Supporting Appendix

Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation

Mary I. O’Connor, John F. Bruno, Steven D. Gaines, Benjamin S. Halpern, Sarah E. Lester, Brian P. Kinlan, Jack M. Weiss

This *Supporting Appendix* describes in detail the statistical methods of the main text. It gives a general overview of multilevel modeling, an explanation for the various models that were considered, a rationale for the centering scheme that was used, and a description of the model diagnostics that were used in assessing the fit of the final model. It also provides additional information on caterpillar plots, a graphical tool that plays an important role in motivating our claim for the existence of a universal temperature dependence model of larval development in marine species. The final section displays model fits for all species used in the analysis.

I. General Modeling Strategies

The main text uses data from multiple published studies on many different marine species in an attempt to develop a quantitative model of the relationship between planktonic larval duration time (PLD) and temperature. Analysis is complicated by the fact that the data are hierarchical consisting both of multiple temperature measurements made on the same species in the same study as well as observations on different species from different studies. Because different studies used different species as well as different experimental designs and laboratory techniques, the data are heterogeneous. Observations coming from the same study would be expected to show a different degree of variability than observations coming from different studies.

Statistically the different studies are blocks. Blocks can be included in an analysis as fixed effects or as random effects. They are random if the individual studies can be viewed as a sample from a population of similar such studies and the primary research goal is to draw inference about the basic phenomenon that the individual studies represent. The purpose for including random effects in a model is to account for observational heterogeneity. Observations that share the same random effect will necessarily be more similar (and thus correlated) than will observations with different random effects.

The incorporation of random effects in a hierarchical design leads to what is generally referred to in the social sciences as a multilevel analysis. An extensive literature exists describing multilevel models (1–7). In other disciplines such models are called mixed models or random coefficient models (8–10). To model the relationship between PLD and temperature we employ a mixed model to synthesize the results from multiple published studies, thus carrying out a form of meta-analysis. Mixed models have been used for meta-analysis in many disciplines, e.g.,

Table 10. The fundamental set of level-2 equations for the level-1 model shown in Eq. 19

Model	Level-2 equations	Assumptions made for the random effects	Composite equation
A	$\beta_{0i} = \beta_0$ $\beta_{1i} = \beta_1$	—	$\log(PLD_{ij}) = \beta_0 + \beta_1 \log T_{ij} + \varepsilon_{ij}$
B	$\beta_{0i} = \beta_0 + u_{0i}$ $\beta_{1i} = \beta_1$	$u_{0i} \sim N(0, \tau_0^2)$	$\log(PLD_{ij}) = (\beta_0 + u_{0i}) + \beta_1 \log T_{ij} + \varepsilon_{ij}$
C	$\beta_{0i} = \beta_0$ $\beta_{1i} = \beta_1 + u_{1i}$	$u_{1i} \sim N(0, \tau_1^2)$	$\log(PLD_{ij}) = \beta_0 + (\beta_1 + u_{1i}) \log T_{ij} + \varepsilon_{ij}$
D	$\beta_{0i} = \beta_0 + u_{0i}$ $\beta_{1i} = \beta_1 + u_{1i}$	$\begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_0^2 & \tau_{01} \\ \tau_{01} & \tau_1^2 \end{bmatrix}\right)$	$\log(PLD_{ij}) = (\beta_0 + u_{0i}) +$ $(\beta_1 + u_{1i}) \log T_{ij} + \varepsilon_{ij}$

T_{ij} is temperature (°C) and PLD_{ij} is planktonic larval development time for species *i* at time *j*.

agriculture (11), sports science (12), medicine (13), and fisheries (14).

We find the multilevel way of viewing mixed models to be an especially appealing one. In the multilevel formulation parameters at one level of a hierarchy are in turn modeled at the next higher level using predictors measured at that level. Consequently in a 2-level model a distinction is made between variables measured at the individual and group levels. This distinction affects both the degrees of freedom used in statistical tests as well as the way in which main effects and interactions are interpreted.

Multilevel modeling is a parsimonious way of dealing with observational heterogeneity. In a multilevel model the many parameters that would need to be estimated in a fixed effects model are replaced by the far smaller number of parameters needed to describe the distribution of the random effects. Multilevel models readily handle unbalanced and missing data, so that even units with one or two observations can contribute useful information in a multilevel analysis. The basic strategy in multilevel modeling is to construct an equation that describes the behavior of individuals (the level-1 model) and then to formulate additional equations that explain how each parameter appearing in the level-1 equation varies across individuals. These additional equations comprise the level-2 model.

As an illustration, one of the models that we consider describes planktonic larval duration time (PLD) of an individual species as a linear function of the ambient temperature (T) when both variables are measured on a log scale[*]. The level-1 model is

$$\log(PLD_{ij}) = \beta_{0i} + \beta_{1i} \log T_{ij} + \varepsilon_{ij}. \qquad \textbf{[19]}$$

Here *i* indexes the species and *j* the individual observation on that species. The term ε_{ij} denotes the error, which is assumed to arise from some probability distribution. Typically this distribution is taken to be normal with mean zero and a variance to be estimated, i.e., $\varepsilon_{ij} \sim N(0, \sigma^2)$. (In Section II, we relax the normality assumption.) As is indicated by the subscript *i* on the parameters β_{0i} and β_{1i}, each species has its own intercept and slope.

The different level-1 equations are linked together by the level-2 model. Since there are two model parameters in Eq. **19**, β_{0i} and β_{1i}, there are potentially two equations at level 2, one for each parameter. Thus in this example four fundamental sets of level-2 equations are possible.[†] These are listed in SI Table 10.[‡]

[*] In this *Supporting Text* we follow standard mathematical convention and use the notation log to denote the natural logarithm function.

[†] There are only four fundamental level-2 equations because we are modeling interspecies variability solely by the inclusion of random effects. We don't consider adding level-2 predictors at this point.

[‡] SI here denotes *Supporting Information.* We use this notation to indicate that the table or figure in question appears in this *Supporting Text* or in one of the other online supporting documents rather than in the main text itself.

The terms u_{0i} and u_{1i} represent the random effects (also called level-2 residuals). As is indicated by the index *i*, they are unique to individual species. The level-2 residual for a species is a surrogate for all those unmeasured variables that make that species different from all other species with respect to a specific model parameter (and thus make the multiple observations on a single species at different temperatures similar to each other). The univariate or joint normal distributions that are assumed for the random effects are standard assumptions that need to be examined as part of model assessment. The composite equation shown in the last column of SI Table 10 is the equation that results from plugging the level-2 equations into the level-1 model given in Eq. **19**.

The linear mixed model in composite form can be expressed compactly using vector notation to yield what's called the Laird-Ware formulation of the model (15; ref. 16, p. 327). For species *i* any of the models in SI Table 10 can be expressed as follows.

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\varepsilon}_i$$

Here $\boldsymbol{\beta}$ and $\mathbf{u}_i$ are the vectors of fixed and random effects respectively while $\mathbf{X}_i$ and $\mathbf{Z}_i$ are the corresponding design matrices in which the observed values of the different predictors appear as columns. Typically the columns of $\mathbf{Z}_i$ are a subset of the columns of $\mathbf{X}_i$, as is the case in our model, although this is not necessary. As an illustration, for Model B of SI Table 10 the terms $\boldsymbol{\beta}$, $\mathbf{u}_i$, $\mathbf{X}_i$, and $\mathbf{Z}_i$ are the following.

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}, \ \mathbf{u}_i = u_{0i}$$

$$\mathbf{X}_i = \begin{bmatrix} 1 & \log T_{i1} \\ 1 & \log T_{i2} \\ \vdots & \vdots \\ 1 & \log T_{im_i} \end{bmatrix}, \text{ and } \mathbf{Z}_i = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Here $\mathbf{X}_i$ is $m_i \times 2$ and $\mathbf{Z}_i$ is $m_i \times 1$ where m_i is the number of observations on species *i*.

Continuing with this formulation, the conditional mean response for an individual species is

$$E(\mathbf{Y}_i|\mathbf{u}_i) = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i$$

and is referred to as the subject-specific model. Notice that the subject-specific model includes both fixed and random effects. The mean response averaged over all species is

$$E(\mathbf{Y}_i) = \mathbf{X}_i\boldsymbol{\beta}\ .$$

This is usually referred to as the marginal model or population-averaged model and includes only fixed effects.

The four fundamental level-2 models given in SI Table 10 are interpreted as follows. Model A contains no random effects. As a result each species is assigned the same intercept and slope. Accordingly the composite equation for Model A is just an ordinary regression model in which the hierarchical structure of the data is ignored. Models B and C incorporate random effects for one parameter but not the other. Model B assumes individual species differ in their intercepts but have a common slope, while Model C assumes species differ in their slopes but have a common intercept. Model D incorporates random effects for both the slope and intercept permitting individual species to differ in both.

Assuming that the level-1 model specified in Eq. **19** is an adequate description of the relationship between PLD and temperature, the various level-2 models allow us to examine our primary research question: is there a uniform relationship between PLD and temperature across species? For this Models A, B, and D are of greatest interest. We can compare these models using null hypothesis significance testing (likelihood ratio tests) and/or information-theoretic methods (AIC). If the evidence favors Model D then this would suggest that all species are unique with respect to PLD. In Model D not only does overall PLD vary at any given temperature across species (different intercepts), but the effect of temperature on PLD varies across species too (different slopes). If the evidence favors Model B then we have support for a common temperature effect across species (common slope) but with individual species still differing in their PLD at any given temperature (different intercepts). If the evidence favors Model A then a common PLD-temperature model is appropriate for all species.

Once the appropriate level-2 structure is determined, the next obvious step would be to include level-2 predictors (variables measured at the species level) into the level-2 equations with the goal of explaining some of the variability currently accounted for by the level-2 random effects. SI Table 4 (*Supporting Text* 1) and Fig. 4

(main text) make a preliminary attempt at this by including developmental mode and the normal temperature range of a species as predictors in the level-2 intercept equation. Further development of the level-2 model will be described in a future publication.

II. Choice of Level-1 Model

While a number of different models have been used to relate planktonic larval duration time with temperature (see, e.g., ref. 17), the general consensus is that the relationship is well-approximated by a power law function (18), i.e., an equation of the form

$$Y = aT^b. \qquad \textbf{[20]}$$

A model of this form was originally used in biology by Huxley (19) to describe allometric growth. In Eq. **20** Y is PLD, T is temperature, and a and b denote parameters to be estimated. An algebraically equivalent linear model can be obtained by log transforming both sides of Eq. **20** to obtain

$$\log Y = \log a + b \log T. \qquad \textbf{[21]}$$

Statistically Eqs. **20** and **21** are not the same and it's not always clear from the literature exactly which of these equations was fit. The statistical versions of Eqs. **20** and **21** are

$$Y = aT^b + \varepsilon \qquad \textbf{[22a]}$$
$$\log Y = \beta_0 + \beta_1 \log T + \varepsilon \qquad \textbf{[22b]}$$

where typically it is assumed $\varepsilon \sim N\left(0, \sigma^2\right)$ in both. Eq. **22a** models the arithmetic mean planktonic larval duration time while Eq. **22b** models the geometric mean. When converted to an arithmetic scale by exponentiating, Eq. **22b** induces a multiplicative error structure yielding a response that is lognormally distributed. Although these distinctions are rarely discussed in the ecological literature (but see refs. 20–22), they have generated a considerable amount of heat in other disciplines (see, e.g., refs. 23–24 for a history of the debate in paleobiology).

To select an appropriate error structure for our level-1 model we examined the mean-variance relationship of the response. SI Fig. 9A plots PLD versus temperature and superimposes a nonparametric smooth (lowess) estimate of the mean. As can be seen the variance does increase with the mean; the vertical spread of the data is greater on the left side of the plot where the mean is also larger. In a lognormal distribution the variance of the distribution is proportional to the square of the mean.[§] Since mean larval duration time is a function of temperature we can approximate the mean-variance relationship by grouping larval duration times by temperature. SI Fig 9B plots the means and variances of PLD in groups formed from sextiles of the temperature distribution. (Sextiles were chosen because they provide enough data in each group to yield a stable variance estimate while still leaving enough data points to fit a regression curve.) A least squares fit of a quadratic model (in which the intercept is constrained to be zero) is superimposed. As is clear from the plot a quadratic model seems to approximate the mean-variance relationship quite well.

SI Fig. 9B clearly rules out using Eq. **22a** with normally distributed errors (in which case the mean and variance would be independent) and instead supports the use of Eq. **22b** with normally distributed errors (yielding a lognormally distributed response on an arithmetic scale). The lognormal is not the only distribution characterized by having a constant coefficient of variation. Another is the gamma distribution. Thus an alternative to fitting Eq. **22b** with normal errors is to fit Eq. **22a** with gamma errors.

A theoretical, a priori case for a lognormal distribution can be made, e.g., if the process being modeled is decomposable into a product of many independent components. Taking the log of this product yields a sum of numerous independent log-transformed random variables, which, from the central limit theorem, should have a normal distribution in the limit. A gamma distribution, on the other hand, can arise theoretically as a sum of independent exponential random variables, each of which represents the waiting time to a randomly occurring event as dictated by a Poisson process.[¶]

Wiens (26) showed that the two competing models, log-normal and gamma, can yield very different results. He recommended fitting both a lognormal and a gamma model when possible to serve as a check for whether the results obtained are model-independent. McCullagh & Nelder (27)

[§] Another way of saying this is that in a lognormal distribution the coefficient of variation is constant.

