Reduced plasticity and opportunity for selection in terrestrial ectotherm populations under climate change

Daniel W.A. Noble, Fonti Kar, Frank Seebacher, Alex Bush, & Shinichi Nakagawa

#### Affliations:

Division of Ecology and Evolution, Research School of Biology, The Australian National University, Canberra, ACT 2600, Australia Ecology and Evolution Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia SOLES, University of Sydney, Sydney, NSW, Australia Department of Biology, Lancaster University, Liverpool, UK

## **Abstract**

## **Introduction**

Variable thermal environments are expected to result in strong selection pressures that lead to adaptation or the evolution of phenotypic plasticity – both of which are considered critical for population resilience to human-induced climate change (Seebacher *et al.* 2015). Phenotypic plasticity is predicted to evolve when environmental variability is high but predictable. Despite this theoretical expectation, empirical support for this prediction is scant [], likely because many organisms can behaviorally adjust micro-habitat selection to offset thermal stress, the costs of plasticity are high [], and/or the prediction is likely only supported for specific plastic responses (i.e., active and developmental plasticity). Reversible forms of phenotypic plasticity, such as acclimatization, may not be expected to adhere to such predictions.

Periods of past climatic change have had disproportionate impacts on some ecosystems over others leading to debates over which ecosystems will be most vulnerable to contemporary climate change. Studies have highlighted species occupying terrestrial ecosystems as being particularly vulnerable given their weak acclimation abilities and greater probability of experiencing thermal extremes that overwhelm physiological homeostasis (Gunderson & Stillman 2015; Seebacher *et al.* 2015; Pinsky *et al.* 2019). Despite marine and freshwater ecosystems appearing to have greater physiological acclimation capacity (e.g., see Seebacher *et al.* 2015), it is unclear if the magnitude of physiological adjustment is sufficient. In addition, low oxygen availability has been suggested as a major factor influencing the vulnerability of aquatic ecosystems.

Applying new effect sizes that allow us to make use of powerful meta-analytic models we: 1) re-evaluate the degree to which aquatic and terrestrial ectotherms are capable of physiological plasticity; 2) test whether the opportunity for selection on physiological traits changes as temperatures rise, by applying new effect sizes that capture changes in physiological trait variance, and 3) test whether climate variability and predictability explain a populations acclimation capacity and affect changes in trait variance.

## **Materials and Methods**

### *Literature collection*

We compiled literature on ectothermic animals that measured physiological rates (e.g., metabolic rate) at two or more temperatures after having been acclimated (or acclimatized) at these temperatures. We used data from a previous meta-analysis (Seebacher *et al.* 2015) and updated Seebacher *et al.* (2015)’s data by extracting data from suitable studies from our own searches that followed the same search protocol. More specifically, we performed a literature search on the 28th of June 2017 using the Web of Science database. We limited our search to articles or proceedings papers published in English from 2013 to 2017 (the date after Seebacher *et al.* 2015 searches were conducted) using the following topic search string: *“(acclimat* AND (therm\* OR temp*) NOT (plant* OR tree\* OR forest\* OR fung\* OR mammal\* OR marsup\* OR bird\* OR human OR exercis\* OR train\* OR hypoxi*))“*. We further limited to the following research areas: Anatomy Morphology; Biodiversity Conservation; Biology; Ecology; Endocrinology Metabolism; Entomology; Evolutionary Biology; Marine Freshwater Biology; Physiology; Respiratory System, Reproductive Biology, Zoology.

Our search resulted in 1,321 papers for screening in Rayyan (Ouzzani *et al.* 2016). We also cross-checked papers we found in our searches with a recent paper by Havird *et al.* (2020), which also updates Seebacher *et al.* (2015)’s dataset. We included any papers that were missed between our searches and those of Havird *et al.* (2020) from the dates 2013-2017. Havird *et al.* (2020) added 7 new studies between 2013-2017 (mainly because they were focused on metabolic rates), and our searches differed from theirs by only a single paper (i.e., Bulgarella *et al.* 2015). Given the physiological traits we included were broader, we had a substantial increase in additional papers that we added to Seebacher *et al.* (2015)’s dataset. More specifically, in addition to the 191 papers we included from the Seebacher *et al.* (2015) dataset, we extracted data from an extra 65 papers (with a total of 238 effects) that were published between 2013 - 2017 (a 34.0314136% increase in the number of published articles). Note that Seebacher *et al.* (2015) included a total of 205 publications, however, not all these contained the necessary statistics we needed to derive effect sizes and associated sampling variances (see below). While we may have missed papers, our goal was to obtain a large representative (and unbiased) sample of acclimation research rather than a comprehensive dataset. As such, our database represents the most up-to-date dataset used by Seebacher *et al.* (2015) to answer questions on acclimation across ectotherms.

