ORIGINAL PAPER

Impact of developmental temperatures on thermal plasticity and repeatability of metabolic rate

Fonti Kar[1] · Shinichi Nakagawa[1] · Daniel W. A. Noble[2]

Abstract

Phenotypic plasticity is an important mechanism that allows populations to adjust to changing environments. Early life experiences can have lasting impacts on how individuals respond to environmental variation later in life (i.e., individual reaction norms), altering the capacity for populations to respond to selection. Here, we incubated lizard embryos (*Lampropholis delicata*) at two fluctuating developmental temperatures (cold $= 23$ ºC $+/- 3$ ºC, hot $= 29$ ºC $+/- 3$ ºC, $n_{cold} = 26$, $n_{hot} = 25$) to understand how it affected metabolic plasticity to temperature later in life. We repeatedly measured individual reaction norms across six temperatures 10 times over ~ 3.5 months ($n_{obs} = 3,818$) to estimate the repeatability of average metabolic rate (intercept) and thermal plasticity (slope). The intercept and the slope of the population-level reaction norm was not affected by developmental temperature. Repeatability of average metabolic rate was, on average, 10% lower in hot incubated lizards but stable across all temperatures. The slope of the thermal reaction norm was overall moderately repeatable ($R = 0.44$, 95% CI $= 0.035 – 0.93$) suggesting that individual metabolic rate changed consistently with short-term changes in temperature, although credible intervals were quite broad. Importantly, reaction norm repeatability did not depend on early developmental temperature. Identifying factors affecting among-individual variation in thermal plasticity will be increasingly more important for terrestrial ectotherms living in changing climate. Our work implies that thermal metabolic plasticity is robust to early developmental temperatures and has the capacity to evolve, despite there being less consistent variation in metabolic rate under hot environments.

Keywords Reaction norm · Repeatability · Metabolic rate · Incubation temperature · Thermal performance curve · Thermal sensitivity · Phenotypic flexibility

✉ Fonti Kar
  fonti.kar@gmail.com

[1] School of Biological Earth and Environmental Sciences, Ecology and Evolution Research Centre, University of New South Wales, Sydney, NSW, Australia

[2] Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, ACT, Australia

Introduction

A substantial amount of variation in an individual's phenotype is determined by formative processes experienced throughout embryonic development. Environmental perturbations during this critical period can have persistent effects on an individual's physiology, morphology, behaviour and life history (Noble et al. 2021; Eyck et al. 2019; O'Dea et al. 2019). Developmental shifts in phenotypes may be adaptive if it allows organisms to better cope in similar environments later in life (Beldade et al. 2011). However, environment-phenotype mismatches can occur when developmental cues fail to predict late life conditions (Auld et al. 2010; Bonamour et al. 2019). A multitude of traits throughout an animal's life are labile; such as metabolism (Bronikowski and Vleck 2010); behavioural traits (Mitchell and Biro 2017); growth (Rodgers et al. 2018) and hormone levels (Lendvai et al. 2014), that can reversibly respond to environmental change. Reversible plasticity in phenotypic traits allows individuals to adjust to immediate changes in their surroundings (Piersma and Drent 2003), and can broadly be classified into two categories, acclimation and phenotypic flexibility (Piersma and Drent 2003; Havird et al. 2020). Acclimation is generally a slower form of reversible plasticity that involves remodelling of physiological systems from chronic exposure to a particular environment (Seebacher 2005). Phenotypic flexibility, in contrast, describes changes in traits that are induced by short-term environmental exposure, such as changes in metabolic rate in response to ambient temperature (Piersma and Lindström 1997; Piersma and Drent 2003).

Reversible plasticity may be able to alleviate the costs associated with phenotype mismatches induced by early life environments (Angilletta et al. 2003; Ghalambor et al. 2007). When environments shift predictably, flexibility in the phenotype would be advantageous because individuals can compensate for the effects of prevailing conditions to avoid discrepancies between the environment and the phenotype (Botero et al. 2015). However, reversible plasticity in a labile trait later in life can change depending on early environmental conditions – altering how organisms respond to environmental variation (Beaman et al. 2016) (Fig. 1). The interaction between early- and late life
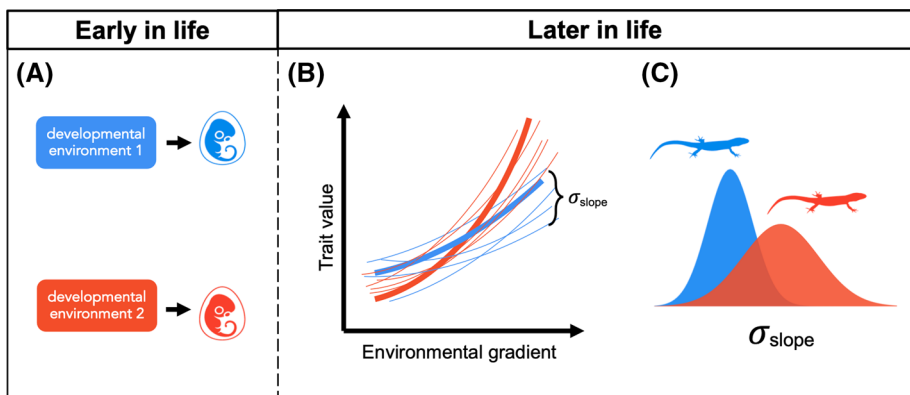

Fig. 1 Conceptual diagram illustrating the effects of (**A**) developmental environments experienced early in life on (**B**) reversible plasticity of labile traits expressed later in life ('reaction norm', Via et al. 1995). Developmental environment can also influence (**C**) repeatability of plastic responses (consistent among individual variation in slopes), this is typically represented by the variation in the slope of the reaction norm, Via et al. 1995)"

plasticity has been supported by a few studies that show developmental differences in plasticity for a variety of traits including mitochondrial function (Shama et al. 2014), metabolic rate (Seebacher et al. 2014) and locomotor performance (Kazerouni et al. 2016). However, these studies solely focus on how developmental thermal effects affect acclimation, whereas the influence on phenotypic flexibility and variability of plastic responses is poorly known.

It has long been recognised that individuals vary in their plasticity, with some responding more flexibly than others (Nussey et al. 2007; Dingemanse and Wolf 2013). Consistent among individual variation in plasticity (also known as repeatability in plasticity) may be heritable and provides the phenotypic substrate for selective forces to act upon (Nussey et al. 2007; Araya-Ajoy and Dingemanse 2017). Developmental environments can influence phenotypic variation available for selection (Sultan and Stearns 2005). For example, zebra finches (*Taeniopygia guttata*) that experience nutritional stress as nestlings weigh less and have reduced growth rates contributing to increases in the repeatability of metabolism and behavioural traits (Careau et al. 2014a). Repeatability of plasticity has also been reported in other labile traits including aggressiveness in great tits (*Parsus major*) (Araya-Ajoy and Dingemanse 2017), explorative behaviour in chickadees (Thompson et al. 2018) and metabolic rate in amphipods (Réveillon et al. 2019). Whether developmental environments affect the repeatability of plasticity per se is still not well understood (Fig. 1). Identifying the factors that impact repeatability is necessary for understanding the evolution of plasticity in changing environments.

