Ecologically relevant thermal fluctuations enhance offspring fitness: biological and methodological implications for studies of thermal developmental plasticity

Joshua M. Hall* and Daniel A. Warner

ABSTRACT

Natural thermal environments are notably complex and challenging to mimic in controlled studies. Consequently, our understanding of the ecological relevance and underlying mechanisms of organismal responses to thermal environments is often limited. For example, studies of thermal developmental plasticity have provided key insights into the ecological consequences of temperature variation, but most laboratory studies use treatments that do not reflect natural thermal regimes. While controlling other important factors, we compared the effects of naturally fluctuating temperatures with those of commonly used laboratory regimes on development of lizard embryos and offspring phenotypes and survival. We incubated eggs in four treatments: three that followed procedures commonly used in the literature, and one that precisely mimicked naturally fluctuating nest temperatures. To explore context-dependent effects, we replicated these treatments across two seasonal regimes: relatively cool temperatures from nests constructed early in the season and warm temperatures from late-season nests. We show that natural thermal fluctuations have a relatively small effect on developmental variables but enhance hatchling performance and survival at cooler temperatures. Thus, natural thermal fluctuations are important for successful development and simpler approximations (e.g. repeated sine waves, constant temperatures) may poorly reflect natural systems under some conditions. Thus, the benefits of precisely replicating real-world temperatures in controlled studies may outweigh logistical costs. Although patterns might vary according to study system and research goals, our methodological approach demonstrates the importance of incorporating natural variation into controlled studies and provides biologists interested in thermal ecology with a framework for validating the effectiveness of commonly used methods.

KEY WORDS: Seasonality, Nest temperatures, Phenotype, Reptile developmental plasticity, Egg incubation

INTRODUCTION

Temperature has great potential to explain variation across biological and ecological scales. Consequently, considerable effort is given to measuring thermal variation in nature and replicating this variation in controlled environments (i.e. laboratory or field manipulations; e.g. Hagstrum and Hagstrum, 1970;

Auburn University, Department of Biological Sciences, 101 Rouse Life Sciences Building, Auburn, AL 36849, USA.

*Author for correspondence (jmh0131@auburn.edu)

J.M.H., 0000-0002-5587-3402

Folguera et al., 2011; Mitchell et al., 2015; Xing et al., 2019). Such studies uncover mechanisms that link thermal variation to important ecological and evolutionary phenomena in wild populations (Worner, 1992; Georges et al., 1994); however, the success of such work may depend heavily on the ability to precisely measure temperatures in the field and accurately replicate them in a controlled study (Faye et al., 2016; Carter et al., 2018). This is challenging because temperature varies spatially and temporally at different scales (Colinet et al., 2015; Hall and Warner, 2018; Tiatragul et al., 2019; Xing et al., 2019). Such complexity results in a myriad of context-dependent thermal effects on biological processes (Liu et al., 1995; Stillwell and Fox, 2005; De Jong et al., 2010; Bannerman and Roitberg, 2014; Manenti et al., 2014; Warner and Shine, 2011; Buckley et al., 2017) that may contribute as much to phenotypic variation as genetic factors (Noble et al., 2018). Owing to technological or logistical constraints, researchers cannot perfectly replicate field temperatures, and therefore, use less complex manipulations that partially correlate with natural conditions (e.g. constant mean temperatures). Regardless, ecologists strongly advocate for recreating real-world thermal variation in the laboratory (Folguera et al., 2011; Bowden et al., 2014; Greenspan et al., 2016; Burggren, 2018; Mickley et al., 2019) because this can enhance our understanding of biological phenomena (Carter et al., 2018). However, replicating real-world temperatures has practical and statistical challenges (Colinet et al., 2015; Burggren, 2018). For example, incubators that reproduce field temperatures in a near-perfect way may be prohibitively costly or require custom-built equipment (Greenspan et al., 2016; Mickley et al., 2019). Moreover, increasing thermal complexity of treatments makes it difficult to interpret findings, replicate studies and synthesize results across the literature (e.g. meta-analysis). Given these challenges, it is critical to understand how increasing levels of thermal complexity influence results under various contexts (e.g. across temperature).

Studies of reptiles have contributed greatly to our knowledge of the ecological and evolutionary significance of thermal variation (see Warner et al., 2018; Refsnider et al., 2019). Because most oviparous reptiles (excluding archosaurs) provide no parental care, developing embryos are often subjected to wide variation in temperature, which can substantially influence fitness-relevant phenotypes of embryos and hatchlings (Janzen, 1993; Pearson and Warner, 2018; Booth, 2018). Therefore, reptiles are often used as indicator species when assessing thermal effects of global change (Telemeco et al., 2009; Tiatragul et al., 2020; Hall and Warner, 2018). For these reasons, reptile developmental plasticity is of broad interest in ecology and evolution (reviewed by Du and Shine, 2015; Noble et al., 2018; While et al., 2018; Refsnider et al., 2019; Du et al., 2019) and has great potential to advance our understanding of thermal biology using experimental approaches (Warner et al.,

2018). Early studies consisted of measuring nest temperatures with a thermometer and incubating eggs at constant mean temperatures (e.g. Licht and Moberly, 1965). Such studies may poorly reflect natural embryo environments for many species (Booth, 2018) but were logistically simple and served as the foundation for future work. Technological advances (e.g. temperature loggers, programmable incubators) have allowed researchers to precisely record nest temperatures and utilize complex incubation treatments in the laboratory (e.g. Tiatragul et al., 2020). For example, temperature loggers can be programmed to record hourly, daily and seasonal thermal variation inside nests (Pearson and Warner, 2018). Moreover, researchers can purchase or build incubators capable of near-perfectly replicating fluctuations in nest temperature (Greenspan et al., 2016) and associated stochastic temperature changes (e.g. heat waves) that typify real thermal environments (Hall and Warner, 2018; Pearson and Warner, 2018; Carter et al., 2018).

These technological and methodological advances have greatly increased our understanding of the mechanisms that underlie developmental plasticity and its ecological and evolutionary relevance (Georges et al., 2005; Warner and Shine, 2011; Bowden et al., 2014; Carter et al., 2018). For example, because developmental rate increases with temperature, embryos spend a greater portion of development at warmer versus cooler temperatures when nest temperatures fluctuate. In addition, natural thermal fluctuations frequently exceed the optimal developmental range of temperatures (Hall and Warner, 2020), whereas constant temperatures, even with similar means, remain within this range throughout development (Les et al., 2009). As a result, a host of phenotypes differ between eggs incubated at fluctuating versus constant temperatures, demonstrating the importance of replicating natural temperatures (Booth, 2018). However, most studies still use constant temperature treatments (While et al., 2018; but see Massey and Hutchings, 2020). Considering the potential benefits of more accurately reproducing real-world thermal environments along with the logistical and empirical costs, researchers should (but rarely do) invest great effort into determining the most appropriate level of thermal complexity to use in their research.

Our goal is to determine the effects of using various levels of thermal complexity in studies of developmental plasticity and how these effects change under different contexts (e.g. seasonal temperature differences). We used temperature data collected from lizard nests across a broad reproductive season to create three commonly used incubation methods and compared their effects on development with those of naturally fluctuating temperatures. We replicated each treatment across two seasonal temperature regimes: relatively cool versus relatively warm temperatures that typify nests early and late in the season, respectively. We incubated eggs under these conditions and measured a host of embryo and hatchling phenotypes. This unique two (seasons) by four (incubation methods) full-factorial design allows us to assess the context-dependent effects of incubation methods that vary substantially in thermal complexity. Our work evaluates the importance of replicating real-world temperatures in studies of thermal ecology and evolution. Because many taxa exhibit thermal developmental plasticity and inhabit thermally complex environments, our results have implications that span various evolutionary and ecological scales.

MATERIALS AND METHODS
Lizard collection and husbandry
The brown anole (*Anolis sagrei* Duméril and Bibron 1837) is a subtropical lizard native to islands in the Caribbean, but is naturalized in Florida, USA (Fig. 1). Females lay a single-egg clutch once every 4–12 days from March to October (Hall et al., 2020). Eggs are deposited in shallow nests (1–4 cm deep) or laid on the soil surface beneath leaf litter or cover objects (e.g. rocks, logs) across a diversity of habitats (Pruett et al., 2020); thus, embryos are subjected to a wide range of mean temperatures and thermal fluctuations during development (Pearson and Warner, 2018; Hall and Warner, 2020). Moreover, nest temperatures are substantially cooler for eggs laid earlier (e.g. March–April) versus later (e.g. July–August), resulting in season-specific effects of nest temperature on egg survival and hatchling phenotypes (Pearson and Warner, 2018).

We captured adult male [49–65 mm snout–vent length (SVL); n=25] and female (42–52 mm SVL; n=100) *A. sagrei* during 22–23 June 2018 from Palm Coast, FL (29.602199, –81.196211). Lizards were transported to Auburn University to form a captive breeding colony. We housed females individually, in plastic cages (29×26×39 cm) illuminated with Reptisun 5.0 UVB and Tropic Sun 5500 K daylight bulbs (Zoo Med Inc., San Luis Obispo, CA, USA) set to a 14 h:10 h light:dark cycle. Mean cage temperature was 27.6°C and, owing to light sources, oscillated daily from 26 to 30°C. Cages included reptile cage carpet (Zoo Med Inc.) as a substrate, two bamboo perches, an artificial plant and a nesting pot (i.e. a plant pot filled with potting soil and peat moss and covered with an artificial leaf). Males were rotated between the same four cages of females every 3 days for the first 2 weeks and then every
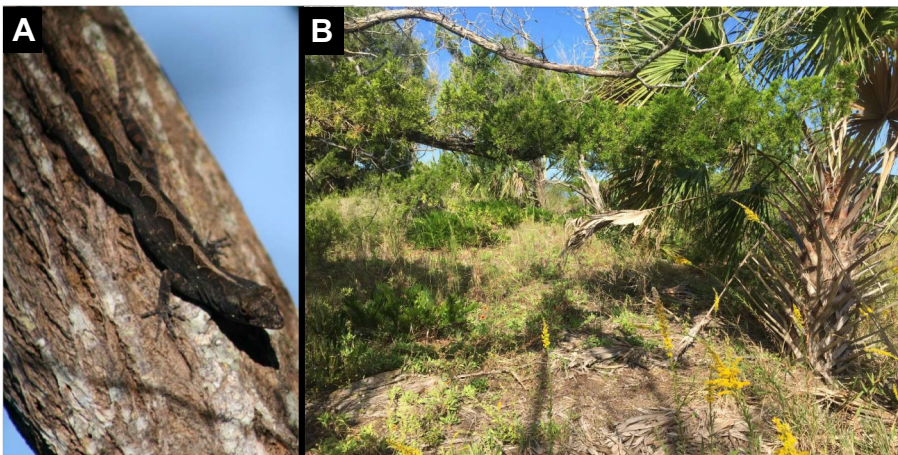
Fig. 1. Example photos of the study species and habitat. A female brown anole (A) and an example of typical brown anole habitat at our field site in Florida, USA (B).

