

Abstract

Global change may not only affect mean temperatures, but also their variability and predictability. Species vulnerability to extinction will thus depend on the evolvability of multiple strategies, including the evolution of means, phenotypic plasticity and bet-hedging strategies. The evolutionary potential of these three strategies is little studied, but current patterns of adaptation to local conditions may provide further information. We thus predicted that reaction norm shapes of populations from widely distributed species depend on environmental conditions. We conducted a meta-analysis on 57 studies of insect diapause and extracted 447 logistic reaction norms, and correlated means and variance components with climate parameters. Three patterns emerged: First, mean phenology correlated only weakly with mean winter onset, as populations at high latitudes failed to track early winter onsets. Secondly, the degree of canalization did not depend on winter severity, likely caused by strong study bias towards non-canalized genotypes. Thirdly, bet-hedging was rare and correlated only weakly with environmental predictability. Genotypes are thus most vulnerable to decreases in climate predictability.

Introduction

Anthropogenic greenhouse gas emissions change the environment at an unprecedented rate (IPCC 2014). Changes in temperature are especially problematic, because temperature is the most important abiotic variable to determine a species' niche (Clarke 2017). Some species profit from temperature changes and expand their phenology (Bell *et al.* 2015) or range (Monceau *et al.* 2014), but the majority of species faces higher extinction risks from anthropogenic climate change (Urban 2015). While rises in mean temperature are under intense scrutiny, they are not the only potential cause of climate-change induced biodiversity loss - increasing climate variability imposes further extinction risk (Bathiany *et al.* 2018). For example, increasing variability changes the thermal performance of ectotherms and represses the thermal optimum (Vasseur *et al.* 2014; Kingsolver & Buckley 2017). While the effects of changing means are relatively well researched, it is inherently more difficult to study effects of changing climate variability.

One of the most commonly observed responses to climate change is a shift in phenology (seasonal timing) (Parmesan & Yohe 2003), which may be either a plastic response to the environment or genetic adaptation (Gienapp *et al.* 2013; Lane *et al.* 2018). Standing levels of plasticity, however, reduce extinction risks only on the short term; on the long term only genetic adaptation can prevent extinction (Visser 2008).

Genetic adaptation to changing environments is traditionally thought to occur in four ways (e.g. Tufto 2015): First, the mean phenotype may shift so that the phenotype retains highest arithmetic mean fitness. For example, the photoperiodic response of the pitcher-plant mosquito *Wyeomyia smithii* has shifted towards shorter day lengths (later diapause) with global warming (Bradshaw & Holzapfel 2001). Secondly, a phenotype with lower arithmetic mean fitness but also lower long-term fitness variance may evolve (conservative bet-hedging, Seger & Brockmann 1987; Starrfelt & Kokko 2012), as the reduction in fitness variance increases geometric mean fitness (e.g. Childs *et al.* 2004). Thirdly, phenotypic plasticity may evolve: Development is often guided by predictive environmental cues, and when the nature of the cue – environment relationship changes, it imposes selection on the cue – response relationship (reaction norm). Lastly, variance in phenotype expression may evolve, such that reduced fitness correlations among the offspring minimize long-term fitness variance (diversified bet-hedging, Seger & Brockmann 1987; Starrfelt & Kokko 2012). These four modes of genetic adaptation affect the match between phenotype and environment and are hence important in a changing climate.

The four strategies are closely related. Starrfelt and Kokko (2012) noted that fitness optimization involves finding a trade-off among arithmetic mean optimization, conservative bet-hedging and diversified bet-hedging, but left the integration of phenotypic plasticity unattended. The distinction between plasticity and diversified bet-hedging is indeed not trivial, especially because there is substantial terminological confusion between plasticity as evolutionary strategy and plasticity as physiological process. We therefore recently integrated plasticity and related strategies in Starrfelt & Kokko's framework (Joschinski & Bonte 2019; Fig. 1). In our view the question which phenotype to express is a decision-making process, and hence inherently dichotomous. It can thus be represented by a binary reaction norm (Fig. 1A) in which the probability of expressing a phenotype is a function of an environmental cue, and the evolutionary strategies represent different shapes of such a reaction norm (Fig. 1B). Reaction norm shapes can be described along three axes: By the allocation of variance within vs among environments (Fig. 1B, x-axis), by the overall degree of phenotypic variance (Fig. 1B, y-axis), and by the mean (inflection point in logistic reaction norms; Fig. 1C). There are

three according axes of strategies, which we refer to as *variance avoidance*, *responsiveness*, and *information use*. These axes represent trade-offs among conservative bet-hedging and arithmetic mean optimization, fixed vs. variable development; and diversified bet-hedging vs. phenotypic plasticity. The classification of the strategies by their reaction norm shape allows predictions of their evolution in a changing climate.

