

Human disturbance and the magnitude and predictability of phenotypic divergence

Santiago Ortega^{1*}, Eduardo S.A. Santos¹, Kiyoko Gotanda², Sarah Sanderson³,
Lucas Damian Gorné^{4,5}, Andrew Hendry³, and Shinichi Nakagawa^{1,6}

*Corresponding author: Santiago Ortega (santiago.ortega@ualberta.ca)

¹Collaboration for Open Science and Synthesis in Ecology and Evolution (COSSEE),
Department of Biological Sciences, Faculty of Science, University of Alberta, Edmonton,
Canada

²Brock University, Ontario, Canada

³McGill University, Quebec, Canada

⁴Universidad Nacional de Córdoba, Facultad de Ciencias Exactas, Físicas y Naturales,
Córdoba, Argentina

⁵Consejo Nacional de Investigaciones Científicas y Técnicas, CONICET, IMBiV, Córdoba,
Argentina

⁶Evolution and Ecology Research Centre, School of Biological, Earth and Environmental
Sciences, University of New South Wales, Sydney, New South Wales, Australia

October 6, 2026

1 Abstract

Environmental change can drive phenotypic change within species, but it remains unclear whether such responses produce a general and predictable degree of phenotypic divergence across populations. We conducted a preregistered reanalysis of PROCEED v6.2, comprising 7,186 contrasts from 254 references and 253 species, to test whether the magnitude and predictability of phenotypic divergence differ among disturbance contexts and across temporal, biological, and methodological characteristics of studies. We used $\ln M$ to quantify phenotypic differences between paired population samples relative to the variation retained within them and fitted phylogenetic location-scale meta-regressions to model average divergence and residual heterogeneity jointly. Across the dataset, average phenotypic divergence was substantial but did not consistently exceed within-population variation, while total heterogeneity reached 96.6%, providing little support for a universal magnitude of divergence. Disturbance contexts differed more strongly in predictability than in average divergence. Landscape change and introduced populations showed greater divergence than *in situ* natural variation, while landscape change was also comparatively predictable. By contrast, responses to introduced species consistently showed greater variation among contrasts than most other contexts, whereas the lower predictability of hunting or harvesting and *in situ* natural variation depended partly on a single large Pacific salmon dataset. Moderator associations should be interpreted cautiously because study characteristics were unevenly represented and partly confounded, and some results were sensitive to robustness analyses. Overall, environmental change can produce substantial phenotypic divergence, but neither its magnitude nor its predictability is uniform across systems. Accounting for heterogeneity therefore reveals when broad comparative responses are likely to generalize and when population-specific evidence is required.

1.1 Keywords

contemporary evolution; human disturbance; intraspecific variation; phenotypic heterogeneity; location-scale meta-analysis; global change

2 Introduction

Wild populations experience ecological conditions that are increasingly reshaped by human activities, altering their selective, developmental, and demographic environments (Alberti et al., 2017; Hendry et al., 2017). As a consequence of human activities, within-species phenotypic divergence has been documented across diverse taxa and traits, but its magnitude varies considerably among populations (Sanderson et al., 2022). Disturbances differ in intensity, duration, and spatial extent, while populations differ in their evolutionary histories and capacities for plastic and evolutionary change (Clement et al., 2023; Diamond and Martin, 2016; Turner, 2010). Accordingly, although previous syntheses found only small average differences in phenotypic change between human-disturbed and natural systems, they also revealed substantial variation among disturbance contexts and individual biological systems (Hendry et al., 2008; Sanderson et al., 2022). The central question is therefore not only how strongly populations diverge phenotypically, but whether disturbances produce a predictable phenotypic signature across populations.

Variation around the average response may contain important biological information itself. Similar average phenotypic divergence can arise because most populations respond by comparable amounts or because responses range from negligible to extreme (Sanderson et al., 2022, 2023). These alternatives may reflect different underlying processes and provide very different levels of predictability for conservation and management (Des Roches et al., 2017; Palkovacs et al., 2012). Moreover, the biological importance of a difference in trait means depends on the variation retained within the populations being compared: the same difference represents stronger phenotypic divergence when within-population variation is low (Lynch, 1990; Nakagawa et al., 2026a). Variation around the average may therefore be important not only biologically, but also for how comparative evidence is synthesized. Conventional meta-regressions typically estimate moderator effects on the average while assuming that residual heterogeneity is constant across disturbance contexts, traits, and study designs. If heterogeneity differs among these groups, imposing a common variance can conceal differences in cross-system predictability and change both estimated average effects and their uncertainty (Nakagawa et al., 2025; Yang et al., 2025). Previous conclusions may therefore be incomplete because they describe average phenotypic divergence without establishing how reliably that average generalizes across systems.

Here, we reanalyse the Phenotypic Rates of Change Evolutionary and Ecological Database (PROCEED) (Gorné et al., 2025; Gotanda and Gorné, 2026) following a preregistered protocol (Ortega et al., 2026). Throughout, we use phenotypic divergence to mean the magnitude of the phenotypic difference between

paired samples relative to the variation within them, irrespective of its direction. The term applies equally to different populations compared a known time after they split (synchronic contrasts) and to the same population compared at different times (allochronic contrasts). We quantify this divergence using $\ln M$, the log ratio of the between-sample standard deviation to the within-sample standard deviation, and fit multilevel phylogenetic location-scale meta-regressions. Our primary analysis compares six forms of human disturbance—climate change, hunting or harvesting, introduced populations, landscape change, pollution, and native responses to introduced species—with *in situ* natural variation, defined as contrasts arising from naturally occurring environmental differences without an identified human disturbance. For each moderator, we ask whether the expected magnitude of divergence differs among its levels and how predictable that magnitude is across contrasts. Additional models examine comparison design, elapsed years and generations, trait type, and study type to assess whether the magnitude and predictability of divergence also vary with temporal structure, biological attributes, and study methodology. By modelling average phenotypic divergence and residual heterogeneity jointly, we reassess both the magnitude of the phenotypic signature associated with human disturbance and the extent to which it generalizes across wild populations.

3 Methods

3.1 Data source and analytical contrasts

We used PROCEED version 6.2, a global compilation of quantitative within-species phenotypic comparisons, either between populations or between samples of the same population (Gorné et al., 2025; Gotanda and Gorné, 2026). We retained contrasts reporting the mean, standard deviation, and sample size for both samples and, following Sanderson et al. (2022), restricted the dataset to comparisons with known elapsed intervals of ≤ 300 generations. We imposed no minimum sample size in the primary analyses (smallest combined sample size, $n_1 + n_2 = 5$); restricting the data to contrasts with $n_1 + n_2 \geq 40$ was instead evaluated as a prespecified sensitivity analysis. PROCEED distinguishes two comparison designs. Allochronic contrasts compare the same population or study system at different time points and therefore estimate phenotypic change along a temporal trajectory. Synchronic contrasts compare different populations or conditions sampled a known time after they split and therefore estimate the phenotypic divergence between them (Gotanda and Gorné, 2026). For allochronic contrasts, elapsed years represent the interval between the earlier and later samples; for synchronic contrasts, they represent the interval from the populations’ split to sampling. Elapsed generations were calculated by dividing elapsed years by species generation time. PROCEED orders

83 samples chronologically or, when identifiable, from the source, ancestral, or original condition to the derived
84 or novel condition. This ordering does not affect $\ln M$, which quantifies the magnitude rather than the
85 direction of divergence.

86 Disturbance type was the primary moderator. It distinguished climate change, hunting or harvesting, intro-
87 duced populations, landscape change, pollution, and native responses to introduced species from contrasts
88 representing *in situ* natural variation. Secondary models examined comparison design, elapsed time in years
89 or generations, and trait type, classified as behaviour, body size, growth, other life-history, other morphol-
90 ogy, phenology, physiology, or response traits. We also examined study type, distinguishing measurements
91 of individuals that developed under natural conditions from estimates based on common-garden breeding or
92 quantitative-genetic approaches. This classification did not imply that field-observed phenotypic differences
93 were caused exclusively by plasticity or lacked a genetic basis. Definitions and coding criteria for all mod-
94 erators are provided in the PROCEED metadata (Gorné et al., 2025). The primary models included 7,186
95 contrasts from 254 source references and 253 species (Fig. 1).



Figure 1: **Taxonomic representation of the 7,186 contrasts included in the analyses.** Rectangle area is proportional to the number of contrasts contributed by each taxonomic group. Labels report the total numbers of contrasts and canonical species. Animal silhouettes were provided by Anna Lenz.

3.2 Effect-size estimation

We quantified phenotypic divergence using the log-magnitude effect size, $\ln M$, which expresses the phenotypic difference between two population samples relative to the variation within them (Nakagawa et al., 2026a). For two samples with means $\bar{x}_1$ and $\bar{x}_2$, standard deviations s_1 and s_2 , and sample sizes n_1 and n_2 , the between-group and pooled within-group mean squares are:

$$\begin{aligned} \text{MS}_B &= \frac{n_1 n_2}{n_1 + n_2} (\bar{x}_1 - \bar{x}_2)^2, \\ \text{MS}_W &= \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}. \end{aligned} \tag{1}$$

The estimated between-group variance component and pooled within-group variance are:

$$\begin{aligned} s_B^2 &= \frac{\text{MS}_B - \text{MS}_W}{n_0}, \quad n_0 = \frac{2n_1 n_2}{n_1 + n_2}, \\ s_W^2 &= \text{MS}_W, \end{aligned} \tag{2}$$

where n_0 is the harmonic mean of the two sample sizes. The effect size is then defined as:

$$\ln M = \ln \left(\frac{s_B}{s_W} \right). \tag{3}$$

Thus, $\ln M$ is the log ratio of the between-group, or mean-difference, standard deviation to the pooled within-group standard deviation. Positive values indicate that the between-sample standard deviation exceeds the within-sample standard deviation, whereas negative values indicate the reverse. Importantly, $\ln M = 0$ means that $s_B = s_W$, not that the two sample means are equal. Because MS_B depends on the squared difference between means, $\ln M$ quantifies the magnitude rather than the direction of phenotypic divergence.

The analytical estimator becomes unstable as MS_B approaches MS_W and is undefined when $\text{MS}_B \leq \text{MS}_W$. We therefore used SAFE (Nakagawa et al., 2026b), a single-fit parametric bootstrap implemented in `orchaRd` [v2.2.0], to estimate $\ln M$ (y_i) and its sampling variance (v_i) for each contrast i (Supplementary material S1).