We split the screening of titles and abstracts for the 1,321 papers found in our search among all authors evenly. To ensure consistency among authors in title and abstracts that should be included, prior to screening all authors went through a randomly selected set of papers together - agreeing on those that were relevant and those that were not based on our inclusion criteria (see below). Where any authors were uncertain about whether to include a paper in the sub-sample they screened, we conservatively included the paper for full text screening and discussed uncertain papers among authors to come to a decision on whether to include the paper. After title and abstract screening, we were left with a total of 149 papers for full text screening. Papers were included only if they: 1) measured a physiological rate acutely at two temperatures on a sample of animals chronically exposed to the same two temperatures for at least 1 week; and 2) where physiological rates measured were burst and sustained locomotion, metabolic rates (standard, resting, routine and maximal), heart rates, and/or enzyme activities.

### *Data Compilation*

We extracted means, standard deviations, and sample sizes for physiological rates at the two test temperatures. If there were more than two test temperatures, we choose only the test temperatures that fell within the most likely natural range of temperatures experienced by the species in question. We extracted these data from text, tables or figures of a given paper. Data were extracted from figures using the R package *metaDigitise* (Pick *et al.* 2019). We also recorded the phylum, class, order, genus and species under study and the latitude and longitude of the population that was being studied. For studies that did not provide latitude and longitude for the population, we searched for similar studies by the lab group to identify where the population was likely to have been sourced or derived from when needed. If the population was derived from the wild, we recorded the nearest latitude and longitude of the population to the field collection site. If the animals had been sourced from a commercial supplier, we took the latitude and longitude of the supplier that the paper identified the animals to have originated from. When it was not possible to find latitude and longitude using these methods, we looked up the distribution of the species in question and took the latitude and longitude of the centroid of the species’ distributional range.

### *Based Effect Sizes and Sampling Variances for Means and Variances*

Following Noble *et al.* (2022) we calculated a series of temperature corrected effect sizes that compared mean physiological rates () as well as the variability in physiological rates ( and ). These effect sizes are similar to the traditional temperature coefficient (), but with formal analytical approximations for their sampling variances. Sampling variances for effect sizes allowed us to make use of traditional meta-analytic modelling approaches.

#### *Comparing changes in mean physiological rates*

To compare mean physiological rates, we calculated the log response ratio, (Noble *et al.* 2022) as follows:

Where, and are mean physiological rates and and are the temperatures that these rates are measured. Log transformation of this ratio makes the effect size normally distributed. Equation @ref(eq:lnq10) is essentially a temperature corrected equivalent to the log response ratio (lnRR) (Hedges *et al.* 1999; Lajeunesse 2011) when the numerator and denominator are measured at different temperatures. This allows one to compare the mean of two temperature treatments directly regardless of the temperatures that these groups have been measured. The sampling variance for equation @ref(eq:lnq10) can be computed as follows (as described in Noble *et al.* (2022)):

Here, and are the standard deviations and and are the sample sizes in group 1 and 2, respectively.

#### *Comparing variance in physiological rates*

Nakagawa *et al.* (2015) recently proposed analogous effect size estimates to *lnRR* that allow for comparisons of changes in variance between two groups, the log variance ratio (*lnVR*) and the log coefficient of variation (*lnCVR*). *lnVR* and *lnCVR* are ratios that describe the relative difference in trait variability between two groups. We refer readers to Nakagawa *et al.* (2015) for the equations describing *lnVR* and *lnCVR*, but these can easily be extended to their analogues (and associated sampling variance) as follows:

Equations @ref(eq:lnq10VR) and @ref(eq:slnq10VR) describe the change in physiological rate variance (eqn @ref(eq:lnq10VR)) across a 10°C temperature change along with its sampling variance (eqn @ref(eq:slnq10VR)). While this is a useful metric, as discussed by Nakagawa *et al.* (2015) there is often a strong mean-variance relationship that needs to be accounted for in analysing changes in variance. As such, we calculated the coefficient of variation, which standardizes changes in variance for changes in means as follows:

where is the coefficient of variation defined as .

