Developmental temperature and repeatability of metabolic rate across temperatures

Fonti Kar1, Shinichi Nakagawa1,2, Danie l W.A. Noble1-4

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

2 *Diabetes and Metabolism Division, Garvan Institute of Medical Research, 384 Victoria Street, Darlinghurst, Sydney, NSW 2010, Australia*

4 *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. Theory predicts that plastic responses induced early in life experiences can have lasting impacts on how individuals respond to environmental variation later in life (reversible plasticity). Cues experienced during developmental can also influence variability of reversible plastic responses. Environment environments can thus alter the capacity for reaction norms to respond to selection which may facilitate evolutionary change in threatened populations. Here, we compared thermal reaction norms of metabolic rate in lizards (*Lampropholis delicata*) that were incubated at two developmental temperatures (ncold = 26, nhot = 25). We repeated assayed individual reaction norms across six acute temperatures to estimate the repeatability of the slope and repeatability of average metabolic rate. The elevation and the slope of the thermal reaction norm was not susceptible to changes in developmental temperatures. The slope of the reaction norm was repeatable (*R* = 0.43) suggesting that individuals were consistent in their response to acute temperature variation, however consistency did not depend on developmental temperatures. Repeatability of average metabolic rate was stable across acute temperatures and did not differ between developmental temperatures. Our results suggest that thermal extremes in natural nest temperatures maintains constancy in thermal plasticity. Moreover, selective processes may be able to operate on consistent expression of metabolic rate as well as thermal plasticity. Understanding how environments at different life stages impact plasticity and its capacity to evolve will become increasingly more relevant for terrestrial ectotherms.

Keywords

reaction norm, repeatability, heritability, metabolic rate, incubation temperature

Introduction

A substantial amount of variation in an individual’s phenotype is determined by formative processes that occur during embryonic development. As such, 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., 2017; O’Dea et al., 2019). These phenotypic responses can have significant impacts such as changing population dynamics and facilitating the evolution of novel traits (Forsman, 2014; Moczek et al., 2011; West-Eberhard, 2003). Development plasticity allows embryos to prime themselves to environments they will eventually survive in (Beldade et al., 2011), however environment-phenotype mismatches can occur when developmental cues fails to predict eventual environments (Auld et al., 2010; Bonamour et al., 2019). This can incur high fitness costs as individuals would express maladaptive phenotypes relative to the selective environment (DeWitt et al., 1998). However, reversible plasticity that is modulated by developmental cues may allow individuals to fine tune their phenotype later in life to avoid mismatches (Beaman et al., 2016).

After birth, an individual’s phenotype can exhibit reversible plasticity in response to environmental variation (I x E) (Nussey et al., 2007). Such adjustments may confer adaptive benefits, allowing individuals to compensate for environmental changes (Angilletta Jr et al., 2003; Ghalambor et al., 2007). Reversible plasticity can be broadly classified into two categories, namely phenotypic flexibility and acclimation (ref). Phenotypic flexibility describes short-term changes in labile traits induced by acute exposure to an environmental cue such as heart rate in response to hypoxia (Piersma & Drent, 2003; Piersma & Lindström, 1997). This form of plasticity is typically represented as a reaction norm or performance curve across different environments (Huey & Kingsolver, 1989; Via et al., 1995). Acclimation, on the other hand, requires chronic exposure to an environmental cue and involves remodelling of physiological systems over longer periods. Acclimation results in shifting of the reaction norm optimum (Seebacher, 2005; Seebacher et al., 2015). Traditionally, developmental plasticity and reversible plasticity are considered separate biological processes, yet theory suggests that early developmental cues may actually underpin the capacity for phenotypes to change later in life (Beaman et al., 2016; Seebacher et al., 2014).

Reversible plasticity may be able to alleviate costs associated with phenotype mismatches induced by early life environments. Cues experienced during development not only conveys information about average changes in the environment such as decreases in precipitation but also its variability i.e. how often rainfall occurs (Bonamour et al., 2019). When environments shift predictably, flexibility in the average phenotype would be advantageous because individuals can adjust to prevailing conditions accordingly to avoid discrepancies between the environment and the phenotype (Botero et al., 2015). For example, seasonal variation in temperature during development reliably predicts dispersal strategies in adult spiders (*Erigone atra*), likely in response to changes in resource availability (Bonte et al., 2008). The interaction of early- and later life plasticity is supported by a growing number of studies that show developmental differences in reaction norm of a variety of traits such as mitochondrial function (Shama et al., 2014), metabolic rate (Seebacher et al., 2014), locomotor performance (Kazerouni et al., 2016). However, these studies solely focus on developmental regulation of acclimation, whereas the influence on phenotypic flexibility is largely unknown. Moreover, these studies have neglected to consider potential confounds such as local adaptation (Amarillo-Suárez & Fox, 2006; Stillwell & Fox, 2009) and parental effects (Bentz et al., 2013; Polačik et al., 2017) that could also affect the shape as well as the variability of reaction norms later in life (George et al., 2017).