3.3 Phylogeny

We built a phylogenetic scaffold from the Open Tree of Life (Hinchliff et al., 2015) using `rot1` [v3.1.1] (Michonneau et al., 2016) and `prepR4pcm` [v1.0.0] (Nakagawa et al., 2026c). Taxonomic reconciliation was performed across the complete PROCEED dataset rather than separately for each model. Within the analytical dataset, 261 of 267 submitted taxon names were resolved. Synonyms and infraspecific names, including subspecies, were then collapsed to their accepted species-level canonical names, yielding 253 unique species for the phylogeny-adjusted models. The six unresolved names, representing 51 contrasts, were excluded. We retrieved the induced subtree containing the resolved taxa and used `ape` [v5.8-0] (Paradis et al., 2004) to resolve multifurcations randomly into bifurcations with zero-length branches. When branch lengths were unavailable, we assigned Grafen branch lengths with power $\rho = 1$ (Grafen, 1989). Finally, we converted the resulting tree into a species-level phylogenetic correlation matrix under a Brownian-motion model of trait evolution using the function `prepR4pcm::pr_phylo_cor`.

3.4 Phylogenetic location–scale meta-regression

We fitted Bayesian multilevel location–scale meta-regressions using `brms` [v2.23.0] with `cmdstanr` [v0.9.0.9000] (Bürkner, 2017; Gabry et al., 2025; Nakagawa et al., 2025; Stan Development Team, 2025) in R [v4.6.1] (R Core Team, 2026). We analysed moderators in separate models because data availability differed among moderators and several moderators were strongly associated within the dataset (Figs. S1–S3). Fitting them jointly would therefore reduce the available data and estimate conditional effects from highly uneven combinations of moderator levels. Each model included the focal moderator in both the location and scale components. We modelled the average magnitude of phenotypic divergence (y_i) as:

$$y_i = \mu_i + u_i^{(\text{ref})} + a_i + \epsilon_i, \quad (4)$$

where

$$\mu_i = \beta_0^{(l)} + \sum_{k=1}^p \beta_k^{(l)} x_{ki}. \quad (5)$$

Here, μ_i is the expected $\ln M$, $\beta_0^{(l)}$ is the location intercept, and $\beta_k^{(l)}$ is the coefficient for moderator x_k . The

terms $u_i^{(\text{ref})}$ and a_i represent random effects for source reference and phylogenetic species, respectively. We modelled the residual term as:

$$\epsilon_i \sim N(0, v_i + \sigma_i^2), \quad (6)$$

where v_i is the known sampling variance and σ_i^2 is the residual heterogeneity. To test whether residual heterogeneity differed among moderator levels, we modelled the residual standard deviation as:

$$\log \sigma_i = \beta_0^{(s)} + \sum_{m=1}^q \beta_m^{(s)} z_{mi}. \quad (7)$$

Positive $\beta_m^{(s)}$ values indicate greater residual heterogeneity (i.e., less predictable magnitude of divergence between groups). We estimated location and scale components jointly. As a prespecified sensitivity analysis, we refitted each model with study system included as an additional random effect to account for potential dependence among contrasts from the same population or research system (Table S1).

3.5 Posterior inference, heterogeneity, robustness, and reproducibility

We based inference for continuous moderators on posterior slopes and for categorical moderators on population-level marginal means and pairwise contrasts. We report posterior means with 95% equal-tailed credible intervals (CrI) and treated intervals excluding zero as well supported. Using the intercept-only model, we calculated total and component-specific I^2 values to partition heterogeneity among source references, phylogenetically structured species effects, and residual variation (Nakagawa et al., 2025). Models were fitted using four chains of 4,000 iterations, including 2,000 warm-up iterations, and convergence was assessed from $\hat{R}$, effective sample sizes, divergent transitions, and maximum tree-depth warnings (Vehtari et al., 2021). Robustness was evaluated using three prespecified sensitivity analyses: restricting the data to contrasts with a combined sample size of at least 40, removing the phylogenetic random effect, and adding a study-system random effect. The study-system analysis is summarized in Table S1, and complete coefficient-level results and convergence diagnostics for all sensitivity models are available in the online supplement (https://santiago-Ortega.github.io/Magnitude_pace_modern_life/).

We additionally assessed small-study effects using the half harmonic-mean sample size, $\tilde{n}_0 = n_0/2$, where n_0 was defined above. Following the modified Egger-type approach developed for $\ln M$ (Nakagawa et al.,

2026a), we used $1/\sqrt{\hat{n}_0}$ as a sample-size-based precision predictor rather than the sampling standard error, because the latter is structurally correlated with $\ln M$ and would therefore induce an artefactual slope. This concern is specific to regressing $\ln M$ on its own standard error; in the primary models, v_i entered only as the known sampling variance (Eq. 6) and not as a moderator. We included $1/\sqrt{\hat{n}_0}$ in both the location and scale components: the location slope tested whether average $\ln M$ varied systematically with study size, whereas the scale slope tested whether residual heterogeneity, and hence the predictability of phenotypic divergence, varied with study size. These analyses were treated as diagnostics of small-study effects rather than direct tests of publication bias (Egger et al., 1997; Nakagawa et al., 2022).

We followed the PRISMA-EcoEvo (O’Dea et al., 2021) reporting guidance and provide the completed checklist as supplementary material. All analytical code is available at https://github.com/Santiago-Ortega/Magnitude_pace_modern_life, and a step-by-step reproducible account of data preparation, effect-size estimation, model fitting, diagnostics, sensitivity analyses, and figure generation is available at https://santiago-Ortega.github.io/Magnitude_pace_modern_life/. Figures were produced in R using `ggplot2` [v4.0.3] (Wickham, 2011), `ggbeeswarm` [v0.7.3] (Clarke et al., 2025), and `patchwork` [v1.3.2] (Pender-
sen, 2025). To aid interpretation, we converted the location estimates to approximate d -equivalent values, d_{eq} (Nakagawa et al., 2026a). These values are reported in the text and displayed on secondary axes in the location panels; pairwise contrasts and regression slopes were retained on the $\ln M$ scale because they represent differences or changes rather than divergence itself. On the conventional standardized mean-difference scale, values of $d = 0.2, 0.5$, and 0.8 are commonly described as small, medium, and large, respectively (Cohen, 1988; Lakens, 2013). By comparison, $\ln M = 0$ corresponds approximately to $d_{eq} = 1.41$, indicating a substantial difference between group means relative to within-group variation.

3.6 Deviations from the preregistered protocol

Following inspection of the data structure, we conducted additional analyses that were not specified in the preregistered protocol to assess potential confounding between categorical moderators and elapsed time. For each of the main categorical moderators, we refitted the location–scale model with elapsed time included as an additive covariate, using $\log_{10}$ elapsed years and, separately, $\log_{10}$ elapsed generations. These models were used to assess whether moderator associations were robust to differences in the elapsed intervals represented within moderator categories. The time-adjusted models for comparison design and study type converged satisfactorily and were interpreted as supplementary robustness analyses. Equivalent models for disturbance context and trait type did not meet our prespecified convergence criteria and were therefore not interpreted.

All preregistered analyses were retained as originally specified. We additionally conducted a non-preregistered leave-one-reference-out sensitivity analysis for Oke et al. (2020), a long-term study of Pacific salmon body size based on data compiled by Clark et al. (2018), which enters PROCEED as a single reference. This reference contributed 811 of the 913 hunting or harvesting contrasts (88.8%), all from five *Oncorhynchus* species (of the 17 species in that category), and 473 of the 1,316 *in situ* natural-variation contrasts (35.9%). Excluding it left 102 hunting or harvesting contrasts from 12 species and 18 references, none involving *Oncorhynchus*. We refitted the disturbance-context model without these contrasts to assess whether conclusions for both categories were driven by this single reference.

4 Results

4.1 Data overview

Most contrasts compared populations or samples measured synchronically (65.1%, $n = 4,678$), whereas the remaining 34.9% were allochronic comparisons ($n = 2,508$). Introductions were by far the most frequently represented disturbance context (57.8%, $n = 4,156$). *In situ* natural variation (18.3%, $n = 1,316$) and hunting or harvesting (12.7%, $n = 913$) were the next most common contexts. Climate change, landscape change, responses to introduced species, and pollution each contributed less than 4% of the available contrasts. Comparison design was strongly associated with disturbance context (Cramér’s $V = 0.90$; Fig. S1), indicating that synchronic and allochronic contrasts were not distributed independently across contexts. Studies of phenotypic differences expressed in wild populations predominated (63.4%, $n = 4,556$), while common-garden, quantitative-genetic, and related designs intended to distinguish persistent or heritable phenotypic differences from environmentally induced variation accounted for 36.6% of the dataset ($n = 2,630$; Fig. 2). The evidence base was also concentrated in morphological traits: body size (37.4%, $n = 2,690$) and other morphological traits (33.3%, $n = 2,393$) together accounted for 70.7% of the analytical sample. Other life-history traits represented 18.1% ($n = 1,299$), whereas physiological ($n = 422$), growth ($n = 215$), phenological ($n = 114$), response ($n = 34$), and behavioural traits ($n = 19$) were comparatively uncommon.

