1 **Global environmental drivers of marine fish egg size**

2 Diego R. Barneche[1,§,*], Scott C. Burgess[2], Dustin J. Marshall[1]

3 [1]Centre for Geometric Biology / School of Biological Sciences, Monash University, Clayton Vic 3800, Australia

4 [2]Department of Biological Science, Florida State University, Tallahassee, FL, 32306, USA

5 [§]Present address: School of Life and Environmental Sciences, The University of Sydney, Sydney NSW 2006,
6 Australia

7 *Running title:* Biogeography of marine fish egg size

8 *Keywords:* Geometric Biology, predictability, life history, trait biogeography, maternal effects,

9 ocean warming

10 **Type of article:** Research paper

11 **Word Count:** 4,448 (Main text); 297 (Abstract)

12 **Number of references:** 51

13 **Number of figures:** 2

14 **Correspondence should be sent to:** Diego R. Barneche, Centre for Geometric Biology /

15 School of Biological Sciences, Monash University, Clayton, Vic. 3800, Australia (Tel: +61

16 (3) 9905 5100; Email: barnechedr@gmail.com).

Abstract

Aim: To test long standing theory on the role of environmental conditions (both mean and predictability) in shaping global patterns in the egg sizes of marine fishes.

Location: Global (50˚ S to 50˚ N).

Time period: 1880 to 2015.

Major taxa studied: Marine fish.

Methods: We compiled the largest geo-located dataset of marine fish egg size (diameter) to date ($n = 1,078$ observations; 192 studies; 288 species; 242 localities). We decomposed sea surface temperature (SST) and chlorophyll-a time-series into mean and predictability (seasonality and colour of environmental noise – i.e. how predictable the environment is between consecutive time steps), and used these as predictors of egg size in a Bayesian phylogenetic hierarchical model. We test four specific hypotheses based on the classic discussion by Rass (1941), as well as contemporary life-history theory, and the conceptual model of Winemiller & Rose (1992).

Results: Both environmental mean and predictability correlated with egg size. Our parsimonious model indicated that egg size decreases by *c*. 2.0-fold moving from 1˚C to 30˚C. Environments that were more seasonal with respect to temperature were associated with larger eggs. Increasing mean chlorophyll-a, from 0.1 to 1 mg m^{-3}, was associated with a *c*. 1.3-fold decrease in egg size. Lower chlorophyll-a seasonality and reddened noise were also associated with larger egg sizes – aseasonal but more temporally autocorrelated resource regimes favoured larger eggs.

Main conclusions: Our findings support results from Rass (1941) and some predictions from Winemiller & Rose (1992). The effects of environmental means and predictability on marine fish egg size are largely consistent with those observed in marine invertebrates with feeding larvae, suggesting important commonalities in how ectotherm egg size responds to environmental change. Our results further suggest that anthropogenically-mediated changes in the environment will have profound effects on the distribution of marine life histories.

Introduction

Life-history traits vary systematically along environmental gradients, particularly with latitude. For example, latitude covaries with adult body size (e.g. mammals, birds, fish, insects; Bergamnn's rule; Meiri, 2011), offspring size (Rass's rule; Rass, 1941; Marshall, 1953; Marshall *et al.*, 2012), developmental mode (Thorson's rule; Thorson, 1950; Marshall *et al.*, 2012), and dispersal capacity (O'Connor *et al.*, 2007). For years, biogeographers and life-history theoreticians have assumed that the latitudinal gradients in life history are driven by differences in average conditions (e.g. mean temperature), but means are not the only moments of the distribution that vary in space and influence life-history evolution (e.g. Vasseur *et al.*, 2014; Marshall & Burgess, 2015).

Although ecologists have traditionally emphasised the role of environmental means across latitudes as drivers of latitudinal variation, classic biogeography treatises recognised that other components of environmental factors also vary systematically with latitude, particularly environmental predictability (Cohen, 1966; Slobodkin & Sanders, 1969; Brown, 1973; Crump, 1981). Environments can be more or less predictable in multiple ways. First, environments can be predictable if they remain relatively constant over time – temperature in much of the deep sea is a good example of such conditions. Second, environments can be predictable if they change in very regular ways – for example, highly seasonal environments, while variable, are also predictable in the timing of events – temperate latitudes are likely to be warmer and drier in summer relative to winter for example. Finally, environments can also be predictable if conditions at one time are strongly correlated with conditions at some point into the future – i.e. there is a strong temporal autocorrelation in local conditions. For example, if an environment is highly autocorrelated, then conditions now are a good indicators of conditions later. This type of predictability occurs when the colour of environmental noise is *reddened* (Vasseur & Yodzis, 2004; Burgess & Marshall, 2014). Importantly, all of these types of predictability (constancy, seasonality and autocorrelation) can differ across latitudes, and each type of predictability will affect life-history evolution differently (Travis, 2001; Burgess & Marshall, 2014; Marshall & Burgess, 2015; Rubio de Casas *et al.*, 2017).

Theory predicts that environmental predictability should shape evolutionary ecology over

73 and above the effects of environmental means (Parker & Begon, 1986; Travis, 2001). For ex-

74 ample, theory predicts that fecundity will be optimised by producing many small eggs in en-

75 vironments with predictable food regimes (Smith & Fretwell, 1974; Parker & Begon, 1986;

76 Rollinson *et al.*, 2013). Moreover, if environmental differences across space are relatively

77 stable and predictable over time, theory predicts, and evidence suggests, that dispersal will

78 not be favoured (Hastings, 1983; Fronhofer *et al.*, 2014). On the other hand, mothers in un-

79 predictable environments should produce fewer and larger offspring (i.e. a conservative bet-

80 hedging approach Einum & Fleming, 2004). Contradictions are common however, for exam-

81 ple, where populations inhabiting increasingly harsh streams produced more numerous and

82 smaller eggs (Morrongiello *et al.*, 2012). Initial studies of trait biogeography support these

83 theoretical predictions – latitudinal gradients in marine invertebrate life histories are driven

84 strongly by changes in predictability of both food and temperature from the tropics to the

85 poles (Marshall & Burgess, 2015). But overall, biogeographical studies of how environmen-

86 tal predictability shapes life-history variation at large spatial scales remain exceedingly rare

87 (Marshall & Burgess, 2015; Rubio de Casas *et al.*, 2017).