2 weeks thereafter to ensure insemination. However, because we captured females in the middle of the breeding season and they can store sperm for many weeks, all females were likely gravid at time of capture. We fed lizards four crickets each, dusted with vitamins and calcium, twice per week and misted cages with water daily.

Research was approved by the Auburn University Institutional Animal Care and Use Committee (no. 2018-3233).

Egg collection and treatment allocation

We collected eggs (n=415) daily from 25 June to 5 August, recorded their mass (to 0.0001 g), date of oviposition, and maternal identity, and randomly distributed them among eight incubation treatments (see below). Eggs were placed individually in Petri dishes (60×15 mm) that were half-filled with moist vermiculite (−150 kPa) and wrapped with parafilm to prevent desiccation and control moisture across treatments. To evenly distribute each female's eggs across treatments, we randomly allocated her first egg to a treatment and then randomly allocated each additional egg to a remaining treatment. No female had more than one egg per treatment. Randomizations were conducted using the RAND function in Excel.

Creation of incubation treatments

We used nest temperatures from Pearson and Warner (2018) to create our incubation regimes. Briefly, 22 Thermochron iButtons were placed in various nesting microhabitats at our field site and recorded temperatures every 2.5 h from 28 March to 15 October 2013. We averaged each 2.5 h timepoint across iButtons to create two naturally fluctuating regimes: one cooler regime used temperatures from 28 March to 5 May (henceforth 'early season') and one warmer regime used temperatures from 27 June to 4 August (henceforth 'late season'). These season-specific incubation temperatures induce variation in egg and hatchling phenotypes (Pearson and Warner, 2018). We refer to these henceforth as the 'natural' treatments. Importantly, each natural treatment represents only one potential replication of field temperatures. Indeed, the 22 nests studied exhibit a wide diversity in both mean and thermal variance (Fig. S1). Natural nest fluctuations are not symmetrical (i.e. not a sine curve); thus, eggs typically spend more time each day at minimum than maximum temperatures. Thus, our natural regimes better mimic minimum versus maximum nest temperatures (see Fig. S1). Accurately representing both mean and variance of nest temperature in laboratory studies is notoriously challenging (Georges, 1989; Georges et al., 2004) and is a common limitation. Regardless, our natural treatments have similar mean temperatures to field nests and include a range of temperatures that represent the majority of those recorded from the field (Fig. S1). Moreover, unlike commonly used incubation treatments (described below), our natural treatments allow temperatures to fluctuate widely and stochastically, like real nests.

We used the mean temperatures and variance from these natural treatments to create six additional treatments: two constant temperature treatments, two sine fluctuations and two hourly means fluctuations. These treatments represent commonly used incubation methods in studies of reptile developmental plasticity (While et al., 2018; Booth, 2018). The constant temperature treatments were the raw average of the early and late season natural treatments (Fig. 2A). For the sine wave, we created a daily repeating sine fluctuation with an amplitude equal to the mean range of the natural regimes (2.4°C; Fig. 2B). For the repeated hourly means, we created a daily repeating fluctuation by averaging the temperature at each hour of the day across all iButtons for each season (Fig. 2C). Thus, for each season (early versus late) there were four methods, resulting in eight thermal treatments (2×4 factorial design; Fig. 2).
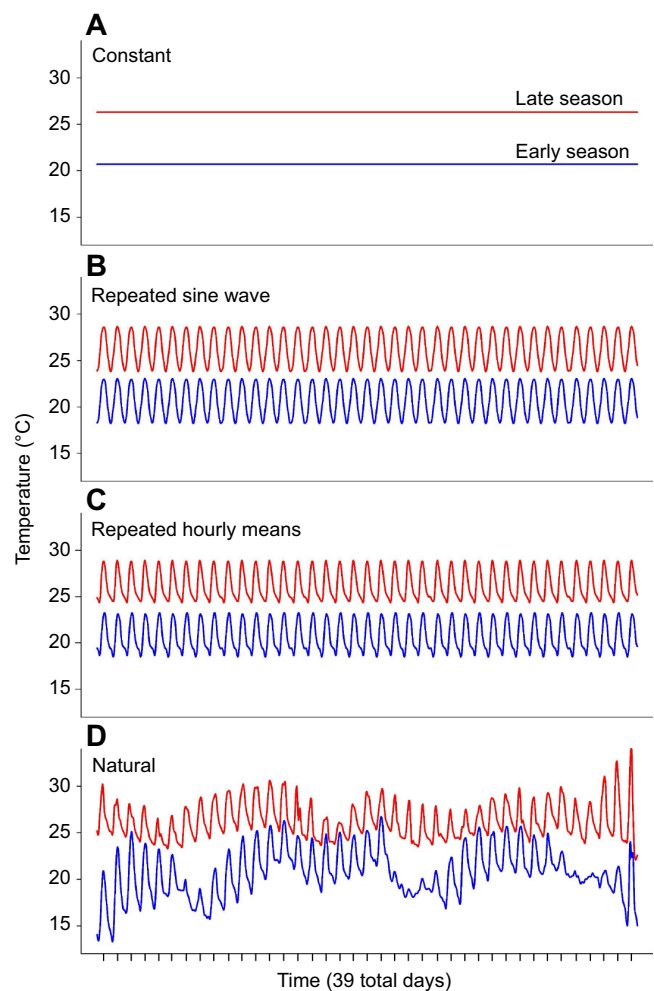
Fig. 2. Incubation treatments. (A–D) Temperatures for the constant, sine wave, hourly means and natural treatments, respectively. Blue and red lines show early- and late-season temperatures, respectively. Each rise and fall of temperature is a daily fluctuation. See Fig. 3 for comparison of panels B and C.

Treatments were programmed into eight Memmert IPP55 plus incubators (i.e. one incubator per season by treatment incubation regime). Ideally, multiple incubators would be used per season by incubation treatment (i.e. at least 16 incubators; Greenspan et al., 2016); however, this was not possible. There was, however, replication of incubators with respect to season (n=4 per season) and treatment (n=2 per treatment).

The sine and hourly means treatments were programmed to loop daily for the duration of the study. The two natural regimes (early versus late) were each 39 days long, so we programmed incubators to loop this 39-day cycle. Owing to temporal variation in egg production, not all eggs in the natural treatments experienced the same temperatures; however, this is a necessary consequence of utilizing a natural thermal regime and is what eggs experience in the wild. Regardless, we estimated all possible mean temperatures that eggs could experience in the natural regimes by using the average incubation period for each treatment (see Results) and calculating the mean temperature that eggs would experience if they were oviposited on each of the 39 days. Mean temperatures ranged from 20.5 to 20.8°C (average=20.7°C) and 26.2 to 26.4°C (average=26.3°C) for early- and late-season natural treatments, respectively. Therefore, differences in mean incubation temperature owing to temporal variation in egg production were slight.

All early-season treatments had the same mean temperature (20.7°C) and all late-season treatments had the same mean temperature (26.3°C). The sine and hourly mean fluctuations treatments had the same daily temperature range (4.8°C), which was equal to the mean daily range of temperatures for both the early- and late-season natural treatments; however, maximum and minimum daily temperatures and the number of hours spent above and below the mean temperature differed (Fig. 3). Thus, within seasons, all treatments had the same means, and all fluctuating treatments had the same amount of daily thermal variation but differed according to daily maximum and minimum temperatures. The thermal characteristics used here are important for ectotherm development (Georges et al., 2005), and they allowed critical tests of season by method interactions on offspring traits (i.e. context-dependent effects of methodology).

Developmental variables

We measured developmental rate, water uptake, oxygen consumption ($\dot{V}_{O_2}$) and heart rate (f_H) to assess effects on embryo physiology. We calculated developmental rate by taking the reciprocal of the incubation period (i.e. number of days from oviposition to hatching) and multiplying by 1000 (Viets et al., 1993). For 12 eggs per treatment, we evaluated water uptake by subtracting egg mass at oviposition from egg mass at approximately 65% of development (estimated from Pearson and Warner, 2018). Water uptake during incubation is vital for successful development and has fitness-relevant effects on hatchling phenotypes (Janzen, 1993).

For another 12 eggs per treatment (n=96 total), we measured $\dot{V}_{O_2}$ at two temperatures (mean seasonal temperatures: 20.7 and 26.3°C) at ~72% development completed. We used dynamic injection to measure $\dot{V}_{O_2}$ of eggs using a Qubit Q-box RP1LP respirometer (Qubit Biology Inc., Kingston, ON; see Lighton, 2008, pp. 29–31). Each egg was placed in a 10 ml syringe with Luer-Lok tip attached to a three-way stopcock (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). A 5 µl drop of tap water was placed inside the syringe via a micropipette to prevent desiccation. The plunger was inserted so the volume of the chamber was 6 ml. The stopcock was attached to the respirometer, and the syringe was flushed with CO_2-free room air for 2 min at a rate of 100 ml min^{-1}. Two previously drilled holes (between the 5- and 6-ml marks) allowed air to exit the syringe (see fig. 4.3 in Lighton, 2008). After flushing, the syringe plunger was moved down to the 4 ml mark and the stopcock was twisted, sealing the egg in the syringe in a 4 ml

volume of air (minus the volume of the egg and the drop of water). The time was noted, and eggs were placed in a constant temperature incubator set to the target temperature (either 20.7 or 26.3°C) for 3 h. At the end of 3 h, the syringe was removed from the incubator, a needle was affixed to the stopcock and a 2 ml sample of air was injected into an injection port in the respirometer. The sample bolus was injected into a stream of dry, CO_2-free air flowing at 50 ml min^{-1}. The time of injection was noted so we could calculate the exact length of time (in seconds) that each egg was respiring. Control syringes (containing a drop of water but no egg) were treated identically to those described previously and were injected into the respirometer at the beginning and end of the experiment. Calculations for oxygen consumption were made according to Lighton (2008) and corrected for standard temperature and pressure.

All eggs produced from 30 July to 5 August were used to assess embryo f_H after completing approximately 70% of development (n=80 eggs; 10 per treatment). For each egg, we measured embryo f_H at 23°C using the Buddy® egg monitoring system (Hulbert et al., 2017). Because temperatures fluctuated by 0.5°C, we used a thermocouple to measure the temperature inside the heart rate monitor for each measure of f_H.