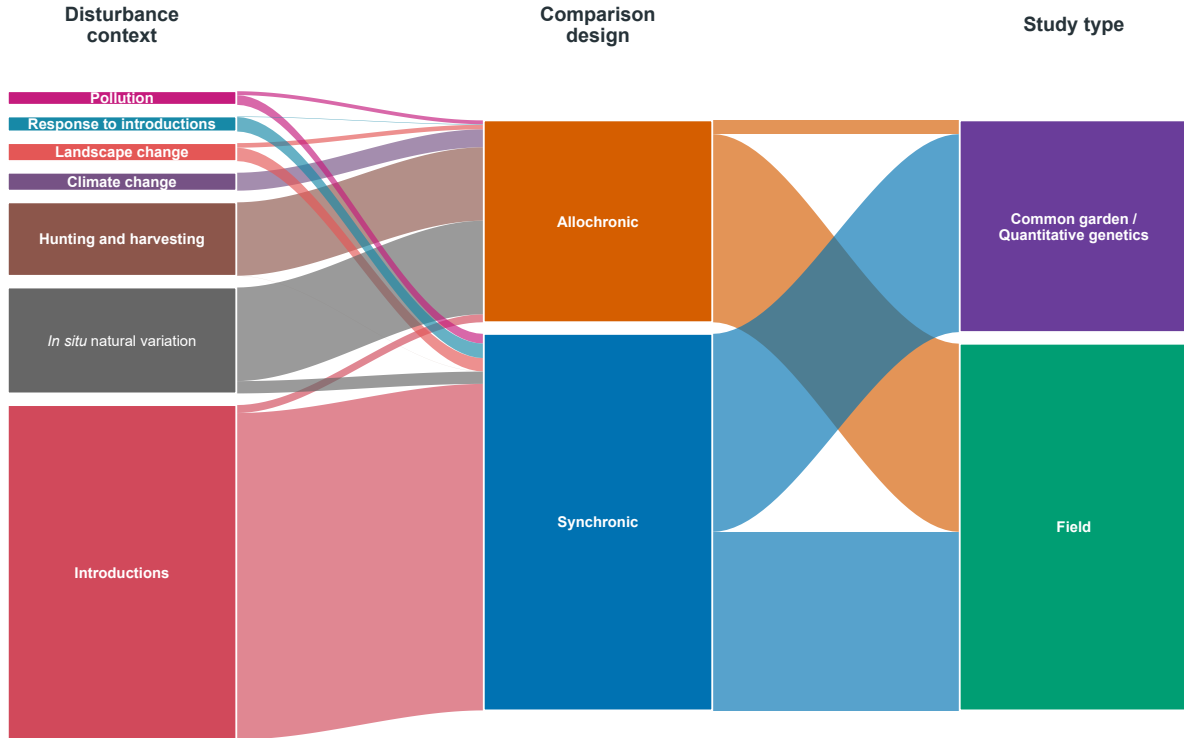


Figure 2: **Structure of the evidence base across comparison designs, study types, and disturbance contexts.** The alluvial diagram represents the 7,186 contrasts included in analyses. The height of each block and the width of each connecting ribbon are proportional to the number of contrasts. Ribbons show how contrasts were distributed across disturbance contexts—introductions involving comparisons of introduced populations, *in situ* natural variation, hunting or harvesting, climate change, landscape change, responses of other species to introductions, and pollution—and, within these contexts, between synchronic and allochronic comparisons and between field-based studies and common-garden or quantitative-genetic studies.

4.2 Overall magnitude and predictability of phenotypic divergence

Across the dataset, there was no clear tendency for phenotypic divergence between paired samples to exceed within-population variation ($\beta_{\text{Overall}}^{(l)} = -0.55$, 95% CrI: $[-1.41, 0.30]$). This reference point should not be interpreted as a threshold for biological importance. On the conventional standardized mean-difference scale, the same estimate corresponded to approximately $d_{\text{eq}} = 0.81$ (95% CrI: $[0.34, 1.92]$), indicating substantial divergence between sample averages. The magnitude of divergence also varied substantially among contrasts ($\sigma = 0.55$, 95% CrI: $[0.54, 0.57]$), resulting in low overall predictability across comparisons (Fig. 3). Furthermore, we detected a small-study effect in the location component ($\beta_{1/\sqrt{\tilde{n}_0}}^{(l)} = 1.75$, 95% CrI: $[1.55, 1.95]$), indicating that smaller studies tended to report larger $\ln M$ values. Small-study effects were also evident in the scale component ($\beta_{1/\sqrt{\tilde{n}_0}}^{(s)} = -1.37$, 95% CrI: $[-1.60, -1.15]$), corresponding to an approximately 75% reduction in residual standard deviation per one-unit increase in $1/\sqrt{\tilde{n}_0}$. Smaller studies therefore showed lower residual heterogeneity and apparently greater predictability. Accordingly, the absolute $\ln M$ values reported below should be interpreted cautiously, as divergence may be somewhat weaker in larger studies. Total heterogeneity was $I_{\text{total}}^2 = 96.6\%$ (95% CrI: $[95.5, 97.7]$), comprising variation attributable to source references (37.3%, $[23.8, 52.4]$), phylogenetically structured species effects (38.1%, $[16.2, 58.9]$), and residual differences among contrasts (21.0%, $[14.6, 28.4]$).

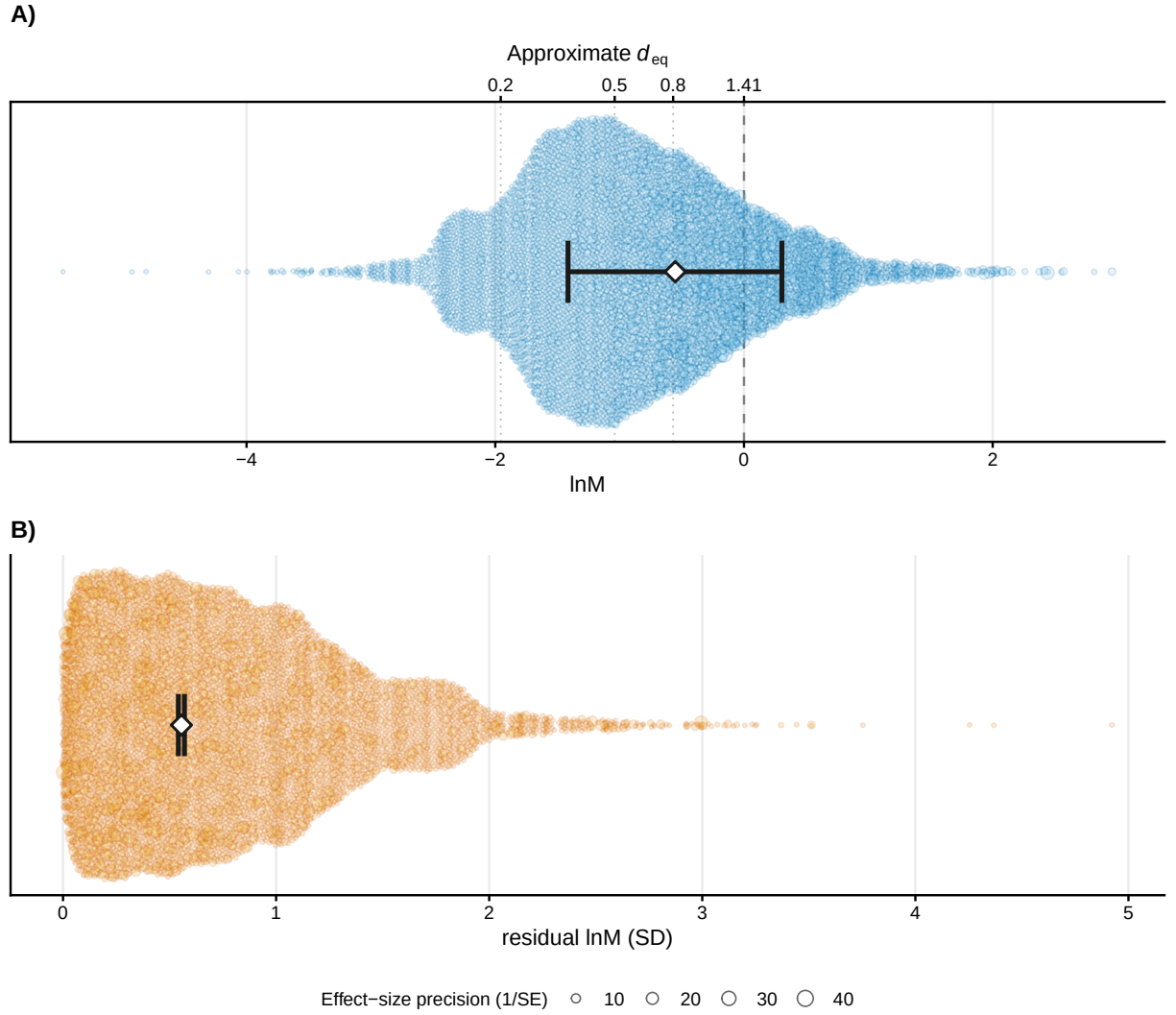


Figure 3: Overall magnitude and predictability of within-species phenotypic divergence across 7,186 contrasts. **(A)** Average magnitude of divergence between paired samples. The dashed line marks $\ln M = 0$ (approximately $d_{eq} = 1.41$), indicating substantial divergence between group means; the upper axis provides approximate d_{eq} values. **(B)** Predictability in the magnitude of divergence among contrasts, quantified by the residual standard deviation, σ . Larger σ values indicate lower predictability. White circles and black lines denote posterior estimates and 95% credible intervals, respectively, while bubble size indicates precision ($1/SE$).

4.3 Disturbance context

In the preregistered analysis, hunting or harvesting and *in situ* natural variation were the only contexts in which phenotypic divergence between samples was clearly smaller than variation within populations (hunting or harvesting: $\beta_{\text{Hunting}}^{(l)} = -0.81$, 95% CrI: $[-1.61, -0.04]$; natural variation: $\beta_{\text{Natural}}^{(l)} = -0.81$, $[-1.61, -0.04]$; Fig. 4A). Nevertheless, both corresponded to moderate standardized divergence between sample averages (hunting or harvesting: $d_{\text{eq}} = 0.63$, $[0.28, 1.36]$; natural variation: $d_{\text{eq}} = 0.63$, $[0.28, 1.36]$). For the remaining contexts, credible intervals overlapped $\ln M = 0$, so between-sample divergence could not be distinguished from within-sample variation; the two were of broadly comparable magnitude.

Pairwise comparisons showed that landscape change produced greater divergence than hunting or harvesting ($\beta_{\text{Landscape-Hunting}}^{(l)} = 0.483$, 95% CrI: $[0.133, 0.830]$) and *in situ* natural variation ($\beta_{\text{Landscape-Natural}}^{(l)} = 0.488$, $[0.148, 0.826]$). Introductions likewise showed greater divergence than hunting or harvesting ($\beta_{\text{Introduction-Hunting}}^{(l)} = 0.282$, $[0.026, 0.543]$) and *in situ* natural variation ($\beta_{\text{Introduction-Natural}}^{(l)} = 0.287$, $[0.038, 0.537]$). Evidence for the remaining pairwise location contrasts was inconclusive. However, Oke et al. (2020) contributed 811 of the 913 hunting or harvesting contrasts, and not all of the salmon populations it represents may be subject to hunting or harvesting (Sanderson et al., 2022). In the non-preregistered sensitivity analysis excluding contrasts from Oke et al. (2020), the hunting or harvesting estimate shifted from $\beta^{(l)} = -0.81$ to -0.22 (95% CrI: $[-1.13, 0.66]$), and its pairwise differences from landscape change and introductions were no longer supported. Hunting or harvesting instead showed greater average divergence than *in situ* natural variation ($\beta_{\text{Hunting-Natural}}^{(l)} = 0.646$, $[0.352, 0.935]$). The differences between *in situ* natural variation and both landscape change and introductions were retained.

Predictability in the magnitude of divergence differed more broadly among disturbance contexts (Fig. 4B).

