Impact of developmental temperatures on the repeatability of thermal plasticity in metabolic rate

Fonti Kar1, Shinichi Nakagawa1, Daniel W.A. Noble2

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*

Corresponding author: Fonti Kar

Correspondence email: [fonti.kar@gmail.com](mailto:fonti.kar@gmail.com)

Abstract

Phenotypic plasticity is an important mechanism that allows populations to adjust to changing environments. Plastic responses induced by early life experiences can have lasting impacts on how individuals respond to environmental variation later in life (i.e., reversible plasticity). Developmental environments can also influence repeatability of plastic responses thereby altering the capacity for reaction norms to respond to selection. Here, we compared metabolic thermal reaction norms in lizards (*Lampropholis delicata*) that were incubated at two developmental temperatures (ncold = 26, nhot = 25). We repeatedly measured individual reaction norms across six acute temperatures10 times over ~3.5 months (nobs = 3,818) to estimate the repeatability of average metabolic rate (intercept) and thermal plasticity (slope). The intercept and the slope of the thermal reaction norm did not change with developmental temperatures. Repeatability of average metabolic rate was on average, 10% lower in hot incubated lizards and was stable across acute temperatures. The slope of the reaction norm was repeatable (*R* = 0.43) suggesting that individuals were moderately consistent in their response to acute temperature variation, however repeatability did not depend on early developmental temperature. Our work implies that thermal plasticity has the capacity to evolve, despite there being less consistent variation in metabolic rate under hot environments. This will be increasingly more important for terrestrial ectotherms living in changing climate.

Keywords

reaction norm, repeatability, metabolic rate, incubation temperature, thermal performance curve, thermal sensitivity, phenotypic flexibility

Introduction

A substantial amount of variation in an individual’s phenotype is determined by formative processes throughout embryonic development. Environmental perturbations during this critical period can have persistent effects on an individual’s physiology, morphology, behaviour and life history (Eyck et al., 2019; Noble et al., 2018; O’Dea et al., 2019). Developmental shifts in phenotypes may be adaptative 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 later life conditions (Auld et al., 2010; Bonamour et al., 2019). A multitude of traits throughout an animal’s life are labile; reversibly responding to environmental change. Reversible plasticity in phenotypic traits allow individuals to adjust to moment-to-moment changes in their surroundings (Piersma & Drent, 2003), and can broadly be classified into two categories, acclimation and phenotypic flexibility (Havird et al., 2020; Piersma & Drent, 2003). 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 short-term changes in traits that are induced by acute environmental exposure, such as changes in metabolic rate in response to acute temperature (Piersma & Drent, 2003; Piersma & Lindström, 1997).

Reversible plasticity may be able to alleviate costs associated with phenotype mismatches induced by early life environments (Angilletta Jr et al., 2003; Ghalambor et al., 2007). When environments shift predictably, flexibility in the average 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 can change depending on early environmental conditions and might alter phenotypic responses to environmental variation (Beaman et al. 2016).The interaction between early- and late life 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 the developmental effects on acclimation, whereas the influence on phenotypic flexibility and variability of plastic responses is largely unknown.

It has long been recognised that individuals vary in their plasticity, with some

responding more flexibly than others (Dingemanse & Wolf, 2013; Nussey et al., 2007). Consistent variation among individuals may be heritable, but importantly, provides the phenotypic substrate for selective forces to act upon (Araya-Ajoy & Dingemanse, 2017; Nussey et al., 2007). Developmental environments, however, can influence variation available for selection (Sultan & Stearns, 2005). For example, zebra finches that experience nutritional stress as nestlings weigh less and have reduced growth rates contributing to increases in consistent among individual differences in metabolism and behaviour (Careau, Buttemer, et al., 2014). Consistent among individual variation in plasticity has also been reported in other labile traits including aggressiveness in great tits (Araya-Ajoy & Dingemanse, 2017), explorative behaviour in chickadees (Thompson et al., 2018) and metabolic rate in amphipods (Réveillon et al., 2019). Whether developmental environments affect consistent variation in plasticity *per se* is still poorly understood. 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 & Franco, 2007; Norin & Metcalfe, 2019). All organisms require energy for growth, maintenance and reproduction (Careau, Killen, et al., 2014). Numerous studies have investigated the influence of various developmental environments, such as temperature (Gangloff et al., 2015; Noble et al., 2018), UV exposure (Kazerouni et al., 2016), and dietary restriction (Careau, Buttemer, et al., 2014) on metabolic rate, however, the impacts on plasticity of metabolic rate is not well established (but see Seebacher et al., 2014). Developmental cues could influence metabolic plasticity, possibly through modifications in metabolic enzymes or cellular membrane structure that influence their function in different environments (Angilletta Jr, 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 the capacity to respond to selection might also be specific to early life conditions. Understanding how early life environments shape metabolic plasticity will be important for ectotherms that develop in environments that fluctuate throughout their life.