#### *Calculating acute and acclimation , and estimates*

Using the mean, standard deviation, and sample size for all acute and acclimation treatments of studies in our databases we derived acute and acclimation , and estimates. For all effect sizes the higher acute or acclimation temperature was in the numerator and the lower of the two temperatures in the denominator. As such, positive effect sizes suggest that the mean or variance is larger at the higher of the two temperatures, standardized to 10°C.

### *Moderator Variables*

We recorded or derived a series of moderator variables from each study that are expected to have an impact on our effect size estimates. These included the duration of acclimation in days and acclimation type (“acclimation” or “acclimatization”) given that acclimation responses are expected to depend on how long chronic temperature exposure occurs (longer exposure = better acclimation response) (Seebacher *et al.* 2015). We also recorded if the sample of animals were derived from captive or wild stocks, the life-history stage of the animals used (“adult” or “juvenile”) and the habitat type (“freshwater”, “marine” or “terrestrial”) given that Seebacher *et al.* (2015) show that these factors can impact estimates. Physiological rate measures varied widely across the studies but could generally be grouped into discrete trait categories (Seebacher *et al.* 2015). As such, using the detailed information on the trait type, and its associated units from a given study, we categorized each effect size into one of 12 trait categories. These categories included measures of whole organism performance measures including cardiac (i.e., ‘cardiac’) and muscle (‘muscle’) function, sprint speed (‘sprint’) and endurance (‘endurance’) and metabolic rates (i.e., maximal and resting metabolic rate; max MR’, ‘rest MR’, respectively). Studies also quantified various enzymatic reaction rates, including enzymes involved in general metabolic responses (categorized as ‘metabolic enzyme’), various parts of the electron transport chain, including ATPase activity (‘ATPase’), mitochondrial leak (‘mito\_leak’) and oxidation (‘mito\_oxidation’) as well as antioxidant enzymes (‘antiox’). All other traits not falling within these categories were placed into ‘other’.

### *Climate Data*

To understand how climate has impacted species’ physiological acclimation abilities we used the coordinates reported by each study to extract temperature data from terrestrial and aquatic environments. It was unclear whether climate at the locations of captive reared organisms would be representative of a population’s climate history - particularly for species reared under captive condition for many generations. Given that we were interested in understanding climate driven effects on acclimation capacity we only used studies on wild populations were used for climate analyses.

Temperature data was extracted using the monthly averages provided by the ERA5 climate model, available from the Copernicus climate data store (Hersbach *et al.* 2020). For each population and species in the dataset we extracted a 30-year period (1958-2022) of either surface temperature at 2 meters for both terrestrial and freshwater taxa, or sea surface temperature for the marine taxa using the *ncdf4* R package (vers. 1.19). We chose a 2-meter resolution because we believed that it more likely to reflects the micro-thermal environment experienced by terrestrial and freshwater ectotherms at those locations.

Using the thermal time-series data for each location we summarised various metrics of thermal variability across months and years as well as estimates of thermal predictability (i.e., autocorrelation). To estimate thermal variability, we calculated the coefficient of variation (, where SD = standard deviation in temperature and M = the mean temperature for each year). To estimate thermal predictability, we calculated the auto-regressive time lag across the entire dataset. Theoretical and empirical studies of plasticity evolution have emphasised the importance of both climate variability and predictability in plasticity evolution.

### *Meta-Analysis*

We analysed our data using multilevel meta-analytic (MLMA) and meta-regression (MLMR) models in R (vers. 4.2.1) using *brms* (Bürkner 2017; vers. 2.18.0 Bürkner 2018; “Stan development team. RStan” 2021) and *metafor* (vers. 3.8.1 Viechtbauer 2010). We fit both Bayesian and frequentist approaches to ensure that our results were consistent, and to create orchard plots more easily (vers. 2.0, Nakagawa *et al.* 2021a). In addition, Bayesian methods better protect against type I errors in the presence of complex sources of non-independence (Nakagawa *et al.* 2021b; Song *et al.* 2021). For our Bayesian models, we ran 4 MCMC chains, each with a warm-up of 1000 followed by 4000 sampling iterations keeping every 5 iterations for a total of 3200 samples from the posterior distribution. We used flat Gaussian priors for ‘fixed’ effects (i.e., ) and a student t-distribution for ‘random’ effects (i.e., ). We checked that all MCMC chains were mixing and had converged (i.e., ). We compared any competing models using Akaike’s Information Criteria (AIC) (if frequentist) or Wantabe Information Criteria (WIC) (if Bayesian). We deemed models with the lowest IC value to be best supported if there was a between the competing models of 2 or more. If two models were within 2 units we went with the most parsimonious model.