88   In fishes, Rass (1941) and Marshall (1953) first noted that lower latitudes and warmer tem-

89 peratures are often associated with smaller egg sizes. Winemiller & Rose (1992) proposed

90 a conceptual argument that life histories encompass a continuum space among three major

91 strategies, and that the interactions among juvenile survival, adult fecundity, and maturity are

92 shaped not only by the average environmental state, but also by its predictability (see hypothe-

93 ses below). Twenty-five years later, formal tests on how (and if) environmental mean and

94 predictability consistently affect marine fish eggs over broad spatial scales are still lacking,

95 despite a rich theoretical body of literature (Cohen, 1966; Smith & Fretwell, 1974; Parker &

96 Begon, 1986; Winemiller & Rose, 1992; Travis, 2001; Einum & Fleming, 2004).

97   Species-level spawning mode is also known to correlate with egg size. For instance,

98 species that exhibit parental care (i.e. brooders) are known to have larger eggs when compared

99 to many pelagic- and demersal-spawning species (Winemiller & Rose, 1992). Also, demersal

100 spawners have on average larger eggs than pelagic spawners, followed by scatterer spawners

101 (i.e. species who deposit their eggs on the benthos, but have a long pelagic stage when

102 compare to ordinary demersal spawners) (Kasimatis & Riginos, 2016). These correlations

4

between spawning mode and egg size might exhibit a strong phylogenetic effect (e.g. most Tetraodontiformes are scatterers, Syngnathidae and Apogonidae are brooders; Kasimatis & Riginos, 2016). However, it is unknown whether the effects of spawning mode remain substantial after accounting for the effects of environmental variables, and phylogeny on egg size.

Here we present a formal analysis on the role of the environment in shaping marine fish egg size. In doing so, we first compile the most comprehensive dataset of marine fish egg sizes which dates back to 1880, and use it to explicitly test the following hypotheses: egg size responds to changes in average temperature (*H1*), productivity (*H2*), some form of environmental predictability (*H3*), and spawning mode (*H4*). Particularly, within *H1*, we predict that egg size will decrease with increasing temperature. Within *H2*, we predict that egg size will decrease with increasing average food regimes. For *H3*, we derive predictions inspired by the conceptual model of Winemiller & Rose (1992): high seasonality and low average food regimes drive declines in egg size (*Periodic* strategy); white-noised environments (i.e. unpredictable) with low seasonality drive declines in egg size (*Opportunistic* strategy); predictable environments with low average food regime drive egg size to increase (*Equilibrium* strategy). For *H4*, we predict that after correcting for phylogeny, spawning mode will not significantly affect egg size. We apply this hypothesis-testing approach using Bayesian phylogenetic hierarchical models in order to account for shared evolutionary history among species, as well as independent evolutionary changes unique to each species.

Materials and Methods

Egg size data compilation

We began by compiling data from multiple previous compilations that provided references on egg size (diameter in mm) for different species of marine fishes (Fritzsche, 1978; Duarte & Alcaraz, 1989; Elgar, 1990; Moser, 1996; Einum & Fleming, 2002; Hixon *et al.*, 2014; Kasimatis & Riginos, 2016). We used only raw data from original references (data extracted from tables or directly from figures using DataThief software). We also investigated papers cited in these original references, repeating this process exhaustively. Importantly, because

we are primarily interested in the role of the environment, we only collected data from publications that reported the origin of specimen collection for populations directly collected from the wild or that were transferred from the wild to an aquarium facility immediately before the measurements. Eggs that were measured immediately following fertilisation (i.e. no change in size) were also included. For eggs that present ellipsoidal, elongated shapes, we used both diameters (long, l, and short, s) to first calculate the volume, V (mm^3; $V = (4/3) \times pi \times (l/2) \times (s/2)^2$), and then back transformed to diameter of sphere with equivalent volume ($[((V \times 3)/(4 \times pi))^{(1/3)}] \times 2$). Some papers reported mean egg size for each female while others reported means across multiple females, so we are not able to properly consider the coefficient of variation in egg size at the same scale across all data (see Einum & Fleming, 2002). Our dataset includes 1,078 observations from 192 studies between 1880 to 2015, and includes 288 species.

Environmental predictability and data

We here consider two components of environmental predictability (Colwell, 1974; Marshall & Burgess, 2015; *cf.* Rubio de Casas *et al.*, 2017). The first component is *seasonality*, which entails the regularity in the timing and magnitude of fluctuations in the average environmental state over seasons; it is expected to influence the evolution of egg size (Winemiller & Rose, 1992). The second type is the *colour of environmental noise* (Halley & Kunin, 1999; Vasseur & Yodzis, 2004), which is defined by how predictable and similar the environment is between successive time points, or how far into the future the environmental state is likely to stay the same, independent of the mean environmental state (Marshall & Burgess, 2015).

