* PM = -0.81, CI = -1.43, 0.05, *pMCMC* = 0.03. Live bearing Hedges' *g* PM = -0.75, CI = -1.88, -0.09, *pMCMC* = 0.03. Table S10). As a result, Hedges' *g* did not differ between the ancestors of oviparous and viviparous species (Figure 4B; Table S9-S11), suggesting that behavioural plasticity was likely present prior to the emergence of live birth.

Our results illustrate that an ancient and evolutionarily conserved ability of egg laying females to adjust their body temperature during pregnancy provides a general resolution to the conflict over adult and embryo thermal optima that paves the way for the evolution of live bearing. The down-regulation of female body temperature in egg-laying species while gravid may appear surprising considering that most of these species lay their eggs within the early stages of development (commonly around the time of limb bud formation (20)). However,

early developmental stages, involving gastrulation, neurulation and organogenesis, are potentially even more sensitive to temperature stress than later stages, which are predominantly associated with growth (8, 10). The temperature sensitivity of early-stage embryos therefore suggests that resolutions to mother-offspring thermal conflicts are required in both egg-laying and live-bearing species. If true, the key innovation of live birth may owe its evolutionary origin to mechanisms of temperature regulation put in place long before live birth emerged. We suggest that this behavioural flexibility of female lizards and snakes enabled the gradual evolution of live birth by helping traverse potential fitness valleys associated with embryo retention. In many live-bearing lineages, temperature mismatches have subsequently been eliminated through the evolution of a lower preferred body temperature in females. However, behavioural flexibility is frequently maintained in viviparous lineages that have colonized cool climates, possibly as a means for mothers to maintain embryos at temperatures significantly warmer than the external environment (2, 5, 6, 21, 22).

In conclusion, female thermoregulatory behaviour is a key adaptation that overcomes the constraints imposed by thermal mismatches between mothers and embryos, and has facilitated both the evolution of live birth, and expansion of reptiles into a variety of environments.

Methods

Literature search and data collection

To investigate the relationship between maternal behavioural plasticity and embryo thermal sensitivity, we collected existing data on reproductive mode (oviparous versus viviparous), the female preferred body temperature (P_{bt}) ("P_{bt} dataset"), the temperature at which hatching success was maximised (T_{opt}) ("T_{opt} dataset"), and female body temperature when gravid ($P_{bt\text{-}g}$) and not gravid ($P_{bt\text{-}ng}$) ("Hedges' *g* dataset"). Complete data are presented in Table S1. Data were collected for each of the variables from published literature. We conducted a literature search using ISI *Web of Science* (v.5.30) with search terms specific to each dataset.

The P_{bt} dataset provided preferred body temperatures (P_{bt}) of adult females measured at one point in time. We used the search terms 'body temperature*', along with one of the following: 'squamat*', 'lizard*', 'snake*' which yielded a total of 1075 papers. We only used data from studies where P_{bt} for females was stated explicitly, unless pooled male/female data stated no significant effect of sex. Data on females that were described to be gravid or data collected during the reproductive season were excluded. This provided a final dataset of female preferred body temperature for 163 species that was independent of the gravid and non-gravid measures used to calculate Hedges' *g* (live bearing: n = 61 and egg laying: n = 102).

The T_{opt} dataset quantified the temperature that maximises embryonic survival (measured as hatching success; T_{opt}) in 50 egg laying species. Hatching success data were taken from the Reptile Developmental Database (23) and supplemented with additional studies using the same method described for the P_{bt} dataset (search terms: 'temperature* AND incubat* AND hatch* OR surv*', along with one of the following: 'squamat*', 'lizard*', 'snake*'), yielding a total of 671 papers. We only included studies where three or more constant temperature treatments were used under controlled laboratory conditions, resulting in 661 papers being

179 rejected due to irrelevance or overlap with the Reptile Development Database. Given that T_{opt}
180 had a strong phylogenetic signature (Phylogenetic Heritability (H^2) = 91.93%, 95% CI: 71.65
181 – 98.16%; Table S3) we fitted a Bayesian Phylogenetic Mixed Effects Model (BPMM) to
182 hatching success data for all species jointly. Including phylogenetic information allowed for
183 the T_{opt} of each species to be estimated with greater accuracy and precision given that the
184 range and number of temperatures across species varied (range: 10-40 ºC, mean number of
185 temperatures per species ± SD: 7.62 ± 4.29). Compared to non-phylogenetic models, T_{opt}
186 estimates produced from phylogenetic models (BPMM) showed smaller sampling error and
187 avoided convergence problems in estimating model parameters. Importantly, this approach
188 was not used to estimate T_{opt} data for species without data, only to better predict T_{opt} values
189 for species for which there were data. The T_{opt} BPMM model was run for 1,100,000 iterations
190 with a burn-in of 10,000 iterations and thinning rate of 500, leaving us with 2,000 samples
191 from the posterior distribution. Autocorrelation was low (lag values < 0.1) and trace plots
192 showed chains mixed well for all parameters. Our model included temperature as a fixed
193 effect (estimating both a linear and quadratic slopes) and random slopes of temperature
194 (linear and quadratic slopes) fitted at the phylogenetic level. From our BPMM we estimated
195 T_{opt}, and its corresponding sampling variance, using the posterior distribution of fixed effects
196 and best linear unbiased predictors (BLUPs) for the random slopes (linear and quadratic) for
197 each species as follows:

