Received: 6 December 2018 | Revised: 21 May 2019 | Accepted: 30 June 2019

DOI: 10.1111/faf.12394

ORIGINAL ARTICLE

FISH and FISHERIES WILEY

# Developmental temperature affects phenotypic means and variability: A meta-analysis of fish data

Rose E. O'Dea[1,2] | Malgorzata Lagisz[1] | Andrew P. Hendry[2] | Shinichi Nakagawa[1]

[1]Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, New South Wales, Australia

[2]Department of Biology & Redpath Museum, McGill University, Montréal, Québec, Canada

**Correspondence**
Rose E. O'Dea, Evolution & Ecology Research Centre, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia.
Email: rose.eleanor.o.dea@gmail.com

**Funding information**
2017 Endeavour Australian Postgraduate Scholarship; ARC Future Fellowship, Grant/Award Number: FT130100268; Discovery, Grant/Award Number: DP180100818

## Abstract

Fishes are sensitive to their thermal environment and face an uncertain future in a warming world. Theoretically, populations in novel environments might express greater levels of phenotypic variability to increase the chance of surviving—and eventually thriving—in the new conditions. Most research on the effect of the early thermal environment in fish species focuses on average phenotypic effects rather than phenotypic variability, but to understand how fishes will respond to rising temperatures we need to consider both the average response of the population, as well as the breadth of individual responses. Here we present the first meta-analysis on the effects of developmental temperature in fishes. Using data from 43 species and over 6,000 individual fish, we show that a change in developmental temperature induces a significant change in phenotypic means and variability, but differently depending on whether the temperature is increased or decreased. Decreases in temperature (cool environments) showed a significant decrease in phenotypic means and no change in phenotypic variability. Increases in temperature (warm environments) showed a non-significant increase in phenotypic means and a marginally significant increase in phenotypic variability. Larger increases in temperature saw greater increases in phenotypic variability, but no increase in the mean phenotypic response. Together, our results suggest that fishes exhibit both directed and stochastic developmental plasticity in response to warming temperatures, which could facilitate or accelerate adaptation to a changing environment.

KEYWORDS

bet-hedging, canalization, genetic compensation, non-adaptive plasticity, spreading reaction norms, systematic review

## 1 | INTRODUCTION

Fish populations are threatened by warming temperatures due to climate change. If these threats are realized the economic impact will be profound; fisheries represent a multibillion dollar industry and support a large fraction of the human population (Dulvy, Sadovy, & Reynolds, 2003; Sumaila, Cheung, Lam, Pauly, & Herrick, 2011). Fishes, like all species, are adapted to survive within a restricted range of temperatures. When temperatures shift beyond this range, populations could become maladapted and must either adapt or perish (Robertson, Rehage, & Sih, 2013).

Plasticity—the expression of different phenotypes when the same genotype is exposed to different environments—can help populations survive rising temperatures (Ghalambor, McKay, Carroll, & Reznick, 2007). Temperature change that predictably occurred within the ancestral history of a population might induce "adaptive

  

developmental plasticity" (sensu Nettle & Bateson, 2015). In this case, the developmental temperature is a cue that triggers a phenotypic change in the direction of the new optimum, but if the temperature change is severe or unprecedented, then this might merely impose developmental stress. In stressful conditions, plasticity is likely to be maladaptive, so selection should favour a reduction in plasticity (this phenomenon is called "genetic compensation"; Grether, 2005). Selection could favour a reduction in plasticity in the population *mean*, however, without decreasing individual plasticity; this scenario is predicted in Figure 1.

A rapid change in temperature could induce greater levels of phenotypic variation within a population, which could facilitate or accelerate adaptation to a new environment (O'Dea, Noble, Johnson, Hesselson, & Nakagawa, 2016). Ordinarily, when a population is well adapted to its environment, we expect high "adaptive precision" (sensu Hansen, Carter, & Pélabon, 2006) so that the genotype reliably produces a near-optimal phenotype. In a changing environment, however, lowering adaptive precision (i.e. increasing phenotypic variability) could increase the chance of an unusual and beneficial phenotype arising in the population (Hansen et al., 2006). Populations with greater phenotypic variance might recover more quickly after an environmental perturbation, as they are more likely to harbour individuals who, by chance, tolerate or even thrive in the changed conditions. The fitter individuals could allow the population to persist in a novel environment, and provide the material for selection to act upon (Ghalambor et al., 2007). This scenario is reminiscent of "bet-hedging"—if it is unclear which single phenotype will maximize fitness in the next generation, betting on a wide range of phenotypes might pay off (Franch-Gras, García-Roger, Serra, & Carmona, 2017; Starrfelt & Kokko, 2012). Potentially, variability itself could be heritable, which might allow these variants to keep up with rapidly changing environments via "heritable bet-hedging" (O'Dea et al., 2016; Pal & Miklos, 1999).

Novel environments could increase phenotypic variance by exposing previously hidden (cryptic) genetic variation, or by inducing new epigenetic changes. Under normal conditions, the genotype slowly accumulates genetic changes that are not expressed. When the temperature changes, some of this variation can be revealed and exposed to selection (McGuigan & Sgrò, 2009; Paaby & Rockman, 2014; Wood & Brodie, 2015). Any variants with a selective advantage could increase in frequency via natural selection and spread through the population. Alternatively, a change in temperature can induce changes in gene expression via epigenetic modifications. While these changes will only be heritable in the short term (if at all), they may still increase the likelihood that the phenotypes become genetically encoded, via genetic assimilation (Crispo, 2007; Pal & Miklos, 1999). Despite a theoretical basis behind the adaptive potential of increased phenotypic variance (Ghalambor et al., 2007), the effect of temperature on phenotypic variance is largely unexplored in empirical studies.

Fishes should reveal whether temperature changes do increase phenotypic variance because, as ectotherms, they are particularly sensitive to their external temperature (Neuheimer, Thresher, Lyle,

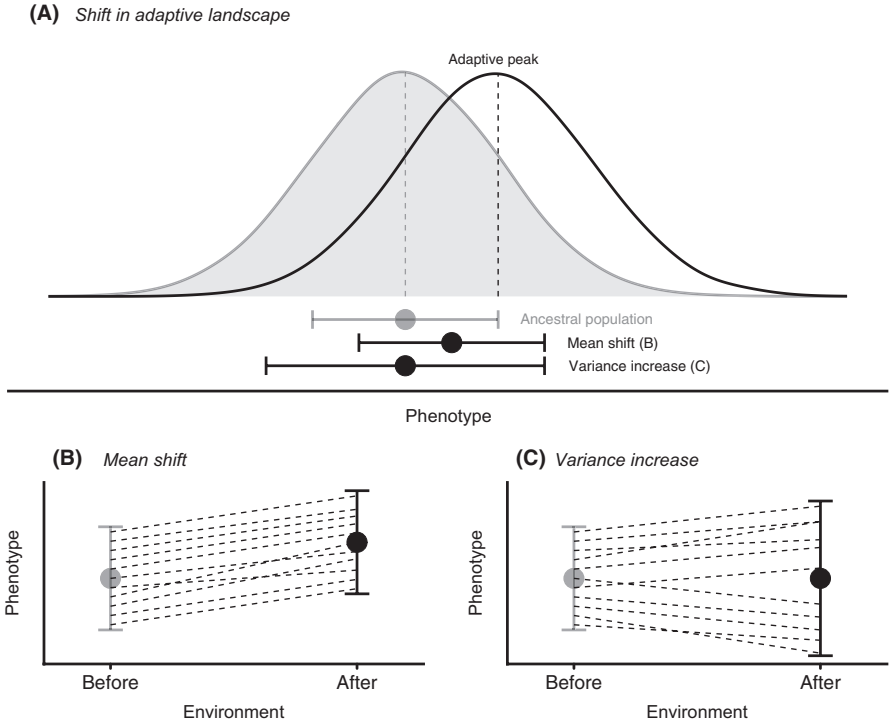


**FIGURE 1** A change in the environment changes the optimal phenotype. Subsequently, phenotypic plasticity changes the phenotypic mean (b) or phenotypic variance (c), which shifts some individuals closer to the new optimum (i.e. adaptive plasticity). Transparent points and lines denote previous conditions and phenotypes, while black denotes the present (after an environmental change). (a) Fitness density curves show the old and new adaptive landscape, with old and new population phenotypes shown underneath (points and error bars represent mean ± *SD*). Under normal conditions (before the change), the phenotype of a hypothetical population is centred on the adaptive peak (i.e. the point of maximum fitness), but after an environmental change the adaptive landscape shifts. The dashed vertical lines represent the optimal phenotypes in each environment. Both a plastic shift in the mean or an increase in variability can push more of the population towards the new adaptive optima. In (b,c), the changes in phenotypes are represented as either (b) consistent shifts in the intercept of reaction norms, or (c) stochastic changes in the intercept and slope of reaction norms (i.e. spreading of the reaction norms)

& Semmens, 2011). Previous fish studies have shown the phenotypic average of many phenotypic traits is affected by the developmental environment (i.e. there is developmental plasticity) (Jonsson & Jonsson, 2014). However, while phenotypic variance is at the heart of evolutionary theory, statistical analyses have historically focussed on testing for differences in phenotypic means. Few, if any, studies explicitly test whether the developmental temperature changes phenotypic variance, but the statistical tools now exist to approach this question using a meta-analysis.

Here we present the first meta-analysis on the phenotypic effects of developmental temperature in fish, and the first in any species to test for the effects of developmental temperature on phenotypic variability. We test 10 a priori predictions, which we registered before data exploration and analysis (O'Dea, Lagisz, Hendry, & Nakagawa, 2018a). We predict that: (a) fish reared in warmer temperatures will have greater phenotypic variability than fish experiencing control temperatures, controlling for any effect of temperature on the phenotypic mean (Figure 1c). (b) Changing the developmental temperature will impact the mean of traits, according to the studies reviewed in Jonsson and Jonsson (2014). Warm temperatures will increase growth rate and metabolic rate, but reduce size, muscle fibre number and heart volume (c.f. Hesse's rule; Müller et al., 2014) (Figure 1b). (c) Cool temperature treatments will cause differences in phenotypic means and variability that are similar in magnitude to warm temperature treatments, but with the same direction in variability and opposing directions in means. This prediction assumes that the developmental temperature and optimal phenotypic mean are linearly correlated, and fishes have evolved adaptive developmental plasticity. (d) Larger differences between control and treatment temperatures will result in larger differences in phenotypic means and variability. (e) Longer treatment durations will cause a larger difference in phenotypic means and variability. (f) An earlier start in treatment will cause larger differences in phenotypic means and variability. (g) A permanent treatment will have a larger effect on phenotypic means and variability than a transient treatment. (h) Fish groups that express greater phenotypic variability in the control environment will also show a greater shift in the average phenotype when exposed to the experimental temperature. (i) Experimental populations with greater amounts of genetic diversity will show more phenotypic variability, and respond more to temperature treatments. To test this prediction, we use a crude, but simple proxy for genetic diversity: the number of fish who contributed gametes to the experimental population. (j) Temperature treatments

that approach or exceed the optimal thermal limits of the species will elicit larger phenotypic effects than temperature treatments within the normal thermal range.

## 2 | METHODS

### 2.1 | Availability of data, code and materials

Data, analysis code and lists of screened studies are available to download from https://osf.io/e2tyw/ (O'Dea, Lagisz, Hendry, & Nakagawa, 2018b). Data include additional variables that we did not make specific predictions for, but could be used by other researchers in exploratory analyses.

### 2.2 | Finding data

#### 2.2.1 | Protocol and registration

We report details of this systematic meta-analysis following the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Moher, Liberati, Tetzlaff, & Altman, 2009). We registered our study protocol prior to data exploration and analysis. The registration includes details of our a priori hypotheses, search methods and planned analyses, and can be viewed on the Open Science Framework (https://osf.io/8ymh9; O'Dea, Lagisz, Hendry, & Nakagawa, 2018a).

#### 2.2.2 | Eligibility criteria

*Study design*

We included experimental studies that reared fish eggs or newly hatched fish embryos under at least two temperature conditions: (a) normal temperature (control treatment; defined below under *Data collection process*) and (b) warmer than normal temperature (warm treatment). If data were available for a cool temperature treatment in addition to the warm treatment, we extracted those data too. Those data allowed us to test whether any phenotypic differences were simply caused by a temperature change, or whether the direction of the temperature changes mattered. When data on multiple temperature treatments were presented at increasing temperature differences from the control, we took each control–treatment pairwise comparison. We excluded treatments that caused survival lower than 10%, so that we were testing for the effects of viable temperature changes, not extreme developmental stress. The cut-off value of 10% was chosen because studies that reported very low survival fell below this value.