We wrote an R package (*noaaErddap*, available on www.github.com/dbarneche/noaaErddap) to download time-series data from NOAA on chlorophyll-a (mg m^{-3}; 0.08˚ resolution), which is used as proxy for environmental food regime, and sea surface temperature (hereafter SST; ˚C; 0.25˚ resolution), covering a 10-year period from 1997 to 2007. Chlorophyll-a data were recorded every eight days (Δt = 8 days, unevenly distributed) and SST data were recorded every day (Δt = 1 day, evenly distributed). For both SST and chlorophyll-a, we extracted the median environmental value within a 200-km radius around each coordinate in our dataset. Following the approach adopted in Marshall & Burgess (2015), only coordinates that had <

160 13% of the time series missing were used, hence all the data from coordinates above 50° of

161 latitude were excluded for the analysis (i.e. temporal coverage < 87%).

162   Our calculations to decompose predictability into seasonality and colour of environmental

163 noise are identical to those described by Marshall & Burgess (2015). We first removed linear

164 trends by extracting the residuals from a linear regression model fitted to the raw time series.

165 For each coordinate, seasonality was estimated as the fraction of the total variance that is due

166 to predictable seasonal periodicities, $a/(a+b)$, where a is the variance of the seasonal trend,

167 and b is the variance of the residual time series (i.e. the time series after the seasonal trend was

168 removed). The seasonal trend was estimated by binning the time-series data into monthly in-

169 tervals, averaging each month across the 10 years, then re-creating a seasonal time-series data

170 set on the same time scale of the original data using a linear interpolation between the monthly

171 midpoints. The colour of environmental noise was estimated using the $1/f^{\theta}$ family of noise

172 models (Halley & Kunin, 1999; Vasseur & Yodzis, 2004), where white noise ($\theta = 0$) occurs

173 when there is no correlation between one measurement and the next, while for reddened noise

174 ($\theta > 0$), there is some correlation between measurements separated by a finite timescale. To

175 do so, we first calculated a residual time series by subtracting the corresponding seasonal

176 value from each data point in the time series. The spectral density (i.e. variance in the resid-

177 ual time series) was assumed to scale with frequency, f, according to an inverse power law,

178 $1/f^{\theta}$. The spectral exponent θ was estimated as the negative slope of the linear regression of

179 $\log_e$ spectral density as a function of $\log_e$ frequency. Spectral density was estimated in R using

180 the *spectrum* function form the *stats* R package v. 3.4.3 for the evenly-distributed SST time

181 series, and the Lomb-Scargle function *lsp* from the *lomb* R package v. 1.0 for the unevenly-

182 distributed chlorophyll-a time series (Glynn *et al.*, 2006). Spectral densities and subsequent

183 θ's were calculated between the frequencies of $2/n\Delta t$ and $1/2\Delta t$ (n = number of observations

184 in the time series), which translates to periods of 16 days to 5.05 years for chlorophyll-a, and

185 2 days to 5.05 years for SST. Both seasonality and colour of environmental noise metrics are

186 implemented in an R package (*envPred*, available on www.github.com/dbarneche/envPred).

Phylogeny and spawning mode

Phylogenetic relatedness might influence broad-scale variation in fish life-history traits (Kasimatis & Riginos, 2016). From an evolutionary perspective, closely related species might have a higher likelihood of sharing some ancestral-state trait (Pagel, 1999), such as egg-type (pelagic, demersal, brooded). At the same time, species may present unique variations in traits that are independent of phylogenetic non-independence. From a technical perspective, both of these unmeasured biological factors (i.e. species uniqueness and non-independence), as well as the possibility for different species composition to underlie differences in egg size across latitudes, likely contribute to variance in a particular life-history trait (Hadfield & Nakagawa, 2010) and, consequently, it is necessary to account for these possible effects. To do so, we created a tree from the Open Tree of Life (OTL) using the *rotl* R package v. 3.0.3 (Michonneau *et al.*, 2016) in order to test for significant phylogenetic heritability in our models (Housworth *et al.*, 2004). We first downloaded the full Actinopterygii tree from OTL (n = 38,939 tips) and then added species from our dataset that were missing in the tree: the Pomacentridae species relationship followed a recent consensus topology (Frédérich *et al.*, 2013) and we inserted them as a sister group to the family Labridae; *Heterostichus rostratus* (family Clinidae), and *Sphyraena argentea* (family Sphyraenidae) were respectively inserted right next to the families Blenniidae and Carangidae (Betancur-R. *et al.*, 2013). This tree was pruned to retain focal species only, and then used to derive a variance-covariance matrix based on Brownian evolution. The tree included a total of 30 polytomies. Branch lengths are unknown for the phylogeny, so the arbitrary method of Grafen (1989) was applied, whereby branch lengths are set to a length equal to the number of descendant tips minus one.

We compiled spawning mode for all 288 species in our dataset. We classified species according to 5 categories: Demersal: n = 188 obs., 65 spp; Mouth brooder: n = 15 obs., 13 spp; Pelagic: n = 836 obs., 180 spp; Pouch brooder: n = 19 obs., 15 spp; Scatterer: n = 20 obs., 15 spp. Information on spawning modes were obtained from the literature (e.g. Kasimatis & Riginos, 2016), www.FishBase.org, and by consulting with experts. We were able to reliably reference the spawning mode of 262 species (91 %). The information on spawning mode for the remaining 26 species was approximated based on other species within the same family. Removing these species had no effect in our results (Table S1). The spawning mode in-

217 formation and associated references can be directly downloaded from our GitHub repository

218 (https://github.com/dbarneche/fishEggSize/blob/master/data/spawningMode.csv).

Testing hypotheses H1–H4

220 Before running our statistical analysis, we first checked for any systematic correlation among

221 our predictor variables (Table 1). Temperature mean and its predictability components were

222 strongly correlated (i.e. ~ 0.7; Dormann *et al.*, 2013) and, given this caveat, we interpret our

223 temperature model coefficients as being non-independent. We note however that the colour of

224 temperature noise was not a significant variable and therefore dropped from our parsimonious

225 model (see Results section).