198 $$Topt = -\frac{(T_f + T_{sp})}{\left(2\left(T_f^2 + T_{sp}^2\right)\right)}$$

199 Where T_f and T_f^2 are the posterior linear and quadratic fixed effect estimates for temperature
200 and T_{sp} and T_{sp}^2 are the posterior BLUPs for a given species extracted from the phylogenetic
201 random slopes. Calculating T_{opt} using the posterior distribution of fixed and random effects

meant that sampling error for a given species could be propagated to subsequent analyses (see below).

For the final "Hedges' *g*" dataset, we collected mean, standard deviation and sample sizes for female preferred body temperatures when gravid ($P_{bt\text{-}g}$) and non-gravid ($P_{bt\text{-}ng}$) (independent from the P_{bt} dataset) from the same study/population. Within the 'title', 'abstract' or 'keywords' we used search terms 'body temperature* AND gravid* OR reproduct*' along with one of the following: 'squamat*', 'lizard*', 'snake*' which yielded a total of 721 papers. We included both field and laboratory measures of female body temperature, comprising studies that directly compared $P_{bt\text{-}g}$ and $P_{bt\text{-}ng}$. Only studies that provided both sample size and error around mean of preferred body temperature were included. Of the 721 papers 648 papers were excluded due to irrelevance or insufficient data. The majority of studies used artificial temperature gradients in the laboratory (n = 37) or measured preferred basking temperature in the field (n = 36). Laboratory studies generally measured body temperature in the same female during gestation ($P_{bt\text{-}g}$) and either before or after gestation ($P_{bt\text{-}ng}$) as repeated measures. In contrast, field studies often measured body temperature on a population during the reproductive season, comparing body temperatures in gravid and non-gravid females at a single time point. Combined, this yielded a total of 73 studies published up to July 2019 from which effect sizes were calculated for 52 species (live bearing: n = 32 and egg laying: n = 20). An effect size of female adjustment of body temperature when gravid for each species was measured as the standardised mean differences (Hedges' *g*) in preferred body temperatures between non-gravid and gravid females ($P_{bt\text{-}g}$ - $P_{bt\text{-}ng}$), adjusting for small sample sizes (24). We found no significant differences in means or variances of Hedges' *g* between laboratory or field studies ("Study type").

The final full dataset (Table S1) contained a total of 224 squamate species. Two species in our data are both oviparous and viviparous (*Saiphos equalis* and *Zootoca vivpara)*.

While for *S. equalis* all data were from an oviparous population, for *Z. vivipara* there were data for both oviparous and viviparous populations and we treated these different reproductive modes as being distinct in our dataset (25).

Quantification and Statistical Analysis

Data were analysed using Phylogenetic Generalised Least Squares (PGLS), Bayesian Phylogenetic Mixed Effects Models with single (BPMM) and multiple response variables (MR-BPMM), and Stochastic Character Mapping (SCM). All analyses were conducted in R version 4.0.1 (26). Ancestral states were reconstructed by initially examining the mode of evolution that best explained how traits have evolved (a model of Brownian motion best explained all traits: Table S2), after which Bayesian Phylogenetic Mixed models (BPMM) were used to simultaneously reconstruct ancestral states of all traits. The robustness of ancestral reconstructions was examined using stochastic character mapping (SCM; Table S12).

Assessing modes of evolution using Phylogenetic Generalised Least Squares (PGLS)

An assumption of many phylogenetic comparative methods is that traits evolve via a Brownian motion process (27). This assumption was assessed by estimating the mode of evolution that best describes the phylogenetic distribution of reproductive mode, *PBT* and *Topt* using PGLS implemented with R packages 'ape' and 'nlme' (28, 29). This was done by comparing PGLS models with different correlation structures constructed from the phylogeny that corresponded to BM, Ornstein-Uhlenbeck (OU: corMartins in 'ape') and Accelerated Decelerated (ACDC: corMartins in 'ape') modes of evolution. In OU models' traits undergo a BM process but evolve towards a central tendency with a strength proportional to parameter alpha (α), and in accelerated/decelerated models of evolution (ACDC) traits diverge under a

