

Developmental temperatures and phenotypic plasticity in reptiles: a systematic review and meta-analysis

Daniel W. A. Noble^{1*}, Vaughn Stenhouse² and Lisa E. Schwanz^{1*}

¹*School of Biological, Earth and Environmental Sciences, Ecology and Evolution Research Centre, The University of New South Wales, Sydney, 2052, Australia*

²*School of Biological Sciences, Victoria University, Wellington, 6037, New Zealand*

ABSTRACT

Early environments can profoundly influence an organism in ways that persist over its life. In reptiles, early thermal environments (nest temperatures) can impact offspring phenotype and survival in important ways, yet we still lack an understanding of whether general trends exist and the magnitude of impact. Understanding these patterns is important in predicting how climate change will affect reptile populations and the role of phenotypic plasticity in buffering populations. We compiled data from 175 reptile studies to examine, and quantify, the effect of incubation temperature on phenotype and survival. Using meta-analytic approaches (standardized mean difference between incubation treatments, Hedges' g), we show that across all trait types examined there is, on average, a moderate to large magnitude of effect of incubation temperatures (absolute effect: $|g| = 0.75$). Unsurprisingly, this influence was extremely large for incubation duration, as predicted, with warmer temperatures decreasing incubation time overall ($g = -8.42$). Other trait types, including behaviour, physiology, morphology, performance, and survival experienced reduced, but still mostly moderate to large effects, with particularly strong effects on survival. Moreover, the impact of incubation temperature persisted at least one-year post-hatching, suggesting that these effects have the potential to impact fitness in the long term. The magnitude of effect increased as the change in temperature increased (e.g. 6°C versus 2°C) in almost all cases, and tended to decrease when temperatures of the treatments fluctuated around a mean temperature compared to when they were constant. The effect also depended on the mid-temperature of the comparison, but not in consistent ways, with some traits experiencing the greatest effects at extreme temperatures, while others did not. The highly heterogeneous nature of the effects we observe, along with a large amount of unexplained variability, indicates that the shape of reaction norms between phenotype and temperature, along with ecological and/or experimental factors, are important when considering general patterns. Our analyses provide new insights into the effects of incubation environments on reptile phenotype and survival and allow general, albeit coarse, predictions for taxa experiencing warming nest temperatures under climatic change.

Key words: incubation temperature, lizard, snake, turtle, crocodile, thermal plasticity, phenotypes, meta-regression, climate change, parental effects.

CONTENTS

I. Introduction	2
II. Methods	3
(1) Literature search and data collection	3
(2) Effect size calculation	5
(3) Moderator variables and predictions	5

* Address for correspondence (Tel.: +61 (0)2 9385 0034; E-mail: daniel.wa.noble@gmail.com or lisa.schwanz@gmail.com).

(a)	Order	5
(b)	Trait category	5
(c)	Temperature difference between the paired temperature treatments	5
(d)	Mid-temperature between the paired temperature treatments	5
(e)	Thermal fluctuations in temperature treatments	5
(f)	Age of measurement	6
(g)	Egg randomization strategy	6
(h)	Data type	6
(i)	Sex ratio of sample	6
(4)	Meta-analysis	7
(5)	Multi-level meta-regression models	8
(a)	Do the effects of incubation temperature differ among orders and trait categories?	8
(b)	Does the temperature difference, mid-temperature, and the fluctuation in temperature between incubation treatments impact effect size?	8
(c)	Do the effects of incubation temperature decrease with age?	8
(6)	Publication bias and methodological considerations	9
III.	Results	9
(1)	Overall impact of incubation temperature on reptile phenotypes	9
(2)	Do the effects of incubation temperature differ among orders and trait categories?	9
(3)	Does the temperature difference, mid-temperature, and the fluctuation in temperature between incubation treatments impact effect size?	10
(4)	Do the effects of incubation temperature decrease with age?	12
(5)	Publication bias and methodological considerations	13
IV.	Discussion	13
(1)	Differential effects of incubation temperatures across traits	14
(2)	Long-term effects of incubation temperatures on reptile phenotypes	16
(3)	Ecological, evolutionary, parental and methodological mediators of incubation temperature	17
(4)	Climate change and conservation implications	18
(5)	Future research directions	18
V.	Conclusions	19
VI.	Acknowledgements	19
VII.	References	19
VIII.	Supporting Information	25

I. INTRODUCTION

The majority of living organisms spend the most formative part of their life, embryonic development, exposed to the vagaries of the environment. Despite the environmental buffering afforded by parental resources, protective membranes, and favourable microhabitats (Roach & Wulff, 1987; Refsnider & Janzen, 2010), an embryo has no alternative but to develop in the microhabitat in which it settles, coping with often unpredictable or fluctuating environmental conditions. These conditions influence development (developmental plasticity), producing phenotypic variation that contributes crucially to ecological and evolutionary dynamics (Sultan, 2007; Uller, 2008; Moczek *et al.*, 2011).

While many vertebrate embryos are conspicuously shielded from environmental fluctuations *via* parental care (i.e. egg-tending and viviparity; Reynolds, Goodwin & Freckleton, 2002; Mank, Promislow & Avise, 2005), reptiles typically deposit their eggs in an untended nest (Reynolds *et al.*, 2002). Terrestrial nests of reptiles vary and fluctuate in temperature and moisture content (e.g. Mrosovsky, 1992; Wood & Bjorndal, 2000; Morjan, 2003; Glen & Mrosovsky, 2004; Warner & Shine, 2008; Telemeco, Elphick & Shine,

2009; Schwanz *et al.*, 2010). Moreover, incubation conditions can influence the phenotype of hatchling reptiles, impacting survival, morphology and behaviour (Deeming & Ferguson, 1991; Deeming, 2004; Booth, 2006). The basic assumption that incubation temperatures vary and are important for reptile phenotype and survival provide a foundation for many fields of research. For example, it underpins research on maternal nesting behaviour (Schwanz & Janzen, 2008; Warner & Shine, 2008; Refsnider & Janzen, 2010) and the predicted impacts of climate change (Janzen, 1994; Hawkes *et al.*, 2007; Mitchell *et al.*, 2008; Boyle *et al.*, 2016). Additionally, it features in understanding demographic and evolutionary dynamics in reptiles (Girondot *et al.*, 2004; Schwanz *et al.*, 2010; Boyle *et al.*, 2014) and forms the basis for adaptive explanations of variation in sex-determining mechanisms (i.e. genotypic or temperature-dependent; Shine, 1999b; Pezaro, Doody & Thompson, 2016; Schwanz *et al.*, 2016), maternal effects, and the evolution of viviparity [i.e. the maternal manipulation hypothesis (MMH); Shine, 1995; Schwarzkopf & Andrews, 2012].

Despite an extensive literature, general patterns of how incubation temperature impacts embryonic and post-hatching phenotype are unknown. Qualitative reviews

of the topic suggest that warmer temperatures can both increase *or* decrease trait values, or have no significant impact, depending on the species and study (Deeming & Ferguson, 1991; Deeming, 2004; Booth, 2006). While species-specific patterns are exceptionally useful in understanding a single species, describing general patterns would expand the field in many ways. For example, we could predict the direction or magnitude of phenotypic change when nest temperatures warm in understudied species. In addition, if we knew which phenotypes are most strongly impacted by incubation temperature and whether this phenotypic variation persists as individuals age, we could estimate fitness consequences. Moreover, a basic understanding of developmental plasticity provides insight into the adaptive significance of maternal nesting behaviour and sex-determining mechanisms.

It is unsurprising that the impact of increasing incubation temperatures is so variable when considering several crucial biological factors. First, specific traits may be sensitive in different ways to temperature or have counter-intuitive links with fitness (Deeming, 2004; ‘propensity to run’ *sensu* Janzen, 1995). Second, we expect nonlinear relationships between developmental rate, phenotype and temperature (i.e. reaction norms; Georges *et al.*, 2005). For example, the relationship may monotonically increase or decrease (Fig. 1A). This type of relationship is common, for example, for incubation duration (e.g. Lowenborg *et al.*, 2010; Rodriguez-Diaz *et al.*, 2010). Alternatively, the relationship may be concave down (or concave up), with a temperature of peak phenotypic value (i.e. an optimal temperature) and lower phenotypic values as temperatures become increasingly hot or cold (Fig. 1B). This type of reaction norm is expected for hatching success (e.g. Spotila *et al.*, 1994; Du & Ji, 2003). Whereas linear reaction norms allow simple predictions regarding effect sizes as temperature change increases, the same is not necessarily true of nonlinear reaction norms (Table 1). For a monotonic, nonlinear relationship, the magnitude of the effect becomes greater as the temperature increase becomes larger; that is, we predict a 6°C increase in temperature to have a greater effect on phenotype compared with a 2°C increase in temperature (Fig. 1A). Moreover, the effect of a given change in temperature (e.g. 2°C) is always in the same direction (positive or negative), but the effect magnitude depends on whether the temperatures studied fall in an area of the reaction norm characterized by a steep relationship or a shallow tail. By contrast, for a concave reaction norm (Fig. 1B), greater increases in temperature (e.g. 5°C) may often lead to larger phenotypic effects than smaller increases (e.g. 1°C), but very large increases in temperature (e.g. 10°C) may have a small effect. Similarly, the magnitude and sign (positive or negative) of the effect depends entirely on whether the mid-temperature falls along a steep or shallow section of the reaction norm, and whether it is a relatively cold or hot part of the reaction norm.

To examine whether general patterns exist in the impact of incubation temperature and which biological factors contribute variation to this impact, we conducted a systematic review and meta-analysis of the literature.

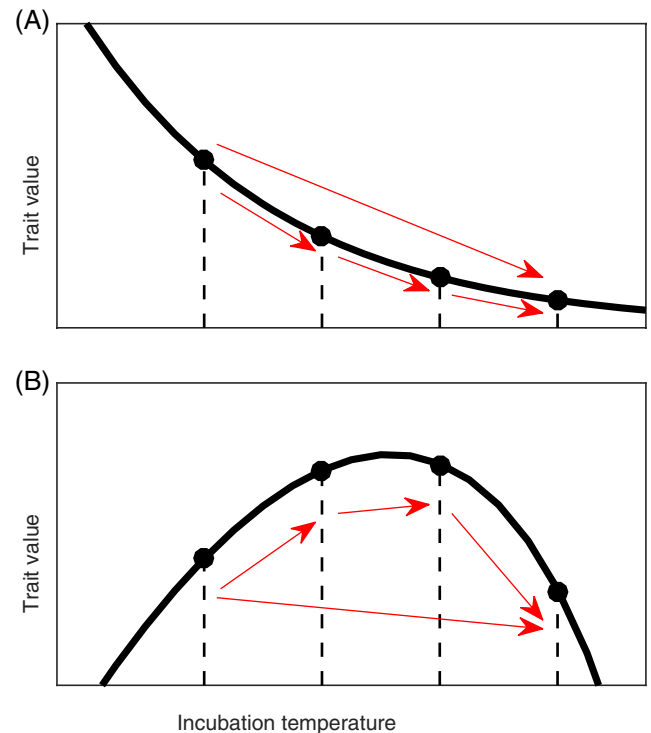


Fig. 1. Illustration of two commonly observed relationships (reaction norms) of hatchling trait value across different incubation temperatures: (A) a non-linear, monotonic function typically observed for incubation duration; (B) a concave down function often assumed to describe hatching success and many phenotypes. Arrows depict the expected variability in direction and magnitude of ‘effect’ on trait value when comparing two incubation temperatures.

For each pairwise temperature comparison in each paper, we calculated ‘the effect’ (see Section II.2) on phenotypic value of increasing incubation temperature. A standardized meta-analytic ‘effect size’ allows comparison across a wide variety of measurement units, permitting consideration of diverse phenotypes. Specifically, we asked: (i) whether the effect of incubation temperature varies across reptilian order and the type of phenotype measured; (ii) how the effect depends on the specific temperature regime, including thermal fluctuations of incubation treatments (e.g. Georges *et al.*, 2005; Booth, 2006); and (iii) whether the effect persists as animals age.

II. METHODS

(1) Literature search and data collection

We report our systematic literature search following, in part, the preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement (see online Fig. S1; Liberati *et al.*, 2009). We conducted a literature search using ISI *Web of Science* (v.5.13.2) with the ‘title’ or ‘abstract’ search terms, temperature* AND incubat*, along with one of the

Table 1. Predictions in effect sizes depending on the shape of the reaction norm

Prediction	Linear	Nonlinear, monotonic	Concave
As the change in temperature increases, the absolute effect size will increase	Yes	Yes	Not always
A given change in temperature always leads to the same effect size	Yes	No	No
The effect size of a given temperature change depends on the mid-temperature	No	Yes	Yes
The sign (positive or negative) of the effect will remain constant across mid-temperatures and temperature differences	Yes	Yes	No

following: reptil*, lizard*, squamat*, snake*, turtle*, chelon*, testudin*, crocodil*, alligator*, tuatara*, sphenodon*. We cross-referenced this search with articles published in three major qualitative reviews: Deeming & Ferguson (1991), Deeming (2004) and Booth (2006), and included any papers missed in our initial search. While it is possible that our search missed or excluded some papers, it yielded a large, unbiased sample of studies indexed up to 15 January 2016.

Our search generated 1319 unique papers. To be included in the meta-analysis, the paper had to conduct experimental incubation of reptile eggs (turtles, snakes, lizards, crocodiles, tuatara) under two or more incubation temperatures within 48 h of oviposition and to report data on incubation duration, hatching success, or post-hatching traits. We considered only papers published in 2014 or earlier, as prior database searches revealed that there is a substantial lag time in publication date and indexing in *Web of Science*, and we wanted to have a complete sample over a given time interval. We eliminated 658 studies based on the title and abstract as they did not meet the inclusion criteria.

The remaining 661 papers were further scrutinized. In addition to the above criterion, treatments, pairwise comparisons or entire papers were excluded if: (i) the species studied was viviparous; (ii) hormones were added to the eggs; (iii) the incubation period was incomplete or some other alteration of the egg was performed; (iv) the study used a temperature-shift experiment (i.e. moved eggs from one temperature to another during incubation); (v) incubation temperature comparisons were confounded with another factor such as year of study, population, maternal environment, or offspring post-hatching environment; (vi) the study reported reproductive traits of adults that had been incubated under different temperatures as these data were available in only one study; (vii) incubation temperature pairs differed in another incubation factor (e.g. moisture, fluctuation) in multi-factorial experiments (we only extracted relevant comparisons from thermal treatments that controlled for these effects). Following these exclusion steps, we had 179 papers with extractable data. Finally, we decided not to use data on sex ratio for the present analyses, as it is given that the effect sizes would differ dramatically between species with temperature-dependent sex determination (TSD) and genotypic sex determination (GSD), leaving a total of 175 studies from which data were extracted.

Hypotheses for the evolution of viviparity in reptiles are based on the importance of developmental (e.g. gestational) temperature for offspring traits (Shine, 1995, 2014); thus examination of effect sizes under different gestational temperature treatments would be useful in testing these hypotheses. However, we excluded viviparous taxa from our study for a number of reasons. First, it has been argued that effective tests of adaptive hypotheses require an understanding of how offspring traits relate to fitness directly and how maternal fitness is impacted (Schwarzkopf & Andrews, 2012). While many of the studies did quantify traits indirectly related to fitness, almost none quantified more direct estimates, such as survival. Second, it was clear that thermal manipulations during gestation were highly variable across studies, making comparisons within viviparous taxa and with oviparous taxa challenging. Specifically, some studies held pregnant females at constant (forced) temperatures with no fluctuation during gestation ($N = 7$), whereas others depressed the temperature at night, leading to uneven levels of temperature fluctuations among treatments ($N = 6$). Alternatively, many studies did not manipulate body temperature directly, but rather manipulated thermoregulatory opportunities for females ($N = 13$); these treatments would be difficult to compare to oviparous taxa. Lastly, the smaller number of studies made it challenging to account for, or understand, how effect size may be impacted by temperatures within thermal treatments or different experimental designs. Given these limitations, we felt that a quantitative synthesis of studies on viviparous taxa would best be addressed in a completely separate analysis where the unique challenges and opportunities afforded by viviparous research could be addressed appropriately. Nonetheless, Table S1 provides readers with a qualitative overview of studies conducting early thermal manipulations during gestation in viviparous taxa.

From the 175 studies, we extracted either (i) means, standard deviations (or standard errors/confidence intervals) and sample sizes, or (ii) proportions and sample sizes for each incubation temperature and response variable reported in the text, tables and/or figures. When both raw and adjusted (e.g. least-square) means were available we used the raw data. Data presented in figures were extracted using DataThief v.1.7 (<http://datathief.org/>). We contacted authors requesting clarification, raw data or summary statistics where necessary. Any data where means, errors

and sample sizes could not be obtained were excluded. Additionally, we took a conservative approach by excluding data with replicate samples in a treatment group, but that had a standard deviation/error of zero ($N_{\text{effects}} = 52$).

(2) Effect size calculation

Traits varied in their units of measurement so we calculated unbiased standardized mean differences, Hedges' g (Borenstein *et al.*, 2009), as our 'effect size'. Each effect size is a pairwise difference between two incubation temperatures and allows comparisons of trait types that vary in their unit of measurement as mean differences are standardized by their pooled standard deviation (see Appendix S1). In all cases, the mean trait value at the colder temperature was subtracted from the mean trait at the warmer temperature. Therefore, the value represents the effect of increasing incubation temperatures (Fig. 1). Positive effect sizes represent larger mean values for the hotter incubation treatment, whereas negative effect sizes represent incidences where the colder incubation temperature resulted in larger mean responses. While most traits were presented as means and errors, some were provided as proportions (e.g. hatching success, survival, proportion fleeing), therefore, we calculated log odds ratios ($\ln OR$) and converted these estimates to Hedges' g for analysis so that effects could be compared (Borenstein *et al.*, 2009).

We used Cohen's (1988) benchmarks when interpreting the magnitude of effect sizes classifying small, medium and large effects as 0.2, 0.5 and 0.8, respectively. We recognize that these benchmarks should be used as a guide rather than strict thresholds (Fritz, Morris & Richler, 2012). In our study, a single paper often presented results from multiple incubation treatments (range = 2–15 incubation treatments) and we calculated Hedges' g for all pairwise comparisons. Although this approach introduces a correlation structure between thermal treatments for a given trait it ensured that the maximum amount of data was used (Noble *et al.*, 2017a). We discuss the results of sensitivity analyses (Nakagawa *et al.*, 2017; Noble *et al.*, 2017a) to test the impact of within-study correlations in Appendix S4, and some caveats of using Hedges' g in Appendix S1.

(3) Moderator variables and predictions

Studies were highly heterogeneous in their methods, species, temperature treatments applied and response variables measured. As such, in addition to recording the species and study for each effect size, we recorded a number of moderator variables likely to explain heterogeneity in effect sizes, the first six of which we were primarily interested in and the last three we considered confounding variables.

(a) Order

We predicted that the effect of incubation temperature would differ across the four reptilian orders.

(b) Trait category

We predicted that the effect of incubation temperature would differ depending on the phenotypic trait being measured, therefore we categorized effect sizes as being derived from one of seven different trait types (Table 2): (i) incubation duration; (ii) behaviour; (iii) development; (iv) morphology; (v) performance; (vi) physiology; or (vii) survival. These categories largely followed those described by Deeming (2004). We categorized incubation duration as a separate trait category because we predicted that incubation temperature should have a strong negative overall Hedges' g (Birchard, 2004). We therefore considered incubation duration as a positive control that allowed comparisons with other trait types. We recognize that some trait values carry obvious directionality (e.g. bigger, faster, higher concentrations) that can be generalized and summarized, whereas other trait values (e.g. gaping behaviour, cognition) do not. For some traits directionality can vary, depending on how they were presented in the paper. For example, papers may report 'running speed' (m/s) or 'time to move 1 m' (s). These both measure locomotor performance, but in reciprocal units. Without the raw data points, measures of time (t) cannot be converted to $(1/t)$ to have comparable directionality to 'speed'. Thus, higher values of locomotion nearly always indicate 'faster', but sometimes include 'required more time'. This variation complicates interpretation of the direction of the effect size, but does not affect its magnitude (see Section II.4).

(c) Temperature difference between the paired temperature treatments

We predicted that this variable would affect the strength of effect sizes, in linear and nonlinear reaction norms. Larger differences between thermal treatments should result in larger effect sizes compared with more similar thermal treatments (but see Fig. 1).

(d) Mid-temperature between the paired temperature treatments

Due to our expectation that many developmental reaction norms would be nonlinear, we predicted that the effect size would vary as a function of the mid-temperature between treatments, in a linear or nonlinear fashion (Fig. 1). However, due to potential variation in the reaction norms, we had no *a priori* prediction of the exact nature of these impacts.