[¶] Gamma random variables can be numerically generated by summing the logarithms of a finite number of uniformly generated random variables (25). Thus in this special instance the gamma distribution is a finite analog of the lognormal distribution.

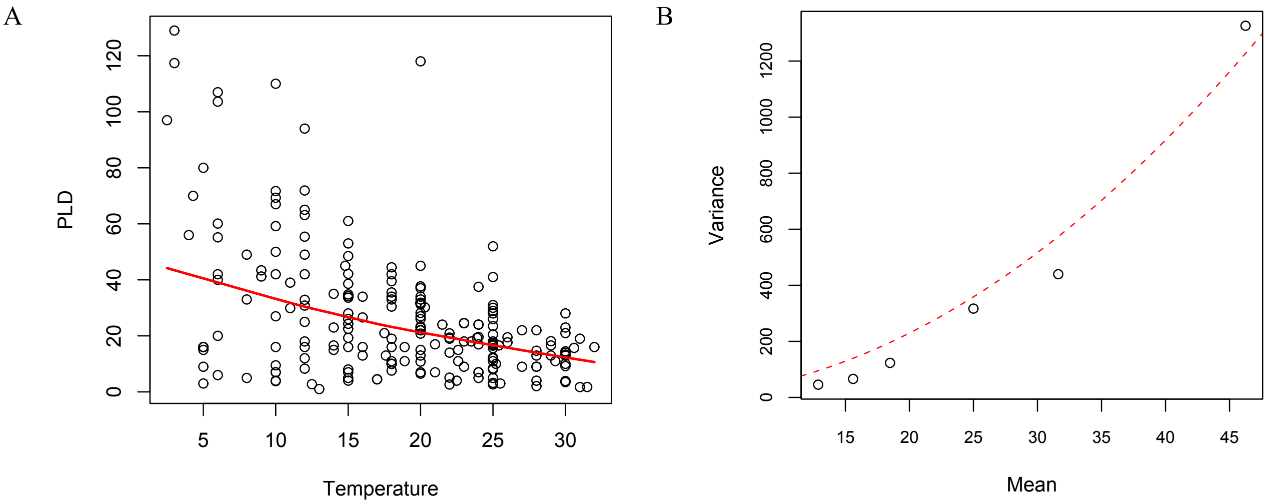

Fig. 9. Choosing an error structure for the level-1 model. (A) Scatter plot of planktonic larval duration time (PLD) as a function of temperature illustrating heteroscedasticity. The variance of PLD appears to increase with mean PLD. The superimposed lowess curve indicates the trend in the mean. (B) Means and variances of PLD for the observations shown in (A) but grouped into sextiles of temperature. A least squares estimate of the theoretical mean-variance relationship for the log-normal distribution is superimposed for comparison.

note that when the variance of the response is small enough such that

$$\mathrm{Var}(\log Y) \approx \frac{\mathrm{Var}(Y)}{\mu^2},$$

the two analyses should produce similar results. For our data we find the following.

$$\mathrm{Var}(\log Y) = 0.8052$$
$$\frac{\mathrm{Var}(Y)}{\mu^2} = 0.7842$$

suggesting that the choice of a gamma or lognormal distribution here may not matter.

While the lognormal model has often been used to model the relationship between larval duration time and temperature, a nonlinear model with gamma errors to our knowledge has not. In all the basic models described in this *Supporting Text*, we found that the choice of probability distribution for the response, lognormal versus gamma, made only trivial differences in fit. Because our final lognormal model provides an adequate fit to the data and because of the additional complexity that arises when fitting a nonlinear random effects model with gamma errors to data, we present only the results based on a lognormal probability distribution here. Consequently in all the log-transformed level-1 equations presented below we assume $\varepsilon_{ij} \sim N(0, \sigma^2)$.

III. Centering the Regressor

Centering a regressor is often crucial to successfully fitting a multilevel model. Not only does centering improve the interpretability of the resulting model (6), it can also reduce the correlations between parameter estimates (9). Because parameter estimates are obtained using numerical optimization routines, this latter effect can be essential for ensuring computational stability and convergence to a proper solution. Furthermore, the presence of a high degree of correlation between the random effects in the model would undermine our primary objective to determine if a common PLD-temperature model suffices for most marine taxa.

As an illustration of this last point suppose the random effects of Model D in SI Table 10 were highly correlated. Using the notation of SI Table 10, the correlation, ρ, of the random effects in this model is the following.

$$\rho = \frac{\tau_{01}}{\sqrt{\tau_0^2} \cdot \sqrt{\tau_1^2}}$$

Let $Y_{ij} = \log(PLD_{ij})$, $x_{ij} = \log T_{ij}$, and suppose $\varepsilon_{ij} \sim N(0\ \sigma^2)$. By using properties of the expectation operator, it is easy to show that for Model D

$$\mathrm{Var}\left(Y_{ij}\right)=\tau_0^2+x_{ij}^2\tau_1^2+\sigma^2+2x_{ij}\tau_{01}\,.$$

Observe that the response variance does not partition neatly into separate, non-overlapping variance components. The term τ_{01} is the joint variability of β_{0i} and β_{1i} and its presence in the equation modifies the estimates of the individual intercept and slope variance components that are obtained. A high correlation between the intercept and slope random effects prevents the variance of the response from being decomposed into components that account separately for the variability in random intercepts and slopes. Consequently it isn't possible to distinguish variability in slopes from variability in intercepts in level-2 model D of SI Table 10.

Since distinguishing between intercept and slope variability is fundamental to understanding interspecific differences in the PLD-temperature relationship, some form of centering is mandatory. By centering we mean subtracting a constant from each observed value of the model regressor. Thus the centered version of Eq. **19** would be the following:

$$\log\left(PLD_{ij}\right)=\beta_{0i}+\beta_{1i}\left(\log T_{ij}-c\right)+\varepsilon_{ij} \qquad \textbf{[23]}$$

for some choice of constant *c*. For the quadratic model considered in Section IV (and also in the main text) the centered level-1 equation is

$$\begin{aligned}\log\left(PLD_{ij}\right)=\beta_{0i}+\beta_{1i}\left(\log T_{ij}-c\right)&\\ +\beta_{2i}\left(\log T_{ij}-c\right)^2+\varepsilon_{ij}\,.&\qquad \textbf{[24]}\end{aligned}$$

The centered UTD model used in the main text is the following.

$$\log\left(PLD_{ij}\right)=\beta_{0i}+\beta_{1i}\left(\frac{1}{k\left(T_{ij}+273\right)}-c\right)+\varepsilon_{ij}$$

where *k* is the Boltzmann constant. In this section we demonstrate three things.

A. A judicious choice of centering constant *c* can improve model interpretation.

B. Centering has no effect on model fit. The loglikelihood is the same whether the model is centered or not. Furthermore there is a one-to-one mapping between the parameter estimates of the centered and uncentered models.

C. Centering can dramatically reduce the correlation of the parameter estimates in the model.

A. Centering Improves Interpretation

To simplify notation, we suppress subscripts in this section and let $y=\log\left(PLD_{ij}\right)$ and $x=\log T_{ij}$. With these identifications Eq. **24** becomes the following.

$$\begin{aligned}y&=\beta_0+\beta_1\left(x-c\right)+\beta_2\left(x-c\right)^2\\ &=p_2\left(x,\,c\right) \qquad \textbf{[25]}\end{aligned}$$

The notation $p_n\left(x,\,c\right)$ is used to denote an n^{th}-degree polynomial in variable *x* with centering constant *c*. Consider first a linear equation with the predictor centered at *c*. After some algebra we obtain the following.

$$\begin{aligned}p_1(x,c)&=\beta_0+\beta_1\left(x-c\right)\\ &=\beta_0+\beta_1x-c\beta_1\\ &=\left(\beta_0-c\beta_1\right)+\beta_1x\\ &=\gamma_0+\gamma_1x \qquad \textbf{[26]}\end{aligned}$$

We see that the centered model with parameters β_0, β_1, and centering constant *c* is equivalent to an uncentered linear model with parameters γ_0 and γ_1. From the last equality in Eq. **26** by matching the coefficients of corresponding terms we obtain the following mapping between the two sets of parameters.

$$\begin{aligned}\gamma_0&=\beta_0-c\beta_1\\ \gamma_1&=\beta_1\end{aligned} \qquad \textbf{[27]}$$

Thus centering the equation has no effect on the value of the slope, but it does change the value of the intercept. From the definition of $p_1\left(x,\,c\right)$ we have, when $x=c$,

$$p_1(c,c)=\beta_0+\beta_1\left(c-c\right)=\beta_0\,.$$

So the intercept in the centered model is the value of the response variable when *x* = *c*. In the uncentered model, the intercept is the value of the response when *x* = 0. Since $x=\log T_{ij}$, the intercept in the uncentered model corresponds to a temperature of 1ºC, a value completely outside the range of our data set and outside the thermal range of most of the species being considered.

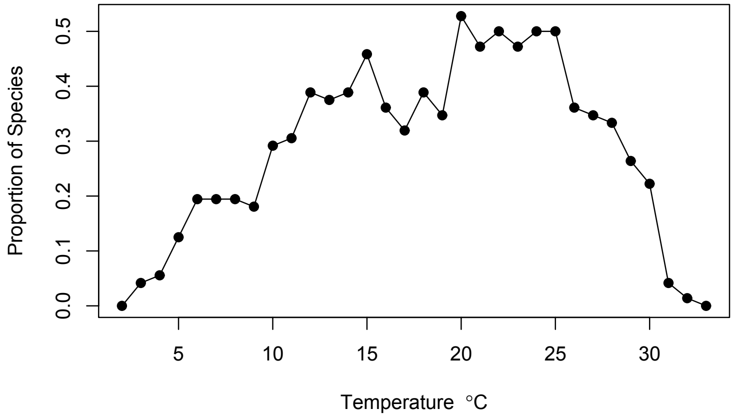

Fig. 10. Fraction of species used in the current study for which a given temperature occurred within the tested range of temperatures

Thus the value of the intercept in the uncentered equation is generally not biologically meaningful.

Next consider the quadratic model of Eq. **25**. Expanding terms we find

$$\begin{aligned} p_2(x,c) &= \beta_0 + \beta_1(x-c) + \beta_2(x-c)^2 \\ &= \beta_0 + \beta_1 x - c\beta_1 + \beta_2\left(x^2 - 2cx + c^2\right) \\ &= \left(\beta_0 - c\beta_1 + c^2\beta_2\right) + \left(\beta_1 - 2c\beta_2\right)x + \beta_2 x^2 \\ &= \gamma_0 + \gamma_1 x + \gamma_2 x^2 \end{aligned}$$

where γ_0, γ_1, and γ_2 are the corresponding parameters from the uncentered model. From the last equality we can identify the following mapping between the parameters of the uncentered and centered models.

$$\begin{aligned} \gamma_0 &= \beta_0 - c\beta_1 + c^2\beta_2 \\ \gamma_1 &= \beta_1 - 2c\beta_2 \\ \gamma_2 &= \beta_2 \end{aligned} \qquad \textbf{[28]}$$

Only the coefficient of the quadratic term is the same in the centered and uncentered models. Just as with the linear model, the intercept represents the value of the response at $x = c$.

$$\begin{aligned} p_2(c,c) &= \beta_0 + \beta_1(c-c) + \beta_2(c-c)^2 \\ &= \beta_0 \end{aligned}$$

In a quadratic model β_1 is no longer interpretable as the slope because the slope varies with x due to the presence of the quadratic term. To understand the role of β_1 we differentiate Eq. **25** with respect to x and evaluate the result at $x = c$.

$$\frac{d}{dx}p_2(x,c) = \beta_1 + 2\beta_2(x-c)$$

$$\left.\frac{d}{dx}p_2(x,c)\right|_{x=c} = \beta_1 + 2\beta_2(c-c) = \beta_1$$

Thus β_1 is the instantaneous rate of change of the response (with respect to temperature) at $x = c$.

In the uncentered quadratic model the intercept and linear terms reflect properties of the response when $x = 0$, but in the centered model the interpretation switches to $x = c$. By choosing c to represent a temperature of interest, the coefficients β_0 and β_1 gain biological meaning beyond their algebraic role as fitting constants. Typical choices for c in applications are the sample mean or median of the predictor, because under random sampling these values estimate the corresponding population values. Because the studies from which we obtained our data are not a random sample, the sample mean or median are not meaningful choices here. Instead we chose the logarithm of 15ºC as the centering constant. It turns out 15ºC is within the range of tested temperatures for a large fraction of the species we considered (SI Fig. 10).

B. Centering Does Not Alter the Fit

We first examine centered and uncentered versions of a regression equation for log(*PLD*) that is linear in log(temperature). For centering constant we use $\log \mathrm{T_c} = \log 15$. In each equation we allow both level-1 parameters to be random at level 2. This is Model D of SI Table 10. The model summary is shown in SI Table 11.

