# Methods

### *Effect size*

To quantify the impacts of developmental stress on mitochondrial function we calculated the standardized mean difference, assuming heteroscedastic population variances () (Bonett, 2008, 2009) along with its associated sampling error. We used the *escalc* function of the *metafor* package (vers. 4.8.0) (Viechtbauer, 2010) in R (vers. 4.5.1) which implements a correction to account for possible bias resulting from effect sizes calculated with small sample sizes (Borenstein & Hedges, 2009). We choose to use as our effect size because our data: 1) contained substantial amounts of ratio data (e.g., respiratory control ratios, RCR, relative gene expression) which makes interpretation with alternative effect size measures, such as log response ratio, challenging; 2) percentages and zero measurement variables and 3) skewed measurement variables. is more robust to these types of measurement variables.

Given the diversity of measurement variables we paid particular attention to effect size direction to ensure that positive increase of means all represent increased in mitochondrial function. ONDI ELABORATES AND TABLE

We calculated as the mean difference between control and treatment groups divided by the pooled standard deviation. As such, positive effect sizes represent situations where the mean of the experimental group was larger than the control group. Again, we corrected the direction of effect size (multiplied by -1) to ensure their interpretation remained consistent with respect to mitochondrial function.

### *Meta-Analysis*

We analysed our data using multilevel meta-analytic (MLMA) and meta-regression (MLMR) models in R using *metafor* (vers. 4.8.0 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, tissue type, and phylogeny as random effects to account for non-independence and identify sources of variability. We also explored models that included species and sample dependency as random effects but these were highly confounded with phylogeny and study so we simplified models by removing these terms. Given *metafor* does not estimate a residual variance by default we also included an observation-level random effect in our models. The size of datasets varied in the number of species and tissues ([Table 1](#tbl-heterogeneity)). As such, we simplified the random effect structure for some datasets; including only random effects with six or more levels. Our MLMA models allowed us to partition variation in among these key sources while accounting for total sampling variance. We took a pluralistic approach to heterogeneity reporting, by reporting absolute (), relative (i.e., , *sensu* Nakagawa & Santos (2012); Noble *et al.* (2022)) and magnitude measures (both and ) of heterogeneity as suggested by Yang *et al.* (2023). We also report each of their stratified versions for each random effect level (Nakagawa & Santos, 2012; Yang *et al.*, 2023). 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 plotting taxa using *ggtree* to visualize the tree (vers. 3.16.0) (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 species names and built a time calibrated phylogeny for the species in our dataset using TimeTree [timetree.org; Kumar *et al.* (2022)]. We used the R packages *ape* (vers. 5.8.1) (Paradis & Schliep, 2019) and *phytools* (vers. 2.4.4) (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 variance. A number of moderators were relevant to all stressor manipulation types and so were included in all models. These included: 1) measurement trait category (i.e., antioxidants, oxidative damage etc.), 2) relative total duration of stressor manipulation, 3) We included relative total duration of stressor manipulation as a continuous moderator because we expected that longer durations of stressor manipulations would have larger impacts on mitochondrial function. We also fit a model comparing stage of development (i.e., pre- or post-natal) as a categorical moderator to test whether pre- or post-natal manipulations had larger impacts on mitochondrial function. However, this moderator did not show differences between pre- and post-natal manipulations, was not relevant for some manipulations (e.g., social deprivation because only one stage was manipulated), and was redundant with relative duration of stressor. As such, we include results and plots for this model in the supplementary materials [Figure 5](#fig-orchard_prepost). We fit models assuming heteroscedastic variation among levels of moderators to capture differences in effect variation.

# Results

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1- Heterogeneity estimates for developmental stressor datasets   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Corticosterone | 140 | 8 | 10 | 86 | 12.2 | 0 | 68 | 6.6 | | Disturbance | 190 | 8 | 16 | 80 | 57.2 | 11 | 0 | 11.8 | | Nutrition | 645 | 15 | 51 | 80 | 8.6 | 13 | 10 | 47.9 | | Social Deprevation | 172 | 2 | 10 | 82 | 36.8 | – | – | 45.1 | |

|  |
| --- |
| Figure 1- Phylogeny of species included in the meta-analysis. |

In total, we collected data from 1147 effect sizes across 21 species ([Figure 1](#fig-phylo)) and 86 studies across the different stressors ([Table 1](#tbl-heterogeneity)).