The temperature treatments needed to commence on or before the day of hatching, and the treatments needed to be simultaneous (i.e. families or groups split between treatments; between-subject design). We included studies where the treatment was maintained for the duration of the experiment, and also studies where fish were brought back to a common, control temperature before being measured.

*Phenotypic measurements*

Studies needed to report means, sample sizes, and variance or a measure of dispersion (standard deviations, standard errors, interquartile ranges or coefficient of variations) for ratio-scale phenotypic traits (i.e. traits measured on a continuous scale with a lower-bound at zero). Where sample sizes or variance was missing, we attempted to contact authors for this information. All contacted authors ($n$ = 8) were asked whether they could provide additional data (published or unpublished) that could be used in our meta-analysis. Five authors replied to requests for data, and two provided data used in analysis. We excluded measurements taken before the day of hatching (e.g. egg mass). We also excluded proportion (binary) data (e.g. sex ratio) because these data types would not allow us to test for variability differences (the distribution of these data is described by the mean and sample size, and any variance that is reported for these measurements is from the group level rather than individual level, making it not comparable to other types of traits). The minimum sample size for inclusion was three fish per treatment group. Genetic and molecular data were outside the scope of this meta-analysis.

### 2.3 | Information sources

#### 2.3.1 | Search

We performed a systematic search using the Scopus and Web of Science online databases on 8th November 2017, removed duplicate results, and obtained 1,316 studies for screening. Both databases were accessed through McGill Library's subscription. The exact search strings were as follows:

Scopus: TITLE-ABS-KEY (“*fish*” OR “bass” OR “carp” OR “char” OR “cod” OR “salmon*” OR “sole” OR “tetra” OR “*trout”) AND TITLE-ABS-KEY (“high* temperature*” OR “elevated temperature*” OR “high* water temperature*” OR “elevated water temperature*” OR “rearing temperature*” OR “effects of temperature*” OR “temperature challenge*” OR “thermal stress*” OR “embr* temper*”).

AND TITLE-ABS-KEY ( “rear*” OR “incubat*” OR “developmental temperature*”) AND LANGUAGE (english) AND NOT SRCTITLE ( “Japanese Edition”).

AND NOT TITLE-ABS-KEY (“cryo*” OR “triploidy” OR “jellyfish”).

AND ( LIMIT-TO ( SUBJAREA, “AGRI”) OR LIMIT-TO ( SUBJAREA, “BIOC”) OR LIMIT-TO ( SUBJAREA, “ENVI”) OR LIMIT-TO ( SUBJAREA, “MEDI”) OR LIMIT-TO ( SUBJAREA, “VETE”) OR LIMIT-TO ( SUBJAREA, “IMMU”)).

Web of Science: (TS=(“*fish*” OR “bass” OR “carp” OR “char” OR “cod” OR “salmon*” OR “sole” OR “tetra” OR “*trout”) AND TS=(“high* temperature*” OR “elevated temperature*” OR “high* water temperature*” OR “elevated water temperature*” OR “rearing temperature*” OR “effects of temperature*” OR “temperature challenge*” OR “thermal stress*” OR “embr* temper*”) AND TS=(“rear*” OR “incubat*” OR “developmental temperature*”) NOT SO=(“Japanese Edition”) NOT TS= (“cryo*” OR “triploidy” OR “jellyfish”)) AND (SU=(Agriculture OR Behavioral Sciences OR

Biochemistry & Molecular Biology OR Biodiversity & Conservation OR Developmental Biology OR Endocrinology & Metabolism OR Environmental Sciences & Ecology OR Evolutionary Biology OR Fisheries OR Genetics & Heredity OR Marine & Freshwater Biology OR Reproductive Biology OR Research & Experimental Medicine OR Veterinary Sciences OR Zoology)).

The first term in our search string—"*fish*" OR "bass" OR "carp" OR "char" OR "cod" OR "salmon*" OR "sole" OR "tetra" OR "*trout"—was designed to include studies on fish that do not necessarily mention fish in the title, abstract or keywords. To decide on the fish names to include, we compiled a list of the most common fish names from the "list of common fish names" page on Wikipedia. We then performed our search with the individual addition of each of these names and recorded the number of hits. For the names that added >10 hits, we downloaded the titles of these papers and scanned them to see which were suitable. We excluded names that generated many hits for studies that were not on fish, such as "ray".

In addition, on 31 January 2018, we performed a backward and forward search to find the studies cited in, and studies that subsequently cited, Jonsson & Jonsson, 2014. This additional search yielded 294 results. All search results can be downloaded from https://osf.io/e2tyw/

## 2.4 | Study selection

The exact numbers of screened and included studies are shown in Figure S1, and the list of included studies is presented in Table S1.

We used Rayyan software to screen titles and abstracts (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016).

Three people (REO, ML and SN) screened the abstracts, using a decision tree (Figure S2). We had a partial overlap of decisions (36% abstracts screened by more than one person, among which 24% of abstracts had conflicting decisions). Conflicting decisions were discussed and resolved.

Nearly 85% of the 1,610 abstracts were excluded after screening.

We performed full-text screening for the remaining 247 papers included after abstract screening, from which 62 papers were included for data extraction. The full list of screened studies is available from https://osf.io/e2tyw/

## 2.5 | Extracting data

### 2.5.1 | Data collection process

Data were extracted from text, tables, or figures. To extract data from figures, we used the *metaDigitse* package (v.1.0; Pick, Nakagawa, & Noble, 2018) in *R* (v. 3.4.3; R Development Core Team, 2018). All data were extracted by one author (REO), but to verify these extractions half of the data (50% of papers) were checked by other authors. We extracted data as control–treatment pairwise comparisons. For laboratory fish strains, the control temperature was taken as the usual rearing conditions for the system. For wild-caught fish, the temperature used as control was either specified in the paper or was inferred from other studies on the same species. The data were excluded if the "control" temperature was outside the reported optimal temperature range for the species, as reported from the websites Fishbase (Froese & Pauly, 2000) and Animal Diversity Web (University of Michigan Museum of Zoology, 2018). Each pairwise comparison was given a unique ID (unit of analysis), a group ID (the group of eggs that had been split between temperatures), a paper ID (the paper reporting the data) and a species ID (the species that was measured). To minimize errors, data were entered into a relational database, built using Filemaker Pro software (v. 12). Data exported from this software are available from https://osf.io/e2tyw/, and a copy of the relational database is available on request from REO.

### 2.5.2 | Data items

For the full list of moderator variables, see Table S2. For each pairwise comparison, we extracted information about the type, magnitude and length of the temperature treatment. Phenotypic traits were divided into 11 fine categories, which we grouped into four broad categories: (a) behaviour (behaviour); (b) life history (growth); (c) morphology (bone number, condition, morphology, scale number, size); and (d) physiology (heart, metabolism, muscle fibre, swim performance). For each species represented in the database, we gathered information on thermal tolerance and life history from the websites Fishbase (Froese & Pauly, 2000) and Animal Diversity Web (University of Michigan Museum of Zoology, 2018) (when information from these two sources was conflicting, we took the average of the two values). Where available, we recorded the number of parental fish that contributed gametes to each experimental group of fish. When survival data were reported for different incubation temperatures of each group, we extracted this information as well (note that these survival data were collected after the first round of peer review and were therefore not included in the registered analysis plan).

## 2.6 | Analysing data

### 2.6.1 | Effect sizes

To test for phenotypic differences between a treatment and a control group of fish, we calculated two effect sizes for each pairwise comparison, along with their associated sampling variance: the log response ratio (*lnRR*; Hedges, Gurevitch, & Curtis, 1999) and the log coefficient of variation ratio (*lnCVR*; Nakagawa et al., 2015). To test for mean phenotypic differences, we used *lnRR*, which is the natural logarithm of the ratio between the mean phenotype in the treatment and control groups. To test for differences in phenotypic variance, we used *lnCVR*, where the ratio represents the difference between the coefficients of variation (i.e. standard deviations divided by means) for the treatment and control. We used *lnCVR* because, as expected, our data showed a strong positive correlation between mean and variance. We calculated each effect size in *R*, using the *escalc* function in the *metafor* package (v. 2.1-0; Viechtbauer, 2010). For both logged ratios, we specified the treatment group as the numerator and the

control group as the denominator, so that positive values indicate the trait value increased in the treatment group, whereas negative values indicate the trait value decreased in the treatment group.

In addition to calculating the phenotypic differences between the treatment and control groups directly, we also estimated them from random-slope meta-regression models. In this alternative method for estimating variability differences, logged standard deviations are modelled directly while controlling for their corresponding logged mean values (*lnSD*; Raudenbush & Bryk, 1987). This *lnSD* method has greater statistical power to test for differences in variability between a control and treatment (c.f. Nakagawa et al., 2015). More details are given below under *Sensitivity analyses*.

### 2.6.2 | Meta-analysis

We fit meta-analytic and meta-regression multilevel linear mixed-effects models, using the *rma.mv* function in the *metafor* package (v. 2.0-0; Viechtbauer, 2010) in *R* (v. 3.5.2; R Development Core Team, 2018), specifying the Nelder-Mead method of optimization. Model estimates were considered statistically significant if their 95% confidence intervals did not cross zero. Our data contained multiple levels and different types of non-independence (Noble, Lagisz, O'Dea, & Nakagawa, 2017). We partially accounted for this non-independence in two main ways: with random-effects, and with sampling variance–covariance matrices.

To decide on the random-effects structure, we compared null models, which were run using the maximum likelihood method, with combinations of five random effects: unit ID, paper ID, group ID, species, and phylogeny (modelled with a phylogenetic relatedness correlation matrix; to generate the phylogeny (shown in Figure 2) we searched for species names in the Open Tree Taxonomy (Hinchliff et al., 2015), using the *tnrs_match_names* function in the *R* package *rotl* (v. 3.0.4; Michonneau, Brown, & Winter, 2016). We computed branch lengths using the default settings of the *compute.brlen* function in the *R* package *ape* (v. 5.1; Paradis & Schliep, 2019)). The data were structured so that group ID and paper ID were roughly equivalent (as few papers presented data for multiple groups of fish), so only one of these random effects could be fit at a time. Subsequent model selection was based on comparing the models variance components and AIC values, which are shown in Table S3. We chose a model with group ID and unit ID as random effects. Here, the variance component for group ID represents between-group variance, and the variance component for unit ID represents residual (within-group) variance.

We specified sampling variance as variance–covariance matrices, with the sampling variance for each effect size on the diagonal, and the covariance between these measures as off-diagonal elements at appropriate locations. We ran two types of models: "conservative" and "non-conservative". The conservative model assumed a 0.5 correlation between the effect size sample variances with the same group ID. The "non-conservative" model assumed no correlation (i.e. independent sample variances). These two approaches yielded qualitatively similar results; here, we present results from the non-conservative models, but the results for conservative models are presented in the SI.

*Meta-regression*

We estimated the amount of heterogeneity in our data set ($I^2_{total}$) for the multilevel meta-analytic models, using the method described by Viechtbauer (2018). Most meta-analyses in ecology and evolution find high levels of heterogeneity (Senior et al., 2016), and ours were no exception (87% and ~100% for *lnRR* and *lnCVR* meta-analytic models, respectively). We therefore turned to meta-regression models to both explain some of this heterogeneity (the between-study variance and within-study variance), and test our a priori predictions. Missing data were removed for meta-regression models (complete cases analysis). The "full model" included all significant and marginally non-significant (i.e. $p$ value < .1) predictors, after first checking for multicollinearity between the predictors.

*Transformations*

All regression coefficients for continuous moderator variables were estimated at the average values of those predictors, by mean-centring continuous inputs (i.e. subtracting the means from each value of the input variable). Where both the type of treatment (cool or warm) and a continuous variable were fit in the same model, the continuous variable was mean-centred separately for each treatment type (Nakagawa, Kar, O'Dea, Pick, & Lagisz, 2017). In addition, the amount of variation expressed by fish in the control group (*lnCV*) was *z*-scaled to be expressed in standard deviation units.

### 2.6.3 | Sensitivity analyses

To determine the robustness of our results, we performed a number of sensitivity analyses.