#### *Multi-level Meta-analysis (MLMA) Models*

We first fit multi-level meta-analysis (MLMA) models (i.e., intercept only models) for both and , that included study, species, and phylogeny as random effects to account for non-independence. We also included trait as a random effect to account for trait variation within the data. Our MLMA models allowed us to partition the variation in and among these key sources while accounting for total sampling variance in each. This allowed us to calculate total heterogeneity [i.e., ; *sensu* Nakagawa & Santos (2012); Noble *et al.* (2022)] along with various metrics describing the proportion of variance explained by each random effect level (Nakagawa & Santos 2012). We also present 95% prediction intervals which describe the expected distribution of effects from future studies (Nakagawa *et al.* 2021a; Noble *et al.* 2022).

A phylogeny was derived using the Open Tree of Life (OTL) with the *rotl* package in R (vers. 3.0.14, Michonneau *et al.* 2016), and plotted using *ggtree* (vers. 3.6.2, Yu *et al.* 2017). We resolved all polytomies in the tree. Any missing taxa were replaced with closely related species and branch lengths were computed using Grafen’s method (using power = 0.7, Grafen 1989). We used the R packages *ape* (vers. 5.6.2, Paradis & Schliep 2019) and *phytools* (vers. 1.2.0, Revell 2012) to prune the tree for individual analyses and calculate phylogenetic covariance (or correlation) matrices used in meta-analytic models.

#### *Multi-level Meta-regression (MLMR) Models*

After quantifying levels of heterogeneity, we fit a series of multi-level meta-regression (MLMR) models to test our key questions. In all models, we included the same random effects as we used in our MLMA models. Acclimation time varied from 4 to 408 days (mean (SD) = 37.9815271 45.1903321 days), and terrestrial ectotherms were acclimated for a much shorter duration (mean (SD) = 23.5284553 15.5570474, n = 125) than both freshwater (mean (SD) = 36.8087167 28.7077629, n = 430) and marine species (mean (SD) = 46.1775362 67.2126693, n = 313). Rates of acclimation have been shown to be faster for many terrestrial groups compared to aquatic organisms [e.g., amphibians and reptiles have faster rates of acclimation than fishes; See Einum & Burton (2023)], which would make it more likely that terrestrial ectotherms would show lower post acclimation . To control for these possible differences, acclimation time was mean-centered (mean = 0) and included in all our models. As such, all estimates can be interpreted as values for an average level of acclimation time (i.e., 37.9815271 days).

We first tested the degree to which acute and acclimation and effects varied by habitat type (i.e., terrestrial, freshwater, and marine ecosystems). Models included an interaction between effect type (i.e., acute or acclimation) and habitat. Reduced mean relative to indicates that acclimation to thermal environments results in partial compensation of physiological rates (i.e., phenotypic plasticity), whereas no differences between and suggests organisms are not capable of physiological plasticity (Havird *et al.* 2020). In contrast, a difference in relative to would imply that changes in between individual variation in physiological rates across 10C differ depending on whether acute or acclimation responses are measured. If the interaction between effect type and habitat was not supported, then we fit a model that only contained additive effects of effect type and habitat. Following on from these models, we subset each habitat type and explored how mean changed across traits. Within each habitat (marine, freshwater, and terrestrial) we fit a series of models that included an interaction between effect type (acute / acclimation) and trait category (as defined above). Variance in effects within trait categories appeared to vary depending on the trait type in question. Comparison of a model with and without heteroscedastic residual variance favored a model with heteroscedastic residual variance across trait categories (; marine = 58, freshwater = 120, and terrestrial = 12). To ensure models converged we limited to trait categories for each habitat with six or more effect sizes.

Second, we tested whether different life-stages are more or less likely to acclimate by fitting a model for each habitat type and including an interaction between life-stage (‘adult’ or ‘juvenile’) and effect type. We predicted that acclimation responses would be more likely early in development compared to later in development, but that this should depend on the habitat type given the different constraints faced by different early life stages across major habitat types.