In the preregistered analysis, hunting or harvesting showed greater residual variation, and therefore lower predictability, than climate change ($\beta_{\text{Hunting-Climate}}^{(s)} = 0.365$, 95% CrI: $[0.189, 0.554]$), introductions ($\beta_{\text{Hunting-Introduction}}^{(s)} = 0.350$, $[0.278, 0.421]$), natural variation ($\beta_{\text{Hunting-Natural}}^{(s)} = 0.227$, $[0.142, 0.311]$), landscape change ($\beta_{\text{Hunting-Landscape}}^{(s)} = 0.796$, $[0.563, 1.054]$), and pollution ($\beta_{\text{Hunting-Pollution}}^{(s)} = 0.549$, $[0.307, 0.821]$). After excluding contrasts from Oke et al. (2020), the differences from climate change and introductions were no longer supported, whereas hunting or harvesting remained less predictable than landscape change, pollution, and *in situ* natural variation.

Responses to introduced species showed lower predictability than climate change ($\beta_{\text{Response-Climate}}^{(s)} = 0.256$, $[0.028, 0.486]$), introductions ($\beta_{\text{Response-Introduction}}^{(s)} = 0.241$, $[0.090, 0.393]$), landscape change ($\beta_{\text{Response-Landscape}}^{(s)} =$

257 0.687, [0.419, 0.976]), and pollution ($\beta_{\text{Response-Pollution}}^{(s)} = 0.440, [0.166, 0.746]$), and all four differences were
 258 retained after excluding Oke et al. (2020). Climate change ($\beta_{\text{Climate-Landscape}}^{(s)} = 0.431, [0.142, 0.737]$) and in-
 259 troductions ($\beta_{\text{Introduction-Landscape}}^{(s)} = 0.447, [0.222, 0.702]$) were also less predictable than landscape change,
 260 and these differences were likewise retained after that exclusion. In the preregistered analysis, *in situ* natu-
 261 ral variation showed lower predictability than introductions ($\beta_{\text{Natural-Introduction}}^{(s)} = 0.123, [0.053, 0.191]$),
 262 landscape change ($\beta_{\text{Natural-Landscape}}^{(s)} = 0.570, [0.340, 0.826]$), and pollution ($\beta_{\text{Natural-Pollution}}^{(s)} = 0.322,$
 263 $[0.080, 0.597]$). These differences were, however, sensitive to the 473 natural-variation contrasts contributed
 264 by Oke et al. (2020). Excluding these contrasts reduced the residual standard deviation for natural variation
 265 from 0.58 to 0.42; natural variation then remained less predictable than landscape change ($\beta^{(s)} = 0.252,$
 266 $[0.009, 0.525]$), no longer differed from pollution ($\beta^{(s)} = -0.001, [-0.262, 0.283]$), and became more pre-
 267 dictable than introductions ($\beta_{\text{Natural-Introduction}}^{(s)} = -0.198, [-0.305, -0.094]$) and responses to introduced
 268 species ($\beta_{\text{Natural-Response}}^{(s)} = -0.440, [-0.617, -0.262]$).

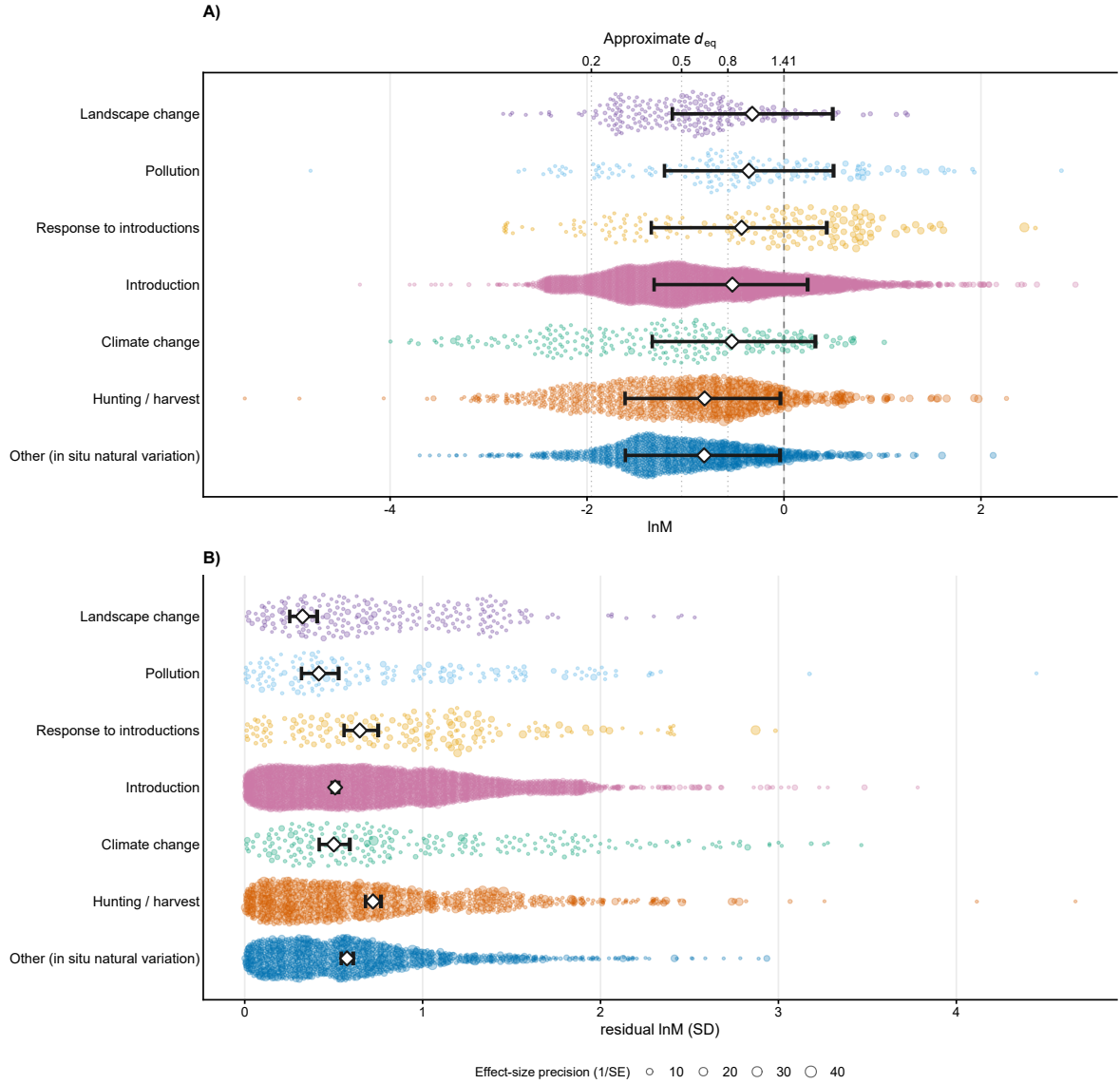


Figure 4: **Magnitude and predictability of within-species phenotypic divergence across disturbance contexts.** (A) Average phenotypic divergence between paired samples for each context. The reference line at $\ln M = 0$ corresponds to an absolute standardized mean difference of approximately $d_{eq} = 1.41$; the upper axis translates $\ln M$ into these approximate values. (B) Predictability of this divergence across contrasts within each context, where a larger σ indicates lower predictability. White circles and black lines represent posterior estimates and 95% credible intervals, respectively. Bubble size is proportional to precision (1/SE).

4.4 Other moderator analyses

In the preregistered analysis, comparison design was associated with both the magnitude and predictability of phenotypic divergence (Fig. S4). For both designs, average $\ln M$ was negative, but the credible intervals overlapped zero, only marginally so for allochronic contrasts (allochronic: $\beta_{\text{Allochronic}}^{(l)} = -0.72$, 95% CrI: $[-1.47, 0.02]$; synchronic: $\beta_{\text{Synchronic}}^{(l)} = -0.47$, $[-1.22, 0.28]$). Synchronic contrasts showed greater average divergence than allochronic contrasts ($\beta_{\text{Synchronic-Allochronic}}^{(l)} = 0.250$, 95% CrI: $[0.077, 0.418]$). This difference was attenuated but remained supported after adjusting for elapsed time, both when expressed in years ($\beta^{(l)} = 0.21$, 95% CrI: $[0.08, 0.35]$) and generations ($\beta^{(l)} = 0.18$, $[0.04, 0.31]$). Synchronic contrasts also showed lower residual variation in the preregistered model ($\beta_{\text{Synchronic-Allochronic}}^{(s)} = -0.208$, 95% CrI: $[-0.260, -0.157]$). This difference remained after adjustment for elapsed years ($\beta^{(s)} = -0.23$, $[-0.30, -0.16]$), but not after adjustment for elapsed generations ($\beta^{(s)} = 0.06$, $[-0.02, 0.14]$).

Elapsed time was positively associated with average divergence and negatively associated with residual variation among contrasts (Fig. 5). Because both components are modelled on the log scale, a slope β corresponds to a multiplicative change of e^β per tenfold increase in elapsed time. A tenfold increase in elapsed years was thus associated with a 36% increase in the ratio of the between-group to the within-group standard deviation ($e^{0.307} = 1.36$; $\beta_{\log_{10}(\text{years})}^{(l)} = 0.307$, 95% CrI: $[0.235, 0.381]$) and a 15% reduction in the residual standard deviation ($e^{-0.167} = 0.85$; $\beta_{\log_{10}(\text{years})}^{(s)} = -0.167$, $[-0.215, -0.119]$). Elapsed generations produced a similar pattern: a tenfold increase was associated with a 33% increase in this ratio ($\beta_{\log_{10}(\text{generations})}^{(l)} = 0.286$, $[0.215, 0.359]$) and a 14% reduction in the residual standard deviation ($\beta_{\log_{10}(\text{generations})}^{(s)} = -0.149$, $[-0.182, -0.117]$).