*Bayesian meta-analysis*

We re-ran all *lnCVR* and *lnRR* models with an alternative Bayesian approach, using the *MCMCglmm* package (v. 2.26; Hadfield, 2010). In order to weight the studies correctly, we fixed the sampling variance for each effect size by using an inverse-Wishart prior ($V$ = 1, *fix* = 1), so that the sampling variances were taken directly from the studies (which matches the assumption of known sampling variances of meta-analytic models). For the random effect of group ID, we used a parameter expanded prior that was minimally informative ($V$ = 1, *nu* = 0.002, *alpha.mu* = 0, *alpha.V* = 1,000). All MCMC chains were run for 100,000 iterations, with a 10,000 burn and 100 thinning interval, and we visually checked that these chains were mixing well. All results were very similar to those produced using *metafor*, and they are available in the SI.

*lnSD instead of lnCVR*

We used an alternative method to test for differences in variability between the control and treatment groups, where the logged standard deviation (*lnSD*) for each group of fish was the response variable. To account for the mean–variance relationship, the logged mean was included as a fixed effect. We tested for the effect of the treatment by including the treatment factor (either "control" or "treatment") as a fixed effect. In addition to the random effects of unit ID and group ID, we also included a random slope for each control–treatment pairwise

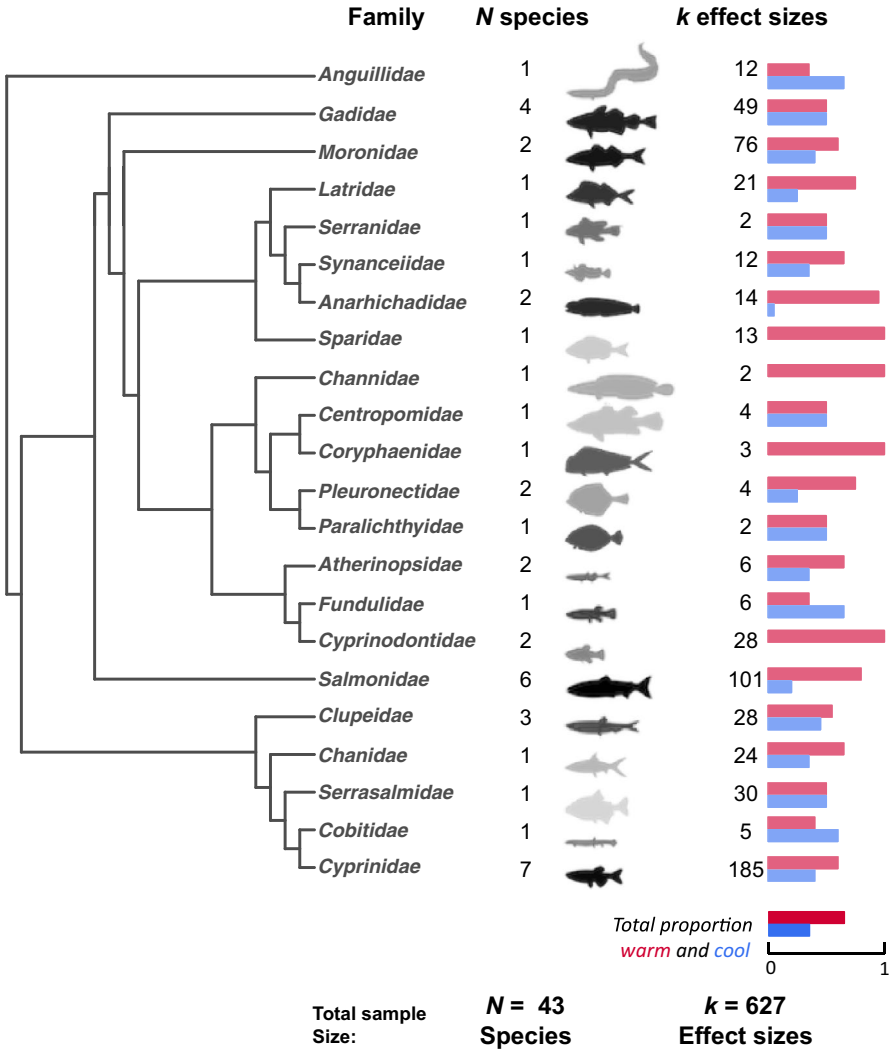


**FIGURE 2** The number of effect sizes for each family of fish, and the number of species representing each family, shown alongside the estimated phylogeny from the Open Tree of Life (Michonneau et al., 2016). The size of fish silhouettes depicts the order of maximum length for the species in that family (range = 6–210 cm), and the silhouette shading level represents the total number of fish measured for species in that family (range = 10–1,960 fish; darker shades depict higher sample sizes). The lengths of horizontal bars correspond to the percentage of effect sizes that originate from warm treatments (red; top bar) and cool treatments (blue; bottom bar) [Colour figure can be viewed at wileyonlinelibrary.com]

comparison. In order to include this random slope, we ran these models using *MCMCglmm* (at the time of writing, this model specification is not supported in *metafor*). This Bayesian approach also allowed us to set 1 as the coefficient of the fixed effect of logged mean, which makes the coefficient for the treatment fixed effect equivalent to *lnCVR* (Nakagawa et al., 2015). For the other fixed effects, we set the prior at 0 with large uncertainty (variance of 10,000,000).

*Publication bias*

As none of our data originates from unpublished studies, the results are at risk of publication bias (a bias towards significant differences). This bias is likely to be a greater issue for mean differences than variance differences, because most studies did not explicitly test for differences in variability. We took three steps to explore whether publication bias was an issue in our data set: first, we first plotted *lnRR* and *lnCVR* against their standard errors (square-root of sampling variances), to look for asymmetry in these funnel plots (Figure S3). Next, we ran Egger's regression on the "meta-analytic residuals" (sensu Nakagawa & Santos, 2012) of effect sizes and their sampling errors. These residuals were calculated from full Bayesian models, including the type of treatments and the interactions with trait type and treatment magnitude for *lnCVR* (Table S17), with the addition of the treatment condition and variability in the control group for *lnRR* (Table S18). Finally, we tested whether studies with larger effects tend to be published earlier (known as the time-lag effect), by including publication year as a moderator variable in meta-regression models (Jennions & Møller, 2002) (Table S21).

*Leave-one-out analyses*

To test how robust our main results were to the exclusion of individual experimental groups of fish, we performed leave-out-one analyses, where we ran the same models multiple times, each time leaving out one subset of data (i.e. excluding one group ID).

## 3 | RESULTS

### 3.1 | Description of data set

Our data set (available from https://osf.io/e2tyw/) includes 62 papers reporting data on 43 species. We analysed 627 effect sizes for the difference between 84 groups of control and treatment fish. The median and mean sample size in each sample (control or treatment group of

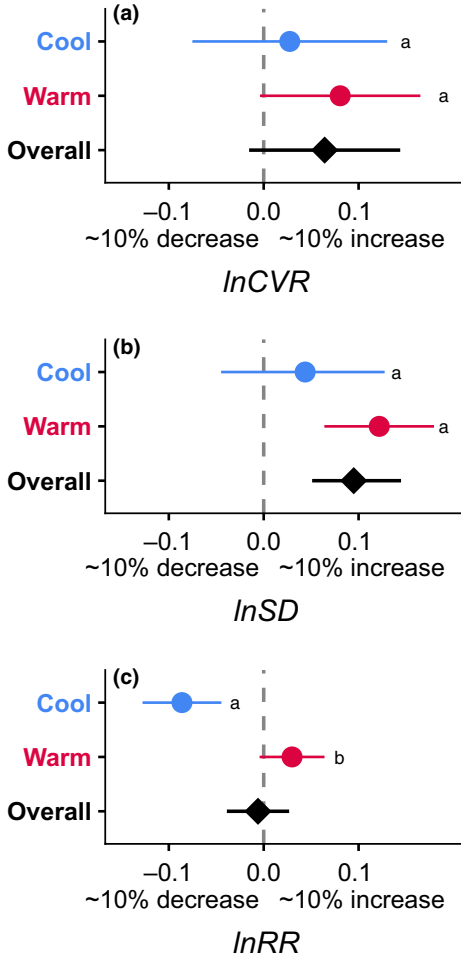


**FIGURE 3** Main meta-analytic and meta-regression results for phenotypic differences between the control and treatment group in (a, b) variability, and (c) means within different trait categories. Whiskers denote 95% confidence intervals; estimates are statistically significant when the confidence intervals do not cross the dashed vertical line at zero. Blue and red circles represent meta-regression intercepts for cool and warm treatments, respectively ($n$ = 217 and $n$ = 410 effect sizes), and lowercase "a" and "b" symbols indicate whether these estimates are significantly different from each other (using a significance threshold of alpha = 0.05). Black diamonds represent meta-analytic intercepts ($n$ = 627). (a, b) Warm treatments tend to increase phenotypic variance, whereas variance in cool treatments is unchanged. Treatment differences in (a) are analysed with *lnCVR*, using a redistricted maximum likelihood model. Treatment differences in (b) are analysed in Bayesian random slope models using *lnSD*, with the log of phenotypic means fixed to 1. (c) Cool treatments tend to decrease phenotypic means, whereas warm treatments show a smaller and non-significant increase in means [Colour figure can be viewed at wileyonlinelibrary.com]

fish) was 30 and 41.6, respectively. Figure 2 shows the spread of data across the phylogeny of species represented in the data set. Warm treatments comprised 65.4% of the effect sizes, and the average increase in temperature for these treatments was 4.4°C. The magnitude of the temperature difference for the cool treatment (34.6% of effect sizes) was 3.1°C. The vast majority of studies measured morphological and physiological response variables (76.2% and 21.1% of effect sizes, respectively). The duration of the temperature treatment was very positively skewed: the median day that fish were measured was 12 days after the treatment started, whereas the average was 35 days. Similarly, the median and mean of the treatment start date was 0 and 1.5 days post-fertilization. In a minority of cases, the fish in the treatment group were brought back to the control temperature before they were measured (i.e. transient treatments: 16% of effect sizes). The median and average number of parents who contributed eggs or sperm to a given experimental group of fish was 10 and 15, based on information available for 70% of effect sizes.

### 3.1.1 | Did warm temperatures increase phenotypic variability?

Fish reared in warmer temperatures expressed 8.4% more variable phenotypes than fish reared in normal temperatures, albeit this difference was marginally non-significant (*lnCVR*: 0.081, 95% confidence interval, CI: −0.004 to 0.165; Figure 3a, Table S4). Using the alternative *lnSD* method of analysis, we estimated a statistically significant 13.0% increase in phenotypic variability in warm temperature treatments ($lnSD_{\text{control–warm treatment slope}}$: 0.122, 95% confidence interval, CI: 0.064–0.179; Figure 3b, Table S5).

### 3.1.2 | Did changed developmental temperatures change phenotypic means?

Warm temperatures tended to show a statistically non-significant 3.1% increase in the means of phenotypic traits ($lnRR_{\text{warm intercept}}$: 0.030, CI: −0.004 to 0.064; Figure 3c, Table S7). Among different types of phenotypic traits, only growth rate (which was classified as life history) showed a statistically significant 38% increase in warmer temperatures, but note that this estimate is based on very little data so it is not informative ($n$ = 4; Figure 4; Table S9). Warm temperature treatments did not reduce mean values in any of the broad categories of phenotypic traits (Table S9).

Cool treatments showed a larger effect on phenotypic means than warm treatments and significantly reduced trait means by 8.3% ($lnRR_{\text{cool intercept}}$: −0.086, CI: −0.128 to −0.045; Figure 4, Table S7).

Because cool and warm treatments had opposing effects on phenotypic means, in the combined meta-analysis these treatments effectively cancelled each other out. The overall meta-analytic mean therefore found no change in the mean phenotype as a result of changes in the developmental temperature ($lnRR_{\text{intercept}}$: −0.006, CI: −0.039 to 0.027; Figure 3c, Table S7).