Post-hatching variables

To assess effects on morphology, we measured the SVL, body mass, body condition and tail length of every hatchling (n=395). Body condition was each hatchling's residual score from a regression of log mass and log SVL. Thus, higher residuals represent individuals that are relatively heavy for their body length. To assess effects on performance, we measured sprint speed and endurance for approximately 35 hatchlings per treatment (sprint speed – range: n=31–43 per treatment, total sample size: n=287; endurance – range: n=30–47 per treatment, total sample size: n=298) with a circular racetrack. The racetrack consisted of two circles of aluminum flashing on top of a wooden panel. The outer and inner pieces of flashing were 1.35 and 1.06 m in circumference, respectively. The racetrack was heated from underneath to keep it within the range of preferred body temperatures for *A. sagrei*. We used a thermocouple to measure the surface of the racetrack during each trial (range=28.8–33.1°C; mean=31.2°C). Lizards were raced within 2 days of hatching: after hatching, each lizard was placed in a cage (see below) and misted with water. On the following day, lizards were raced. No food was provided prior to racing. For racing, hatchlings were placed individually in 50 ml centrifuge tubes
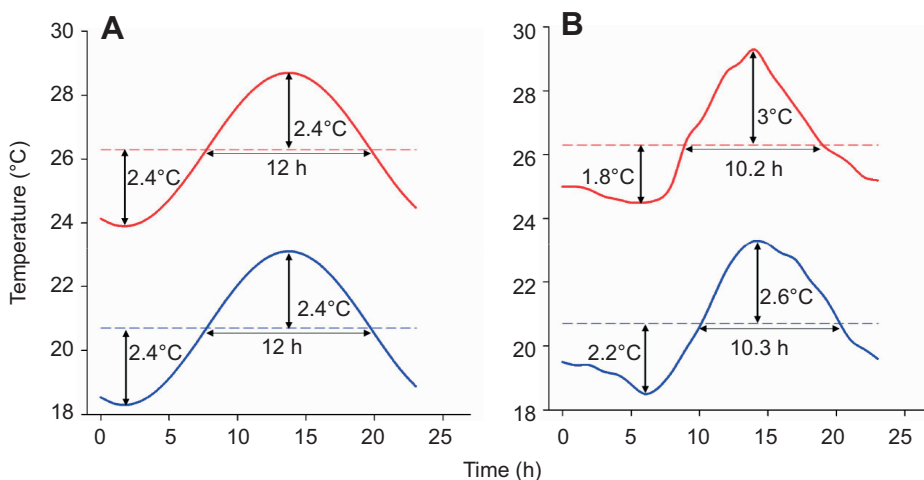
Fig. 3. Comparison between sine wave and hourly means temperatures. (A) Sine waves, (B) hourly means. Blue and red lines show the early-and late-season temperatures, respectively. Temperature values show the difference between the maximum and minimum temperatures of the fluctuation and the mean temperature (i.e. broken lines). Time values show the total time the fluctuation was warmer than the mean temperature.

wrapped with duct tape to prevent them from seeing out (to reduce stress). Tubes were placed on the racetrack for 45 min to allow the lizard's body temperature to equilibrate with the racetrack. Lizards were quickly placed on the racetrack, one at a time, and chased in a circle using a paintbrush until exhaustion. Exhaustion was the point at which we tapped the lizard on the base of the tail 11 times without it making any movement. All racing trials were conducted by J.M.H. We recorded the total distance each lizard ran and the total time for each trial. To calculate total distance, we considered the distance of one lap to be the circumference of the center of the raceway (1.21 m). Each hatchling was raced once. The total distance each hatchling ran was our measure of endurance. Sprint speed for each lizard was calculated as the distance of the first lap divided by the time it took to complete that lap.

We measured hatchling growth and survival in the laboratory over 36 days for 25 hatchlings per treatment. Starting on 6 August, we kept every hatchling until we reached our target sample size for each treatment. Hatchlings were raced and then housed individually in plastic cages (13×21×17 cm). Each cage contained reptile carpet (Zoo Med Inc.) as a substrate, several plastic leaves and one wooden perch. Lighting sources, light cycles and cage temperatures were as described for adults. We provided each lizard 20 fruit flies dusted with vitamins and calcium twice weekly and misted cages with water daily. At the end of 36 days, we measured the SVL of survivors (i.e. final SVL).

All egg and hatchling phenotypes were selected because (1) they are commonly measured phenotypes in studies of reptile developmental plasticity and (2) they are sometimes correlated with fitness (Booth, 2018; While et al., 2018).

Statistical analysis

All analyses were performed in R (https://www.r-project.org/). Owing to variation in egg production among females, the number of females that contributed eggs to each analysis ranged from 63 to 95 (mean 84.8), with some females contributing only one egg and others contributing as many as seven. Therefore, to assess the potential for maternal effects to influence results, for each response variable, we compared two models with a likelihood ratio test: one that included maternal ID as a random effect and one that did not (Table 1). In all models, the fixed effects were season, treatment and their interaction. For some analyses, we included appropriate covariates (see Results). Once the best model was selected, we

dropped covariates and interaction terms that were not statistically significant ($P>0.05$). Because developmental rates differed between seasons and treatments (see Results) and because eggs often varied in age at time of measurements owing to the nature of egg production (i.e. single-egg clutches), we calculated a developmental age to use as a covariate for some analyses (see Results). To calculate developmental age, we divided the age (days since oviposition) at the time an egg was measured by the mean incubation period (in days) for its treatment and then multiplied by 100.

For egg and hatchling survival, we conducted generalized linear models with a binomial distribution using the stats package in R (i.e. glm function) and mixed effects models using the lme4 package (i.e. glmer function; https://CRAN.R-project.org/package=lme4). For developmental and hatchling variables, we conducted general linear models using the stats package in R (i.e. lm function) and linear mixed models using the lme4 package (i.e. lmer function). For mixed models, maternal ID was a random effect, and we used the lmerTest package (Kuznetsova et al., 2017) to assess statistical significance of fixed effects by calculating denominator degrees of freedom via the Satterthwaite approximation. We used the emmeans package (https://CRAN.R-project.org/package=emmeans) to make *post hoc* comparisons and adjusted *P*-values via false discovery rate correction. Assumptions of statistical tests were evaluated by visually inspecting model residuals. We log-transformed developmental rate and hatchling endurance to reduce heteroscedasticity and normalize residuals.

RESULTS
Developmental variables

Maternal ID improved model fit for developmental rate and $\dot{V}_{O_2}$ consumption at 26.3°C (Table 1). In general, the effects of treatment were minor compared with the seasonal effect and only developmental rate was influenced by the interaction (Table 2, Fig. 4). Egg survival and water uptake were affected by seasonal temperatures: eggs incubated at early-season temperatures were 2.99 (1.06–8.44; 95% confidence limits) times more likely to die and absorbed 93.77 mg (±10.6 mg s.e.m.) more water than late-season incubated eggs (Tables 2 and 3, Fig. 4A,B). Season and treatment interactively influenced developmental rate (Table 2): at early-season temperatures, eggs in the constant treatment had slower developmental rates than all fluctuating treatments, and the natural treatment increased developmental rate compared with hourly

Table 1. Comparisons of models that did or did not include maternal ID as a random effect

Response	Without maternal ID			With maternal ID			χ^2	d.f.	P
	k	AIC	logLik	k	AIC	logLik			
Egg survival	**8**	**160.9**	**−72.5**	9	162.9	−72.5	0	1	0.99
Water uptake	**10**	**1031.9**	**−505.9**	11	1033.3	−505.7	0.6	1	0.43
Developmental rate	9	−1407.8	712.9	**10**	**−1435.9**	**727.9**	30.1	1	<0.0001
Heart rate	**10**	**515.6**	**−247.8**	11	517.2	−247.6	0.4	1	0.54
O₂ consumption (20.7°C)	**10**	**495.4**	**−237.7**	11	495.3	−236.6	2.1	1	0.14
O₂ consumption (26.3°C)	10	551.6	−265.8	**11**	**547.7**	**−262.8**	5.9	1	0.02
Initial SVL	10	723.6	−351.8	**11**	**718.9**	**−348.4**	6.8	1	0.009
Initial mass	10	3121.1	−1550.5	**11**	**3109.9**	**−1543.9**	13.1	1	0.0003
Initial body condition	10	−1043.6	531.8	**11**	**−1056.9**	**539.4**	15.3	1	0.0001
Initial tail length	10	1563.2	−771.6	**11**	**1544.6**	**−761.3**	20.6	1	<0.0001
Sprint speed	**10**	**−637.9**	**328.9**	11	−635.9	328.9	0	1	1
Endurance	**10**	**275.7**	**−127.8**	11	275.8	−126.9	1.9	1	0.17
Final SVL	**10**	**313.9**	**−146.9**	11	315.7	−146.9	0.1	1	0.71
Hatchling survival	**8**	**189.1**	**−86.6**	9	191.1	−86.6	0	1	1

Bold type denotes the model we selected based on a likelihood ratio test.
k, number of model parameters; AIC, Akaike information criterion; logLik, loglikelihood; SVL, snout–vent length.

Table 2. Results of season, treatment and their interaction on developmental and post-hatching variables

Response variable	Covariate	Covariate results	Season	Treatment	Season×treatment
Egg survival[b]	Egg mass	–	χ^2_1=4.89; **P=0.027**	χ^2_3=6.17; P=0.10	–
Water uptake[b]	Developmental age	$F_{1,90}$=37.01; **P<0.0001**	$F_{1,90}$=76.54; **P<0.0001**	$F_{3,90}$=0.29; P=0.83	–
Developmental rate[a,*]	Egg mass	–	$F_{1,323}$=41,721.99; **P<0.0001**	$F_{3,328}$=34.04; **P<0.0001**	$F_{3,326}$=21.93; **P<0.0001**
Heart rate[b]	Temperature	$F_{1,73}$=10.44; **P=0.002**	$F_{1,73}$=0.16; P=0.69	$F_{3,73}$=2.30; P=0.08	–
O$_2$ consumption (20.7°C)[b]	Developmental age	$F_{1,87}$=25.73; **P<0.0001**	$F_{1,87}$=0.03; P=0.39	$F_{3,87}$=1.08; P=0.22	–
O$_2$ consumption (26.3°C)[a]	Developmental age	$F_{1,81}$=43.6; **P<0.0001**	$F_{1,64}$=1.80; P=0.18	$F_{3,65}$=3.60; **P=0.02**	–
Initial SVL[a]	Egg mass	$F_{1,348}$=172.99; **P<0.0001**	$F_{1,342}$=41.27; **P<0.0001**	$F_{3,348}$=2.64; **P=0.049**	–
Body mass[a]	Egg mass	$F_{1,364}$=453.85; **P<0.0001**	$F_{1,335}$=26.10; **P<0.0001**	$F_{3,341}$=1.58; P=0.19	–
Body condition[a]	Egg mass	$F_{1,367}$=152.42; **P<0.0001**	$F_{1,331}$=131.42; **P<0.0001**	$F_{3,337}$=3.49; **P=0.016**	–
Tail length[a]	Egg mass	$F_{1,373}$=51.92; **P<0.0001**	$F_{1,336}$=101.51; **P<0.0001**	$F_{3,342}$=1.11; P=0.35	–
Sprint speed[b]	Initial SVL	$F_{1,281}$=9.16; **P=0.003**	$F_{1,281}$=33.96; **P<0.0001**	$F_{3,281}$=1.80; P=0.15	–
Endurance[b,*]	Initial SVL	$F_{1,289}$=10.21; **P=0.002**	$F_{1,289}$=70.53; **P<0.0001**	$F_{3,289}$=4.49; **P=0.004**	$F_{3,289}$=3.27; **P=0.021**
Final SVL[b]	Initial SVL	$F_{1,148}$=37.81; **P<0.0001**	$F_{1,148}$=0.25; P=0.62	$F_{3,148}$=0.36; P=0.78	–
Hatchling survival[b]			χ^2_1=27.43; **P<0.0001**	χ^2_3=17.61; **P=0.0005**	–