### *Sensitivity Analyses*

### *Publication Bias*

## **Results**

The final dataset included a total of 91 freshwater (fishes = 48, Molluscs = 4, Amphibians = 19, Reptiles = 8, Arthropods = 10, and a single Crustacean and Nematode species), 90 marine (fishes = 47, Annelids = 2, Molluscs = 21, Echinoderms = 7, Reptiles = 1, Arthropods = 10, and a single Crustacean and Cnidarian species), and 45 terrestrial species (Annelids = 1, Molluscs = 5, Arthropods = 14, Reptiles = 12 and Amphibians = 12 along with a single Tardigrade species) ([Figure 1](#fig-1) A). We had more data on acute thermal responses (n = 1115) compared to thermal responses after an acclimation period (n = 798) because both acclimation temperatures had separate acute responses ([Figure 1](#fig-1)). While the two acute effect sizes did differ significantly from each other, on average (Acute responses were higher for animals acclimated to high temperatures – = 0.0692282, 95% CI: 0.0352333 to 0.1032231, p < 0.001), they were in the same direction and only differed by ~10%. As such, we averaged the two acute effect sizes in subsequent analyses.

Most of the effect size estimates came from measurements of metabolic rates (both resting and maximal – = 190, = 3069, considering acute and acclimation effects together), metabolic enzymes ( = 61, = 2394) and whole-organism performance traits (i.e., measures of speed and endurance – = 73, = 963).

### *Do terrestrial and aquatic ectotherms differ in their capacity to acclimate?*

Overall, and differed by only 7.942601% across all habitats (95%CI: 4.7858502 to 11.2428758%). Ectotherms in marine and freshwater environments showed partial compensation of physiological rates ([Figure 1](#fig-1) B) amounting to reduced of 10.1151427% (95% CI: 6.0391131 to 14.1246045) in freshwater and 9.2635907% (95% CI: 2.6165698 to 15.7998426) in marine environments. In contrast, terrestrial ectotherms showed no acclimation (possibly even inverse acclimation) – showing a 4.4490697% increase in (95% CI: 0.1937765 to 11.4392369, [Figure 1](#fig-1) B). Considering acute responses of animals acclimated to high temperatures are generally slightly elevated compared to cold acclimated animals (~7%; = 0.0692282, 95% CI: 0.0352333 to 0.1032231, p < 0.001), acclimatization is not likely going to provide adaptive benefits under climate change.

|  |
| --- |
| **Fig** 1**.** Taxonomic distribution of acute and acclimation Q10 estimates across major habitats. **A)** Phylogenetic distribution of taxa contained within the data. The total number of acute and acclimation type Q10 effect sizes are highlighted as well as whether the taxa is marine, freshwater or terrestrial. Silouettes are representative taxa of major clades within the tree. **B)** Acute and Acclimation lnRR Q10 across marine, freshwater, and terrestrial environments. **C)** lnCVR Q10 across traits for marine, freshwater and terrestrial systems. Note there were no differences between acute and acclimation Q10 types. k = total number of effect size estimates while the numbers in brackets indicate the number of species. |

|  |
| --- |
| Acute and Acclimation lnRR q10 across traits for A) marine, B) freshwater and C) terrestrial systems |

### *Does the opportunity for selection differ across terrestrial and aquatic ectotherms?*

### *Does climate variability predict acclimation capacity among aquatic and terrestrial ectotherms?*

## **Discussion**

One explanation for why terrestrial ectotherms show minimal acclimation capacity may be related to the fact that terrestrial ectotherms, were, on average acclimated for significantly less time than ectotherms from aquatic habitats . Rates of acclimation have been shown to be faster for many terrestrial groups compared to aquatic organisms [e.g., amphibians and reptiles have higher rates of acclimation than fishes; See Einum & Burton (2023)]. However, faster rates of acclimation would result in opposite patterns to those we observed – in other words, terrestrial species would be more likely to exhibit lower compared to when controlling for acclimation time.

## References

Bulgarella, M., Trewick, S.A., Godfrey, A.J.R., Sinclair, B.J. & Morgan-Richards, M. (2015). Elevational variation in adult body size and growth rate but not in metabolic rate in the tree weta hemideina crassidens. *Journal of Insect physiology*, 75, 30–38.