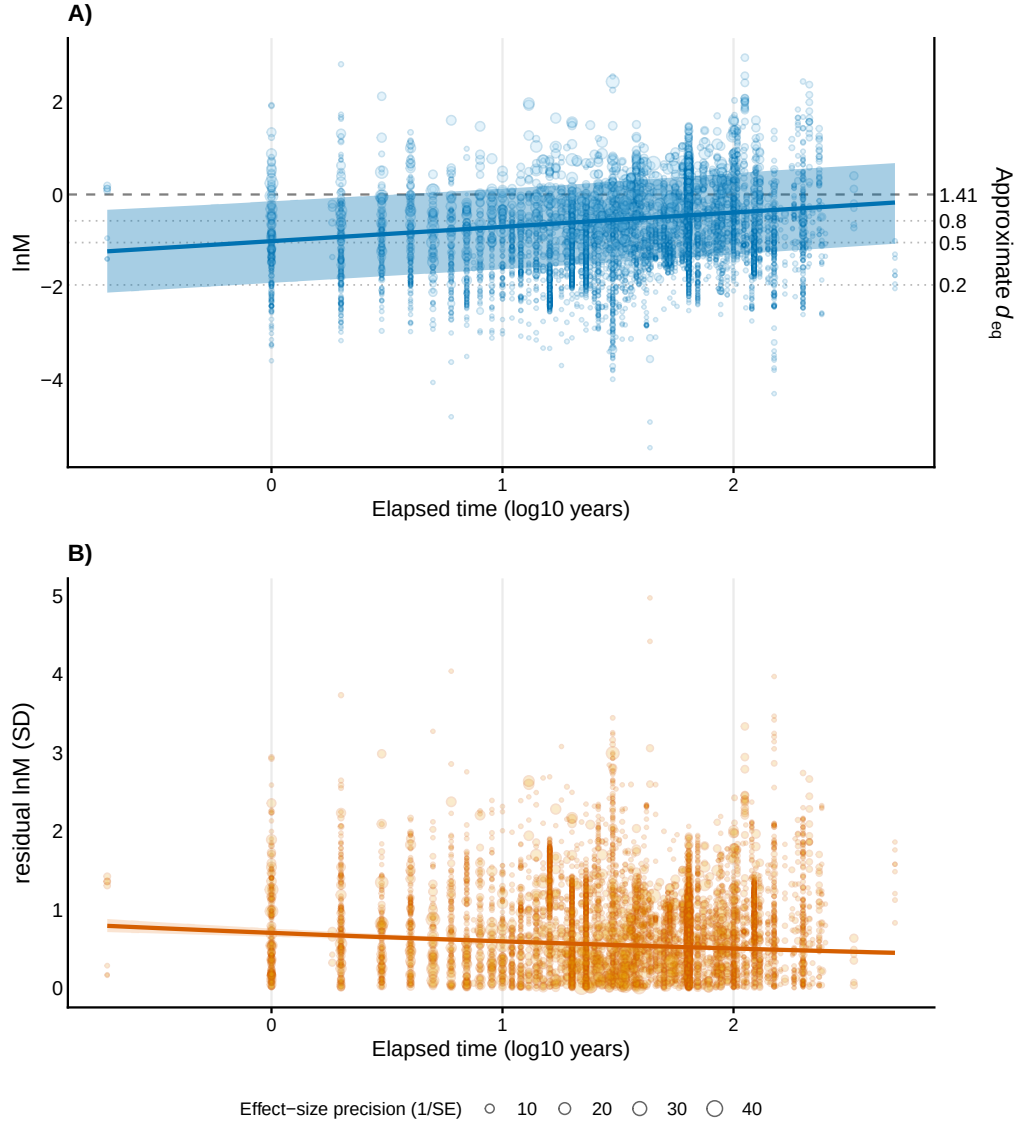


Figure 5: **Elapsed time is associated with the magnitude and predictability of within-species phenotypic divergence.** (A) Average divergence between paired samples over elapsed time (log₁₀ years). The reference line at $\ln M = 0$ corresponds to an absolute standardized mean difference of approximately $|d| = 1.41$, which is also indicated on the right axis. (B) Predictability of divergence among contrasts over elapsed time. A larger σ indicates lower predictability. The solid line and shaded ribbon show the posterior estimate and 95% credible interval, respectively. One horizontal unit represents a tenfold increase in time, and bubble size reflects precision (1/SE).

Trait types also differed in their average magnitude and predictability of phenotypic divergence (Fig. 6A). Behavioural traits were the only category for which average $\ln M$ was clearly below zero, indicating that average divergence between samples was smaller than the phenotypic variation retained within them, although this category comprised only 19 contrasts. For the remaining trait types, credible intervals overlapped $\ln M = 0$, so between-sample divergence could not be distinguished from within-sample variation; the two were of broadly comparable magnitude. Pairwise comparisons showed that behavioural traits had lower average divergence than all other trait types except phenology and physiology. Other life-history and response traits also showed greater divergence than several morphological and physiological categories. Differences were more extensive in the scale component (Fig. 6B). Behavioural and body-size traits were among the least predictable categories, whereas growth traits produced comparatively predictable estimates. Other life-history traits also showed greater residual variation than other morphological, growth, and physiological traits.

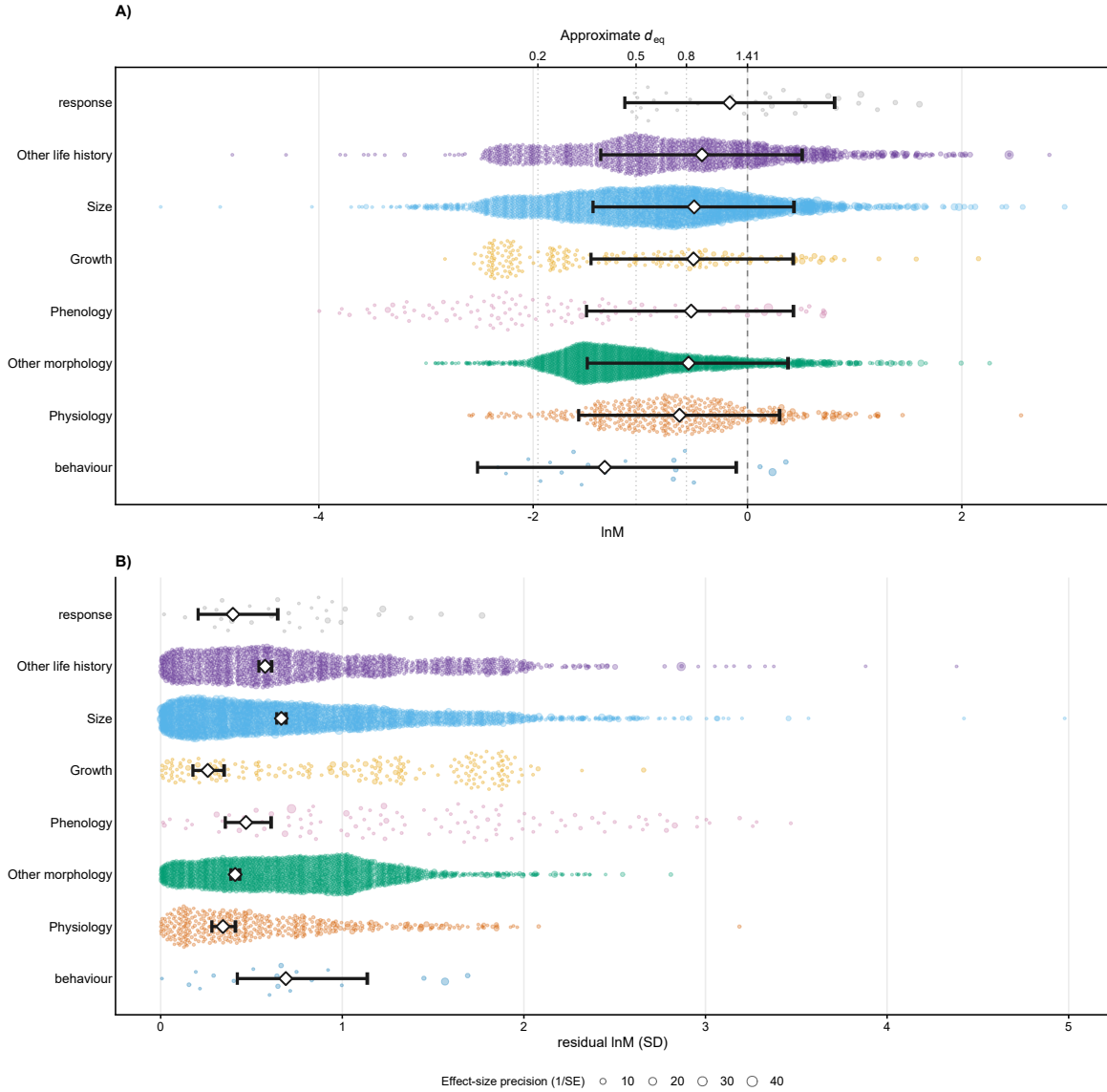


Figure 6: **Magnitude and predictability of within-species phenotypic divergence across trait types** (A) Average phenotypic divergence between paired population samples for each trait type. The upper axis displays approximate absolute standardized mean differences. (B) Predictability of divergence among contrasts within each trait type. White circles and black intervals represent posterior estimates and 95% credible intervals, respectively. Bubble size indicates precision (1/SE). Response traits denote measures of performance expressed as ratios.

Study type was similarly associated with both model components in the preregistered analysis. Field-based contrasts involving individuals that developed under natural conditions showed smaller average divergence than contrasts based on common-garden breeding, quantitative-genetic methods, or related approaches ($\beta_{\text{Common garden/QG-Field}}^{(l)} = 0.210$, 95% CrI: [0.063, 0.355]), and greater residual heterogeneity ($\beta_{\text{Field-Common garden/QG}}^{(s)} = 0.338$, [0.278, 0.398]; parameterized as field relative to common-garden/QG). However, this location difference was no longer supported after accounting for elapsed time, whether expressed in years ($\beta^{(l)} = -0.07$, 95% CrI: [-0.17, 0.03]; parameterized as field relative to common-garden/QG) or generations ($\beta^{(l)} = -0.06$, [-0.15, 0.04]). In contrast, the difference in predictability persisted after time adjustment. Field-based contrasts retained greater residual heterogeneity than common-garden or quantitative-genetic contrasts when controlling for elapsed years ($\beta^{(s)} = 0.40$, [0.33, 0.47]) and elapsed generations ($\beta^{(s)} = 0.34$, [0.27, 0.41]). Thus, the apparent difference in average divergence between study types was not robust to adjustment for elapsed time, whereas the difference in predictability was.

4.5 Excluded analyses and sensitivity analyses

Four preregistered moderator analyses were excluded because they did not meet the convergence criteria of $\hat{R} \leq 1.01$, effective sample sizes ≥ 400 , and no divergent transitions (Vehtari et al., 2021). These models tested whether the average magnitude and predictability of divergence differed among broad taxonomic groups, mathematical forms of trait measurement, transformed and untransformed trait values, and ongoing versus novel environmental change; their estimates were not interpreted.