Table 11. Comparing the fit of uncentered and centered linear models.

Model 1: Uncentered

$$\log\left(PLD_{ij}\right)=\beta_0+u_{0i}+\left(\beta_1+u_{1i}\right)\log T_{ij}+\varepsilon_{ij}$$

Model 2: Centered

$$\log\left(PLD_{ij}\right)=\beta_0+u_{0i}+\left(\beta_1+u_{1i}\right)\left(\log T_{ij}-\log \mathrm{T_c}\right)+\varepsilon_{ij}$$

Model	$\hat{\beta}_0$	$\hat{\beta}_1$	Loglikelihood
1	7.118208	– 1.447185	– 84.3156
2	3.199160	– 1.447186	– 84.3156

$\mathrm{T_c}$ is the centering value 15°C and T_{ij} is temperature. Observe that the estimated linear coefficients are the same but that the intercepts are different in the two models. The loglikelihoods are identical.

Table 12. Comparing the fit of uncentered and centered quadratic models.

Model 3: Uncentered

$$\log\left(PLD_{ij}\right)=\beta_0+u_{0i}+\left(\beta_1+u_{1i}\right)\log T_{ij}+\beta_2\left(\log T_{ij}\right)^2+\varepsilon_{ij}$$

Model 4: Centered

$$\log\left(PLD_{ij}\right)=\beta_0+u_{0i}+\left(\beta_1+u_{1i}\right)\left(\log T_{ij}-\log \mathrm{T_c}\right)+\beta_2\left(\log T_{ij}-\log \mathrm{T_c}\right)^2+\varepsilon_{ij}$$

Model	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	Loglikelihood
3	4.978570	0.1055712	– 0.2823850	– 76.13833
4	3.193582	– 1.4238550	– 0.2823822	– 76.13833

$\mathrm{T_c}$ is the centering value 15°C and T_{ij} is temperature. Observe that the estimated quadratic coefficients are the same (within numerical accuracy) but that the linear coefficients and intercepts are different in the two models. The loglikelihoods are identical.

The loglikelihoods in SI Table 11[‖] are identical and as was predicted in Section IIIA the estimate for the slopes remains unchanged. The estimated intercepts can be converted from one to the other using Eq. **27**.[**]

```
> fixef(Model2)[1]-log(15)*
      fixef(Model2)[2]
(Intercept)
 7.118213
```

which we see is the reported estimate from the uncentered model.

Next consider centered and uncentered versions of a regression model for log(PLD) that is quadratic in log(temperature), again using $\log \mathrm{T_c}=\log 15$ as the centering constant (SI Table 12). To minimize convergence problems we only allow two of the level-1 parameters, the intercept and the coefficient of the linear term, to be random at level 2, a constraint we relax later.

The reported loglikelihoods are identical and as expected the estimated coefficient of the quadratic term is unchanged (accurate to the fourth decimal). The estimated intercept and linear terms can be converted from one to the other using Eq. **28**.

```
> fixef(Model4)[2]-2*log(15)*
    fixef(Model4)[3]
I(log(temp) - log(15))
                  0.1055555

> fixef(Model4)[1]-log(15)*
    fixef(Model4)[2]+(log(15))^2*
    fixef(Model4)[3]
(Intercept)
     4.978592
```

[‖] All numerical results displayed in this *Supporting Text* were obtained using R 2.1.1 (28). Mixed models were fit using either the lme function of the nlme package or the lmer function of the lme4 package.

[**] R syntax is displayed here. The R function fixef extracts a vector of estimates of the fixed effects (β_0 and β_1 in this case) from a model object. The bracket notation [ ] references a particular element in that vector.

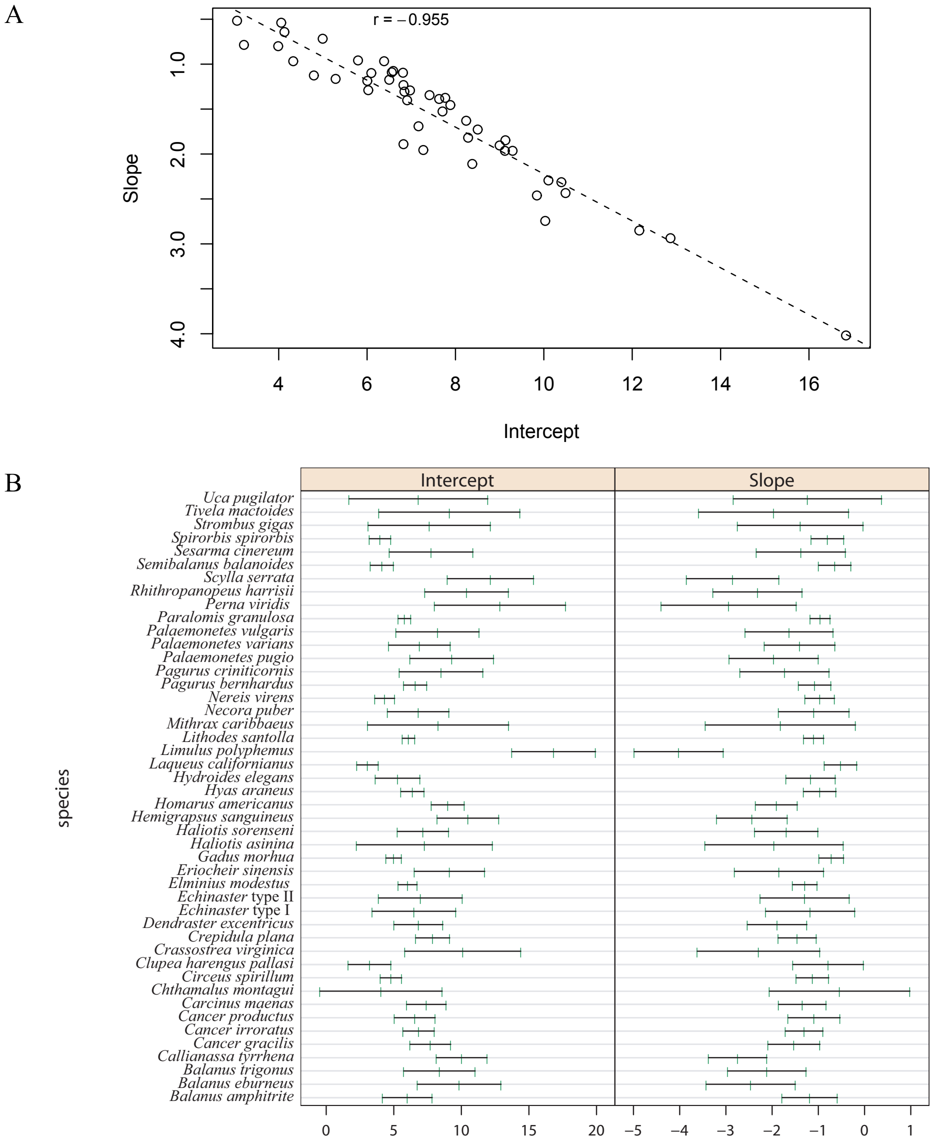

Fig. 11. The correlation of slopes and intercepts in an uncentered model. (A) Plots of slopes and intercepts for individual linear regressions of log(PLD) versus uncentered log(temperature). The high degree of negative correlation in the estimates is readily apparent. (B) The negative correlation is manifested in the fact that the 95% confidence interval plots for the intercept and slope are mirror images of each other.

The results are identical to four decimal places to estimates obtained for the uncentered model. Thus centering does no damage. The fit is the same and the coefficients for the uncentered model are readily recoverable if desired.

C. Centering Can Reduce Correlations Among Parameters

Consider again a linear model relating log(PLD) to log(temperature). Restricting the data to species with three or more temperature observations, we fit individual linear regression models of log(PLD) versus log(temperature), Eq. **19**, for each species separately without centering the predictor. SI Fig. 11A plots the individual slopes and intercepts that were obtained for each species while SI Fig. 11B plots point estimates and 95% confidence intervals for slopes and intercepts in a pairwise fashion for each species.

As is clear from SI Fig. 11A the correlation of slopes and intercepts is quite large and negative. A negative correlation between slopes and intercepts in uncentered models is typically observed when zero is a value of the predictor only as an extrapolation beyond the range of the

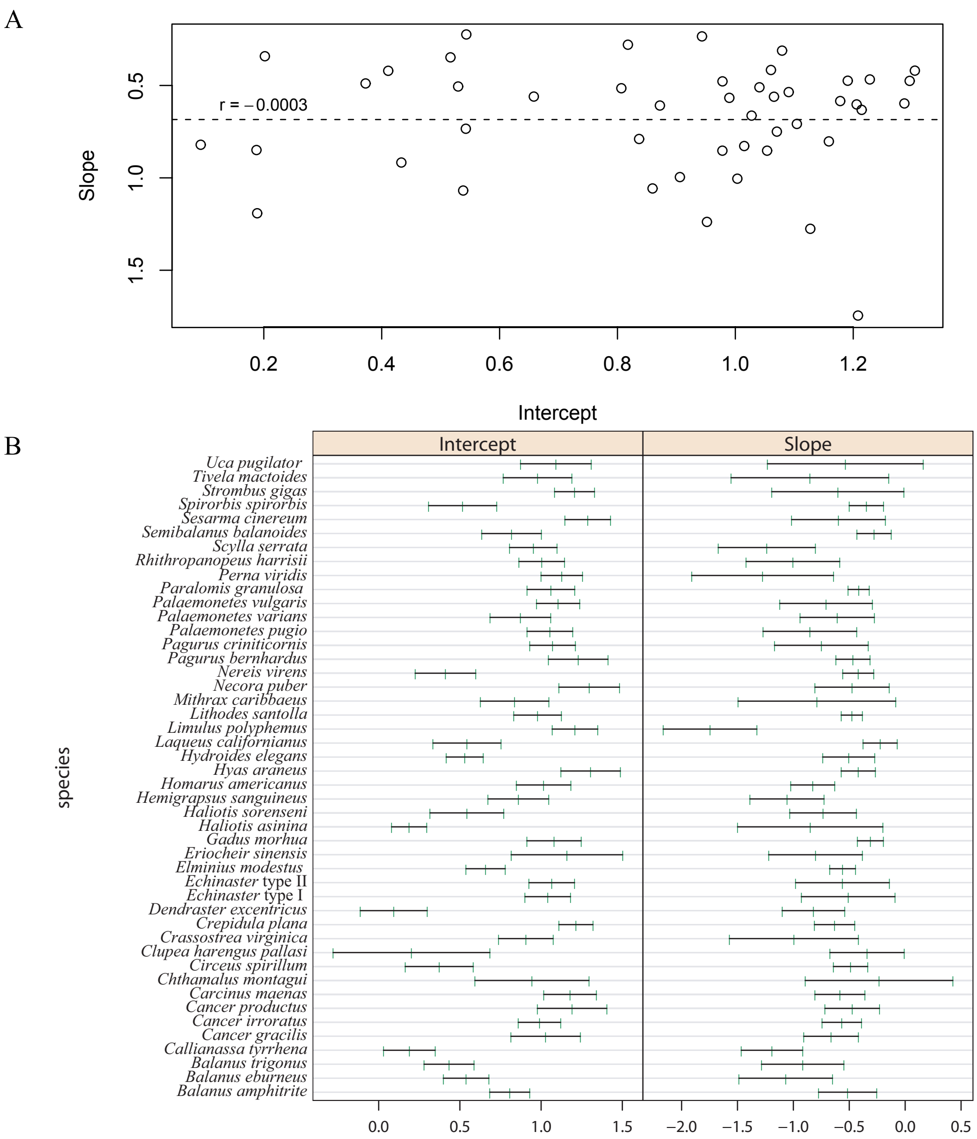

Fig. 12. The effect of centering on the correlation of slopes and intercepts. (A) Plots of slopes and intercepts for individual regressions with a centered regressor. Centering about log 33 has made these estimates essentially uncorrelated. (B) 95% confidence intervals are shown.

data, as is the case here (29; ref. 9, p. 34). Observe that in the interval plot (SI Fig. 11B) the left half representing the intercepts is essentially the mirror image of the right half representing the slopes. Thus in the uncentered model it is clear that slopes and intercepts cannot be interpreted independently of each other. When one is high the other is low.

SI Fig. 12 shows how centering can improve things. A centering constant of log 33 was used, a choice that causes the slope and intercepts to be essentially uncorrelated. In SI Fig. 12B the now uncorrelated intercepts and slopes clearly show a very different behavior from before. In particular notice that while most of the confidence intervals for the slopes overlap, many of the confidence intervals for the intercepts do not. This is evidence for greater interspecific variation among the intercepts than among the slopes, an interpretation that is now permissible given the low correlation that is present.