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, 80, 1–28., doi:10.18637/jss.v080.i01.

Bürkner, P.-C. (2018). Advanced bayesian multilevel modeling with the r package brms. *The R Journal*, 10, 395–411.

Einum, S. & Burton, T. (2023). Divergence in rates of phenotypic plasticty among ectotherms. *Ecology Letters*, 26, 147–156.

Grafen, A. (1989). The phylogenetic regression. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 326, 119–157.

Gunderson, A.R. & Stillman, J.H. (2015). Plasticity in thermal tolerance has limited potential to buffer ectotherms from global warming. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20150401.

Havird, J.C., Neuwald, J.L., Shah, A.A., Mauro, A., Marshall, C.A. & Ghalambor, C.K. (2020). Distinguishing between active plasticity due to thermal acclimation and passive plasticity due to Q10 effects: Why methodology matters. *Functional Ecology*, 0, 1–14.

Hedges, L.V., Gurevitch, J. & Curtis, P.S. (1999). The meta-analysis of response ratios in experimental ecology. *Ecology*, 80, 1150–1156.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., *et al.* (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146, 1999–2049.

Lajeunesse, M.J. (2011). On the meta-analysis of response ratios for studies with correlated and multi-group designs. *Ecology*, 92, 2049–2055.

Michonneau, F., Brown, J.W. & Winter, D.J. (2016). Rotl: An r package to interact with the open tree of life data. *Methods in Ecology and Evolution*, 7, 1476-1481. doi:10.1111/2041-210X.12593.

Nakagawa, S., Lagisz, M., O’Dea, R.E., Rutkowska, J., Yang, Y., Noble, D.W.A., *et al.* (2021a). The orchard plot: Cultivating forest plots for use in ecology, evolution and beyond. *Research Synthesis Methods*, 12, 4–12.

Nakagawa, S., Poulin, R., Mengersen, K., Reinhold, K., Engqvist, L., Lagisz, M., *et al.* (2015). Meta-analysis of variation: Ecological and evolutionary applications and beyond. *Methods in Ecology and Evolution*, 6, 143–152.

Nakagawa, S. & Santos, E.S. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology*, 26, 1253–1274.

Nakagawa, S., Senior, A.M., Viechtbauer, W. & Noble, D.W.A. (2021b). An assessment of statistical methods for non-independent data in ecological meta-analyses: comment. *Ecology*, in press., https://doi.org/10.1002/ecy.3490.

Noble, D.W.A., Pottier, P., Lagisz, M., Burke, S., Drobniak, S.M., O’Dea, R.E., *et al.* (2022). Meta-analytic approaches and effect sizes to account for “nuisance heterogeneity” in comparative physiology. *Journal of Experimental Biology*, 225, jeb243225.

Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, 5, 210–220.

Paradis, E. & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526–528.

Pick, J.L., Nakagawa, S. & Noble, D.W.A. (2019). Reproducible, flexible and high throughput data extraction from primary literature: The metaDigitise r package. *Methods in Ecology and Evolution*, 10, 426–431.

Pinsky, M.L., Eikeset, A.M., McCauley, D.J., Payne, J.L. & Sunday, J.M. (2019). Greater vulnerability to warming of marine versus terrestrial ectotherms. *Nature*, 569, 108–111.

Revell, L.J. (2012). Phytools: An r package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, 3, 217–223.

Seebacher, F., White, C.R. & Franklin, C.E. (2015). Physiological plasticity increases resilience of ectothermic animals to climate change. *Nature Climate Change*, 5, 61–66.

Song, C., Peacor, S.D., Osenberg, C.W. & Bence, J.R. (2021). An assessment of statistical methods for nonindependent data in ecological meta-analyses. *Ecology*, e03184.

Stan development team. RStan: The r interface to stan. (2021). *R package version 2.21.3. https://mc-stan.org/.*

Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, 36, 1–48. URL: https://www.jstatsoft.org/v36/i03/.

Yu, G., Smith, D., Zhu, H., Guan, Y. & Lam, T.T.-Y. (2017). Ggtree: An r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8, 28–36, doi:10.1111/2041–210X.12628.

## Supplemental Results and Figures

|  |
| --- |
| I2 estimates |

|  |
| --- |
| I2 estimates lnCVR |