Much of existing research on the influence of developmental environments focuses on changes in the phenotypic mean. However, in order to understand the adaptive potential of developmental responses we need to also consider its influence on phenotypic variability (Nakagawa et al., 2015). Developmental environments can alter patterns of gene expression which can manifest as changes in phenotypic variability (Colinet & Hoffmann, 2012; Jones, 2012). Developmental stressors can trigger the release of hidden genetic variation which may harbour beneficial phenotypic variants that can survive under stressful conditions (McGuigan & Sgrò, 2009). Variation induced by the environment may become heritable (via epigenetic modifications) which could ultimately lead to beneficial traits become genetically assimilated (Crispo, 2007), allowing populations to persist and adapt to novel environments (Ghalambor et al., 2007). Phenotypic variation is therefore essential for evolutionary change as it provides new material for selection to operate on (Falconer, 1952), however this depends on the degree to which variability is consistent over time (Nakagawa & Schielzeth, 2010). Despite its importance, few studies have empirically measured the effects of developmental plasticity on consistent expression of phenotypic variance (Careau et al., 2014; Kaiser et al., 2019; but see O’Dea et al., 2019).

Energy metabolism is a key fitness related trait that is both repeatable and highly labile (Nespolo & Franco, 2007; Norin & Metcalfe, 2019). All organisms require energy to undertake all biological processes (Careau, Killen, et al., 2014). Across broad scales, metabolism dictates energy flow which shapes complex community structures (Barneche & Allen, 2018). At the individual level, metabolic rate determines energy expenditure which has important consequences on resource allocation and life history evolution (Biro & Stamps, 2010; Réale et al., 2010; Ricklefs & Wikelski, 2002). Metabolic rate is also strongly linked with other fitness components such as body size (Gillooly et al., 2001) and reproductive senescence (Friesen et al., 2017) which implies that phenotypic changes in metabolic rate may compromise fitness and survival (Burton et al., 2011; Pettersen et al., 2016). As such, numerous studies have investigated the influence of various developmental cues such as temperature (Gangloff et al., 2015; Noble et al., 2017), UV exposure (Kazerouni et al., 2016), nutrition (Careau, Buttemer, et al., 2014) on metabolic rate and in some instances, the impact can be long lasting on individuals (Noble et al., 2017). However, in order to understand how animals adjust to both early and late life environments, we need to turn our focus to understanding how *plasticity* of metabolic rate is impacted.

Here we examined how developmental temperature impacts phenotypic flexibility to acute temperature change in an oviparous skink (*Lampropholis delicata*). Specifically, we were interested in whether developmental temperature affects the overall thermal reaction norm of metabolic rate, as well phenotypic variation of metabolic rate at different temperatures.. Over three and a half months, we repeatedly measured routine metabolic rate at six temperatures for lizards (nobs = 3818) that hatched from two incubation treatments (nhot = 25, ncold = 26) to address the following key questions: (1) How does developmental temperature change: (1) the overall thermal reaction norm of metabolic rate (elevation and slope); (2) the repeatability of the slope of the reaction norm and (3) temperature-specific repeatability? We expect lizards that hatched from the hot developmental temperature would have on average higher metabolic rates and steeper reaction norms. Moreover, we expect increases in repeatability under high thermal stress for average metabolic rate as well as the slope. Our experimental approach will provide important insights on how development cues mediate the capacity for animals to respond to temperature variation.

Materials and Methods

*Lizard Collection and Husbandry*

Since 2015, 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 side of the enclosure to 28-32 ºC. Each enclosure was lined with newspaper and lizards had constant access to water and tree bark 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 egg laying seasons (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 second 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. As soon as eggs were found, 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 using an elastic band. Each clutch was pseudo-randomly assigned to one of two fluctuating developmental 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 period for the ‘hot’ treatment was 29.36 days (SD = 2.17, range = 15 - 49) days and 48.48 days (SD = 4.18, range = 25 - 56) for the ‘cold’ treatment (Kar et al unpublished – Chapter 3).