Here we employed a ‘reaction norm approach’ to examine the impact of developmental temperature on plasticity of metabolic rate in an oviparous skink (*Lampropholis delicata*) (Via et al., 1995). 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 (nobs = 3,818) that hatched from two incubation treatments (nhot = 25, ncold = 26) to address the following key questions: (1) How does developmental temperature change the intercept and slope of the thermal reaction norm?; (2) How does the repeatability of metabolic plasticity (i.e. slope of the reaction norm) change with developmental temperatures? (3) Does developmental temperature treatments differ in their repeatability of metabolic rate (intercept) at each acute temperature (i.e. temperature-specific repeatability)? Our experimental approach will bring important insights on how development environments mediate the capacity for ectotherms to respond to thermal variation.

Materials and Methods

*Lizard Collection and Husbandry*

We established a breeding colony of adult *L. delicata* (nfemales = 144, nmales = 50) using wild individuals collected across five sites throughout the Sydney region between 28 August and 8 September 2015. Three females were housed with a single male in opaque plastic enclosures measuring 35cm 25cm 15cm (L W H). Enclosures were kept under UV lights (12L:12D) in a temperature-controlled room set to 24ºC. Lizards had access to a heat lamp that elevated temperatures on one side of the enclosure to 32 ºC. 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 *ad libitum* (*Acheta domestica*) 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 cm 17.5 cm 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) and all procedures 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 over October 2017 – March 2018. When eggs were discovered, they were weighed using a digital scale to the nearest 0.01g (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 0.01mm. Following measurements, each egg was placed in a plastic cup (80ml) containing three grams of vermiculite and four grams of water and covered using cling wrap which was secured by an elastic band. 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 29ºC whereas the ‘cold’ treatment was exposed to a mean temperature of 23ºC. Both incubators fluctuated +/- 3ºC the mean temperature over a 24-hour period. These treatments represent the temperature extremes of natural nest sites 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 47.7days (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). We used closed-system respirometry instead of flow-through respirometry because it was more logistically feasible given our need to measure a large number of hatchlings at a range of temperatures. 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 & Dingemanse 2015). MR was measured as the volume of CO2 production per unit time ( mL min-1) as CO2 production is less susceptible to fluctuations in water vapour and more feasible to detect in smaller organisms (Lighton, 2008; Tomlinson et al., 2018). Nonetheless, CO2 production was strongly correlated with O2 consumption (*r* =0.81, p < 0.05]) with RQ values averaged 0.77 (SD = 0.41). Due to logistical constraints, lizards were randomly assigned to one of two blocks for MR measurements (block 1: n =26, block 2: n = 25). We sampled lizards once a week for two-weeks consecutively and then allowed them to rest for one week before the next week of measurements. Each week of measurements 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 (+/- 1ºC).

Metabolic rate was measured at 24ºC, 26ºC, 28ºC, 30ºC, 32ºC and 34ºC in a randomised order. However, at each sampling session we intentionally missed measurements at two randomly selected temperatures using a planned missing data design (Nakagawa, 2015; Noble & Nakagawa, 2018). Missing data was imputed during analysis (see Statistical analysis). 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 minutes to allow body temperatures to equilibrate. The lids of the chambers were left ajar during this time to minimise CO2 build up. After 30 minutes, each chamber was flushed with fresh air and sealed. A 3 mL ‘control/baseline’ air sample was immediately taken via a two-way valve to account for any residual CO2 that was not flushed from the chambers. The chambers were left in the incubator at the set temperature for lizards to respire for 90 minutes. After this time, two replicate air samples (3mL) were taken from each chamber in order to estimate the change in CO2. Two samples were taken so we could explicitly estimatemeasurement error (see Statistical analysis, 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.

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 and *.* 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 CO2 produced by an individual was calculated following (Core Team, 2013):

Equation: 1

where %CO2 is the maximum percentage of CO2 in air sample above baseline, which was corrected by subtracting any ‘residual’ CO2 from the initial flush from the larger of the two air samples; Vchamber is the volume of the chamber (70 mL); Vlizard is the volume of the lizard, assuming that the mass of the lizard is the same as its volume, and *t* is the duration of time in minutes after where the chamber has been sealed and the first air sample was taken (90 minutes).