(e) Thermal fluctuations in temperature treatments

Incubation treatments differed in how variable temperatures were around a mean treatment temperature. Eggs were either incubated at 'constant' temperatures ($\pm \leq 0.5^\circ\text{C}$) or thermally 'fluctuating' incubation temperatures ($\pm > 0.5\text{--}7.5^\circ\text{C}$). Given that most studies were incubated at constant incubation temperatures, this distinction led to a highly-skewed distribution when looking at fluctuation as a continuous variable. To avoid problems when modelling, we categorized studies as coming from either a 'constant' or 'fluctuating' incubation regime. Not all studies reported

Table 2. Trait categories used as moderator variables

Trait category	Examples
Incubation duration	Time to hatching
Behaviour	Basking, thermoregulation, activity, exploration, aggression, anti-predator behaviour, number of stops, display, feeding, response time
Development	Growth rate, skin shedding, abnormalities, asymmetries
Morphology	Length of body/limbs, mass (live, wet, dry, carcass), organ size, body condition, head dimensions, colour patterns
Performance	Sprint or terrestrial speed, swimming speed, righting time, endurance, stroke frequency
Physiology	Body components (e.g. ash, energy, lipids), residual yolk, hormones, metabolic rate, thermal limits
Survival	Hatching success, post-hatching survival

temperature variability in treatments and we assumed that these were constant (i.e. 0). We predicted that effect sizes in fluctuating thermal treatments would be lower compared to treatments where eggs were incubated at constant temperatures. This prediction is based on the fact that, when temperatures fluctuate, they experience a greater range of temperatures, and two thermal regimes thus share more temperatures in common. This is akin to comparing two frequency distributions where either: (a) the standard deviation is very low and the distributions are very different (constant treatments) or (b) the standard deviation is very high and the distributions overlap considerably (fluctuating treatments).

(f) Age of measurement

Effects of early environments have been shown to have an impact on phenotypes in early ages (Burton & Metcalfe, 2014), and while some studies show long-term effects of early environments throughout life (Kerr *et al.*, 2007), these are predicted to decrease as individuals age (Deeming, 2004). We therefore recorded the age, in days, that phenotypes were measured and categorized them into one of four age categories: (i) 0–10 days; (ii) 11–180 days; (iii) 181–365 days; and (iv) >365 days post-hatching. The age for survival data was the end of the interval over which survival was assessed. We modelled age as a categorical variable, as opposed to a continuous one, given the few studies measuring phenotypes beyond 6 months and because this made it easier to calculate the mean magnitude (absolute value) of effects. Different reptilian orders vary in their overall longevity, possibly making comparisons across disparate groups challenging using absolute age measurements. While using a proportional value of age (i.e. a proportion of total longevity for each species) may overcome this challenge and test interesting biological questions, it is difficult to acquire such data for long-lived and cryptic taxa, thus limiting our sample size. Importantly, all orders were represented across all age categories in our study, preventing confounds between age and order.

(g) Egg randomization strategy

Studies differed in how eggs were allocated to each incubation treatment and we categorized effect sizes as coming from one of four types of egg-incubation design: ‘split clutch pseudo’, ‘split clutch’, ‘independent pseudo’ and ‘other’. Studies categorized as ‘split clutch pseudo’ distributed eggs from a single clutch evenly across multiple incubation treatments regardless of the size of the clutch. This practice resulted, at times, in high levels of non-independence because many eggs from a single female were represented in multiple incubation treatments (Noble *et al.*, 2017a) (see online Appendix S1 for further discussion). This contrasted with studies that utilized a ‘split clutch’ design where a single egg from a clutch was represented in each of the incubation treatments. In some studies, each incubation treatment was composed of eggs from independent females (i.e. clutches in one treatment only), although eggs from the same female were often treated as replicates (‘independent pseudo’). When it was not possible to determine the egg-allocation strategy, we categorized studies as ‘other’. We acknowledge that levels of this categorical variable in some cases have a taxonomic bias by virtue of some groups (e.g. geckos, small skinks) having small (often only two eggs) clutch sizes and so most studies categorized as ‘split-clutch’ were biased towards these groups.

(h) Data type

Studies presented both raw and adjusted means for incubation treatments. While we chose raw data over adjusted data when possible, in some cases only adjusted data were presented. To test whether this difference explained variation in effect sizes we categorized data presented as least-square means or adjusted (‘adj’) means separately from raw data (‘raw’).

(i) Sex ratio of sample

Incubation temperatures affect sex ratio in many reptile groups which can lead to differences in phenotypes between incubation treatments being driven by sex differences rather

than incubation treatment *per se*. Sexes often show little phenotypic differences at early ages (reviewed in Badyaev, 2002) and most studies presented data from hatchlings. As such, sex ratio changes were unlikely to explain much variation in effect sizes. Nonetheless, we attempted to control for these effects by categorizing the sex of the samples for each incubation treatment ('male', 'female' or 'mixed') when sex ratio data were available, and then created a three-level categorical predictor that accounted for same sex, mixed sex and opposite sex comparisons (i.e. 'mixed.sex' 'same.sex', 'diff.sex'). This categorical moderator variable also dealt with situations where male and female data were presented separately within papers. If the species was known to have TSD we attempted to account for sex differences of samples if sex ratio was known. Where no sex ratio data were available we assumed a 'mixed' sex sample for each treatment.

(4) Meta-analysis

We performed meta-analyses using Hedges' g as well as the absolute value of Hedges' g ($|g|$), which we interpret as the 'magnitude' of the effect size. We analysed $|g|$ because we were specifically interested in testing whether incubation temperature drove large effects on phenotypes regardless of directionality, particularly because *a priori* predictions on directionality were not always possible. We implemented Bayesian meta-analytic random effects and meta-regression models in *R* (v.3.2.1; R Core Development Team) with the *MCMCglmm* package (Hadfield, 2010) using traditional meta-analytic approaches that accounted for measurement error and sources of non-independence. Absolute effect sizes follow a folded-normal distribution (Hereford, Hansen & Houle, 2004; Kingsolver *et al.*, 2012), so we applied posterior distributions of parameters from Gaussian models to the folded normal distribution in order to obtain mean and credible intervals for absolute magnitudes (i.e. 'analyse and transform' *sensu* Morrissey, 2016a,b). Doing so provides estimates and credible intervals for parameters that relate to absolute effect sizes. However, we were also interested in modelling and predicting the effect of incubation temperature on absolute effect sizes as a function of temperature-related variables (Section II.3). Given that the folded normal distribution is not yet implemented as a family for modelling continuous predictors, we modelled absolute effect size assuming a log-normal error distribution (see this section below).

For all our models, we first ran three independent MCMC chains and tested chain convergence using the Gelman & Rubin (1992) convergence diagnostic (PRS < 1). MCMC chains were run for 510000 iterations with a 10000 iteration burn in and a thinning interval of 1000. In total across the three chains, we ran 1500000 iterations sampling 1500 iterations from the posterior distribution. We ensured there was low autocorrelation (lag values < 0.1) within our MCMC chains and explored plots of MCMC chains to ensure they were mixing well. We used non-informative uniform priors for our fixed effects and inverse-Wishart priors for our random effects (Hadfield, 2010). We also used

parameter-expanded priors for our random effects, however, these results did not differ from other parameterizations and so we only present models that use an inverse-Wishart prior ($V = 1$, $\nu = 0.002$). Once we verified that chains were not affected by starting values, we re-ran our models using a single chain (burn in = 30000, iterations = 1500000, thinning = 1000) to obtain posterior estimates of our parameters. Throughout, we present the mean and 95% highest posterior density intervals (95% credible intervals) of the posterior distribution. Credible intervals not overlapping each other and the value zero suggest statistical significance.

Obtaining a fully resolved phylogenetic tree with branch lengths for diverse taxa is notoriously difficult given that homologous character data are often not available across disparate lineages, and tree topology can be unresolved (Chamberlain *et al.*, 2012; Hinchcliff *et al.*, 2015). Nonetheless, ignoring phylogenetic relatedness among species can affect meta-analytic results (Chamberlain *et al.*, 2012). In an attempt to account for phylogenetic non-independence, we derived a phylogenetic tree describing the hypothesized relationships among taxa using the open tree of life database (Hinchcliff *et al.*, 2015; <https://tree.opentreeoflife.org>). The open tree of life (OTL) constructs tree topology by synthesizing existing phylogenetic trees along with taxonomic data from published research (Hinchcliff *et al.*, 2015), where taxonomy contributes to tree topology when no phylogenetic trees are available. The OTL does not provide branch lengths and so we were not able to account for evolutionary divergence times between taxa in our analysis. We obtained a phylogenetic tree for the taxa in our data set by resolving synonymous taxonomic names and obtaining tree topology for existing taxa in the OTL database. We resolved polytomies and the placement of un-matched taxa by supplementing tree topology with existing phylogenetic relationships presented in Pyron, Burbrink & Weins (2013) and Jin & Brown (2013). The 'resolved' phylogenetic tree used for analyses is provided in Fig. S2. Not all species were present for all analyses, therefore we pruned the tree by dropping taxa not present in respective data sets. Given that we were not able to include divergence times, our tree should be considered only a rough estimate of phylogenetic relationships. We derived a phylogenetic correlation matrix (\mathbf{A}) by estimating branch lengths using Grafen's method (Grafen, 1989), where lengths are assigned to each edge on the phylogenetic tree and the correlation between species is obtained by calculating the total length from the root of the tree to each species' node. In addition to including phylogenetic relationships, as a sensitivity analysis, we also included a nested taxonomic random effect in a separate model; this gave nearly identical results to models with phylogeny and so is not presented here.

We first ran an intercept-only multi-level meta-analytic model (MLMA) to obtain an overall mean estimate for all trait types (except incubation duration) and determine the relative amount of heterogeneity in effect sizes explained by various random effects. We ran a separate intercept-only model for incubation duration data because we had strong

a priori predictions that incubation duration should be inversely related to incubation temperature, leading to large negative effect sizes (Birchard, 2004). In these models, we included a number of sources of non-independence between our effect sizes including study, species, and phylogeny (Nakagawa & Santos, 2012; Nakagawa *et al.*, 2017; Noble *et al.*, 2017a). Details on model structure are provided in Appendix S2.

In addition to phylogeny-, species- and study-level non-independence, our data contained three sources of within-study covariance: (i) shared treatment comparisons, (ii) shared traits, and (iii) shared samples across treatments (e.g. split clutch designs) (Noble *et al.*, 2017a). We attempted to account for within-study covariance in our main analysis using the full data set, but this approach was computationally prohibitive. Thus, we conducted a sensitivity analysis using data subsets. Importantly, the addition of the within-study covariance matrix using smaller data sets did not impact our results greatly (see online Appendix S2 and Fig. S3). Therefore, in our main analyses we estimated model parameters and credible intervals assuming within-study covariance to be negligible. Ignoring the covariance will lead to decreased credible intervals around parameter estimates, increasing the probability of type I errors (Noble *et al.*, 2017a). However, this approach should not lead to biased parameter estimates. Finally, we quantified heterogeneity measures (I^2 ; Higgins & Thompson, 2002; Nakagawa & Santos, 2012) from intercept-only models; these can be interpreted as the proportion of variation explained by study, species and phylogenetic effects. See online Appendix S2 for further details.

(5) Multi-level meta-regression models

After estimating overall heterogeneity, we explored the biological drivers that we anticipated would moderate the effect of incubation temperature. In all models, we included sample sex as a covariate to account for same-sex, mixed-sex and opposite-sex comparisons. Our moderator variables allowed us to ask three major questions and explore how different study methods and temporal trends impact effect sizes. See online Appendix S3 for further information.

(a) Do the effects of incubation temperature differ among orders and trait categories?

To test whether effect sizes varied between trait categories, we extended our MLMA model by including a trait category moderator variable and interacting it with taxonomic order to account for the possibility that traits may respond differently to incubation temperatures in different orders. The posterior distribution for the mean effect size for each trait category across orders was calculated and estimates were applied to the folded normal to estimate the mean and 95% CIs of the absolute effect size for each trait category (Morrissey, 2016b). Given that traits varied in their variance we estimated a residual variance for each trait category

within each order, particularly given that the folded normal distribution is sensitive to estimates of residual variance. Incubation duration had a markedly different distribution with mainly all negative effects with extreme variance. As such, we modelled it separately by log-transforming the absolute effect sizes and assuming a Gaussian error distribution. The posterior distribution for each of the orders was then back-transformed to obtain the mean magnitude of effect on the absolute scale.

(b) Does the temperature difference, mid-temperature, and the fluctuation in temperature between incubation treatments impact effect size?

The incubation temperature eggs experience is predicted to be one of the major explanatory variables causing heterogeneity in effect sizes across studies (Birchard, 2004; Deeming, 2004). Additionally, species likely have different thermal reaction norm shapes (only part of which may be quantified by a given study), and eggs can be incubated under both fluctuating and constant thermal conditions, complicating simple comparisons (Deeming, 2004). To characterise the impact of these modifiers on the effect size, we tested for linear, non-linear and interactive effects of treatment mid-temperature, temperature difference and the fluctuation in treatment temperature on effect size. We included quadratic and interaction terms in this model to account for the possibility that thermal reaction norms for traits in response to treatment temperature differences and mid-temperature may not be linear and may depend on their combined effects. We also included an interaction between linear and non-linear effects with trait type to account for the possibility that reaction norms for traits may be different. In addition to modelling Hedges' g , we also modelled absolute effect size ($|g| + 0.001$) using a log-normal distribution, as the folded normal is difficult to implement for continuous parameter estimates. We calculated conditional and marginal R^2 for these models to understand the amount of variance in effect sizes explained by temperature-related moderators (Nakagawa & Schielzeth, 2013).

(c) Do the effects of incubation temperature decrease with age?

We tested whether the effect of incubation temperature decreases as the age of trait measurement increases. While we were interested in testing whether the effect of age shows similar patterns across trait categories, this was not possible because several trait categories had very few studies that measured traits at later ages. However, morphological and physiological traits did have at least two or more studies across all age categories and so we also ran two models to estimate whether effect size for these traits mirrored overall patterns (i.e. pooled across all trait types) across age. In these models, we also estimated different residual variances in each trait category and applied the folded normal conversion to the posterior distributions.

(6) Publication bias and methodological considerations

Publication bias results when studies that do not find statistically significant results are less likely to be published. This process may bias the sample of studies available and can impact results of a meta-analysis (Borenstein *et al.*, 2009). To test for publication bias we explored the relationship between meta-analytic residuals from our MLMA model and precision ($SE_{g,i}^{-1}$), where $SE_{g,i}$ is the standard error of effect size i , using funnel plots. If publication bias exists, we expect to see a ‘missing set’ of effect sizes from studies with small sample size (low precision) (Borenstein *et al.*, 2009). In the absence of publication bias, studies should be distributed symmetrically around the effect-size mean regardless of study precision.

Funnel plots may not, however, provide a clear indication of publication bias on their own. Therefore, we also tested for publication bias using a modified version of Egger’s regression (Egger *et al.*, 1997; Nakagawa & Santos, 2012; see online Appendix S3). Residuals from our models account for non-independence resulting from shared study, phylogeny and species effects. Lack of publication bias leads to a non-significant intercept (β_0) from Egger’s regression suggesting that effects sizes come from a homogeneous, symmetrical distribution (Egger *et al.*, 1997; Sutton *et al.*, 2011). We also tested for the number of missing studies using the trim-and-fill method (Duval & Tweedie, 2000) in *metafor* (Viechtbauer, 2010). We additionally explored the influence of methodological moderators including differences in experimental design and the presentation of raw or adjusted statistics. We also tested for temporal trends by including a publication-year moderator in models. Code and data can be accessed from Noble, Stenhouse & Schwanz (2017b).

III. RESULTS

We used data from 175 publications (published from 1974 to 2014), totaling 6440 pairwise effect sizes from 92 species. Each paper yielded an average of 36.8 ± 44.43 (mean \pm S.D.) effect sizes, with a range of 1–200. Papers covered 1–6 trait categories (2.65 ± 1.5), with 1–26 individual traits (6.05 ± 5.44) compared across 1–36 temperature pairs (5.06 ± 6.5). The greatest taxonomic coverage was in Squamata, while Crocodylia and Rhynchocephalia showed poor coverage (Fig. 2A).

(1) Overall impact of incubation temperature on reptile phenotypes

As predicted, there was a significant negative weighted mean effect size of incubation temperature on incubation duration ($g = -8.42$, 95% CI = -10.73 to -6.63 , $N = 703$). Accounting for large outlying effect sizes (i.e. constraining data to three S.D. above the mean) gave qualitatively similar results (not shown). Hot incubation temperatures strongly

decreased incubation duration across all reptiles (Figs 3 and S4). Overall, differences between studies explained a significant proportion of heterogeneity in effect sizes for incubation duration ($I_{st}^2 = 0.35$, 95% CI = 0.19 – 0.45), whereas phylogeny ($H^2 = 0$, 95% CI = 0 – 0.22) and species differences ($I_{sp}^2 = 0$, 95% CI = 0 – 0.05) explained negligible amounts of variation.

Across all other trait types (behaviour, morphology, performance, survival, development, physiology), the overall weighted mean effect size of incubation temperature on phenotypic traits was negligible ($g = -0.03$, 95% CI = -0.19 to 0.10 , $N = 5737$) (Fig. 3). The *magnitude* of the effect of incubation temperature on phenotypic traits, however, was moderate to large ($|g| = 0.75$, 95% CI = 0.72 – 0.79) (Fig. 2B). This result derives from strong effects in both positive and negative directions for all trait types contributing to a very small overall effect, demonstrating the challenges of focusing on the raw effect sizes. Importantly, however, heterogeneity in raw effect sizes could not simply be explained by chance given significant variation in effect sizes beyond sampling variability ($I_t^2 = 0.91$, 95% CI = 0.90 – 0.91). Again, differences between studies explained a significant proportion of heterogeneity in effect sizes ($I_{st}^2 = 0.13$, 95% CI = 0.09 – 0.2), whereas phylogeny ($H^2 = 0$, 95% CI = 0 – 0.07) and species-specific differences ($I_{sp}^2 = 0.02$, 95% CI = 0 – 0.07) explained essentially no heterogeneity.

(2) Do the effects of incubation temperature differ among orders and trait categories?

The magnitude of the effect of incubation temperature on traits was moderate to large (i.e. absolute effect size, $|g|$) in most cases (Fig. 2B). Unsurprisingly, the magnitude of the effect on incubation duration was substantially larger than for any other trait category (Fig. 2B). In Squamata and Testudines, survival was strongly impacted by incubation temperatures and differed significantly from all other traits. Behavioural and physiological traits in Squamata were also strongly affected by incubation temperature (Fig. 2B), differing significantly from other traits. Across orders, traits were mostly impacted in a similar way, although there was more variation among trait categories in Squamata compared to Testudines.

In comparison, the raw effect size (g) for traits across orders ranged from -10.98 to 0.55 (Fig. 3). Incubation duration across all orders showed strong negative effects (-10.98 to -3.75 ; Fig. 3) and development in Testudines showed a significant positive effect size (0.55 , 95% CI = 0.12 – 0.92), indicating that developmental traits in hot-incubated eggs were, on average, 0.55 S.D. units higher than cold-incubated eggs. Physiological traits in Squamata also showed a significant positive effect size (0.31 , 95% CI = 0.02 – 0.62). However, these should be interpreted cautiously given that credible intervals are close to zero and these analyses do not control for within-study correlations.

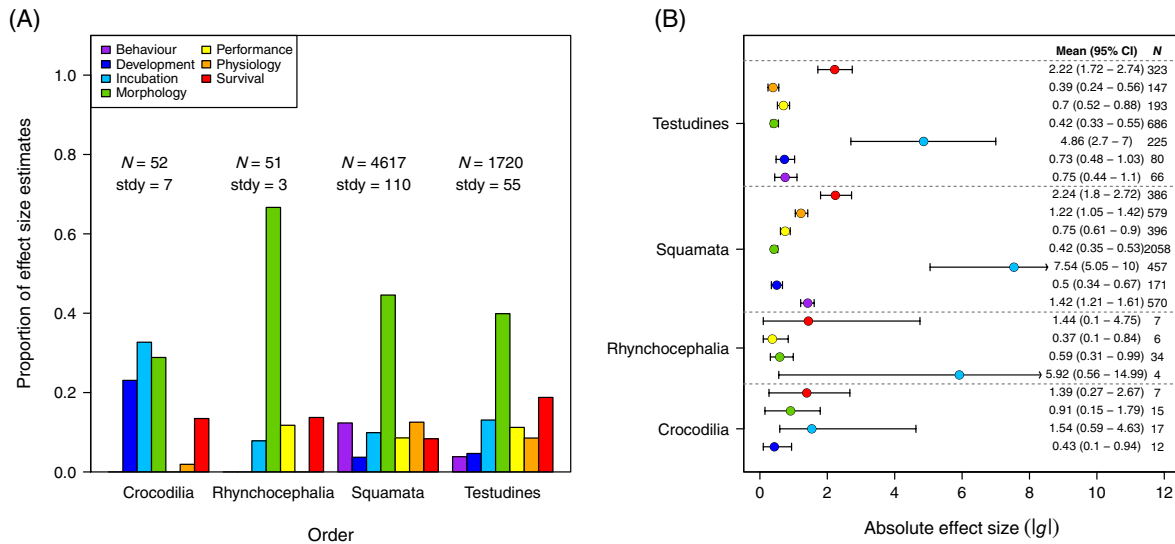


Fig. 2. (A) Proportion of effect size estimates across trait categories and taxonomic orders for the entire data set. N is the total number of effects in each order and ‘study’ is the total number of studies. (B) Magnitude of effect of incubation temperatures on phenotypic traits across orders. Colour corresponds to trait category colours in (A). N is number of effect sizes in the analysis and the posterior mean and 95% credible intervals are provided for each estimate. Posterior distributions of effects were taken from a model that included species, phylogeny and study as random effects and means in each group were applied to the folded normal distribution. Angled arrows indicate that credible intervals have been shortened to simplify visual presentation.