*Planned Missing Data and Metabolic Rate At Different Temperatures*

Metabolic measurements commenced in April 2018 and went on till July 2018. At the start of measurements, hatchlings were approximately on average 88.68 days old (SD = 23.75, range = 26 - 131). Given the scope of our experiment, we used closed-system respirometry. We quantified routine metabolic rate (hereafter referred to as metabolic rate [MR]) as our measurements included the energetic costs of random movements that we were not able to control for (Withers 1992; Mathot & Dingemanse 2015). MR was measured as the volume of CO2 production per unit time ( mL min-1) as CO2 production is more sensitive to change in smaller organisms, and is less susceptible to fluctuations in water vapour. Nonetheless, CO2 production was strongly correlated with O2 consumption (*r* =0.81, p = <0.05]). 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).

MR was measured at 24ºC, 26ºC, 28ºC, 30ºC, 32ºC and 34ºC in a randomised order however, at each sampling session we purposely missed measurements at two random temperatures which were imputed during 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. 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 measurement error (see Statistical analysis). Chambers were then reopened and flushed with fresh air before placed back into the incubator for the second measurement temperature (2 temperatures / day) following the same procedure.

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

All analyses were conducted in *R* (Core Team, 2013)*.* We checked the data for potential input or mechanical errors using density and Cleveland plots, for more details see ESM. MR and mass was log transformed. We fitted linear mixed models in *brms* (Bürkner, 2017). For all models we used noninformative priors with 4000 iterations with a burn in 1500, sampling from the posterior distribution every fifth iteration. We ensured proper mixing by inspecting trace plots and checked that scale reduction factors were less than 1.01. We report posterior means and 95% credible intervals for all parameters throughout.

A previous study with the same sampling design demonstrated that measurement error decreased with increasing temperature. As such, we wanted to statistically account for its measurement error as it may conflate repeatability and heritability estimates (Ponzi et al., 2018). We did this by fitting a nested random effect of lizard identity, sampling session and temperature in all our models (e.g. ID001\_s1\_temp24). This nested random effect (hereafter referred to as measurement error) groups the two replicates together and partitions out the variance attributed to difference among replicates.

A previous study in the same species using a similar experimental design found that that individual responses to acute temperature change was moderately repeatable (Kar et al unpublished, Chapter 2). The same study also showed that measurement error decreased with temperature (Kar et al unpublished, Chapter 2). We therefore attempted to use model selection to determine the most appropriate random effects structure for our analysis. In all models, temperature and body mass were included as fixed effects. Despite our efforts in running more iterations and setting stronger priors, we encountered convergence issues for estimating random temperature slopes for measurement error. As such, we were unable to use a model selection approach and opted to fit random intercepts for lizard identity, sampling session number and measurement error and a random slope for lizard identity only for all subsequent analyses unless stated otherwise.

Heterogenous residual variance may influence estimates of repeatability. We therefore explicitly modelled residual variance to change over temperature in and verified if it was better supported than our homogenous variance model using WAIC values. Homogenous variance was better supported by our data, as such we did not incorporated heterogenous variance our subsequent models (Table S1).

One benefit of using *brms* is its in-built function to perform data imputation during model fitting using the function *mi* (See Data accessibility). This not only retains the hierarchical structure of the imputed data but also ultimately increases statistical power. We performed imputation during model fitting in all of analyses described below and also performed the same analyses using complete case data which are presented in the ESM. Overall, conclusions matched across imputation and complete case analyses and we therefore present the imputation analysis in the main text

First, we investigated whether developmental temperatures influenced the elevation and slope of the reaction norm. We fitted a model with MR as the response and included an interaction term between treatment and temperature to test for treatment differences in reaction norm shape.