### 3.1.3 | The differences between cool and warm treatments

In meta-regressions of mean phenotypic differences, the type of treatment (cool or warm temperatures) was an important moderator variable to account for heterogeneity in the size and magnitude of effects (*lnRR* $Q_m$ = 35.53, $df$ = 1, $p$ < 0.000; Table S19). Warm

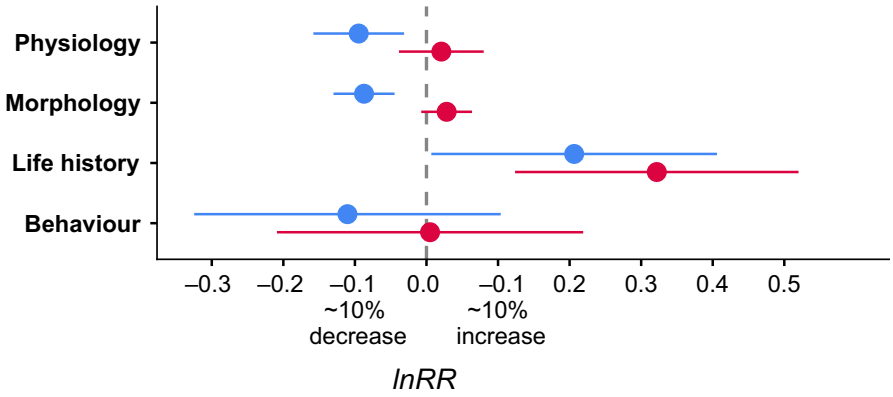


**FIGURE 4** Effects of temperature treatments on phenotypic means, within different types of trait categories. Whiskers denote 95% confidence intervals; estimates are statistically significant if the confidence intervals do not cross the dashed vertical line. The confidence intervals for physiology and morphology are narrower than for life history and behaviour, because they are estimated from more data. Sample sizes are shown in Table S9 [Colour figure can be viewed at wileyonlinelibrary.com]

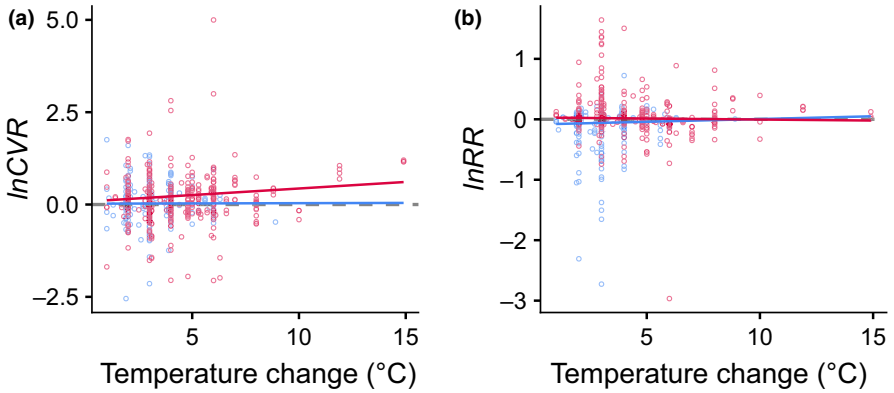


**FIGURE 5** The relationship between the magnitude of temperature change (absolute values) and phenotypic effects for (a) differences in variability and (b) differences in means. Note the scale of the *y*-axis is wider than the *x*-axis in Figure 3. Open circles are raw values, and solid lines show the intercept and slope estimates for meta-regression models. Cool treatments are shown in blue, and warm treatments are shown in red. (a) Warm treatments tend to increase phenotypic variability, and the magnitude of this effect significantly increases as the treatment moves further from the control temperature. Cool treatments do not affect phenotypic variability, regardless of the size of the temperature difference. (b) The magnitude of temperature change had no impact on the size of the phenotypic mean difference. Sample sizes are shown in Tables S10 and S12 [Colour figure can be viewed at wileyonlinelibrary.com]

treatments caused a marginal increase in means, cool treatments caused a statistically significant decrease in means, and warm and cool treatment mean effects were significantly different from each other by 11.0% ($lnRR_{warm-cool\ slope}$: −0.116 CI: −0.154 to −0.078; Figure 3c, Figure 4, Table S7).

In contrast to phenotypic differences in means, the type of treatment was less important for meta-regressions of phenotypic differences in variability ($lnCVR\ Q_m$ = 1.236, *df* = 1, *p* = 0.266; Table S19). While the tendency for variability to increase was driven by warm temperature treatments, the 5.2% contrast between the treatment types was non-significant ($lnCVR_{warm-cool\ slope}$: −0.053, CI: −0.147 to 0.041; Figure 3a, Table S4).

### 3.1.4 | Do larger changes in temperature cause larger effects?

Larger temperature treatments caused greater variability increases in the warm temperature treatments, but not the cool temperature treatments. The slope of the meta-regression model indicated that a 1-degree increase in warm temperatures caused a significant 3.5% increase in variability ($lnCVR_{warm\ degree\ difference\ slope}$: 0.036, CI: 0.009–0.062; Figure 5a, Table S10). To illustrate the effect of the magnitude of the temperature change, we ran post hoc meta-regression models where the intercept was estimated at different distances from the control temperature (Figure 6; Table S11). At the average magnitude of warm temperature treatments (4.4°C), the model predicted an 8.2% increase in variability. When we shifted the intercept to 3 degrees warmer, at 7.4°C, the phenotypic variability increased by 20.3% ($lnCVR_{warm\ average\ magnitude\ +\ 3\ degrees\ intercept}$: 0.185, CI: 0.071–0.300 Figure 6a; Table S11).

A change in the developmental temperature caused a change in phenotypic means, but the magnitude of this difference did not increase as the temperature moved further away from the control ($lnRR_{warm\ degree\ difference\ slope}$: −0.003 CI: −0.014 to 0.008; $lnRR_{cool\ degree\ difference\ slope}$: 0.009 CI: −0.015 to 0.033; Figure 5b, Table S12). Post hoc meta-regression models confirmed that increasing the magnitude of temperature change did not increase the magnitude, or statistical significance, of the mean difference estimates (Figure 6b; Table S13).

To illustrate the combined effects of temperature treatments on the mean and variance of phenotypic traits, we present simulated normal distributions of a phenotypic trait, based on our model estimates, in Figure 7.

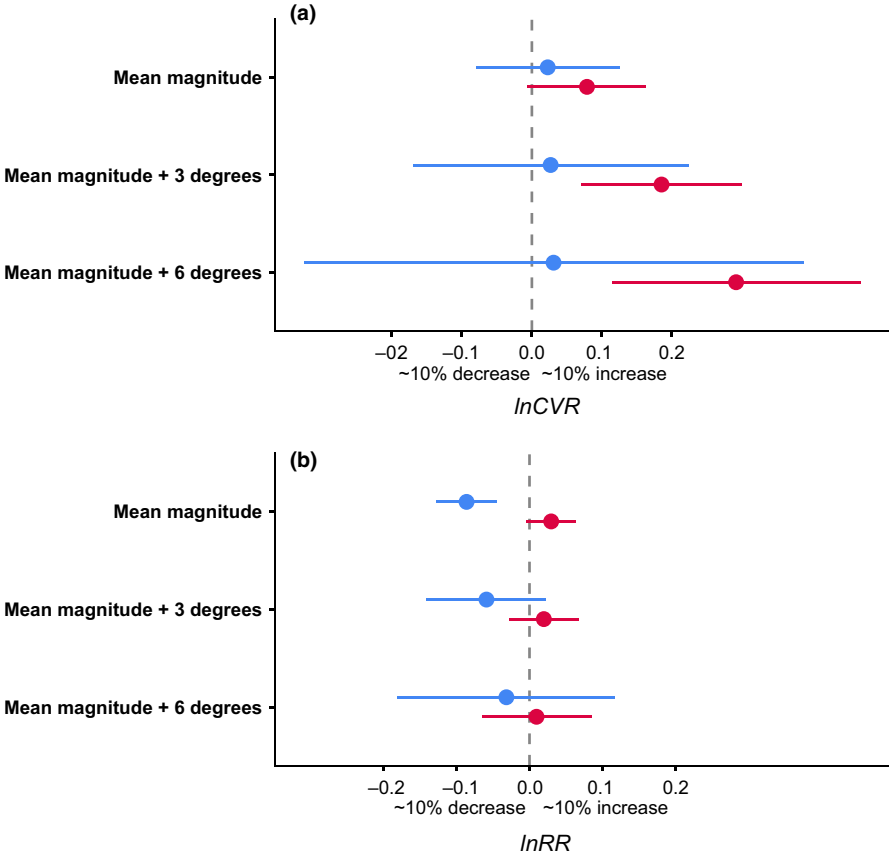


**FIGURE 6** Estimates of the phenotypic differences in (a) means and (b) variability between control and treatment groups, estimated from the intercept of meta-regression models with the type of treatment—cool (blue) or warm (red)—and the magnitude of temperature change. Each model includes the interaction term, and the magnitude of temperature change is shifted separately for each type of treatment, to estimate the intercept at different magnitudes of temperature change. The "mean magnitude" is 3.1 degrees for cool treatments, and 4.4 degrees for warm treatments. Whiskers denote 95% confidence intervals; estimates are statistically significant if the confidence intervals do not cross the dashed vertical line [Colour figure can be viewed at wileyonlinelibrary.com]

### 3.1.5 | Do longer treatment durations cause larger effects?

The length of the treatment (mean ± *SD* = 35.5 ± 44.9 days) had no effect on the magnitude of phenotypic differences between the treatment and control groups, in either mean or variability (variance: $lnCVR_{\text{treatment duration slope}}$: 0.000, CI: −0.002 to 0.001; mean: $lnRR_{\text{treatment duration slope}}$: 0.000, CI: 0.000–0.001; Figure S6, Table S14).

### 3.1.6 | Did changing the temperature earlier cause larger effects?

In order to be included in our meta-analysis, the temperature treatment had to start before or on the day of hatching (i.e. the time between eggs being fertilized and developing into larvae). Within this limited range of time for the treatment to start, we found little effect of the timing of the treatment on the magnitudes of phenotypic differences (variance: $lnCVR_{\text{treatment start slope}}$: −0.006, CI: −0.012 to 0.000; mean: $lnRR_{\text{treatment start slope}}$: −0.001, CI: −0.003 to 0.002; Figure S7, Table S14).

### 3.1.7 | Do permanent treatments cause larger effects than transient treatments?

The permanence of the treatment condition (permanent or transient) had no effect on phenotypic variability ($lnCVR_{\text{permanent–transient difference}}$: 0.042, CI: −0.124 to 0.207; Figure S8, Table S15), and a significant effect on phenotypic means (7.4% difference between permanent and transient treatments; $lnRR_{\text{permanent-transient difference}}$: −0.077, CI: −0.143 to −0.012; Figure S8, Table S15). Because transient treatments were only a small portion of our data set (16% of effect sizes), it is possible that this difference was not due to the treatment conditions per se, but rather due to uneven sampling of other moderator variables. For example, cool treatments were over-represented in the transient data subset (49% cool treatments in transient treatments compared with 32% in permanent treatments). We therefore included both treatment type and treatment condition as fixed effects in the "full model" (which also included trait type and variability of the control group; Table S18). In the full model, both cool and warm treatments showed a reduction in mean phenotype in transient compared with permanent conditions. Because cool treatments tended to decrease the phenotypic mean, this suggests that transient treatments had a larger phenotypic effect than permanent treatments. In contrast, warm treatments tended to increase phenotypic means (albeit not statistically significantly), suggesting that warm treatments had larger phenotypic effects when the treatment condition was permanent rather than transient.