For covariates, egg mass is mass at oviposition, temperature is the temperature inside the heart rate monitor, and initial SVL is SVL at time of hatching. See Materials and Methods for a description of developmental age. See Table S1 for covariate effect sizes. A dash indicates that a covariate or interaction was removed from the model owing to a lack of statistical significance. Bold text denotes statistical significance (P<0.05). Asterisks indicate the response variable was log-transformed; superscripts a and b denote models that did and did not include maternal ID as a random effect, respectively.

means (Table 3, Fig. 4C). At late-season temperatures, the difference between the constant temperature treatment and natural treatment was not statistically significant, but both resulted in slower developmental rates than the sine and hourly means treatments (Table 3, Fig. 4C). There were no statistically significant effects of treatment, season or their interaction on embryo f_H or $\dot{V}_{O_2}$ at 20.7°C (Table 2, Fig. 4D,E); however, we did observe a treatment effect on $\dot{V}_{O_2}$ at 26.3°C (Table 2). Eggs incubated at constant temperatures had greater rates of oxygen consumption than all other treatments; however, the difference between the constant and sine treatments was marginally not significant (P=0.05; Table 3, Fig. 4F). See Table S3 for sample sizes, raw means and standard deviations of each developmental variable.

Post-hatching variables

Maternal ID improved model fit for all measures of hatchling morphology but not for performance, growth or survival (Table 1). In general, the effects of treatment were minor compared with seasonal effects, and only hatchling endurance was influenced by the interaction (Table 2, Figs 5 and 6). Late-season temperatures produced shorter (Fig. 5A) but heavier offspring (Fig. 5B,C), with longer tails (Fig. 5D) compared with early-season temperatures (Table 4). Initial analysis revealed that SVL was significantly affected by treatment (Table 2); however, this effect was not statistically significant after P-values were corrected for *post hoc* comparisons (Table 4). Body condition was the only morphological variable that was clearly affected by treatment (Table 2). Hatchlings from the hourly means treatment had greater body condition than those from the sine wave and constant temperature treatments, but no other differences were statistically significant (Table 4, Fig. 5C).

Sprint speed was greater for hatchlings from late-season temperatures (Table 4, Fig. 6A), but neither treatment nor the interaction were statistically significant (Table 2). We observed an interaction for hatchling endurance: at early-season temperatures, hatchlings from the natural treatment had greater endurance than those from the constant temperature and sine treatments, and those from the hourly means treatment had greater endurance than those from the constant treatment; however, there were no significant treatment effects for late-season incubated eggs (Table 4, Fig. 6B). There were no statistically significant effects of season, treatment or their interaction on hatchling final SVL, indicating that growth did not differ among treatments (Table 2, Fig. 6C).

We observed season and treatment effects on hatchling survival, but the interaction was not statistically significant (P=0.42; Table 2). Hatchlings from early-season incubation temperatures were 7.54 (3.24–17.5; 95% confidence limits) times as likely to die as those from late-season temperatures (Table 4); however, this was primarily driven by high mortality in the early-season constant, sine and hourly means treatments (Fig. 6D). Hatchlings from the natural and hourly means treatments were 8.70 (1.9–39.93; 95% confidence limits) and 4.15 (1.11–15.51; 95% confidence limits) times more likely to survive than those from the constant temperature treatment, respectively, but no other differences were statistically clear (Table 4, Fig. 6D). See Table S3 for sample sizes, raw means and standard deviations of each post-hatching variable.

DISCUSSION

Several studies have compared the effects of constant versus repeated temperature fluctuations on development [e.g. birds (Olson et al., 2006), lizards (Du and Shine, 2010), frogs (Niehaus et al., 2006; Arrighi et al., 2013), fish (Eme et al., 2018); insects (Liu et al., 1995; Folguera et al., 2011; Xing et al., 2019)]. A few studies have measured or modeled the effects of natural thermal regimes on biological processes (e.g. Niehaus et al., 2012; Bannerman and Roitberg, 2014) or tested differences in constant temperature, repeated daily fluctuations and random changes in diurnal fluctuations (e.g. Schaefer and Ryan, 2006; Manenti et al., 2014). To our knowledge, however, no study of developmental plasticity has compared constant temperature and repeated fluctuations with natural nest fluctuations under different contexts (e.g. season) in the laboratory. Some have compared eggs incubated in the laboratory with eggs incubated in the field (e.g. St Juliana and Janzen, 2007; Paitz et al., 2010); however, these studies cannot pinpoint thermal effects because there are many other variables in natural nests that influence development (Warner et al., 2018).

Developmental variables

Seasonal effects on development were typical: warmer temperatures increase rates of survival and development and reduce total water uptake (Gutzke and Packard, 1987; Pearson and Warner, 2018). Water uptake is greater for eggs incubated at cool temperatures, likely because they have a longer incubation period and more time to absorb water (Booth et al., 2000). The absence of a season effect on $\dot{V}_{O_2}$ and f_H indicates there was no thermal acclimation of embryo
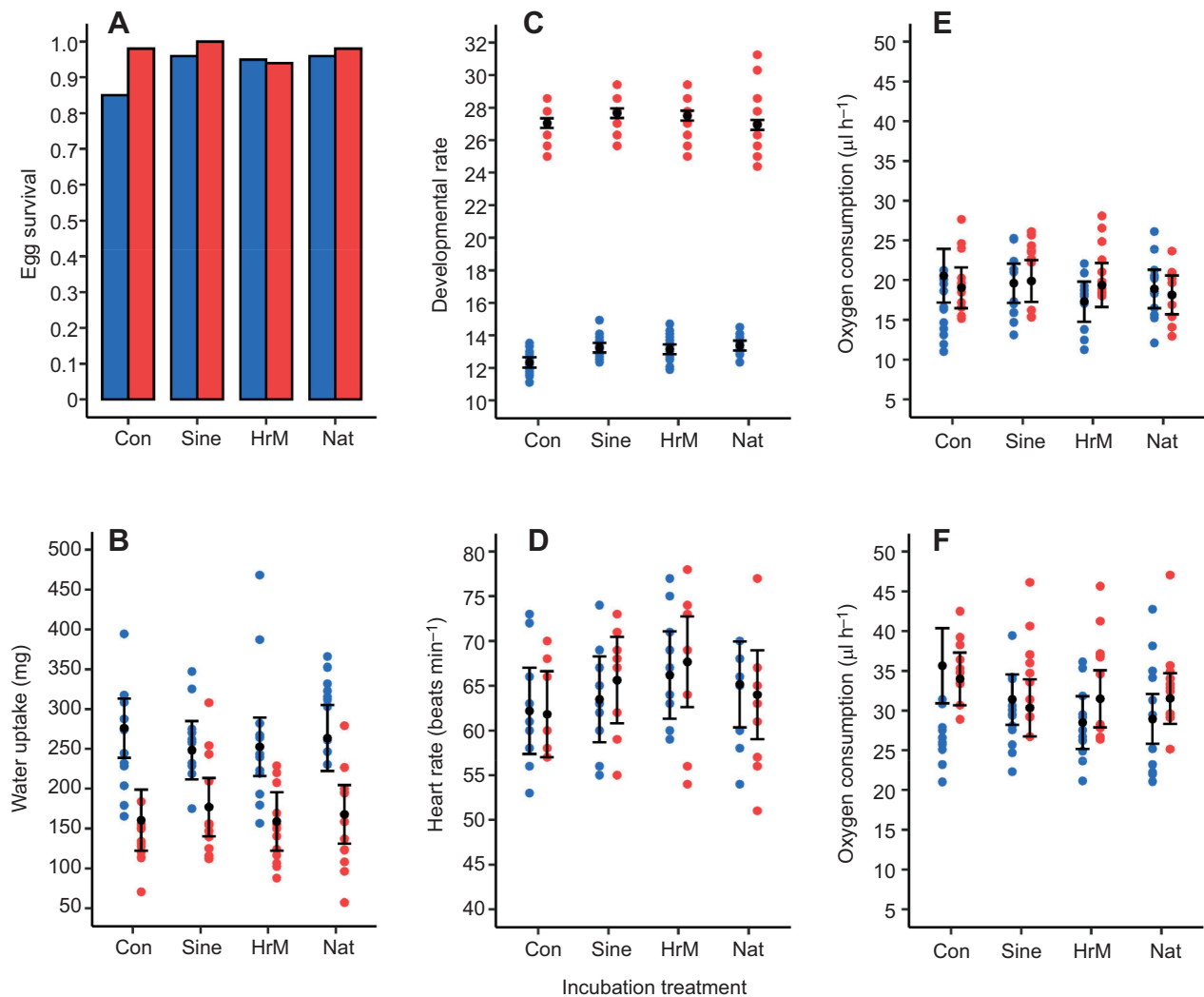
Fig. 4. Effects of treatments on developmental variables. (A–D) Treatment effects on egg survival, egg water uptake, embryo developmental rate and embryo heart rate, respectively. E and F show oxygen consumption rates at 20.7 and 26.3°C, respectively. Blue and red bars or circles represent early- and late-season temperatures, respectively. Colored circles show raw data, black circles show estimated marginal means (via package 'emmeans'), and vertical bars show 95% confidence intervals. Con, constant; Sine, sine wave; HrM, hourly means; Nat, natural. See Tables 2 and 3 for statistical tests and Table S3 for sample sizes.

physiology, which is typical for shallow-nesting species (such as anoles) but not for species that nest deeper in the soil (Du et al., 2010). The effects of incubation treatment were minimal for most egg phenotypes, but some fitness-relevant variables were affected (e.g. developmental rate). Moreover, treatment effects were small compared with those of seasonal temperature. In studies of thermal developmental plasticity, the largest effects of temperature are on developmental rates (Stillwell and Fox, 2005; Noble et al., 2018), and our data indicate this is also true for incubation methods. Developmental rates can have important effects on fitness, given that early hatching is often associated with increased hatchling survival (Pearson and Warner, 2018) and faster development reduces embryo exposure to adverse conditions (Doody, 2011). At cooler temperatures, the natural treatment increased developmental rates relative to other treatments; however, at warmer temperatures, the natural treatment somewhat slowed development relative to other treatments. This likely results from the curvilinear relationship between developmental rates and temperature (see fig. 3 in Bowden et al., 2014): at early-season temperatures, relatively small reductions in developmental rate owing to cold temperatures are offset by relatively greater increases at warmer temperatures and vice versa for

the late-season treatment. This is particularly important for species with temperature-dependent sex determination (TSD), where the proportion of development spent at a given temperature determines sex (Georges et al., 1994). Thus, slight differences in incubation methods could influence sex ratios (Neuwald and Valenzuela, 2011; Bowden et al., 2014) and fitness (Warner and Shine, 2011). Studies that compare the different incubation methods used here to examine sex ratios in species with TSD would provide further insight into the ecological relevance of the effects of laboratory incubation on this important demographic trait.