Prespecified sensitivity analyses were run for the six reported moderator models under each of the three alternative specifications (18 models); sensitivity models for the excluded environmental-change model were not interpreted. Most conclusions were stable, particularly the location effects of comparison design, elapsed time, study type, and disturbance context. Of 34 non-intercept coefficients per specification, 3 (minimum sample size), 5 (no phylogeny), and 5 (study system) changed sign or changed whether their 95% CrI included zero. These changes were concentrated in trait-type location coefficients (growth, other morphology, and body size relative to behaviour; phenology under the minimum-sample-size specification) and a small number of scale effects: the scale coefficients for introductions and responses to introductions relative to climate change under the minimum-sample-size and no-phylogeny specifications, and the scale effects of comparison design and elapsed years under the study-system specification. Several sensitivity models did not meet our convergence criteria (two of six under the minimum-sample-size specification, three under the no-phylogeny specification, and five under the study-system specification), so departures from the primary results were

interpreted cautiously. Model-specific exclusions, coefficient-level comparisons, and convergence diagnostics are reported in the online supplement (https://santiago-Ortega.github.io/Magnitude_pace_modern_life/).

5 Discussion

5.1 Phenotypic divergence is heterogeneous rather than universal

We asked whether environmental change leaves a general and predictable signature of within-species phenotypic divergence. Across 7,186 contrasts, paired population samples showed substantial average divergence, but this divergence did not consistently exceed the phenotypic variation retained within populations (Fig. 3A). This does not imply that phenotypic differences among samples were small or biologically unimportant. Many individual contrasts showed strong phenotypic divergence, but its magnitude varied substantially across the evidence base. Rather, the overall estimate shows that even appreciable differences between sample averages often coexist with substantial variation among individuals within populations. The key result was therefore not an absence of phenotypic change, but the lack of a universal relationship between differences among samples and variation within populations. This result complements previous syntheses showing that contemporary phenotypic change is widespread but highly variable across taxa, traits, and environmental contexts (Hendry et al., 2008; Sanderson et al., 2022, 2023).

The clearest pattern was the extent of this variation (Fig. 3B). Total heterogeneity exceeded 96%, meaning that sampling error accounted for less than 4% of the variation in $\ln M$ and that nearly all of it reflected genuine differences in the magnitude of divergence. These differences were distributed across levels: comparable shares were attributable to source references and to phylogenetically structured species effects, and about one fifth remained among individual contrasts. Residual variation in $\ln M$ also differed among disturbance contexts, comparison designs, elapsed intervals, trait types, and study types. Similar average divergence can therefore reflect either comparable differences across systems or a mixture of weak and extreme responses (Sanderson et al., 2023). By modelling both the expected $\ln M$ and its residual variation, the location-scale models showed not only where phenotypic divergence tended to be greater, but also where it was more predictable across the evidence base (Nakagawa et al., 2025).

5.2 Disturbance context shapes the magnitude and predictability of phenotypic divergence

Disturbance contexts differed in the average phenotypic divergence between population samples. In the preregistered analysis, landscape-change contrasts showed greater divergence than hunting or harvesting and *in situ* natural variation, while introduced populations also showed greater divergence than these two contexts (Fig. 4A). However, the hunting or harvesting category was dominated by Oke et al. (2020), and the corresponding pairwise differences with landscape change and introduced populations were no longer supported after excluding that reference. In contrast, the location differences involving *in situ* natural variation were retained. These results complement previous syntheses based on rates of phenotypic change, typically expressed in darwins or haldanes (Gingerich, 1993; Haldane, 1949; Hendry and Kinnison, 1999; Hendry et al., 2008; Sanderson et al., 2022). Darwins measure proportional change in the average trait value per unit time, whereas haldanes measure change in the average trait value in units of the within-population standard deviation per generation (Hendry and Kinnison, 1999). A population average can change rapidly without producing strong phenotypic divergence when substantial variation among individuals is retained, although directional selection and other processes producing average shifts may also alter within-population variation. For example, two populations could differ markedly in average body size while still overlapping extensively if both retain broad phenotypic distributions. Rates of change and $\ln M$ therefore describe related but distinct features of population responses. Rates scale the change in trait averages by elapsed time and, for haldanes, also by within-population variation. By contrast, $\ln M$ is not scaled by time: it quantifies how distinct two samples are relative to the variation retained within them, irrespective of the direction of change.

Differences among disturbance contexts were even clearer in the predictability of divergence (Fig. 4B). Landscape change produced comparatively predictable divergence across both the preregistered and leave-one-reference-out analyses. Hunting or harvesting appeared less predictable in the preregistered analysis, but its differences from climate change and introduced populations were no longer supported after excluding Oke et al. (2020), indicating that this pattern should be interpreted cautiously. *In situ* natural variation also appeared comparatively unpredictable in the preregistered analysis, but this depended on the salmon contrasts contributed by the same reference: once they were excluded, natural variation was more predictable than introduced populations and responses to introduced species, and was less predictable only than landscape change. Responses of resident populations to introduced species consistently showed relatively high variation among contrasts. Introduced populations themselves were intermediate, emphasizing that phenotypic diver-

gence following an introduction differs from the response of resident populations exposed to an introduced species. Thus, disturbance categories with comparable average divergence can still differ substantially in how well that average generalizes across populations.

Several biological mechanisms could generate these differences. Harvesting can impose strong selection, but the traits targeted, the direction and intensity of selection, associated demographic effects, and management histories differ among systems (Campbell-Staton et al., 2021; Conover and Munch, 2002; Festa-Bianchet, 2017). Such variation could contribute to heterogeneous phenotypic responses, although the present evidence for particularly low predictability under hunting or harvesting is sensitive to the strong representation of a single reference. Responses to introduced species may likewise depend on whether the introduced organism acts as a predator, competitor, parasite, host, or prey, and on its ecological overlap with the focal population (Strauss et al., 2006). Landscape change, in contrast, can create persistent differences in habitat structure, connectivity, cover, and resource distribution (Fischer and Lindenmayer, 2007; Smokorowski and Pratt, 2007), potentially generating more repeatable degrees of phenotypic divergence across populations. Moreover, each contrast was assigned to a single disturbance context, whereas many populations are exposed to several disturbances simultaneously, such as harvesting combined with climate change or habitat alteration. Such co-occurring pressures could themselves contribute to the low predictability of some contexts. These mechanisms remain hypotheses, however, because our models describe differences among disturbance contexts rather than the processes generating them.

5.3 Temporal, biological, and methodological correlates of phenotypic divergence

Comparison design and elapsed time captured related but different aspects of phenotypic divergence. Synchronic contrasts showed greater average divergence than allochronic contrasts, and this difference persisted after adjustment for elapsed time expressed in either years or generations, although its magnitude was reduced. This attenuation suggests that part of the design difference reflects the longer intervals typically represented by synchronic comparisons, but that elapsed time alone does not explain it. Synchronic contrasts also showed greater predictability in the preregistered model and after adjustment for elapsed years, but this difference was not supported when elapsed time was expressed in generations. The difference in predictability between designs therefore appears to reflect differences in the number of generations represented by synchronic and allochronic contrasts, rather than a property of the designs themselves. This interpre-

tation is plausible because neither design is expected a priori to be more predictable: synchronic contrasts integrate the distinct histories of two populations since they split, including uncertainty in both trajectories, whereas allochronic contrasts may be sensitive to short-term fluctuations within a single population. Moreover, comparison design was strongly associated with disturbance context (Cramér's $V = 0.90$; Fig. S1), and therefore should not be interpreted as an independent test of spatial versus temporal divergence.

Longer elapsed intervals were associated with greater and more predictable divergence (Fig. 5). Similar results were obtained when time was expressed in years or generations. This pattern is consistent with persistent environmental differences becoming clearer as plastic responses, demographic turnover, and selection act over longer periods (Crozier et al., 2011; Losos et al., 1997; Ozgul et al., 2010; Reznick et al., 1997). This pattern also contrasts informatively with the well-established finding that rates of phenotypic change decline as the interval over which they are measured increases (Gingerich, 1993; Hendry and Kinnison, 1999): although change per unit time tends to be lower over longer intervals, the divergence accumulated between samples tended to increase. However, years and generations are closely related measures, so their agreement is supportive rather than an independent replication. Moreover, these models do not estimate a causal rate of divergence. Elapsed time describes the interval between endpoints or since the populations split, not the path of change between them, and it was unevenly distributed across disturbance contexts and study designs.

Trait type and study type further showed how biological variation and environmental control influence $\ln M$. Behavioural traits had low average divergence and some of the least predictable estimates (Fig. 6), although they were represented by only 19 contrasts. Although behaviour can be repeatable, it can also change rapidly with environmental conditions, social context, and individual state, and individuals may differ in their responses to these conditions (Bell et al., 2009; Dingemanse and Wolf, 2013; Dingemanse et al., 2010). Behavioural contrasts may therefore combine persistent population differences with short-term variation at the time of measurement. Growth traits, by contrast, produced the most predictable estimates. Growth integrates physiological and environmental conditions over a developmental period and is limited by trade-offs among energy acquisition, maintenance, development, and reproduction (Dmitriew, 2011). These shared constraints may restrict the range of viable responses, although this explanation was not tested directly.

A related pattern appeared across study types. In the preregistered analysis, field-based contrasts showed smaller and less predictable divergence than common-garden and quantitative-genetic contrasts. However, the difference in average divergence was no longer supported after adjustment for elapsed time, whether expressed in years or generations, showing that the original location difference was sensitive to the temporal intervals represented by the two study types. In contrast, field-based contrasts remained less predictable

after adjustment for either measure of elapsed time. Individuals developing in natural conditions experience variation in climate, food, density, age, condition, and developmental history, which can maintain substantial variation within populations (Gienapp et al., 2008; Merilä and Hendry, 2014). Common-garden and quantitative-genetic approaches standardize or model some of this variation, potentially producing more consistent estimates of phenotypic divergence across systems. Because $\ln M$ measures between-group divergence relative to within-group variation, reducing environmentally induced variation within groups could contribute to greater predictability. The study-type results therefore support a difference in predictability rather than evidence that persistent or heritable divergence is generally greater than phenotypic divergence observed in the field.

5.4 Limitations and future opportunities

Several limitations provide context for these findings. First, disturbance context, comparison design, trait type, and study type were closely associated with one another and with elapsed time within the dataset. Additional analyses showed that some associations were robust to adjustment for elapsed time, whereas others were not. In particular, the average difference between comparison designs persisted after adjustment, while the average difference between study types did not. However, the effects of these moderators could not be fully separated, and equivalent time-adjusted models for disturbance context and trait type did not meet our convergence criteria. Moderator associations should therefore be interpreted as patterns in the available evidence rather than independent causal effects. Better-balanced data across moderator combinations would allow these influences to be separated more effectively. Second, the evidence base was unevenly distributed across contexts, designs, traits, and taxa, with introductions, synchronic comparisons, morphological traits, and a small number of taxonomic groups contributing disproportionately to the dataset (Figs. 1 and 2). More balanced representation of sparsely sampled disturbance contexts, trait categories, and taxonomic groups would help test whether the patterns identified here extend beyond the best-sampled portions of the dataset. Third, most records compared only two endpoints. Long-term studies that repeatedly measure both trait averages and variation among individuals are needed to distinguish gradual accumulation from short-term plasticity, abrupt shifts, and reversals.