Clearly estimating separate slopes and intercepts for each species is inefficient. A better approach is to fit a common population-averaged model about which individual species are allowed to vary randomly—a multilevel model. In a multilevel model, centering the regressors affects the correlation of the random effects. SI Table 13 gives estimates of the parameters of the random

Table 13. The effect of centering on the correlation of random effects

Model	Random effect	Variance	Correlation
Uncentered, $c = 0$	Intercept Slope	4.206 0.271	–0.923
Centered, $c = \log 15$	Intercept Slope	0.856 0.271	–0.523
Centered, $c = \log 33$	Intercept Slope	0.628 0.271	–0.092

effects distribution, τ_0^2, τ_1^2, and ρ, for three different multilevel models. All three models share a common level-1 linear model between log(PLD) and log(temperature) and a common level-2 model with random effects for slopes and intercepts (model D in SI Table 10) but differ in the choice of centering constant. The first model is uncentered, the second model is centered at log 15, the centering constant used in the main text, and the third centers at log 33, the constant that yielded the uncorrelated intercepts and slopes of SI Fig. 12.

Just as with the individual regressions in SI Fig. 12, using a centering constant of log 33 in a multilevel model yields random effects for the slope and intercept that are nearly uncorrelated.[††] The centering constant, $\log \mathrm{T_c} = \log 15$, that we use in the main text produces an intermediate reduction in correlation.[‡‡]

IV. The Need for a Quadratic Term in the Level-1 Model

SI Table 14 shows the results of fitting the four models of SI Table 10 using maximum likelihood (ML) with independent, normally distributed errors. These models are all built on the same level-1 equation, Eq. **29,** a model linear in log(temperature), but each makes different assumptions for the random effects. A centering constant of $\log \mathrm{T_c} = \log 15$ is used for all four models.

$$\text{Level 1: } \log\left(PLD_{ij}\right) = \beta_{0i} + \beta_{1i}\left(\log T_{ij} - \log \mathrm{T_c}\right) + \varepsilon_{ij} \qquad \textbf{[29]}$$

Model D is the clear winner in terms of AIC. Boundary-adjusted likelihood ratio tests (described in the legend to SI Table 18) can also be used to demonstrate that Model D is a significant improvement over model B ($p < .0001$), which in turn is a significant improvement over Model A ($p < .0001$).

As with ordinary regression models, residual plots can be used to examine whether the structural specification of the model is correct. In particular we can assess whether the assumption that log(PLD) is linearly related to log(temperature) is tenable. SI Fig. 13 plots the level-1 residuals (estimates of ε_{ij}) versus the predictor temperature, both on a log-transformed scale and an arithmetic scale. A locally weighted regression curve (lowess) is included to better detect any systematic trend in the scatter. Both plots show that the model overestimates log(PLD) at both low and high temperatures suggesting that the multilevel model could be improved by including a quadratic term in the level-1 model.

As a result we modify the level-1 model of Eq. **29** to the following.

[††] A centering constant of log 38 yields predicted random effects with the smallest amount of correlation. This value is also beyond the range of the data. The maximum temperature in our data set is 32ºC.

[‡‡] This choice may seem counterintuitive. Why not use a centering constant that makes the random effects uncorrelated? For graphical displays (such as SI Fig. 11a) and for basic interpretation a model with uncorrelated random effects is ideal. But if the purpose is to compare models then centering is far less important because the model diagnostics for centered and uncentered models are the same (Section IIIB). Still centering can play a role even here because, as noted in the introduction to Section III, the algorithms used to estimate mixed models can fail to converge when correlations between the different random effects are high. We've chosen log 15 as a centering constant because it facilitates convergence in all the models we fit, it reduces the correlation enough so that the different random effects are not totally confounded (Section IIIC), and it allows us to interpret model parameters with reference to a meaningful temperature value (Section IIIA).

Table 14. Fit statistics for the four models of SI Table 10

Model	Model description	No. of parameters	Loglikelihood	AIC
A	No random effects	3	–263.51	533.02
B	Random intercepts	4	–103.72	215.45
C	Random slopes	4	–241.53	491.06
D	Random slopes and intercepts	6	–84.32	180.63

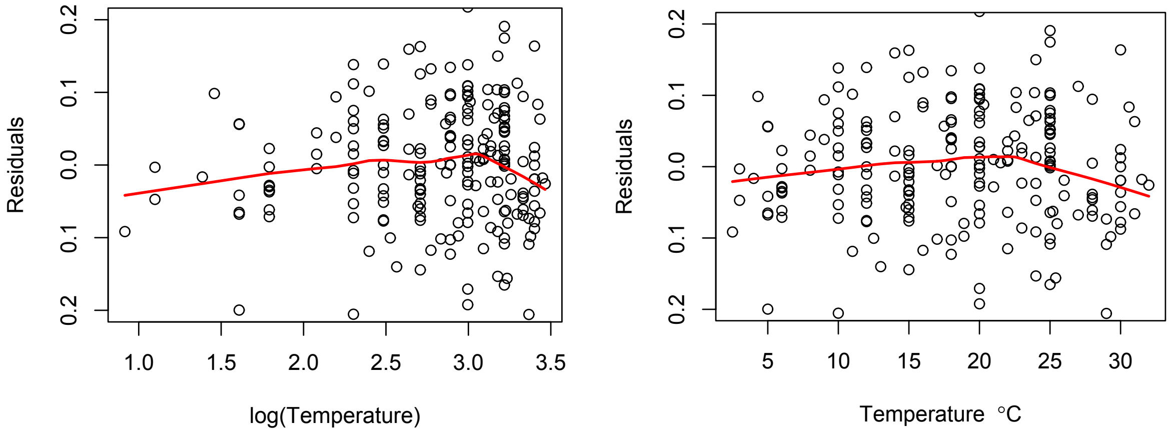

Fig. 13. Plot of level-1 residuals versus temperature on a log scale and an arithmetic scale for Model D in SI Table 10. The superimposed lowess curve reveals a quadratic pattern to the scatter in both plots.

Model E

$$\log(PLD_{ij}) = \beta_{0i} + \beta_{1i}(\log T_{ij} - \log \mathrm{T_c}) + \beta_{2i}(\log T_{ij} - \log \mathrm{T_c})^2 + \varepsilon_{ij} \qquad \textbf{[30]}$$

with level-2 model

$$\beta_{0i} = \beta_0 + u_{0i}$$
$$\beta_{1i} = \beta_1 + u_{1i}$$
$$\beta_{2i} = \beta_2$$

Comparing Model E to Model D we find that including a quadratic term yields a significant improvement (SI Table 15). In addition, the residual plot no longer shows any systematic trend (SI Fig. 14).

How should this latest modification of the level-1 model be interpreted? Although an allometric power law $y = aT^b$ can be a useful starting point in some instances, it is often found in practice that a constant power law relationship is an inadequate description of relative growth for biological data. A growth process that satisfies an allometric relationship but only if one or more parameters are allowed to vary is said to exhibit complex allometry (30). Bervian *et al.* (31) review the various modeling strategies for dealing with complex allometry. One of the simplest approaches (32) is to assume that the allometric exponent can be expressed as $b = f(T)$ for some function *f*. It is this form of complex allometry that is assumed in Eq. **30** as we now demonstrate.

Suppressing the centering constant $\log \mathrm{T_c}$ for simplicity, consider again Eq. **21** but this time extended to include a term quadratic in log(temperature) as in Eq. **30**.

$$\log Y = \log a + b \log T + c(\log T)^2$$
$$= \log a + (b + c \log T)\log T$$

Exponentiating both sides yields the following.

Table 15. Testing for a quadratic term in the level-1 equation

Model	df	AIC	Loglikelihood	Likelihood ratio	*p*
D	6	180.63	–84.32		
E	7	166.28	–76.14	16.35	<0.0001

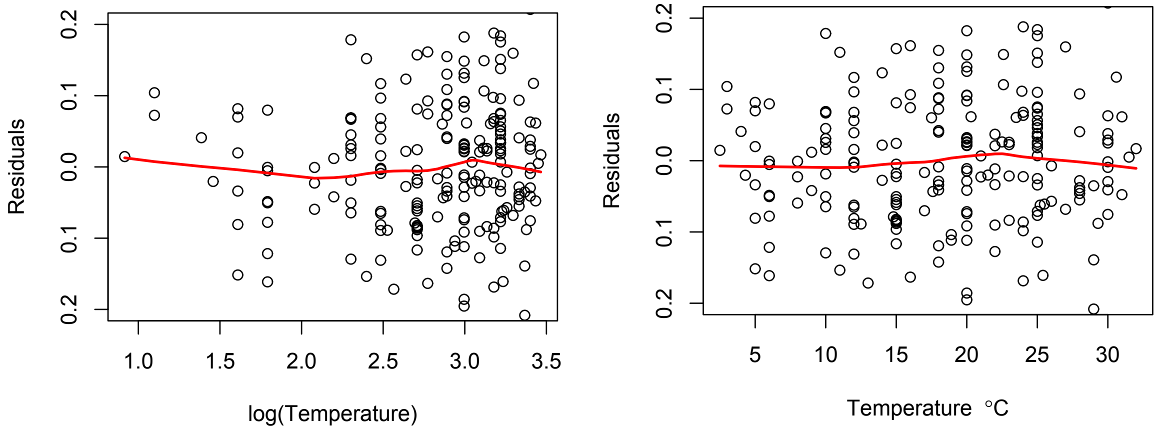

Fig. 14. Plot of level-1 residuals versus temperature on a log scale and an arithmetic scale for Model E in which the level-1 model includes a quadratic term. The superimposed lowess curve fails to reveal any pattern to the scatter in either plot.

Table 16. Testing for a quadratic term random effect

Model	df	AIC	Loglikelihood	Likelihood ratio	p
E	7	166.28	–76.14		
F	10	163.87	–71.94	8.41	0.027

$$
\begin{aligned}
Y &= \exp\left[\log a + \left(b + c\log T\right)\log T\right] \\
&= \exp\left(\log a\right)\cdot \exp\left[\left(b + c\log T\right)\log T\right] \\
&= a\exp\left[\log T^{\left(b + c\log T\right)}\right] \\
&= aT^{\left(b + c\log T\right)}
\end{aligned}
\qquad \textbf{[31]}
$$

By adding a quadratic term in $\log T$ to the log-transformed allometric equation shown in Eq. **21**, the constant b is replaced by the function $f(T) = b + c\log T$ to yield what is now the complex allometric equation, Eq. **31**.

Because the level-1 model has been modified to include a new regressor and its associated parameter, new modeling possibilities arise at level 2. In addition to permitting intercepts and/or linear terms to be random, a random quadratic term is also a possibility. We call this Model F.

Model F

Level 1:

$$
\begin{aligned}
\log\left(PLD_{ij}\right) &= \beta_{0i} + \beta_{1i}\left(\log T_{ij} - \log \mathrm{T_c}\right) \\
&\quad + \beta_{2i}\left(\log T_{ij} - \log \mathrm{T_c}\right)^2 + \varepsilon_{ij}
\end{aligned},
$$

$$
\varepsilon_{ij} \sim N\left(0, \sigma^2\right)
$$

Level 2:

$$
\begin{aligned}
\beta_{0i} &= \beta_0 + u_{0i} \\
\beta_{1i} &= \beta_1 + u_{1i}\;, \\
\beta_{2i} &= \beta_2 + u_{2i}
\end{aligned}
$$

$$
\text{where } \begin{bmatrix} u_{0i} \\ u_{1i} \\ u_{2i} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_0^2 & \tau_{01} & \tau_{02} \\ \tau_{01} & \tau_1^2 & \tau_{12} \\ \tau_{02} & \tau_{12} & \tau_2^2 \end{bmatrix}\right)
$$

Comparing Model E and Model F we obtain the results shown in SI Table 16.[§§] Although the evidence isn't strong, using either AIC or significance testing to compare models we should prefer a quadratic model in which all three level-1 parameters are allowed to be random at level 2.

V. Other Aspects of Model Fit

As with any regression model, residual analysis can be used to test model assumptions. One difference with multilevel models is that there is more than one kind of residual to

[§§] The legend of Table 18 explains how to determine the p-value that is reported here for a sequential likelihood ratio test comparing two nested models with different variance components.

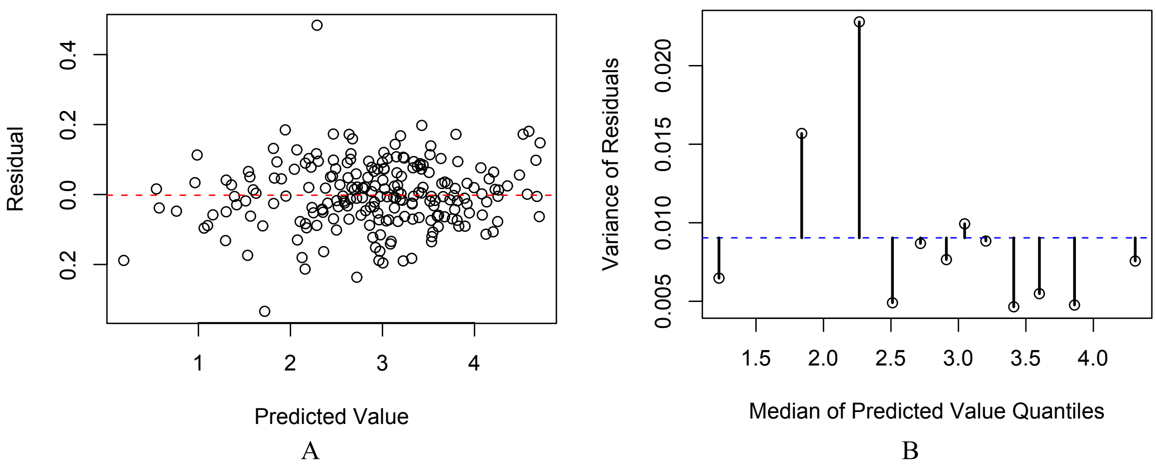

Fig. 15. Examining the level-1 residuals for heteroscedasticity. (A) Scatter plot of residuals against model predicted values. (B) Variance of residuals within 12 quantile groupings of the predicted values. The horizontal line denotes the residual variance calculated using all the data.

consider. The level-1 residuals estimate ε_{ij} and are assumed to be independent, normally distributed random variates with constant variance. The level-2 residuals of model F are empirical Bayes predictions of u_{0i}, u_{1i}, and u_{2i} and are assumed to have a joint multivariate normal distribution. Because the level-2 residuals are always confounded with the level-1 residuals (ref. 4, p. 132), a level-1 analysis is typically more useful as a diagnostic tool.