The adaptive value of phenotypic plasticity, bet-hedging and arithmetic mean optimization depends on environmental conditions (Tufto 2015), so one would predict that reaction norm shapes are locally adapted. Climate change will rapidly alter local prevailing conditions, and it is unknown how quickly suitable evolutionary strategies can emerge. Nevertheless, some widespread species have undergone rapid adaptation upon range expansions to novel climates (Gomi & Takeda 1996), and these patterns of historical reaction norm evolution may provide information about future resilience to climate change. Taking these examples of local adaptation in relatively wide-spread species, we predict that the position on the *variance avoidance* axis evolves in response to the mean environment and its predictability; *responsiveness* should depend on the amplitude of environmental change (fixed development being only adaptive in stable environments); and the *information use* axis should correlate with environmental predictability. While some of these predictions have been tested in isolation, rigorous test over all axes at once are missing.

Insect diapause has historically received ample attention (e.g. Danilevskii 1965; Bradshaw 1976; Urbanski *et al.* 2012), as insects can be conveniently reared under laboratory conditions. Insect diapause is a polyphenism that is controlled by day length, and the reaction norm (photoperiodic response curve) is generally logit-shaped. Several excellent case studies exist, in which photoperiodic response curves from multiple populations were recorded, and their critical day length (inflection point) correlated with latitude, in order to infer the adaptive value of insect diapause (Danilevskii 1965). This correlation has, however, limited ecological meaning, because the selective agent is not latitude but timing of winter onset, and the relevant phenotypic response is not the day length reaction, but species phenology. Moreover, this approach does not incorporate adaptive bet-hedging responses in the face of changing climate predictability. Hence, we collected day length reaction norms from 447 populations (57 studies) and converted them to phenological reaction norms (diapause percentage vs day of year). We then derived the inflection point, the phenotypic variance and the variance composition of the reaction norms and correlated them with winter onset and winter predictability as derived from climate data.

Because studies are (understandably) biased to temperate climates and nearly no canalized phenotypes were available, there is insufficient information about the responsiveness axis. We therefore concentrated on the other two axes and predicted that:

- 1) Inflection points (critical day lengths) correlate with mean winter onset (*arithmetic mean optimization*)
- 2) The variance composition correlates with environmental predictability (*information use*)
- 3) Deviation from model 1) correlates with environmental predictability (*conservative bet-hedging*)

For comparison, we also estimated by how much diapause timing and critical photoperiod change with latitude.