Second, in order to estimate the repeatability of the slope of the reaction norm (*Rslope*), we fitted separate models for each treatment group containing a nested random effect of individual identity and sampling session, hereafter referred to as series (Araya-Ajoy et al., 2015). We fitted series in place of the random intercept of sampling session. Series groups together all the measurements from an individual at a given sampling session and allows partitioning variance that is attributed to within an individual, among sampling sessions. We fitted a random slope of temperature for series which allowed the model to estimate an ‘overall’ among sampling session slope. The repeatability of the slope is thus the proportion of variance in slopes explained among individual differences and is calculated as:

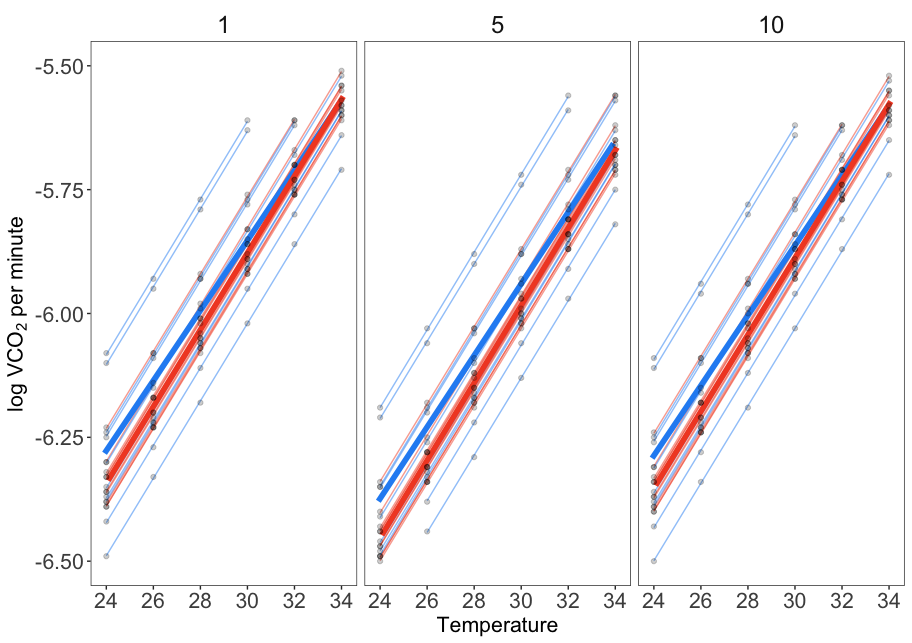
where: is the among individual variance in the temperature slope term and the is the among sampling session within individual variance in the temperature slope

Lastly, we ran separate models for each treatment group to test whether developmental temperature may have impacted temperature-specific repeatability of average metabolic rate. Each model had MR as the response and temperature, body mass and age as fixed effects and the random effects structure described above. We first calculated among individual variance in metabolic rate at each temperature *It*  following Schielzeth and Nakagawa (n.d., in review):

where is the among individual variance in intercepts, is the specific temperature at which repeatability is calculated for, is the among individual 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

Results



**Figure 1.** Predicted thermal reaction norms of metabolic rate for the ‘cold’ developmental temperature group (thick blue line, n = 26) and the ‘hot’ developmental temperature group (thick red line, n = 25). Predictions were made from an imputation model. There were no significant difference among treatment in the elevation or slope of the reaction norm (see Table 2). Thin lines present individual reaction norms for a subset of 10 individuals from each treatment. Grey points represents model predictions for individual’s mean log metabolic rate. Each panel represents a distinct sampling sessions to illustrate the consistency of individual reaction norms. Note that a slight ‘jitter’ was added to each treatment’s reaction norms to highlight the presence of two reaction norms.

Overall, we found no evidence to suggest that elevation or the slope of the thermal reaction norms of metabolic rate differed between developmental temperatures (Fig. 1, Table 1, Table S2). Both temperature and body mass had positive effects on metabolic rate. Model coefficients for the main effect model is presented in Table S3-4

**Table 1** Model coefficients of full model testing whether developmental temperature affects the elevation and slope of the thermal reaction norm of metabolic rate. This model used an imputed dataset of n = 6000. The intercept is the cold developmental temperature. Note that the imputation model also estimates an intercept and residual variance for mass as it was also missing data. Mass and MR was log transformed and Age was z-transformed. Bolded estimates are significantly different from zero. Values with \* indicate very small values that are still greater than zero.