Energy metabolism is a key fitness related trait that is both consistently different among individuals and highly labile within individuals (Nespolo and Franco 2007; Norin and Metcalfe 2019). All organisms require energy for growth, maintenance and reproduction (Careau et al. 2014c). Numerous studies have investigated the influence of various developmental environments, such as temperature (Gangloff et al. 2015; Noble et al. 2021), ultra-violet (UV) exposure (Kazerouni et al. 2016), and dietary restriction (Careau et al. 2014a) on metabolic rate, however, the impacts on plasticity of metabolic rate is not well established (but see Seebacher et al. 2014). Developmental environments are expected to influence metabolic plasticity, possibly through modifications in metabolic enzymes or cellular membrane structure that influence their function in different environments (Angilletta 2016). Such changes imply that tolerance to environmental perturbations may be determined by the developmental environment a given cohort experiences. Furthermore, if repeatability of metabolic plasticity is also affected, then a population's capacity to respond to selection might also depend on early life conditions. Understanding how early life environments shape metabolic plasticity will be particularly important for organisms where metabolic rate is closely intertwined with prevailing environmental conditions.

Here we employed a 'reaction norm approach'(sensu Via et al. 1995) to examine the impact of developmental temperature on metabolic rate plasticity in an oviparous skink (*Lampropholis delicata*). Specifically, we were interested in testing whether developmental temperature affects the shape and repeatability of metabolic thermal reaction norms. Over 3.5 months, we repeatedly measured routine metabolic rate at six temperatures for lizards ($n_{obs} = 3{,}818$) that hatched from two incubation treatments (cold = 23 ºC ± 3 ºC, hot = 29 ºC ± 3 ºC, $n_{cold} = 26$, $n_{hot} = 25$) to address the following key questions: (1) How does developmental temperature change the intercept and slope of the thermal reaction norm?; (2) Does the repeatability of metabolic plasticity (i.e. slope of the reaction norm) differ between the two developmental temperatures? (3) Do developmental temperature treatments differ in their repeatability of metabolic rate (intercept) at each temperature (i.e. temperature-specific repeatability)? Our experimental approach provides important insights

into how developmental environments mediate the capacity for ectotherms to metabolically respond to thermal variation.

Materials and methods

Lizard collection and husbandry

We established a breeding colony of adult *L. delicata* ($n_{females} = 144$, $n_{males} = 50$) using wild individuals collected across three sites throughout the Sydney region between 28 August and 8 September 2015 (UNSW Kensington Campus: -33.92, 151.24; Sydney Park: -33.91, 151.18, Macquarie Park: -33.77, 151.10). These three sample locations represent the same genetic lineage of *Lampropholis delicata*, it is therefore unlikely that our experiments would be impacted by collection at different sampling sites (Chapple et al. 2011). Furthermore, all adults were kept in common lab conditions for a year prior to the commencement of this study, therefore any environmental effects are likely to be minimal. Post-hoc analyses show that the population origin of dams and sires did not explain a significant amount of variation in metabolic rate (see below and Table S2). Three females were housed with a single male in opaque plastic enclosures measuring $35 \times 25 \times 15$ cm (L×W×H). Enclosures were kept under UV lights on a 12 h light: 12 h dark cycle in a temperature-controlled room set to 24 ℃. Lizards had access to a heat lamp (IKEA Kvart fitted with a E14 2700 k reflector bulb 36.6 cm x 10 cm) that elevated temperatures on one side of the enclosure to 32 ℃. Each enclosure was lined with newspaper and lizards had constant access to water and tree bark was used as refuge. Adult lizards were fed medium sized crickets (*Acheta domestica*) ad libitum dusted with calcium powder and multi-vitamin every two days. From the beginning of the egg laying season (October of each year), we replaced the newspaper lining with garden potting mix and placed an opaque plastic box ($12 \times 17.5 \times 4.3$ cm) containing moistened vermiculite in each enclosure for females to oviposit their eggs. During this time, enclosures and vermiculite boxes were sprayed gently with water every other day to maintain a relatively humid environment. From October to November, vermiculite boxes were checked every day for eggs. Animal collection was approved by the New South Wales National Parks and Wildlife Service (SL101549). All procedures adhere to the 8th edition of the Australian Code for the Care and Use of Animals for Scientific Purposes and were approved by the Macquarie University Ethics committee (ARA 2015/015) and University of New South Wales Animal Care and Ethics committee (ACEC 15/51A).

Developmental temperature manipulations

Eggs were collected between October 2017–March 2018. When eggs were discovered, they were weighed using a digital scale to the nearest 0.01 g (Ohaus Scout SKX123). We also measured egg length (distance between the furthest points along the longest axis of the egg) and egg width (distance between the widest points along the axis perpendicular to the longest axis of the egg) using digital callipers to the nearest 0.01 mm in case any egg features differed between treatments. This is important because egg size can impact suites of offspring traits (Lancaster et al. 2010; Krist 2010). Preliminary analyses showed that none of the egg dimensions differed between treatments and egg dimensions were all positively correlated (Supplementary Materials Sect. 1, Table S1). Following measurements, each egg was placed in a plastic cup (80 ml) containing 3 g of vermiculite and 4 g

of water and covered using cling wrap which was secured by an elastic band. We used a split-clutch design such that eggs from each clutch were pseudo-randomly assigned to one of two fluctuating incubation temperature treatments. We used two incubators to precisely control the temperature of eggs (LabWit, ZXSD-R1090). The 'hot' treatment was exposed to a mean temperature of 29 ℃ whereas the 'cold' treatment was exposed to a mean temperature of 23 ℃. Both incubators fluctuated $+/-3$ ℃ the mean temperature over a 24-h period. These treatments represent the average thermal minimum and maximum of natural nest sites in Sydney populations of *L. delicata* (Cheetham et al. 2011). Egg cups were rotated within each incubator weekly to avoid uneven heat circulation within incubators. Incubators were also checked daily for hatchlings. On average, the incubation duration for the 'hot' treatment was 30 days (SD = 1.40, range = 27–33) days and 47.7 days (SD = 5.90, range = 25–53) for the 'cold' treatment.

Planned missing data and metabolic rate at different temperatures

Metabolic measurements commenced in April 2018 and continued until August 2018. At the beginning of measurements, hatchlings were on average 88.68 days old (SD = 23.75, range = 26–131). Due to the scale of our experiment, we opted for closed-system respirometry instead of flow-through respirometry. The former approach allowed us to obtain high-throughput measurements of MR across many individuals, whereas the latter approach provides a more in-depth characterisation of metabolic changes over time for a much smaller subset of individuals.

We quantified routine metabolic rate (hereafter referred to as metabolic rate [MR]) as our measurements likely included the energetic costs of random movements (Withers 1992; Mathot and Dingemanse 2015). MR was measured as the volume of CO_2 production per unit time ($\dot{V}_{CO_2}$ mL min^{-1}) as CO_2 production is less susceptible to fluctuations in water vapour and signals are more easily detected in smaller organisms (Lighton 2008; Tomlinson et al. 2018). CO_2 production was strongly correlated with O_2 consumption in any case ($r = 0.81$, $p < 2.2e^{-16}$) with RQ values averaging 0.77 (SD = 0.41, $n_{obs} = 198$).

Due to logistical constraints, lizards were randomly assigned to one of two blocks for MR measurements (block 1: n = 26, block 2: n = 25). Each block was measured two days apart. We sampled lizards once a week for two-weeks consecutively and then allowed them to rest for one week before the next set of measurements. Each measurement week was considered a sampling session (ten sampling sessions in total over the course of 14 weeks). We used the same incubators described above to precisely control the temperature at which MR measurements were taken ($+/-1$ ℃).