226   The full model was constructed using all the original six variables with the following struc-

227 ture:

$$\ln E = \beta X + \gamma_{spp} + \gamma_{phy} + \varepsilon, \tag{1}$$

228 where $\ln E$ represents log-transformed egg size, β is a vector of 7 fixed-effect coefficients,

229 with 1 intercept and 6 slopes each corresponding to an environmental variable (i.e. both SST

230 and chlorophyll-a averages, seasonalities, and colours of environmental noise) in the model

231 matrix X. γ_{spp} and γ_{phy} are respectively vectors of random-effect coefficients that account for

232 residual intercept deviations attributable to species uniqueness and patterns of relatedness as

233 described by the phylogeny, and ε is the model unexplained residual variation.

234   Fixed effects were assigned weakly informative priors following a Gaussian distribution.

235 Standard deviations ($\sigma_{\gamma_{spp}}$, $\sigma_{\gamma_{phy}}$ and σ_{ε}) were assigned more informative priors following

236 a Student-*t* distribution to speed-up model convergence (notice, though, that the full model

237 was first fitted with weakly informative priors in order to obtain estimates for the more in-

238 formative priors). We account for the phylogenetic non-independence among species follow-

239 ing the method of Hadfield & Nakagawa (2010), where the non-independence random-effect

240 coefficients (γ_{phy}) are distributed following a multivariate normal distribution with means

241 of zero (such that γ_{phy} represents actual deviations from the model intercept) and a single

242 standard deviation ($\sigma_{\gamma_{phy}}$) which is weighted by the variance-covariance matrix, A, obtained

243 from the tips of the phylogeny (i.e. $\sigma_{\gamma_{phy}} \times A$, with A being obtained using the R package

244 *ape* version 5.0; Paradis *et al.*, 2004). The phylogenetic heritability (equivalent to Pagel's

245 λ; Pagel, 1999; Hadfield & Nakagawa, 2010), was estimated as the proportion of total vari-

246 ance, conditioned on the fixed effects, attributable to the random effect of phylogeny (i.e.

247 $\sigma^2_{\gamma_{phy}}/(\sigma^2_{\gamma_{spp}} + \sigma^2_{\gamma_{phy}} + \sigma^2_{\varepsilon})$). The posterior distributions of model parameters were estimated

248 using Markov chain Monte Carlo (MCMC) methods with NUTS sampler using the R pack-

249 age *brms* version 2.1.0 (Bürkner, 2017) by constructing three chains of 15,000 steps, including

250 7,500-step warm-up periods, so a total of 22,500 steps were retained to estimate posterior dis-

251 tributions (i.e. (15,000 - 7,500) $\times$ 3 = 22,500).

252   We performed a model selection procedure by first dropping any non-significant slope

253 (i.e. those whose 95% posterior credible intervals overlap zero) and then testing whether this

254 simpler model was significantly better than the original full model. After this procedure, we

255 also tested whether the best of the two models was significantly improved by adding egg

256 spawning mode. Model comparisons were done using leave-one-out cross-validation (LOO),

257 which, similarly to widely applicable information criterion (WAIC), is a fully Bayesian model

258 selection procedure for estimating pointwise out-of-sample prediction accuracy (Vehtari *et*

259 *al.*, 2016). We calculated the expected log pointwise predictive density ($\widehat{elpd}_{loo}$) using the

260 log-likelihood evaluated at the posterior simulations of the parameter values (Vehtari *et al.*,

261 2016). We calculated *p*-values for the pairwise differences in $\widehat{elpd}_{loo}$ ($\Delta\widehat{elpd}_{loo}$) using stan-

262 dard errors (s.e.) and a normal probability density function. This method of calculating s.e.'s

263 is reliable for data sets with many observations (n = 1,078 in our analysis) because the dis-

264 tribution of $\Delta\widehat{elpd}_{loo}$ is well approximated by a Gaussian distribution (Vehtari *et al.*, 2016).

265 Results were similar using K-fold cross-validation (K = 10). Model comparison was imple-

266 mented using the *loo* R package (Vehtari *et al.*, 2016) version 1.1.0. All analyses, tables and

267 figures are fully reproducible in R. Data and code are available via the GitHub repository

268 (https://github.com/dbarneche/fishEggSize).

269 **Results**

270 Fish egg size varied in diameter from 0.24 mm to 6.5 mm, spanning 1.4 orders of magnitude

271 in egg size variation. There was a significant relationship between latitude and egg size ($p <$

0.0001), although only 3.5% of variation in egg size is explained by latitude alone (Fig. 1). For every 10 degrees of latitude moving poleward, egg size increases by 7% on average.

Our parsimonious model contains all original variables except the colour of temperature noise; the full model had a lower average predictive accuracy over the parsimonious model ($\Delta \widehat{elpd}_{loo}$ = -1.6; p = 0.399). Although adding spawning mode to this model improved the average predictive accuracy ($\Delta \widehat{elpd}_{loo}$ = 6.1), it was not significant (p = 0.136), and we therefore retained the simpler model as the parsimonious model (i.e. rejected hypothesis *H4*; Table 2). After accounting for the fixed effects, the majority of the variance observed in the residuals is explained by substantial heritability among species (95.5%; 95% CI: 91.6% – 97.7%), suggesting a substantial effect of phylogeny, and a minor contribution of non-heritable variation within species (2.4%; 95% CI: 1.1% – 5.0%).

In agreement with our prediction of hypothesis *H1*, the parsimonious model indicated that egg size decreases systematically with increasing mean temperature (Fig. 2a). Particularly, the mean temperature slope (β_{SST} = -0.02; 95% CI: -0.03 – -0.01) suggests that egg size will decrease by *c*. 2.0-fold moving from 1°C to 30°C after accounting for the effects of environmental predictability and mean food regime. Environments that were more seasonal with respect to temperature had lower mean temperatures, and were therefore associated with larger egg sizes ($\beta_{SeasonSST}$ = 0.82; 95% CI: 0.47 – 1.19), which is opposite to our prediction of *H3*.