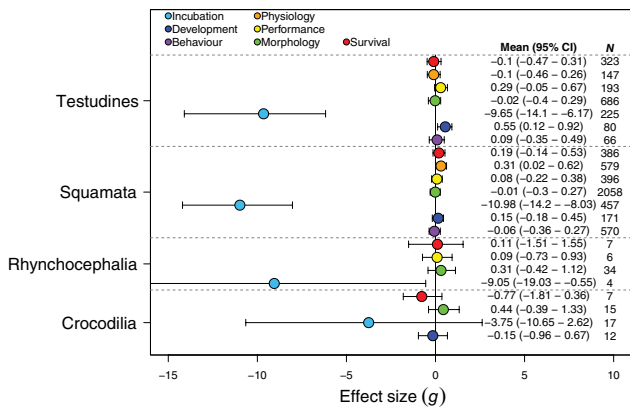


Fig. 3. Raw effect size (Hedges’ g) across trait types within each major reptilian order. Means are calculated for a mixed-sex sample and control for study, phylogeny and species non-independence. Credible intervals close to zero should be interpreted with caution as analyses were not able to control for all sources of non-independence. Credible intervals with arrowed heads indicate that large CIs have been shortened for clarity. N number of effect sizes in the analysis.

(3) Does the temperature difference, mid-temperature, and the fluctuation in temperature between incubation treatments impact effect size?

Most studies incubated eggs at constant temperatures ($N_{\text{studies}} = 148$); fewer studies (34 studies across 22 species) incubated eggs at fluctuating temperatures ($\pm >0.5$ to 7.5°C fluctuation). We predicted that fluctuating temperatures

Table 3. Model coefficients (*Est.*, parameter estimate) and lower (*L.*) and upper (*U.*) 95% credible intervals (*CI*) for the absolute effect ($|g|$) of treatment mid-temperature (T_{mid}), temperature difference (T_{diff}) and fluctuations in temperature (T_F) on effect size for incubation duration. Coefficients are presented on a log-transformed scale. Continuous coefficients are z -transformed $[(x - \bar{x}) / sd(x)]$. Bold type indicates credible intervals that do not overlap zero

	Incubation duration		
	<i>Est.</i>	<i>L. CI</i>	<i>U. CI</i>
Intercept	2.06	1.70	2.39
Sex (mix)	-0.05	-0.39	0.14
Sex (same)	-0.62	-0.84	-0.25
T_F	-0.45	-0.86	0.02
T_{mid}	-0.53	-0.62	-0.41
T_{mid}^2	-0.07	-0.11	-0.03
T_{diff}	0.66	0.56	0.74
T_{diff}^2	-0.19	-0.25	-0.12
$T_{mid} * T_{diff}$	0.04	-0.06	0.10
$T_{diff} * T_F$	-0.17	-0.45	0.16

would diminish the impact of increased incubation temperatures due to greater overlap in developmental temperatures. Fluctuating treatments had a tendency to decrease the effect across traits compared to constant temperatures and there was a slight tendency for this effect to be exacerbated when the temperature difference between treatments was large, but note that credible intervals overlapped zero for fluctuating temperature (T_F) and the interaction term between T_F and temperature difference ($T_{diff} * T_F$) (Tables 3–5 and Fig. S5).

Table 4. Model coefficients (*Est.*, parameter estimate) and lower (*L.*) and upper (*U.*) 95% credible intervals (*CI*) for the absolute effect ($|g|$) of treatment mid-temperature (T_{mid}), temperature difference (T_{diff}) and fluctuations in temperature (T_F) on effect size for behavioural, physiological and morphological traits. Coefficients are presented on a log-transformed scale. Continuous coefficients are z -transformed $[(x - \bar{x}) / sd(x)]$. Note that sex, T_F and $T_{diff} * T_F$ did not interact with trait type, and so these are the same across traits. Bold type indicates credible intervals that do not overlap zero

	Behaviour			Development			Morphology		
	<i>Est.</i>	<i>L. CI</i>	<i>U. CI</i>	<i>Est.</i>	<i>L. CI</i>	<i>U. CI</i>	<i>Est.</i>	<i>L. CI</i>	<i>U. CI</i>
Intercept	-0.87	-1.22	-0.59	-0.93	-1.25	-0.55	-1.24	-1.47	-1.02
Sex (mix)	0.00	-0.21	0.16	0.00	-0.21	0.16	0.00	-0.21	0.16
Sex (same)	0.03	-0.14	0.25	0.03	-0.14	0.25	0.03	-0.14	0.25
T_F	-0.29	-0.53	0.01	-0.29	-0.53	0.01	-0.29	-0.53	0.01
$T_{diff} * T_F$	-0.17	-0.37	0.05	-0.17	-0.37	0.05	-0.17	-0.37	0.05
T_{mid}	-0.37	-0.54	-0.14	0.06	-0.22	0.46	0.09	0.01	0.18
T_{mid}^2	0.00	-0.1	0.13	0.15	-0.05	0.47	0.13	0.08	0.19
T_{diff}	0.09	-0.17	0.3	0.48	0.28	0.77	0.25	0.18	0.33
T_{diff}^2	-0.03	-0.16	0.09	-0.07	-0.27	0.11	-0.01	-0.07	0.03
$T_{mid} * T_{diff}$	-0.15	-0.31	0.05	0.16	-0.06	0.57	0.04	-0.07	0.10

Table 5. Model coefficients (*Est.*, parameter estimate) and lower (*L.*) and upper (*U.*) 95% credible intervals (*CI*) for the absolute effect ($|g|$) of treatment mid-temperature (T_{mid}), temperature difference (T_{diff}) and fluctuations in temperature (T_F) on effect size for performance and development traits and survival. Coefficients are presented on a log-transformed scale. Continuous coefficients are z -transformed $[(x - \bar{x}) / sd(x)]$. Note that sex, T_F and $T_{diff} * T_F$ did not interact with trait type, and so these are the same across traits. Bold type indicates credible intervals that do not overlap zero

	Performance			Physiology			Survival		
	<i>Est.</i>	<i>L. CI</i>	<i>U. CI</i>	<i>Est.</i>	<i>L. CI</i>	<i>U. CI</i>	<i>Est.</i>	<i>L. CI</i>	<i>U. CI</i>
Intercept	-0.88	-1.19	-0.65	-0.88	-1.16	-0.59	-0.27	-0.54	-0.03
Sex (mix)	0.00	-0.21	0.16	0.00	-0.21	0.16	0.00	-0.21	0.16
Sex (same)	0.03	-0.14	0.25	0.03	-0.14	0.25	0.03	-0.14	0.25
T_F	-0.29	-0.53	0.01	-0.29	-0.53	0.01	-0.29	-0.53	0.01
$T_{diff} * T_F$	-0.17	-0.37	0.05	-0.17	-0.37	0.05	-0.17	-0.37	0.05
T_{mid}	-0.03	-0.20	0.16	-0.02	-0.25	0.21	-0.04	-0.15	0.11
T_{mid}^2	0.10	0.04	0.22	0.22	0.00	0.38	0.06	0.03	0.13
T_{diff}	0.22	0.06	0.38	0.43	0.29	0.62	0.37	0.23	0.49
T_{diff}^2	-0.08	-0.23	0.01	-0.12	-0.30	0.02	-0.03	-0.08	0.01
$T_{mid} * T_{diff}$	-0.32	-0.45	-0.15	-0.08	-0.27	0.16	-0.04	-0.11	0.06

We predicted that, in general, the impact of increased incubation temperatures would increase as the change in temperatures became greater. If reaction norms were concave, however, we expected that there would also be a strong, nonlinear influence of mid-temperature between the treatments and an interaction between mid-temperature and temperature difference (e.g. a small increase in temperature would have a much bigger effect at extreme temperatures than at intermediate temperatures where phenotypic values peak). Temperature differences between the treatments ranged from 0.2 to 29°C (mean = 4.52°C, S.D. = 2.47°C), allowing a substantial range to address this question. Mid-temperatures of incubation treatments ranged from 10.5 to 39°C (mean = 27.75°C, S.D. = 2.97°C). Rhynchocephalia were incubated at the lowest mid-temperatures (20.41°C), whereas Crocodylia (mean = 31.18°C), Squamata (mean = 27.41°C) and Testudines (mean = 28.76°C) were much warmer and similar to each other.

The parameter estimates and predictions for the magnitude of the effect (Tables 3–5, Figs 4 and 5) and for raw effect sizes (see online Tables S2 and S3, Fig. 6) across temperature regimes provide insight into the shape of the reaction norms for each trait category. However, because examining the effect size is a rather indirect method for describing reaction norms and because the apparent reaction norms vary across trait categories, we focus mainly on highlighting broad and salient results.

An illustration of how the parameter estimates relate to the reaction norm can be found by considering the assumed reaction norm for incubation duration (Fig. 1A): incubation duration consistently declines as temperature increases, but flattens out at very high temperatures (Figs 1A and 4). With a reaction norm of this shape, we predicted: (i) that the magnitude of the effect increases as the temperature difference increases (Fig. 4; positive estimate for T_{diff} in Table 3); (ii) that this increase in magnitude levels off

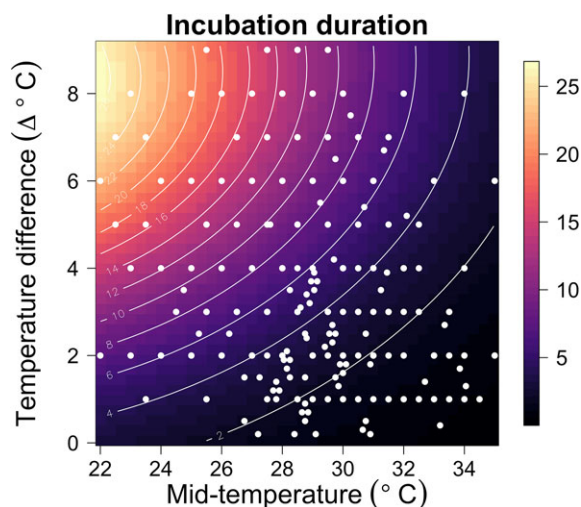


Fig. 4. Predictions of how mid-temperature and temperature difference impact the magnitude of effect ($|g|$) between hot and cold incubation treatments for incubation duration. Predictions are made over ± 2 S.D. of the mean temperature difference between treatments and their mid-temperature. Dots indicate actual mid-temperature and temperature difference combinations for each trait category. Absolute effect size ($|g|$) is modelled using log-transformed data and back-calculated to absolute effect sizes presented in the colour key.

at increasingly greater temperature changes due to the flattening of the curve (Fig. 4; negative estimate for T_{diff}^2 in Table 3); (iii) that a given increase in temperature has a greater effect at cold temperatures than at hot temperatures [negative estimate for mid-temperature (T_{mid}) in Table 3]; and (iv) that larger temperature increases will have a greater effect at cold temperatures than at hot temperatures (Table 3, not supported by $T_{mid} * T_{diff}$). We also found a particularly low impact of increasing incubation temperatures at high mid-temperatures (negative estimate for T_{mid}^2 in Table 3), suggesting a fairly flat relationship between incubation duration and temperature at high temperatures. These effects remained even when accounting for the largest outliers (i.e. setting all values greater than 3 standard deviations to 3S.D. – results not shown). Thus, the significance of the temperature terms considered here is driven by the shape of the reaction norm.

The magnitude of the effect of incubation temperature ($|g|$) on other trait categories also depended on mid-temperature and temperature difference, but this differed among the traits. Temperature differences between two incubation temperatures had a large impact on the magnitude of the effect across traits, with increased temperature differences leading to larger magnitude of effect for all traits except behaviour (T_{diff} ; Tables 4 and 5; Figs 5 and S6). The mid-temperature between the two incubation treatments had a less-consistent impact on the magnitude of effect, and there was evidence for non-linear patterns (T_{mid} and T_{mid}^2 ; Tables 4 and 5; Figs 5 and S7); however, some clear results were apparent. For behavioural traits, model predictions suggest that the impact of increasing incubation temperature

is greater at lower temperatures than at higher ones (T_{mid}). By contrast, for morphological traits, the impact is much higher at higher mid-temperatures than at lower ones (T_{mid} and T_{mid}^2). For performance and survival, the magnitude of the effect is greatest at cold and hot mid-temperatures and lowest at intermediate mid-temperatures. The predicted magnitude of effects for each trait category were moderate to large (see intercepts in Tables 4 and 5), although the full model explained very little variance in absolute effect sizes (marginal $R^2 = 10\%$).

Similar conclusions can be drawn when considering the influence of temperature regime on raw effect sizes (g) with additional insight into the shape of thermal reaction norms (Figs 6, S8 and S9; Tables S2 and S3). Performance, morphology and survival largely showed similar predicted patterns, with positive effects of increasing temperature difference (hot > cold) at lower mid-temperatures and negative effects (hot < cold) at high mid-temperatures; these trends are as one would predict for reaction norms with optimal phenotypes at intermediate temperatures (Fig. 1B). More complicated patterns were evident for behavioural, developmental and physiological traits, although all demonstrated negative effect sizes (decreasing phenotypic values) at high mid-temperatures. Marginal R^2 for the full model was 29%, whereas when considering the variance explained by both the fixed and random effects R^2 (conditional R^2) was 42%.

(4) Do the effects of incubation temperature decrease with age?

Most studies measured traits on offspring between 0 and 10 ($N = 4075$ effects across 140 studies) and 11–180 ($N = 1228$ effects across 50 studies) days of age. By contrast, only 20 ($N = 345$ effects) and 8 ($N = 89$ effects) studies measured traits between 181 and 365 and greater than 365 days, respectively (Fig. 7A). Forty-two papers presented measurements of traits at hatching and at least one other age category. All trait categories were represented across all age groups except performance, which was never measured at greater than 365-days post-hatching (Fig. 7A). Only 14 of 88 species (excluding incubation duration) were measured beyond 180 days, but representatives from each order were measured beyond this stage. To ensure that these effects were not the result of temperature differences between age categories we tested whether mid-temperatures were different across age categories. Comparing deviance information criteria (DIC) of models with and without age as a categorical variable (temperature as response and study as random effect) showed that temperature did not differ between age categories ($\Delta DIC = 5.14$). Additionally, temperature differences between age categories also did not differ ($\Delta DIC = 2.76$).

The effect of incubation temperature did not decrease across age categories as predicted (but see Figs S10 and S11), and the magnitude of the effect across age categories was moderate to large ($|g|$, black diamonds in Fig. 7B). Only morphology and physiology were measured with

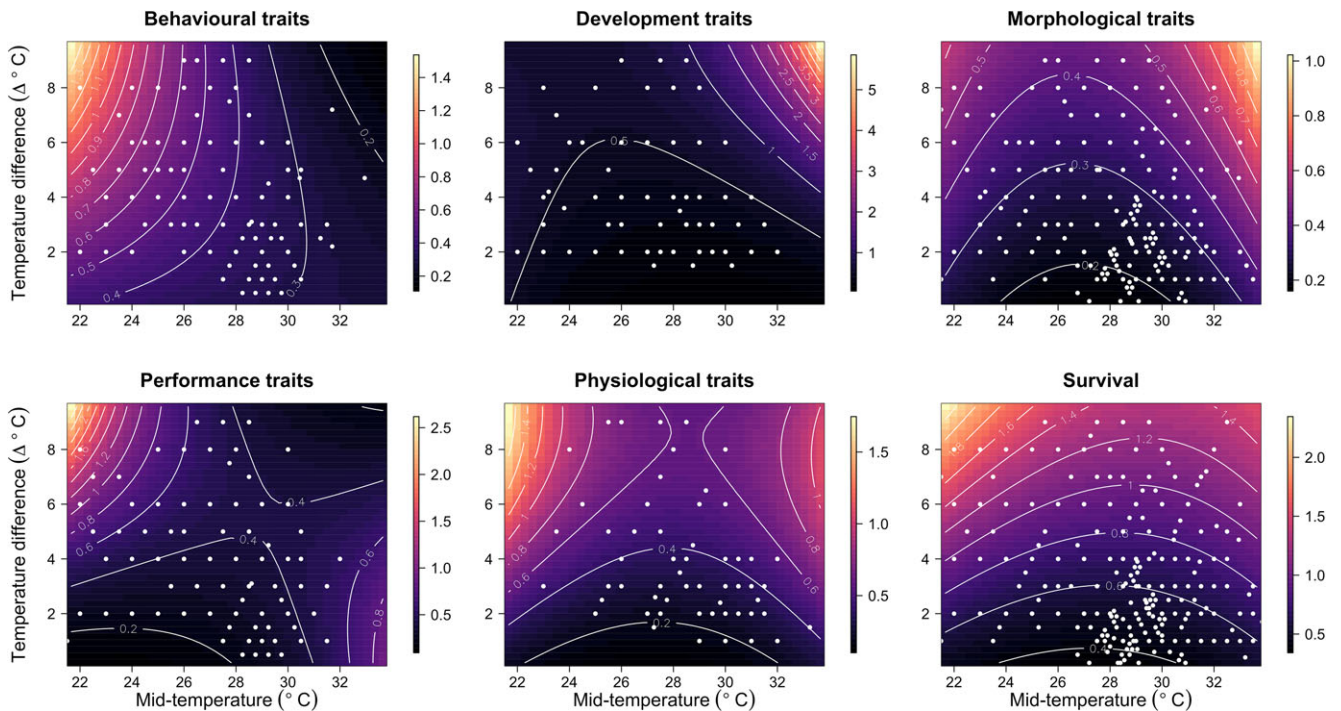


Fig. 5. Predicted absolute effect ($|g|$) of mid-temperature and temperature difference on the magnitude of effects across six trait categories (behaviour, development, morphology, performance, physiology and survival). Predictions are made for mixed-sex comparisons at constant incubation temperatures. Dots indicate actual mid-temperature and temperature difference combinations for each trait category. Predictions are marginalized over the random effects. Absolute effect size ($|g|$) is modelled using log-transformed data and back-calculated to absolute effect size presented in the colour key.

sufficient sample sizes across all age categories to explore patterns within trait categories. The magnitude of effects observed for physiological traits was large (Fig. 7B, orange circles) but mirrored overall effects across age, whereas morphological traits (Fig. 7B, green circles) across age showed a weaker effect, particularly at the 181–365-day age category.

Significant overall effects (g) of incubation temperature were apparent in the 181–365-day age category (Fig. 7B, black circles), with a small to moderate effect size where eggs in warmer incubation temperatures had, on average, 0.42 S.D. larger trait values than cold-incubated eggs. Effects later in age (>365 days) tended to show a similar positive mean effect (i.e. hot > cold), although the late age category did not differ significantly from zero.

(5) Publication bias and methodological considerations

There was little evidence of publication bias across trait types and effects were symmetrically distributed around the mean (Fig. 8A). The intercept from Egger's regression, however, was significant (0.25, 95% CI = 0.12–0.37), indicating possible publication bias. Even so, this effect was small and statistical significance was likely due to our large sample size. Trim-and-fill analyses support a lack of publication bias, and suggested that no effects were missing on the left or right side of the funnel plot ($P = 0.5$). While there was a very slight

tendency for effect size to decrease across years, credible intervals overlapped zero (−0.08; 95% CI = −0.14 to 0).

The data set contained 643 effect sizes provided as least-square-adjusted means and therefore accounting for covariates (e.g. body size, egg mass), whereas the vast majority of estimates were raw descriptive statistics ($N = 5797$). Effect sizes calculated with raw descriptive statistics were not different from effect sizes calculated with adjusted values (parameter estimate = 0; 95% CI = −0.13 to 0.11).

Most study designs consisted of splitting all eggs within clutches across incubation treatments (i.e. 'split-clutch pseudo'; $N = 126$); fewer studies split clutches such that only one egg from a clutch was in each incubation treatment (i.e. 'split-clutch'; $N = 17$). Three studies were classified as having 'independent-pseudo' designs and the last 29 studies could not be classified (i.e. 'other'). Study design did not impact overall effect-size magnitude, although there was a slight tendency for split clutch and independent pseudo study designs to generate positive effect sizes (Fig. 8B).