Statistical analysis

We fitted Bayesian linear mixed effect models in *R* (Core Team, 2013) using the package ‘*brms’* (Bürkner, 2017). Metabolic rate was log transformed and body mass, age and temperature were z-transformed so parameter estimates of main effects and interaction terms were more interpretable (Schielzeth, 2010). Our planned missing data design resulted in random missingness across temperatures (36% missingness in MR and body mass) The package *brms* is capable of performing model-based data imputation. As such, we performed imputation during model fitting in all of analyses. Model-based imputation not only retains the hierarchical structure of the dataset but also increases statistical power (P. Bürkner, personal communication 25 October 2020, Nakagawa, 2015). Sensitivity analyses suggest that models with imputed data resulted in similar conclusions to complete case analyses. However, we present results from the imputation analysis in the main text as parameter estimates were more precise (See ESM). For all models we used default priors and ran four Markov Chain Monte Carlo (MCMC) chains; taking 800 samples from the posterior distribution after discarding the first 1500 iterations. This gave a total of 3200 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, suggesting 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 and an interaction between treatment and temperature as fixed effects. 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 measurement error can vary by temperature (Chapter 2), here we assumed that measurement error was constant across temperatures by fitting it as a random intercept as estimating a random slope for resulted in model convergence issues. Heterogeneous variance across temperatures can also influence parameter estimates (Careau, Buttemer, et al., 2014). However, heterogeneous residual variance was not supported by our data, therefore homogenous variance was used in all models (Table S1). 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 (*Rslope*) 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 as fixed effects. 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 individual across all sampling sessions. A random temperature slope was estimated for lizard identity and series which allowed us to calculate slope repeatability. 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):

where: is the among-individual variance in the temperature slope term and the is the within-individual variance in the temperature slope.

We estimated adjusted repeatability of average metabolic rate (i.e. intercept of the reaction norm) at each acute 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 fixed effects. 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 *It*  following Schielzeth and Nakagawa (2020):

where is the among individual variance in intercepts, is the specific temperature at which repeatability is calculated for, is the among individual variance in slope and is the covariance between the intercept and slope at the among individual level. Temperature specific repeatability () is then calculated as follows:

where: is the variance due to sampling session and is residual variance.

We also wanted to estimate overall repeatability of average metabolic rate across all acute temperatures. We 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 acute temperatures. Similarly, we calculated repeatability as per the equation above but using just the single estimate for 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 developmental treatment from the hot (Hot – Cold). To test whether differences among treatments were significant, we calculated probabilities of direction (*pd*) using the package ‘bayestestR’ (Makowski, Ben-Shachar, & Lüdecke, 2019). The probability of direction is calculated relative to the posterior median therefore it ranges from 50 -100%. The value of *pd* describes whether an effect is either positive or negative as is always relative to the sign of the median (Makowski, Ben-Shachar, Chen, et al., 2019). If the median is positive, then *pd* describes the proportion of the posterior distribution that is also positive (Makowski, Ben-Shachar, Chen, et al., 2019). A *pd* value of 95% can be interpreted as the effect is positive with a probability of 95%.

Results



**Figure 1.** Predicted thermal reaction norm of metabolic rate (VCO2 min-1 g-1) for the ‘cold’ developmental temperature group (blue line, nlizards = 26) and the ‘hot’ developmental temperature group (red line, nlizards = 25) Points are raw data and are coloured according to treatment groups nobs = 3818. Dashed lines represent the upper and lower bounds of 95% credible intervals.

We found no evidence to suggest that metabolic rate or its response to acute temperature was influenced by early developmental temperature (Fig. 1, Table 1, Table S2). Congruently, there were no treatment differences in thermal reaction norms at the first sampling session when acclimation effects is likely to have the least effect (see ESM). We therefore refitted the model with just the main effects (Table S3-4). Across all models, temperature and body mass had positive effects on metabolic rate (Table 1, Table S3-4). Nonetheless, reaction norms slopes were significantly repeatable but repeatability of slopes (*Rslope*) did not depend on developmental temperature treatments (Hot: *Rslope* = 0.42, 95% CI: 0.04 – 0.91; Cold: *Rslope* = 0.46, 95% CI: 0.03 – 0.95; *pd* = 53.5%, Fig. 2, Table S6-9). A *pd* value of 53.5% indicates that there is roughly equal probability that the difference in *Rslope* is positive or negative, indicating little difference among treatment groups.