Metabolic rate was measured at six acute (i.e., short-term; < two hours) temperatures, 24, 26, 28, 30, 32 and 34 ℃ in a randomised order. We employed a planned missing data design to minimise the stress of lizards during the experiment as the lizards were also part of an ongoing growth rate experiment. Metabolic rate data is particularly amenable to a planned missing data design because it is strongly correlated with temperature and body mass. For each block of lizards, we intentionally missed measurements at two randomly selected temperatures at each sampling session (Nakagawa 2015; Noble and Nakagawa 2021). For more details see Supplementary Materials (Fig. S1). Missing MR data temperatures were then recovered using data imputation techniques during analysis (see Statistical analyses).

At ~ 06:00, lizards were gently encouraged into an opaque respiratory chamber and then weighed. After which, chambers were placed inside preheated incubators set at the randomised temperature for 30 min to allow body temperatures to equilibrate. The lids of the chambers were left ajar during this time to minimise CO_2 build up. After 30 min, each chamber was flushed with fresh air and sealed. A 3 mL 'control' air sample was immediately taken via a two-way valve to account for any residual CO_2 that was not flushed from the chambers. The chambers were left in the incubator at the set temperature for lizards to respire for 90 min. After this time, two replicate air samples (3 mL) were taken from each chamber in order to estimate the change in CO_2 relative to the control sample. Two samples were taken so we could explicitly estimate measurement error (see Statistical analyses, Ponzi et al. 2018). Chambers were then reopened and flushed with fresh air before being placed back into the incubator for the second measurement temperature (2 temperatures / day) following the same procedure approximately two hours later. Overall, this sampling design enabled us to characterise the thermal reaction norm (four out of six temperatures for our planned missing data design) for each lizard 10 times while accounting for measurement error. This resulted in n = 4,080 measurements of MR ([2 air samples × 4 temperatures] × 10 sampling sessions = 80 samples per lizard). However, additional missing data from equipment malfunction or human error meant that our total sample size was n = 3,818.

All air samples were injected into the inlet line of a Sables System FMS (Las Vegas NV, USA) with the flow rate set to 200 mL min^{-1} to measure $\dot{V}_{CO_2}$ and $\dot{V}_{O_2}$. Water vapour was scrubbed from the inlet air with Drierite. Output peaks were processed using the R package 'metabR' (https://github.com/daniel1noble/metabR). The rate of CO_2 produced by an individual was calculated following (Lighton 2008):

$$\dot{V}_{CO_2} \, \mathrm{mL} \, min^{-1} = \frac{\Delta\%CO_2 \times (V_{chamber} - V_{lizard})}{t}$$

where $\Delta\%CO_2$ is the maximum percentage of CO_2 in air sample above baseline, which was corrected by subtracting any 'residual' CO_2 from the initial flush from the larger of the two air samples; $V_{chamber}$ is the volume of the chamber (70 mL) and V_{lizard} is the volume of the lizard. We used the mass of the lizard as a proxy for its volume (1 g = 1 ml) because of their high correlation and increased accuracy and precision in mass measurements (Friesen et al. 2017; Kar et al. 2021), and t is the duration of time in minutes after where the chamber had been sealed and the first air sample was taken (90 min).

Statistical analyses

We fitted Bayesian linear mixed effect models in *R* (Core Team 2013) using the package 'brm''s' (Bürkner 2017). Metabolic rate was log transformed then body mass was log transformed and then z-transformed. Age and temperature were z-transformed so parameter estimates of main effects and interaction terms were more interpretable (Schielzeth 2010). Post-hoc analyses revealed that including dam ID, sire ID and population origin of either parents as random effects did not improve model performance. As such these were not included in our final analyses (Supplementary Materials Sect. 1, Table S2). Our planned missing data design resulted in random missingness across temperatures (36% missingness in MR and body mass). The 'brms' package is capable of performing model-based data imputation. As such, we performed imputation during model fitting in all of our analyses. Model-based imputation not only retains the hierarchical structure of the dataset but also increases statistical power and

precision by recovering missing data values (Nakagawa 2015; Noble and Nakagawa 2021). We found that models with imputed data resulted in similar conclusions to complete case analyses (Supplementary Materials 2–4, Table S5, S7 – S9, S11 – S14). We present results from the imputation analysis in the main text as parameter estimates were more precise (Supplementary Materials). For all models we used default priors which are presented in Table S3. We ran four Markov Chain Monte Carlo (MCMC) chains; taking 800 samples from the posterior distribution after discarding the first 1,500 iterations. This gave a total of 3,200 samples from the posterior distribution across all chains. We ensured chains were mixing by inspecting trace plots and checked that scale reduction factors were less than 1.01, which indicates that all chains had converged. Throughout we report posterior means and 95% credible intervals for all parameters. All data and code to reproduce our results are provided (see Data accessibility).

To test whether developmental temperatures changed the shape of reaction norms, we fitted a full model with MR as the response and temperature, treatment (developmental temperatures, 23 ºC as the baseline) and an interaction between treatment and temperature as predictors. The model also included a random intercept for lizard identity and sampling session. We wanted to account for measurement error in all our models as it may conflate parameter estimates (Ponzi et al., 2018). Using the two replicate air samples, we estimated measurement error variance by including a nested random effect of lizard identity, sampling session and temperature in all our models (e.g. ID001_session1_temp24). This nested random effect (hereafter referred to as measurement error) estimates the variance attributed to differences among replicates. While we show in a previous study that measurement error can vary by temperature (Kar et al. 2021), here we assumed that measurement error was constant across temperatures by fitting it as a random intercept as estimating a random slope resulted in model convergence issues. Heterogeneous residual variance across temperatures can also influence parameter estimates (Careau et al. 2014a). However, Watanabe-Akaike information criterion (WAIC) and expected log predictive density (ELPD) values indicated that a heterogeneous residual variance model was not well supported, therefore homogenous variance was used in all models (Table S4). Acclimation can influence metabolic plasticity and its effects can take place throughout the course of our study. Unfortunately, it was not possible to measure MR at hatching. However, we still tested whether there were treatment differences in thermal reaction norms in the first sampling session (~2.5 months of age) where acclimation effects were likely to have the weakest effect.

We estimated adjusted repeatability of the reaction norm slope (R_{slope}) in each developmental temperature treatment by fitting separate models for each treatment group. MR was fitted as the response and temperature, body mass and age (days since hatching) as predictors. We included lizard identity, measurement error and a nested random effect of individual identity and sampling session (hereafter referred to as series, Araya-Ajoy et al. 2015). Lizard identity estimates among individual variance, whereas series partitions variance within individuals across all sampling sessions which allows the estimation of slope repeatability. A random temperature slope was estimated for lizard identity and series. The repeatability of the slope is calculated as the proportion of total variance in slopes explained by among individual differences (Araya-Ajoy et al. 2015):

$$R_{slope} = \frac{V_{I,slope}}{\left(V_{I,slope} + V_{series,slope}\right)}$$

where $V_{I,slope}$ is the among-individual variance in the temperature slope term and the $V_{series,slope}$ is the within-individual variance in the temperature slope.