252 BM model, but rates of evolution accelerate or decelerate over time according to parameter g
253 (Acceleration: g < 1. BM: g = 1. Deceleration g >1). Estimating parameters, such as α and g,
254 directly from the data can be unstable (30, 31). Therefore, we estimated values of α (OU) and
255 Bloomberg's g (ACDC) by fixing α and g at 0.1 increments within the range of 0 to 10 and
256 used AIC values to identify best-fit models. For all traits there was significant phylogenetic
257 signature and BM best explained the way traits have evolved (Table S2: R code: Section 0).
258
259 *Bayesian Phylogenetic Mixed Effects models* (BPMM)
260 We implemented BPMMs in *R* with the *MCMCglmm* package (32). Unlike other generalised
261 linear mixed model packages, *MCMCglmm* can incorporate inverse phylogenetic correlation
262 matrices with missing data in response variables (i.e., trait values for both T_{opt} and P_{bt} were
263 not available for all species). BPMMs allow missing values to be predicted with an accuracy
264 related to the phylogenetic signature in traits and the strength of phylogenetic correlations
265 between traits. All traits had high phylogenetic signature (phylogenetic heritability >80%)
266 producing high correspondence between raw and predicted values (Figure S1). As a result,
267 our BPMMs enabled us to deal with the fact that not all traits have been measured in all
268 species.
269
270 Phylogenetic relationships were modelled using a published phylogeny comprising over 4000
271 species of squamates (33) that was pruned to the 224 species in our dataset using the *ape*
272 package in *R* (28). The random effect *animal*, with a covariance matrix derived from the
273 phylogenetic tree was included in all models (32). We calculated the phylogenetic signature
274 (equivalent to heritability, H^2, in the terminology of *MCMCglmm*) for each trait as the
275 variance explained by *animal* relative to total variance (animal and residual). In the case of
276 binary variables (reproductive mode) the residual variance is not identifiable and we

therefore used the intraclass coefficient of variation (ICC: phylogenetic variance / (phylogenetic variance + residual variance +$pi^{2/3}$)*100) to examine phylogenetic signature (34). More traditional methods were also used to validate estimates of phylogenetic signature. Pagel's lambda (λ) from *pgls* models (35, 36) and Blomberg's K (K) in *phytools* using the function *phylosig* (37, 38), which gave comparable estimates to BPMMs (see R code section 0).

Multi-response BPMMs (MR-BPMMs) can also be fitted with *MCMCglmm* that allow the phylogenetic and within-species (residual) correlations between traits to be estimated. Correlations between traits (e.g., A & B) were calculated as:

$$\frac{cov(A,B)}{\sqrt{(var(A)\cdot var(B))}}$$

Missing data imputation for missing sampling variances

The accuracy of measures of T_{opt}, P_{bt}, and Hedges' g varied across species due to study design and sample sizes which can be accounted for by weighting data points by their inverse sampling variance. However, missing values in sampling variances are not permitted in *MCMCglmm*. As data on the error and sample size was missing for T_{opt}, P_{bt} and Hedges' g it would not have been possible to account for sampling error in our analyses without drastically reducing the size of our dataset. Consequently, we used multiple imputation with predictive mean matching in the *mice* package in R to impute missing error and sample sizes (39). To incorporate uncertainty in imputations, 20 complete datasets were generated, and all analyses were conducted by sampling across these datasets. Each model sampled through the 20 datasets 75 times (1500 sampling events) for 2000 iterations with only the last iteration being saved. Estimates from the last iteration of each sampling event i were used as the starting parameter values for the next $i + 1$. This led to a total posterior sample of 1500

301 iterations, the first 500 iterations were discarded as a burn-in and the remaining 1000 (50 per
302 dataset) were used to estimate parameters. Pooling of posterior distributions from model
303 parameters from across each of the *m = 20* datasets enabled imputation uncertainty in
304 sampling variances to be accounted for.
305
306 *Model convergence, prior settings and characterisation of posterior distributions*
307 Non-informative uniform priors were used for fixed effects and inverse-Wishart priors for
308 random effects ($V = 1$, nu = 0.002; (32)) apart from models with binary response variables
309 (e.g. reproductive mode). For binary variables the residual variance was fixed to 1 and we
310 specified a fixed effect prior that is relatively flat on the logit scale (mu = 0, V = residual
311 variance x $\pi^{2/3}$) and parameter-expanded random effect priors ($V = 1$, nu = 0.002,
312 alpha.mu=0, alpha.V = 1000). To examine model convergence we ran three independent
313 MCMC chains and examined autocorrelation, which was low (lag values < 0.1), trace plots,
314 which showed chains mixed well, and Gelman and Rubin's convergence diagnostic that
315 models converged (potential scale reduction factors were all below 1.1: R function
316 gelman.diag (40). The convergence of models of the probability of viviparity (binary
317 response variable) occurred more slowly and therefore were run for longer (burn-in =
318 1000000, iterations = 6000000, thinning = 5000). All other models used the run settings
319 specified in the data imputation section. Posterior distributions of all parameters were
320 characterised using modes and 95% credible intervals (CIs) throughout. Effects were
321 regarded as significant where CIs did not span 0. *pMCMC* (number of iterations above or
322 below 0 / total number of iterations) are also presented to facilitate general interpretation.
323
324 *Specific analyses*