We've already used the level-1 residuals to investigate the structural form of the level-1 model (Section IV). Here we'll examine the assumptions of homoscedasticity and normality. The plot in SI Fig. 15A of the residuals versus the predicted values for Model F reveals no obvious change in variability from left to right. SI Fig. 15B explores this further. Here the residuals have been placed into twelve equal-sized groups whose boundaries are defined by quantiles of the predicted values,

$$q_{\min},\ q_{1/12},\ q_{1/6},\ldots,\ q_{11/12},q_{\max}.$$

Within each group the variance of the corresponding residuals has been calculated. Once again there's no obvious trend (linear and quadratic regressions are not significant, $p >$ 0.19).

SI Fig. 16A is a normal probability plot of the level-1 residuals along with a 95% confidence envelope. With 214 level-1 residuals we'd expect on average roughly 11 observations to fall outside of these bands. In the plot there is only one residual that plots on or outside the bands.

Turning to the level-2 residuals, SI Figs. 16B–16D display univariate normal probability plots for the three different sets of predicted random effects. Univariate normality doesn't guarantee joint multivariate normality but it can be used to locate obvious problems. Based on the plots neither the predictions of u_{2i}, SI Fig. 16D, nor the predictions of u_{1i}, SI Fig. 16C, show any obvious problems, but the predictions of u_{0i}, SI Fig. 16B, are highly deviant.

Model assumptions for the level-2 residuals are expected to hold only after all relevant explanatory variables and parameters have been included. Since at this point we have included no predictors in the level-2 equations it's not surprising that problems are seen in the residual plots. These most likely indicate that systematic differences among the species exist that are not accounted for by our simple probability model of a common mean of zero for each random effect. The fact that the problems appear to be most severe with random effects for the intercept and not at all for the random effects for the linear and quadratic terms will prove to be important in Section VII where we examine the level-2 residuals more closely in light of our thesis that a common planktonic larval duration time model with respect to temperature holds for all planktonic species.[¶¶]

[¶¶] A thorough development of a level-2 explanatory model for the level-1 parameters will be detailed elsewhere (O'Connor *et al.*, in preparation).

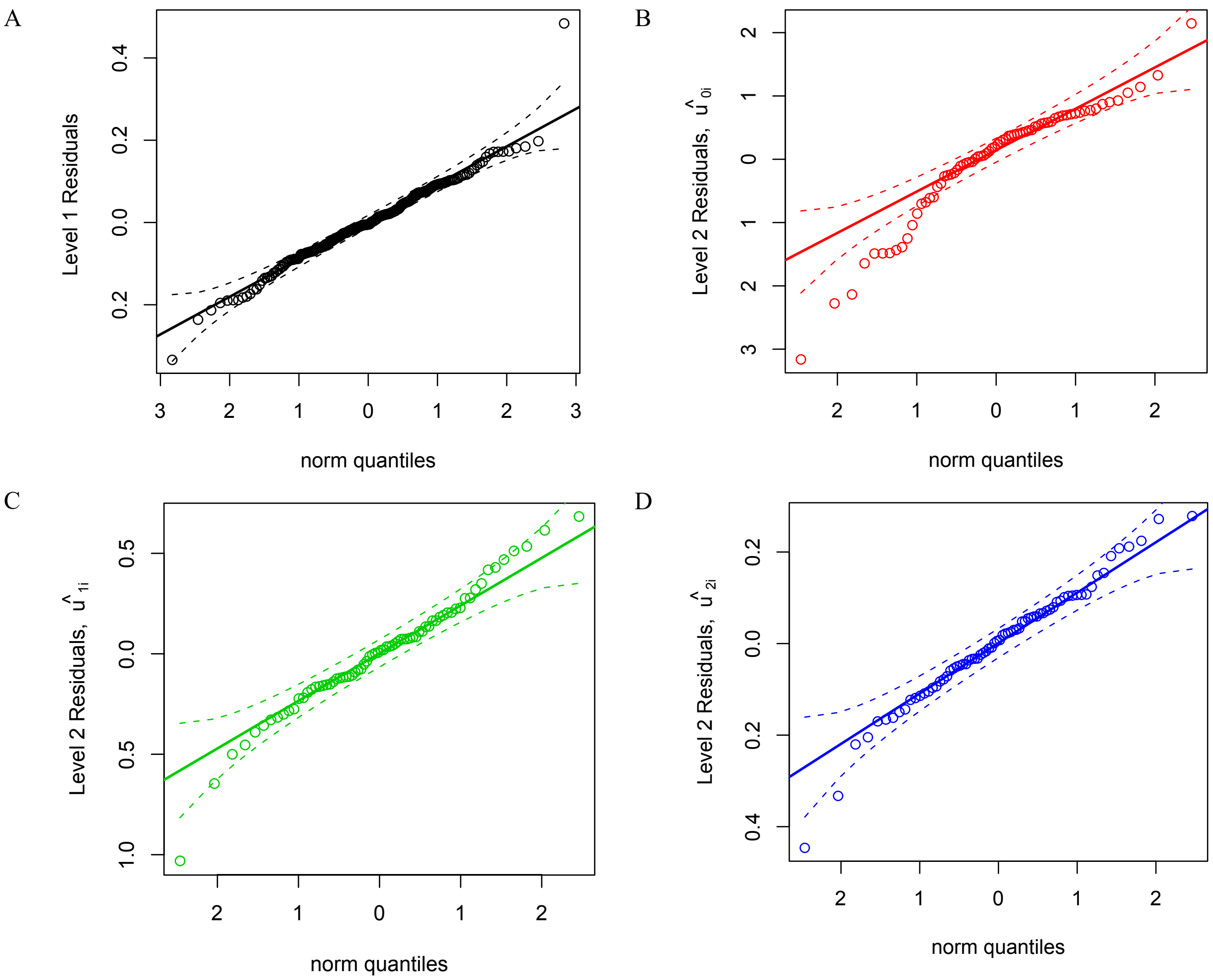

Fig. 16. Univariate normal probability plots of (A) level-1 residuals and predictions of the level-2 residuals for (B) u_{0i}, (C) u_{1i}, and (D) u_{2i}.

VI. The Influence and Fit of Level-2 Units on the Log(PLD) Quadratic Model

In their chapter on "Assumptions of the Hierarchical Linear Model", Snijders & Bosker (ref. 4, pp. 134–139), give a protocol for identifying those level-2 units that are poorly described by a fitted model. They recommend looking at two criteria, influence and fit, and argue that worrisome level-2 units are those that simultaneously have a large effect on model estimates (big influence) and are poorly described by the model (bad fit). In this section we summarize their approach and apply it to our model (Model F). We use their notation in what follows.

A. The Influence of Level-2 Units

To describe the influence of the i^{th} level-2 unit Snijders & Bosker (4) propose using a weighted average of C_i^F and C_i^R, defined as the influence of the i^{th} level-2 unit on the model's fixed effects and random effects, respectively. Each is calculated similarly. Let $\boldsymbol{\gamma}$ be the vector of fixed effects. This consists of β_0, β_1, and β_2, from the PLD quadratic model, model F. Let $\boldsymbol{\varphi}$ be the vector of random effects parameters consisting of the level-2 variances τ_0^2, τ_1^2, and τ_2^2, the level-2 covariances of the random effects, τ_{01}, τ_{02}, and τ_{12}, and the level-1 variance σ^2. We fit model F with and without the i^{th} level-2 unit and calculate the differences in the estimates of the fixed and random effects that result. If $\hat{\boldsymbol{\gamma}}_{(-i)}$ and $\hat{\boldsymbol{\varphi}}_{(-i)}$ are the vectors of estimates obtained without the i^{th} level-2 unit, we then compute $\hat{\boldsymbol{\gamma}}-\hat{\boldsymbol{\gamma}}_{(-i)}$ and $\hat{\boldsymbol{\varphi}}-\hat{\boldsymbol{\varphi}}_{(-i)}$. If the i^{th} level-2 unit is not influential on model fit, then the observed deviations in these vectors should look like random noise.

Let $\hat{\Sigma}_F$ and $\hat{\Sigma}_R$ be the estimated variance-covariance matrices for the fixed and random effects estimates obtained using all the data (without omitting any level-2 units). We could extract the diagonal elements from these matrices and use them to standardize the observed deviations by dividing them by the square root of their respective variances thus forming *z*-scores. Large *z*-scores would indicate deviations that are unusual for random noise. To get a single influence score for a level-2 unit we could then square the individual *z*-scores of each parameter (to prevent cancellation due to sign differences) and add them up. But because the estimates are correlated merely dividing by the square root of the variance is not enough. We need to use the entire variance-covariance matrix in the standardization process constructing what's called a quadratic form. The relevant quadratic forms here are $\frac{1}{r+1}\left(\hat{\boldsymbol{\gamma}}-\hat{\boldsymbol{\gamma}}_{(-i)}\right)^T\hat{\Sigma}_F^{-1}\left(\hat{\boldsymbol{\gamma}}-\hat{\boldsymbol{\gamma}}_{(-i)}\right)$ for the fixed effects and $\frac{1}{q}\left(\hat{\boldsymbol{\varphi}}-\hat{\boldsymbol{\varphi}}_{(-i)}\right)^T\hat{\Sigma}_R^{-1}\left(\hat{\boldsymbol{\varphi}}-\hat{\boldsymbol{\varphi}}_{(-i)}\right)$ for the random effects. Here r is the number of predictors in the model and q is the number of estimated variance components and correlations.

These expressions are what Snijders & Bosker (4) call C_i^F and C_i^R respectively. In order to have a single summary statistic, Snijders & Bosker (4) recommend averaging these two quantities yielding what they call C_i.

$$C_i=\frac{1}{r+q+1}\left((r+1)C_i^F+qC_i^R\right)$$

They don't recommend formally testing C_i but instead suggest examining the empirical distribution of C_i for all the level-2 units and flagging those units that have unusually large values.

B. The Fit of Level-2 Units

To assess fit, Snijders & Bosker (4) recommend using the level-1 residuals, $y_{ij}-\hat{y}_{ij}$, of each level-2 unit. Here $\hat{y}_{ij}$ is the estimate of y_{ij} obtained using only the estimated fixed effects. Let $\mathbf{y}_i$ be the vector of observed values for the i^{th} level-2 unit and let $\hat{\mathbf{y}}_i$ be its vector of estimated values. Let $\hat{\Sigma}_{y_i}$ be the estimated variance-covariance matrix for the observations coming from the i^{th} level-2 unit. (Note: This matrix is denoted $\mathbf{V}_h$ in the Appendix of this document where it is used in a different context. A sample estimate $\hat{\Sigma}_{y_i}$ can be obtained by using the theoretical formula given there and replacing the theoretical quantities by their corresponding sample estimates.)

Construct the quadratic form $S_i^2=\left(\mathbf{y}_i-\hat{\mathbf{y}}_i\right)^T\hat{\Sigma}_{y_i}^{-1}\left(\mathbf{y}_i-\hat{\mathbf{y}}_i\right)$ as a scaled measure of deviation. S_i^2 is called a standardized multivariate residual and has a chi-square distribution with m_i degrees of freedom, the number of observations made on the i^{th} level-2 unit, and can be used in a lack-of-fit test. A small *p*-value for this test is taken as evidence of lack-of-fit. Because the test is repeated for each level-2 unit thus inflating the nominal Type I error, Snijders & Bosker (4) recommend carrying out the tests using an adjusted α-level based on the Bonferroni correction.