Intriguingly, constant temperature treatments resulted in greater metabolic rates than other treatments. There are biological and methodological explanations for this result. Embryo metabolic acclimation to temperature is widespread across reptiles and likely serves to compensate for environmental conditions that reduce development rates (e.g. cool temperatures; Du et al., 2010). Constant temperatures result in relatively slow development compared with fluctuating conditions. Thus, embryos may compensate by upregulating metabolism when temperatures are thermostable. We think this biological explanation is unlikely because we did not observe metabolic acclimation with respect to

Table 3. Pairwise comparisons and log odds ratios for developmental variables

Response	Season	Contrast/ratio	Estimate	s.e.	d.f.	Statistic	P
Survival (odds ratio)	Early/Late	**Early/Late**	**0.33**	**0.18**	**Inf**	**−2.07**	**0.038**
Water uptake (mg)	Early–Late	**Early–Late**	**93.89**	**10.73**	**90**	**8.75**	**<0.0001**
Developmental rate*	Early	**Con–HrM**	**−0.064**	**0.007**	**336.4**	**−8.571**	**<0.0001**
	Early	**Con–Nat**	**−0.081**	**0.007**	**325.0**	**−10.831**	**<0.0001**
	Early	**Con–Sine**	**−0.071**	**0.007**	**322.56**	**−9.766**	**<0.0001**
	Early	**HrM–Nat**	**−0.017**	**0.007**	**322.13**	**−2.376**	**0.027**
	Early	HrM–Sine	−0.008	0.007	324.08	−1.105	0.27
	Early	Nat–Sine	0.009	0.007	320.29	1.284	0.24
	Late	**Con–HrM**	**−0.018**	**0.007**	**331.01**	**−2.404**	**0.025**
	Late	Con–Nat	0.004	0.007	336.31	0.577	0.56
	Late	**Con–Sine**	**−0.024**	**0.007**	**320.44**	**−3.306**	**0.003**
	Late	**HrM–Nat**	**0.022**	**0.007**	**330.18**	**2.97**	**0.006**
	Late	HrM–Sine	−0.006	0.007	330.01	−0.82	0.50
	Late	**Nat–Sine**	**−0.028**	**0.007**	**330.29**	**−3.844**	**0.001**
O$_2$ consumption at 26.3°C (μl h^{-1})	Combined	**Con–HrM**	**4.13**	**1.40**	**76.25**	**2.94**	**0.01**
	Combined	**Con–Nat**	**3.95**	**1.35**	**74.19**	**2.92**	**0.01**
	Combined	Con–Sine	3.18	1.38	74.39	2.31	0.05
	Combined	HrM–Nat	−0.18	1.24	77.04	−0.15	0.88
	Combined	HrM–Sine	−0.94	1.22	71.80	−0.78	0.60
	Combined	Nat–Sine	−0.76	1.13	49.03	−0.68	0.60

Bold text denotes statistical significance (*P*<0.05). *P*-values are adjusted via the false discovery rate method. See Table S2 for estimated marginal means of each variable. Con, constant; HrM, hourly means; Nat, natural; Statistic, *t*-statistic or *z* ratio for contrasts and log odds ratios, respectively; Inf, infinity degrees of freedom; asterisks denote that a variable was log transformed.

temperature (i.e. early versus late treatments). An alternate explanation is that embryos in the constant treatment were younger and smaller at the time $\dot{V}_{O_2}$ was measured because of treatment-specific variation in developmental rates. Because $\dot{V}_{O_2}$ scales with embryo size and, thus, age, using developmental age as a covariate may have caused the model to overestimate mean values of $\dot{V}_{O_2}$ for the constant treatment. Note that the 95% confidence interval of the estimated marginal mean $\dot{V}_{O_2}$ for the early-season constant treatment excludes nearly all the raw data in Fig. 4F. Although several studies have examined metabolic acclimation of embryos to temperature (see Du and Shine, 2015), to our knowledge, no study has explicitly considered the effects of fluctuating versus constant temperatures on acclimation.

Post-hatching variables

Seasonal effects on post-hatching variables were typical: warmer temperatures have small effects on SVL and hatchling growth but increase body mass, body condition, tail length, performance and survival in the laboratory (Pearson and Warner, 2018). Treatment effects on hatchling phenotypes were relatively small and most were not statistically clear. However, there was a relatively large effect on hatchling survival – a critical component of fitness. Although the interaction between season and treatment was not statistically significant, the natural regime vastly improved hatchling survival compared with other early-season treatments (i.e. the effect was large; Fig. 6D). This may be related, in part, to performance because at colder temperatures, hatchlings from the natural treatment had greater endurance than other groups (Table 4). Additionally, sprint speed and growth rates were slowest for hatchlings from the early-season constant temperature treatment, though these effects were not statistically significant. In general, hatchling performance and survival are lower at extreme incubation temperatures compared with intermediate temperatures, which are within the optimal thermal range (OTR; Andrews and Schwarzkopf, 2012; Noble et al., 2018). Although eggs in the early-season, natural treatment briefly experienced temperatures as low as 13°C, there were many days that temperatures peaked at 24–26°C (Fig. 2D), which is well within the

OTR for this species (21–31°C; Hall and Sun, 2020). Because developmental rates increase with temperature, in the early-season natural treatment, eggs spent a greater proportion of development at temperatures within the OTR than in other treatments, and this may have enhanced hatchling performance and survival. Evidence for this speculation is that, for early-season temperatures, the natural treatment resulted in a faster developmental rate than all other treatments (Table 3). Therefore, at colder temperatures, using a naturally fluctuating regime may optimize development compared with repeated fluctuations or constant temperatures, whereas at warmer temperatures, which are within the OTR, it may have little effect. Importantly, because hatchlings were housed singly, we measured intrinsic mortality; however, competition and predation (i.e. sources of extrinsic mortality) exert strong selective pressures on this species (Calsbeek and Cox, 2010). Therefore, incubation at cool, constant temperatures would likely result in even lower hatchling survival in the wild.

There are two important conclusions to draw from the hatchling survival results. First, these effects must have resulted from treatment-specific patterns of embryo development because the treatments were applied to eggs. Thus, conditions appropriate for successful embryo development may not necessarily be best for hatchlings (e.g. compare early-season constant egg versus hatchling survival). Consequently, researchers should assess the effects of egg environments on hatchling performance and survival. Although assessing the influence of developmental environments on later life stages will be a critical advancement (Refsnider et al., 2019), the importance of this could depend on whether research goals are focused on proximate mechanisms of thermal effects versus the fitness consequences. Second, although incubation treatments had minimal effects on egg phenotypes and hatchling morphology, they had relatively large effects on performance and survival (i.e. fitness), indicating that many commonly measured phenotypes may not be good fitness proxies. Although numerous studies demonstrate that incubation temperature influences locomotor performance in reptiles, very little work has been done to unearth the mechanisms driving these patterns (Booth, 2018). Most likely,
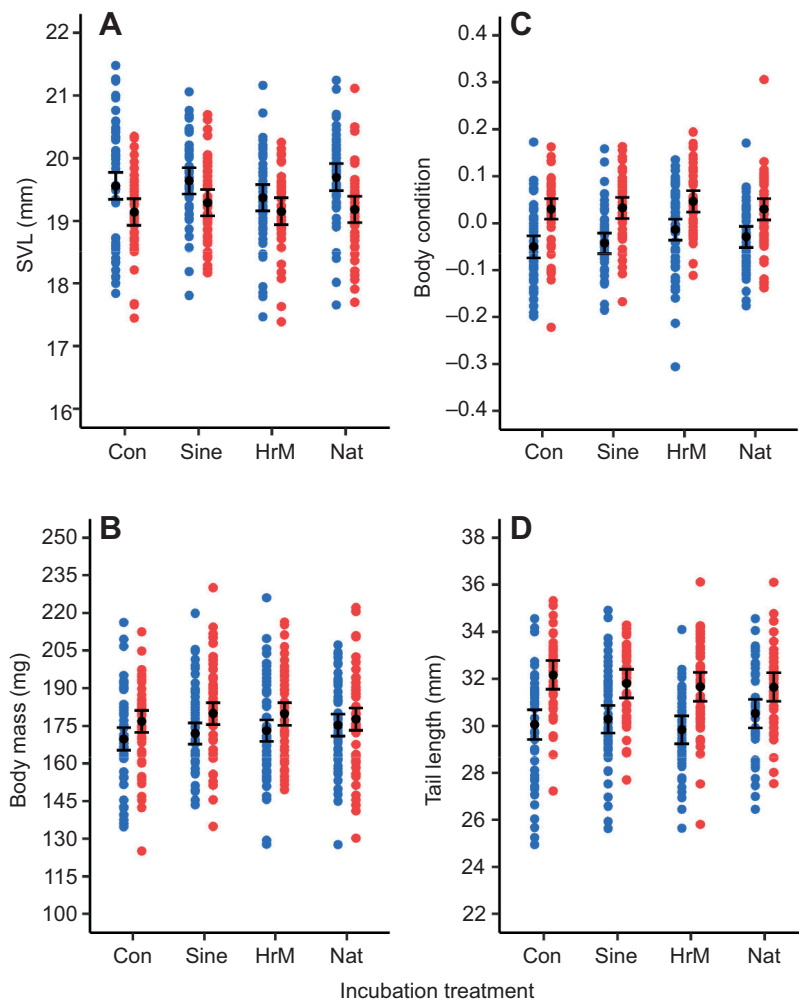
Fig. 5. Effects of treatments on hatchling morphology. (A–D) Treatment effects on hatchling SVL, body mass, body condition and tail length, respectively, at the time of hatching. Blue and red circles represent early-and late-season temperatures, respectively. Colored circles show raw data, black circles show estimated marginal means (via package 'emmeans'), and bars show 95% confidence intervals. SVL, snout–vent length. Con, constant; sine, sine wave; HrM, hourly means; Nat, natural. See Tables 2 and 4 for statistical tests and Table S3 for sample sizes.

thermal effects on performance are driven by morphological and physiological variation at cellular and subcellular levels (e.g. number and type of muscle fibers; mitochondrial density; oxygen transport; reviewed by Booth, 2017, 2018); therefore, we should not necessarily expect gross measures of hatchling morphology to coincide with performance and fitness. Our study adds a new layer of complexity by demonstrating that the mean and variance of incubation temperature can potentially interact to determine locomotor performance and survival, independent of gross morphology (e.g. body size).