We conducted six main analyses: **1)** Phylogenetic signature in P_{bt} and T_{opt} was estimated using a MR-BPMM of P_{bt} and T_{opt} with heterogenous phylogenetic and residual covariance matrices fitted as random effects (R code model M1.1. Table S3). **2)** The phylogenetic correlation between P_{bt} and T_{opt} was estimated using a MR-BPMM with unstructured phylogenetic and residual covariance matrices fitted as random effects (R code model M1.4. Table S4). **3)** To quantify the number of ancestral egg-laying with and without mismatched adult female and embryo thermal optima a two-step approach was used. The ancestral states of reproductive mode were first estimated for each node in the phylogenetic tree using a BPMM with the probability of live bearing as the response variable (R code model M3.0). Next a MR-BPPM with P_{bt}, T_{opt} and Hedges' *g* as response variables was used (R code model M3.1) to reconstruct ancestral states for each trait. From this model mismatches in thermal optima (CI of $P_{bt} - T_{opt}$ not overlapping 0) and levels of female behavioural plasticity were estimated (Table S5). Separate models were used for reproductive mode and P_{bt}, T_{opt} and Hedges' *g* to ensure that ancestral estimates were not influenced by the strong correlation between live bearing and P_{bt} in extant species. **4)** The phylogenetic correlation between the probability of live bearing and P_{bt} was estimated using a MR-BPMM with unstructured phylogenetic and residual covariance matrices fitted as random effects (R code model M2.1. Table S6). **5)** To examine if female plasticity differed between the ancestors of oviparous and viviparous lineages with and without mismatched mother-offspring thermal optima, we examined estimates of Hedges' *g* in relation to predicted states of reproductive mode (R code model M3.1. Table S7 & S10-S11). **6)** The phylogenetic correlation between P_{bt} and Hedges' *g* was estimated across all species (R code model M3.1. Table S8), as well as for egg-laying and live-bearing species separately using an MR-BPMM with separate unstructured phylogenetic and residual covariance matrices for each reproductive mode (R code model M4.1. Table S9).

Stochastic character mapping (SCM)

The robustness of ancestral state reconstructions of reproductive mode was examined by assessing the correspondence between the estimates from MCMCglmm with those obtained from SCM. The SCM approach calculates the conditional likelihood that each ancestral node is oviparous or viviparous based on an estimated transition rate matrix (Q) between reproductive modes and the length of the branch associated with that node. Based on these conditional likelihoods, ancestral states at each node are stochastically simulated and used in combination with observations at the tips to reconstruct a character history along each branch. Each character history is simulated using a continuous-time Markov model where changes between states and the time spent in each state is modelled as a Poisson process (see (41) for details). We found extremely good correspondence between estimates of ancestral states obtained from MCMCglmm models and SCM showing inferences of ancestral states were robustness to methodological details (Table S2).

364 **Data and code availability**

365 All data and code are publicly available at: https://doi.org/10.17605/OSF.IO/JT28V

References

1. J. S. Doody, J. A. Moore, Conceptual model for thermal limits on the distribution of reptiles. *Herpetological Conservation and Biology* **5**, 283–289 (2010).

2. R. Shine, "The evolution of viviparity in reptiles: An ecological analysis" in *Biology of the Reptilia*, Development B., (John Wiley & Sons, Ltd, 1985), pp. 605–694.

3. S. M. Lambert, J. J. Wiens, Evolution of viviparity: a phylogenetic test of the cold-climate hypothesis in Phrynosomatid lizards. *Evolution* **67**, 2614–2630 (2013).

4. R. Shine, A new hypothesis for the evolution of viviparity in reptiles. *The American Naturalist* **145**, 809–823 (1995).

5. D. A. Warner, R. Shine, Fitness of juvenile lizards depends on seasonal timing of hatching, not offspring body size. *Oecologia* **154**, 65–73 (2007).

6. M. Le Henanff, S. Meylan, O. Lourdais, The sooner the better: reproductive phenology drives ontogenetic trajectories in a temperate squamate (Podarcis muralis). *Biol J Linn Soc* **108**, 384–395 (2013).

7. L. Ma, L. B. Buckley, R. B. Huey, W.-G. Du, A global test of the cold-climate hypothesis for the evolution of viviparity of squamate reptiles. *Global Ecology and Biogeography* **27**, 679–689 (2018).

8. C. A. Beuchat, Temperature effects during gestation in a viviparous lizard. *Journal of Thermal Biology* **13**, 135–142 (1988).

9. C. Monasterio, L. P. Shoo, A. Salvador, P. Iraeta, J. A. Díaz, High temperature constrains reproductive success in a temperate lizard: implications for distribution range limits and the impacts of climate change. *Journal of Zoology* **291**, 136–145 (2013).