IV. DISCUSSION

The impact of reptilian incubation environments on phenotypes remains an important research topic with evolutionary and conservation implications. While the general conclusion from qualitative reviews is that

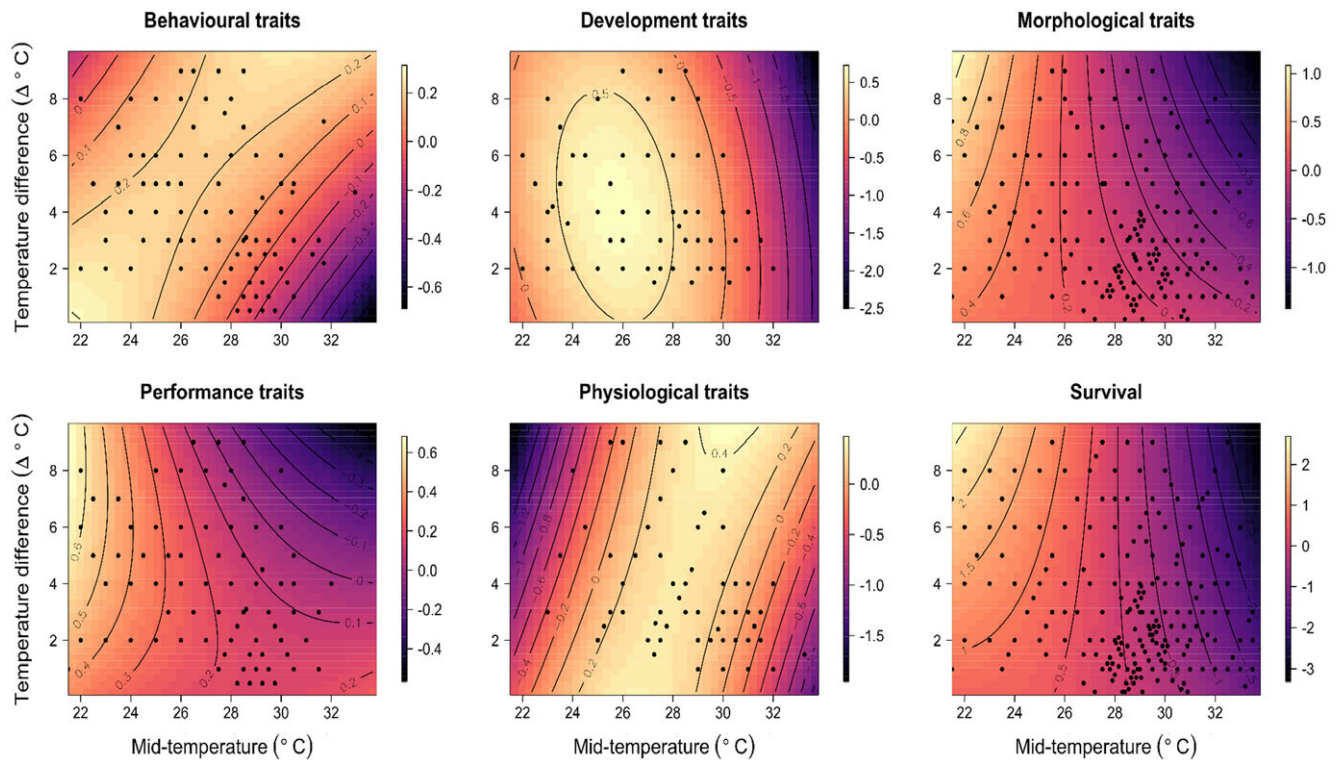


Fig. 6. Predictions of how mid-temperature and temperature difference impact effect size (g) between hot and cold incubation treatments for each trait category. Predictions are made over ± 2 S.D. of the mean temperature difference between treatments and their mid-temperature. Dots indicate actual mid-temperature and temperature difference combinations for each trait category. Predictions are marginalized over the random effects. Effect size is presented in the colour key.

temperature can have dramatic consequences on phenotypes (Birchard, 2004; Booth, 2004; Deeming, 2004), effective generalization across contexts has been hindered by the lack of quantitative syntheses. Here, we provide the first quantitative synthesis on the influence of incubation environments on reptile phenotypes across 92 different species representing all orders; attempting to understand general patterns and gain new insights. We have clarified the overall strength of effects as well as current gaps and limitations in this field. The insights generated should set the foundation for future research directions, essential if we are to appreciate more fully how early thermal environments orchestrate phenotypic development.

Our results are largely congruent with those of Deeming (2004) in that incubation environments can impact offspring phenotypes in important, but in many cases unpredictable, ways across reptiles. Because the reaction norms of phenotype *versus* temperature are likely to be nonlinear, examining the ‘effect’ (*via* meta-analysis) of increasing incubation temperatures is challenging given the existence of both positive and negative effects (both increasing and decreasing phenotypes). Similar challenges arise for any synthetic approach to nonlinear relationships (Kingsolver *et al.*, 2012). Here, we have taken the dual approach of considering both the raw and absolute value of a pairwise effect size.

The raw effect size includes information on the direction of the effect, and thus allows clearer interpretation of the shape of the underlying reaction norm. For example, the existence of positive effects at cold temperatures and negative effects at warm temperatures is suggestive that intermediate temperatures may be more optimal for the phenotype (Fig. 1B). Of course, this would need to be validated by testing whether fitness is also maximized at these temperatures. However, focusing on the overall mean of the raw effect sizes gives the false impression that incubation temperature has a negligible impact on phenotypes (see Fig. 3, except incubation duration). By contrast, considering the absolute value of effect size estimates ($|g|$) reveals that the magnitude of the impact of temperature across many traits was in the proximity of what Cohen (1988) considered large effects (i.e. $g > 0.8$) (Fig. 2). Moreover, examining the magnitude of the effect revealed the importance of the difference in temperature treatments and informs comparisons across ages and trait categories.

(1) Differential effects of incubation temperatures across traits

Developmental plasticity should result in traits responding to early thermal environments differently, yielding a characteristic phenotypic change (or lack of change) as embryos experience different thermal environments (i.e.

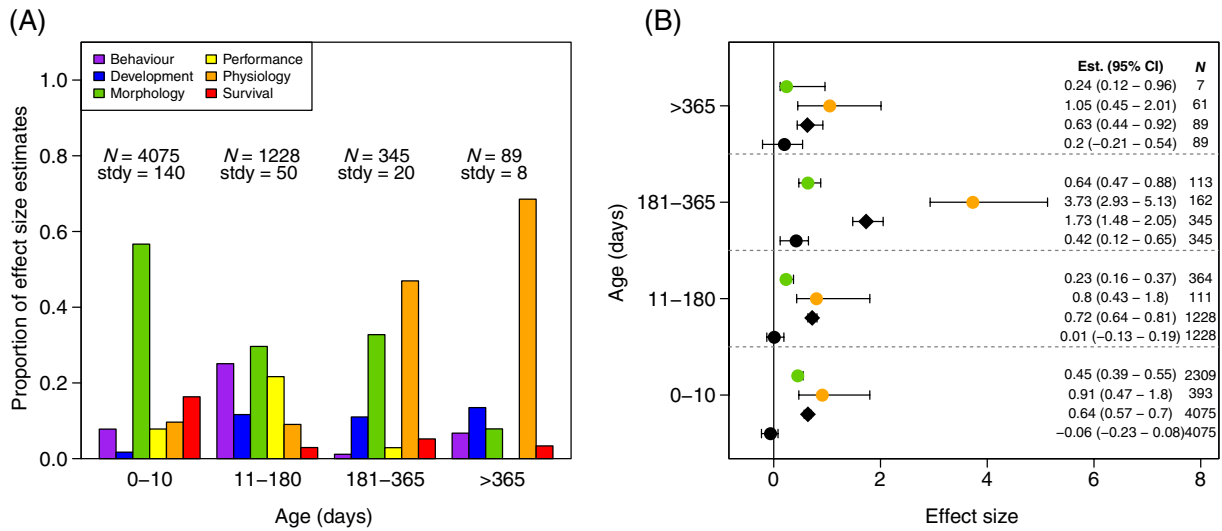


Fig. 7. (A) Proportion of effect sizes across trait and age categories. N is the total number of effects in each age category and 'stdy' is the total number of studies. (B) Weighted mean effect size (g , black circles) and absolute effect size ($|g|$, black diamonds) across age. Average absolute effect size is also provided for morphological ($|g|$, green circles) and physiological ($|g|$, orange circles) traits for which sufficient data across all age categories existed (trait colours match legend in (A)). Values are estimated means for a mixed-sex sample. Estimates (95% credible intervals) are also provided.

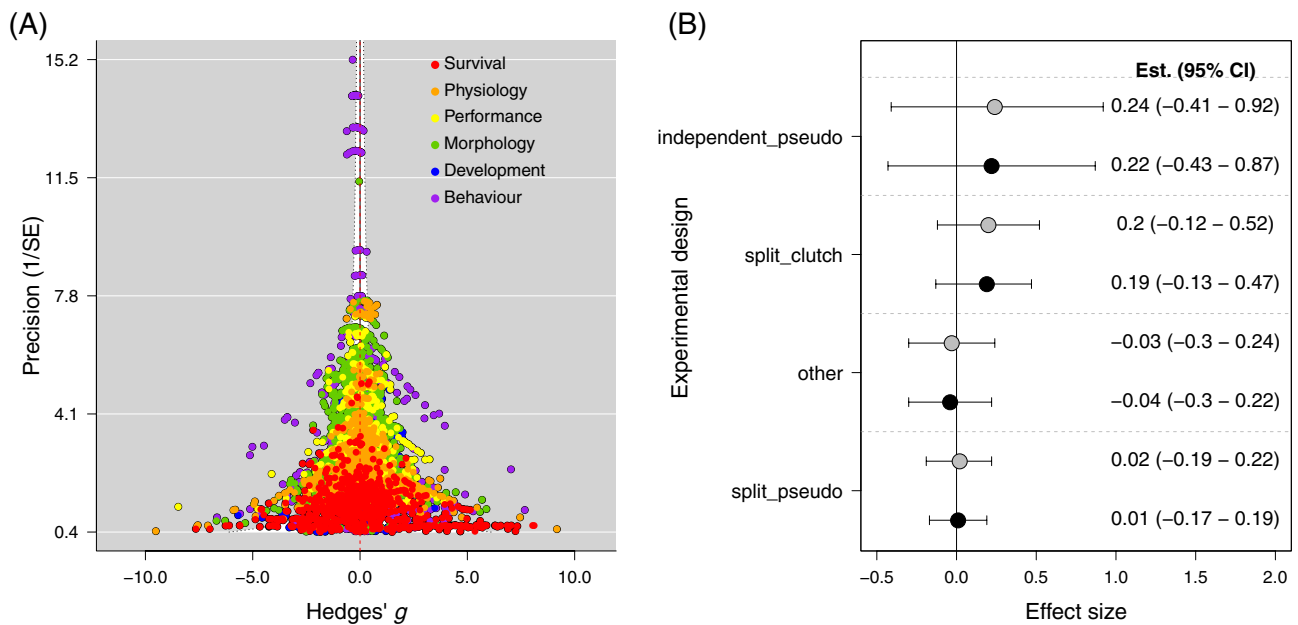


Fig. 8. (A) Funnel plot across all traits using meta-analytic residuals from multi-level meta-analytic model controlling for species, study and phylogenetic variance. 99% sampling confidence intervals (white area) indicate the deviation in effects expected from sampling variability. (B) Predicted mean effect size for a mixed-sex comparison for different experimental designs. Circles indicate raw (black) and adjusted (grey) means. See Section II.3 for definitions of the different experimental designs.

thermal reaction norm; Angilletta *et al.*, 2003). Reaction norms may show characteristic non-linear or linear patterns or be constrained over a wide thermal range. In general, how traits respond to incubation environment should depend on how past selection has shaped the thermal reaction norm, and is thus difficult to predict *a priori* when only a portion of the thermal reaction norm is characterized; as was the case across most studies in our meta-analysis.

Nonetheless, some traits were more strongly impacted than others, a result that was observed across orders. Specifically, the magnitude of survival was quite large across all orders, but particularly so for Squamata and Testudines. Behaviour and physiology were also strongly impacted by incubation temperature in Squamata suggesting that incubation temperatures can impact the development of the hatching phenotype in important, yet unpredictable, ways.

To understand these patterns more fully, it will be important to characterize a greater range of incubation temperatures in the same population, and determine how developmental systems change in response to temperature. Such studies are necessary in order to make more definitive broad-scale comparisons on the impact of incubation temperatures on phenotypes.

Temperature and its impact on different traits explained 10 and 29% of the variability in effects in $|g|$ and g , respectively. There was substantial variation among trait categories in how the temperature regime (i.e. mid-temperature, temperature difference and fluctuation) impacted phenotypes. We hypothesize that this pattern primarily reflects variation in the phenotypic reaction norms in response to temperature. While reaction norm shape is best examined with the trait values themselves, we can gain insight using our pairwise comparisons between mean phenotype at different temperatures. Most notably, the parameter estimates and figures support a nonlinear, declining curve shape for incubation duration as incubation temperature increases (Fig. 1A). For survival, performance and morphology, the results are congruent with expectations of a concave curve, with poorest phenotypic values at the coldest and hottest temperatures (Fig. 1B). We expect that direct analysis of reaction norms across species will demonstrate these consistent patterns where sufficient data are available.

In all cases except behaviour, a greater increase in temperature difference between treatments led to a greater impact on phenotypes. In addition, fluctuation in temperatures tended to diminish the impact of increased temperatures in laboratory-based incubation experiments, but this effect was weak and our confidence in examining the effect would greatly increase with additional data on temperature fluctuations and a greater understanding of embryo traits that maintain development despite fluctuations (Du & Shine, 2015). In particular, we might expect that the impact would diminish as fluctuations increased on a continuous scale. Due to small sample sizes of ‘non-constant’ treatments, we categorized treatments into ‘constant’ and ‘fluctuating’, which may mask the impact of very high fluctuations. Natural nest temperatures are known to fluctuate strongly and increased thermal variation is predicted to increase with a warming climate (Vázquez *et al.*, 2017), suggesting that the effects of incubation temperature on phenotypes in the wild may be weaker than that reported here.

It is worth noting that these patterns were apparent despite the fact that we examined numerous species, which may have species-specific reaction norms, and based on studies with varied temperature treatments. Considering the data in aggregate in this way highlights central aspects of evolutionary ecology, particularly selection on maternal nesting behaviour (Mainwaring *et al.*, 2016). Given that extreme temperatures appear to impact hatchling survival and phenotype the most, maternal nesting behaviour may reflect selection to avoid extreme (particularly extremely hot) nest temperatures (Refsnider & Janzen, 2010).

(2) Long-term effects of incubation temperatures on reptile phenotypes

Early environments have long-lasting effects on phenotypic development across a diverse set of species (Lindström, 1999; West-Eberhard, 2003; Kerr *et al.*, 2007; Monaghan, 2008; Burton & Metcalfe, 2014). However, one might expect the strength of these effects to decrease over time as developmental systems are impacted more by current environmental experiences (Deeming, 2004). We did not find evidence that phenotypic impacts of incubation temperatures diminished over time. By contrast, effects later in life (181–365 days post-hatching) had the strongest impact on phenotypes compared to earlier age categories, particularly for physiological traits (Fig. 7B). While strong effects between 181 and 365 days post-hatching may result from specific studies that happened to measure animals beyond 6 months, they do suggest that the effect of incubation temperature can persist well beyond hatching, supporting a host of studies that have shown persistent effects of early life experiences (reviewed in: Lindström, 1999; Monaghan, 2008; Burton & Metcalfe, 2014). The possibility that incubation effects could even become exacerbated later in life provides a strong motivation for studies collecting data longitudinally in order to test this pattern more robustly.

The potential for incubation effects to persist late into life has particular relevance for adaptive evolutionary hypotheses of sex-determining mechanisms (Shine, 1999b). Across reptiles, some populations have genotypic sex determination (GSD), largely mediated by sex chromosomes, while other populations exhibit temperature-dependent sex determination (TSD), whereby sex is permanently established during embryonic incubation by temperature (Valenzuela & Lance, 2004). Adaptive hypotheses to explain this variation assume that some phenotypic effect established by temperature during embryonic development impacts the fitness of offspring in a sex-specific fashion (Shine, 1999b). While the persistence of effects is not strictly necessary, many adaptive hypotheses assume that phenotype and fitness effects persist into adulthood (Shine, 1999b; Schwanz *et al.*, 2016). Our finding that effects can persist lends plausibility to hypotheses that rely on persistent effects throughout life.

How might incubation temperatures in reptiles induce persistent effects throughout life? Different incubation temperatures in reptiles likely elicit organizational changes to cells and tissues through epigenetic modifications, and thus changes in gene expression. These effects can include permanent histone modification, changes to DNA methylation or effects on non-coding RNA molecules (Burton & Metcalfe, 2014; O’Dea *et al.*, 2016). Indeed, Paredes *et al.* (2016) found that low incubation temperatures in wall lizards (*Podarcis muralis*) result in developmental stress eliciting genome-wide DNA hypomethylation. The important role of temperature in orchestrating developmental shifts is also known for species with TSD, including organizational changes in bearded dragons (*Pogona vitticeps*) that ultimately lead to sex-reversal (Quinn *et al.*, 2007).

Although early developmental environments can have profound organizational effects on the phenotype, the mechanistic basis for these effects is still unclear in most species.

(3) Ecological, evolutionary, parental and methodological mediators of incubation temperature

Our analyses detected much unexplained variability (R^2 values were only between 10 and 42%) that cannot be accounted for by the main biological factors examined here. Total heterogeneity among effects was also large (~90%), indicating that sampling error (i.e. chance) alone cannot simply explain the variability in effect sizes we observed. This level of heterogeneity is not uncommon in most biological meta-analyses (Senior *et al.*, 2016), and suggests that effects may not be easily generalizable – even when accounting for temperature differences among studies. We also found little support for phylogenetically conserved or species-specific responses, as was hypothesized to be important by Deeming (2004, p. 249), although this may change with the addition of more species and more-robust phylogenetic analyses. By contrast, our analyses imply that the particular characteristics of the study (e.g. local environments or particular population attributes) may be more important mediators of the effect of incubation environment on phenotype.

Several potential drivers exist. First, the effects of incubation temperature may be strongly mediated by maternal and/or parental effects, especially given the experimental nature of the studies included in our meta-analysis (i.e. in many cases eggs were not directly exposed to nesting environments). Parental effects, particularly maternal effects, have substantial impact on offspring phenotypes (Mousseau & Fox, 1998; Marshall & Uller, 2007; Räsänen & Kruuk, 2007; Uller, 2008; Noble *et al.*, 2014; Schwanz, 2016). Maternal environment can lead to complex developmental programming, such that mothers adaptively match offspring to local environments, buffer them against environmental stressors, or in the case of a mismatch, even negatively impact offspring fitness (Marshall & Uller, 2007; Miller *et al.*, 2012; Allan *et al.*, 2014). For example, in one viviparous lizard species, the effect of early thermal environment is mediated by female food resource availability (Itonaga, Jones & Wapstra, 2012), and similar processes prior to, or during ovulation, may be at work in oviparous taxa. Parental effects can include changes in yolk hormone concentrations, transfer of mRNA and immunological factors, changes in lipid and protein concentrations in the egg, clutch size and epigenetic programming (Mousseau & Fox, 1998; Morgan *et al.*, 1999; Bowden *et al.*, 2004; Groothuis & Schwabl, 2008; Hesselquist & Nilsson, 2009; Burton & Metcalfe, 2014; Tschirren, 2015). The impact of parental effects on phenotypic development can be complex, particularly when considering the interaction between parental effects and early offspring environment (Stratmann & Taborsky, 2014). These interactions may account for the large unpredictable

effects of incubation environment on offspring phenotype and may even mediate the magnitude of response to early thermal conditions (Miller *et al.*, 2012; Allan *et al.*, 2014). This hypothesis should be considered carefully in the future. Experiments manipulating, in a fully factorial way, both the maternal and offspring environment will go a long way to elucidating whether parental effects play an important role.

Second, local selection pressures are likely responsible for moderating the impact of thermal environments on phenotypes by moulding thermal reaction norms. Indeed, we anticipate variation both among species and populations in thermal reaction norms, some of which may be adaptive. For example, survival curves for green sea turtles (*Chelonia mydas*) differ between populations that nest on black *versus* pale sand beaches (Weber *et al.*, 2012). Selection based on local climate, nesting conditions or post-hatching ecology is expected to shape optimal temperatures, developmental sensitivity and canalization (Angilletta *et al.*, 2003). Different reaction norms between populations would explain the lack of phylogenetic conservatism and species-specific responses in our data, however, we still know very little about how much within- and between-population variation exists in thermal reaction norms and this would be a fruitful area of research.