C. Combining Level-2 Unit Influence and Fit

We calculate these diagnostics for the PLD data set. The model being used is Model F:

$$\text{Level 1: } \log\left(PLD_{ij}\right)=\beta_{0i}+\beta_{1i}\left(\log T_{ij}-\log \mathrm{T_c}\right)+\beta_{2i}\left(\log T_{ij}-\log \mathrm{T_c}\right)^2+\varepsilon_{ij}$$

$$\text{Level 2: } \beta_{0i}=\beta_0+u_{0i}$$
$$\beta_{1i}=\beta_1+u_{1i}$$
$$\beta_{2i}=\beta_2+u_{2i}$$

with $\varepsilon_{ij}\sim N\left(0,\sigma^2\right)$ and

$$\begin{bmatrix}u_{0i}\\u_{1i}\\u_{2i}\end{bmatrix}\sim N\left(\begin{bmatrix}0\\0\\0\end{bmatrix},\begin{bmatrix}\tau_0^2&\tau_{01}&\tau_{02}\\\tau_{01}&\tau_1^2&\tau_{12}\\\tau_{02}&\tau_{12}&\tau_2^2\end{bmatrix}\right).$$

A plot of C_i versus the *p*-value of the standardized multivariate residual for each level-2 unit is shown in SI Fig. 17. The two displayed

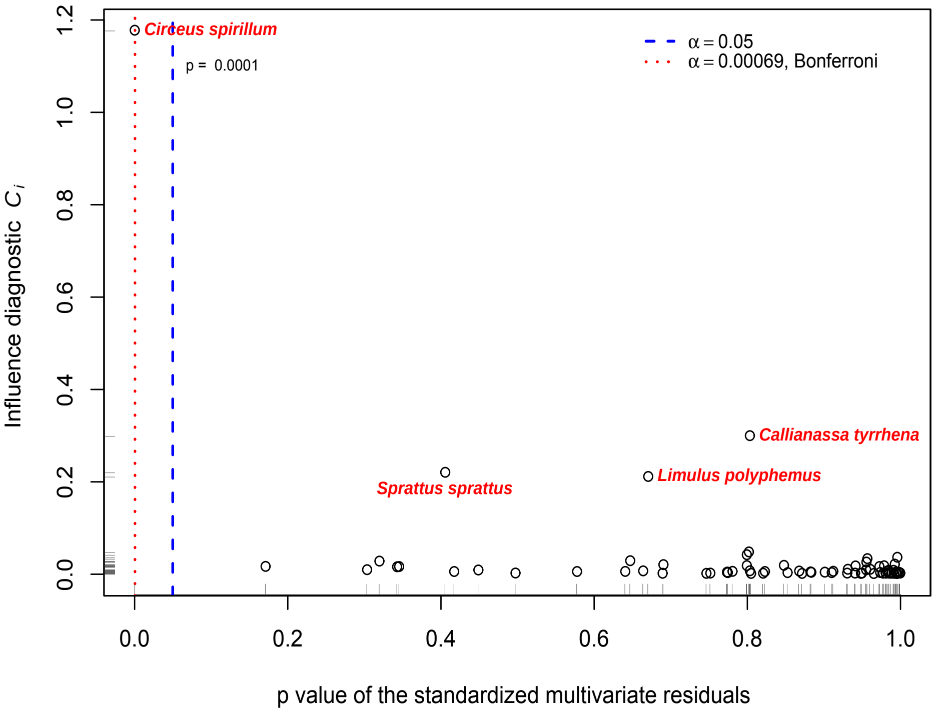

Fig. 17. The influence and fit of level-2 units. A diagnostic plot for level-2 units using the methodology described in ref. 4. The *p*-value displayed on the *x*-axis corresponds to a goodness of fit test. Level-2 units with large influence diagnostics and significant *p*-values should be viewed as problematic. Only one species, *Circeus spirillum*, satisfies both of these criteria.

vertical lines indicate the usual significance level α = .05 (blue and dashed) and the Bonferroni-adjusted α-level for 72 tests, $\alpha = {}^{.05}\!/_{72} = .00069$, (red and dotted) corresponding to the 72 different level-2 units.

The marginal rug plot on the *y*-axis shows that there are four species whose C_i influence scores might be unusual. They are identified by name on the plot. All four of these would be deemed extreme outliers (in that they would be located beyond the outer fence in a box plot).[‖‖] Based on the *p*-values displayed on the *x*-axis, only one of these four also exhibits a significant lack of fit (when tested at α = .05 and also at the Bonferroni-adjusted α-level). Snijders & Bosker (4) require a level-2 unit to be both influential and exhibit a significant lack of fit in order to be considered worrisome. Following them we would conclude that there is only one level-2 unit to worry about, *Circeus spirillum*.

SI Table 17 lists the fixed and random effects parameter estimates obtained for the four species that yielded the largest values of C_i. Also listed are C_i^F, C_i^R, and the *p*-value for the multivariate residual lack-of-fit test. The last column of the table contains the corresponding parameter estimates for the full model, a model in which all species are included.

From the table we see that what distinguishes *Circeus spirillum* from the rest is that it has a large impact on the estimate of β_2. When this species is included β_2 decreases roughly 12% (from –0.238 to –0.270). No other species has a comparable effect on any of the fixed effects. (This is also apparent from the reported value of C_i^F which is larger for *Circeus spirillum* than for any other species although *Sprattus sprattus* comes close.) All four of the species do have large effects on one or more of the variance components, but these effects are harder to interpret. It's worth noting that *Limulus polyphemus* has a very large effect on the estimate

[‖‖] The outer fences in a box plot occur at $q_{.25} - 3 \times \text{IQ}$ and $q_{.75} + 3 \times \text{IQ}$ where $q_{.25}$ and $q_{.75}$ are the first and third quartiles and IQ is the interquartile range, $q_{.75} - q_{.25}$.

Table 17. Model results for species that are extreme outliers in the C_i distribution

Parameter	Omitted species				All species in model
	Circeus spirillum	*Callianassa tyrrhena*	*Limulus polyphemus*	*Sprattus sprattus*	
$\hat{\beta}_0$	3.230	3.208	3.160	3.253	3.203
$\hat{\beta}_1$	–1.401	–1.374	–1.393	–1.428	–1.404
$\hat{\beta}_2$	–0.238	–0.259	–0.257	–0.257	–0.270
$\hat{\tau}_0^2$	0.838	0.846	0.774	0.698	0.843
$\hat{\tau}_1^2$	0.199	0.127	0.106	0.168	0.154
$\hat{\tau}_2^2$	0.034	0.026	0.017	0.036	0.034
$\hat{\tau}_{01}$	–0.196	–0.180	–0.070	–0.148	–0.156
$\hat{\tau}_{02}$	–0.090	–0.077	0.014	0.005	–0.053
$\hat{\tau}_{12}$	0.082	0.056	0.039	0.069	0.071
$\hat{\sigma}^2$	0.013	0.019	0.018	0.017	0.021
C_i^F	0.217	0.066	0.054	0.187	—
C_i^R	0.746	0.172	0.398	0.542	—
C_i	0.553	0.134	0.273	0.413	—
p	0.0001	0.803	0.671	0.406	—
n	3	4	3	2	

of τ_2^2, the variance of the quadratic random effects, causing a roughly 50% increase in the estimate when it is included. This fact will loom large in Section VII where we evaluate the evidence for a common temperature model of planktonic larval duration.

SI Fig. 18 displays the individual regression "lines" computed using model F (a quadratic level-1 model with random effects for each coefficient) in which the influential species shown in SI Fig. 17 and listed in SI Table 17 are identified. Individual trajectories are obtained from the composite equation that combines the estimated fixed effects and predicted random effects for individual species (the so-called empirical Bayes estimates). The population model, one that contains only the fixed effects portion of the multilevel model, is shown for comparison.

In the plot both *Limulus polyphemus* and *Circeus spirillum* look deviant, while *Callianassa tyrrhena* does not look particularly unusual. *Sprattus sprattus* looks as if it fits worse than *Circeus spirillum* until one realizes that its curve is based on only two observations. Sample sizes for the four most influential species are given in the last row of SI Table 17.

VII. Caterpillar Plots and the Search for a Common Temperature Model

An extremely useful graphical tool for exploring the relationships between level-2 units (species in our model) is the caterpillar plot, a terminology used in the multilevel modeling package MLWin (ref. 33, p. 39). The theory for these plots was developed by Goldstein and co-workers (1, 34–35). A caterpillar plot displays 95% confidence intervals for the predicted level-2 residuals for a given parameter plotted against the rank order of the point predictions. Typically the 95% confidence intervals are adjusted for multiple testing so that pairwise tests can be carried out and deemed significant at the 5% level. Formulas for calculating the standard errors are described in refs. 1 and 36. Specific details of these calculations for our data are presented in the Appendix of this document.

SI Figs. 19A–19C display caterpillar plots for the three predicted level-2 residuals for a quadratic level-1 model with random effects for each coefficient, model F. Each plot displays one of $\hat{u}_{2i}$, $\hat{u}_{1i}$, or $\hat{u}_{0i}$ in rank order from smallest to largest, along with error bars representing 95% confidence intervals for each random effect not

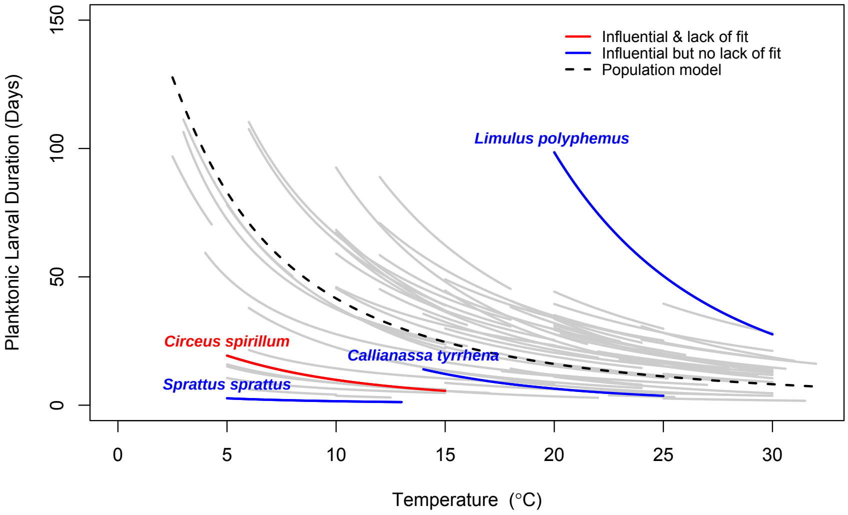

Fig. 18. Locations of influential species (as determined by C_i) among all estimated trajectories. The individual species trajectories are computed from the estimated fixed effects and predicted random effects obtained using Model F.

adjusted for multiple testing. Note: The striking similarity of SI Figs. 19A and 19B is due to the high correlation that exists between these two sets of random effects.

What's apparent from these plots is that while many of the confidence intervals for u_{0i} fail to overlap zero, SI Fig. 19C, nearly all of the confidence intervals for u_{1i} and u_{2i} do, SI Figs. 19A–19B. In fact the caterpillar plots suggest that there may only be a few species driving the need for random effects for the linear and quadratic terms. To test this we sequentially drop the most deviant species, as reflected in the caterpillar plots of SI Figs. 19A–19B, each time comparing the following three models that share the same level-1 model but differ in their level-2 models.

Model 1:	Random intercepts only
Model 2:	Random intercepts and linear coefficients
Model 3:	Random intercepts, linear coefficients, and quadratic coefficients

Likelihood ratio tests require nested models so we compare model 3 against model 2 and model 2 against model 1. The results of these tests are summarized in SI Table 18 where we report *p*-values for both the maximum likelihood (ML) and restricted maximum likelihood (REML) estimates of the test statistic.

Observe that by dropping a single species, *Limulus polyphemus*, the likelihood ratio test for the inclusion of a quadratic random effect is no longer significant at α = .05. Based on significance testing alone we should favor a random intercept and random linear coefficients model (model 2) over either of the other two. AIC agrees with this ranking. Recall from Section VI that *Limulus polyphemus* was flagged as an influential level-2 unit. If the second and third most deviant species in the caterpillar plots, *Laqueus californianus* and *Callianassa tyrrhena*, are also dropped, then neither likelihood ratio test is significant, causing us to prefer the random intercepts model, model 1. Using AIC leads to a similar conclusion (the random intercepts model, model 1, and the random intercepts and linear coefficients model, model 2, are essentially tied). Dropping additional deviant species continues to yield results that support model 1.

Thus using 69 of the 72 species we conclude that a temperature model with constant linear and quadratic terms, but random intercepts, adequately describes the relationship between PLD and temperature. This supports the hypothesis that the PLD-temperature relationship is uniform across most species. Further analysis reveals that additional simplification is not possible. Random intercepts must be retained in the model. Continuing with the protocol of SI Table 18 we would need to remove 82% of the most deviant species shown in SI Fig. 19C before a constant intercept model would be preferred over a random intercepts model (details not shown).