### 3.1.8 | Does having more variation allow for larger average plastic responses?

The amount of phenotypic variability in normal temperatures affected the amount of developmental plasticity expressed in experimental temperatures. In warm temperature treatments, at the average magnitude of the temperature change (4.4°C), an increase in baseline variability of one standard deviation estimated a 10%

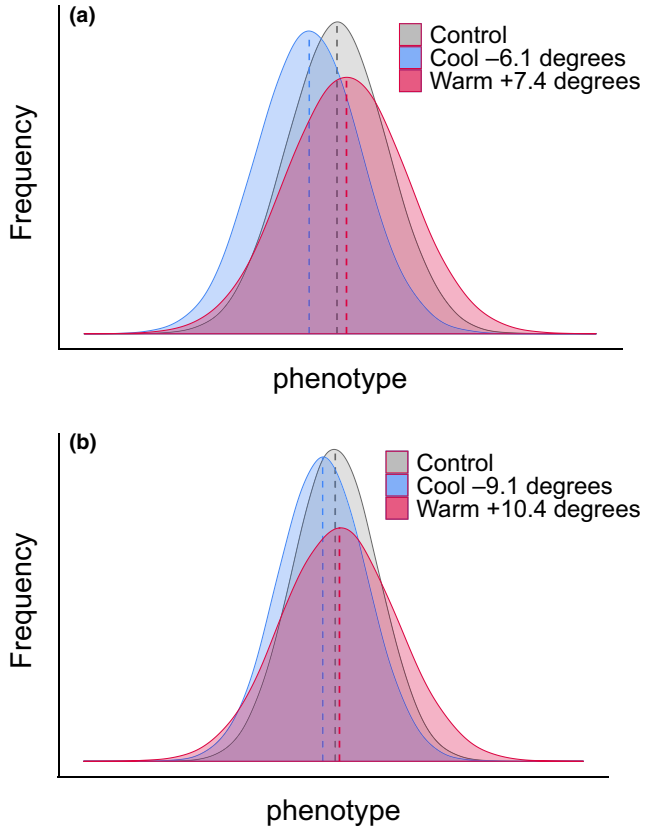


**FIGURE 7** Phenotypic distributions for fish in the control and treatment groups when the temperature is changed beyond the average treatment magnitude (a) by 3 degrees and (b) by 6 degrees celsius, based on simulations of model estimates. Control group = grey, cool treatment = blue, and warm treatment = red. Dashed vertical lines show the phenotypic means. (a) At 6.1 degrees below normal, cool treatments decrease trait means but cause no noticeable difference in variance. At 7.4 degrees above normal, warm treatments show a smaller increase in means, but show a noticeable increase in variance. (b) If the treatment moves 3 degrees further from the control, the differences in means do not increase (if anything, they decrease), but the warm treatment variability continues to increase [Colour figure can be viewed at wileyonlinelibrary.com]

increase in the phenotypic mean ($lnRR_{\text{warm control variability slope}}$: 0.095, CI: 0.068–0.122; Figure S9, Table S16). The same increase in variation was associated with a 7.0% decrease in the phenotypic mean at the average magnitude of cool temperature treatments (3.1°C) ($lnRR_{\text{cool control variability slope}}$: −0.072, CI: −0.108 to −0.037; Figure S9, Table S16). The relationship between the magnitude of variability of the control group and the magnitude of plasticity was particularly consistent for warm treatments, with the slope remaining statistically significant in all full models (Table S18).

### 3.1.9 | Does the number of parents impact the plastic response of a population?

We predicted that groups of fish from greater numbers of parents would show greater plastic responses to changes in the developmental temperature, but this was not the case (Table S14). The basis for our prediction was that greater genetic diversity would lead to greater phenotypic diversity. This assumption was not statistically supported: a post hoc meta-regression found no significant relationship between the number of parents and the amount of phenotypic variability expressed in normal temperatures ($lnCV_{\text{number parents slope}}$: 0.153, CI: −0.036 to 0.342; Figure S10c). We also recognize that, in this data set, groups of fish from greater numbers of parents might not be more genetically diverse, as different groups were kept or collected under different experimental conditions, and encompassed a wide range of species.

### 3.1.10 | Does the distance of the treatment from the species' thermal limit matter?

The magnitude of the temperature change had no impact on mean phenotypic differences, regardless of the distance of the temperature change from the optimal thermal limit of the species (Tables S12 and S14). The magnitude of the temperature change did matter for variability: larger differences in temperature induced greater increases in variability (Table S10). However, pushing temperatures beyond the thermal limit of the species did not induce greater phenotypic variability. An increase in distance from the thermal limit of 1 degree tended to reduce the variability difference by 1.2% ($lnCVR_{\text{distance from thermal limit slope}}$: −0.012, CI: −0.023 to −0.001; Table S14).

## 3.2 | Are phenotypic differences correlated with survival?

We ran post hoc exploratory analyses to see whether phenotypic differences were correlated with survival differences between a control and treatment group, using the subset of studies that reported survival information (survival data were available for 38 groups of fish, which represented 44.3% of the whole data set). The reason for collecting survival data is that phenotypic differences in the temperature treatments could be caused by differential survival, rather than phenotypic plasticity.

Differences in phenotypic variability showed a non-significant tendency to be greater when the temperature treatment decreased survival, in both warm and cool treatments ($lnCVR_{\text{warm survival slope}}$: 0.23, CI: −0.359 to 0.819; $lnCVR_{\text{cool survival slope}}$: 0.76, CI: −0.368 to 1.888; Table S22; Figure S11a).

The phenotypic mean in the temperature treatment tended to decrease as survival decreased as well, with the trend being most noticeable for cool treatments. For cool treatments, which showed a negative estimate for *lnRR*, decreased survival was associated with an increase in phenotypic differences between the control and temperature treatment ($lnRR_{\text{warm survival intercept}}$: −0.023, CI: −0.068 to 0.023; $lnRR_{\text{warm survival slope}}$: −0.661, CI: −0.971 to −0.351; Table S22; Figure S11b). For warm treatments, which showed a positive estimate for *lnRR*, decreased survival slightly reduced phenotypic differences between the control and temperature treatment ($lnRR_{\text{warm survival intercept}}$: 0.037, CI: 0.001–0.072; $lnRR_{\text{warm survival slope}}$: −0.206, CI: −0.35 to −0.062; Table S22; Figure S11b).

### 3.3 | Publication bias and sensitivity analyses

#### 3.3.1 | Funnel plots and leave-one-out

Visual inspection of funnel plots indicated some asymmetrical distribution of effect sizes around the meta-analytic mean (Figure S3a and S3b). The usefulness of these funnel plots for multivariate meta-analyses is debatable, however, and when we averaged the effect sizes within groups of fish (the main random effect), the funnel plots looked more symmetrical (Figure S3c and S3d). To test the sensitivity of our meta-analytic and meta-regression means to exclusion of certain levels of the group-ID random effect, we ran "leave-one-out" analyses. We re-ran the meta-analytic model and meta-regression of treatment type (cool and warm) after removing one experimental group of fish ($n$ = 84 models, for the 84 groups of fish in the data set), and compared the estimates and confidence intervals to our overall results. The estimates for mean and variance differences overall, and in cool and warm treatments, appeared fairly robust (Figures S4 and S5).

#### 3.3.2 | Publication bias

The results of Egger's regression on the meta-analytic residuals indicated the presence of publication bias in the data set for phenotypic differences in means, but not variability (Table S20). However, we did not find evidence of a time-lag bias (larger effect sizes were not published earlier; Table S21).

## 4 | DISCUSSION

Changing fish's developmental temperature changed their average phenotype, but in opposing directions depending on whether the temperature was increased or decreased. These shifts in average phenotype could indicate any combination of adaptive plasticity (Figure 1b), maladaptive responses to thermal stress, or selection. Increases in temperature, but not decreases, also increased phenotypic variability, which suggests a reduction in genotype precision that causes a spreading in reaction norms (Hansen et al., 2006; Snell-Rood, Kobiela, Sikkink, & Shephard, 2018) (Figure 1). These effects were not significantly moderated by the starting date or duration of the temperature change, the distance of the new temperature from the thermal limit of the species, or the number of parents who contributed to the spawning, and there was a great deal of unexplained variation in effects between studies. The shift in average phenotype was limited, as the average phenotype did not continue to change as the temperature changed (although warmer temperatures did induce more variation around this mean). Populations that expressed more variability in normal conditions showed larger plastic responses in both mean and variance. Combined, these results demonstrate warmer-than-standard developmental temperatures can increase the frequency of rare phenotypes in fish populations, and potentially induce novel phenotypes.

### 4.1 | Increased variability in novel environments

We found that warm environments increase phenotypic variability, despite no change in the phenotypic mean in most trait categories (except growth rate), with larger changes in the environment causing greater expression of stochastic plasticity. The evolutionary consequences of phenotypic plasticity have long been debated and depend on the extent to which the plasticity is heritable (Crispo, 2007). Hansen et al.'s, 2006 literature survey suggests an underappreciated source of phenotypic variation is genotype imprecision (whereby genotypes do not precisely produce their target phenotype, so this variation is not heritable). Even non-heritable phenotypic variation could be beneficial in challenging environments, if this improves the odds of some individuals enduring temporary warming periods. On an evolutionary timescale, phenotypic variation caused by cryptic genetic variation could quickly spread through the population. A permanent environmental change might eventually select for the plastic phenotype to be produced regardless of the developmental temperature (i.e. genetic assimilation; Crispo, 2007). There is some evidence, based on genotype-by-environment interactions, for the adaptive potential of warm-induced variants in a coral reef fish and a salmon species (*Acanthochromis polyacanthus*: Munday, Donelson, & Domingos, 2017, *Onchorhynchus nerka*: Burt, Hinch, & Patterson, 2012). Warm temperatures in early life also cause epigenetic changes (e.g. sex determination; Piferrer, Ribas, & Diaz, 2012), and these epigenetic effects could range from short term (Campos, Valente, Conceicao, Engrola, & Fernandes, 2014) to transgenerational (Burton & Metcalfe, 2014). Of course, the beneficial effects of increased variability are entwined with population size, because small populations cannot afford to lose a large fraction of their population (e.g. Devils Hole pupfish, *Cyprinodon diabolis*; Jones et al., 2016).

### 4.2 | The direction of the temperature change matters

Compared to warm treatments, and contrary to our predictions, cool treatments had no significant effect on phenotypic variability and caused a larger shift in the phenotypic mean. An artificial explanation for this result is that the "cool" and "warm" categories in our data could have been inaccurate; perhaps the experimental fish represented in the meta-analyses were kept closer to their thermal maximum than their thermal minimum to accelerate development, as is common in aquaculture (Arguello-Guevara, Bohorquez-Cruz, & Silva, 2017). If we accept the temperature categories, then there are competing explanations for their differences, depending on whether the temperature-induced shift in average phenotype is interpreted as (a) an adaptive response to a shift in the optimal phenotype; or (b) a maladaptive response to thermal stress (either maladaptive plasticity or selection). Under the first scenario, assuming that the optimal phenotype is linearly correlated with the environmental temperature, fishes seemed to respond better to cool rather than warm temperature changes—perhaps in cool temperatures, they are relieved of constraints that exist at warm temperatures (e.g. metabolic

constraints; Pörtner & Lannig, 2009). More speculatively, the ancestral history of fishes might have occurred in cooler environments, so that cool temperatures represent a familiar change that fish can adaptively respond to (Figure 1b), whereas warm temperatures are a novel stressor that triggers an increase in developmental noise (Figure 1c) (Ghalambor et al., 2007). The second scenario leads to the opposite interpretation; cool temperatures are more likely than warm temperatures to cause a slide in the population mean away from the adaptive optimum (i.e. fishes in warm temperatures show greater "genetic compensation", sensu Grether, 2005). While we cannot distinguish between these competing explanations, an exploratory analysis did suggest that mean phenotypic differences in cool temperatures were higher with increasing mortality (potentially indicating maladaptation and selection). Future research is needed to test whether or not phenotypic changes are adaptive.

### 4.3 | Limited plasticity in population means

Contrary to our expectations, larger temperature changes did not cause larger shifts in the phenotypic mean (i.e. limited directed plasticity). Similar results have been found for thermal acclimation at later developmental stages in coral reef fishes (Donelson, 2015; Grenchik, Donelson, & Munday, 2013). Again, alternative interpretations depend on whether a shift in the average phenotype is considered adaptive. Adaptive plasticity might be constrained at more extreme temperatures. For example, the oxygen and capacity limited thermal tolerance hypothesis predicts fishes growth and aerobic scope will be constrained in warm temperatures, as increases in basal metabolic demands outpace resource consumption and the availability of dissolved oxygen (Donelson, Munday, McCormick, & Nilsson, 2011; Pörtner & Lannig, 2009). Alternatively, plasticity is likely to be costly in novel environments (Snell-Rood et al., 2018), so the average phenotype might be "fixed" in order to prevent a maladaptive slide in the population average in response to environmental perturbations. This canalization is seen in examples of genetic compensation and counter-gradient variation (Grether, 2005).

### 4.4 | Average differences could be over-estimated

We found some evidence that studies reporting large average differences between treatment and control groups were over-represented in our data set. This pattern could reflect publication bias and selective reporting within studies, whereby "positive" findings are more likely to be published and reported by authors than null results (Jennions, Lortie, Rosenberg, & Rothstein, 2013). In contrast to mean differences, our effect sizes for variance differences did not show evidence of selective reporting and/or publication bias. This is not surprising, because studies typically do not test hypotheses based on variability differences (an exception in our data set was Burt et al., 2012). Unfortunately, while there has been a recent push towards increasing transparency in scientific publications, our review found low uptake of these initiatives within this field (Nosek et al., 2015). Only two studies included in the data set had data readily available to download online, and many studies were excluded due to low reporting standards of essential information. We therefore urge future studies on the effects of developmental temperature to make all data publicly available, to reduce the adverse effects of selective reporting in research synthesis (Parker et al., 2016).