|  |  |  |  |
| --- | --- | --- | --- |
| Parameter | Estimate | Lower | Upper |
| Intercept MR | **-7.618** | **-7.84** | **-7.397** |
| Intercept Mass | **-1.442** | **-1.449** | **-1.436** |
| Treatment 29 | 0.135 | -0.069 | 0.344 |
| Temperature | **0.077** | **0.072** | **0.081** |
| Age | -0.035 | -0.078 | 0.009 |
| Treatment 29 Temperature | -0.005 | -0.011 | 0.002 |
| Mass | **0.622** | **0.507** | **0.733** |
| VI, Intercept | **0.012** | **0.001** | **0.038** |
| VI, Slope | **0\*** | **0\*** | **0\*** |
| Vsession, Intercept | **0.01** | **0.003** | **0.029** |
| Vmeasurement error, Intercept | **0.044** | **0.04** | **0.049** |
| COVI, Intercept – I, Slope | -0.000115 | -0.000823 | 6.63e-05 |
| Residual MR | **0.041** | **0.038** | **0.043** |
| Residual Mass | **0.043** | **0.041** | **0.045** |

Individual slopes of the thermal reaction norm were repeatable in both treatment groups, however there were no treatment differences (Fig. 1). This result should be interpreted with caution as repeatability of the slope was estimated with a large degree of uncertainty (Hot: Rslope = 0.44 , 95% CI: 0.03 – 0.95; Cold: Rslope = 0.42, 95% CI: 0.03 – 0.94). Coefficients for the models that were used to estimate repeatability of the slope are presented in Table S6-9.

A picture containing table, group

Description automatically generated

**Figure 2.** Adjusted repeatability for average metabolic rate for the ‘cold’ developmental temperature group (blue) and the ‘hot’ developmental temperature group (red). Estimates were calculated from an imputation model. There were no significant differences among treatment in repeatability estimates (see Table 2). Repeatability did not change with acute temperature. Error bars represent 95% credible intervals.

Across both treatment groups, repeatability did not change across acute temperatures (Fig. 2, Table 2). There was a trend for the cold developmental treatment to have on higher repeatability compared to the hot developmental treatment however credible intervals overlapped partially (Fig. 2, Fig S2, Table 2). Model coefficients for each treatment group are presented in Table S10-13.

**Table 2** Temperature specific, adjusted repeatability estimates of log transformed metabolic rate for lizards from two developmental temperatures (nhot = 25, ncold = 26). These values were estimated from an imputation analysis, nobs = 6000. Bolded values are significantly different from zero

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Cold development temperature  n = 26 | | | | Hot development temperature  n = 25 | | |
| Temperature | Repeatability | Lower | Upper | Repeatability | Lower | Upper |
| 24 | **0.22** | **0.11** | **0.37** | **0.09** | **0.03** | **0.18** |
| 26 | **0.22** | **0.12** | **0.37** | **0.09** | **0.03** | **0.19** |
| 28 | **0.22** | **0.11** | **0.36** | **0.1** | **0.04** | **0.2** |
| 30 | **0.22** | **0.12** | **0.36** | **0.11** | **0.04** | **0.21** |
| 32 | **0.22** | **0.11** | **0.36** | **0.12** | **0.04** | **0.23** |
| 34 | **0.22** | **0.11** | **0.37** | **0.13** | **0.05** | **0.25** |

Discussion

The thermal reaction norm of metabolic rate was not susceptible to developmental temperature changes. Congruently, we found no differences among developmental temperatures in the repeatability of slope. This suggests that while individuals displayed consistent variation in their plasticity (I x E), early thermal environments did not impact individual variation in phenotypic flexibility of metabolic rate. Consistent individual variation in average metabolic rate were also unaffected by developmental temperatures, as well as acute temperatures.

The influence of developmental temperature on thermal reaction norms of metabolic rate

Thermal reaction norms of metabolic rate were robust to changes in developmental temperature. Our incubation treatments represent thermal extremes of natural nest sites and may not be distinctive enough to elicit a change in phenotypic flexibility (Cheetham et al., 2011). Among the few studies that investigated the effects of pre- and post-hatching temperature on reversible plasticity of metabolic rate, results have been mixed and lacked generality (Table 1, Beaman et al., 2016). For example, wild caught mosquitofish that developed in either spring or summer temperatures have different thermal reaction norms for metabolic scope (Seebacher et al., 2014). Whereas, there were no significant interaction between incubation temperature and reversible plasticity of metabolic rate in tadpoles of striped marsh frogs (Seebacher & Grigaltchik, 2014). The evolution of developmental control on reversible plasticity is thus likely very species specific. Past selection regimes that have optimised each species’ thermal reaction norms may allow some species to better withstand fluctuations in developmental temperatures than others. The impacts of early life environments on later-life plasticity should therefore be examined in the context of each species recent and past thermal history (Roelofs et al., 2010). While experimental studies with wild animals is valuable to understand how natural populations respond under controlled settings. Common-garden experiments may be necessary to rule out potential shared environmental effects that could affect phenotypic measurements made in the lab (de Villemereuil et al., 2016; Munday et al., 2013). Despite limited knowledge of a species’ ancestral exposure to temperature (Roelofs et al., 2010), selecting incubation temperatures based on critical thermal limits or breadth of thermal performance curves that have been shaped by evolutionary processes may allow for better detection of developmental effects on reversible plasticity.