Small-study effects provide an additional reason for caution. Smaller studies tended to report larger $\ln M$ values but lower residual heterogeneity, suggesting that they may portray phenotypic divergence as both stronger and more predictable than larger studies (Nakagawa et al., 2022). These patterns do not demonstrate publication bias, because small-study effects can also arise from differences in study design, taxa, traits, or

biological systems, but they suggest that the absolute magnitude and predictability of divergence reported here may be somewhat overstated.

Finally, several preregistered moderator models did not meet our convergence criteria, and the phylogenetic reconstruction required approximations for unresolved relationships and branch lengths. Conclusions were generally stable across the 18 prespecified sensitivity models, particularly for the location effects of comparison design, elapsed time, and disturbance context. Changes were concentrated in trait-type comparisons and a small number of scale effects, including the lower predictability of responses to introduced species relative to climate change and the scale effects of comparison design and elapsed years when study system was modelled. The additional non-preregistered time-adjustment analyses further showed that some apparent moderator differences depended on elapsed-time structure, particularly the average difference between study types. Similarly, the leave-one-reference-out analysis showed that the location and scale results for hunting or harvesting and *in situ* natural variation depended substantially on a single large salmon dataset. Several sensitivity models, particularly under the study-system specification, did not meet our convergence criteria, so their departures from the primary results should be interpreted cautiously. Our findings support broad comparative patterns but should not be treated as precise predictions for particular species or systems.

5.5 Conclusions

Phenotypic responses to environmental change can be substantial, but their magnitude relative to within-population variation is neither uniform nor uniformly predictable across systems. Similar average differences can arise from broadly repeatable responses or from mixtures of weak and extreme divergence, making an average response a limited guide to what should be expected in any particular population. By combining $\ln M$ with phylogenetic location-scale meta-regression, we showed that disturbance context, elapsed time, comparison design, trait type, and study type were associated with variation in the magnitude, predictability, or both of phenotypic divergence. Studies of global change should therefore retain and report within-population variation, while syntheses should treat heterogeneity as information about how reliably responses generalize across systems rather than simply as statistical noise. Understanding and anticipating phenotypic responses requires knowing not only how much populations are expected to differ, but also how confidently that expectation can be extended to another population.

References

- Marina Alberti, Cristian Correa, John M. Marzluff, Andrew P. Hendry, Eric P. Palkovacs, Kiyoko M. Gotanda, Victoria M. Hunt, Travis M. Apgar, and Yuyu Zhou. Global urban signatures of phenotypic change in animal and plant populations. *Proceedings of the National Academy of Sciences of the United States of America*, 114:8951–8956, 8 2017. ISSN 10916490. doi: 10.1073/PNAS.1606034114.
- Alison M. Bell, Shala J. Hankison, and Kate L. Laskowski. The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, 77:771–783, 4 2009. ISSN 0003-3472. doi: 10.1016/J.ANBEHAV.2008.12.022.
- Paul Christian Bürkner. brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, 80:1–28, 8 2017. ISSN 1548-7660. doi: 10.18637/JSS.V080.I01. URL <https://www.jstatsoft.org/index.php/jss/article/view/v080i01>.
- Shane C. Campbell-Staton, Brian J. Arnold, Dominique Gonçalves, Petter Granli, Joyce Poole, Ryan A. Long, and Robert M. Pringle. Ivory poaching and the rapid evolution of tusklessness in african elephants. *Science*, 374:483–487, 10 2021. ISSN 10959203. doi: 10.1126/SCIENCE.ABE7389.
- Jeanette Clark, Rich Brenner, and Bert Lewis. Compiled age, sex, and length data for Alaskan salmon, 1922–2017. KNB Data Repository, 2018.
- Erik Clarke, Scott Sherrill-Mix, and Charlotte Dawson. *ggbeeswarm: Categorical Scatter (Violin Point) Plots*, 2025. URL <https://github.com/eclarke/ggbeeswarm>. R package version 0.7.3.
- Dale T. Clement, Isabelle P. Neylan, Nicholas J. Roberts, Sebastian J. Schreiber, Pete C. Trimmer, and Andrew Sih. Evolutionary history mediates population response to rapid environmental change through within-generational and transgenerational plasticity. *The American Naturalist*, 201(5):E90–E109, 2023. doi: 10.1086/723624.
- Jacob Cohen. *Statistical power analysis for the behavioral sciences*, Rev. Lawrence Erlbaum Associates, Inc, 1988.
- David O. Conover and Stephan B. Munch. Sustaining fisheries yields over evolutionary time scales. *Science*, 297:94–96, 7 2002. ISSN 00368075. doi: 10.1126/SCIENCE.1074085.
- Lisa G. Crozier, Mark D. Scheuerell, and Richard W. Zabel. Using time series analysis to characterize evolutionary and plastic responses to environmental change: A case study of a shift toward earlier migration date in sockeye salmon. *American Naturalist*, 178:755–773, 12 2011. ISSN 00030147. doi: 10.1086/662669.

- Simone Des Roches, David M. Post, Nash E. Turley, Joseph K. Bailey, Andrew P. Hendry, Michael T. Kinnison, Jennifer A. Schweitzer, and Eric P. Palkovacs. The ecological importance of intraspecific variation. *Nature Ecology & Evolution*, 2:57–64, 12 2017. ISSN 2397-334X. doi: 10.1038/s41559-017-0402-5. URL <https://www.nature.com/articles/s41559-017-0402-5>.
- Sarah E. Diamond and Ryan A. Martin. The interplay between plasticity and evolution in response to human-induced environmental change. *F1000Research*, 5:2835, 2016. doi: 10.12688/f1000research.9731.1.
- Niels J. Dingemanse and Max Wolf. Between-individual differences in behavioural plasticity within populations: causes and consequences. *Animal Behaviour*, 85:1031–1039, 5 2013. ISSN 0003-3472. doi: 10.1016/J.ANBEHAV.2012.12.032.
- Niels J. Dingemanse, Anahita J.N. Kazem, Denis Réale, and Jonathan Wright. Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology & Evolution*, 25:81–89, 2 2010. ISSN 0169-5347. doi: 10.1016/J.TREE.2009.07.013.
- Caitlin M. Dmitriew. The evolution of growth trajectories: What limits growth rate? *Biological Reviews*, 86:97–116, 2 2011. ISSN 14647931. doi: 10.1111/J.1469-185X.2010.00136.X. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1469-185X.2010.00136.x>.
- Matthias Egger, George Davey Smith, Martin Schneider, and Christoph Minder. Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315:629–634, 9 1997. ISSN 0959-8138. doi: 10.1136/BMJ.315.7109.629.
- Marco Festa-Bianchet. When does selective hunting select, how can we tell, and what should we do about it? *Mammal Review*, 47:76–81, 1 2017. ISSN 1365-2907. doi: 10.1111/MAM.12078. URL <https://onlinelibrary.wiley.com/doi/10.1111/mam.12078>.
- Joern Fischer and David B. Lindenmayer. Landscape modification and habitat fragmentation: A synthesis. *Global Ecology and Biogeography*, 16:265–280, 5 2007. ISSN 1466822X. doi: 10.1111/J.1466-8238.2007.00287.X. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1466-8238.2007.00287.x>.
- Jonah Gabry, Rok Češnovar, Andrew Johnson, and Steve Bröder. *cmdstanr: R Interface to 'CmdStan'*, 2025. URL <https://mc-stan.org/cmdstanr/>. R package version 0.9.0, <https://discourse.mc-stan.org>.
- P. Gienapp, C. Teplitsky, J. S. Alho, J. A. Mills, and J. Merilä. Climate change and evolution: Disentangling environmental and genetic responses. *Molecular Ecology*, 17:167–178, 1 2008. ISSN 09621083. doi: 10.1111/J.1365-294X.2007.03413.X. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-294X.2007.03413.x>.

- Philip D. Gingerich. Quantification and Comparison of Evolutionary Rates. *American Journal of Science*, 293:453–478, jan 1 1993. doi: 10.2475/ajs.293.A.453.
- Lucas D. Gorné, Andrew P. Hendry, Fanie Pelletier, Sarah Sanderson, Cristian Correa, Carlos Arias, Marc Olivier Beausoleil, Maryse Boisjoly, Erika Crispo, Daniel Berner, Luis F. De León, Joseph D. DiBattista, Grant E. Haines, Benjamin C. Haller, Michael T. Kinnison, Shahin Muttalib, Ann E. McKellar, Rose E. O’Dea, Winer Daniel Reyes-Corral, Yanny Ritchot, Krista B. Oke, Zachary T. Wood, Thomas Farrugia, and Kiyoko M. Gotanda. Proceed v6.1: Phenotypic rates of change evolutionary and ecological database. *Ecology*, 106:e70009, 3 2025. ISSN 19399170. doi: 10.1002/ECY.70009.
- Kiyoko M Gotanda and Lucas D. Gorné. Phenotypic Rates of Change Evolutionary and Ecological Database, 2 2026. URL <https://onlinelibrary.wiley.com/doi/10.1111/mec.16299>.
- A. Grafen. The Phylogenetic Regression. *Philosophical Transactions of the Royal Society of London. Series B*, 326(1233), 12 1989. URL <https://www.jstor.org/stable/2396904?seq=1>.
- J. B. S. Haldane. Suggestions as to quantitative measurement of rates of evolution. *Evolution*, 3:51, 3 1949. ISSN 00143820. doi: 10.2307/2405451.
- Andrew P. Hendry and Michael T. Kinnison. Perspective: The pace of modern life: Measuring rates of contemporary microevolution. *Evolution*, 53:1637–1653, 10 1999. ISSN 0014-3820. doi: 10.1111/J.1558-5646.1999.TB04550.X.
- Andrew P. Hendry, Thomas J. Farrugia, and Michael T. Kinnison. Human influences on rates of phenotypic change in wild animal populations. *Molecular Ecology*, 17:20–29, 1 2008. ISSN 09621083. doi: 10.1111/J.1365-294X.2007.03428.X. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1365-294X.2007.03428.x>.
- Andrew P. Hendry, Kiyoko M. Gotanda, and Erik I. Svensson. Human influences on evolution, and the ecological and societal consequences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372, 1 2017. ISSN 14712970. doi: 10.1098/RSTB.2016.0028/30286.
- Cody E. Hinchliff, Stephen A. Smith, James F. Allman, J. Gordon Burleigh, Ruchi Chaudhary, Lyndon M. Coghill, Keith A. Crandall, Jiabin Deng, Bryan T. Drew, Romina Gazis, Karl Gude, David S. Hibbett, Laura A. Katz, H. Dail Laughinghouse, Emily Jane McTavish, Peter E. Midford, Christopher L. Owen, Richard H. Ree, Jonathan A. Rees, Douglas E. Soltisc, Tiffani Williams, and Karen A. Cranston. 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(41):12764–12769, 10 2015. ISSN 10916490.

doi: 10.1073/PNAS.1423041112.