Overall, temperature-specific repeatability was relatively low and the cold developmental treatment tended to have higher repeatability estimates compared to the hot developmental treatment (Fig. 3, Fig S1, Table 2). Irrespective of acute temperature, repeatability of average metabolic rate was on average 10% higher in cold incubated lizards (*pd* = 95.7%, Fig. 3B, C). There was a 95.7% probability that difference in overall repeatability was negative, which indicates that lizards from the cold treatment are more likely to have higher repeatability. Higher repeatability in the cold treatment was associated with significant differences among individual variation and residual variance (Fig. S2).



**Figure 2.** Thermal reaction norms of mass-adjusted metabolic rate for lizards reared at ‘hot’ developmental temperatures (top panel, red lines, nlizards = 25) and ‘cold’ developmental temperatures (bottom panel, blue lines, nlizards = 26) at session number one, five and ten. Lines represent individual reaction norms for a random subset of 10 individuals from each treatment.

**Table 1** Model coefficients of full model testing whether developmental temperature affects the elevation (intercept) and slope of the thermal reaction norm of metabolic rate. This model used an imputed dataset of nobs = 6000, 36% of observations were imputed. The intercept is the cold developmental temperature. MR was log transformed and mass, age and temperature were z-transformed. Bolded estimates are significantly different from zero. Lower and upper bound of estimates represent 95% credible intervals. COV represents covariance. Main effects model is presented in Table S3

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Estimate** | **Lower** | **Upper** |
| *Fixed effects* |  |  |  |
| Intercept MR | **-6.292** | **-6.372** | **-6.218** |
| Treatment 29 | -0.003 | -0.062 | 0.058 |
| Acute Temperature | **0.262** | **0.246** | **0.278** |
| Treatment 29 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 |
| COVIntercept – 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** |
| Residuals | **0.041** | **0.038** | **0.043** |



**Figure 3. A)** Temperature-specificadjusted repeatability for average metabolic rate for the ‘cold’ developmental temperature group (blue, nlizards = 26) and the ‘hot’ developmental temperature group (red, nlizards = 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 acute temperature. **C)** The difference in repeatability (Hot – Cold) overall and at each acute temperature. Difference between treatments are represented as posterior distributions and median point 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 an imputation model. Contrasts are presented in Table S5.

Discussion

Contrary to our expectations, early developmental temperature did not change the intercept or slope of the population reaction norm. Thermal plasticity of metabolic rate (slope) was significantly repeatable but its repeatability was also unaffected by developmental temperature. In line with our prediction, temperature-specific repeatability of metabolic rate (i.e., intercept) was lower among lizards that were reared in hot developmental temperatures. Our results suggest that, while individuals displayed consistent variation in their plasticity (I x E), how metabolic rate responds to acute temperature variation later in life was robust to thermal extremes 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 in fluctuating environments

Thermal reaction norms of metabolic rate are robust to developmental temperature

Developmental environments that affect later life plasticity may alter how 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 & Barrett, 2017; McCaw et al., 2020). However, our results suggest instead that thermal reaction norms for metabolic rate were robust to changes in incubation temperature. Results have been mixed among the few studies that have investigated the effects of pre- and post-hatching temperature on the plasticity of metabolic rate (Table 1, Beaman et al., 2016). For example, wild caught mosquitofish 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 & Grigaltchik, 2014). Given that our lizards were reared in a common environment post hatching, the lack of difference we observed may be the result of reversible plasticity resulting from acclimation in metabolic rate to the laboratory conditions. It is possible that acclimation capacities may have overwhelmed any developmental differences in thermal reaction norms. Acclimation of physiological function takes approximately 3-4 weeks to complete, so it is likely that acclimation had already taken place by the time we began the study when lizards were about ~2.5 months old (Seebacher, 2005). Future studies should employ cross factorial designs where post-hatch environments are deliberately matched and mismatched with early environmental conditions to disassociate acclimation effects (Kazerouni et al., 2016; Schnurr et al., 2014).