Third, different life histories and/or demographic structure of populations may also contribute to the unexplained variation in effects across reptilian taxa. Reptile populations can vary substantially in their mode of reproduction, age at sexual maturity, offspring size, and offspring number (Shine, 2005). Life-history traits such as lifespan, generation time and age at sexual maturity – traits falling along a fast–slow pace-of-life continuum – may impact the magnitude of effect by shaping the strength of selection on thermal reaction norms. Populations with short generation times or higher juvenile survival may be predicted to have more-canalized responses to incubation environment, particularly if survival is strongly affected. By contrast, populations with greater longevity, later age of sexual reproduction and lower juvenile survival may exhibit greater plasticity because early-life effects on phenotypes may be offset through post-hatching development. However, we acknowledge that arguments for the opposite prediction could be made, and clear *a priori* hypotheses regarding the specific role of life-history and demographic effects on thermal reaction norms is not entirely clear. While these traits may have some explanatory power, we were not able to extract sufficient detail from most studies to test whether population life history or demography play a role in explaining effect-size heterogeneity. Additionally, it is unlikely that using species-level data would be sufficient given that variation in effect-size estimates could not be explained by between-species differences (i.e. small variance estimate for species-level random effect). Future comparative work will be necessary to understand the role of population-level characteristics in mediating the impact of incubation temperature.

Finally, effect-size heterogeneity can result from methodological differences between studies (e.g. different incubators, experimental design, moisture levels) (Higgins

& Thompson, 2002; Higgins *et al.*, 2003). However, while we were not able to include moderators for all aspects of experimental design, we did not find evidence that certain experimental designs or statistical approaches (i.e. raw *versus* adjusted statistics) led to different overall effects. Nonetheless, methodological differences such as moisture levels, which have a major impact on phenotypes in reptiles with flexible-shelled eggs, may be a possible methodological factor mediating the strength of effect and should be investigated in future meta-analyses (Booth, 2004).

(4) Climate change and conservation implications

Climate change is expected to have important, and often negative, consequences for ectotherms (Deutsch *et al.*, 2008; Sinervo *et al.*, 2010; Munday *et al.*, 2013). Nest temperatures (and their thermal variance) are predicted to increase with global climate change, impacting reptile phenotype and fitness (Hays *et al.*, 2003; Mitchell *et al.*, 2008; Vázquez *et al.*, 2017). The ability to cope with climate stressors is thought to be mediated primarily through phenotypically plastic responses, allowing populations to shift phenotypic distributions to new optima (Visser, 2008; Chevin, Lande & Mace, 2010; Merilä, 2012). Behavioural compensation by egg-laying females (e.g. the use of shaded areas for nests or deeper nest sites; Schwanz & Janzen, 2008; Telemeco *et al.*, 2009; Refsnider *et al.*, 2013; Mainwaring *et al.*, 2016) or interactions between thermal environments and parental effects [e.g. see Miller *et al.* (2012) in response to CO₂] have the potential to dampen or exacerbate negative effects. However, it is largely unclear whether this will be the case. Based on our results, increased nest temperatures will impact populations through three likely pathways: phenology, hatching success, and post-hatching phenotypic effects.

First, the extraordinarily large effect of incubation temperature on incubation duration will alter the phenology of hatchings, strongly advancing hatching date under warming temperatures. However, it is important to note that the magnitude of this effect levels off as very hot temperatures are reached. Early hatching often has positive impacts on post-hatching fitness (Warner & Shine, 2007), although the potential for a mismatch with the seasonal timing of resource availability and predation pressure could create unanticipated challenges for populations (Visser & Both, 2005; Willette, Tucker & Janzen, 2005).

Second, we found that survival was strongly impacted by incubation temperature, where records of 'survival' were largely hatching success. In this trait category, we found the largest effect of incubation temperature at cold and hot temperatures, and a lower effect at intermediate temperatures. This finding has important implications for our understanding of the mechanisms contributing to widespread population extinctions observed in oviparous reptiles (Sinervo *et al.*, 2010; Walker, Stuart-Fox & Kearney, 2016), suggesting that more extreme nest temperatures experienced by eggs may also result in increased egg mortality

that could have important consequences on population viability.

Finally, the moderate to strong impacts of incubation temperature on post-hatching phenotypes have potential fitness consequences. For example, performance and morphology were affected more strongly at warm temperatures, suggesting developmental instability at extreme temperatures (Telemeco *et al.*, 2013; Dahlke *et al.*, 2016; Sinclair *et al.*, 2016), whereas behaviour showed decreased effects. Information about which phenotypes are most closely tied to fitness will be necessary to appreciate fully the consequences of phenotypic change on fitness and population viability. Additionally, incubation temperatures can also have strong impacts on hatchling sex ratios, and this combined effect needs to be incorporated for species with TSD (Mitchell *et al.*, 2008; Holleley *et al.*, 2015; Boyle *et al.*, 2016).

Climate change is only one of many factors leading to reptile declines worldwide. Habitat destruction along with persecution and illegal (or legal) harvesting for medicine, food and/or the pet trade are recognized as major threats (Gibbons *et al.*, 2000; Böhm *et al.*, 2013). Alleviating these stressors on populations is a major goal of conservation initiatives. Some of these initiatives rely heavily on captive breeding and management programs to bolster populations (Gibbons *et al.*, 2000). Manipulating early environments has thus been suggested as an important tool to modify phenotypes for positive conservation outcomes (Crews *et al.*, 1994). It is easy to see the practical advantages of such an approach; modifying early thermal environments is an easy way to allow commercial or conservation organizations to produce larger, more robust offspring with faster growth rates and increased survival. Although this approach may seem useful and practical (e.g. for species with TSD), our analysis troublingly shows that incubation temperatures can produce unpredictably strong effects on phenotypes. Additionally, the long-term consequences of such manipulations are still far from clear. It is therefore important that such programs demonstrate repeatable developmental outcomes of incubation temperatures given a set of experimental conditions and try to understand the long-term fitness consequences of these manipulations if such approaches are to be a success.

(5) Future research directions

Future work will be needed to understand fully the causal biotic and abiotic factors mediating the effects of incubation temperatures on reptile phenotypes and survival. Additional data, more detailed experiments and new analytical techniques will no doubt shed important insights, but it is clear that much work is needed to clarify the developmental consequences of thermal environments for fitness and population persistence. Elucidating such effects has great urgency in order to appreciate the impact of climate change on reptile populations worldwide, and to ensure consistent success for conservation programs. We see a number of extremely important research questions

worth pursuing: (i) how do phenotypic reaction norms vary across taxa, populations and traits (survival and performance may be particularly useful targets)? (ii) Do transgenerational parental effects, including *via* epigenetic, hormonal and RNA-based mechanisms, exacerbate or dampen the impact of incubation temperatures on hatching phenotypes? (iii) For how long do effects of incubation temperature on phenotypes manifest and what are their consequences for long-term fitness? (iv) Do the long-term impacts of incubation temperature differ across traits and, if so, why? (v) How will natural nest temperatures (mean and their variance) be impacted by climate change?

This research will be facilitated by gathering more data on thermal reaction norms (or thermal performance curves) across a wide range of plausible temperatures and a number of independent and functionally relevant phenotypic traits (e.g. escape behaviour, performance, metabolism). Population-level comparative analyses may be particularly fruitful in elucidating answers to the above questions.

V. CONCLUSIONS

(1) The magnitude of the phenotypic effect of incubation temperature is moderate to large across orders, trait categories and ages. There is no evidence that this effect is substantially larger in any single order of reptiles, although data are sparse for Rhynchocephalia (a species-poor order) and Crocodylia.

(2) Effects of incubation temperature can persist for many months post-hatching. Sampling is poor for ages >1 year, thus more data would be useful in increasing our confidence in the persistence of effects.

(3) The effect of temperature on incubation duration is much stronger than on any other trait category. Survival also stands out with particularly strong effect sizes, while the relative strength of other trait categories varies in ways that compels more detailed comparison of reaction norms.

(4) Temperature fluctuations in the incubation environment potentially decrease the phenotypic effect of different mean temperatures, particularly when the temperature differences between treatments are large (although not significantly). More data are needed from fluctuating temperature regimes to assess more rigorously whether this tendency is real, and to quantify the impact of increasing fluctuation.

(5) On average, increased temperature changes lead to greater phenotypic effects. Despite expectations that the exact impact of warming incubation temperatures will depend on the trait studied and the shape of the reaction norm, we can say that, on average, nest temperatures that increase by 4°C would have a greater impact on nearly all phenotypes than would an increase of 2°C.

(6) The effect of increased incubation temperature depends on the temperatures experienced (mid-temperature), and this dependence varies according to trait type. Survival, morphology and performance were affected more strongly

at extreme temperatures compared to intermediate temperatures indicating that increasingly warmer nest temperatures will accelerate change in these traits. Thus, collecting phenotypic data from extreme incubation temperatures is important.

(7) Substantial variation in the magnitude and direction of the phenotypic effects of incubation temperature remain unexplained. Future research should quantify the shape of the reaction norm to explore interspecific variation along with how parental and/or ecological effects might mediate responses.

VI. ACKNOWLEDGEMENTS

We thank Shinichi Nakagawa, Alistair Senior and Malgorzata Lagisz for helpful code and discussion, and two anonymous reviewers along with R. Andrews, F. Janzen, S. Nakagawa, and A. Senior for feedback on previous versions of the manuscript. We also thank Tobias Uller for insightful discussions and a host of colleagues who provided data or clarification on their papers including: Robin Andrews, Geoffrey Birchard, Mark Blumberg, Gabriel Blouin-Demers, David Booth, David Crews, Sharyon Downes, Sean Doody, Wei-Guo Du, Leah Fisher, Scott Freedberg, Mattias Hagman, Duncan Irschick, Xiang Ji, Day Ligon, Nicola Mitchell, Gary Packard, Ryan Paitz, Turk Rhen, Tania Rodriguez-Diaz, Rick Shine, Tony Steyermark, Rory Telemeco, Tobias Uller, Nicole Valenzuela, Dan Warner, and Grahame Webb. The research was supported by an ARC Discovery Early Career Research Award (LES: DE120101446) and a start-up grant from the School of Biological, Earth, and Environmental Sciences at UNSW (L.E.S.). D.W.A.N. was supported by an ARC Discovery Early Career Research Award (DE150101774) and a UNSW VC Research Fellowship.

VII. REFERENCES

- References marked with an asterisk were included in the meta-analysis and in the supporting information.*
- *ALBERTS, A. C., PERRY, A. M., LEMM, J. M. & PHILLIPS, J. A. (1997). Effects of incubation temperature and water potential on growth and thermoregulatory behavior of hatchling Cuban rock iguanas (*Cyclura nubila*). *Copeia* **1997**, 766–776.
 - ALLAN, B. J. M., MILLER, G. M., McCORMICK, M. I., DOMENICI, P. & MUNDAY, P. L. (2014). Parental effects improve escape performance of juvenile reef fish in a high-CO₂ world. *Proceedings of the Royal Society B: Biological Sciences* **281**, 1–7.
 - *AMIEL, J. J. & SHINE, R. (2012). Hotter nests produce smarter young lizards. *Biology Letters* **8**, 372–374.
 - *ANDREWS, R. M. (2005). Incubation temperature and sex ratio of the Veiled Chameleon (*Chamaeleo calyptratus*). *Journal of Herpetology* **39**, 515–518.
 - *ANDREWS, R. M., MATHIES, T., QUALLS, C. P. & QUALLS, F. J. (1999). Rates of embryonic development of *Sceloporus* lizards: do cold climates favor the evolution of rapid development? *Copeia* **1999**, 692–700.
 - *ANDREWS, R. M., MATHIES, T. & WARNER, D. A. (2000). Effect of incubation temperature on morphology, growth, and survival of juvenile *Sceloporus undulatus*. *Herpetological Monographs* **14**, 420–431.
 - ANGILLETTA, M. J., WILSON, R. S., NAVAS, C. A. & JAMES, R. S. (2003). Trade-offs and the evolution of thermal reaction norms. *Trends in Ecology & Evolution* **18**, 234–240.

- *ANGILLETTA, M. J., WINTERS, R. S. & DUNHAM, A. E. (2000). Thermal effects on the energetics of lizard embryos: implications for hatchling phenotypes. *Ecology* **81**, 2957–2968.
- *ARNOLD, S. J. & PETERSON, C. R. (2002). A model for optimal reaction norms: the case of the pregnant garter snake and her temperature sensitive embryos. *The American Naturalist* **160**, 306–316.
- BADYAEV, A. V. (2002). Growing apart: an ontogenetic perspective on the evolution of sexual size dimorphism. *Trends in Ecology & Evolution* **17**, 369–378.
- *BEAL, A. & WEBB, G. (1989). Effect of incubation temperature on tooth eruption in embryos of *Crocodylus johnstoni*. *Copeia* **1989**, 325–331.
- *BESSON, A. A., NELSON, N. J., NOTTINGHAM, C. M. & CREE, A. (2012). Is cool egg incubation temperature a limiting factor for the translocation of tuatara to southern New Zealand? *New Zealand Journal of Ecology* **36**, 90–99.
- *BEUCHAT, C. A. (1988). Temperature effects during gestation in a viviparous lizard. *Journal of Thermal Biology* **13**, 135–142.
- *BILCKE, J., DOWNES, S. & BUSCHER, I. (2006). Combined effect of incubation and ambient temperature on the feeding performance of a small ectotherm. *Austral Ecology* **31**, 937–947.
- BIRCHARD, G. F. (2004). Effects of incubation temperature. In *Reptilian Incubation: Environment, Evolution and Behaviour* (ed. D. C. DEEMING), pp. 103–123. Nottingham University Press, Nottingham.
- *BIRCHARD, G. F. & REIBER, C. L. (1995). Growth, metabolism, and chorioallantoic vascular density of developing Snapping Turtles (*Chelydra serpentina*): influence of temperature. *Physiological Zoology* **68**, 799–811.
- *BIRCHARD, G. F. & REIBER, C. L. (1996). Heart rate during development in the turtle embryo: effect of temperature. *Journal of Comparative Physiology B* **166**, 461–466.
- *BLOUIN-DEMERS, G., KISSNER, K. J. & WEATHERHEAD, P. J. (2000). Plasticity in preferred body temperature of young snakes in response to temperature during development. *Copeia* **2000**, 841–845.
- *BLOUIN-DEMERS, G., WEATHERHEAD, P. J. & ROW, J. R. (2004). Phenotypic consequences of nest-site selection in black rat snakes (*Elaphe obsoleta*). *Canadian Journal of Zoology* **82**, 449–456.
- *BLUMBERG, M. S., LEWIS, S. J. & SOKOLOFF, G. (2002). Incubation temperature modulates post-hatching thermoregulatory behavior in the Madagascar ground gecko, *Pareodura pictus*. *The Journal of Experimental Biology* **205**, 2777–2784.
- *BOBYN, M. L. & BROOKS, R. J. (1994). Incubation conditions as potential factors limiting the northern distribution of snapping turtles, *Chelydra serpentina*. *Canadian Journal of Zoology* **72**, 28–37.
- BÖHM, M., COLLEN, B., BAILLIE, J. E., BOWLES, P., CHANSON, J., COX, N., HAMMERSON, G., HOFFMANN, M., LIVINGSTONE, S. R. & RAM, M. (2013). The conservation status of the world's reptiles. *Biological Conservation* **157**, 372–385.
- *BOOTH, D. T. (1998). Effects of incubation temperature on the energetics of embryonic development and hatchling morphology in the Brisbane river turtle *Emydura signata*. *Journal of Comparative Physiology B* **168**, 399–404.
- *BOOTH, D. T. (2000). Incubation of eggs of the Australian broad-shelled turtle, *Chelodina expansa* (Testudinata: Chelidae), at different temperatures: effects on pattern of oxygen consumption and hatchling morphology. *Australian Journal of Zoology* **48**, 369–378.
- BOOTH, D. T. (2004). Post-hatching phenotypic effects of incubation in reptiles. In *Reptilian Incubation: Environment, Evolution and Behaviour* (ed. D. C. DEEMING), pp. 229–251. Nottingham University Press, Nottingham.
- BOOTH, D. T. (2006). Influence of incubation temperature on hatchling phenotype in reptiles. *Physiological and Biochemical Zoology* **79**, 274–281.
- *BOOTH, D. T., BURGESS, E., MCCOSKER, J. & LANYON, J. M. (2004). The influence of incubation temperature on post-hatching fitness characteristics of turtles. *International Congress Series* **1275**, 226–233.
- *BOOTH, D. T., THOMPSON, M. B. & HERRING, S. (2000). How incubation temperature influences the physiology and growth of embryonic lizards. *Journal of Comparative Physiology B* **170**, 269–276.
- BORENSTEIN, M., HEDGES, L. V., HIGGINS, J. P. T. & ROTHSTEIN, H. R. (2009). *Introduction to Meta-analysis*. John Wiley & Sons, Ltd, West Sussex.
- BOWDEN, R. M., HARMS, H. K., PAITZ, R. T. & JANZEN, F. J. (2004). Does optimal egg size vary with demographic stage because of a physiological constraint? *Functional Ecology* **18**, 522–529.
- BOYLE, M., HONE, J., SCHWANZ, L. E. & GEORGES, A. (2014). Under what conditions do climate-driven sex ratios enhance versus diminish population persistence? *Ecology and Evolution* **4**, 4522–4533.
- BOYLE, M., SCHWANZ, L. E., HONE, J. & GEORGES, A. (2016). Dispersal and climate warming determine range shift in model reptile populations. *Ecological Modelling* **328**, 34–43.
- *BRANA, F. & JI, X. (2000). Influence of incubation temperature on morphology, locomotor performance, and early growth of hatchling Wall lizards (*Podarcis muralis*). *Journal of Experimental Zoology* **286**, 422–433.
- *BROOKS, R. J., BOBYN, M. L., GALBRAITH, D. A., LAYFIELD, J. A. & NANCEKIVELL, E. G. (1991). Maternal and environmental influences on growth and survival of embryonic and hatchling snapping turtles (*Chelydra serpentina*). *Canadian Journal of Zoology* **69**, 2667–2676.
- *BUCKLEY, C. R., JACKSON, M., YOUSSEF, M., IRSCHICK, D. J. & ADOLPH, S. C. (2007). Testing the persistence of phenotypic plasticity after incubation in the western fence lizard, *Sceloporus occidentalis*. *Evolutionary Ecology Research* **9**, 169–183.
- *BURGER, J. (1989). Incubation temperature has long-term effects on behaviour of young Pine snakes (*Pituophis melanoleucus*). *Behavioral Ecology and Sociobiology* **24**, 201–207.
- *BURGER, J. (1990). Effects of incubation temperature on behavior of young black racers (*Coluber constrictor*) and Kingsnakes (*Lampropeltis getulus*). *Journal of Herpetology* **24**, 158–163.
- *BURGER, J. (1991a). Effects of incubation temperature on behavior of hatchling pine snakes: implications for reptilian distribution. *Behavioral Ecology and Sociobiology* **28**, 297–303.
- *BURGER, J. (1991b). Response to prey chemical cues by hatchling pine snakes (*Pituophis melanoleucus*): effects of incubation temperature and experience. *Journal of Chemical Ecology* **17**, 1069–1078.
- *BURGER, J. (1998a). Antipredator behaviour of hatchling snakes: effects of incubation temperature and simulated predators. *Animal Behaviour* **56**, 547–553.
- *BURGER, J. (1998b). Effects of incubation temperature on hatchling pine snakes: implications for survival. *Behavioral Ecology and Sociobiology* **43**, 11–18.
- *BURGER, J. & ZAPPALORTI, R. T. (1988). Effects of incubation temperature on sex ratios in pine snakes: differential vulnerability of males and females. *The American Naturalist* **132**, 492–505.
- *BURGER, J., ZAPPALORTI, R. T. & GOCHFELD, M. (1987). Developmental effects of incubation temperature on hatchling pine snakes (*Pituophis melanoleucus*). *Comparative Biochemistry and Physiology Part A Physiology* **87**, 727–732.
- BURTON, T. & METCALFE, N. B. (2014). Can environmental conditions experienced in early life influence future generations? *Proceedings of the Royal Society B: Biological Sciences* **281**, 1–8.
- *CADBY, C. D., JONES, S. A. & WAPSTRA, E. (2014). Geographical differences in maternal basking behaviour and offspring growth rate in a climatically widespread viviparous reptile. *Journal of Experimental Biology* **217**, 1175–1179.
- *CALSBEEK, R. & SINERVO, B. (2007). Correlational selection on lay date and life-history traits: experimental manipulations of territory and nest site quality. *Evolution* **61**, 1071–1083.
- CHAMBERLAIN, S. A., HOVIK, S. M., DIBBLE, C. J., RASMUSSEN, N. L., VAN ALLEN, B. G., MAITNER, B. S., AHERN, J. R., BELL-DERESKE, L. P., ROY, C. L., MEZA-LOPEZ, M., CARRILLO, J., SIEMAN, E., LAJEUNESSE, M. J. & WHITNEY, K. D. (2012). Does phylogeny matter? Assessing the impact of phylogenetic information in ecological meta-analysis. *Ecology Letters* **15**, 627–636.
- *CHEN, Y.-H., HUANG, S.-P., CHANG, M.-H. & TU, M.-C. (2010). Thermal effects on embryogenesis and hatchlings of the Grass Lizard *Takydromus stejnegeri* (Squamata: Lacertidae) and implications of their potential for limiting its altitudinal distribution in Taiwan. *Zoological Studies* **49**, 374–380.
- *CHEN, Y., ZHAO, B., SUN, B.-J., WANG, Y. & DU, W.-G. (2011). Between-population variation in body size and growth rate of hatchling Asian yellow pond turtles, *Mauremys mutica*. *Herpetological Journal* **21**, 113–116.
- CHEVIN, L.-M., LANDE, R. & MACE, G. M. (2010). Adaptation, plasticity and extinction in a changing environment: towards a predictive theory. *PLoS Biology* **8**, e1000357.
- *CHOO, B. & CHOU, L. (1987). Effect of temperature on the incubation period and hatchability of *Trionyx sinensis* weigmann eggs. *Journal of Herpetology* **21**, 230–232.
- *CLARK, B. F., AMIEL, J. J., SHINE, R., NOBLE, D. W. A. & WHITING, M. J. (2014). Colour discrimination and associative learning in hatchling lizards incubated at 'hot' and 'cold' temperatures. *Behavioral Ecology and Sociobiology* **68**, 239–247.
- COHEN, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Hillsdale.
- *COOMBER, P., CREWS, D. & GONZALEZ-LIMA, F. (1997). Independent effects of incubation temperature and gonadal sex on the volume and metabolic capacity of brain nuclei in the Leopard Gecko (*Eublepharis macularius*), a lizard with temperature-dependent sex determination. *The Journal of Comparative Neurology* **380**, 409–421.
- CREWS, D., BERGERON, J. M., BULL, J. J., FLORES, D., TOUSIGNANT, A., SKIPPER, J. K. & WIBBELS, T. (1994). Temperature-dependent sex determination in reptiles: proximate mechanisms, ultimate outcomes and practical applications. *Developmental Genetics* **15**, 297–312.
- *CREWS, D., COOMBER, P. & GONZALEZ-LIMA, F. (1997). Effect of age and sociosexual experience on the morphology and metabolic capacity of brain nuclei in the leopard gecko (*Eublepharis macularius*), a lizard with temperature-dependent sex determination. *Brain Research* **758**, 169–179.
- DAHLKE, F. T., POLITIS, S. N., BUTTS, I. A. E., TRIPPEL, E. A. & PECK, M. A. (2016). Fathers modify thermal reaction norms for hatching success in Atlantic cod, *Gadus morhua*. *Journal of Experimental Marine Biology and Ecology* **474**, 148–155.
- DEEMING, D. C. (2004). Post-hatching phenotypic effects of incubation in reptiles. In *Reptilian Incubation: Environment, Evolution and Behaviour* (ed. D. C. DEEMING), pp. 229–252. Nottingham University Press, Nottingham.
- DEEMING, D. C. & FERGUSON, M. W. J. (1991). Physiological effects of incubation temperature on embryonic development in reptiles and birds. In *Egg Incubation: Its Effects on Embryonic Development in Birds and Reptiles* (eds D. C. DEEMING and M. W. J. FERGUSON), pp. 147–171. Cambridge University Press, Cambridge.