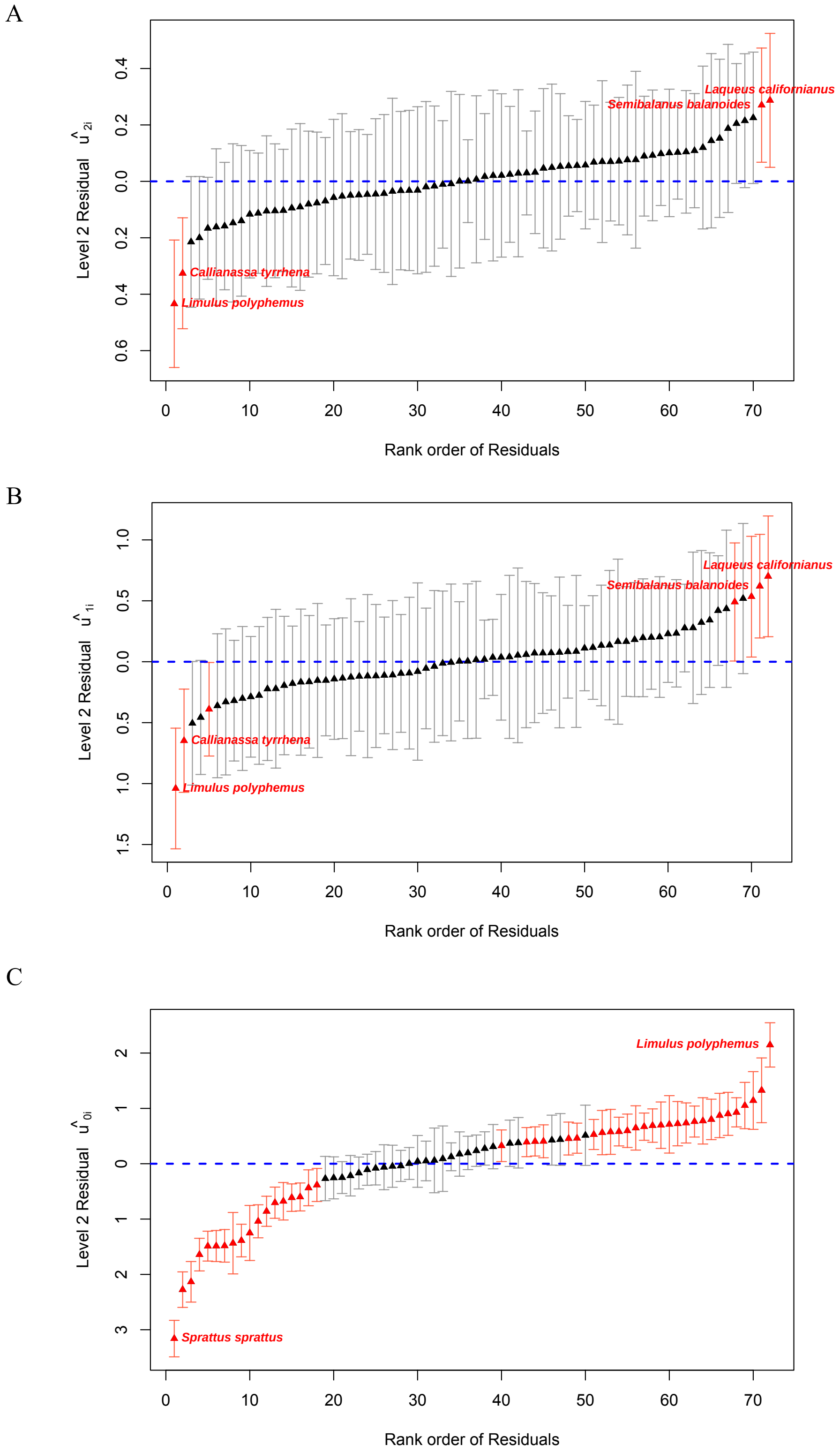

Fig. 19. Caterpillar plots for the three sets of random effects, (A) u_{2i}, (B) u_{1i}, and (C) u_{0i}

Table 18. Model results after dropping the most aberrant species shown in caterpillar plots (SI Figs. 19A and 19B).

Omitted species	Model	Random effects included in model	AIC	Log-likelihood (ML)	LR test *p*		AIC-best model
					(ML)	(REML)	
None	3	u_{0i} , u_{1i} , & u_{2i}	163.87	–71.93	–	–	✱
	2	u_{0i} & u_{1i}	166.28	–76.14	0.027	0.035	
	1	u_{0i}	172.21	–81.10	0.004	0.003	
Limulus polyphemus	3	u_{0i} , u_{1i} , & u_{2i}	150.00	–65.00	–	–	
	2	u_{0i} & u_{1i}	148.29	–67.15	0.174	0.214	✱
	1	u_{0i}	153.03	–71.51	0.008	0.005	
L. polyphemus & *Laqueus californianus*	3	u_{0i} , u_{1i} , & u_{2i}	143.76	–61.88	–	–	
	2	u_{0i} & u_{1i}	142.95	–64.47	0.117	0.189	✱
	1	u_{0i}	143.75	–66.88	0.059	0.037	
L. polyphemus, *L. californianus*, & *Callianassa tyrrhena*	3	u_{0i} , u_{1i} , & u_{2i}	136.22	–58.11	–	–	
	2	u_{0i} & u_{1i}	132.00	–59.00	0.515	0.554	✱
	1	u_{0i}	132.00	–61.00	0.090	0.067	✱

The ✱ in the last column identifies the model(s) with the lowest AIC value in each block of three. The sequential LR test being used compares the current model with the model in the row immediately above it and hence tests the need for the omitted random effect. Because the null hypothesis is that one of the random effects has zero variance, zero being a value that lies on the boundary of the parameter space, the usual regularity conditions required in classical likelihood theory do not hold. As a result when standard likelihood ratio tests are used here the *p*-values that are obtained tend to be overestimated. The asymptotic distribution of the likelihood ratio statistic (LR) for a comparison of models differing in a single random effect is better represented as a mixture of chi-squared distributions with *p*-value given by $\frac{1}{2}P\left(\chi_k^2 > \text{LR}\right) + \frac{1}{2}P\left(\chi_{k+1}^2 > \text{LR}\right)$, where *k* is the number of random effects in the "smaller" model. Simulations suggest that rejection proportions based on the REML likelihood ratio test statistic come closer to achieving the prescribed nominal significance level than those based on the ML test statistic. We follow the guidelines in ref. 10, pp. 70–71, and ref. 37 for carrying out these tests. The original derivation of this test can be found in refs. 38–40. A dissenting viewpoint on the reliability of this test is ref. 41.

VIII. Visualizing the Final Model

SI Table 19 displays the estimates for a level-1 quadratic model in which only the intercepts are allowed to be random (model 1 of Section VII). The composite version of this model, designated Model G, is also shown.

SI Fig. 20 shows the model predictions for individual species superimposed on the raw data. The common population-averaged trajectory that is displayed in each panel is calculated using only fixed effects. The model-based empirical Bayes trajectories combine the fixed effects with the predicted random effects (empirical Bayes estimates), here just $\hat{u}_{0i}$ for the intercept, and thus vary from panel to panel.

Model G:

$$\log\left(PLD_{ij}\right) = \beta_0 + u_{0i} + \beta_1\left(\log T_{ij} - \log \mathrm{T_c}\right) + \beta_2\left(\log T_{ij} - \log \mathrm{T_c}\right)^2 + \varepsilon_{ij}$$

$$u_{0i} \sim N\left(0,\, \tau^2\right),\ \varepsilon_{ij} \sim N\left(0,\, \sigma^2\right)$$

Table 19. Parameter estimates for a quadratic level-1 model with random intercepts

Parameter	Estimate	Standard error
β_0	3.167	0.107
β_1	–1.344	0.046
β_2	–0.276	0.041
σ^2	0.0230	–
τ^2	0.7530	–

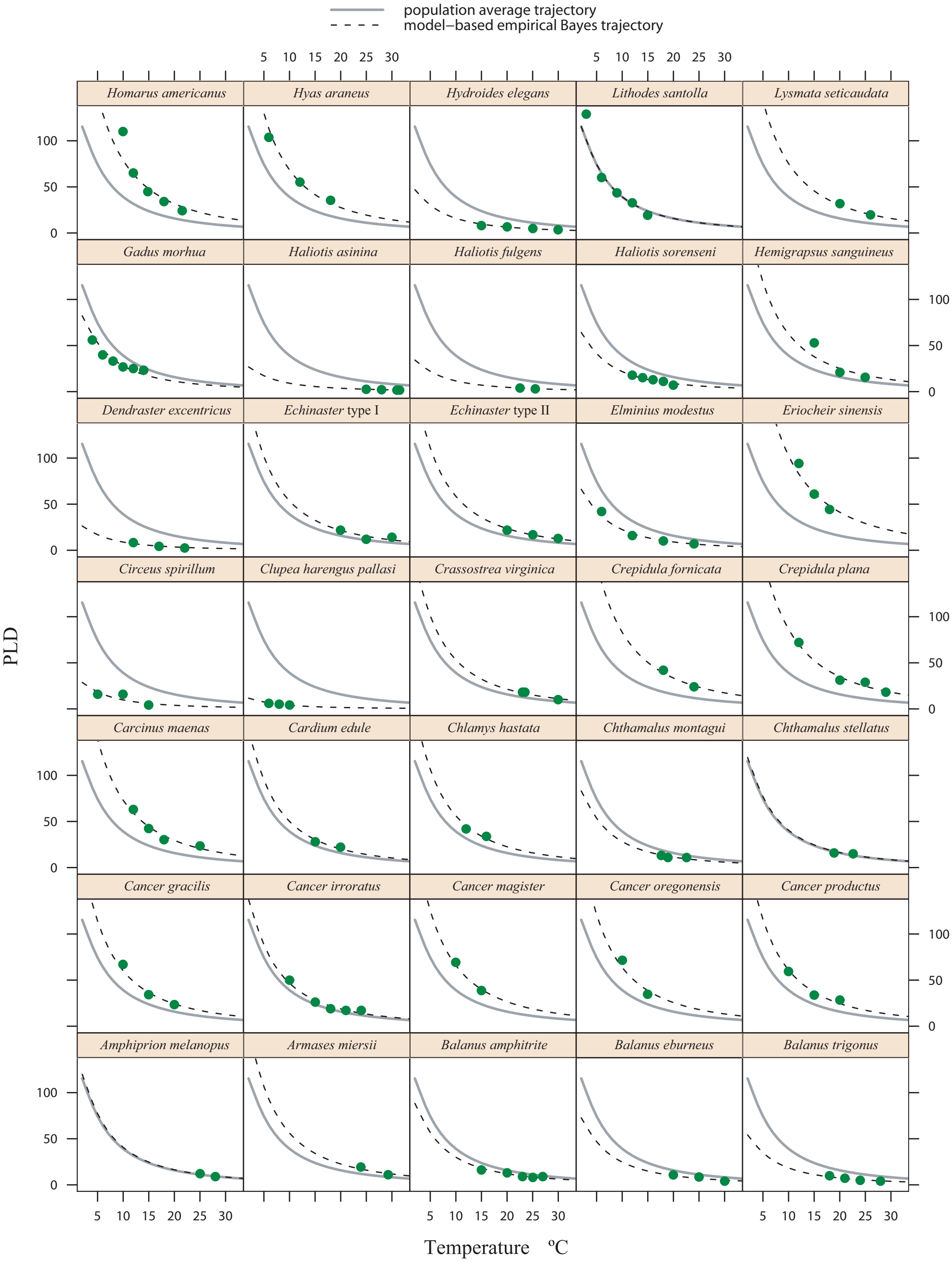

Fig. 20. (Part 1) The fitted model for individual species *Amphiprion melanopus—Lysmata seticaudata*. Population-averaged (fixed effects only) and empirical Bayes (fixed and random effects) fitted curves are shown (using the random intercepts model, Model G, p. 21, of this *Supporting Text*) as well as the raw data values.

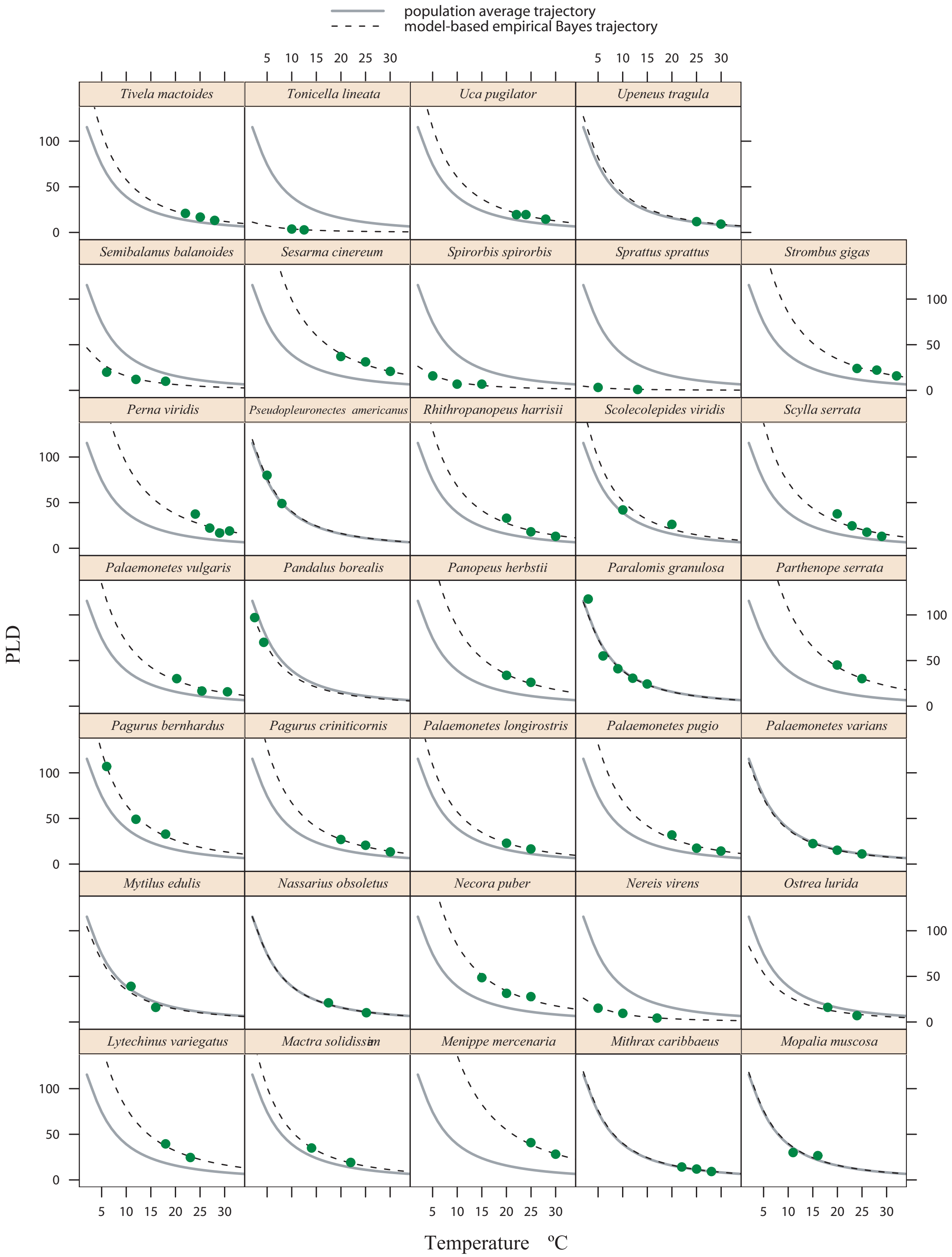

Fig. 20. (Part 2) The fitted model for individual species *Lytechinus variegatus—Upeneus tragula*. Population-averaged (fixed effects only) and empirical Bayes (fixed and random effects) fitted curves are shown (using the random intercepts model, Model G, p. 21, of this *Supporting Text*) as well as the raw data values.