We also estimated adjusted repeatability of average metabolic rate (i.e. intercept of the reaction norm) at each temperature by fitting separate models for each treatment group. Similar to above, MR was included as the response and temperature, body mass and age as predictors. We included lizard identity, sampling session and measurement error as random intercepts and temperature as a random slope for lizard identity. We calculated among individual variance in metabolic rate at each temperature I_t following Schielzeth and Nakagawa (2020):

$$I_t = V_I + \left(t^2.V_S\right) + \left(2t.Cov_{I,S}\right)$$

where V_I is the among individual variance in intercepts, t is the specific temperature at which repeatability is calculated for, V_S is the among individual variance in slope and $COV_{I,S}$ is the covariance between the intercept and slope at the among individual level. Temperature specific repeatability (R_t) is then calculated as follows:

$$R_t = \frac{I_t}{\left(I_t + V_{session} + V_e\right)}$$

where $V_{session}$ is the variance due to sampling session and V_e is residual variance.

We also wanted to estimate overall repeatability of average metabolic rate across all temperatures. We therefore fitted the same model as above for each treatment, but we omitted the random temperature slope for lizard identity, this estimates an average among individual variance across all temperatures. Similarly, we calculated repeatability as per the equation above but using just the single estimate of among individual variance.

In order to test for differences in repeatability among the two developmental temperatures, we calculated contrasts by subtracting the posterior distributions of repeatability estimates of the cold treatment from the hot treatment (Hot – Cold). To test whether the magnitude of differences among treatments were significant, we calculated probabilities of direction (pd) using the package 'bayestestR'(Makowski et al. 2019b). The probability of direction is calculated relative to the posterior median and ranges from 50 to 100%. The value of pd describes whether an effect is either positive or negative as it is always relative to the sign of the median (Makowski et al. 2019a). If the median is positive, then pd describes the proportion of the posterior distribution that is also positive (Makowski et al. 2019a). A pd value of 95% can be interpreted as the effect is positive with a probability of 95%.

Results

We found no evidence to suggest that metabolic rate or its response to short-term temperature changes was influenced by early developmental temperature (Fig. 2, Table 1, Supplementary Materials Sect. 2 Table S5-7). Congruently, there were no treatment differences in thermal reaction norms at the first sampling session when acclimation effects are likely to have the least effect (Supplementary Materials, Sect. 2). We therefore refitted the model with just the main effects (Supplementary Materials, Sect. 2, Table S6–S7). Across all models, temperature and body mass had positive effects on metabolic rate (Table 1, Supplementary Materials, Sect. 2). Reaction norm slopes were significantly repeatable in both treatments, albeit estimated with a large degree of error. Repeatability of slopes (R_{slope}) did not depend on developmental temperature treatments (Hot: $R_{slope} = 0.42$, 95%

Fig. 2 Predicted thermal reaction norm of metabolic rate (VCO_2 min^{-1} g^{-1}) for the 'cold' developmental temperature group (blue line, $n_{lizards} = 26$) and the 'hot' developmental temperature group (red line, $n_{lizards} = 25$) Points are raw data and are coloured according to treatment groups, $n_{obs} = 3818$. Dashed lines represent the upper and lower bounds of 95% credible intervals
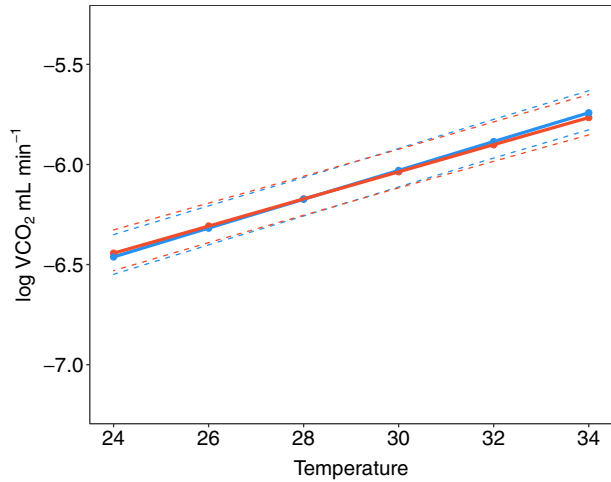
Table 1 Model coefficients of the full model testing whether developmental temperature (cold = 23 ℃ ± 3 ℃, hot = 29 ℃ ± 3 ℃) affects the elevation (intercept) and slope of the thermal reaction norm of metabolic rate

Parameter	Estimate	Lower	Upper
Fixed effects			
Intercept MR	**− 6.292**	**− 6.372**	**− 6.218**
Treatment (hot)	− 0.003	− 0.062	0.058
Temperature	**0.262**	**0.246**	**0.278**
Treatment × Temperature	− 0.016	− 0.039	0.007
Age	− 0.035	− 0.079	0.006
Mass	**0.128**	**0.105**	**0.151**
Random Effects			
Lizard Identity			
Intercept	**0.009**	**0.006**	**0.015**
Temperature Slope	$9.53e^{-5}$	$1.54e^{-7}$	0.000479
$COV_{Intercept - Slope}$	− 0.00018	− 0.00122	0.000599
Sampling Session			
Intercept	**0.01**	**0.003**	**0.026**
Measurement Error			
Intercept	**0.044**	**0.04**	**0.049**
Residual	**0.041**	**0.038**	**0.043**

This model used an imputed dataset of $n_{obs} = 6,000$, 36% of observations were imputed. The intercept is the cold developmental temperature. MR was log transformed and mass, age and temperature were z-transformed. Credible intervals of bolded estimates do not overlap zero. Lower and upper bound of estimates represent 95% credible intervals. COV represents covariance. Main effects model is presented in Table S4

There is not particularly definition for the "significance" for bolded estimates as we are employing Bayesian statistics. These values are "significant" because the credible intervals does not overlap zero

CI 0.04 − 0.91; Cold: $R_{slope} = 0.46$, 95% CI 0.03 − 0.95; $pd = 53.5\%$, Fig. 3, Supplementary

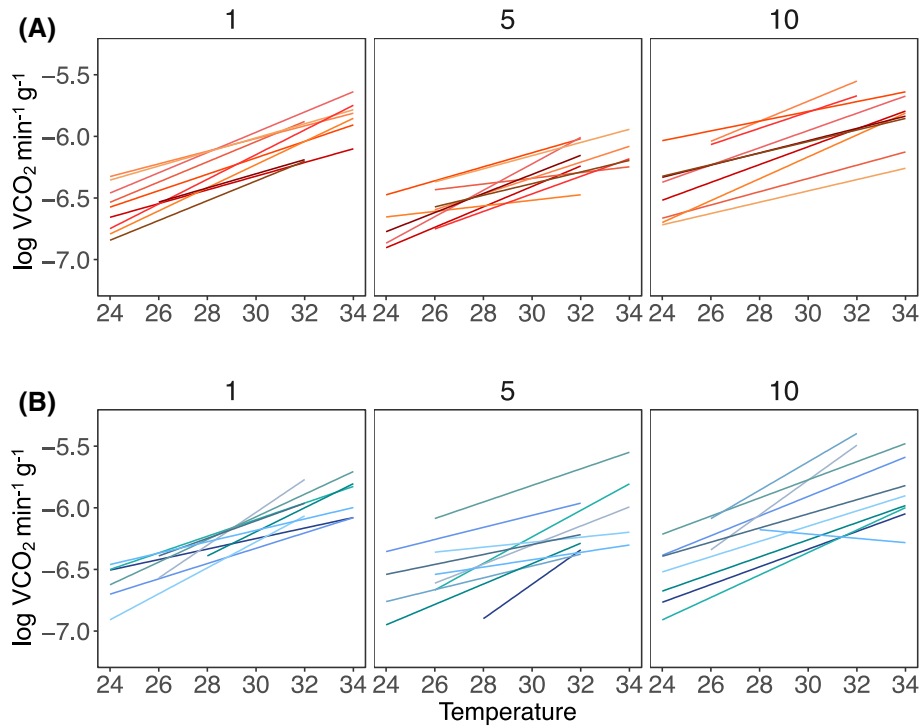**Fig. 3** Thermal reaction norms of mass-adjusted metabolic rate for lizards reared at (**A**) 'hot' developmental temperatures (top, red lines, $n_{lizards}=25$) and (**B**) 'cold' developmental temperatures (bottom, blue lines, $n_{lizards}=26$) at session number one, five and ten. Each uniquely coloured line represents an individual reaction norm. A random subset of 10 individuals from each treatment are presented

Materials, Sect. 3). A *pd* value of 53.5% indicates that there is roughly equal probability that the difference in R_{slope} is positive or negative, indicating little difference among treatment groups.