Implications for experimental design and study comparisons
Most studies that incorporate fluctuating temperatures use repeated sine waves (e.g. Folguera et al., 2011; Xing et al., 2019), few use an hourly means fluctuation (e.g. Davis et al., 2006; Tiatragul et al., 2017; Hall and Warner, 2018) and almost none use a natural regime (Pearson and Warner, 2018; Tiatragul et al., 2020). The largest effects of incubation method were, generally, between constant and fluctuating treatments. For example, the effects of sine waves or hourly means were similar and neither differed substantially from the natural treatment for most phenotypes. Thus, researchers should use fluctuating rather than constant temperatures; however, the type of fluctuation may be of little importance if temperatures are within the OTR. Sine waves are particularly advantageous because they are easy to standardize and replicate. Conversely, recreating natural regimes is costly with respect to equipment and time. For example, to create our natural treatments,

we made programs with 39 different fluctuations, and, owing to limited memory in incubators, these were assimilated into six individual programs that were each manually uploaded approximately once per week throughout the study. This extra effort is only worthwhile under certain contexts (i.e. colder temperatures). If we had conducted a study using only early- and late-season constant temperatures (or repeated sine waves), we would have concluded that early-season temperatures are less conducive to proper development owing to relatively large treatment effects on hatchling survival; however, using natural fluctuations demonstrates that early and late temperatures each result in high hatchling survival (Fig. 6D).

Finally, great effort is given to describing general trends in ecological and evolutionary phenomena (e.g. meta-analyses and literature reviews on reptile developmental plasticity: Noble et al., 2018; Warner et al., 2018; Mitchell et al., 2018; While et al., 2018; Refsnider et al., 2019). Such broad syntheses of published data depend on methodological consistency across studies and repeatability of results. We found mixed support for the assumption that data are comparable across incubation methods. For example, increasing thermal complexity (from constant to natural treatments) had little influence on egg survival, but it vastly improved survival of hatchlings at colder temperatures. Owing to these context-dependent effects, researchers should exercise caution when making comparisons across studies that utilize both different methods and temperatures, and seriously consider the thermal physiology of their organism along with methodology.
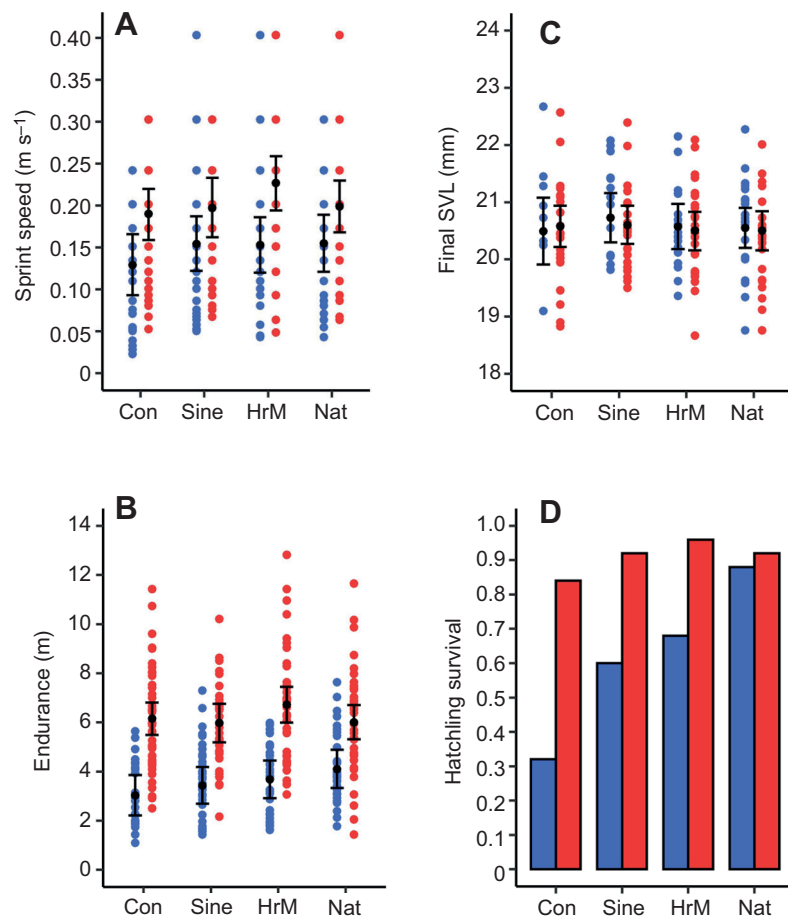
Fig. 6. Effects of treatments on post-hatching variables.
(A,B) Treatment effects on hatchling sprint speed and endurance, respectively, at the time of hatching. (C,D) Hatchling final SVL and survival, respectively, after 36 days in the laboratory. Blue and red bars or circles represent early- and late-season temperatures, respectively. Colored circles show raw data, black circles show estimated marginal means (via package 'emmeans'), and horizontal bars show 95% confidence intervals. Con, constant; sine, sine wave; HrM, hourly means; Nat, natural. See Tables 2 and 4 for statistical tests and Table S3 for sample sizes.

Biological relevance versus statistical significance

Owing to the misuse and misinterpretation of null hypothesis significance testing in ecology and evolution, researchers must always evaluate biological relevance and statistical significance (Martinez-Abrain, 2008). Our results afford opportunity to discuss this issue with respect to incubation methods. Slight differences in treatment occasionally resulted in statistically significant, but potentially biologically trivial differences among groups. For example, hatchlings from early-season incubation temperatures were longer in SVL than late-season incubated hatchlings; however, this extra length only equates to a 1.9% difference. Additionally, it did not afford greater locomotor performance to early-season incubated hatchlings, even though longer lizards, in general, run faster and have more endurance (Table S1). Moreover, past studies using similar incubation temperatures show no effect of temperature on SVL (Pearson and Warner, 2018) or the opposite effect of that observed in our study (Pearson and Warner, 2016), indicating our result may be spurious.

Conversely, treatment effects on hatchling survival were much greater at early- versus late-season temperatures, even though the interaction term was not statistically significant (Fig. 6D). Given the large effect size, ignoring this potential interaction may lead to false conclusions. Indeed, power for detecting interactions is low for binomial data even when sample sizes are relatively large (i.e. greater type II error rate; Marshall, 2007). Normally, we would not conduct *post hoc* tests in the absence of a statistically significant interaction; however, we do so here to emphasize biological relevance versus statistical significance. *Post hoc* inspection indicates that the natural and constant treatments differ for early-

but not late-season temperatures (Figs S2 and S3) and that early-season temperatures resulted in lower survival than late-season temperatures for constant, sine and hourly means treatments but not the natural treatment (Fig. S3). Therefore, despite a lack of statistical significance, the biological relevance seems clear: natural fluctuations greatly enhance offspring survival at cooler temperatures, particularly in comparison with constant temperatures; however, at warmer temperatures, thermal fluctuations are less important (Fig. 6D). Ultimately, these data indicate that considering biological relevance and statistical significance is vital when comparing results across studies that use different incubation methods, even if mean temperatures are comparable.

Limitations and caveats

One important limitation is that our study species constructs shallow nests that are prone to extreme thermal fluctuations (Hall and Warner, 2020). Many species, however, construct nests in thermostable environments and constant incubation temperatures are likely optimal for development (e.g. Andrews, 2018). Thus, our results are not generalizable across some species. Another caveat is that we used a single natural regime per seasonal temperature to represent 'natural' conditions, when there are large differences in mean and variance of temperature across natural nests (Fig. S1). Moreover, our natural regimes were created from taking mean temperatures across nests, which reduces the thermal variation present in nature. Alternate natural regimes that replicate more extreme conditions may generate different results. A future study could potentially improve upon our methods by replicating the coldest and warmest nests in addition to a nest of intermediate

Table 4. Pairwise comparisons and log odds ratios for post-hatching variables

Response	Season	Contrast/ratio	Estimate	s.e.	d.f.	Statistic	P
Initial SVL (mm)	Early–Late	**Early–Late**	**0.37**	**0.06**	**341.86**	**6.42**	**<0.0001**
	Combined	Con—HrM	0.09	0.08	356.58	1.1	0.33
	Combined	Con—Nat	−0.09	0.08	348.23	−1.08	0.33
	Combined	Con–Sine	−0.11	0.08	339.42	−1.4	0.32
	Combined	HrM–Nat	−0.18	0.08	352.38	−2.2	0.08
	Combined	HrM–Sine	−0.21	0.08	349.34	−2.53	0.07
	Combined	Nat–Sine	−0.03	0.08	342.2	−0.31	0.76
Initial body mass (mg)	Early–Late	**Early–Late**	**−6.02**	**1.18**	**335.19**	**−5.11**	**<0.0001**
Initial body condition	Early–Late	**Early–Late**	**−0.069**	**0.006**	**330.82**	**−11.46**	**<0.0001**
	Combined	**Con–HrM**	**−0.026**	**0.009**	**345.29**	**−3.02**	**0.02**
	Combined	Con–Nat	−0.01	0.009	337.04	−1.2	0.35
	Combined	Con–Sine	−0.005	0.008	328.43	−0.54	0.59
	Combined	HrM–Nat	0.016	0.009	341.79	1.84	0.13
	Combined	**HrM–Sine**	**0.021**	**0.008**	**338.44**	**2.54**	**0.04**
	Combined	Nat–Sine	0.006	0.008	330.85	0.68	0.59
Tail length (mm)	Early–Late	**Early–Late**	**−1.64**	**0.16**	**336.4**	**−10.08**	**<0.0001**
Sprint speed (m s^{-1})	Early–Late	**Early–Late**	**−0.056**	**0.01**	**281**	**−5.828**	**<0.0001**
Endurance (m)*	Early	**Con–HrM**	**−0.217**	**0.094**	**289**	**−2.313**	**0.045**
	Early	**Con–Nat**	**−0.334**	**0.095**	289	**−3.525**	**0.003**
	Early	Con–Sine	−0.130	0.093	289	−1.397	0.24
	Early	HrM–Nat	−0.117	0.091	289	−1.283	0.24
	Early	HrM–Sine	0.087	0.089	289	**0.982**	0.3
	Early	**Nat–Sine**	**0.204**	**0.089**	289	**2.292**	**0.045**
	Late	Con–HrM	−0.091	0.082	289	−1.110	0.54
	Late	Con–Nat	0.046	0.080	289	0.575	0.81
	Late	Con–Sine	0.021	0.086	289	0.239	0.81
	Late	HrM–Nat	0.137	0.084	289	1.630	0.54
	Late	HrM–Sine	0.111	0.089	289	1.246	0.54
	Late	Nat–Sine	−0.026	0.088	289	−0.291	0.81
Survival (odds ratio)	Early/Late	**Early/Late**	**0.13**	**0.06**	**Inf**	**−4.68**	**<0.0001**
	Combined	**Con/HrM**	**0.24**	**0.12**	**Inf**	**−2.77**	**0.02**
	Combined	**Con/Nat**	**0.11**	**0.07**	**Inf**	**−3.65**	**0.002**
	Combined	Con/Sine	0.36	0.18	Inf	−2.1	0.07
	Combined	HrM/Sine	1.5	0.79	Inf	0.77	0.44
	Combined	Nat/HrM	2.1	1.31	Inf	1.19	0.28
	Combined	Nat/Sine	3.14	1.9	Inf	1.9	0.09