### 4.5 | Other limitations and future directions

Our data had limited coverage over some moderator variables for which we tested predictions, which highlights areas warranting future research. The vast majority of traits represented in our data set were morphological (mostly length and mass). Future studies should focus on other types of traits, such as behaviour, for which little data were available (this gap was also identified by a more general review of the effects of fishes' rearing environments; Jonsson & Jonsson, 2014). These data could facilitate better predictions of how fishes respond to climate change. Longer-term studies are also required to assess the lifelong implications (i.e. fitness) of different developmental environments; the majority of our data set consists of phenotypic traits measured in juvenile fish. It is valuable to measure adult fish because the phenotypic response of fish to different temperatures can vary depending on the measured life stage, due to changes in both optimal temperature conditions (Arguello-Guevara et al., 2017) and different resource requirements (e.g. endogenous vs. exogenous feeding; Baras et al., 2012). For example, the dominance of juvenile measurements in our data set could account for why we found an overall increase in body size at warm temperatures, despite a decrease being generally expected in adult fishes (Burt et al., 2012; Kim, Metcalfe, Silva, & Velando, 2017; Munday et al., 2017). Another explanation could be the generally benign conditions experienced in laboratory settings, which could mask resource limitation trade-offs that would be expected in nature (Munday, Kingsford, O'Callaghan, & Donelson, 2008). To see whether the patterns of increasing variability with increasing temperature are seen in the field as well as the laboratory, a potential study could use existing long-term data sets of wild fish populations.

### 4.6 | Practical implications

As the world warms, and temperature fluctuations become more frequent and severe (Bathiany, Dakos, Scheffer, & Lenton, 2018), how will fishes respond? In the short term, our results suggest minimal responses of the average population phenotypes, but (as predicted by Ghalambor et al., 2007) an increase in phenotypic variants in the population. If the initial population size is large enough, then increased variability should increase the likelihood of that population surviving and adapting to the new environment. Importantly, the potential for increased variation in warm environments is predicted by a population's underlying level of phenotypic variability. To reduce the economic impact of climate change on fisheries, therefore, our results reaffirm that it is important that harvested populations maintain phenotypic variation. The importance of maintaining phenotypic variation could affect management strategies for harvested fish populations (Villegas-Ríos,

Moland, & Olsen, 2016). Large and diverse populations will stand the best chance of adapting to environmental change. Additional sources of stress should be reduced as much as possible; for example, survival of hatchery-reared *Salmo salar* during a heatwave was improved through minimizing larval stress by mimicking natural rearing conditions (Bamberger, 2009).

## 5 | CONCLUSIONS

We found proof-of-concept support for an increase in phenotypic variability in warm environments, especially for large changes in temperature. Unusual phenotypes that are induced by the environment could facilitate adaptation to novel environments. We encourage future studies to report and consider the implications of this variation. Further empirical research will be needed to determine how often variants induced by the environment are heritable and stable, whether heritable variants are caused by underlying cryptic genetic variation or epigenetic modifications, and whether the propensity for variability is itself heritable (i.e. heritable bet-hedging, sensu O'Dea et al., 2016). Future theoretical work should consider the implications of environmental effects on intraspecific variation for evolutionary and ecological models (Bolnick et al., 2011). As environmental conditions are becoming increasingly unpredictable, the capacity of species to produce and maintain phenotypic variability might be a crucial determinant of long-term population survival.

### ACKNOWLEDGEMENTS

REO was funded for this project by a 2017 Endeavour Australian Postgraduate Scholarship. SN was supported by an ARC Future Fellowship (FT130100268), and ML and SN were supported by a Discovery grant (DP180100818).

### DATA ACCESSIBILITY STATEMENT

Data, analysis code, and lists of screened studies are available to download from https://osf.io/e2tyw/ (O'Dea et al., 2018b).

### ORCID

*Rose E. O'Dea* https://orcid.org/0000-0001-8177-5075
*Malgorzata Lagisz* https://orcid.org/0000-0002-3993-6127
*Andrew P. Hendry* https://orcid.org/0000-0002-4807-6667
*Shinichi Nakagawa* https://orcid.org/0000-0002-7765-5182

### REFERENCES

*Indicates articles used in the meta-analysis.

*Abdel, I., Abellán, E., López-Albors, O., Valdés, P., Nortes, M. J., & García-Alcázar, A. (2004). Abnormalities in the juvenile stage of sea bass (*Dicentrarchus labrax* L.) reared at different temperatures: Types, prevalence and effect on growth. *Aquaculture International*, *12*(6), 523–538. https://doi.org/10.1007/s10499-004-0349-9

*Ackerly, K. L., & Ward, A. B. (2015). How temperature-induced variation in musculoskeletal anatomy affects escape performance and survival of zebrafish (*Danio rerio*). *Journal of Experimental Zoology. Part A, Ecological Genetics and Physiology*, *325*(1), 25–40. https://doi.org/10.1002/jez.1993

*Alami-Durante, H., Rouel, M., & Kentouri, M. (2006). New insights into temperature-induced white muscle growth plasticity during *Dicentrarchus labrax* early life: A developmental and allometric study. *Marine Biology*, *149*(6), 1551–1565. https://doi.org/10.1007/s00227-006-0304-6

*Albokhadaim, I., Hammond, C. L., Ashton, C., Simbi, B. H., Bayol, S., Farrington, S., & Stickland, N. (2007). Larval programming of post-hatch muscle growth and activity in Atlantic salmon (*Salmo salar*). *Journal of Experimental Biology*, *210*(10), 1735–1741. https://doi.org/10.1242/jeb.003194

*Anastasiadi, D., Diaz, N., & Piferrer, F. (2017). Small ocean temperature increases elicit stage-dependent changes in DNA methylation and gene expression in a fish, the European sea bass. *Scientific Reports*, *7*(1), 68. https://doi.org/10.1038/s41598-017-10861-6

Arguello-Guevara, W., Bohorquez-Cruz, M., & Silva, A. (2017). Effect of two temperatures on yield and increase in cranial skeletal abnormalities during early development of palm ruff, *Seriolella violacea* (Guichenot 1848). *Aquaculture Research*, *48*(1), 298–310. https://doi.org/10.1111/are.12882

*Arul, V. (1991). Effect of temperature on yolk utilization of *Channa striatus*. *Journal of Thermal Biology*, *16*(1), 1–5. https://doi.org/10.1016/0306-4565(91)90043-2

*Aydin, I., Küçük, E., Şahin, T., & Kumlu, M. (2015). Effect of temperature on reversed asymmetry in hatchery-reared flounder (*Platichthys flesus* luscus Pallas, 1811). *Turkish Journal of Fisheries and Aquatic Sciences*, *15*, 737–740. https://doi.org/10.4194/1303-2712-v15_3_17

Bamberger, A. (2009). Evaluation of a new semi-natural incubation technique for Atlantic salmon *Salmo salar* fry. *Journal of Fish Biology*, *74*(7), 1419–1433. https://doi.org/10.1111/j.1095-8649.2009.02208.x

Baras, E., Ginanjar, R., Ahmad, M., Permana, A., Priyadi, A., Legendre, M., et al. (2012). Biology and culture of the clown loach *Chromobotia macracanthus* (Cypriniformes, Cobitidae): 4-thermal biology of embryos and larvae. *Aquatic Living Resources*, *25*(2), 131–142. https://doi.org/10.1051/alr/2012012

*Barrionuevo, W. R., & Burggren, W. W. (1999). O2 consumption and heart rate in developing zebrafish (*Danio rerio*): Influence of temperature and ambient O2. *American Journal of Physiology, Regulatory, Integrative and Comparative Physiology*, *276*(2), R505–R513. https://doi.org/10.1152/ajpregu.1999.276.2.R505

Bathiany, S., Dakos, V., Scheffer, M., & Lenton, T. M. (2018). Climate models predict increasing temperature variability in poor countries. *Science Advances*, *4*(5), https://doi.org/10.1126/sciadv.aar5809

*Berlinsky, D. L., Taylor, J. C., Howell, R. A., Bradley, T. M., & Smith, T. I. J. (2004). The effects of temperature and salinity on early life stages of Black Sea Bass *Centropristis striata*. *Journal of the World Aquaculture Society*, *35*(3), 335–344. https://doi.org/10.1111/j.1749-7345.2004.tb00097.x

Bolnick, D. I., Amarasekare, P., Araújo, M. S., Bürger, R., Levine, J. M., Novak, M., … Vasseur, D. A. (2011). Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution*, *26*(4), 183–192. https://doi.org/10.1016/j.tree.2011.01.009

*Brooks, S., & Johnston, I. A. (1993). Influence of development and rearing temperature on the distribution, ultrastructure and myosin sub-unit composition of myotomal muscle-fibre types in the plaice *Pleuronectes platessa*. *Marine Biology*, *117*(3), 501–513. https://doi.org/10.1007/BF00349326

*Brown, C. A., Gothreaux, C. T., & Green, C. C. (2011). Effects of temperature and salinity during incubation on hatching and yolk utilization of

Gulf killifish *Fundulus grandis* embryos. *Aquaculture*, *315*(3–4), 335–339. https://doi.org/10.1016/j.aquaculture.2011.02.041

*Burt, J. M., Hinch, S. G., & Patterson, D. A. (2012). Parental identity influences progeny responses to incubation thermal stress in sockeye salmon *Onchorhynchus nerka*. *Journal of Fish Biology*, *80*(2), 444–462. https://doi.org/10.1111/j.1095-8649.2011.03190.x

Burton, T., & Metcalfe, N. B. (2014). Can environmental conditions experienced in early life influence future generations? *Proceedings of the Royal Society B: Biological Sciences*, *281*(1785), 20140311. https://doi.org/10.1098/rspb.2014.0311

Campos, C., Valente, L. M. P., Conceicao, L. E. C., Engrola, S., & Fernandes, J. M. O. (2014). Molecular regulation of muscle development and growth in Senegalese sole larvae exposed to temperature fluctuations. *Aquaculture*, *432*, 418–425. https://doi.org/10.1016/j.aquaculture.2014.04.035

*Canino, M. F. (1994). Effects of temperature and food availability on growth and RNA/DNA ratios of walleye pollock *Theragra chalcogramma* (Pallas) eggs and larvae. *Journal of Experimental Marine Biology and Ecology*, *175*(1), 1–16. https://doi.org/10.1016/0022-0981(94)90173-2

*Carey, G. R., & Franklin, C. E. (2009). Effect of incubation and rearing temperature on locomotor ability in barramundi, *Lates calcarifer* Bloch, 1790. *Marine and Freshwater Research*, *60*(3), 203–208. https://doi.org/10.1071/MF07250

*Carmichael, G. J. (1983). Scale-number differences of central stonerollers incubated and reared at different temperatures. *Transactions of the American Fisheries Society*, *112*(3), 441–444. https://doi.org/10.1577/1548-8659(1983)112<441:SDOCSI>2.0.CO;2

*Colchen, T., Teletchea, F., Fontaine, P., & Pasquet, A. (2017). Temperature modifies activity, inter-individual relationships and group structure in fish. *Current Zoology*, *63*(2), 175–183. https://doi.org/10.1093/cz/zow048

Crispo, E. (2007). The Baldwin effect and genetic assimilation: Revisiting two mechanisms of evolutionary change mediated by phenotypic plasticity. *Evolution*, *61*(11), 2469–2479. https://doi.org/10.1111/j.1558-5646.2007.00203.x

*de Assis, J. M. F., Carvalho, R. F., Barbosa, L., Agostinho, C. A., & Pai-Silva, M. D. (2004). Effects of incubation temperature on muscle morphology and growth in the pacu (*Piaractus mesopotamicus*). *Aquaculture*, *237*(1–4), 251–267. https://doi.org/10.1016/j.aquaculture.2004.04.022

*DiMaria, R. A., Miller, J. A., & Hurst, T. P. (2010). Temperature and growth effects on otolith elemental chemistry of larval Pacific cod, *Gadus macrocephalus*. *Environmental Biology of Fishes*, *89*(3–4), 453–462. https://doi.org/10.1007/s10641-010-9665-2

Donelson, J. M. (2015). Development in a warm future ocean may enhance performance in some species. *Journal of Experimental Marine Biology and Ecology*, *472*, 119–125. https://doi.org/10.1016/j.jembe.2015.07.008