Food regime (indexed by mean chlorophyll-a) showed a negative relationship with egg size, in agreement with our prediction of hypothesis *H2* (β_{Chl} = -0.27; 95% CI: -0.37 – -0.16; Fig. 2b). Increasing mean chlorophyll-a from 0.1 to 1 mg m^{-3} was associated with a *c*. 1.3-fold decrease in egg size after holding all the remaining variables constant. Chlorophyll-a seasonality had a stronger effect than that observed for mean chlorophyll-a ($\beta_{SeasonChl}$ = -0.69; 95% CI: -0.87 – -0.51), in agreement with our prediction of *H3*: higher seasonality was associated with smaller egg sizes. Additionally, increasing the colour of chlorophyll-a noise (i.e. more temporally-autocorrelated food regime) was associated with a larger egg size ($\beta_{\theta Chl}$ = 0.27; 95% CI: 0.14 – 0.40), also in line with our prediction of *H3* (Fig. 3).

Discussion

Here we present a formal and comprehensive test of the role of the environment in shaping marine fish egg size at a global scale, after controlling for potential effects that are attributable to species and phylogenetic relatedness. Importantly, our dataset is composed of geo-located data only, instead of averages for species, thus it is well suited to evaluate the direct effects of the environment on egg size. Our approach builds on decades of empirical (Rass, 1941; Thorson, 1950; Marshall, 1953; Kasimatis & Riginos, 2016), experimental (Kokita, 2003; Shama, 2015), theoretical (Smith & Fretwell, 1974; Parker & Begon, 1986) and conceptual work (Winemiller & Rose, 1992) documenting the effects of the environment on egg size. On the one hand, our parsimonious model largely confirms predictions based on previous findings, indicating that higher temperatures are associated with smaller egg sizes (Fig. 2a; Rass, 1941; Marshall, 1953; Kokita, 2003; Kasimatis & Riginos, 2016). On the other hand, for the first time, our results support previously untested concepts relating the effects of environmental predictability (here measured as seasonality and the colour of environmental noise) on marine fish egg size (Winemiller & Rose, 1992).

Higher mean temperatures were home to fish with smaller egg size, largely supporting Rass's rule (Rass, 1941), and hence our results add to the vast literature reporting the effects of temperature on fish egg size (Rass, 1941; Marshall, 1953; Kokita, 2003; Laptikhovsky, 2006; Shama, 2015; Kasimatis & Riginos, 2016) and effects of temperature on egg size in aquatic organisms more generally (Marshall *et al.*, 2012). Environments that are seasonal with respect to temperature were associated with a larger egg size, but we note we cannot disentangle this effect from those of mean temperature because these two variables were substantially correlated. Importantly, temperature is predicted to have multiple effects, within and among species. Within species, for example, increasing temperatures are associated with an earlier onset of maturity and declines in maximum adult body size (i.e. temperature-size rule; Zuo *et al.*, 2012). Considering that, within a population, smaller females may produce smaller eggs (e.g. Braga Goncalves *et al.*, 2011), increasing ocean warming will likely decrease the reproductive contribution of populations via effects on both female and egg size (Kokita, 2003). Moreover, warming oceans are expected to alter current distributions of some tropical and temperate fish species (e.g. Feary *et al.*, 2014). Thus, changes in temperature will likely

12

329 change the dynamics and quality of larvae provided in the system by altering their richness,

330 abundance, size, and dispersal capacity (Rass, 1941; Marshall, 1953; Kokita, 2003; O'Connor

331 *et al.*, 2007; Hixon *et al.*, 2014).

332 Our results highlight congruences between latitudinal patterns in fish and marine inverte-

333 brate egg size. Particularly, both groups exhibit a general increase in egg size with increasing

334 latitude (Fig. 1; Marshall *et al.*, 2012) despite a substantial variation around this central ten-

335 dency. A portion of this variation is explained by incorporating the additional moments of

336 distributions of environmental factors. Our results indicate that productivity (here measured

337 as chlorophyll-a as a proxy) mean, seasonality and noise colour influence patterns in marine

338 fish egg size, in agreement with the early propositions of Thorson (1950), which were re-

339 cently corroborated by Marshall & Burgess (2015). Interestingly, our results (i.e. directions

340 of effects) are consistent with the recent analysis of environmental predictability on marine

341 invertebrates egg size with feeding larvae (Marshall & Burgess, 2015). Our findings may

342 therefore reflect fundamental commonalities regarding how feeding larvae of ectotherms (both

343 zoo-planktivorous fish and phyto-planktivorous invertebrates) respond to environmental pre-

344 dictability.

345 To the best of our knowledge, our study presents the first formal test of the effects of en-

346 vironmental predictability on marine fish egg size at the global scale. Specifically, we dis-

347 cuss our results in light of the three extreme strategies in the trilateral continuum proposed

348 by Winemiller & Rose (1992). Environments with autocorrelated fluctuations (i.e. reddened

349 noise) with a low mean food regime exhibit eggs that are 86.6% larger than the average, con-

350 sistent with the *Equilibrium strategy* suggested by Winemiller & Rose (1992). Similarly,

351 our results indicate that environments with high seasonality and low mean food regime will

352 have eggs that are on average 10.5% smaller, consistent with a *Periodic strategy* sensu Wine-

353 miller & Rose (1992). However, our results indicate that environments with low seasonality

354 and white-coloured noise (i.e. not temporally autocorrelated) have eggs that are 15.2% larger

355 on average, contrary to the expectations of Winemiller and Rose's suggested *Opportunistic*

356 *strategy*. Instead, this finding is consistent with Einum & Fleming (2004) suggestion that

357 mothers employ a conservative bet-hedging strategy under such conditions. Einum & Flem-

358 ing (2004) suggested that when mothers cannot anticipate the environment of their offspring,

they should insulate them from poor conditions through better provisioning. Overall our findings strongly support the notion that environmental quality directly affects egg size (Smith & Fretwell, 1974; Parker & Begon, 1986; Rollinson *et al.*, 2013; Shama, 2015). Importantly, all three components of environmental variation: mean, seasonality, and colour of environmental noise drove the patterns in egg size.

