Results

# Methods

### *Effect size*

We calculated the standardized mean difference with heteroscedastic population variances () (Bonett, 2008, 2009) as our effect size metric for each study along with its associated sampling error. We choose to use as our effect size because our data: 1) contained ratios (e.g., respiratory control ratios, RCR) which makes interpretation with alternative effect size measures, such as log response ratio, challenging; 2) contained non-ratio-scale data and 3) We calculated as the difference in means between control and treatment groups divided by the pooled standard deviation. We used the Hedge’s g correction to account for small sample sizes and bias in the estimation of the pooled standard deviation . We calculated for each study using the *escalc* function in *metafor* package in R (vers. 4.6.0) (Viechtbauer, 2010).

### *Meta-Analysis*

We analysed our data using multilevel meta-analytic (MLMA) and meta-regression (MLMR) models in R (vers. 4.3.1) using *metafor* (vers. 4.6.0 Viechtbauer, 2010). Meta-analytic mean estimates and meta-regression models were plotted using orchard plots which convey overall meta-analytic means, alongside raw effect sizes and prediction intervals, which describe effect heterogeneity (vers. 2.0, Nakagawa *et al.*, 2021, 2023).

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

Using data subsets for each of our major stressor categories, we first fit multi-level meta-analysis (MLMA) models (i.e., intercept-only models) with , that included study, species, trait type, and phylogeny as random effects to account for non-independence and identify sources of variability. Our MLMA models allowed us to partition variation in among these key sources while accounting for total sampling variance. This allowed us to calculate the proportion of 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). As a second measure of heterogeneity we also present 95% prediction intervals which describe the expected distribution of effects that are expected from future studies (Nakagawa *et al.*, 2021; Noble *et al.*, 2022).

A phylogeny was derived by first cross-checking taxa names using the Open Tree of Life (OTL) with the *rotl* package in R (vers. 3.1.0) (Michonneau, Brown, & Winter, 2016), and plotted using *ggtree* to visualize the tree (vers. 3.9.0) (Yu *et al.*, 2017). We resolved any taxa names that were outdated or changed. Once names in the dataset and *rotl* database matched we exported taxa names and built a time calibrated phylogeny for the species in our dataset using TimeTree [timetree.org; ]. We used the R packages *ape* (vers. 5.7.1) (Paradis & Schliep, 2019) and *phytools* (vers. 1.9.16) (Revell, 2012) to prune the tree for individual analyses and calculate phylogenetic correlation matrices used in meta-analytic models.

#### Multi-level Meta-Regression (MLMR) Models

After quantifying levels of heterogeneity in each of the different stressor datasets, 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, however, MLMR models differed in the moderators included because each stressor type had unique sources of heterogeneity that we expected would explain effect size heterogeneity.

# Results

Bonett, D.G. (2008) Confidence intervals for standardized linear contrasts of means. *Psychological Methods* **13**, 99. American Psychological Association.

Bonett, D.G. (2009) Meta-analytic interval estimation for standardized and unstandardized mean differences. *Psychological methods* **14**, 225. American Psychological Association.

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

Nakagawa, S., Lagisz, M., O’Dea, R.E., Pottier, P., Rutkowska, J., Senior, A.M., Yang, Y. & Noble, D.W. (2023) orchaRd 2.0: An r package for visualising meta-analyses with orchard plots. *Methods in Ecology and Evolution* **14**, 2003–2010. Wiley Online Library.

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

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

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

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

Revell, L.J. (2012) Phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223.

Viechtbauer, W. (2010) Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **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 Ecol. Evol.* **8**, 28–36, doi:10.1111/2041–210X.12628.