Variability in developmental cues is also an important factor for the evolution of reversible plasticity (Bonamour et al., 2019). The magnitude as well as the variability of developmental temperatures may affect how individuals perceive the signal (Bonamour et al., 2019). For instance, increased temperature fluctuations during development might imply that future temperatures are may also predictable vary. Under this scenario, the benefits of reversible plasticity should increase as plastic strategies to offset the potential costs of an developmental environment-phenotype mismatch (Beaman et al., 2016). In support of this idea, zebrafish reared in a temperatures that shifted stochastically throughout development had greater thermal tolerance compared to fish that were reared at constant thermal regimes (Schaefer & Ryan, 2006). In the case of our study, both incubation treatments experienced the same level temperature variability (+/- 3ºC) over a 24-hr period which may explain there were no differences in their capacity to reversible adjust their MR as hatchlings. Future studies that manipulate the developmental cue variation in conjunction with magnitude would be an insightful avenue to explore how stochasticity in the environment might drive phenotypic responses.

Acclimatory responses enable organisms to maintain similar physiological rates across different environments. Acclimation involves remodelling physiological systems which causes shifts in thermal reaction norms (Seebacher et al., 2015). Both of our treatment groups were housed at the same temperature and may have acclimated to the same temperature which resulted in a convergence of their reaction norms. Indeed, studies that have shown a significant interaction between developmental environments and reversible plasticity have used a cross factorial design to disassociate such acclimation effects (Kazerouni et al., 2016; Schnurr et al., 2014). That being said, terrestrial organisms generally are more limited in their ability to acclimate compare to freshwater or marine organisms (Seebacher et al., 2015). If acclimation effects did in fact overwhelm the influence of developmental temperatures, we expected there to be treatment differences upon hatching or at very young ages. While it was not logistically possible to measure MR upon hatching, we tested for treatment differences in thermal reaction norms at the first sampling session (~2.5 months of age) compared to the last sampling session (~6 months of age). To our surprise, we found marginal differences among treatments in both the elevation and slope of the reaction norm in sampling session 10. This suggests that treatment differences may manifest later in life (Bize et al., 2003). However, the effect size was relatively small and should be interpreted carefully as we each analysis had smaller sample sizes which might contribute to an increased risk of type 1 errors.

Developmental temperatures and among-individual plasticity of metabolic rate

The repeatability of thermal plasticity and average metabolic rate did not depend on developmental temperature. Our developmental temperatures may not have been stressful enough to ‘decanalize’ cryptic genetic variation which could lead to changes in phenotypic variation and thus repeatability (Crispo, 2007). Indeed, one study found that embryos in Great Plains skinks are be able to withstand much more elevated temperatures (~42ºC) (Fitch, 1964). Moreover, the critical thermal limit of adult *L.delicata* is ~40.8ºC which suggests that the temperature extremes of nest sites may be relatively tolerable for embryo development (Greer, 2005). Although the lack of difference between developmental temperatures implies that the potential for selection to act on the metabolic reaction norms is the same across treatments, subtle differences in the underlying variance components can be masked by the ratio nature of repeatability (even heritability) calculations (Rowiński & Rogell, 2017; Wilson, 2018). Given that repeatability is a proportion of total phenotypic variance explained by individual differences (*R = VI / VI + VE*), developmental changes in residual variance (VE) can also influence repeatability. In zebra finches, nutritional stress during the juvenile stage lead to an increase in repeatability of average metabolic rate when modelling with homogenous residuals (Careau, Buttemer, et al., 2014). After accounting for heterogenous residuals among treatment groups, there was a tendency for both residual variance and among individual variance to increase in stressed birds. As such, comparisons using repeatability alone gave the impression that both treatment groups were the same (Careau, Buttemer, et al., 2014). It is worth noting that residual variance encapsulates within-individual variance which describes the stability of within individual responses which is biologically relevant for many evolutionary studies (Westneat et al., 2014), however it also captures aspects of the environment such as temporal or shared environmental effects (Kruuk, 2004; Kruuk & Hadfield, 2007). Thus, interpretations on changes in residual variance as changes in within individual variance may be misleading and researchers should be more aware of other non-genetic sources of variance. We emphasise that future studies comparing repeatability across different developmental contexts need to consider heterogenous residual variances among treatment groups. In this study, we used model selection to determine homogenous residuals was best supported by our data therefore we can conclude the among individual variance in metabolic rate is robust to changes in developmental temperatures.