10. T. J. Sanger, J. Kyrkos, D. J. Lachance, B. Czesny, J. T. Stroud, The effects of thermal stress on the early development of the lizard Anolis sagrei. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology* **329**, 244–251 (2018).

11. W.-G. Du, R. Shine, L. Ma, B.-J. Sun, Adaptive responses of the embryos of birds and reptiles to spatial and temporal variations in nest temperatures. *Proceedings of the Royal Society B: Biological Sciences* **286**, 20192078 (2019).

12. R. M. Andrews, T. Mathies, D. A. Warner, Effect of Incubation Temperature on Morphology, Growth, and Survival of Juvenile Sceloporus undulatus. *Herpetological Monographs* **14**, 420–431 (2000).

13. D. G. Blackburn, Squamate Reptiles as Model Organisms for the Evolution of Viviparity. *Herpetological Monographs* **20**, 131–146 (2006).

14. D. G. Blackburn, Evolution of viviparity in squamate reptiles: Reversibility reconsidered. *J Exp Zool B Mol Dev Evol* **324**, 473–486 (2015).

15. R. Shine, Reptilian viviparity in cold climates: testing the assumptions of an evolutionary hypothesis. *Oecologia* **57**, 397–405 (1983).

16. R. M. Andrews, B. R. Rose, Evolution of Viviparity: Constraints on Egg Retention. *Physiological Zoology* **67**, 1006–1024 (1994).

17. D. W. A. Noble, V. Stenhouse, L. E. Schwanz, Developmental temperatures and phenotypic plasticity in reptiles: a systematic review and meta-analysis. *Biological Reviews* **93**, 72–97 (2018).

18. C. M. Watson, R. Makowsky, J. C. Bagley, Reproductive mode evolution in lizards revisited: updated analyses examining geographic, climatic and phylogenetic effects support the cold-climate hypothesis. *Journal of Evolutionary Biology* **27**, 2767–2780 (2014).

19. C. A. Beuchat, Reproductive Influences on the Thermoregulatory Behavior of a Live-Bearing Lizard. *Copeia* **1986**, 971–979 (1986).

20. R. M. Andrews, T. Mathies, Natural History of Reptilian Development: Constraints on the Evolution of Viviparity. *BioScience* **50**, 227–238 (2000).

21. R. M. Andrews, Evolution of viviparity in squamate reptiles (Sceloporus spp.): a variant of the cold-climate model. *Journal of Zoology* **250**, 243–253 (2000).

22. T. Uller, M. Olsson, Offspring size and timing of hatching determine survival and reproductive output in a lizard. *Oecologia* **162**, 663–671 (2010).

23. D. W. A. Noble, *et al*., A comprehensive database of thermal developmental plasticity in reptiles. *Sci Data* **5**, 1–7 (2018).

24. M. Borenstein, L. V. Hedges, J. P. Higgins, H. R. Rothstein, "Effect Sizes Based on Means" in *Introduction to Meta-Analysis*, (John Wiley & Sons, Ltd, 2009), pp. 21–32.

25. H. Recknagel, N. A. Kamenos, K. R. Elmer, Common lizards break Dollo's law of irreversibility: Genome-wide phylogenomics support a single origin of viviparity and re-evolution of oviparity. *Molecular Phylogenetics and Evolution* **127**, 579–588 (2018).

26. R Core Team, *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria*. (2020).

27. L. Z. Garamszegi, Ed., *Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology: Concepts and Practice* (Springer-Verlag, 2014) https:/doi.org/10.1007/978-3-662-43550-2 (December 4, 2020).

28. E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

29. J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, R. C. Team, nlme: Linear and nonlinear mixed effects models. *R package version* **3**, 111 (2013).

30. E. Paradis, *Analysis of Phylogenetics and Evolution with R*, 2nd Ed. (Springer-Verlag, 2012) https:/doi.org/10.1007/978-1-4614-1743-9 (December 4, 2020).

31. T. F. Hansen, Stabilizing selection and the comparative analysis of adaptation. *Evolution* **51**, 1341–1351 (1997).

32. J. D. Hadfield, MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software* **33**, 1–22 (2010).

33. R. A. Pyron, F. T. Burbrink, J. J. Wiens, A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology* **13**, 93 (2013).

34. S. Nakagawa, P. C. D. Johnson, H. Schielzeth, The coefficient of determination R2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of The Royal Society Interface* **14**, 20170213 (2017).

35. D. Orme, *et al*., CAPER: comparative analyses of phylogenetics and evolution in R. *Methods in Ecology and Evolution* **3**, 145–151 (2013).

36. R. P. Freckleton, P. H. Harvey, M. Pagel, Phylogenetic analysis and comparative data: a test and review of evidence. *Am. Nat.* **160**, 712–726 (2002).

37. L. J. Revell, phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217–223 (2012).

38. S. P. Blomberg, T. Garland, A. R. Ives, Testing for Phylogenetic Signal in Comparative Data: Behavioral Traits Are More Labile. *Evolution* **57**, 717–745 (2003).