After accounting for the effects of the environment and phylogeny, a portion of the variation in the data was left unexplained (i.e. residuals in Fig. 2). For the majority of our records, we were not able to properly assess the depth in which the egg was collected. Depth is known to affect patterns in egg size in marine organisms (Laptikhovsky, 2006). So it is possible that some of our records for "tropical" species were obtained at greater depths were temperatures are much lower. Another aspect that could influence part of the unexplained variance is spawning mode (Kasimatis & Riginos, 2016). However, our phylogenetic hierarchical model already accounts for clade-specific deviations that drive substantial changes in egg size (e.g. brooding in seahorses and cardinalfishes), and adding spawning mode to the model, as tested through our hypothesis *H4*, did not significantly improve model fit. This could be due to the fact that some spawning modes are under-represented in the database (e.g. brooders and scatterers), and therefore future studies with more estimates for these groups might be able to resolve the parameter estimates with greater precision.

Here we synthesised large amounts of data to uncover the role of multiple environmental components in shaping marine fish egg size. By doing so, we have confirmed previous results reported in the literature as well as revealed novel patterns. Particularly, our analysis formally decomposed two axes of environmental predictability, showing that they both affect egg size, and in ways consistent to those observed in marine invertebrates, suggesting commonalities in how ectotherm egg size responds to environmental change. Future environmental change in the world's oceans will not only drive the average environmental state at different locations, but also their degree of environmental predictability, so our study suggests that anthropogenically-mediated changes in the environment will have profound effects on the distribution of marine life histories.

Statement of authorship

All authors conceived the study; DRB compiled and standardised the dataset; DRB and SCB analysed the data; DRB wrote the first draft of the manuscript, and all authors contributed substantially to revisions.

Data accessibility statement

All data, R packages, and code (data manipulation, analyses, figures and tables) can be downloaded from our GitHub repositories (www.github.com/dbarneche/noaaErddap; www.github.com/dbarneche/envPred; https://github.com/dbarneche/fishEggSize).

Biosketches

Diego Barneche studies theoretical and empirical approaches that aim to unify evolutionary ecology across levels of biological organisation. He is passionate about data, code, and reproducible science.

Scott Burgess studies life history evolution, dispersal, and population dynamics of coastal marine species using experimental and modelling approaches.

Dustin Marshall works on marine life histories and is increasingly interested in mechanistic explanations for the diversity of life histories found in the sea.

411 **References**

412   Betancur-R., R., Broughton, R.E., Wiley, E.O., Carpenter, K., López, J.A., Li, C., Holcroft,

413   N.I., Arcila, D., Sanciangco, M., Cureton II, J.C., Zhang, F., Buser, T., Campbell, M.A.,

414   Ballesteros, J.A., Roa-Varon, A., Willis, S., Borden, W.C., Rowley, T., Reneau, P.C., Hough,

415   D.J., Lu, G., Grande, T., Arratia, G. & Ortí, G. (2013) The tree of life and a new classification

416   of bony fishes. *PLOS Currents Tree of Life*, **1**.

417   Braga Goncalves, I., Ahnesjö, I. & Kvarnemo, C. (2011) The relationship between female

418   body size and egg size in pipefishes. *Journal of Fish Biology*, **78**, 1847–1854.

419   Brown, J.H. (1973) Species diversity of seed-eating desert rodents in sand dune habitats.

420   *Ecology*, **54**, 775–787.

421   Burgess, S.C. & Marshall, D.J. (2014) Adaptive parental effects: The importance of estimat-

422   ing environmental predictability and offspring fitness appropriately. *Oikos*, **123**, 769–776.

423   Bürkner, P. (2017) brms: An R package for Bayesian multilevel models using Stan. *Journal*

424   *of Statistical Software, Articles*, **80**, 1–28.

425   Cohen, D. (1966) Optimizing reproduction in a randomly varying environment. *Journal of*

426   *Theoretical Biology*, **12**, 119–129.

427   Colwell, R.K. (1974) Predictability, constancy, and contingency of periodic phenomena.

428   *Ecology*, **55**, 1148–1153.

429   Crump, M.L. (1981) Variation in propagule size as a function of environmental uncertainty

430   for tree frogs. *The American Naturalist*, **117**, 724–737.

431   Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G.,

432   Gruber, B., Lafourcade, B., Leitão, P.J., Münkemüller, T., McClean, C., Osborne, P.E.,

433   Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D. & Lautenbach, S. (2013) Collinearity:

434   A review of methods to deal with it and a simulation study evaluating their performance.

435   *Ecography*, **36**, 27–46.

436   Duarte, C.M. & Alcaraz, M. (1989) To produce many small or few large eggs: A size-

437   independent reproductive tactic of fish. *Oecologia*, **80**, 401–404.

438   Einum, S. & Fleming, I.A. (2002) Does within☐population variation in fish egg size reflect

439   maternal influences on optimal values? *The American Naturalist*, **160**, 756–765.

440   Einum, S. & Fleming, I.A. (2004) Environmental unpredictability and offspring size: Conser-

441   vative versus diversified bet-hedging. *Evolutionary Ecology Research*, **6**, 443–455.