Overall, temperature-specific repeatability was relatively low, with the cold developmental treatment tending to have higher repeatability estimates compared to the hot developmental treatment (Fig. 4, Fig S2, Supplementary Materials, Sect. 4 Table S12). Irrespective of temperature, repeatability of average metabolic rate was on average 10% higher in cold incubated lizards (*pd*=95.7%, Fig. 4b, c). There was a 95.7% probability that the difference in overall repeatability was negative, indicating that lizards from the cold treatment are more likely to have higher repeatability. Higher repeatability in the cold treatment was associated with significant increases in among individual and residual variance (Fig. S3).

Discussion

Early developmental temperature did not change the intercept or slope of the population reaction norm in delicate skinks. Thermal plasticity of metabolic rate was unaffected by developmental temperature, however; variation in slope had relatively high repeatability ($R>0.4$). However, temperature-specific repeatability of metabolic rate (i.e., intercept)
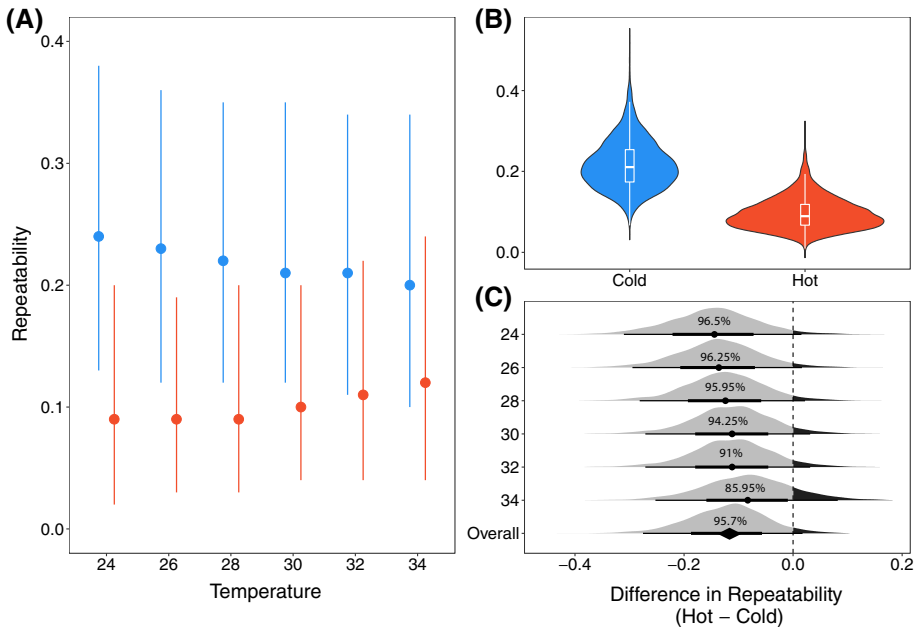
Fig. 4 (**A**) Temperature-specific adjusted repeatability for average metabolic rate for the 'cold' developmental temperature group (blue, $n_{lizards} = 26$) and the 'hot' developmental temperature group (red, $n_{lizards} = 25$). Error bars represent 95% credible intervals. (**B**) Violin and boxplot showing the posterior distribution of overall adjusted repeatability of each treatment group irrespective of temperature. (**C**) Posterior distribution of the difference in repeatability (Hot–Cold) overall and at each temperature. Point represents the median; thicker lines represent the interquartile range and thin lines represent the 95% credible intervals. The probability of direction is presented on each distribution and describes the probability that the difference in repeatability is either positive or negative. Grey regions of the distribution represent negative estimates indicating repeatability was greater in the cold treatment, whereas black regions represent positive estimates which indicates that repeatability was greater in the hot treatment. All values were calculated from imputation models (Supplementary Materials, Sect. 4 Table S15–16). Contrasts are presented in Table S12

was lower among lizards that were reared in hot developmental temperatures. Our results suggest that, while individuals displayed consistent variation in their plasticity, how metabolic rate responds to short-term temperature variation later in life was robust to simulated thermal minimum and maximum of natural nest sites. Developmental temperatures did not have an impact on average metabolic rate but rather it changed the amount of consistent individual variation in average metabolic rate. Below we discuss the implications of our results for the evolution of thermal reaction norms.

Thermal reaction norms of metabolic rate are robust to developmental temperature

Developmental environments that affect later life plasticity may affect how individuals and populations respond to environmental fluctuations (Beaman et al. 2016). Epigenetic modifications during development that influence the physiological system are likely responsible for shaping plastic responses in complex ways (Hu and Barrett 2017; McCaw et al. 2020). However, our results suggest instead that thermal reaction norms for metabolic rate were robust to changes to the incubation temperatures we selected which were based on the thermal minimal and maximal of natural nests found in Sydney (Range = 22.5 ℃ –34.5 ℃,

mean = 27.4 ºC, SD = 2.66, N = 59, Cheetham et al. 2011). It should also be noted that our treatments were fluctuating, and at times, temperatures overlapped which may contribute to the lack of differences observed.

Among the few studies that have investigated the effects of pre- and post-hatching temperature on the plasticity of metabolic rate, results have been mixed (Table 1, Beaman et al., 2016). For example, wild caught mosquitofish (*Gambusia holbrooki*) developing under more variable spring conditions exhibited steeper thermal reaction norms for metabolic scope compared to fish born in summer (Seebacher et al. 2014). In contrast, incubation temperature did not affect plasticity in metabolic rate of striped marsh frog tadpoles (Seebacher and Grigaltchik 2014). Given that our lizards were reared in a common environment the lack of difference we observed may be due to the waning effects of developmental environments on metabolism as individuals age and acclimate to post-hatching environments. Generally, acclimation of physiological function takes approximately three to four weeks to complete (Seebacher 2005). Given that we began the study when lizards were ~2.5 months old, it is possible that metabolic acclimation was an important factor as young lizards were able to behaviourally thermoregulate in their enclosures and may have aligned to their preferred temperatures (~26 ºC in adults, Goulet et al. 2016). Future studies should employ cross factorial designs where post-hatch environments are deliberately matched and mismatched with early environmental conditions to disassociate acclimation and thermal preference effects (Schnurr et al. 2014; Kazerouni et al. 2016).