- *DEMUTH, J. P. (2001). The effects of constant and fluctuating incubation temperatures on sex determination, growth, and performance in the tortoise *Gopherus polyphemus*. *Canadian Journal of Zoology* **79**, 1609–1620.
- DEUTSCH, C. A., TEWKSBURY, J. J., HUEY, R. B., SHELDON, K. S., GHALAMBOR, C. K., HAAK, D. C. & MARTIN, P. R. (2008). Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 6668–6672.
- *DIAS, B. G., ATAYA, R. S., RUSHWORTH, D., ZHAO, J. & CREWS, D. (2007). Effect of incubation temperature and androgens on dopaminergic activity in the leopard gecko, *Eublepharis macularius*. *Developmental Neurobiology* **67**, 630–636.
- *DING, G.-H., YANG, J., WANG, J. & JI, X. (2012). Offspring sex in a TSD gecko correlates with an interaction between incubation temperature and yolk steroid hormones. *Naturwissenschaften* **99**, 999–1006.
- *DOWNES, S. J. & SHINE, R. (1999). Do incubation-induced changes in a lizard's phenotype influence its vulnerability to predators? *Oecologia* **120**, 9–18.
- *DU, W.-G. & FENG, J.-H. (2008). Phenotypic effects of thermal mean and fluctuations on embryonic development and hatchling traits in a lacertid lizard, *Takydromus septentrionalis*. *Journal of Experimental Zoology* **309A**, 138–146.
- *DU, W., HU, L., LU, J. & ZHU, L. (2007). Effects of incubation temperature on embryonic development rate, sex ratio and post-hatching growth in the Chinese three-keeled pond turtle, *Chinemys reevesii*. *Aquaculture* **272**, 747–753.
- *DU, W.-G. & JI, X. (2003). The effects of incubation thermal environments on size, locomotor performance and early growth of hatchling soft-shelled turtles, *Pelodiscus sinensis*. *Journal of Thermal Biology* **28**, 279–286.
- *DU, W.-G. & JI, X. (2006). Effects of constant and fluctuating temperatures on egg survival and hatchling traits in the Northern Grass Lizard (*Takydromus septentrionalis*, Lacertidae). *Journal of Experimental Zoology* **305A**, 47–54.
- *DU, W.-G., JI, X., ZHANG, Y.-P., LIN, Z.-H. & XU, X.-F. (2010a). Geographic variation in offspring size of a widespread lizard (*Takydromus septentrionalis*): importance of maternal investment. *Biological Journal of the Linnean Society* **101**, 59–67.
- *DU, W.-G., WARNER, D. A., LANGKILDE, T., ROBBINS, T. & SHINE, R. (2010b). The physiological basis of geographic variation in rates of embryonic development within a widespread lizard species. *The American Naturalist* **176**, 522–528.
- DU, W.-G. & SHINE, R. (2015). The behavioural and physiological strategies of bird and reptile embryos in response to unpredictable variation in nest temperature. *Biological Reviews* **90**, 19–30.
- *DU, W.-G., SHOU, L. & LIU, J. K. (2003). The effect of incubation temperature on egg survival, hatchling traits and embryonic use of energy in the blue-tailed skink, *Eumeces elegans*. *Animal Biology* **53**, 27–36.
- *DU, W., ZHENG, R. & SHU, L. (2006). The influence of incubation temperature on morphology, locomotor performance, and cold tolerance of hatchling Chinese Three-Keel pond turtles, *Chinemys reevesii*. *Chelonian Conservation and Biology* **5**, 294–299.
- DUVAL, S. & TWEEDIE, R. (2000). Trim and fill: a simple funnel-plot-based method for testing and adjusting for publication bias in meta-analysis. *Biometrics* **56**, 455–463.
- EGGER, M., SMITH, G. D., SCHNEIDER, M. & MINDER, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* **315**, 629–634.
- *EL MOUDEN, E. H., ZNARI, M. & PIEAU, C. (2001). Effects of incubation temperature on embryonic development and sex determination in the north African agamid lizard, *Agama impaleensis*. *Herpetologica* **11**, 101–108.
- *ELPHICK, M. J. & SHINE, R. (1998). Longterm effects of incubation temperatures on the morphology and locomotor performance of hatchling lizards (*Bassiana duperoyi*, Scincidae). *Biological Journal of the Linnean Society* **63**, 429–447.
- *ELPHICK, M. J. & SHINE, R. (1999). Sex differences in optimal incubation temperatures in a scincid lizard species. *Oecologia* **118**, 431–437.
- *ESQUERRE, D., KEOGH, J. & SCHWANZ, L. (2014). Direct effects of incubation temperature on morphology, thermoregulatory behaviour and locomotor performance in jacky dragons (*Amphibolurus muricatus*). *Journal of Thermal Biology* **43**, 33–39.
- *ETCHBERGER, C. R., EWERT, M. A., PHILLIPS, J. B., NELSON, C. E. & PRANGE, H. D. (1993). Environmental and maternal influences on embryonic pigmentation in a turtle (*Trachemys scripta elegans*). *Journal of Zoology* **230**, 529–539.
- *FISHER, L., GODFREY, M. & OWENS, D. (2014). Incubation temperature effects on hatchling performance in the Loggerhead Sea turtle (*Caretta caretta*). *PLoS ONE* **9**, 1–22.
- *FLATT, T., SHINE, R., BORGES-LANDEZ, P. A. & DOWNES, S. J. (2001). Phenotypic variation in an oviparous montane lizard (*Bassiana duperoyi*): the effects of thermal and hydric incubation environments. *Biological Journal of the Linnean Society* **74**, 339–350.
- *FREEDBERG, S., GREIVES, T. J., EWERT, M. A., DEMAS, G. E., BEECHER, N. & NELSON, C. E. (2008). Incubation environment affects immune system development in a turtle with environmental sex determination. *Journal of Herpetology* **42**, 536–541.
- *FREEDBERG, S., STUMPF, A. L., EWERT, M. A. & NELSON, C. E. (2004). Developmental environment has long-lasting effects on behavioural performance in two turtles with environmental sex determination. *Evolutionary Ecology Research* **6**, 739–747.
- FRITZ, C. O., MORRIS, P. E. & RICHLER, J. J. (2012). Effect size estimates: current use, calculations and interpretation. *Journal of Experimental Psychology: General* **141**, 2–18.
- *GAO, J.-F., QU, Y.-F., LUO, L.-G. & JI, X. (2010). Evolution of reptilian viviparity: test of the maternal manipulation hypothesis in a temperate snake, *Gloydus brevicaudus* (Viperidae). *Zoological Science* **27**, 248–255.
- GELMAN, A. & RUBIN, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457–472.
- *GEORGES, A. (1988). Sex determination is independent of incubation temperature in another chelid turtle, *Chelodina longicollis*. *Copeia* **1988**, 248–254.
- GEORGES, A., BEGGS, K., YOUNG, J. E. & DOODY, J. S. (2005). Modelling development of reptile embryos under fluctuating temperature regimes. *Physiological and Biochemical Zoology* **78**, 18–30.
- GIBBONS, J. W., SCOTT, D. E., RYAN, T. J., BULMANN, K. A., TUBERVILLE, T. D., METTS, B. S., GREENE, J. L., MILLS, T., LEIDEN, Y., POPPY, S. & WINNE, C. T. (2000). The global decline of reptiles, déjà vu amphibians. *BioScience* **50**, 653–666.
- GIRONDOT, M., DELMAS, V., RIVALAN, P., COURCHAMP, F., PRÉVOT-JULLIARD, A.-C. & GODFREY, M. H. (2004). Implication of TSD for population dynamics. In *Temperature-dependent Sex Determination in Vertebrates* (eds N. VALENZUELA and V. LANCE). 148–155. Smithsonian Books, Washington.
- GLEN, F. & MROSOVSKY, N. (2004). Antigua revisited: the impact of climate change on sand and nest temperatures at a hawkbill turtle (*Eretmochelys imbricata*) nesting beach. *Global Change Biology* **10**, 2036–2045.
- *GODFREY, M. & MROSOVSKY, N. (2005). Pivotal temperature for green sea turtles, *Chelonia mydas*, nesting in Suriname. *Herpetologica* **16**, 55–61.
- *GOODMAN, R. M. (2008). Latent effects of egg incubation temperature on growth in the lizard *Anolis carolinensis*. *Journal of Experimental Zoology* **309A**, 525–533.
- *GOODMAN, B. A., SCHWARZKOPF, L. & KROCKENBERGER, A. K. (2013). Phenotypic integration in response to incubation environment adaptively influences habitat choice in a tropical lizard. *The American Naturalist* **182**, 666–673.
- *GOODMAN, R. M. & WALGUARNERY, J. W. (2007). Incubation temperature modifies neonatal thermoregulation in the lizard *Anolis carolinensis*. *Journal of Experimental Zoology* **307A**, 439–448.
- GRAFEN, A. (1989). The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **326**, 119–157.
- *GREENBAUM, E. & CARR, J. L. (2001). Sexual differentiation in the spiny softshell turtle (*Apalone spinifera*), a species with genetic sex determination. *Journal of Experimental Zoology* **290**, 190–200.
- GROOTHUIS, T. G. G. & SCHWABL, H. (2008). Hormone-mediated maternal effects in birds: mechanisms matter but what do we know of them? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **363**, 1647–1661.
- *GUO, J.-H., ZHU, X.-P., ZHAO, W.-H., WEI, C.-Q. & CHEN, Y.-L. (2010). Effects of incubation temperature and substrate humidity on embryonic development of *Mauremys mutica*. *Chinese Journal of Applied Ecology* **21**, 215–220.
- *GUTZKE, W. H. N. & CREWS, D. (1988). Embryonic temperature determines adult sexuality in a reptile. *Nature* **332**, 832–834.
- HADFIELD, J. D. (2010). MCMC methods for multi-response generalised linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* **33**, 1–22.
- *HARE, K. M. & CREE, A. (2010). Exploring the consequences of climate-induced changes in cloud cover on offspring of a cool-temperate viviparous lizard. *Biological Journal of the Linnean Society* **101**, 844–851.
- *HARE, K. M., DAUGHERTY, C. H. & CREE, A. (2002). Incubation regime affects juvenile morphology and hatching success, but not sex, of the oviparous lizard *Oligosoma suteri* (Lacertilia: Scincidae). *New Zealand Journal of Zoology* **29**, 221–229.
- *HARE, K. M., LONGSON, C. G., PLEDGER, S. & DAUGHERTY, C. H. (2004). Size, growth and survival are reduced at cool incubation temperatures in the temperate lizard *Oligosoma suteri* (Lacertilia: Scincidae). *Copeia* **2002**, 383–390.
- *HARE, K. M., PLEDGER, S. & DAUGHERTY, C. H. (2008). Low incubation temperatures negatively influence locomotor performance and behavior of the nocturnal lizard *Oligosoma suteri* (Lacertilia: Scincidae). *Copeia* **2008**, 16–22.
- *HARLOW, P. S. (2000). Incubation temperature determines hatchling sex in Australian rock dragons (Agamidae: Genus *Ctenophorus*). *Copeia* **2000**, 958–964.
- *HARLOW, P. S. & SHINE, R. (1999). Temperature-dependent sex determination in the Frillneck lizard, *Chlamydosaurus kingii* (Agamidae). *Herpetologica* **55**, 205–212.
- *HARLOW, P. S. & TAYLOR, J. E. (2000). Reproductive ecology of the jacky dragon (*Amphibolurus muricatus*): an agamid lizard with temperature-dependent sex determination. *Austral Ecology* **25**, 640–652.
- HAWKES, L. A., BRODERICK, A. C., GODFREY, M. H. & GODLEY, B. J. (2007). Investigating the potential impacts of climate change on a marine turtle population. *Global Change Biology* **13**, 923–932.
- HAYS, G. C., BRODERICK, A. C., GLEN, F. & GODLEY, B. J. (2003). Climate change and sea turtles: a 150-year reconstruction of incubation temperatures at a major marine turtle rookery. *Global Change Biology* **9**, 642–646.
- HEREFORD, J., HANSEN, T. F. & HOULE, D. (2004). Comparing strengths of directional selection: how strong is strong? *Evolution* **58**, 2133–2143.
- HESSELQUIST, D. & NILSSON, J.-Å. (2009). Maternal transfer of antibodies in vertebrates: trans-generational effects on offspring immunity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **364**, 51–60.
- HIGGINS, J. P. T. & THOMPSON, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* **12**, 1539–1558.