Bold text denotes statistical significance (P<0.05). P-values are adjusted via the false discovery rate method. See Table S2 for estimated marginal means of each variable. Con, constant; HrM, hourly means; Nat, natural; Statistic, t-statistic or z ratio for contrasts and log odds ratios, respectively; Inf, infinity degrees of freedom; asterisks denote that a variable was log transformed.

temperature (rather than using mean temperatures); however, experimenters are always limited with respect to resources (e.g. incubators) and sample sizes. We emphasize that this experiment required approximately 400 eggs, which limits the number of viable study species and illustrates the difficulty of utilizing an even more complex design (e.g. multiple natural treatments per season).

Conclusions

The degree to which greater levels of treatment complexity influences results is dependent on ecological factors important to the study system, and researchers should invest time in evaluating methods in a context-dependent way. Our results are important for all researchers attempting to achieve greater levels of ecological relevance in studies of thermal ecology. Our design can determine how much thermal complexity is required to effectively reproduce real thermal environments under different conditions, thus minimizing logistical costs of future studies and improving the reliability of results. This work has several important implications for controlled laboratory or field studies of ecology and evolution. First, quantifying the effects of various methods should be done under different ecological contexts so researchers can assess the effects relative to ecological factors important to the study system. Second, depending on the context, using potentially costly, but ecologically meaningful treatments may illuminate mechanisms that drive biological phenomena or may be unnecessary. Third, the context-dependent nature of methodology should be considered when synthesizing published studies because even small differences in methods can result in biologically meaningful or, sometimes, statistically spurious differences. Finally, ecological or methodological effects may not be detectable in the stage at which they are applied (e.g. egg), but rather manifest later (e.g. hatchlings). Thus, researchers should measure a broad range of phenotypes at multiple life stages when aiming to assess the effects of developmental environments on fitness.

Data availability
Data are available through Auburn University's digital repository (AUrora): http://dx.doi.org/10.35099/aurora-8.

Supplementary information
Supplementary information available online at https://jeb.biologists.org/lookup/doi/10.1242/jeb.231902.supplemental

References

Andrews, R. M. (2018). Developmental plasticity in reptiles: Insights into thermal and maternal effects on chameleon phenotypes. *J. Exp. Zool. A Ecol. Integr. Physiol.* **329**, 298-307. doi:10.1002/jez.2160

Andrews, R. M. and Schwarzkopf, L. (2012). Thermal performance of squamate embryos with respect to climate, adult life history, and phylogeny. *Biol. J. Linn. Soc.* **106**, 851-864. doi:10.1111/j.1095-8312.2012.01901.x

Arrighi, J. M., Lencer, E. S., Jukar, A., Park, D., Phillips, P. C. and Kaplan, R. H. (2013). Daily temperature fluctuations unpredictably influence developmental rate and morphology at a critical early larval stage in a frog. *BMC Ecol.* **13**, 18. doi:10.1186/1472-6785-13-18

Bannerman, J. A. and Roitberg, B. D. (2014). Impact of extreme and fluctuating temperatures on aphid–parasitoid dynamics. *Oikos* **123**, 89-98. doi:10.1111/j.1600-0706.2013.00686.x

Booth, D. T. (2017). Influence of incubation temperature on sea turtle hatchling quality. *Integr. Zool.* **12**, 352-360. doi:10.1111/1749-4877.12255

Booth, D. T. (2018). Incubation temperature induced phenotypic plasticity in oviparous reptiles: Where to next? *J. Exp. Zool. A Ecol. Integr. Physiol.* **329**, 343-350.

Booth, D. T., Thompson, M. B. and Herring, S. (2000). How incubation temperature influences the physiology and growth of embryonic lizards. *J. Comp. Physiol. B* **170**, 269-276. doi:10.1007/s003600000097

Bowden, R. M., Carter, A. W. and Paitz, R. T. (2014). Constancy in an inconstant world: moving beyond constant temperatures in the study of reptilian incubation. *Integr. Comp. Biol.* **54**, 830-840. doi:10.1093/icb/icu016

Buckley, L. B., Arakaki, A. J., Cannistra, A. F., Kharouba, H. M. and Kingsolver, J. G. (2017). Insect development, thermal plasticity and fitness implications in changing, seasonal environments. *Integr. Comp. Biol.* **57**, 988-998. doi:10.1093/icb/icx032

Burggren, W. (2018). Developmental phenotypic plasticity helps bridge stochastic weather events associated with climate change. *J. Exp. Biol.* **221**, jeb161984. doi:10.1242/jeb.161984

Calsbeek, R. and Cox, R. M. (2010). Experimentally assessing the relative importance of predation and competition as agents of selection. *Nature* **465**, 613-616. doi:10.1038/nature09020

Carter, A. W., Sadd, B. M., Tuberville, T. D., Paitz, R. T. and Bowden, R. M. (2018). Short heatwaves during fluctuating incubation regimes produce females under temperature-dependent sex determination with implications for sex ratios in nature. *Sci. Rep.* **8**, 3. doi:10.1038/s41598-017-17708-0

Colinet, H., Sinclair, B. J., Vernon, P. and Renault, D. (2015). Insects in fluctuating thermal environments. *Annu. Rev. Entomol.* **60**, 123-140. doi:10.1146/annurev-ento-010814-021017

Davis, J. A., Radcliffe, E. B. and Ragsdale, D. W. (2006). Effects of high and fluctuating temperatures on *Myzus persicae* (Hemiptera: Aphididae). *Environ. Entomol.* **35**, 1461-1468. doi:10.1093/ee/35.6.1461

De Jong, M. A., Kesbeke, F. M. N. H., Brakefield, P. M. and Zwaan, B. J. (2010). Geographic variation in thermal plasticity of life history and wing pattern in *Bicyclus anynana*. *Clim. Res.* **43**, 91-102. doi:10.3354/cr00881

Doody, J. S. (2011). Environmentally cued hatching in reptiles. *Integr. Comp. Biol.* **51**, 49-61. doi:10.1093/icb/icr043

Du, W.-G. and Shine, R. (2010). Why do the eggs of lizards (*Bassiana duperreyi*: Scincidae) hatch sooner if incubated at fluctuating rather than constant temperatures? *Biol. J. Linn. Soc.* **101**, 642-650. doi:10.1111/j.1095-8312.2010.01525.x

Du, W.-G. and Shine, R. (2015). The behavioural and physiological strategies of bird and reptile embryos in response to unpredictable variation in nest temperature. *Biol. Rev.* **90**, 19-30. doi:10.1111/brv.12089

Du, W.-G., Ye, H., Zhao, B., Warner, D. A. and Shine, R. (2010). Thermal acclimation of heart rates in reptilian embryos. *PLoS ONE* **5**, e15308. doi:10.1371/journal.pone.0015308

Du, W.-G., Shine, R., Ma, L. and Sun, B.-J. (2019). Adaptive responses of the embryos of birds and reptiles to spatial and temporal variations in nest temperatures. *Proc. R. Soc. B* **286**, 20192078. doi:10.1098/rspb.2019.2078

Eme, J., Mueller, C. A., Lee, A. H., Melendez, C., Manzon, R. G., Somers, C. M., Boreham, D. R. and Wilson, J. Y. (2018). Daily, repeating fluctuations in embryonic incubation temperature alter metabolism and growth of Lake whitefish (*Coregonus clupeaformis*). *Comp. Biochem. Physiol. A Mol. Integr. Physi.* **226**, 49-56. doi:10.1016/j.cbpa.2018.07.027

Faye, E., Rebaudo, F., Yánez-Cajo, D., Cauvy-Fraunié, S. and Dangles, O. (2016). A toolbox for studying thermal heterogeneity across spatial scales: from unmanned aerial vehicle imagery to landscape metrics. *Methods Ecol. Evol.* **7**, 437-446. doi:10.1111/2041-210X.12488

Folguera, G., Bastías, D. A., Caers, J., Rojas, J. M., Piulachs, M.-D., Bellés, X. and Bozinovic, F. (2011). An experimental test of the role of environmental temperature variability on ectotherm molecular, physiological and life-history traits: implications for global warming. *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **159**, 242-246. doi:10.1016/j.cbpa.2011.03.002

Georges, A. (1989). Female turtles from hot nests: is it duration of incubation or proportion of development at high temperatures that matters? *Oecologia* **81**, 323-328. doi:10.1007/BF00377078

Georges, A., Limpus, C. and Stoutjesdijk, R. (1994). Hatchling sex in the marine turtle *Caretta caretta* is determined by proportion of development at a temperature, not daily duration of exposure. *J. Exp. Zool.* **270**, 432-444. doi:10.1002/jez.1402700504

Georges, A., Doody, S., Beggs, K. and Young, J. (2004). Thermal models of TSD under laboratory and field conditions. In *Temperature-dependent Sex Determination in Vertebrates* (ed. N. Valenzuela and V. Lance), pp. 79-89. Washington, DC: Smithsonian Books.