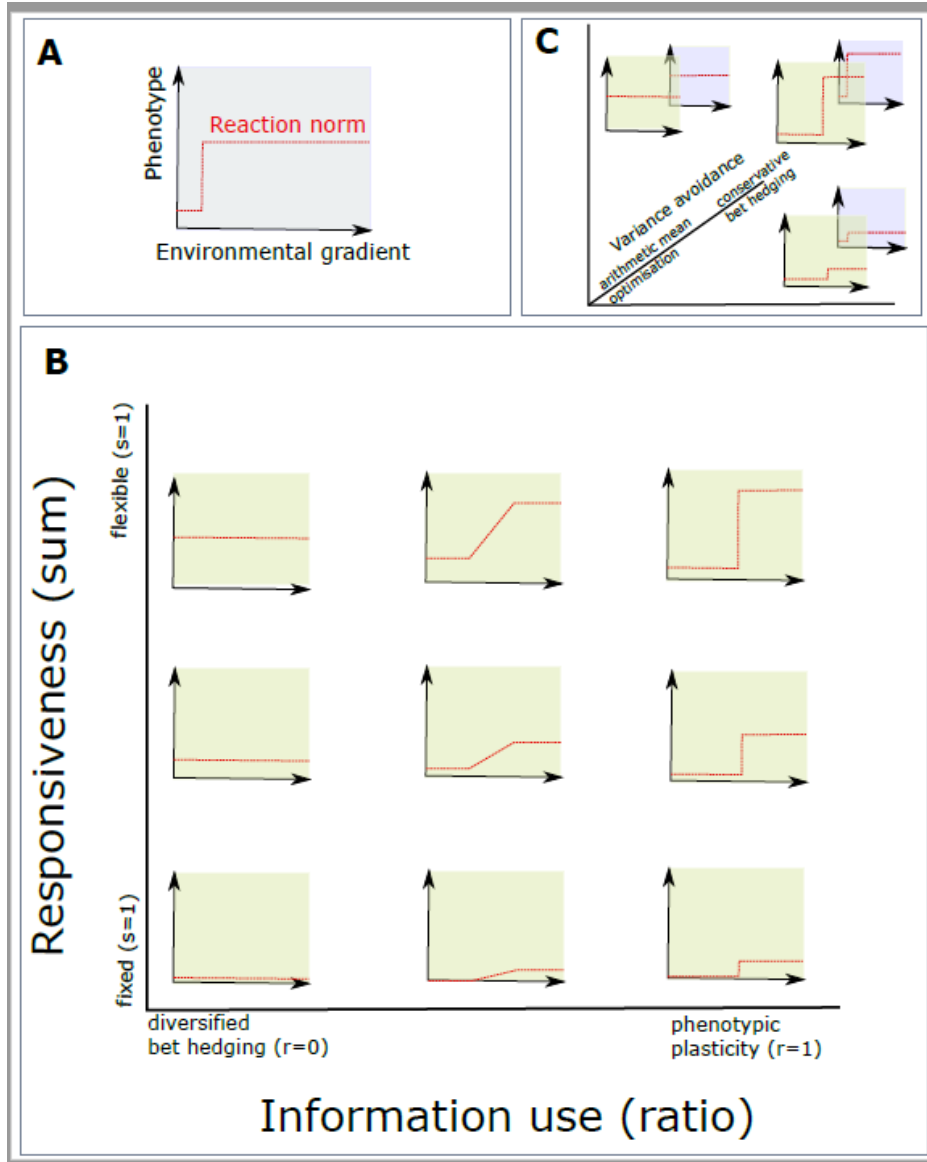


Fig. 1. Relationship of evolutionary strategies with reaction norm properties. Panel A shows a dichotomous reaction norm. The decision to switch phenotypes can be expressed by a steep logistic curve. Reaction norms can divert in various ways from this step function: By changes in the ratio (x-axis) and sum (y-axis) of the variance components σ_{among}^2 and σ_{within}^2 (Panel B), and by changes in the mean (Panel C). Axes describe the resulting evolutionary strategies.

σ_{among}^2 = squared standard deviation along environments (E_i), $\frac{\sum (p_{Ei} - \bar{p}_E)^2}{n}$; σ_{within}^2 = Variance of Bernoulli draws, $\frac{\sum p_{Ei}(1-p_{Ei})}{n}$; r = variance composition, $\frac{\sigma_{among}^2}{\sigma_{among}^2 + \sigma_{within}^2}$; s = phenotypic variance, $\sigma_{among}^2 + \sigma_{within}^2$.

Methods

1. Overview

We used climate station data to calculate mean winter onset for 25,340 locations. We then calculated winter predictability with two methods: First, we calculated the standard deviation in winter onset. This served as estimate of the reliability of day length as cue. Secondly, we determined the "colour of environmental noise" (Vasseur & Yodzis 2004), that is, the relative extent of long-frequency autocorrelations in temperatures.

We then extracted photoperiodic response curves from 57 published studies (447 diapause reaction norms) along with their sampling locations. In these studies multiple populations of a species have been subjected to several day length conditions in climate chambers, and the percentage of diapausing individuals under each day length was reported. To reconstruct continuous reaction norms from this discontinuous data, we applied a Markov chain Monte Carlo algorithm to estimate lower and upper limit, inflection point and slope. Based on these parameters we calculated mean diapause timing, degree of phenotypic variance and variance composition, and correlated them with the climate variables.