- HIGGINS, J. P. T., THOMPSON, S. G., DEEKS, J. J. & ALTMAN, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal* **327**, 557–560.
- HINCHCLIFF, C. E., SMITH, S. A., ALLMAN, J. F., BURLEIGH, J. G., CHAUDHARY, R., COGHILL, L. M., CRANDALL, K. A., DENG, J., DREW, B. T., GAZIS, R., GUDE, K., HIBBETT, D. S., KATZ, L. A., LAUGHINGHOUSE, H. D. IV, MCGTAVISH, E. J., MIDFORD, P. E., OWEN, C. L., REE, R. H., REES, J. A., SOLTIS, D. E., WILLIAMS, T. & CRANSTON, K. A. (2015). Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 12764–12769.
- HOLLELEY, C. E., O'MEALLY, D. O., SARRE, S. D., MARSHALL GRAVES, J. A., EZAZ, T., MATSUBARA, K., AZAD, B., ZHANG, X. & GEORGES, A. (2015). Sex reversal triggers the rapid transition from genetic to temperature-dependent sex. *Nature* **523**, 79–82.
- *HUTTON, J. M. (1987). Incubation temperatures, sex ratios and sex determination in a population of Nile crocodiles (*Crocodilus niloticus*). *Journal of Zoology* **211**, 143–155.
- *TRAETA, P., MONASTERIO, C., SALVADOR, A. & DIAZ, J. A. (2006). Mediterranean hatchling lizards grow faster at higher altitude: a reciprocal transplant experiment. *Functional Ecology* **20**, 865–872.
- ITONAGA, K., JONES, S. A. & WAPSTRA, E. (2012). Effects of maternal basking and food quantity during gestation provide evidence for the selective advantage of matrotrophy in a viviparous lizard. *PLoS ONE* **7**, e41835 (<https://doi.org/10.1371/journal.pone.0041835>).
- *JANZEN, F. J. (1993). The influence of incubation temperature and family on eggs, embryos, and hatchlings of the smooth softshell turtle (*Apalone mutica*). *Physiological Zoology* **66**, 349–373.
- JANZEN, F. J. (1994). Climate change and temperature-dependent sex determination in reptiles. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 7487–7490.
- *JANZEN, F. J. (1995). Experimental evidence for the evolutionary significance of temperature dependent sex determination. *Evolution* **49**, 864–873.
- *JANZEN, F. J. & MORJAN, C. L. (2002). Egg size, incubation temperature, and posthatching growth in Painted turtles (*Chrysemys picta*). *Journal of Herpetology* **36**, 308–311.
- *JI, X. & BRANA, F. (1999). The influence of thermal and hydric environments on embryonic use of energy and nutrients, and hatchling traits, in the wall lizards (*Podarcis muralis*). *Comparative Biochemistry and Physiology Part A Physiology* **124**, 205–213.
- *JI, X., CHEN, F., DU, W. G. & CHEN, H. L. (2003). Incubation temperature affects hatchling growth but not sexual phenotype in the Chinese soft-shelled turtle, *Pelodiscus sinensis* (Trionychidae). *Journal of Zoology* **261**, 409–416.
- *JI, X. & DU, W. G. (2001a). The effects of thermal and hydric environments on hatchling success, embryonic use of energy and hatchling traits in a colubrid snake, *Elaphe carinata*. *Comparative Biochemistry and Physiology Part A Physiology* **129**, 461–471.
- *JI, X. & DU, W. G. (2001b). Effects of thermal and hydric environments on incubating eggs and hatchling traits in the Cobra, *Naja naja atra*. *Journal of Herpetology* **35**, 186–194.
- *JI, X., GAO, J.-F. & HAN, J. (2007a). Phenotypic responses of hatchlings to constant versus fluctuating incubation temperatures in the multi-banded krait, *Bungarus multicinctus* (Elapidae). *Zoological Science* **24**, 384–390.
- *JI, X., LIN, C.-X., LIN, L. H., QIU, Q. B. & DU, Y. (2007b). Evolution of viviparity in warm-climate lizards: an experimental test of the maternal manipulation hypothesis. *Journal of Evolutionary Biology* **20**, 1037–1045.
- *JI, X., HUANG, H., HU, X. & DU, W. G. (2002). Geographic variation in female reproductive characteristics and egg incubation of *Eumeces chinensis*. *Chinese Journal of Applied Ecology* **13**, 680–684.
- *JI, X., LIN, L. H., LUO, L.-G., LU, H. L., GAO, J.-F. & HAN, J. (2006). Gestation temperature affects sexual phenotype, morphology, locomotor performance, and growth of neonatal brown forest skinks, *Sphenomorphus indicus*. *Biological Journal of the Linnean Society* **88**, 453–463.
- *JI, X., QIU, Q. B. & DIONG, C. H. (2002). Influence of incubation temperature on hatchling success, energy expenditure for embryonic development, and size and morphology of hatchlings in the Oriental Garden Lizard, *Calotes versicolor* (Agamidae). *Journal of Experimental Zoology* **292**, 649–659.
- *JI, X. & ZHANG, C. H. (2001). Effects of thermal and hydric environments on incubating eggs, hatchling success, and hatchling traits in the Chinese skink (*Eumeces chinensis*). *Acta Zoologica Sinica* **47**, 256–265.
- JIN, Y.-T. & BROWN, R. P. (2013). Species history and divergence times of viviparous and oviparous Chinese toad-headed lizards (*Phrynocephalus*) on the Qinghai-Tibetan plateau. *Molecular Phylogenetics and Evolution* **68**, 259–268.
- *KEARNEY, M. & SHINE, R. (2004). Developmental success, stability, and plasticity in closely related parthenogenetic and sexual lizards (*Heteronotia*, Gekkonidae). *Evolution* **58**, 1560–1572.
- KERR, T. D., BOUTIN, S., LAMONTAGNE, J. M., MCAHAM, A. G. & HUMPHRIES, M. M. (2007). Persistent maternal effects on juvenile survival in North American red squirrels. *Biology Letters* **3**, 289–291.
- KINGSOLVER, J. G., DIAMOND, S. E., SIEPIELSKI, A. D. & CARISON, S. M. (2012). Synthetic analyses of phenotypic selection in natural populations: lessons, limitations and future directions. *Evolutionary Ecology* **26**, 1101–1118.
- *LE HENANFF, M., MEYLAN, S. & LOURDAIS, O. (2013). The sooner the better: reproductive phenology drives ontogenetic trajectories in a temperate squamate (*Podarcis muralis*). *Biological Journal of the Linnean Society* **108**, 384–395.
- *LES, H. L., PAITZ, R. T. & BOWDEN, R. M. (2009). Living at extremes: development at the edges of viable temperature under constant and fluctuating conditions. *Physiological and Biochemical Zoology* **82**, 105–112.
- *LESHEM, A., AR, A. & ACKERMAN, R. A. (1991). Growth, water, and energy metabolism of the soft-shelled turtle (*Trionyx triunguis*) embryo: effects of temperature. *Physiological Zoology* **64**, 568–594.
- *LEWIS-WINOKUR, V. & WINOKUR, R. M. (1994). Incubation temperature affects sexual differentiation, incubation time, and posthatching survival in desert tortoises. *Canadian Journal of Zoology* **73**, 2091–2097.
- *LI, H., QU, Y.-F., HU, R.-B. & JI, X. (2009). Evolution of viviparity in cold-climate lizards: testing the maternal manipulation hypothesis. *Evolutionary Ecology* **23**, 777–790.
- *LI, H., WANG, Z., CHEN, C. & JI, X. (2012). Does the variance of incubation temperatures always constitute a significant selective force for origin of reptilian viviparity? *Current Zoology* **58**, 812–819.
- LIBERATI, A., ALTMAN, D. G., TETZLAFF, J., MULROW, C., GÖTZSCHE, P. C., IOANNIDIS, J. P. A., CLARKE, M., DEVEREAUX, P. J., KLEIJNEN, J. & MOHER, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine* **6**, e1000100 (<https://doi.org/10.1371/journal.pmed.1000100>).
- *LIGON, D. B., BIDWELL, J. R. & LOVERN, M. B. (2009). Incubation temperature effects on hatchling growth and metabolic rate in the African spurred tortoise, *Geochelone sulcata*. *Canadian Journal of Zoology* **87**, 64–72.
- *LIGON, D. B. & LOVERN, M. B. (2009). Temperature effects during early life stages of the alligator snapping turtle (*Macrochelys temminckii*). *Chelonian Conservation and Biology* **8**, 74–83.
- *LIGON, D., PETERSON, C. & LOVERN, M. (2012). Acute and persistent effects of pre- and posthatching thermal environments on growth and metabolism in the red-eared slider turtle, *Trachemys scripta elegans*. *Journal of Experimental Zoology* **317**, 227–235.
- *LIN, Z. & JI, X. (1998). The effects of thermal and hydric environments on incubating eggs and hatchlings of the grass lizard, *Takydromus septentrionalis*. *Zoological Research* **19**, 439–445.
- *LIN, Z. H., JI, X., LUO, L. G. & MA, X. M. (2005). Incubation temperature affects hatchling success, embryonic expenditure of energy and hatchling phenotypes of a prolonged egg-retaining snake, *Deinagkistrodon acutus* (Viperidae). *Journal of Thermal Biology* **30**, 289–297.
- *LIN, L.-H., MA, X.-M., LI, H. & JI, X. (2010). Phenotypic variation in hatchling Chinese ratsnakes (*Zaocys dhumnades*) from eggs incubated at constant temperatures. *Journal of Thermal Biology* **35**, 28–33.
- LINDSTRÖM, J. (1999). Early development and fitness in birds and mammals. *Trends in Ecology & Evolution* **14**, 343–347.
- *LOWENBORG, K., GOTTHARD, K. & HAGMAN, M. (2012). How a thermal dichotomy in nesting environments influences offspring of the world's most northerly oviparous snake, *Natrix natrix* (Colubridae). *Biological Journal of the Linnean Society* **107**, 833–844.
- *LOWENBORG, K., SHINE, R. & HAGMAN, M. (2011). Fitness disadvantages to disrupted embryogenesis impose selection against suboptimal nest-site choice by female grass snakes, *Natrix natrix* (Colubridae). *Journal of Evolutionary Biology* **24**, 177–183.
- *LOWENBORG, K., SHINE, R., KARVEMO, S. & HAGMAN, M. (2010). Grass snakes exploit anthropogenic heat sources to overcome distributional limits imposed by oviparity. *Functional Ecology* **24**, 1095–1102.
- *LU, H.-L., HU, R.-B. & JI, X. (2009). The variance of incubation temperatures does not affect the phenotype of hatchlings in a colubrid snake, *Xenochrophis piscator*. *Journal of Thermal Biology* **34**, 138–143.
- *LU, H. L., JI, X. A., LIN, L. H. & ZHANG, L. (2006). Relatively low upper threshold temperature in lizards from cool habitats. *Journal of Thermal Biology* **31**, 256–261.
- *LU, H., LIN, Z., LI, H. & JI, X. (2014). Geographic variation in hatchling size in an oviparous skink: effects of maternal investment and incubation thermal environment. *Biological Journal of the Linnean Society* **113**, 283–296.
- *LU, H. L., WANG, Y., TANG, W. & DU, W. G. (2013). Experimental evaluation of reproductive response to climate warming in an oviparous skink. *Integrative Zoology* **8**, 175–183.
- MAINWARING, M. C., BARBER, I., DEEMING, D. C., PIKE, D. A., ROZNIK, E. A. & HARTLEY, I. R. (2016). Climate change and nesting behaviour in vertebrates: a review of the ecological threats and potential for adaptive responses. *Biological Reviews* in press (<https://doi.org/10.1111/brv.12317>).
- MANK, J. E., PROMISLOW, D. E. & AVISE, J. C. (2005). Phylogenetic perspectives in the evolution of parental care in ray-finned fishes. *Evolution* **59**, 1570–1578.
- MARSHALL, D. J. & ULLER, T. (2007). When is a maternal effect adaptive? *Oikos* **116**, 1957–1963.
- *MATHIES, T. & ANDREWS, R. M. (1997). Influence of pregnancy on the thermal biology of the lizard, *Sceloporus jarrovi*: why do pregnant females exhibit low body temperatures? *Functional Ecology* **11**, 498–507.
- MERILÄ, J. (2012). Evolution in response to climate change: in pursuit of the missing evidence. *BioEssays* **34**, 811–818.

- *MICHELLI-CAMPBELL, M. A., CAMPBELL, H. A., CRAMP, R. L., BOOTH, D. T. & FRANKLIN, C. E. (2011). Staying cool, keeping strong: incubation temperature affects performance in a freshwater turtle. *Journal of Zoology* **285**, 266–273.
- MILLER, G. M., WATSON, S.-A., DONELSON, J. M., MCCORMICK, M. I. & MUNDAY, P. L. (2012). Parental environment mediates impacts of increased carbon dioxide on a coral reef fish. *Nature Climate Change* **2**, 858–861.
- MITCHELL, N. J., KEARNEY, M. R., NELSON, N. J. & PORTER, W. P. (2008). Predicting the fate of a living fossil: how will global warming affect sex determination and hatching phenology in tuatara. *Proceedings of the Royal Society B: Biological Sciences* **275**, 2185–2193.
- MOCZEK, A. P., SULTAN, S., FOSTER, S., LEDÓN-RETTIG, C., DWORKIN, I., NIJHOUT, H. F., ABOUHEIF, E. & PFENNIG, D. W. (2011). The role of developmental plasticity in evolutionary innovation. *Proceedings of the Royal Society B: Biological Sciences* **278**, 2705–2713.
- MONAGHAN, P. (2008). Early growth conditions, phenotypic development and environmental change. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **363**, 1635–1645.
- *MONASTERIO, C., SHOO, L. P., SALVADOR, A., IRAETA, P. & DIAZ, J. A. (2013). High temperature constrains reproductive success in a temperate lizard: implications for distribution range limits and the impacts of climate change. *Journal of Zoology* **291**, 163–145.
- *MONASTERIO, C., SHOO, L. P., SALVADOR, A., SILICEO, I. & DIAZ, J. A. (2011). Thermal constraints on embryonic development as a proximate cause for elevational range limits in two Mediterranean lacertid lizards. *Ecography* **34**, 1030–1039.
- MORGAN, H. D., SUTHERLAND, H. G. E., MARTIN, D. I. K. & WHITELAW, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nature Genetics* **23**, 314–318.
- MORJAN, C. L. (2003). Variation in nesting patterns affecting nest temperatures in two populations of painted turtles (*Chrysemys picta*) with temperature-dependent sex determination. *Behavioural Ecology and Sociobiology* **53**, 254–261.
- MORRISEY, M. B. (2016a). Meta-analysis of magnitudes, differences and variation in evolutionary parameters. *Journal of Evolutionary Biology* **29**, 1882–1904.
- MORRISEY, M. B. (2016b). Rejoinder: further considerations for meta-analysis of transformed quantities such as absolute values. *Journal of Evolutionary Biology* **29**, 1922–1931.
- MOUSSEAU, T. A. & FOX, C. W. (1998). The adaptive significance of maternal effects. *Trends in Ecology & Evolution* **13**, 403–407.
- MROSOVSKY, N. A. J. P. (1992). Sex ratio of hatchling loggerhead sea turtles: data and estimates from a 5-year study. *Canadian Journal of Zoology* **70**, 530–538.
- *MROSOVSKY, N., KAMEI, S., REES, A. & MARGARITOU, D. (2002). Pivotal temperature for loggerhead turtles (*Caretta caretta*) from Kyparissia Bay, Greece. *Canadian Journal of Zoology* **80**, 2118–2124.
- *MULLINS, M. A. & JANZEN, F. J. (2006). Phenotypic effects of thermal means and variances on smooth softshell turtle (*Apalone mutica*) embryos and hatchlings. *Herpetologica* **62**, 27–36.
- MUNDAY, P. L., WARNER, R. R., MONRO, K., PANDOLFI, J. M. & MARSHALL, D. J. (2013). Predicting evolutionary responses to climate change in the sea. *Ecology Letters* **16**, 1488–1500.
- *MURRAY, J. D., DEEMING, D. C. & FERGUSON, M. W. J. (1990). Size-dependent pigmentation-pattern formation in embryos of *Alligator mississippiensis*: time of initiation of pattern generation mechanism. *Proceedings of the Royal Society B: Biological Sciences* **239**, 279–293.
- *MUTH, A. (1980). Physiological ecology of Desert Iguana (*Dipsosaurus Dorsalis*) eggs: temperature and water relations. *Ecology* **61**, 1335–1343.
- NAKAGAWA, S., NOBLE, D. W. A., SENIOR, A. M. & LAGISZ, M. (2017). Meta-evaluation of meta-analysis: ten appraisal questions for biologists. *BMC Biology* **15**, 18–32.
- NAKAGAWA, S. & SANTOS, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology* **26**, 1253–1274.
- NAKAGAWA, S. & SCHIELZETH, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution* **4**, 133–142.
- *NELSON, N. J., THOMPSON, M. B., PLEDGER, S., KEALL, S. N. & DAUGHERTY, C. H. (2004). Egg mass determines hatchling size, and incubation temperature influences post-hatching growth, of tuatara *Sphenodon punctatus*. *Journal of Zoology* **263**, 77–87.
- *NELSON, N. J., THOMPSON, M. B., PLEDGER, S., KEALL, S. N. & DAUGHERTY, C. H. (2006). Performance of juvenile Tuatara depends on age, clutch, and incubation regime. *Journal of Herpetology* **40**, 399–403.
- *NEUWALD, J. L. & 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.
- NOBLE, D. W. A., LAGISZ, M., O'DEA, R. E. & NAKAGAWA, S. (2017a). Non-independence and sensitivity analyses in ecological and evolutionary meta-analyses. *Molecular Ecology* in press (<https://doi.org/10.1111/mec.14031>).
- NOBLE, D. W. A., STENHOUSE, V. & SCHWANZ, L. E. (2017b). Early thermal environments and developmental plasticity in reptiles: a systematic review and meta-analysis. Available at osf.io/zbgss (Accessed date 12-April-2017)
- NOBLE, D. W. A., MCFARLANE, S. E., KEOGH, J. S. & WHITING, M. J. (2014). Maternal and additive genetic effects contribute to variation in offspring traits in a lizard. *Behavioral Ecology* **25**, 633–640.
- O'DEA, R. E., NOBLE, D. W. A., JOHNSON, S. L., HESSELSON, D. & NAKAGAWA, S. (2016). The role of non-genetic inheritance in evolutionary rescue: epigenetic buffering, heritable bet hedging and epigenetic traps. *Environmental Epigenetics* **2**, 1–12.
- *O'DONNELL, R. P. & ARNOLD, S. J. (2005). Evidence for selection on thermoregulation: effects of temperature on embryo mortality in the garter snake *Thamnophis elegans*. *Copeia* **2005**, 929–933.
- *O'STEEN, S. (1998). Embryonic temperature influences juvenile temperature choice and growth rate in snapping turtles *Chelydra serpentina*. *The Journal of Experimental Biology* **201**, 439–449.
- *O'STEEN, S. & JANZEN, F. J. (1999). Embryonic temperature affects metabolic compensation and thyroid hormones in hatching snapping turtles. *Physiological and Biochemical Zoology* **72**, 520–533.
- *OUFIERO, C. E. & ANGILLETTA, M. J. JR. (2006). Convergent evolution of embryonic growth and development in the Eastern Fence lizard (*Sceloporus undulatus*). *Evolution* **60**, 1066–1075.
- *PAEZ, V. P., CORREA, J. C., CANO, A. M. & BOCK, B. C. (2009). A comparison of maternal and temperature effects on sex, size, and growth of hatchlings of the Magdalena River Turtle (*Podocnemis lewyana*) incubated under field and controlled laboratory conditions. *Copeia* **2009**, 698–704.
- *PAN, Z.-C. & JI, X. (2001). The influence of incubation temperature on size, morphology, and locomotor performance of hatchling grass lizards (*Takydromus wolteri*). *Acta Oecologica Sinica* **21**, 2031–2038.
- *PARACHU MARCO, M. V., IGNACIO PINA, C., SIMONCINI, M. & ALEJANDRO, L. (2010). Effects of incubation and rearing temperatures on *Caiman latirostris* growth. *Zoological Studies* **49**, 367–373.
- PARÉDES, U., RADERSMA, R., CANNELL, N., WHILE, G. M. & ULLER, T. (2016). Low incubation temperature induces DNA hypomethylation in lizard brains. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology* **325A**, 390–395.
- *PARKER, S. L. & ANDREWS, R. M. (2007). Incubation temperature and phenotypic traits of *Sceloporus undulatus*: implications for the northern limits of distribution. *Oecologia* **151**, 218–231.
- *PATTERSON, L. D. & BLOUIN-DEMERS, G. (2008). The effect of constant and fluctuating incubation temperatures on the phenotype of black ratsnakes (*Elaphe obsoleta*). *Canadian Journal of Zoology* **86**, 882–889.
- *PETT-PARE, C. A. & BLOUIN-DEMERS, G. (2012). Female Eastern Hog-nosed Snakes (*Heterodon platirhinos*) choose nest sites that produce offspring with phenotypes likely to improve fitness. *Canadian Journal of Zoology* **90**, 1215–1220.
- PEZARO, N., DOODY, J. S. & THOMPSON, M. B. (2016). The ecology and evolution of temperature-dependent reaction norms for sex determination in reptiles: a mechanistic conceptual model. *Biological Reviews* in press (<https://doi.org/10.1111/brv.12285>).
- *PINA, C., LARRIERA, A. & CABRERA, M. (2003). Effect of incubation temperature on incubation period, sex ratio, hatching success, and survivorship in *Caiman latirostris* (Crocodylia, Alligatoridae). *Journal of Herpetology* **37**, 199–202.
- *PINA, C., SIROSKI, P., LARRIERA, A., LANCE, V. & VERDADE, L. (2007). The temperature-sensitive period (TSP) during incubation of broad-snouted caiman (*Caiman latirostris*) eggs. *Amphibia-Reptilia* **28**, 123–128.
- *PUTZ, O. & CREWS, D. (2006). Embryonic origin of mate choice in a lizard with temperature-dependent sex determination. *Developmental Psychobiology* **48**, 29–38.
- PYRON, R. A., BURBRINK, F. T. & WEINS, J. J. (2013). A phylogeny and revised classification of Squamata, including 4161 species of lizards and snakes. *BMC Evolutionary Biology* **13**, 1–53.
- *QU, Y.-F., LI, H., GAO, J.-F. & JI, X. (2011). Embryonic thermosensitivity and hatchling morphology differ between two coexisting lizards. *Acta Oecologica* **37**, 375–380.
- *QUALLS, C. P. & ANDREWS, R. M. (1999). Cold climates and the evolution of viviparity in reptiles: cold incubation temperatures produce poor-quality offspring in the lizard, *Sceloporus virgatus*. *Biological Journal of the Linnean Society* **67**, 353–376.
- *QUALLS, C. P. & SHINE, R. (1996). Reconstructing ancestral reaction norms: an example using the evolution of reptilian viviparity. *Functional Ecology* **10**, 688–697.
- *QUALLS, F. J. & SHINE, R. (1998). Geographic variation in lizard phenotypes: importance of the incubation environment. *Biological Journal of the Linnean Society* **64**, 477–491.
- *QUALLS, F. J. & SHINE, R. (2000). Post-hatching environment contributes greatly to phenotypic variation between two populations of the Australian garden skink, *Lampropholis guichenoti*. *Biological Journal of the Linnean Society* **71**, 315–341.
- QUINN, A. E., GEORGES, A., SARRE, S. D., GUARINO, F., EZAZ, T. & MARSHALL GRAVES, J. A. (2007). Temperature sex reversal implies sex gene dosage in a reptile. *Science* **316**, 411.
- *RADDER, R. S., QUINN, A. E., GEORGES, A., SARRE, S. D. & SHINE, R. (2008). Genetic evidence for co-occurrence of chromosomal and thermal sex-determining systems in a lizard. *Biology Letters* **4**, 176–178.