39. S. van Buuren, K. Groothuis-Oudshoorn, mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software* **45**, 1–67 (2011).

40. S. P. Brooks, A. Gelman, General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics* **7**, 434–455 (1998).

41. J. P. Bollback, SIMMAP: Stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics* **7**, 88 (2006).

Figures

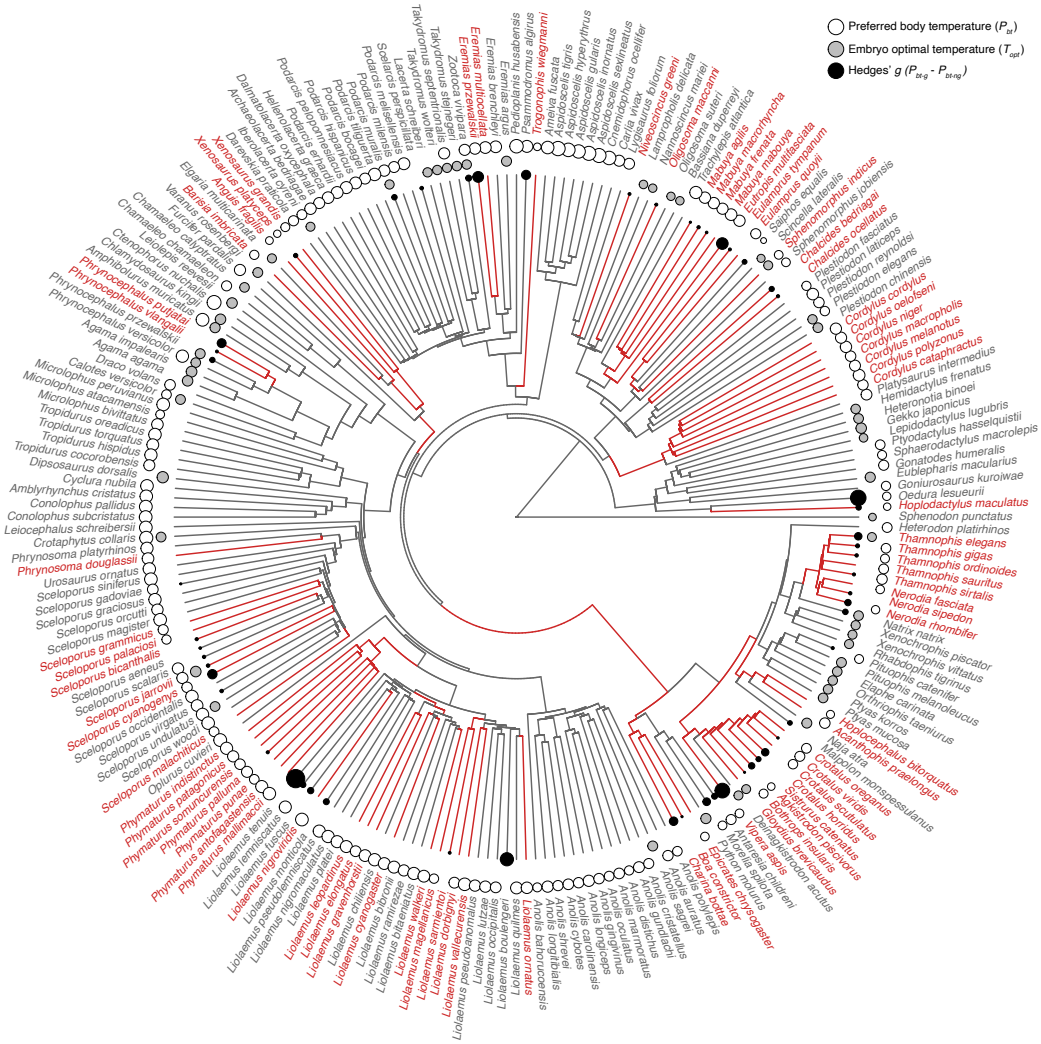
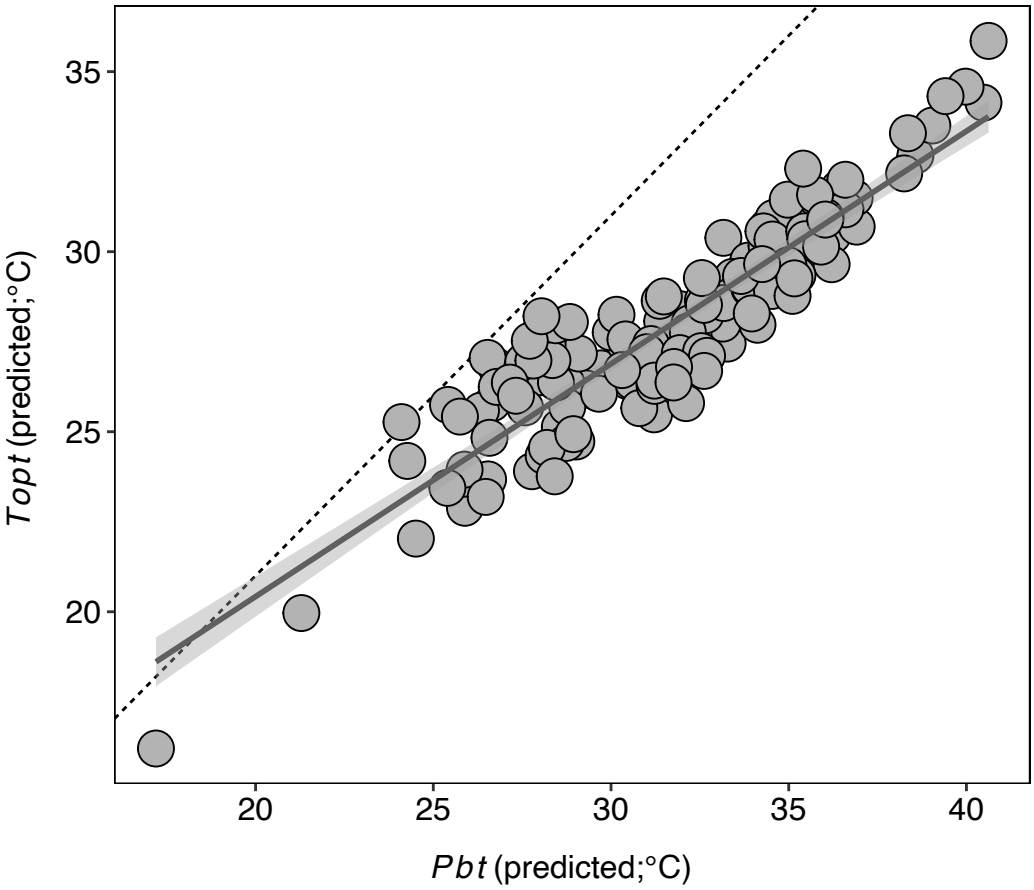
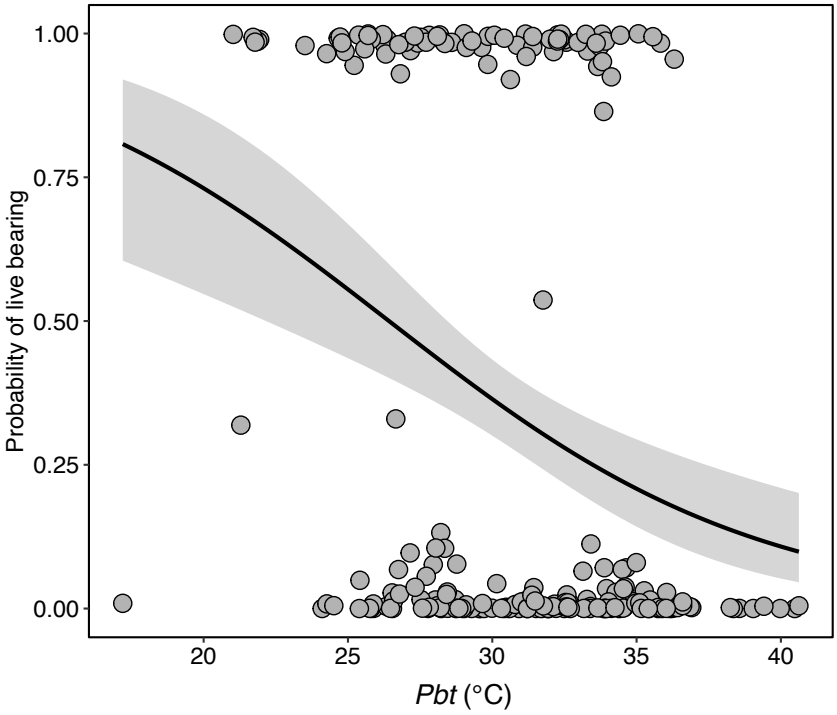