2. climate data

data preparation

We used land surface temperature data from the Global Historical Climatology Network GHCN-Daily (Menne *et al.* 2012a, b). We extracted daily minimum and maximum temperatures from all climate stations with a perl script (~12.6 million months with data, ~34,000 stations). All further analysis was conducted in R version 3.4.3 (R Core Team n.d.), using R base functions and convenience functions (Bache & Wickham 2014; Dowle & Srinivasan 2017; Rinker 2017; Wickahm 2017; Becker *et al.* 2018; Duncan Temple Lang and the CRAN team 2018; Garnier 2018). We removed all days with incomplete data (either minimum or maximum temperature not recorded), all years in which more than half of the data points was missing, all stations that covered less than 3 years of data, and all stations from the southern hemisphere. This procedure left 10,991,727 months (3-244 years) in 26,804 climate stations. We then calculated the average of daily maximum and daily minimum temperature for each station to estimate daily mean temperature.

winter onset and day length predictability

To estimate winter onset in each year and station, we identified cold days with average temperatures below 10°C. We then determined winter onset as the fifth cold day after midsummer. Years in which winter did not arrive according to this definition were excluded, and stations with less than 3 years with winter onset removed. We calculated a weighted mean winter onset and a frequency weighed standard deviation of winter onset to account for differences in reliability (days with eligible data) across years. We obtained 25,340 estimates of mean winter onset, day length at winter onset and winter predictability in the northern hemisphere.

temperature predictability

We calculated the "colour of noise", based on the degree of long-frequency autocorrelation in the temperature data (Vasseur & Yodzis 2004). Randomly fluctuating (white-noise) environments can be defined as time series with no temporal autocorrelation, so preceding

temperatures of any gap length (e.g. 1 day or 1 week earlier) have equally low explanatory power. In predictable (“red-noise”) environments, on the other hand, the explanatory power depends on gap lengths, and this can be demonstrated by decomposing the time series with a Fourier transform, and estimating the influence of each frequency. We hence produced power spectral density plots for all stations (daily temperatures of multiple years). First, we detrended the data by subtracting from each daily temperature estimate its station-wide mean. Missing values were replaced by linear interpolation (Moritz 2018). We then calculated the spectral density within the frequency range of $1/366$ and $1/(2 \text{ months})$ with the function *spectrum*, using a standard smoothing algorithm (*pgram*), and estimated beta as the negative of the slope of $\log_{10}(\text{spectral density}) \sim \log_{10}(\text{frequency})$. Beta was only calculated if there were at least five frequency estimates. In total there were 26,804 estimates of beta, but only 25,340 estimates were based on 25 or more years of data and subsequently used.

3. Empirical data

Eligibility criteria

In our literature search for diapause reaction norms we concentrated on studies that measure photoperiodic response curves of arthropods. We excluded marine and intertidal organisms, because corresponding climate estimates were only available for terrestrial systems. Invertebrates with a larval stage in shallow water (e.g. mosquitoes) were nevertheless included. Studies with estimates for fewer than 3 populations were excluded, because in these cases the variance would be absorbed by the random term “species” which was included in the analysis. We also excluded all studies that measured diapause at less than four photoperiods. To maximize sample sizes, we did not restrict our analysis to any geographic location or publication language.

Search strategy

We conducted two independent literature searches in the Web of Science core collection (Fig. 2). First (26.11.2018) we limited the search terms to:

TS = ((photoperiodic AND (geogr OR range)) OR (photoperiod* AND latitud*) OR (photoperiod* AND longitud*))*

Secondly (28.11.2018), we used a wider range of search terms,

TS = (("day length" OR photoperiod OR diapaus* OR hibern* OR dorman*) AND (geogr* OR "range" OR latitud* OR longitud* OR cline\$ OR clinal))*

but filtered the output by arthropod-related terminology (Supplementary material S1). We found 1683 references in the first search, of which we judged 278 potentially relevant, and 57 met all eligibility criteria. In the second search we found further 2362 references (6179 before filtering), with 355 potentially relevant and 13 eligible articles. We did a forward-citation search on the 70 eligible articles of both searches on 4.12.2018 and found 790 new references, which included 117 potential and 4 eligible articles. A second forward-citation search on these four articles on 5.12.2018 brought 118 new articles, but none were relevant. One further article was found to be wrongly tagged as negative after the search was closed. Altogether there were 75 useful references (623 populations).