442   Elgar, M.A. (1990) Evolutionary compromise between a few large and many small eggs:

443   Comparative evidence in teleost fish. *Oikos*, **59**, 283–287.

444   Feary, D.A., Pratchett, M.S., J Emslie, M., Fowler, A.M., Figueira, W.F., Luiz, O.J., Naka-

445   mura, Y. & Booth, D.J. (2014) Latitudinal shifts in coral reef fishes: Why some species do

446   and others do not shift. *Fish and Fisheries*, **15**, 593–615.

447   Frédérich, B., Sorenson, L., Santini, F., Slater, G.J. & Alfaro, M.E. (2013) Iterative ecological

448   radiation and convergence during the evolutionary history of damselfishes (pomacentridae).

449   *The American Naturalist*, **181**, 94–113.

450   Fritzsche, R.A. (1978) *Development of fishes of the mid-atlantic bight: An atlas of egg, larval*

451   *and juveniles stages*, Fish; Wildlife Service U.S. Department of Interior.

452   Fronhofer, E.A., Stelz, J.M., Lutz, E., Poethke, H.J. & Bonte, D. (2014) Spatially correlated

453   extinctions select for less emigration but larger dispersal distances in the spider mite *Tetrany-*

454   *chus urticae*. *Evolution*, **68**, 1838–1844.

455   Glynn, E.F., Chen, J. & Mushegian, A.R. (2006) Detecting periodic patterns in unevenly

456   spaced gene expression time series using lomb–Scargle periodograms. *Bioinformatics*, **22**,

457   310–316.

458   Grafen, A. (1989) The phylogenetic regression. *Philosophical Transactions of the Royal Soci-*

459   *ety of London B: Biological Sciences*, **326**, 119–157.

460   Hadfield, J.D. & Nakagawa, S. (2010) General quantitative genetic methods for comparative

461   biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical char-

462   acters. *Journal of Evolutionary Biology*, **23**, 494–508.

463   Halley, J.M. & Kunin, W.E. (1999) Extinction risk and the 1/f family of noise models. *Theo-*

464   *retical Population Biology*, **56**, 215–230.

465 Hastings, A. (1983) Can spatial variation alone lead to selection for dispersal? *Theoretical*
466 *Population Biology*, **24**, 244–251.

467 Hixon, M.A., Johnson, D.W. & Sogard, S.M. (2014) BOFFFFs: On the importance of con-
468 serving old-growth age structure in fishery populations. *ICES Journal of Marine Science*, **71**,
469 2171.

470 Housworth, E.A., Martins, E.P. & Lynch, M. (2004) The phylogenetic mixed model. *The*
471 *American Naturalist*, **163**, 84–96.

472 Kasimatis, K. & Riginos, C. (2016) A phylogenetic analysis of egg size, clutch size, spawn-
473 ing mode, adult body size, and latitude in reef fishes. *Coral Reefs*, **35**, 387–397.

474 Kokita, T. (2003) Potential latitudinal variation in egg size and number of a geographically
475 widespread reef fish, revealed by common-environment experiments. *Marine Biology*, **143**,
476 593–601.

477 Laptikhovsky, V. (2006) Latitudinal and bathymetric trends in egg size variation: A new look
478 at thorson's and rass's rules. *Marine Ecology*, **27**, 7–14.

479 Marshall, D.J. & Burgess, S.C. (2015) Deconstructing environmental predictability: Season-
480 ality, environmental colour and the biogeography of marine life histories. *Ecology Letters*, **18**,
481 174–181.

482 Marshall, D.J., Krug, P.J., Kupriyanova, E.K., Byrne, M. & Emlet, R.B. (2012) The biogeog-
483 raphy of marine invertebrate life histories. *Annual Review of Ecology, Evolution, and System-*
484 *atics*, **43**, 97–114.

485 Marshall, N.B. (1953) Egg size in arctic, antarctic and deep-sea fishes. *Evolution*, **7**, 328–
486 341.

487 Meiri, S. (2011) Bergmann's rule – what's in a name? *Global Ecology and Biogeography*, **20**,
488 203–207.

489 Michonneau, F., Brown, J.W. & Winter, D.J. (2016) Rotl: An r package to interact with the
490 open tree of life data. *Methods in Ecology and Evolution*, n/a–n/a.

491 Morrongiello, J.R., Bond, N.R., Crook, D.A. & Wong, B.B.M. (2012) Spatial variation in
492 egg size and egg number reflects trade-offs and bet-hedging in a freshwater fish. *Journal of*

18

Animal Ecology, **81**, 806–817.

Moser, H.G. (1996) *The early stages of fishes in the california current region. atlas no. 33*, Allen Press, Lawrence, Kansas, USA.

O'Connor, M.I., Bruno, J.F., Gaines, S.D., Halpern, B.S., Lester, S.E., Kinlan, B.P. & Weiss, J.M. (2007) Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. *Proceedings of the National Academy of Sciences*, **104**, 1266–1271.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877–884.

Paradis, E., Claude, J. & Strimmer, K. (2004) APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, **20**, 289–290.

Parker, G.A. & Begon, M. (1986) Optimal egg size and clutch size: Effects of environment and maternal phenotype. *The American Naturalist*, **128**, 573–592.

Rass, T.S. (1941) *Geograficheskie parallelizmy v stroenii kostistykh ryb severnykh morei (Analogous or parallel variations in structure and development of fishes in Northern and Arctic seas)*, MOIP, Moscow, Russia.

Rollinson, N., Hutchings, J.A., Sherratt, A.E.T.N. & McPeek, E.M.A. (2013) Environmental quality predicts optimal egg size in the wild. *The American Naturalist*, **182**, 76–90.