## Corticosterone developmental impacts on mitochondrial function

Overall, exposure to corticosterone during development negatively impacted mitochondrial function ( = -0.92, 95% CI = -1.69 to -0.16, = 0.02; [Figure 2](#fig-orchard_int)a). However, effect heterogeneity was high (95% Prediction Intervals: -3.07 to 1.23; [Table 1](#tbl-heterogeneity)), with 12.24% of variation being driven by differences among studies (; [Table 1](#tbl-heterogeneity)) and 67.53% of variation being driven by differences among tissues (; [Table 1](#tbl-heterogeneity)) with little to no variation explained by phylogeny ([Table 1](#tbl-heterogeneity)).

## Social deprivation developmental impacts on mitochondrial function

Social deprevation studies were only ever done with mammals with all studies manipulating environments postnatally. Social deprivation during development had a negative impact on mitochondrial function ( = -0.8, 95% CI = -1.4 to -0.21, = 0.01; [Figure 2](#fig-orchard_int)b). Effect heterogeneity was again high (95% Prediction Intervals: -3.43 to 1.82; [Table 1](#tbl-heterogeneity)), with 36.82% of the variation being driven by differences among studies () ([Table 1](#tbl-heterogeneity)).

## Disturbance developmental impacts on mitochondrial function

Disturbance during development had a positive impact on mitochondrial function ( = -0.8, 95% CI = -1.54 to -0.06, = 0.03; [Figure 2](#fig-orchard_int)c). Effect heterogeneity was high (95% Prediction Intervals: -3.17 to 1.57; [Table 1](#tbl-heterogeneity)), with substantial variation across studies ( = 57.25%)([Table 1](#tbl-heterogeneity)).

Disturbance stress occured only postnatally in fish and birds, whereas in mammals disturbance stressors were also applied prenatally.

## Nutrition developmental impacts on mitochondrial function

Nutritional stress during development had limited impact on mitochondrial function overall ( = -0.38, 95% CI = -1.03 to 0.27, = 0.25; [Figure 2](#fig-orchard_int)d). Effect heterogeneity was high (95% Prediction Intervals: -2.85 to 2.08; [Table 1](#tbl-heterogeneity)), with variation being driven by differences between studies ( = 8.57%), phylogeny ( = 13.21%) and tissue ( = 10.29%) ([Table 1](#tbl-heterogeneity)).

|  |
| --- |
| Figure 2- Orchard plots showing the overall meta-analytic mean effect size and 95% prediction intervals for each stressor category. |

|  |
| --- |
| Figure 3- Orchard plots showing the overall meta-analytic mean effect size and 95% prediction intervals for each stressor category for the different measurement types. |

|  |
| --- |
| Figure 4- Orchard plots showing the overall meta-analytic mean effect size and 95% prediction intervals for each stressor category for the different measurement types. |

|  |
| --- |
| Figure 5- Orchard plots showing the overall meta-analytic mean effect size and 95% prediction intervals for each stressor category for pre-natal versus postnatal manipulations. |

# References

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.

Borenstein, M. & Hedges, L.V. (2009) Effect sizes for meta-analysis. *In: The Handbook of Research Synthesis and Meta-analysis (eds Cooper H, Hedges LV, Valentine JC). Russell Sage Foundation, New York.*, 207–243.

Kumar, S., Suleski, M., Craig, J.M., Kasprowicz, A.E., Sanderford, M., Li, M., Stecher, G. & Hedges, S.B. (2022) TimeTree 5: An expanded resource for species divergence times. *Molecular biology and evolution* **39**, msac174. Oxford University Press.

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/.

Yang, Y., Noble, D.W., Spake, R., Senior, A.M., Lagisz, M. & Nakagawa, S. (2023) Measuring biological generality in meta-analysis: A pluralistic approach to heterogeneity quantification and stratification. *EcoEvoRxiv*.

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.

# Supplementary Materials