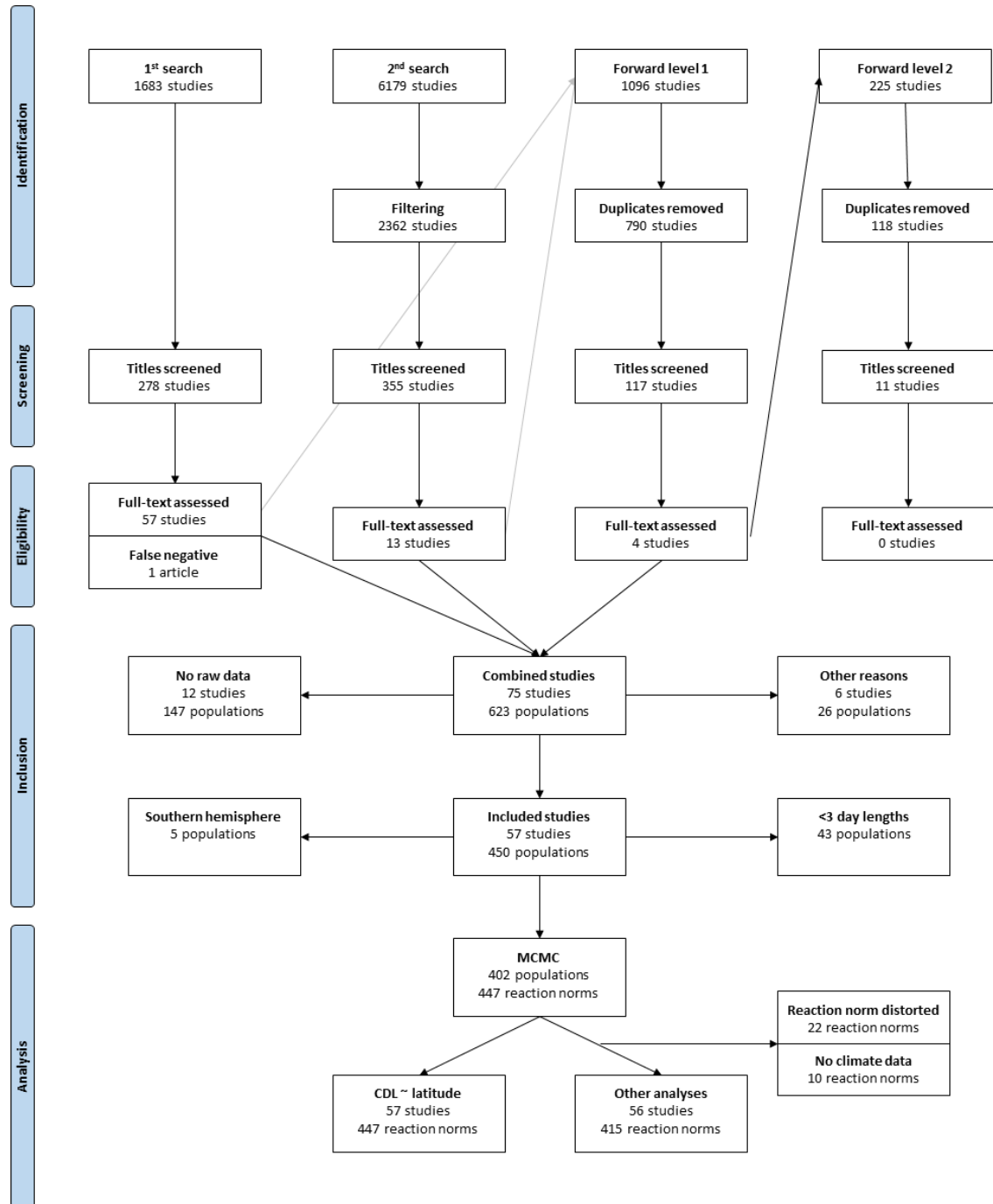


Fig. 2. Prisma Scheme

Inclusion criteria

12 articles (147 populations) were excluded because they were not accompanied by raw data, tables or figures that allowed further analysis, and the authors were deceased, did no longer have the raw data or did not respond to our emails. We further removed six articles (26 populations) that were otherwise not usable. From the remaining 57 studies we removed 43 further individual populations with less than three day length measurements, and five populations from the southern hemisphere, so 402 populations remained. Because some studies reported reaction norms for multiple lines from the same population, there were 447 reaction norms available, and these 447 reaction norms consisted of 3035 individual data points.