Daniël Lakens. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and anovas. *Frontiers in Psychology*, 4:62627, 11 2013. ISSN 16641078. doi: 10.3389/FPSYG.2013.00863. URL www.frontiersin.org.

Jonathan B. Losos, Kenneth I. Warheit, and Thomas W. Schoener. Adaptive differentiation following experimental island colonization in anolis lizards. *Nature*, 387:70–73, 5 1997. ISSN 1476-4687. doi: 10.1038/387070a0. URL <https://www.nature.com/articles/387070a0>.

Michael Lynch. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *The American Naturalist*, 136(6):727–741, 1990. ISSN 00030147, 15375323. URL <http://www.jstor.org/stable/2462164>.

Juha Merilä and Andrew P. Hendry. Climate change, adaptation, and phenotypic plasticity: The problem and the evidence. *Evolutionary Applications*, 7:1–14, 1 2014. ISSN 17524563. doi: 10.1111/EVA.12137. URL <https://onlinelibrary.wiley.com/doi/10.1111/eva.12137>.

François Michonneau, Joseph W. Brown, and David J. Winter. rotl: an R package to interact with the Open Tree of Life data. *Methods in Ecology and Evolution*, 7(12):1476–1481, 12 2016. ISSN 2041210X. doi: 10.1111/2041-210X.12593. URL <https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.12593>.

Shinichi Nakagawa, Malgorzata Lagisz, Michael D. Jennions, Julia Koricheva, Daniel W.A. Noble, Timothy H. Parker, Alfredo Sánchez-Tójar, Yefeng Yang, and Rose E. O’Dea. Methods for testing publication bias in ecological and evolutionary meta-analyses. *Methods in Ecology and Evolution*, 13:4–21, 1 2022. ISSN 2041210X. doi: 10.1111/2041-210X.13724.

Shinichi Nakagawa, Ayumi Mizuno, Kyle Morrison, Lorenzo Ricolfi, Coralie Williams, Szymon M. Drobnik, Malgorzata Lagisz, and Yefeng Yang. Location-Scale Meta-Analysis and Meta-Regression as a Tool to Capture Large-Scale Changes in Biological and Methodological Heterogeneity: A Spotlight on Heteroscedasticity. *Global Change Biology*, 31(5):e70204, 5 2025. ISSN 1365-2486. doi: 10.1111/GCB.70204. URL <https://onlinelibrary.wiley.com/doi/10.1111/gcb.70204>.

Shinichi Nakagawa, Ayumi Mizuno, Coralie Williams, Malgorzata Lagisz, Rose E. O’Dea, Daniel W. A. Noble, Alistair M. Senior, Erick Lundgren, and Santiago Ortega. A new effect size for meta-analysis of magnitude: lnm. *Ecology Letters*, 29:e70428, 9 2026a. ISSN 1461-0248. doi: 10.1111/ELE.70428.

Shinichi Nakagawa, Ayumi Mizuno, Coralie Williams, Santiago Ortega, Szymon M. Drobnik, Malgo-

624 rzata Lagisz, Yefeng Yang, Alistair M. Senior, Daniel W. A. Noble, and Erick Lundgren. Mas-
625 tering a simulation-based parametric method to obtain bias-corrected point estimates and sampling
626 variance for effect sizes. *Research Synthesis Methods*, 0:1–28, 9 2026b. ISSN 1759-2879. doi:
627 10.1017/RSM.2026.10108.

628 Shinichi Nakagawa, Santiago Ortega, Ayumi Mizuno, Eduardo S.A. Santos, Malgorzata Lagisz, Bhavya Jain,
629 Jimuel Jr Celeste, and Sergio Poo Hernandez. *prepR4pcm: Prepare Data and Trees for Phylogenetic*
630 *Comparative Methods*, 2026c. URL <https://github.com/itchyshin/prepR4pcm>. R package version
631 1.0.0.

632 Rose E. O’Dea, Malgorzata Lagisz, Michael D. Jennions, Julia Koricheva, Daniel W.A. Noble, Timothy H.
633 Parker, Jessica Gurevitch, Matthew J. Page, Gavin Stewart, David Moher, and Shinichi Nakagawa.
634 Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology:
635 a PRISMA extension. *Biological Reviews*, 96(5):1695–1722, 10 2021. ISSN 1469185X. doi: 10.1111/brv.
636 12721. URL <https://onlinelibrary.wiley.com/doi/10.1111/brv.12721>.

637 K. B. Oke, C. J. Cunningham, P. A.H. Westley, M. L. Baskett, S. M. Carlson, J. Clark, A. P. Hendry, V. A.
638 Karatayev, N. W. Kendall, J. Kibele, H. K. Kindsvater, K. M. Kobayashi, B. Lewis, S. Munch, J. D.
639 Reynolds, G. K. Vick, and E. P. Palkovacs. Recent declines in salmon body size impact ecosystems and
640 fisheries. *Nature Communications*, 11:4155–, 8 2020. ISSN 2041-1723. doi: 10.1038/s41467-020-17726-z.
641 URL <https://www.nature.com/articles/s41467-020-17726-z>.

642 Santiago Ortega, Shinichi Nakagawa, Eduardo S A Santos, Kiyoko M Gotanda, Sarah Sanderson, Lucas D
643 Gorné, and Andrew Hendry. Human disturbance and the magnitude of phenotypic divergence, May
644 2026. URL osf.io/9fq4n.

645 Arpat Ozgul, Dylan Z. Childs, Madan K. Oli, Kenneth B. Armitage, Daniel T. Blumstein, Lucretia E.
646 Olson, Shripad Tuljapurkar, and Tim Coulson. Coupled dynamics of body mass and population growth
647 in response to environmental change. *Nature*, 466:482–485, 7 2010. ISSN 1476-4687. doi: 10.1038/
648 nature09210. URL <https://www.nature.com/articles/nature09210>.

649 Eric P. Palkovacs, Michael T. Kinnison, Cristian Correa, Christopher M. Dalton, and Andrew P. Hendry.
650 Fates beyond traits: Ecological consequences of human-induced trait change. *Evolutionary Appli-*
651 *cations*, 5:183–191, 2 2012. ISSN 17524563. doi: 10.1111/J.1752-4571.2011.00212.X. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1752-4571.2011.00212.x>.

653 Emmanuel Paradis, Julien Claude, and Korbinian Strimmer. Ape: Analyses of phylogenetics and evolution

in r language. *Bioinformatics*, 20:289–290, 1 2004. ISSN 1367-4803. doi: 10.1093/BIOINFORMATICS/BTG412.

Thomas Pendersen. patchwork: The composer of plots, 2025.

R Core Team. R: A language and environment for statistical computing, 6 2026.

David N. Reznick, Frank H. Shaw, F. Helen Rodd, and Ruth G. Shaw. Evaluation of the rate of evolution in natural populations of guppies (*poecilia reticulata*). *Science*, 275:1934–1937, 3 1997. ISSN 00368075. doi: 10.1126/SCIENCE.275.5308.1934.

Sarah Sanderson, Marc Olivier Beausoleil, Rose E. O’Dea, Zachary T. Wood, Cristian Correa, Victor Frankel, Lucas D. Gorné, Grant E. Haines, Michael T. Kinnison, Krista B. Oke, Fanie Pelletier, Felipe Pérez-Jvostov, Winer D. Reyes-Corral, Yanny Ritchot, Freedom Sorbara, Kiyoko M. Gotanda, and Andrew P. Hendry. The pace of modern life, revisited. *Molecular Ecology*, 31(4):1028–1043, 2 2022. ISSN 1365294X. doi: 10.1111/MEC.16299.

Sarah Sanderson, Daniel I. Bolnick, Michael T. Kinnison, Rose E. O’Dea, Lucas D. Gorné, Andrew P. Hendry, and Kiyoko M. Gotanda. Contemporary changes in phenotypic variation, and the potential consequences for eco-evolutionary dynamics. *Ecology Letters*, 26:S127–S139, 9 2023. ISSN 14610248. doi: 10.1111/ELE.14186. URL <https://onlinelibrary.wiley.com/doi/10.1111/ele.14186>.

K. E. Smokorowski and T. C. Pratt. Effect of a change in physical structure and cover on fish and fish habitat in freshwater ecosystems: A review and meta-analysis. *Environmental Reviews*, 15:15–41, 2007. doi: 10.1139/A06-007.

Stan Development Team. Stan reference manual, 2025.

Sharon Y. Strauss, Jennifer A. Lau, and Scott P. Carroll. Evolutionary responses of natives to introduced species: What do introductions tell us about natural communities? *Ecology Letters*, 9:357–374, 3 2006. ISSN 14610248. doi: 10.1111/J.1461-0248.2005.00874.X. URL <https://onlinelibrary.wiley.com/doi/10.1111/j.1461-0248.2005.00874.x>.

Monica G. Turner. Disturbance and landscape dynamics in a changing world. *Ecology*, 91(10):2833–2849, 2010. doi: 10.1890/10-0097.1.

Aki Vehtarh, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul Christian Burkner. Rank-normalization, folding, and localization: An improved $\hat{r}$ for assessing convergence of mcmc (with discussion). <https://doi.org/10.1214/20-BA1221>, 16:667–718, 6 2021. ISSN 1936-0975. doi: 10.1214/20-BA1221.

- 684 Hadley Wickham. `ggplot2`. *WIREs Computational Statistics*, 3:180–185, 3 2011. ISSN 1939-5108. doi:
685 10.1002/wics.147. URL <https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.147>.
- 686 Ke Yang, Enxuan Lin, Wangli Xu, Liping Zhu, and Tiejun Tong. An alternative measure for quantifying
687 the heterogeneity in meta-analysis. *Statistics in Medicine*, 44:e70089, 5 2025. ISSN 10970258. doi:
688 10.1002/SIM.70089. URL <https://onlinelibrary.wiley.com/doi/10.1002/sim.70089>.