Figure 1. Variation in reproductive mode, preferred body temperature and plasticity in female thermo-regulatory behaviour across 224 species of squamate reptile (lizards and snakes). Tips and branches are coloured according to reproductive modes (red = live bearing/viviparous, grey = egg laying/oviparous; branch colours represent predicted ancestral values; Table S10). The size of white circles corresponds to the preferred body temperature of female squamates (P_{bt} dataset, 163 species), the size of grey circles to the optimal temperature for embryo survival in oviparous species (T_{opt} dataset, 50 species), and the size of the black circles to the discrepancy between preferred body temperature in gravid versus non-gravid females (Hedges' g dataset; absolute values, 52 species).



475

476 **Figure 2. Mismatches in the thermal optima of mothers and embryos across 152**
477 **oviparous species of lizards and snakes.** The relationship between female preferred body
478 temperature (P_{bt}) and the temperature that maximizes hatching success of embryos (T_{opt})
479 predicted by a multi-response Bayesian Phylogenetic Mixed Model across egg laying species
480 (regression lines ± 95% confidence intervals are plotted). Deviations from the dotted line (1:1
481 relationship) indicate the magnitude of the mismatch between adult and embryo thermal
482 optima. Corresponding measurements of T_{opt} and P_{bt} were not available for all species (n_{Pbt} =
483 103; n_{Topt} = 48; n_{Pbt} & n_{Topt} = 12), hence model predictions rather than raw data are shown (for
484 plot of raw data as well as correspondence between predicted and raw data values see Figure
485 S1).



486

487 **Figure 3. Probability of being a live-bearing species in relation to female preferred body**

488 **temperature.** The line and shading represent a logistic regression curve ± 95% confidence

489 interval.

490

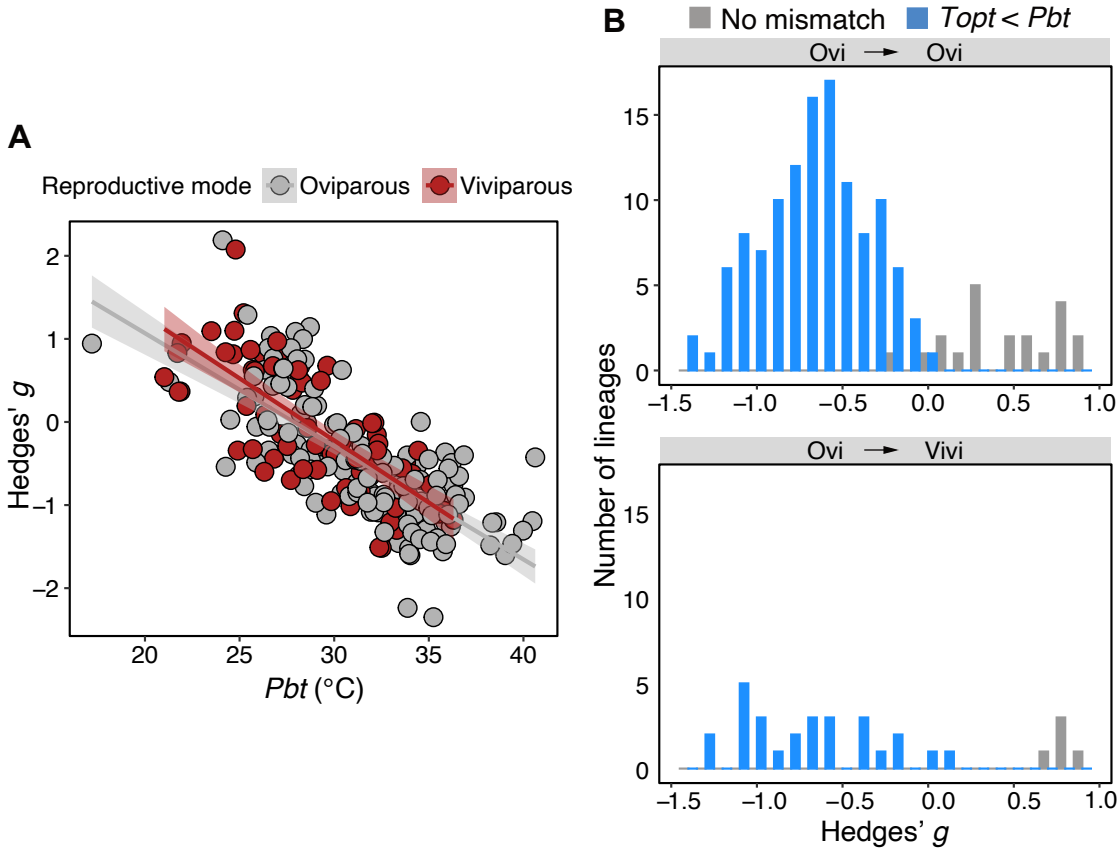


Figure 4. Female behavioural plasticity resolves the constraints imposed by thermal physiology, such that transitions to live-bearing (viviparity) occur despite mismatches between females and embryos. (A) Adjustment of body temperature in gravid females in relation to their non-gravid preferred body temperature (P_{bt}) across 52 extant oviparous and viviparous lizards and snakes (Hedges' g; $P_{bt\text{-}g}$ - $P_{bt\text{-}ng}$). Regression lines ± 95% credible intervals are plotted. High values of Hedges' g signify an increase in body temperature when gravid while low values imply a decrease. (B) The adjustment of body temperature by females in lineages that remained egg laying (top panel) and lineages that evolved live bearing (bottom panel) according to the presence or absence of a thermal mismatch between females and their offspring (inferred from Hedges' g data shown in panel A). Abbreviations: Ovi, oviparous (egg laying); Vivi, viviparous (live bearing).