Data extraction

The reaction norms in 49 of the 57 studies were presented as figures. In these cases we saved the figure and extracted the data with WebPlotDigitizer Version 3.12 (Rohatgi 2017). Where necessary, the day length was then rounded or corrected to match the description in materials and methods of the respective study. Y-values that were slightly above 100% or below 0% were set to 100% and 0% respectively.

Detailed information on numbers of individuals per day length estimate were rarely available (97 reaction norms), as numbers were either given as population-level means (26 reaction norms), as global average or range (291 reaction norms), or missed entirely (33 reaction norms). We wish to emphasize that a lack of detailed information should not be confused with an unweighted (“vote-count”) meta-analysis, because the sample size (day lengths per population) was always known. Rather, the missing information occurred on a lower level (points within population) than the level of replication (population). Where the data was provided, we recorded it for later weighing of the data points.

Calculation of mean and variance composition

The published reaction norms reported the change of diapause percentages with day length. Day length changes, however, more strongly at northern latitudes and thus has no direct bearing on insect phenology. Hence, we converted day lengths into julian days, by using the reported latitude of the sampling location and the *daylength* function from the package *geosphere* (Hijmans 2017). For 728 reaction norms (24% of the data) one or more day length treatments were outside naturally occurring day lengths given the latitude in which the populations were sampled (588 were longer than midsummer days, 140 shorter than midwinter days). We assumed that these artificial day lengths represent diapause incidence at midsummer and midwinter, respectively, but removed 22 reaction norms that became severely distorted by this assumption. All further analysis except the correlation of critical photoperiod with latitude are based on the converted reaction norms.

We specified the reaction norm shape via four parameters:

$$p(x) = c + \frac{(d-c)}{1+\exp(b*(x-e))} \quad (\text{eq. 1})$$

In this equation $p(x)$ is the frequency of diapausing individuals under julian day x . e influences the inflection point of the curve, and hence directly represents the axis *mean*. c and d indicate the lower and upper diapause threshold, and b is the slope of the curve. The *variance composition* can then be defined as the ratio of variance between and among environments

(Fig. 1). The variance within each environment is that of a Bernoulli draw: $p_x * (1-p_x)$. Because a reaction norm consists of multiple p along an environmental gradient (day length treatments), we define the variance within environments as:

$$\sigma_{within}^2 \stackrel{\text{def}}{=} \frac{\sum p_x * (1-p_x)}{n}; n = \text{number of treatments} \quad (\text{eq. 2})$$

σ_{within}^2 constitutes a potential diversified bet-hedging trait and is maximized by a flat reaction norm at the 0.5 level.

We define the variance among treatments as the squared standard deviation,

$$\sigma_{among}^2 \stackrel{\text{def}}{=} \frac{\sum (p_x - \bar{p}_x)^2}{n-1} \quad (\text{eq.3})$$

This component may represent phenotypic plasticity and is maximized by a steep reaction norm. The *variance composition* can then be defined as the ratio of the two components:

$$r = \frac{\sigma_{among}^2}{\sigma_{among}^2 + \sigma_{within}^2} \quad (\text{eq. 4})$$

Lastly, *phenotypic variance* describes the magnitude of phenotypic variation, and can be regarded the sum of the two variance components,

$$s = \sigma_{among}^2 + \sigma_{within}^2 \quad (\text{eq. 5})$$

Phenotypic variance is zero for flat reaction norms at the 0 % level, and increases as the phenotype move away from 0% in some (σ_{among}^2) or all (σ_{within}^2) environments.

To derive midpoints and variance components along with a measure of reliability, we modelled each reaction norm as logistic curve according to eq. 1. Due to the data scarcity (on average seven data points per reaction norm), standard non-linear regression techniques did not always yield reasonable estimates, for example the slope could not be estimated when there was only one data point present on the sloped part of the reaction norm. Nevertheless, the range of the possible parameter space can be estimated with Markov chain Monte Carlo methods. We thus estimated the 4-dimensional credible parameter space and calculated the variance components based on this parameter space.