- *RADDER, R. S., SHANBHAG, B. A. & SAIDAPUR, S. K. (2002). Influence of incubation temperature and substrate on eggs and embryos of the garden lizard, *Calotes versicolor* (Daud.). *Amphibia-Reptilia* **23**, 71–82.
- *RAFFERTY, A. & REINA, R. (2014). The influence of temperature on embryonic developmental arrest in marine and freshwater turtles. *Journal of Experimental Marine Biology and Ecology* **450**, 91–97.
- RÄSÄNEN, K. & KRUK, L. E. B. (2007). Maternal effects and evolution at ecological time-scales. *Functional Ecology* **21**, 408–421.
- REFSNIDER, J. M., BODENSTEINER, B. L., RENEKER, J. L. & JANZEN, F. J. (2013). Nest depth may not compensate for sex ratio skews caused by climate change in turtles. *Animal Conservation* **16**, 481–490.
- REFSNIDER, J. M. & JANZEN, F. J. (2010). Putting eggs in one basket: ecological and evolutionary hypotheses for variation in oviposition-site choice. *Annual Review of Ecology, Evolution, and Systematics* **41**, 39–57.
- *REID, K. A., MARGARITOU, D. & SPEAKMAN, J. R. (2009). Incubation temperature and energy expenditure during development in loggerhead sea turtle embryos. *Journal of Experimental Marine Biology and Ecology* **378**, 62–68.
- REYNOLDS, J. D., GOODWIN, N. B. & FRECKLETON, R. P. (2002). Evolutionary transitions in parental care and live bearing in vertebrates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **357**, 269–281.
- *RHEN, T., SAKATA, I. T. & CREWS, D. (2005). Effects of gonadal sex and incubation temperature on the ontogeny of gonadal steroid concentrations and secondary sex structures in leopard geckos, *Eublepharis macularius*. *General and Comparative Endocrinology* **142**, 289–296.
- *RHEN, T., SAKATA, J. T., ZELLER, M. & CREWS, D. (2000). Sex steroid levels across the reproductive cycle of female Leopard Geckos, *Eublepharis macularius*, from different incubation temperatures. *General and Comparative Endocrinology* **118**, 322–331.
- *RHEN, T., WILLINGHAM, E., SAKATA, J. T. & CREWS, D. (1999). Incubation temperature influences sex-steroid levels in juvenile Red-Eared Slider Turtles, *Trachemys scripta*, a species with temperature-dependent sex determination. *Biology of Reproduction* **61**, 1275–1280.
- ROACH, D. A. & WULFF, R. D. (1987). Maternal effects in plants. *Annual Review of Ecology, Evolution, and Systematics* **18**, 209–235.
- *ROCK, J. & CREE, A. (2003). Intraspecific variation in the effect of temperature on pregnancy in the viviparous gecko *Hoplodactylus maculatus*. *Herpetologica* **59**, 8–22.
- *RODRIGUEZ-DIAZ, T. & BRANA, F. (2012). Altitudinal variation in egg retention and rates of embryonic development in oviparous *Zootoca vivipara* fits predictions from the cold-climate model on the evolution of viviparity. *Journal of Evolutionary Biology* **25**, 1877–1887.
- *RODRIGUEZ-DIAZ, T., GONZALEZ, F., JI, X. & BRANA, F. (2010). Effects of incubation temperature on hatchling phenotypes in an oviparous lizard with prolonged egg retention: are the two main hypotheses on the evolution of viviparity compatible? *Zoology* **113**, 33–38.
- *ROOSENBURG, W. M. & KELLEY, K. C. (1996). The effect of egg size and incubation temperature on growth in the turtle, *Malaclemys terrapin*. *Journal of Herpetology* **30**, 198–204.
- *SAKATA, J. T., COOMBER, P., GONZALEZ-LIMA, F. & CREWS, D. (2000). Funtinal connectivity among limbic brain areas: differential effects of incubation temperature and gonadal sex in the Leopard Gecko, *Eublepharis macularius*. *Brain, Behavior and Evolution* **55**, 139–151.
- *SCHIELZETH, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**, 103–113.
- SCHWANZ, L. E. (2016). Parental thermal environment alters offspring sex ratio and fitness in an oviparous lizard. *Journal of Experimental Biology* **219**, 2349–2357.
- SCHWANZ, L. E., CORDERO, G. A., CHARNOV, E. L. & JANZEN, F. J. (2016). Sex-specific survival to maturity and the evolution of environmental sex determination. *Evolution* **70**, 329–341.
- SCHWANZ, L. E. & JANZEN, F. J. (2008). Climate change and temperature-dependent sex determination: can plasticity in maternal nesting behavior prevent extreme sex ratios? *Physiological and Biochemical Zoology* **81**, 826–834.
- SCHWANZ, L. E., SPENCER, R.-J., BOWDEN, R. M. & JANZEN, F. J. (2010). Climate and predation dominate early life-stages and adult recruitment in a turtle with temperature-dependent sex determination: insight from a long-term study. *Ecology* **91**, 3016–3026.
- SCHWARZKOPF, L. & ANDREWS, R. M. (2012). Are moms manipulative or just selfish? Evaluating the “maternal manipulation hypothesis” and implications for life-history studies of reptiles. *Herpetologica* **68**, 147–159.
- *SCHWARZKOPF, L. & BROOKS, R. J. (1985). Sex determination in the northern painted turtles: effect of incubation at constant and fluctuating temperatures. *Canadian Journal of Zoology* **63**, 2543–2547.
- SENIOR, A. M., GRUEBER, C. E., KAMIYA, T., LAGISZ, M., O'DWYER, K., SANTOS, E. S. A. & NAKAGAWA, S. (2016). Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. *Ecology* **97**, 3293–3299.
- *SEXTON, O. & MARION, K. (1974). Duration of incubation of *Sceloporus undulatus* eggs at constant temperature. *Physiological Zoology* **47**, 91–98.
- SHINE, R. (1995). A new hypothesis for the evolution of viviparity in reptiles. *The American Naturalist* **145**, 809–823.
- *SHINE, R. (1999a). Egg-laying reptiles in cold climates: determinants and consequences of nest temperatures in montane lizards. *Journal of Evolutionary Biology* **12**, 918–926.
- SHINE, R. (1999b). Why is sex determined by nest temperature in many reptiles? *Trends in Ecology and Evolution* **14**, 186–189.
- *SHINE, R. (2002). Reconstructing an adaptationist scenario: what selective forces favor the evolution of viviparity in montane reptiles? *The American Naturalist* **160**, 582–593.
- SHINE, R. (2005). Life-history evolution in reptiles. *Annual Review of Ecology, Evolution, and Systematics* **36**, 23–46.
- SHINE, R. (2014). Evolution of an evolutionary hypothesis: A history of changing ideas about the adaptive significance of viviparity in reptiles. *Journal of Herpetology* **48**, 147–167.
- *SHINE, R. & DOWNES, S. (1999). Can preganant lizards adjust their offspring phenotypes to environmental conditions? *Oecologia* **119**, 1–8.
- *SHINE, R., ELPHICK, M. J. & HARLOW, P. S. (1995). Sisters like it hot. *Nature* **378**, 451–452.
- *SHINE, R. & HARLOW, P. S. (1993). Maternal thermoregulation influences offspring viability in a viviparous lizard. *Oecologia* **96**, 122–127.
- *SHINE, R. & HARLOW, P. S. (1996). Maternal manipulation of offspring phenotypes via nest-site selection in an oviparous lizard. *Ecology* **77**, 1808–1817.
- *SHINE, R., MADSEN, T. R. L., ELPHICK, M. J. & HARLOW, P. S. (1997). The influence of nest temperatures and maternal brooding on hatchling phenotypes in water pythons. *Ecology* **78**, 1713–1721.
- SINCLAIR, B. J., MARSHALL, K. E., SEWELL, M. A., LEVESQUE, D. L., WILLET, C. S., SLOTSBO, S., DONG, Y., HARLEY, C. D. G., MARSHALL, D. J., HELMUTH, B. S. & HUEY, R. B. (2016). Can we predict ectotherm responses to climate change using thermal performance curves and body temperatures? *Ecology Letters* **19**, 1372–1385.
- SINERO, B., MÉNDEZ-DE-LA-CRUZ, F., MILES, D. B., HEULIN, B., BASTIAANS, E., VILLAGRÁN-SANTA CRUZ, M., LARA-RESENDIZ, R., GADSEN, H., AVILA, L. J., MORANDO, M., DE LA RIVA, I. J., SEPULVEDA, P. V., ROCHA, C. F. D., IBARGÜENGOYÍA, N., PUNTRIANO, C. A., et al. (2010). Erosion of lizard diversity by climate change and altered thermal niches. *Science* **328**, 894–899.
- *SPENCER, R.-J., JANZEN, F. J. & THOMPSON, M. B. (2006). Counterintuitive density-dependent growth in a long-lived vertebrate after the removal of nest predators. *Ecology* **87**, 3109–3118.
- *SPOTILA, J. R., ZIMMERMAN, L. C., BINCKLEY, C. A. & GRUMBLES, J. S. (1994). Effects of incubation conditions on sex determination, hatching success, and growth of hatchling Desert Tortoises, *Gopherus agassizii*. *Herpetological Monographs* **8**, 103–116.
- *STEYERMARK, A. C. & SPOTILA, J. R. (2000). Effects of maternal identity and incubation temperature on snapping turtle (*Chelydra serpentina*) metabolism. *Physiological and Biochemical Zoology* **73**, 298–306.
- *STEYERMARK, A. C. & SPOTILA, J. (2001a). Effects of maternal identity and incubation temperature on hatching and hatchling morphology in Snapping Turtles, *Chelydra serpentina*. *Copeia* **2001**, 129–135.
- *STEYERMARK, A. C. & SPOTILA, J. R. (2001b). Effects of maternal identity and incubation temperature on Snapping turtle (*Chelydra serpentina*) growth. *Functional Ecology* **15**, 624–632.
- STRATMANN, A. & TABORSKY, B. (2014). Antipredator defences of young are independently determined by genetic inheritance, maternal effects and own early experience in mouthbrooding cichlids. *Functional Ecology* **28**, 944–953.
- SULTAN, S. E. (2007). Development in context: the timely emergence of eco-devo. *Trends in Ecology & Evolution* **22**, 575–582.
- SUTTON, J. T., NAKAGAWA, S., ROBERTSON, B. C. & JAMIESON, I. G. (2011). Dismantling the roles of natural selection and genetic drift in shaping variation at MHC immunity genes. *Molecular Ecology* **20**, 4408–4420.
- *TAMPLIN, J. W. & CYR, A. B. (2011). Effects of acclimation and egg-incubation temperature on selected temperature by hatchling western painted turtles (*Chrysemys picta bellii*). *Journal of Thermal Biology* **36**, 507–514.
- *TANG, X., YUE, F., MA, M., WANG, N., HE, J. & CHEN, Q. (2012a). Effects of thermal and hydric conditions on egg incubation and hatchling phenotypes in two *Phrynocephalus* lizards. *Asian Herpetological Research* **3**, 184–191.
- *TANG, X.-L., YUE, F., YAN, X. F., ZHANG, D.-J., XIN, Y., WANG, C. & CHEN, Q. (2012b). Effects of gestation temperature on offspring sex and maternal reproduction in a viviparous lizard (*Eremias multiocellata*) living at high altitude. *Journal of Thermal Biology* **37**, 438–444.
- *TELEMCO, R. S., BAIRD, T. A. & SHINE, R. (2011). Tail waving in a lizard (*Bassiana duperreyi*) functions to deflect attacks rather than as a pursuit-deterrent signal. *Animal Behaviour* **82**, 369–375.
- TELEMCO, R. S., ELPHICK, M. J. & SHINE, R. (2009). Nesting lizards (*Bassiana duperreyi*) compensate partly, but not completely, for climate change. *Ecology* **90**, 17–22.
- *TELEMCO, R. S., RADDER, R. S., BAIRD, T. A. & SHINE, R. (2010). Thermal effects on reptile reproduction: adaptation and phenotypic plasticity in a montane lizard. *Biological Journal of the Linnean Society* **100**, 642–655.

- TELEMCO, R. S., WARNER, D. A., REIDA, M. K. & JANZEN, F. J. (2013). Extreme developmental temperatures results in morphological abnormalities in painted turtles (*Chrysemys picta*): a climate change perspective. *Integrative Zoology* **8**, 197–208.
- *THOMPSON, M. B. (1988). Influence of incubation temperature and water potential on sex determination in *Emydura macquarii* (Testudines: Pleurodira). *Herpetologica* **44**, 86–90.
- *TOUSIGNANT, A. & CREWS, D. (1995). Incubation temperature and gonadal sex affect growth and physiology in the Leopard gecko (*Eublepharis macularius*), a lizard with temperature-dependent sex determination. *Journal of Morphology* **224**, 159–170.
- *TRNIK, M., ALBRECHTOVA, J. & KRATOCHVIL, L. (2011). Persistent effect of incubation temperature on stress-induced behavior in the Yucatan Banded Gecko (*Coleonyx elegans*). *Journal of Comparative Psychology* **125**, 22–30.
- TSCHIRREN, B. (2015). Differential effects of maternal yolk androgens on male and female offspring: a role for sex-specific selection? *PLoS ONE* **10**, 1–11.
- ULLER, T. (2008). Developmental plasticity and the evolution of parental effects. *Trends in Ecology & Evolution* **23**, 432–438.
- *ULLER, T., ODIERNA, G. & OLSSON, M. (2008). Sex ratio variation and sex determination in the mallee dragon *Ctenophorus fordi*. *Integrative Zoology* **3**, 157–165.
- *ULLER, T., WHILE, G. M., CADBY, C. D., HARTS, A., O'CONNOR, K., PEN, I. & WAPSTRA, E. (2011). Altitudinal divergence in maternal thermoregulatory behaviour may be driven by differences in selection on offspring survival in a viviparous lizard. *Evolution* **65**, 2313–2324.
- *VALENZUELA, N. (2001). Constant, shift, and natural temperature effects on sex determination in *Podocnemis expansa* turtles. *Ecology* **82**, 3010–3024.
- VALENZUELA, N. & LANCE, V. (2004). *Temperature-dependent Sex Determination in Vertebrates*. Smithsonian Books, Washington.
- *VANDAMME, R., BAUWENS, D., BRANA, F. & VERHEYEN, R. F. (1992). Incubation temperature differentially affects hatching time, egg survival, and hatchling performance in the lizard *Podarcis muralis*. *Herpetologica* **48**, 220–228.
- VÁZQUEZ, D. P., GIANOLI, E., MORRIS, W. F. & BOZINOVIC, F. (2017). Ecological and evolutionary impacts of changing climate variability. *Biological Reviews* **92**, 22–42.
- *VERDÚ-RICOY, J., IRAETA, P., SALVADOR, A. & DÍAZ, J. (2014). Phenotypic responses to incubation conditions in ecologically distinct populations of a lacertid lizard: a tale of two phylogeographic lineages. *Journal of Zoology* **292**, 184–191.
- VEICHTBAUER, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**, 1–48.
- *VIETS, B. E., TOUSIGNANT, A., EWERT, M. A., NELSON, C. E. & CREWS, D. (1993). Temperature-dependent sex determination in the Leopard Gecko, *Eublepharis macularius*. *The Journal of Experimental Zoology* **265**, 679–683.
- VISSER, M. (2008). Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B: Biological Sciences* **275**, 649–659.
- VISSER, M. E. & BOTH, C. (2005). Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B: Biological Sciences* **272**, 2561–2569.
- *VLADIMIROVA, I. G., ALEKSEEVA, T. A. & NECHAEVA, M. V. (2005). Effect of temperature on the rate of oxygen consumption during the second half of embryonic and early postembryonic development of European Pond Turtle *Emys orbicularis* (Reptilia: Emydidae). *Biology Bulletin* **32**, 484–489.
- WALKER, S., STUART-FOX, D. & KEARNEY, M. R. (2016). Has contemporary climate change played a role in population declines of the lizard *Ctenophorus decresii* from semi-arid Australia? *Journal of Thermal Biology* **54**, 66–77.
- *WANG, Z., LU, H.-L., MA, L. & JI, X. (2014). Viviparity in high-altitude *Phrynocephalus* lizards is adaptive because embryos cannot fully develop without maternal thermoregulation. *Oecologia* **174**, 639–649.
- *WANG, Z., MA, L., SHAO, M. & JI, X. (2013). Differences in incubation length and hatchling morphology among five species of oviparous *Phrynocephalus* lizards (Agamidae) from China. *Asian Herpetological Research* **4**, 225–232.
- *WAPSTRA, E. (2000). Maternal basking opportunity affects juvenile phenotype in a viviparous lizard. *Functional Ecology* **14**, 345–352.
- *WARNER, D. A., MOODY, M. A., TELEMCO, R. S. & KOLBE, J. J. (2012). Egg environments have large effects on embryonic development, but have minimal consequences for hatchling phenotypes in an invasive lizard. *Biological Journal of the Linnean Society* **105**, 25–41.
- WARNER, D. A. & SHINE, R. (2007). Fitness of juvenile lizards depends on seasonal timing of hatching not offspring body size. *Oecologia* **154**, 65–73.
- WARNER, D. A. & SHINE, R. (2008). Maternal nest-site choice in a lizard with temperature-dependent sex determination. *Animal Behaviour* **75**, 861–870.
- *WARNER, D. & SHINE, R. (2011). Interactions among thermal parameters determine offspring sex under temperature-dependent sex determination. *Proceedings of the Royal Society B: Biological Sciences* **278**, 256–265.
- *WEBB, J. K., SHINE, R. & CHRISTIAN, K. A. (2006). The adaptive significance of reptilian viviparity in the tropics: testing the maternal manipulation hypothesis. *Evolution* **60**, 115–122.
- WEBER, S. B., BRODERICK, A. C., GROOTHUIS, T. G. G., ELLICK, J., GODLEY, B. J. & BLOUNT, J. D. (2012). Fine-scale thermal adaptation in a green turtle nesting population. *Proceedings of the Royal Society B: Biological Sciences* **279**, 1077–1084.
- WEST-EBERHARD, M. J. (2003). *Developmental Plasticity and Evolution*. Oxford University Press Inc., New York.
- *WHILE, G. M., ULLER, T. & WAPSTRA, E. (2009). Offspring performance and the adaptive benefits of prolonged pregnancy: experimental tests in a viviparous lizard. *Functional Ecology* **23**, 818–825.
- *WHITEHEAD, P. J., SEYMOUR, R. S. & WEBB, G. J. W. (1992). Energetics of development of embryos of the Australian Freshwater Crocodile, *Crocodylus johnstoni*: relation to duration of incubation. *Physiological Zoology* **65**, 360–378.
- WILLETTE, D. A. S., TUCKER, J. K. & JANZEN, F. J. (2005). Linking climate and physiology at the population level for a key life-history stage of turtles. *Canadian Journal of Zoology* **83**, 845–850.
- WOOD, D. W. & BJORNDAAL, K. A. (2000). Relation of temperature, moisture, salinity, and slope to nest site selection in Loggerhead sea turtles. *Copeia* **2000**, 119–128.
- *XU, D.-D. & JI, X. (2007). Sexual dimorphism, female reproduction and egg incubation in the oriental leaf-toed gecko (*Hemidactylus bowringii*) from southern China. *Zoology* **110**, 20–27.
- *YAN, X. F., TANG, X.-L., YUE, F., ZHANG, D. J., XIN, Y., WANG, C. & CHEN, Q. (2011). Influence of ambient temperature on maternal thermoregulation and neonate phenotypes in a viviparous lizard, *Eremias multiocellata*, during the gestation period. *Journal of Thermal Biology* **36**, 187–192.
- *YNTEMA, C. L. (1976). Effects of incubation temperatures on sexual differentiation in the turtle, *Chelydra serpentina*. *Journal of Morphology* **150**, 453–462.
- *YNTEMA, C. L. (1978). Incubation times for eggs of the turtle *Chelydra serpentina* (Testudines: Chelydridae) at various temperatures. *Herpetologica* **34**, 274–277.
- *ZHANG, Y. & JI, X. (2004). Pivotal range and thermosensitive period of the pig-nosed turtle, *Carettochelys insculpta* (Testudines: Carettochelydidae), from northern Australia. *Canadian Journal of Zoology* **82**, 1251–1257.
- *ZHANG, D.-J., TANG, X.-L., YUE, F., CHEN, Z., LI, R.-D. & CHEN, Q. (2010). Effect of gestation temperature on sexual and morphological phenotypes of offspring in a viviparous lizard, *Eremias multiocellata*. *Journal of Thermal Biology* **35**, 129–133.
- *ZHU, X.-P., WEI, C.-Q., ZHAO, W.-H., DU, H.-J., CHEN, Y.-L. & GUI, J.-F. (2006). Effects of incubation temperatures on embryonic development in the Asian yellow pond turtle. *Aquaculture* **259**, 243–248.

VIII. SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article.

Appendix S1. Supplementary methods.

Appendix S2. Meta-analysis.

Appendix S3. Multi-level meta-regression (MLMR) models.

Appendix S4. Sensitivity analyses.

Fig. S1. PRISMA statement.

Fig. S2. Phylogenetic tree used for meta-analytic models.

Fig. S3. Sensitivity analysis testing for changes in model predictions of Hedges' g for traits (development, performance, behaviour and survival) with changes in mid-temperature and their temperature difference.

Fig. S4. Funnel plot of effect size as a function of precision for incubation duration.

Table S1. Studies manipulating thermal environments of gestating viviparous reptiles.

Table S2. Model coefficients and lower (L.) and upper (U.) 95% credible intervals (CI) for the effect (g) of average treatment temperature (T_{mid}), temperature difference (T_{diff}) and fluctuations in temperature (T_F) on effect size for behavioural, physiological and morphological traits.

Table S3. Model coefficients and lower (L) and upper (U) 95% credible intervals (CI) for the effect (g) of average treatment temperature (T_{mid}), temperature difference (T_{diff}) and fluctuations in temperature (T_F) on effect size for behavioural, physiological and survival traits.

Fig. S5. Absolute effect size for continuous temperature fluctuation for performance, physiology, survival, behaviour, development and morphology for each of the four reptilian orders.

Fig. S6. Absolute effect size in relation to difference in treatment temperature for performance, physiology, survival, behaviour, development and morphology for each of the four reptilian orders.

Fig. S7. Absolute effect size in relation to mid-temperature of treatments for performance, physiology, survival, behaviour, development and morphology for each of the four reptilian orders.

Fig. S8. Hedges' g plotted against difference in treatment temperatures for performance, physiology, survival, behaviour, development and morphology for each of the four reptilian orders.

Fig. S9. Hedges' g plotted against difference in treatment mid-temperature for performance, physiology, survival, behaviour, development and morphology for each of the four reptilian orders.

Fig. S10. Absolute effect size across age for performance, physiology, survival, behaviour, development and morphology for each of the four reptilian orders.

Fig. S11. Hedges' g plotted against age of measurement for performance, physiology, survival, behaviour, development and morphology for each of the four reptilian orders.

(Received 15 November 2016; revised 14 March 2017; accepted 17 March 2017)