Stable thermal reaction norms of metabolic rate across both developmental temperatures have key evolutionary implications. Our results imply that population reaction norms may be robust to temperature variation within the thermal range of natural nests (Cheetham et al., 2011). Past thermal regimes encountered by ancestors may have canalized population responses so that they are less sensitive to fluctuations in developmental temperature (Liefting et al. 2009). Canalization may reduce the costs of phenotypic plasticity during development if environmental variation is predictable across generations (Aubret and Shine 2010). In support of this, damselflies undergoing range expansions exhibit geographic variation in thermal reaction norms that align with past climatic conditions (Lancaster et al. 2015). Population comparisons across environmental gradients might reveal whether local adaptation shapes developmental plasticity of population reaction norms that lead to canalisation (Toftegaard et al. 2015). Developmental environments may play a stronger role in shaping population plastic responses in areas that experience greater thermal variability, such as those in temperate or high elevation regions (Rutschmann et al. 2016; While et al. 2018; Bonamour et al. 2019). While our incubation treatments represent the average thermal minimal and maximal of natural nest sites, their thermal variability (+/−3 °C) was the same. The lack of thermal variation in our treatments may be a possible explanation why the thermal reaction norms did not differ. Future studies should attempt to isolate the effects of thermal averages and variability during development to understand how thermal reaction norms are shaped in ectotherms.

Developmental temperatures and repeatable thermal reaction norms of metabolic rate

Repeatability of reaction norm slopes did not change with developmental temperature, but lizards reared in hot temperatures had reduced repeatability in metabolic rate (intercept). Variation in developmental time has important consequences on hatching condition and may contribute to differences in consistent variation in hatchling phenotypes such as MR.

Developmental time exhibits a negative nonlinear relationship with temperature, such that development times are considerably shorter at hotter temperatures (Noble et al. 2021; Marshall et al. 2020). Consequently, eggs reared in warmer environments are expected to be more constrained in their developmental rates, thus hatching phenotypes are more likely to be less variable compared to eggs reared in cooler environments (Pettersen et al. 2019). Indeed, incubation duration was short and less variable in our hot developmental treatment (Hot: 30 days, SD = 1.40, range = 27–33; Cold: 47.7 days, SD = 5.90, range = 25–53). Shortened development times may restrict embryo yolk assimilation that is needed for growth and may have contributed to reduced among-individual variation in MR (Oufiero and Angilletta 2006; Storm and Angilletta 2007). Elevated mitochondrial proton leak at hot developmental temperatures may also lead to less efficient energy production and may explain why metabolic rate did not differ among treatments even though we detected differences in repeatability (Chamberlin 2004). While is not possible to rule out any collection site differences that may contribute to epigenetic differences among our two treatment groups these effects are likely to be minimal because (1) all sites represent the same genetic lineage (Chapple et al. 2011); (2) females were reared in the lab for a year prior to experiments, removing environmental effects on female reproduction and (3) we used a split-clutch design to ensure that all sites were represented in our treatments. Despite there being no developmental differences in the repeatability of metabolic plasticity, we show that it is possible for it to evolve under changing environments (Ghalambor et al., 2007).

We found that individuals consistently vary in metabolic plasticity in response to temperature. While several studies have reported significant among individual variation in thermal plasticity slopes (Careau et al. 2014b; Briga and Verhulst 2017), its repeatability is rarely estimated as it requires a study design that allows partitioning of within and between individual variance of slopes (Araya-Ajoy et al., 2015). Our repeatability estimates for reaction norm slopes were consistent with a previous study of the same species ($R = 0.23$, Kar et al. 2021). Similarly, moderate repeatability of thermal sensitivity of metabolic rate has also been observed in amphipods ($R = 0.38$, Réveillon et al. 2019). Assuming that repeatable reaction norm slopes have a heritable basis (Driessen et al. 2007), our work implies that thermal plasticity has the potential to be selected upon and evolve (Falconer 1952; but see Dohm 2002).

Consistent individual differences in metabolic rate at a given temperature were also present and stable across all temperatures. This result demonstrates that MR is repeatable within the operable range of temperatures in *L. delicata* (Matthews et al. 2016). Overall, our estimates for the repeatability of MR ranged from 0.09–0.22. Our results are in line with a meta-analysis showing that repeatability decreases with time between repeated measurements (White et al. 2013). Specifically, the average repeatability of MR in ectotherms from studies that had a measurement interval that was equal or larger than our study (≥ 8.5 days) was $R = 0.33$ (SD = 0.21, n = 18). Interestingly, repeatability of average MR in wild caught adult *L. delicata* ($R = 0.3 - 0.5$, Kar et al. 2021) was quite comparable or even larger. This is likely due to life stage differences that shape phenotypic among-individual variation. As individuals mature, their experiences in different microhabitats (diet, thermal preferences) can promote among-individual variation in traits (Kruuk and Hadfield 2007). Such common (micro) environment effects could further increase repeatability and may explain differences between lab and wild studies (Auer et al. 2016). Furthermore, the repeatability in metabolic rate we observe may explain consistent variation in thermoregulation, behaviour and life history observed in this species (Goulet et al. 2017).

Conclusion

The role of developmental temperature on phenotypic plasticity exhibited later in life is complex. We show that at the population level, thermal plasticity of metabolic rate was robust to changes in temperature during embryonic development suggesting that thermal reaction norms may be canalised. In contrast, the impact of developmental temperature manifested as a change in the repeatability of temperature-specific metabolic rate. This has important evolutionary implications. Reduced among individual variation in hot temperatures may alter a population's ability to respond to selection under warming climate. However, population thermal reaction norms could still respond to selective processes to some extent (assuming they are heritable), allowing populations to persist. Elucidating the role of developmental environments in shaping plastic responses may require more stressful incubation conditions and cross-factorial experimental designs to disassociate the effects of acclimation from developmental plasticity.

Data availability Datasets and code used to generate results of this study is accessible via Open Science Framework (https://bit.ly/38IzTsp).

Declarations

Conflict of interest The authors declare no conflicts of interest.

References

Angilletta MJ Jr (2016) Thermal adaptation: a theoretical and empirical synthesis. Oxford University Press, Oxford

Angilletta MJ Jr, Wilson RS, Navas CA, James RS (2003) Tradeoffs and the evolution of thermal reaction norms. Trends Ecol Evol 18:234–240. https://doi.org/10.1016/S0169-5347(03)00087-9

Araya-Ajoy YG, Dingemanse NJ (2017) Repeatability, heritability, and age-dependence of seasonal plasticity in aggressiveness in a wild passerine bird. J Anim Ecol 86:227–238. https://doi.org/10.1111/1365-2656.12621

Araya-Ajoy YG, Mathot KJ, Dingemanse NJ (2015) An approach to estimate short-term, long-term and reaction norm repeatability. J Anim Ecol 6:1462–1473. https://doi.org/10.1111/2041-210X.12430

Aubret F, Shine R (2010) Fitness costs may explain the post-colonisation erosion of phenotypic plasticity. J Exp Biol 213:735–739. https://doi.org/10.1242/jeb.040576

Auer SK, Bassar RD, Salin K, Metcalfe NB (2016) Repeatability of metabolic rate is lower for animals living under field versus laboratory conditions. J Exp Biol 219:631–634. https://doi.org/10.1242/jeb.133678

Auld JR, Agrawal AA, Relyea RA (2010) Re-evaluating the costs and limits of adaptive phenotypic plasticity. Proc R Soc B Biol Sci 277:503–511. https://doi.org/10.1098/rspb.2009.1355

Beaman JE, White CR, Seebacher F (2016) Evolution of plasticity: mechanistic link between development and reversible acclimation. Trends Ecol Evol 31:237–249. https://doi.org/10.1016/j.tree.2016.01.004

Beldade P, Mateus ARA, Keller RA (2011) Evolution and molecular mechanisms of adaptive developmental plasticity. Mol Ecol 20:1347–1363. https://doi.org/10.1111/j.1365-294X.2011.05016.x