The minimum requirement for evolutionary change to occur is the presence of consistent variation for selection to act on (Falconer, 1952; but see Dohm, 2002). We found repeatable differences in individual slopes. However, it should be noted that the variance components used to calculate repeatability of the slope were relatively small and repeatability it is was estimated with a substantial degree of uncertainty. Being said, our estimates of repeatability of the slope are consistent with another study of the same species (*R* = 0.23, Kar et al. unpublished). Similarly, moderate repeatability of thermal sensitivity of metabolic rate has also been observed in amphipods (*R* = 0.38). Several studies have reported significant among individual variation in thermal plasticity slopes (Briga & Verhulst, 2017; Careau, Gifford, et al., 2014), however repeatability of the slope is rarely estimated as it requires a study design that allows partitioning of within individual variance of slopes (Araya-Ajoy et al., 2015). Nevertheless, consistency in the slope of thermal reaction norms implies that thermal plasticity itself is heritable to some extent and can be shaped by selective processes (Falconer, 1952; Driessen et al., 2007; but see Dohm, 2002). Indeed, thermal plasticity has been shown to rapidly diverged in invasive populations of cabbage white butterflies, further supporting the idea that thermal reaction norms can evolve as populations experience different thermal environments (Kingsolver et al., 2007; but see Condon et al., 2014). Consistent variation in plasticity may facilitate evolutionary change in thermal reaction norms may allow populations to persist as global temperature and temperature variability continue to increase (Ghalambor et al., 2007).

Our repeatability estimates of average metabolic rate did not change across acute temperature. This result demonstrates that temperatures within the operable range of *L.delicata* maintains consistent individual differences in MR. Overall, our estimates for repeatability of MR are relatively low (*R* = 0.09 – 0.22) compared to values reported for “reptiles” (*R* = 0.86, n = 1) (Nespolo & Franco, 2007; White et al., 2013), however upon close inspection of the original study, it turns out that the repeatability estimate was for maximal MR for garter snakes and therefore not entirely comparable with our results (Garland & Bennett, 1990). We compared our results with a meta-analysis that investigated the relationship between repeatability and time interval between MR measurements (White et al., 2013). Our repeatability estimates were a lot more consistent with ectotherms (invertebrates and fish) from studies that had a measurement interval that was equal or larger than our study ( 8.5 days, *R =* 0.33, SD = 0.21, n = 18). Interestingly, repeatability of average MR in wild caught *L. delicata* (*R =* 0.3 – 0.5, Chapter 2 Kar et al. unpublished) was comparatively larger relative to this study. This is likely due to life stage differences in environmental effects that affects phenotypic variation. As individual mature, their experiences in different microhabitats can promote variation in the population (Kruuk & Hadfield, 2007). For example, some individuals may prefer warmer environments which can lead to elevated metabolic rates, increased oxidative stress, fast maturation and short lifespans (Biro & Stamps, 2010; Réale et al., 2010). Such common (micro) environment effects can bias repeatability estimates and may contribute to differences between lab and wild studies (Auer et al., 2016).

Conclusions

An individual’s ability to adjust in response to environmental change could be determined by its experiences during embryonic development. In order for such plastic responses to evolve, consistent phenotypic variation in reaction norms is required for selection to act on. In this study, we manipulated the temperature at which lizard embryos were incubated and assayed their thermal reaction norms of metabolic rate as juveniles. We demonstrated that thermal plasticity of metabolic rate, as well as the repeatability of reaction norm attributes (slope and elevation) were unfazed by developmental temperature. Thermal sensitivity in ectotherms requires a multifaceted approach. Integration of acclimation responses, thermal preferences, past population thermal experiences could reveal important insights on how different aspects of thermal adaptation can shape variation in plasticity and assist threatened ectotherms to persist in warming climate. Our focus should turn to the interactive effect of mean changes as well as variability changes in early life cues to elucidate the conditions that can induce different forms of 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