Rubio de Casas, R., Willis, C.G., Pearse, W.D., Baskin, C.C., Baskin, J.M. & Cavender-Bares, J. (2017) Global biogeography of seed dormancy is determined by seasonality and seed size: A case study in the legumes. *New Phytologist*, **214**, 1527–1536.

Shama, L.N.S. (2015) Bet hedging in a warming ocean: Predictability of maternal environment shapes offspring size variation in marine sticklebacks. *Global Change Biology*, **21**, 4387–4400.

Slobodkin, L.B. & Sanders, H.L. (1969) *On the contribution of environmental predictability to species diversity. Diversity and stability in ecological systems* Brookhaven symposia in biology. (ed. by G.M. Woodwell) and H.H. Smith), Brookhaven National Laboratory, Upton, NY, USA.

521 Smith, C.C. & Fretwell, S.D. (1974) The optimal balance between size and number of off-

522 spring. *The American Naturalist*, **108**, 499–506.

523 Thorson, G. (1950) Reproductive and larval ecology of marine bottom invertebrates. *Biologi-

524 cal Reviews*, **25**, 1–45.

525 Travis, J.M.J. (2001) The color of noise and the evolution of dispersal. *Ecological Research*,

526 **16**, 157–163.

527 Vasseur, D.A. & Yodzis, P. (2004) The color of environmental noise. *Ecology*, **85**, 1146–

528 1152.

529 Vasseur, D.A., DeLong, J.P., Gilbert, B., Greig, H.S., Harley, C.D.G., McCann, K.S., Savage,

530 V., Tunney, T.D. & O'Connor, M.I. (2014) Increased temperature variation poses a greater risk

531 to species than climate warming. *Proceedings of the Royal Society of London B: Biological

532 Sciences*, **281**.

533 Vehtari, A., Gelman, A. & Gabry, J. (2016) loo: Efficient leave-one-out cross-validation and

534 WAIC for Bayesian models.

535 Winemiller, K.O. & Rose, K.A. (1992) Patterns of life-history diversification in North Ameri-

536 can fishes: Implications for population regulation. *Canadian Journal of Fisheries and Aquatic

537 Sciences*, **49**, 2196–2218.

538 Zuo, W., Moses, M.E., West, G.B., Hou, C. & Brown, J.H. (2012) A general model for effects

539 of temperature on ectotherm ontogenetic growth and development. *Proceedings of the Royal

540 Society of London B: Biological Sciences*, **279**, 1840–1846.

Tables

Table 1. Pearson correlation ($n = 1078$ observations) among predictor environmental variables after excluding data with environmental coverage < 87% (see methods). Bold values represent statistically significant correlations (i.e. $p < 0.001$).

	SST (Season.)	SST (Colour)	SST (Mean)	Chl (Season.)	Chl (Colour)	Chl (Mean)
SST (Season.)	-	**0.68**	**-0.73**	**0.39**	**-0.22**	**0.27**
SST (Colour)	**0.68**	-	**-0.71**	**0.24**	-0.05	**0.39**
SST (Mean)	**-0.73**	**-0.71**	-	**-0.32**	-0.03	**-0.49**
Chl (Season.)	**0.39**	**0.24**	**-0.32**	-	**0.16**	**-0.16**
Chl (Colour)	**-0.22**	-0.05	-0.03	**0.16**	-	-0.02
Chl (Mean)	**0.27**	**0.39**	**-0.49**	**-0.16**	-0.02	-

Table 2. Average estimates and 95% posterior credible intervals (of Bayesian posterior distributions) for fitted parameters in the parsimonious model.

Parameter	Mean Estimate	2.5%	97.5%	Effective sampling
Random Effects				
$\sigma_{\gamma spp}$	0.24	0.20	0.28	5346
$\sigma_{\gamma phy}$	1.57	1.13	2.04	4275
Fixed effects				
β_0 (intercept)	0.43	-1.31	2.24	6309
β_{SST}	-0.02	-0.03	-0.01	17046
β_{Chl}	-0.27	-0.37	-0.16	22500
$\beta_{\theta Chl}$	0.27	0.14	0.40	22500
$\beta_{SeasonChl}$	-0.69	-0.87	-0.51	22500
$\beta_{SeasonSST}$	0.82	0.47	1.19	12568

Figure legends

551 **Figure 1.** Raw data showing the relationship between egg size (diameter in mm) and abso-

552 lute latitude. Blue and red points represent data from the southern and northern hemispheres

553 respectively.

554 **Figure 2.** Fish egg size (diameter in mm) as a function of *(a)* mean SST *(b)* and mean

555 chlorophyll-a. Parameter estimates (parsimonious model in Table 2) were obtained using

556 Bayesian phylogenetic hierarchical models. The effects of environmental predictability and

557 mean chlorophyll-a on egg size were controlled for in *(a)*, and the environmental predictability

558 and mean SST in *(b)*. Random effects attributable to relatedness among species and species

559 uniqueness (γphy and γspp in eqn 1) have also been controlled for. Thus, the deviations from

560 the mean model fit represent the actual model residuals. Top-left equations show average

561 fixed-effect predictions. Dashed thick line represents average posterior fixed-effect fit, and

562 dashed thin lines represent the 95% posterior credible intervals extracted from 22,500 MCMC

563 samples from our Bayesian analysis.

564 **Figure 3.** Predictive surface between chlorophyll-a seasonality and noise colour on the aver-

565 age size of marine fish eggs (diameter in mm). Each black circle represents a separate obser-

566 vation in our dataset and the colour coding represents the predicted surface using parameters

567 from our parsimonious Bayesian phylogenetic hierarchical model (Table 2). Predicted egg

568 sizes were calculated for mean SST and chlorophyll-a, as well as mean SST seasonality.