Bonamour S, Chevin L-M, Charmantier A, Teplitsky C (2019) Phenotypic plasticity in response to climate change: the importance of cue variation. Philos Trans R Soc B Biol Sci 374:20180178–20180212. https://doi.org/10.1098/rstb.2018.0178

Botero CA, Weissing FJ, Wright J, Rubenstein DR (2015) Evolutionary tipping points in the capacity to adapt to environmental change. Proc Natl Acad Sci 112:184–189. https://doi.org/10.1073/pnas.1408589111

Briga M, Verhulst S (2017) Individual variation in metabolic reaction norms over ambient temperature causes low correlation between basal and standard metabolic rate. J Exp Biol 220:3280–3289. https://doi.org/10.1242/jeb.160069

Bronikowski AM, Vleck D (2010) Metabolism, body size and life span: a case study in evolutionarily divergent populations of the garter snake (*Thamnophis elegans*). Integr Comp Biol 50:880–887. https://doi.org/10.1093/icb/icq132

Bürkner PC (2017) brms: an R package for Bayesian multilevel models using Stan J Stat Softw 80:1–28 https://doi.org/10.18637/jss.v080.i01

Careau V, Buttemer WA, Buchanan KL (2014a) Developmental stress can uncouple relationships between physiology and behaviour. Biol Lett 10:20140834. https://doi.org/10.1098/rsbl.2014.0834

Careau V, Gifford ME, Biro PA (2014b) Individual (co)variation in thermal reaction norms of standard and maximal metabolic rates in wild-caught slimy salamanders. Funct Ecol 28:1175–1186. https://doi.org/10.1111/1365-2435.12259

Careau V, Killen SS, Metcalfe NB (2014c) Adding Fuel To The "Fire Of Life": energy budgets across levels of variation in ectotherms and endotherms In: Integrative Organismal Biology, 1st edn. Wiley-Blackwell, p 17

Chamberlin ME (2004) Top-down control analysis of the effect of temperature on ectotherm oxidative phosphorylation. Am J Physiol Regul Integr Comp Physiol 287:R794–R800. https://doi.org/10.1152/ajpregu.00240.2004

Chapple DG, Hoskin CJ, Chapple SN, Thompson MB (2011) Phylogeographic divergence in the widespread delicate skink (*Lampropholis delicata*) corresponds to dry habitat barriers in eastern Australia. BMC Evol Biol 11:191. https://doi.org/10.1186/1471-2148-11-191

Cheetham E, Doody JS, Stewart B, Harlow P (2011) Embryonic mortality as a cost of communal nesting in the delicate skink. J Zool 283:234–242. https://doi.org/10.1111/j.1469-7998.2010.00764.x

Core Team R (2013) Team (2012) R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria

Dingemanse NJ, Wolf M (2013) Between-individual differences in behavioural plasticity within populations: causes and consequences. Anim Behav 85:1031–1039. https://doi.org/10.1016/j.anbehav.2012.12.032

Driessen G, Ellers J, Van Straalen NM (2007) Variation selection and heritability of thermal reaction norms for juvenile growth in Orchesella cincta Collembola: Entomobryidae) Eur J Entomol 104:39–46 https://doi.org/10.14411/eje.2007.006

Dohm MR (2002) Repeatability estimates do not always set an upper limit to heritability. Funct Ecol 16:273–280. https://doi.org/10.1046/j.1365-2435.2002.00621.x

Eyck HJF, Buchanan KL, Crino OL, Jessop TS (2019) Effects of developmental stress on animal phenotype and performance: a quantitative review. Biol Rev 94:1143–1160. https://doi.org/10.1111/brv.12496

Falconer DS (1952) The problem of environment and selection. Am Nat 86:293–298

Friesen CR, Johansson R, Olsson M (2017) Morph-specific metabolic rate and the timing of reproductive senescence in a color polymorphic dragon. J Exp Zool Part Ecol Integr Physiol 327:433–443. https://doi.org/10.1002/jez.2118

Gangloff EJ, Vleck D, Bronikowski AM (2015) Developmental and immediate thermal environments shape energetic trade-offs, growth efficiency, and metabolic rate in divergent life-history ecotypes of the garter snake *Thamnophis Elegans*. Physiol Biochem Zool 88:550–563. https://doi.org/10.1086/682239

Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Funct Ecol 21:394–407. https://doi.org/10.1111/j.1365-2435.2007.01283.x

Goulet CT, Thompson MB, Chapple DG (2016) Repeatability and correlation of physiological traits: do ectotherms have a "thermal type"? Ecol Evol 7:710–719. https://doi.org/10.1002/ece3.2632

Goulet CT, Thompson MB, Michelangeli M et al (2017) Thermal physiology: a new dimension of the pace-of-life syndrome. J Anim Ecol 86:1269–1280. https://doi.org/10.1111/1365-2656.12718

Havird JC, Neuwald JL, Shah AA et al (2020) Distinguishing between active plasticity due to thermal acclimation and passive plasticity due to Q10 effects: why methodology matters. Funct Ecol 34:1015–1028. https://doi.org/10.1111/1365-2435.13534

Hu J, Barrett RDH (2017) Epigenetics in natural animal populations. J Evol Biol 30:1612–1632. https://doi.org/10.1111/jeb.13130

Kar F, Nakagawa S, Friesen CR, Noble DWA (2021) Individual variation in thermal plasticity and its impact on mass-scaling. Oikos 130:1131–1142. https://doi.org/10.1111/oik.08122

Kazerouni EG, Franklin CE, Seebacher F (2016) UV-B radiation interacts with temperature to determine animal performance. Funct Ecol 30:584–595. https://doi.org/10.1111/1365-2435.12520

Krist M (2010) Egg size and offspring quality: a meta-analysis in birds. Biol Rev 86:692–716. https://doi.org/10.1111/j.1469-185X.2010.00166.x

Kruuk LEB, Hadfield JD (2007) How to separate genetic and environmental causes of similarity between relatives. J Evol Biol 20:1890–1903. https://doi.org/10.1111/j.1420-9101.2007.01377.x

Lancaster LT, McAdam AG, Sinervo B (2010) Maternal adjustment of egg size organizes alternative escape behaviors, promoting adaptive phenotypic integration. Evolution 64:1607–1621. https://doi.org/10.1111/j.1558-5646.2010.00941.x

Lancaster LT, Dudaniec RY, Hansson B, Svensson EI (2015) Latitudinal shift in thermal niche breadth results from thermal release during a climate-mediated range expansion. J Biogeogr 42:1953–1963. https://doi.org/10.1111/jbi.12553

Lendvai ÁZ, Ouyang JQ, Schoenle LA et al (2014) Experimental food restriction reveals individual differences in corticosterone reaction norms with no oxidative costs. PLoS ONE 9:e110564. https://doi.org/10.1371/journal.pone.0110564

Liefting M, Hoffmann AA, Ellers J (2009) Plasticity versus environmental canalization: population differences in thermal responses along a latitudinal gradient in *Drosophila serrata*. Evolution 63:1954–1963. https://doi.org/10.1111/j.1558-5646.2009.00683.x

Lighton JRB (2008) Measuring metabolic rates. Oxford University Press, New York, USW

Makowski D, Ben-Shachar MS, Chen SHA, Lüdecke D (2019a) Indices of effect existence and significance in the Bayesian framework. Front Psychol. https://doi.org/10.3389/fpsyg.2019.02767