IX. References Cited

1. Goldstein H (1995) *Multilevel Statistical Models* (Edward Arnold, London).
2. Kreft IGG, de Leeuw J (1998) *Introducing Multilevel Modeling* (Sage Publications, Thousand Oaks, CA).
3. Raudenbush SW, Bryk AS (2002) *Hierarchical Linear Models* (Sage Publications, Thousand Oaks, CA).
4. Snijders TAB, Bosker R (1999) *Multilevel Analysis* (Sage Publications, Thousand Oaks, CA).
5. Leyland AH, Goldstein H (2001) *Multilevel Modelling of Health Statistics* (Wiley, New York).
6. Singer JD, Willett JB (2003) *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence* (Oxford University Press, Oxford, UK).
7. Hox JJ (2002) *Multilevel Analysis: Techniques and Applications* (Lawrence Erlbaum, Mahwah, NJ).
8. Brown H, Prescott R (1999) *Applied Mixed Models in Medicine* (Wiley, New York).
9. Pinheiro JC, Bates DM (2000) *Mixed-Effects Models in S and S-Plus* (Springer, New York).
10. Verbeke G, Molenberghs G (2000) *Linear Mixed Models for Longitudinal Data* (Springer, New York).
11. St-Pierre NR (2000) *J Dairy Sci* 84:741–755.
12. Hopkins WG (2004) *Sportscience* 8:20–24.
13. Arends LR, Vokó Z, Stijnen T (2003) *Stat Med* 22:1335–1353.
14. Gibson AJF, Myers RA (2003) *American Fisheries Society Symposium* 35:211–221.
15. Laird NM, Ware JH (1982) *Biometrics* 38:963–974.
16. Fitzmaurice GM, Laird NM, Ware JH (2004) *Applied Longitudinal Data Analysis* (Wiley, New York).
17. Hamel P, Magnan P, East P, Lapointe M, Laurendeau P (1997) *Can J Fish Aquat Sci* 54:190–197.
18. Gillooly JF, Charnov EL, West GB, Savage VM, Brown JH (2001) *Nature* 417:70–73.
19. Huxley JS (1924) *Nature* 14:896–897.
20. McArdle BH, Anderson MJ (2004) *Can J Fish Aquat Sci* 61:1294–1302.
21. Lee Y-W, Sampson, DB (2005) *Can J Fish Aquat Sci* 62:363–373.
22. Stow CA, Reckhow KH, Qian SS (2006) *Ecology* 87:1472–1477.
23. Siem E (1983) *J Theor Biol* 104:161–168.
24. Gingerich PD (2000) *J Theor Biol* 204:201–221.
25. Ott WR (1995) *Environmental Statistics and Data Analysis* (Lewis Publishers, Boca Raton, FL).
26. Wiens BL (1999) *Am Stat* 53:89–93.
27. McCullagh P, Nelder JA (1989) *Generalized Linear Models* (Chapman and Hall, London, UK).
28. R Development Core Team (2005) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria) ISBN 3-900051-07-0, http://www.R-project.org.
29. Healy MJR (2001) in *Multilevel Modelling of Health Statistics,* eds Leyland AH, Goldstein H (Wiley, New York), pp 1–12.
30. Jolicouer P (1989) *J Theor Biol* 140:41–49.
31. Bervian G, Fontoura NF, Haimovici M (2006) *J Fish Biol* 68:196–208.
32. Strauss RE (1993) in *Problems of Relative Growth*, ed. Huxley JS (John Hopkins University Press, Baltimore, MD), pp xlvii–lxxv.
33. Rasbash J, Steele F, Browne W, Prosser B (2005) *A User's Guide to MLwiN Version 2.0* (University of Bristol, London).
34. Goldstein H, Healy MJR (1995) *J R. Stat Soc Ser A* 158:175–177.
35. Goldstein H, Spiegelhalter DJ (1996) *J R. Stat Soc Ser A* 159:385–443.
36. Snijders TAB, Berkhof J (in press) Chapter 4 in *Handbook of Quantitative Multilevel Analysis,* eds. de Leeuw J, Kreft I (Kluwer, New York).
37. Visscher PM (2006) *Twin Res Hum Genet* 9:490–495.
38. Chernoff H (1954) *Ann Math Stat* 25:573–578.
39. Self SG, Liang KY (1987) *J Am Stat Assoc* 82:605–610.
40. Stram DO, Lee JW (1994) *Biometrics* 50:1171–1177.
41. Crainiceanu CM, Ruppert D (2004) *J R Stat Soc B* 66:165–185.

Appendix. The Variance of the Level-2 Residuals

This section provides additional details of the variance calculations that were used in constructing the caterpillar plots discussed in Section VII. Consider again the random intercepts, random linear coefficients, and random quadratic coefficients model, model F, that was used in constructing the caterpillar plots. This model is shown in multilevel form on p. 13 of this *Supporting Text* and is shown in composite form in Eq. **32** below.

$$\begin{aligned}\log\left(PLD_{ij}\right) = \beta_0 &+ \beta_1\left(\log T_{ij} - \log \mathrm{T_c}\right) \\ &+ \beta_2\left(\log T_{ij} - \log \mathrm{T_c}\right)^2 + u_{0i} \\ &+ u_{1i}\left(\log T_{ij} - \log \mathrm{T_c}\right) \\ &+ u_{2i}\left(\log T_{ij} - \log \mathrm{T_c}\right)^2 + \varepsilon_{ij} \qquad \textbf{[32]}\end{aligned}$$

The distribution of the ε_{ij} and the joint distribution of u_{0i}, u_{1i}, and u_{2i} are as described previously. Here i = 1 to 72, the number of species considered in our analysis and, for species i, j = 1 to m_i, where the number of observations m_i varies from species to species.

The 72 species yield a system of 214 equations that together comprise a composite model that can be written as a single matrix equation. The predictors for the fixed effects form what is called the design matrix for the model. To simplify notation, let $x_{ij} = \log T_{ij} - \log \mathrm{T_c}$. Formally i indicates the level-2 unit (species) and j the observation on that level-2 unit. For these data and the given model the design matrix **X** is the following 214 × 3 matrix.

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{11}^2 \\ 1 & x_{12} & x_{12}^2 \\ 1 & x_{21} & x_{21}^2 \\ 1 & x_{22} & x_{22}^2 \\ \vdots & \vdots & \vdots \\ 1 & x_{72,2} & x_{72,2}^2 \end{bmatrix}$$

Using the design matrix **X** the fixed effect portion of the composite model can be written as the following matrix product.

$$\mathbf{X\beta} = \begin{bmatrix} 1 & x_{11} & x_{11}^2 \\ 1 & x_{12} & x_{12}^2 \\ 1 & x_{21} & x_{21}^2 \\ 1 & x_{22} & x_{22}^2 \\ \vdots & \vdots & \vdots \\ 1 & x_{72,2} & x_{72,2}^2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}$$

Let **Z** denote the design matrix for the random effects part of the model. **Z** contains the predictors that multiply the random effects u_{0i}, u_{1i}, and u_{2i} in Eq. **32**. For model F it contains the same elements as **X**, just organized differently. Here is a portion of **Z** for the first three level-2 units (the first three species in alphabetical order).

$$\mathbf{Z} = \left[\begin{array}{ccc|ccc|ccc|cc} 1 & x_{11} & x_{11}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & x_{12} & x_{12}^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & 0 & 1 & x_{21} & x_{21}^2 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & x_{22} & x_{22}^2 & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & x_{31} & x_{31}^2 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & x_{35} & x_{35}^2 & 0 & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & & \ddots & \end{array}\right]$$

From the matrix we see that the first species has two temperature observations, the second also has two, and the third species has five. Using **Z** the random effects portion of the composite model can be written as the following matrix product.

$$\mathbf{Z}\mathbf{b} = \left[\begin{array}{ccc|ccc|cc} 1 & x_{11} & x_{11}^2 & 0 & 0 & 0 & 0 & \cdots \\ 1 & x_{12} & x_{12}^2 & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & 0 & 1 & x_{21} & x_{21}^2 & 0 & \cdots \\ 0 & 0 & 0 & 1 & x_{22} & x_{22}^2 & 0 & \cdots \\ \hline \vdots & \vdots & \vdots & & & & \ddots & \end{array}\right] \begin{bmatrix} u_{01} \\ u_{11} \\ u_{21} \\ u_{02} \\ u_{12} \\ u_{22} \\ \vdots \\ u_{2,72} \end{bmatrix}$$

We can simplify things further by writing **Z** as the following 214 × 216 block diagonal matrix.

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{Z}_{72} \end{bmatrix}$$

Here, for example,

$$\mathbf{Z}_1 = \begin{bmatrix} 1 & x_{11} & x_{11}^2 \\ 1 & x_{12} & x_{12}^2 \end{bmatrix}$$

and in general $\mathbf{Z}_h$ is the $m_h \times 3$ design matrix for the random effects of level-2 unit *h* (where m_h is the number of level-1 observations of level-2 unit *h*).

Let

$$\mathbf{\Omega} = \begin{bmatrix} \tau_0^2 & \tau_{01} & \tau_{02} \\ \tau_{01} & \tau_1^2 & \tau_{12} \\ \tau_{02} & \tau_{12} & \tau_2^2 \end{bmatrix},$$

the variance-covariance matrix of the random effects. If $\mathbf{Y}_h$ is the response vector for level-2 unit *h* and $\mathbf{V}_h$ is its variance-covariance matrix, then the variances and covariances of the elements of the response vector $\mathbf{Y}_h$ can be written succinctly as the following matrix expression.

$$\mathbf{V}_h = \mathrm{Var}(\mathbf{Y}_h) = \mathbf{Z}_h \mathbf{\Omega} \mathbf{Z}_h^T + \sigma^2 \mathbf{I}_{m_h}$$

Here $\mathbf{I}_{m_h}$ is the $m_h \times m_h$ identity matrix and $\sigma^2 = \mathrm{Var}(\varepsilon_{ij})$, a scalar. Arrange the vectors $\mathbf{Y}_h$ in level-2 unit order in the single response vector **Y**. Since observations from different level-2 units are independent, we can write $\mathrm{Var}(\mathbf{Y})$ as the following block-diagonal matrix **V**.

$$\mathbf{V} = \mathrm{Var}(\mathbf{Y}) = \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{V}_{72} \end{bmatrix}$$

Next form the matrix product $\mathbf{R}_h = \mathbf{Z}_h \mathbf{\Omega}$ and arrange the 72 $m_h \times 3$ matrices that result as the 214 × 216 block diagonal matrix **R**.

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{R}_{72} \end{bmatrix}$$

Finally form the 216 × 216 block-diagonal matrix **S** in which the blocks are identical each consisting of the matrix **Ω**.

$$\mathbf{S} = \begin{bmatrix} \mathbf{\Omega} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{\Omega} \end{bmatrix}$$

Having established the notation we can finally get to the formula of interest. The comparative (also called the conditional) variance of the level-2 residuals is given by the following formula (ref. 1, p. 42; ref. 33, p. 15).

$$\mathbf{S} - \mathbf{R}^T \mathbf{V}^{-1} \left(\mathbf{V} - \mathbf{X} \left(\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X} \right) \mathbf{X}^T \right) \mathbf{V}^{-1} \mathbf{R}$$