Georges, A., Beggs, K., Young, J. E. and Doody, J. S. (2005). Modelling development of reptile embryos under fluctuating temperature regimes. *Physiol. Biochem. Zool.* **78**, 18-30. doi:10.1086/425200

Greenspan, S. E., Morris, W., Warburton, R., Edwards, L., Duffy, R., Pike, D. A., Schwarzkopf, L. and Alford, R. A. (2016). Low-cost fluctuating-temperature chamber for experimental ecology. *Methods Ecol. Evol.* **7**, 1567-1574. doi:10.1111/2041-210X.12619

Gutzke, W. H. N. and Packard, G. C. (1987). Influence of the hydric and thermal environments on eggs and hatchlings of bull snakes *Pituophis melanoleucus*. *Physiol. Zool.* **60**, 9-17. doi:10.1086/physzool.60.1.30158624

Hagstrum, D. W. and Hagstrum, W. R. (1970). A simple device for producing fluctuating temperatures, with an evaluation of the ecological significance of fluctuating temperatures. *Ann. Entomol. Soc. Am.* **63**, 1385-1389. doi:10.1093/aesa/63.5.1385

Hall, J. M. and Sun, B. J. (2020). Heat tolerance of reptile embryos: current knowledge, methodological considerations, and future directions. *J. Exp. Zool. A: Ecol. Integr. Physiol.* doi:10.1002/jez.2402

Hall, J. M. and Warner, D. A. (2018). Thermal spikes from the urban heat island increase mortality and alter physiology of lizard embryos. *J. Exp. Biol.* **221**, jeb181552. doi:10.1242/jeb.181552

Hall, J. M. and Warner, D. A. (2020). Thermal sensitivity of lizard embryos indicates a mismatch between oxygen supply and demand at near-lethal temperatures. *J. Exp. Zool. A: Ecol. Integr. Physiol.* doi:10.1002/jez.2359

Hall, J. M., Mitchell, T. S., Thawley, C. J., Stroud, J. T. and Warner, D. A. (2020). Adaptive seasonal shift towards investment in fewer, larger offspring: Evidence from field and laboratory studies. *J. Anim. Ecol.* **89**, 1242-1253. doi:10.1111/1365-2656.13182

Hulbert, A. C., Mitchell, T. S., Hall, J. M., Guiffre, C. M., Douglas, D. C. and Warner, D. A. (2017). The effects of incubation temperature and experimental design on heart rates of lizard embryos. *J. Exp. Zool. A Ecol. Integr. Physiol.* **327**, 466-476. doi:10.1002/jez.2135

Janzen, F. J. (1993). An experimental analysis of natural selection on body size of hatchling turtles. *Ecology* **74**, 332-341. doi:10.2307/1939296

Kuznetsova, A., Brockhoff, P. B. and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**. doi:10.18637/jss.v082.i13

Les, H. L., Paitz, R. T. and Bowden, R. M. (2009). Living at extremes: development at the edges of viable temperature under constant and fluctuating conditions. *Physiol. Biochem. Zool.* **82**, 105-112. doi:10.1086/590263

Licht, P. and Moberly, W. R. (1965). Thermal requirements for embryonic development in the tropical lizard *Iguana iguana*. *Copeia* **1965**, 515-517. doi:10.2307/1441008

Lighton, J. R. (2008). *Measuring Metabolic Rates: A Manual for Scientists*, pp. 1-26. Oxford University Press.

Liu, S.-S., Zhang, G.-M. and Zhu, J. U. N. (1995). Influence of temperature variations on rate of development in insects: analysis of case studies from entomological literature. *Ann. Entomol. Soc. Am.* **88**, 107-119. doi:10.1093/aesa/88.2.107

Manenti, T., Sørensen, J. G., Moghadam, N. N. and Loeschcke, V. (2014). Predictability rather than amplitude of temperature fluctuations determines stress resistance in a natural population of *Drosophila simulans*. *J. Evol. Biol.* **27**, 2113-2122. doi:10.1111/jeb.12463

Marshall, S. W. (2007). Power for tests of interaction: effect of raising the Type I error rate. *Epidemiol. Perspect. Innovat.* **4**, 4. doi:10.1186/1742-5573-4-4

Martinez-Abrain, A. (2008). Statistical significance and biological relevance: a call for a more cautious interpretation of results in ecology. *Acta Oecol.* **34**, 9-11. doi:10.1016/j.actao.2008.02.004

Massey, M. D. and Hutchings, J. A. (2020). Thermal variability during ectotherm egg incubation: a synthesis and framework. *J. Exp. Zool. A: Ecol. Integr. Physiol.* doi:10.1002/jez.2400

Mickley, J. G., Moore, T. E., Schlichting, C. D., DeRobertis, A., Pfisterer, E. N. and Bagchi, R. (2019). Measuring microenvironments for global change: DIY environmental microcontroller units (EMUs). *Methods Ecol. Evol.* **10**, 578-584. doi:10.1111/2041-210X.13128

Mitchell, T. S., Maciel, J. A. and Janzen, F. J. (2015). Maternal effects influence phenotypes and survival during early life stages in an aquatic turtle. *Funct. Ecol.* **29**, 268-276. doi:10.1111/1365-2435.12315

Mitchell, T. S., Janzen, F. J. and Warner, D. A. (2018). Quantifying the effects of embryonic phenotypic plasticity on adult phenotypes in reptiles: a review of current knowledge and major gaps. *J. Exp. Zool. A Ecol. Integr. Physiol.* **329**, 203-214. doi:10.1002/jez.2187

Neuwald, J. L. and Valenzuela, N. (2011). The lesser known challenge of climate change: thermal variance and sex-reversal in vertebrates with temperature-dependent sex determination. *PLoS ONE* **6**, e18117. doi:10.1371/journal.pone.0018117

Niehaus, A. C., Wilson, R. S. and Franklin, C. E. (2006). Short-and long-term consequences of thermal variation in the larval environment of anurans. *J. Anim. Ecol.* **75**, 686-692.

Niehaus, A. C., Wilson, R. S., Storm, J. J. and Angilletta, M. J. (2012). Fall field crickets did not acclimate to simulated seasonal changes in temperature. *J. Comp. Physiol. B* **182**, 199-207. doi:10.1007/s00360-011-0611-1

Noble, D. W. A., Stenhouse, V. and Schwanz, L. E. (2018). Developmental temperatures and phenotypic plasticity in reptiles: a systematic review and meta-analysis. *Biol. Rev.* **93**, 72-97. doi:10.1111/brv.12333

Olson, C. R., Vleck, C. M. and Vleck, D. (2006). Periodic cooling of bird eggs reduces embryonic growth efficiency. *Physiol. Biochem. Zool.* **79**, 927-936. doi:10.1086/506003

Paitz, R. T., Gould, A. C., Holgersson, M. C. N. and Bowden, R. M. (2010). Temperature, phenotype, and the evolution of temperature–dependent sex determination: how do natural incubations compare to laboratory incubations? *J. Exp. Zool. B Mol. Dev. Evol.* **314B**, 86-93. doi:10.1002/jez.b.21312

Pearson, P. R. and Warner, D. A. (2016). Habitat- and season-specific temperatures affect phenotypic development of hatchling lizards. *Biol. Lett.* **12**, 20160646. doi:10.1098/rsbl.2016.0646

Pearson, P. R. and Warner, D. A. (2018). Early hatching enhances survival despite beneficial phenotypic effects of late-season developmental environments. *Proc. R. Soc. B* **285**, 20180256. doi:10.1098/rspb.2018.0256

Pruett, J. E., Fargevieille, A. and Warner, D. A. (2020). Temporal variation in maternal nest choice and its consequences for lizard embryos. *Behav. Ecol.* **31**, 902-910. doi:10.1093/beheco/araa032

Refsnider, J. M., Clifton, I. T. and Vazquez, T. K. (2019). Developmental plasticity of thermal ecology traits in reptiles: trends, potential benefits, and research needs. *J. Therm. Biol.* **84**, 74-82. doi:10.1016/j.jtherbio.2019.06.005

Schaefer, J. and Ryan, A. (2006). Developmental plasticity in the thermal tolerance of zebrafish *Danio rerio. J. Fish Biol.* **69**, 722-734. doi:10.1111/j.1095-8649.2006.01145.x

St Juliana, J. R. and Janzen, F. J. (2007). Can natural phenotypic variances be estimated reliably under homogeneous laboratory conditions? *J. Evol. Biol.* **20**, 1406-1414. doi:10.1111/j.1420-9101.2007.01343.x

Stillwell, R. C. and Fox, C. W. (2005). Complex patterns of phenotypic plasticity: interactive effects of temperature during rearing and oviposition. *Ecology* **86**, 924-934. doi:10.1890/04-0547

Telemeco, R. S., Elphick, M. J. and Shine, R. (2009). Nesting lizards (*Bassiana duperreyi*) compensate partly, but not completely, for climate change. *Ecology* **90**, 17-22. doi:10.1890/08-1452.1

Tiatragul, S., Kurniawan, A., Kolbe, J. J. and Warner, D. A. (2017). Embryos of non-native anoles are robust to urban thermal environments. *J. Thermal Biol.* 65, 119-124.

Tiatragul, S., Hall, J. M., Pavlik, N. G. and Warner, D. A. (2019). Lizard nest environments differ between suburban and forest habitats. *Biol. J. Linn. Soc.* **126**, 392-403. doi:10.1093/biolinnean/bly204

Tiatragul, S., Hall, J. M. and Warner, D. A. (2020). Nestled in the city heat: urban nesting behavior enhances embryo development of an invasive lizard. *J. Urban Ecol.* **6**, juaa001.

Viets, B. E., Tousignant, A., Ewert, M. A., Nelson, C. E. and Crews, D. (1993). Temperature-dependent sex determination in the leopard gecko, *Eublepharis macularius. J. Exp. Zool.* **265**, 679-683. doi:10.1002/jez.1402650610

Warner, D. A. and Shine, R. (2011). Interactions among thermal parameters determine offspring sex under temperature-dependent sex determination. *Proc. R. Soc. B* **278**, 256-265. doi:10.1098/rspb.2010.1040

Warner, D. A., Du, W.-G. and Georges, A. (2018). Developmental plasticity in reptiles: physiological mechanisms and ecological consequences. *J. Exp. Zool. A Ecol. Integr. Physiol.* **329**, 153-161. doi:10.1002/jez.2199

While, G. M., Noble, D. W., Uller, T., Warner, D. A., Riley, J. L., Du, W. G. and Schwanz, L. E. (2018). Patterns of developmental plasticity in response to incubation temperature in reptiles. *J. Exp. Zool. A Ecol. Integr. Physiol.* **329**, 162-176. doi:10.1002/jez.2181

Worner, S. P. (1992). Performance of phenological models under variable temperature regimes: consequences of the Kaufmann or rate summation effect. *Environ. Entomol.* **21**, 689-699. doi:10.1093/ee/21.4.689

Xing, K., Hoffmann, A. A., Zhao, F. and Ma, C.-S. (2019). Wide diurnal temperature variation inhibits larval development and adult reproduction in the diamondback moth. *J. Therm. Biol.* **84**, 8-15. doi:10.1016/j.jtherbio.2019.05.013