Makowski D, Ben-Shachar MS, Lüdecke D (2019b) bayestestR: describing effects and their uncertainty, existence and significance within the Bayesian framework J Open Source Softw 4:1541. https://doi.org/10.21105/joss.01541

Marshall DJ, Pettersen AK, Bode M, White CR (2020) Developmental cost theory predicts thermal environment and vulnerability to global warming. Nat Ecol Evol 4:406–411. https://doi.org/10.1038/s41559-020-1114-9

Mathot KJ, Dingemanse NJ (2015) Energetics and behavior: unrequited needs and new directions. Trends Ecol Evol 30:199–206. https://doi.org/10.1016/j.tree.2015.01.010

Matthews G, Goulet CT, Delhey K, Chapple DG (2016) The effect of skin reflectance on thermal traits in a small heliothermic ectotherm. J Therm Biol 60:109–124. https://doi.org/10.1016/j.jtherbio.2016.06.013

McCaw BA, Stevenson TJ, Lancaster LT (2020) Epigenetic responses to temperature and climate. Integr Comp Biol 60:1–12. https://doi.org/10.1093/icb/icaa049

Mitchell DJ, Biro PA (2017) Is behavioural plasticity consistent across different environmental gradients and through time? Proc Biol Sci 284:20170893–20170898. https://doi.org/10.1098/rspb.2017.0893

Nakagawa S (2015) Missing data: mechanisms, methods and messages. In: Fox GA, Negrete-Yankelevich S, Sosa VJ (eds) Ecological statistics: contemporary theory and application. Oxford University Press, USA, pp 81–105

Nespolo RF, Franco M (2007) Whole-animal metabolic rate is a repeatable trait: a meta-analysis. J Exp Biol 210:3877–3878. https://doi.org/10.1242/jeb.013110

Noble DWA, Nakagawa S (2021) Planned missing data designs and methods: options for strengthening inference, increasing research efficiency and improving animal welfare in ecological and evolutionary research. Evol Appl 14:1958–1968. https://doi.org/10.1111/eva.13273

Norin T, Metcalfe NB (2019) Ecological and evolutionary consequences of metabolic rate plasticity in response to environmental change. Philos Trans R Soc B Biol Sci 374:20180180–20180189. https://doi.org/10.1098/rstb.2018.0180

Nussey DH, Wilson AJ, Brommer JE (2007) The evolutionary ecology of individual phenotypic plasticity in wild populations. J Evol Biol 20:831–844. https://doi.org/10.1111/j.1420-9101.2007.01300.x

O'Dea RE, Lagisz M, Hendry AP, Nakagawa S (2019) Developmental temperature affects phenotypic means and variability: a meta-analysis of fish data. Fish Fish 20:1005–1022. https://doi.org/10.1111/faf.12394

Oufiero CE, Angilletta MJ (2006) Convergent evolution of embryonic growth and development in the eastern fence lizard(*Sceloporus undulatus*). Evolution 60:1066–1075. https://doi.org/10.1111/j.0014-3820.2006.tb01183.x

Pettersen AK, White CR, Bryson-Richardson RJ, Marshall DJ (2019) Linking life-history theory and metabolic theory explains the offspring size-temperature relationship. Ecol Lett 22:518–526. https://doi.org/10.1111/ele.13213

Piersma T, Drent J (2003) Phenotypic flexibility and the evolution of organismal design. Trends Ecol Evol 18:228–233. https://doi.org/10.1016/S0169-5347(03)00036-3

Piersma T, Lindström Å (1997) Rapid reversible changes in organ size as a component of adaptive behaviour. Trends Ecol Evol 12:134–138. https://doi.org/10.1016/S0169-5347(97)01003-3

Ponzi E, Keller LF, Bonnet T, Muff S (2018) Heritability, selection, and the response to selection in the presence of phenotypic measurement error: effects, cures, and the role of repeated measurements. Evolution 72:1992–2004. https://doi.org/10.1111/evo.13573

Réveillon T, Rota T, Chauvet É et al (2019) Repeatable inter-individual variation in the thermal sensitivity of metabolic rate. Oikos 85:935–938. https://doi.org/10.1111/oik.06392

Rodgers EM, Cocherell DE, Nguyen TX et al (2018) Plastic responses to diel thermal variation in juvenile green sturgeon, Acipenser medirostris. J Therm Biol 76:147–155. https://doi.org/10.1016/j.jtherbio.2018.07.015

Schielzeth H (2010) Simple means to improve the interpretability of regression coefficients. Methods Ecol 1:103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Schielzeth H, Nakagawa S (2020) Conditional repeatability and the variance explained by reaction norm variation in random slope models bioRxiv 2020.03.11.987073 https://doi.org/10.1101/2020.03.11.987073

Schnurr ME, Yin Y, Scott GR (2014) Temperature during embryonic development has persistent effects on metabolic enzymes in the muscle of zebrafish. J Exp Biol 217:1370–1380. https://doi.org/10.1242/jeb.094037

Seebacher F (2005) A review of thermoregulation and physiological performance in reptiles: what is the role of phenotypic flexibility? J Comp Physiol B 175:453–461. https://doi.org/10.1007/s00360-005-0010-6

Seebacher F, Grigaltchik VS (2014) Embryonic developmental temperatures modulate thermal acclimation of performance curves in tadpoles of the frog *Limnodynastes peronii*. PLoS ONE 9:e106492. https://doi.org/10.1371/journal.pone.0106492

Seebacher F, Beaman J, Little AG (2014) Regulation of thermal acclimation varies between generations of the short-lived mosquitofish that developed in different environmental conditions. Funct Ecol 28:137–148. https://doi.org/10.1111/1365-2435.12156

Shama LNS, Strobel A, Mark FC, Wegner KM (2014) Transgenerational plasticity in marine sticklebacks: maternal effects mediate impacts of a warming ocean. Funct Ecol 28:1482–1493. https://doi.org/10.1111/1365-2435.12280

Storm MA, Angilletta MJ (2007) Rapid assimilation of yolk enhances growth and development of lizard embryos from a cold environment. J Exp Biol 210:3415–3421. https://doi.org/10.1242/jeb.005652

Sultan SE, Stearns SC (2005) Environmentally contingent variation: phenotypic plasticity and norms of reaction. In: Hallgrímsson B, Hall BK (eds) Variation. Academic Press, Burlington, pp 303–332

Thompson MJ, Evans JC, Parsons S, Morand-Ferron J (2018) Urbanization and individual differences in exploration and plasticity. Behav Ecol 29:1415–1425. https://doi.org/10.1093/beheco/ary103

Toftegaard T, Posledovich D, Navarro-Cano JA et al (2015) Variation in plant thermal reaction norms along a latitudinal gradient - more than adaptation to season length. Oikos 125:622–628. https://doi.org/10.1111/oik.02323

Tomlinson S, Dalziell EL, Withers PC et al (2018) Measuring metabolic rates of small terrestrial organisms by fluorescence-based closed-system respirometry. J Exp Biol. https://doi.org/10.1242/jeb.172874

Via S, Gomulkiewicz R, De Jong G et al (1995) Adaptive phenotypic plasticity: consensus and controversy. Trends Ecol Evol 10:212–217. https://doi.org/10.1016/S0169-5347(00)89061-8

White CR, Schimpf NG, Cassey P (2013) The repeatability of metabolic rate declines with time. J Exp Biol 216:1763–1765. https://doi.org/10.1242/jeb.076562

Withers PC (1992) Comparative animal physiology. Saunders College Pub, Philadelphia