Donelson, J. M., Munday, P. L., McCormick, M. I., & Nilsson, G. E. (2011). Acclimation to predicted ocean warming through developmental plasticity in a tropical reef fish. *Global Change Biology*, *17*(4), 1712–1719. https://doi.org/10.1111/j.1365-2486.2010.02339.x

*Dou, S. Z., Masuda, R., Tanaka, M., & Tsukamoto, K. (2005). Effects of temperature and delayed initial feeding on the survival and growth of Japanese flounder larvae. *Journal of Fish Biology*, *66*, 362–377. https://doi.org/10.1111/j.1095-8649.2004.00601.x

*Drozd, B., Kouril, J., Blaha, M., & Hamackova, J. (2009). Effect of temperature on early life history in weatherfish, *Misgurnus fossilis* (L. 1758). *Knowledge and Management of Aquatic Ecosystems*, *392*(4), 17. https://doi.org/10.1051/kmae/2009010

Dulvy, N. K., Sadovy, Y., & Reynolds, J. D. (2003). Extinction vulnerability in marine populations. *Fish and Fisheries*, *4*(1), 25–64. https://doi.org/10.1046/j.1467-2979.2003.00105.x

Franch-Gras, L., García-Roger, E. M., Serra, M., & Carmona, M. J. (2017). Adaptation in response to environmental unpredictability. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1868), 20170427. https://doi.org/10.1098/rspb.2017.0427

Froese, R., & Pauly, D. (2000). *FishBase 2000: Concepts, design and data sources*. Retrieved from http://www.fishbase.org

*Galloway, T. F., Kjørsvik, E., & Kryvi, H. (1998). Effect of temperature on viability and axial muscle development in embryos and yolk sac larvae of the Northeast Arctic cod (*Gadus morhua*). *Marine Biology*, *132*(4), 559–567. https://doi.org/10.1007/s002270050421

*Georga, I., & Koumoundouros, G. (2010). Thermally induced plasticity of body shape in adult zebrafish *Danio rerio* (Hamilton, 1822). *Journal of Morphology*, *271*(11), 1319–1327. https://doi.org/10.1002/jmor.10874

Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, *21*(3), 394–407. https://doi.org/10.1111/j.1365-2435.2007.01283.x

Grenchik, M. K., Donelson, J. M., & Munday, P. L. (2013). Evidence for developmental thermal acclimation in the damselfish, *Pomacentrus moluccensis*. *Coral Reefs*, *32*(1), 85–90. https://doi.org/10.1007/s00338-012-0949-1

Grether, G. F. (2005). Environmental change, phenotypic plasticity, and genetic compensation. *The American Naturalist*, *59*(7), 1570–1578. https://doi.org/10.1111/j.0014-3820.2005.tb01806.x

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, *33*(2), 1–22. https://doi.org/10.18637/jss.v033.i02

*Hall, T. E., & Johnston, I. A. (2003). Temperature and developmental plasticity during embryogenesis in the Atlantic cod *Gadus morhua* L. *Marine Biology*, *142*(5), 833–840. https://doi.org/10.1007/s00227-003-1030-y

Hansen, T. F., Carter, A. J. R., & Pélabon, C. (2006). On adaptive accuracy and precision in natural populations. *The American Naturalist*, *168*(2), 168–181. https://doi.org/10.1086/505768

Hedges, L. V., Gurevitch, J., & Curtis, P. S. (1999). The meta-analysis of response ratios in experimental ecology. *Ecology*, *80*(4), 1150–1156. https://doi.org/10.1890/0012-9658(1999)080[1150:TMAORR]2.0.CO;2

*Hernández-Rubio, M. C., & Figueroa-Lucero, G. (2013). Effects of temperature and salinity during the embryonic period of *Chirostoom humboldtainum* and *Chirostoma riojai* (Atherinopsidae) until hatching. *Hidrobiologica*, *23*(3), 365–373. https://doi.org/10.1139/f71-101

Hinchliff, C. E., Smith, S. A., Allman, J. F., Burleigh, J. G., Chaudhary, R., Coghill, L. M., … Cranston, K. A. (2015). Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences*, *112*(41), 12764–12769. https://doi.org/10.1073/pnas.1423041112

Jennions, M. D., Lortie, C. J., Rosenberg, M. S., & Rothstein, H. R. (2013). Publication and related biases. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *Handbook of meta-analysis in ecology and evolution* (pp. 207–236). Princeton, NJ: Princeton University Press.

Jennions, M. D., & Møller, A. P. (2002). Relationships fade with time: A meta-analysis of temporal trends in publication in ecology and evolution. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1486), 43–48. https://doi.org/10.1098/rspb.2001.1832

*Jeuthe, H., Brännäs, E., & Nilsson, J. (2015). Effects of variable egg incubation temperatures on the embryonic development in Arctic charr *Salvelinus alpinus*. *Aquaculture Research*, *47*(12), 3753–3764. https://doi.org/10.1111/are.12825

*Johnston, I. A., Cole, N. J., Abercromby, M., & Vieira, V. L. A. (1998). Embryonic temperature modulates muscle growth characteristics in larval and juvenile herring. *The Journal of Experimental Biology*, *201*(5), 623–646.

*Jones, A. C., Lim, D., Wayne-Thompson, J. J., Urbina, N., Puentedura, G., Hillyard, S., & van Breukelen, F. (2016). Oxygen consumption is limited at an ecologically relevant rearing temperature in pupfish eggs. *Journal of Experimental Zoology. Part A, Ecological Genetics and Physiology*, *325*(8), 539–547. https://doi.org/10.1002/jez.2048

Jonsson, B., & Jonsson, N. (2014). Early environment influences later performance in fishes. *Journal of Fish Biology*, *85*(2), 151–188. https://doi.org/10.1111/jfb.12432

*Kamler, E., Keckei, H., & Bauer-Nemeschkal, E. (1998). Temperature-induced changes of survival, development and yolk partitioning in *Chondrostoma nasus*. *Journal of Fish Biology*, *53*(3), 658–682. https://doi.org/10.1111/j.1095-8649.1998.tb01009.x

Kim, S.-Y., Metcalfe, N. B., da Silva, A., & Velando, A. (2017). Thermal conditions during early life influence seasonal maternal strategies in the three-spined stickleback. *BMC Ecology*, *17*(1), https://doi.org/10.1186/s12898-017-0144-x

*Korwin-Kossakowski, M. (2018). The influence of temperature during the embryonic period on larval growth and development in carp, *Cyprinus carpio* L., and grass carp, *Ctenopharyngodon idella* (Val.): Theoretical and practical aspects. *Archives of Polish Fisheries*, *16*(3), 561–587. https://doi.org/10.2478/s10086-008-0020-6

*Koumoundouros, G., Divanach, P., Anezaki, L., & Kentouri, M. (2001). Temperature-induced ontogenetic plasticity in sea bass (*Dicentrarchus labrax*). *Marine Biology*, *139*(5), 817–830. https://doi.org/10.1007/s002270100635

*Kucharczyk, D., Luczynski, M., Kujawa, R., & Czerkies, P. (1997). Effect of temperature on embryonic and larval development of bream. *Aquatic Sciences*, *59*(3), 214–224. https://doi.org/10.1007/s000270050009

*Lõhmus, M., Sundström, L. F., Björklund, M., & Devlin, R. H. (2010). Genotype-temperature interaction in the regulation of development, growth, and morphometrics in wild-type, and growth hormone transgenic coho salmon. *PLoS ONE*, *5*(4), e9980–e10011. https://doi.org/10.1371/journal.pone.0009980

*MacGregor, R. B., & MacCrimmon, H. R. (1977). Evidence of genetic and environmental influences on meristic variation in the rainbow trout, *Salmo gairdneri* Richardson. *Environmental Biology of Fishes*, *2*(1), 25–33. https://doi.org/10.1007/BF00001413

*Mari, L., Garaud, L., Evanno, G., & Lasne, E. (2016). Higher temperature exacerbates the impact of sediments on embryo performances in a salmonid. *Biology Letters*, *12*(12), 20160745. https://doi.org/10.1098/rsbl.2016.0745

*Martell, D. J., Kieffer, J. D., & Trippel, E. A. (2005). Effects of temperature during early life history on embryonic and larval development and growth in haddock. *Journal of Fish Biology*, *66*(6), 1558–1575. https://doi.org/10.1111/j.0022-1112.2005.00699.x

*Matschak, T. W., Hopcroft, T., Mason, P. S., Crook, A. R., & Stickland, N. C. (1998). Temperature and oxygen tension influence the development of muscle cellularity in embryonic rainbow trout. *Journal of Fish Biology*, *53*(3), 581–590. https://doi.org/10.1006/jfbi.1998.0726

*McCarthy, I., Moksness, E., & Pavlov, D. A. (1998). The effects of temperature on growth rate and growth efficiency of juvenile common wolffish. *Aquaculture International*, *6*(3), 207–218. https://doi.org/10.1023/A:1009202710566

McGuigan, K., & Sgrò, C. M. (2009). Evolutionary consequences of cryptic genetic variation. *Trends in Ecology & Evolution*, *24*(6), 305–311. https://doi.org/10.1016/j.tree.2009.02.001

*McGurk, M. D. (1984). Effects of delayed feeding and temperature on the age of irreversible starvation and on the rates of growth and mortality of Pacific herring larvae. *Marine Biology*, *84*(1), 13–26. https://doi.org/10.1007/BF00394522

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*(12), 1476–1481. https://doi.org/10.1111/2041-210X.12593

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G., The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

*Morehead, D. T., & Hart, P. R. (2003). Effect of temperature on hatching success and size of striped trumpeter (*Latris lineata*) larvae. *Aquaculture*, *220*(1–4), 595–606. https://doi.org/10.1016/S0044-8486(02)00636-1

Müller, J., Bässler, C., Essbauer, S., Schex, S., Müller, D. W. H., Opgenoorth, L., & Brandl, R. (2014). Relative heart size in two rodent species increases with elevation: Reviving Hesse's rule. *Journal of Biogeography*, *41*(12), 2211–2220. https://doi.org/10.1111/jbi.12365

Munday, P. L., Donelson, J. M., & Domingos, J. A. (2017). Potential for adaptation to climate change in a coral reef fish. *Global Change Biology*, *23*(1), 307–317. https://doi.org/10.1111/gcb.13419

Munday, P. L., Kingsford, M. J., O'Callaghan, M., & Donelson, J. M. (2008). Elevated temperature restricts growth potential of the coral reef fish *Acanthochromis polyacanthus*. *Coral Reefs*, *27*(4), 927–931. https://doi.org/10.1007/s00338-008-0393-4

Nakagawa, S., Kar, F., O'Dea, R. E., Pick, J. L., & Lagisz, M. (2017). Divide and conquer? Size adjustment with allometry and intermediate outcomes. *BMC Biology*, *15*(1), 107. https://doi.org/10.1186/s12915-017-0448-5

Nakagawa, S., Poulin, R., Mengersen, K., Reinhold, K., Engqvist, L., Lagisz, M., & Senior, A. M. (2015). Meta-analysis of variation: Ecological and evolutionary applications and beyond. *Methods in Ecology and Evolution*, *6*(2), 143–152. https://doi.org/10.1111/2041-210X.12309

Nakagawa, S., & Santos, E. S. A. (2012). Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology*, *26*(5), 1253–1274. https://doi.org/10.1007/s10682-012-9555-5

Nettle, D., & Bateson, M. (2015). Adaptive developmental plasticity: What is it, how can we recognize it and when can it evolve? *Proceedings of the Royal Society B: Biological Sciences*, *282*(1812), 23–31. https://doi.org/10.1098/rspb.2015.1005

Neuheimer, A. B., Thresher, R. E., Lyle, J. M., & Semmens, J. M. (2011). Tolerance limit for fish growth exceeded by warming waters. *Nature Climate Change*, *1*(5), 110–113. https://doi.org/10.1038/nclimate1084

*Nissling, A. (2004). Effects of temperature on egg and larval survival of cod (*Gadus morhua*) and sprat (*Sprattus sprattus*) in the Baltic Sea – Implications for stock development. *Hydrobiologia*, *514*(1–3), 115–123. https://doi.org/10.1023/B:hydr.0000018212.88053.aa

Noble, D. W. A., Lagisz, M., O'Dea, R. E., & Nakagawa, S. (2017). Nonindependence and sensitivity analyses in ecological and evolutionary meta-analyses. *Molecular Ecology*, *26*(9), 2410–2425. https://doi.org/10.1111/mec.14031