Stable thermal reaction norms of metabolic rate across both developmental temperatures has 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 predecessors may have canalized population responses so that they are less sensitive to fluctuations in developmental temperatures (Liefting et al., 2009). Canalization may reduce the costs of phenotypic plasticity during development if environmental variation is predictable across generations (Aubret & 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 (Bonamour et al., 2019). While our incubation treatments represent thermal extremes of natural nest sites, they may not have been severe enough to induce changes in the thermal reaction norms, particularly given that we used more realistic fluctuating nest temperatures. Developmental stress is thought to lead to recruitment of heat shock proteins thereby changing reversible plasticity later in life (Beaman et al., 2016; Chevin & Hoffmann, 2017). Recent work has shown lizard embryos exposed to extreme heat produce higher levels of heat shock proteins and had greater thermal tolerance, however this subsequently reduced tolerance later in life (Gao et al., 2014). This implies there may be constraints in how phenotypic responses can be shaped by extreme developmental environments.

Developmental temperatures and repeatable thermal plasticity of metabolic rate

Repeatability of reaction norm slopes did not change with developmental temperature, but lizards reared in hot temperatures had reduced temperature-specific 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. Developmental time exhibits a negative nonlinear relationship with temperature, such that development times are considerably shorter at hotter temperatures (Marshall et al., 2020; Noble et al., 2018). 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, we found that incubation duration was short and less variable in the hot developmental treatment. Shortened development may restrict embryo yolk assimilation that is needed for growth (Oufiero & Angilletta, 2006; Storm & Angilletta, 2007). Elevated levels of proton leak at hot developmental temperatures leads to less efficient energy production and may explain why metabolic rate did not differ among treatments despite changes in repeatability (Chamberlin, 2004). Reduced consistent variation under hot nest temperatures may be problematic as global temperatures continue to rise (Botero et al., 2015). Provided that some of the repeatable differences in metabolic rate is heritable (Dohm, 2002; Falconer, 1952), our results suggest that the evolutionary potential of metabolic rate may be dampened for populations living in warming environments. However, populations may be able to evolve metabolic plasticity in order to persist under rising temperatures (Ghalambor et al., 2007).

We found that individuals consistently vary in how their metabolic rate changes with acute temperature. While several studies have reported significant among individual variation in thermal plasticity slopes (Briga & Verhulst, 2017; Careau, Gifford, et al., 2014), 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 another study of the same species (*R* = 0.23, Chapter 1). Similarly, moderate repeatability of thermal sensitivity of metabolic rate has also been observed in amphipods (*R* = 0.38). Assuming that repeatable reaction norm slopes have a heritable basis (Driessen et al., 2007), our work implies that thermal plasticity can be selected upon and therefore evolve (Falconer, 1952; but see Dohm, 2002).

Consistent individual differences in metabolic rate were stable across acute temperatures. This result demonstrates that temperatures within the operable range of *L. delicata* maintains consistent individual differences in MR (Matthews et al., 2016). Repeatability in metabolic rate may be an important mechanism that promotes consistent variation in thermoregulation, behaviour and life history (Goulet et al., 2017; Réale et al., 2010; Sæther, 1987). Overall, our estimates for the repeatability of MR ranged from 0.09 – 0.22. Our results are in line with a meta-analysis that showed that repeatability decreases with time (White et al., 2013). 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, Chapter 2) was comparatively larger relative to this study. This is likely due to life stage differences in environmental effects that shape phenotypic variation. As individuals mature, their experiences in different microhabitats (diet, thermal preferences) can promote among-individual variation in traits (Kruuk & Hadfield, 2007). Such common (micro) environment effects could further increase repeatability and may explain differences between lab and wild studies (Auer et al., 2016).

Conclusions

The role of developmental temperature on phenotypic plasticity exhibited later in life is complex. 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 repeatability in temperature-specific metabolic rate. This has key evolutionary implications as reduced consistent individual variation in hot temperatures may alter a population’s ability to respond to selection under warming climate. However, populations may still have the ability to evolve in their thermal reaction norms as individuals consistently varied in their thermal plasticity in metabolic rate. Elucidating the role of developmental environments on shaping plastic responses may require more stressful incubation conditions and cross-factorial experimental designs to disassociate the effects of acclimation from developmental plasticity.

Data accessibility

Datasets and code used to generate results of this study is accessible via Open Science Framework (DOI: XXXXXXXXXXX)

Acknowledgements

We would like to thank Martin Whiting for the use of his facilities at Macquarie University. We are grateful for the assistance of numerous Lizard Lab members and interns with husbandry duties. Special thanks to Christine Wilson for her commitment to caring for our animals. We thank Timothee Bonnet for his advice on partitioning measurement error from our models.

References