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

O'Dea, R. E., Lagisz, M., Hendry, A. P., & Nakagawa, S. (2018a). The effect of rearing temperature on phenotypic mean and variance of fishes. *Open Science Framework*. Retrieved from https://osf.io/8ymh9

O'Dea, R. E., Lagisz, M., Hendry, A. P., & Nakagawa, S. (2018b). Developmental temperature affects phenotypic means and variability: A meta-analysis of fish data. *Open Science Framework*. Retrieved from https://osf.io/e2tyw/

O'Dea, R. E., Noble, D. W. A., Johnson, S. L., Hesselson, D., & Nakagawa, S. (2016). The role of non-genetic inheritance in evolutionary rescue: Epigenetic buffering, heritable bet hedging and epigenetic traps. *Environmental Epigenetics*, *2*(1), dvv014. https://doi.org/10.1093/eep/dvv014

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210. https://doi.org/10.1186/s13643-016-0384-4

Paaby, A. B., & Rockman, M. V. (2014). Cryptic genetic variation: Evolution's hidden substrate. *Nature Reviews Genetics*, *15*(4), 247–258. https://doi.org/10.1038/nrg3688

Pal, C., & Miklos, I. (1999). Epigenetic inheritance, genetic assimilation and speciation. *Journal of Theoretical Biology*, *200*(1), 19–37. https://doi.org/10.1006/jtbi.1999.0974

*Pan, T.-C.-F., & Hunt von Herbing, I. (2017). Metabolic plasticity in development: Synergistic responses to high temperature and hypoxia in zebrafish, *Danio rerio*. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, *327*(4), 189–199. https://doi.org/10.1002/jez.2092

Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*(3), 526–528. https://doi.org/10.1093/bioinformatics/bty633

Parker, T. H., Forstmeier, W., Koricheva, J., Fidler, F., Hadfield, J. D., Chee, Y. E., … Nakagawa, S. (2016). Transparency in ecology and evolution: Real problems, real solutions. *Trends in Ecology & Evolution*, *31*(9), 711–719. https://doi.org/10.1016/j.tree.2016.07.002

*Pavlidis, M., Koumoundouros, G., Sterioti, A., Somarakis, S., Divanach, P., & Kentouri, M. (2000). Evidence of temperature-dependent sex determination in the European sea bass (*Dicentrarchus labrax* L.). *Journal of Experimental Zoology*, *287*(3), 225–232. https://doi.org/10.1002/1097-010X(20000801)287:3<225::AID-JEZ4>3.0.CO;2-D

*Pavlov, D. A., & Moksness, E. (1997). Development of the axial skeleton in wolffish, *Anarhichas lupus* (Pisces, Anarhichadidae), at different temperatures. *Environmental Biology of Fishes*, *49*(4), 401–416. https://doi.org/10.1023/A:1007352802352

*Peck, M. A., & Buckley, L. J. (2008). Measurements of larval Atlantic cod (*Gadus morhua*) routine metabolism: Temperature effects, diel differences and individual-based modeling. *Journal of Applied Ichthyology*, *24*(2), 144–149. https://doi.org/10.1111/j.1439-0426.2007.01004.x

*Perrichon, P., Pasparakis, C., Mager, E. M., Stieglitz, J. D., Benetti, D. D., Grosell, M., & Burggren, W. W. (2017). Morphology and cardiac physiology are differentially affected by temperature in developing larvae of the marine fish mahi-mahi (*Coryphaena hippurus*). *Biology Open*, *6*(6), 800–809. https://doi.org/10.1242/bio.025692

*Peterson, R. H., Martin-Robichaud, D. J., & Berge, O. (1996). Influence of temperature and salinity on length and yolk utilization of striped bass larvae. *Aquaculture International*, *4*(2), 89–103. https://doi.org/10.1007/BF00140591

Pick, J. L., Nakagawa, S., & Noble, D. W. A. (2018). Reproducible, flexible and high-throughput data extraction from primary literature: The metaDigitise r package. *Methods in Ecology and Evolution*, *10*, 426–431. https://doi.org/10.1111/2041-210X.13118

Piferrer, F., Ribas, L., & Diaz, N. (2012). Genomic approaches to study genetic and environmental influences on fish sex determination and differentiation. *Marine Biotechnology*, *14*(5), 591–604. https://doi.org/10.1007/s10126-012-9445-4

*Politis, S. N., Mazurais, D., Servili, A., Zambonino-Infante, J.-L., Miest, J. J., Sørensen, S. R., … Butts, I. A. E. (2017). Temperature effects on gene expression and morphological development of European eel, *Anguilla anguilla* larvae. *PLoS ONE*, *12*(8), e0182726–e182823. https://doi.org/10.1371/journal.pone.0182726

Pörtner, H.-O., & Lannig, G. (2009). Chapter 4 oxygen and capacity limited thermal tolerance. *Fish Physiology*, *27*, 143–191. https://doi.org/10.1016/S1546-5098(08)00004-6

R Development Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

*Raine, J. C., & Leatherland, J. F. (1999). Ontogeny of thyroid tissue and tissue thyroid hormone clearance in rainbow trout embryos reared at two temperatures. *Fish Physiology and Biochemistry*, *20*(3), 209–217. https://doi.org/10.1023/A:1007775807438

Raudenbush, S. W., & Bryk, A. S. (1987). Examining correlates of diversity. *Journal of Educational Statistics*, *12*(3), 241–269. https://doi.org/10.3102/10769986012003241

*Réalis-Doyelle, E., Pasquet, A., Fontaine, P., & Teletchea, F. (2017). How climate change may affect the early life stages of one of the most common freshwater fish species worldwide: The common carp (*Cyprinus carpio*). *Hydrobiologia*, *805*(1), 365–375. https://doi.org/10.1007/s10750-017-3324-y

Robertson, B. A., Rehage, J. S., & Sih, A. (2013). Ecological novelty and the emergence of evolutionary traps. *Trends in Ecology & Evolution*, *28*(9), 552–560. https://doi.org/10.1016/j.tree.2013.04.004

*Savoie, A., Le François, N. R., Cahu, C., Blier, P. U., & Andreassen, I. (2006). Do protein hydrolysates improve survival and growth of newly-hatched spotted wolffish (*Anarhichas minor*), a non-metamorphic aquaculture fish species? *Aquaculture*, *261*(2), 782–788. https://doi.org/10.1016/j.aquaculture.2006.08.047

*Schnurr, M. E., Yin, Y., & Scott, G. R. (2014). Temperature during embryonic development has persistent effects on metabolic enzymes in the muscle of zebrafish. *Journal of Experimental Biology*, *217*(8), 1370–1380. https://doi.org/10.1242/jeb.094037

*Schönweger, G., Schwerte, T., & Pelster, B. (2000). Temperature-dependent development of cardiac activity in unrestrained larvae of the minnow *Phoxinus phoxinus*. *American Journal of Physiology, Regulatory, Integrative and Comparative Physiology*, *279*(5), R1634–R1640. https://doi.org/10.1152/ajpregu.2000.279.5.R1634

Senior, A. M., Grueber, C. E., Kamiya, T., Lagisz, M., O'Dwyer, K., Santos, E. S. A., & Nakagawa, S. (2016). Heterogeneity in ecological and evolutionary meta-analyses: Its magnitude and implications. *Ecology*, *97*(12), 3293–3299. https://doi.org/10.1002/ecy.1591

*Sfakianakis, D. G., Leris, I., & Kentouri, M. (2010). Effect of developmental temperature on swimming performance of zebrafish (*Danio rerio*) juveniles. *Environmental Biology of Fishes*, *90*(4), 421–427. https://doi.org/10.1007/s10641-010-9751-5

*Sfakianakis, D. G., Leris, I., Laggis, A., & Kentouri, M. (2011). The effect of rearing temperature on body shape and meristic characters in zebrafish (*Danio rerio*) juveniles. *Environmental Biology of Fishes*, *92*(2), 197–205. https://doi.org/10.1007/s10641-011-9833-z

*Sfakianakis, D. G., Papadakis, I. E., Papadaki, M., Sigelaki, I., & Mylonas, C. C. (2013). Influence of rearing temperature during early life on sex differentiation, haemal lordosis and subsequent growth during the whole production cycle in European sea bass *Dicentrarchus labrax*. *Aquaculture*, *412–413*, 179–185. https://doi.org/10.1016/j.aquaculture.2013.07.033

*Silva, P., Valente, L. M. P., Olmedo, M., Álvarez-Blázquez, B., Galante, M. H., Monteiro, R. A. F., & Rocha, E. (2010). Influence of temperature on muscle fibre hyperplasia and hypertrophy in larvae of blackspot seabream, *Pagellus bogaraveo*. *Aquaculture Research*, *42*(3), 331–340. https://doi.org/10.1111/j.1365-2109.2010.02627.x

Snell-Rood, E. C., Kobiela, M. E., Sikkink, K. L., & Shephard, A. M. (2018). Mechanisms of plastic rescue in novel environments. *Annual Review of Ecology, Evolution, and Systematics*, *49*(1), 331–354. https://doi.org/10.1146/annurev-ecolsys-110617-062622

Starrfelt, J., & Kokko, H. (2012). Bet-hedging – A triple trade-off between means, variances and correlations. *Biological Reviews*, *87*(3), 742–755. https://doi.org/10.1111/j.1469-185X.2012.00225.x

Sumaila, U. R., Cheung, W. W. L., Lam, V. W. Y., Pauly, D., & Herrick, S. (2011). Climate change impacts on the biophysics and economics of world fisheries. *Nature Climate Change*, *1*(9), 449–456. https://doi.org/10.1038/nclimate1301

*Sweet, J. G., & Kinne, O. (1964). The effects of various temperature-salinity combinations on the body form of newly hatched *Cyprinodon macularius* (Teleostei). *Helgoländer Wissenschaftliche Meeresuntersuchungen*, *11*(2), 49–69. https://doi.org/10.1007/BF01611131

University of Michigan Museum of Zoology. (2018). *Animal diversity web*. Retrieved from https://animaldiversity.org/

*Usher, M. L., Stickland, N. C., & Thorpe, J. E. (1994). Muscle development in Atlantic salmon (*Salmo salar*) embryos and the effect of temperature on muscle cellularity. *Journal of Fish Biology*, *44*(6), 953–964. https://doi.org/10.1111/j.1095-8649.1994.tb01267.x

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Viechtbauer, W. (2018). *I^2 for multilevel and multivariate models*. Retrieved from http://www.metafor-project.org/doku.php/tips:i2_multilevel_multivariate

Villegas-Ríos, D., Moland, E., & Olsen, E. M. (2016). Potential of contemporary evolution to erode fishery benefits from marine reserves. *Fish and Fisheries*, *18*(3), 571–577. https://doi.org/10.1111/faf.12188

*Walsh, W. A., Swanson, C., & Lee, C. S. (1991). Effects of development, temperature and salinity on metabolism in eggs and yolk-sac larvae of milkfish, *Chanos chanos* (Forsskal). *Journal of Fish Biology*, *39*, 115–125. https://doi.org/10.1111/j.1095-8649.1991.tb04346.x

*Wen, W., Huang, X., Chen, Q., Feng, L., & Wei, L. (2013). Temperature effects on early development and biochemical dynamics of a marine fish, *Inimicus japonicus*. *Journal of Experimental Marine Biology and Ecology*, *442*(C), 22–29. https://doi.org/10.1016/j.jembe.2013.01.025

*Whitney, C. K., Hinch, S. G., & Patterson, D. A. (2014). Population origin and water temperature affect development timing in embryonic sockeye salmon. *Transactions of the American Fisheries Society*, *143*(5), 1316–1329. https://doi.org/10.1080/00028487.2014.935481

Wood, C. W., & Brodie, E. D. I. (2015). Environmental effects on the structure of the G-matrix. *Evolution*, *69*(11), 2927–2940. https://doi.org/10.1111/evo.12795

*Zummo, G., Farina, F., Tota, B., & Johnston, I. A. (1996). Influence of temperature on the development of the heart ventricle in herring (*Clupea harengus*) larvae. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, *275*(2–3), 196–203.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** O'Dea RE, Lagisz M, Hendry AP, Nakagawa S. Developmental temperature affects phenotypic means and variability: A meta-analysis of fish data. *Fish Fish*. 2019;20:1005–1022. https://doi.org/10.1111/faf.12394