MCMC specifications

We used rjags (Plummer 2018) to run Markov chain simulations on each of the 447 reaction norms. We ran 4 replicate chains with lengths of 11,000 iterations and discarded a burn-in of 1,000 iterations. We specified our model with (eq. 1), and consequently chose the binomial density function to estimate the likelihood. If specified in the primary study, we used the sample sizes of each day length treatment as number of trials, otherwise we used the global average of the study. For those studies that did not mention sample sizes, we used a global average of 100 trials for each of the data points. We implemented uninformative priors for all four parameters. These were:

$$\begin{aligned} b &\sim \text{unif } \{-100, 100\} \\ c &\sim \text{unif } \{0, 1\} \\ d &\sim \text{unif } \{c, 1\} \end{aligned}$$

$e \sim \text{unif} \{D_{\min}, D_{\max}\}$, with D_{\min} and D_{\max} being the range of applied day length treatments, converted in julian days.

The upper limit of the logit-function was constrained to be higher than the lower limit, because otherwise switching between the two equal solutions (positive slope, $d > c$ and negative slope, $c < d$) would render the chain unidentifiable. Despite the relative data scarcity, the four replicate Markov chains mixed well in nearly all cases, providing a well-defined frequency distribution (Supp S2). We repeated the analyses on the untransformed day length reaction norms to obtain a critical day length estimate that is comparable to those obtained in earlier studies.

The MCMC algorithms provided a 4-dimensional parameter space to define continuous reaction norms, and we calculated the variance components of those curves that fall within the credible intervals. To do so, we followed the trace of the MCMC algorithm. For each iteration step we sampled 1000 equally spaced day lengths around the proposed inflection point $e \pm 100$ days, and performed the variance calculations (eq. 2-5) on the proposed parameters b, c, d and e . Following the logic of the MCMC algorithm, we reported the 0.025 and 0.975 quantiles of the resulting frequency distribution as credible intervals.

Merging with climate data

To combine climate data and study site locations, we averaged the climate estimates from the 5 closest stations within a 5° radius (weighted by 1/euclidian distance). When the coordinates were not directly provided in the study, we used the coordinates of the quoted town or area. Town and area coordinates were made available by the WikiProject Geographical coordinates (https://en.wikipedia.org/wiki/Wikipedia:WikiProject_Geographical_coordinates) and the Geohack tool (<https://www.mediawiki.org/wiki/GeoHack>). 10 populations did not have any climate station nearby and were only used for correlations with latitude, but not in any further analysis.

4. Analysis

We used linear mixed-effects models with a nested random structure (package metafor, Viechtbauer 2010) to correlate the reaction norm properties with climate variables. The random effects were nested on five levels (population/study/species/genus/order), but we simplified the random structure to populations/species/order, ignoring both study ID and genus. Study ID was disregarded because most species were only represented by a single study, and those species that were represented by multiple studies usually contained the same first or lead authors and applied the same methods (Table 1). Genus was disregarded because there were either only very few genera per order available (e.g. Diptera), or all species within an order were placed in different genera (Lepidoptera, Table 1). We weighed the reaction norm estimates by the inverse of the variance (credible interval ranges, divided by 2×1.96 and squared), but truncated the intervals at the lower limit to prevent some estimates from obtaining nearly infinite weight.

We performed the following models (credible interval truncation in brackets):

0a) Critical photoperiod was correlated with latitude, using the day length reaction norms (10 minutes)

0b) Mean diapause timing was correlated with latitude, using the julian day reaction norms (1 week)

1) Mean diapause timing was correlated with mean winter onset (1 week)

- 2a) The ratio of the variances (e.q. 4) was correlated with day length predictability (5%)
- 2b) The ratio of the variances (e.q. 4) was correlated with temperature predictability (5%)
- 3a) the residuals of model 1 were correlated with day length predictability (5%).
- 3b) the residuals of model 1 were correlated with temperature predictability (5%).

We used the full dataset (447 reaction norms) for model 0a), but removed all reaction norms that were not convertible into julian days or had no nearby climate stations (415 remaining) for all other models.

We assumed a gaussian distribution for all models, though we logit-transformed variance ratios prior to analysis. For all models we report partial R^2 values, calculated as proportion of variance reduction at each random level, $\frac{\partial^2_{random} - \partial^2_{mixed}}{\partial^2_{random}}$. This is an extension of a commonly used pseudo- R^2 value (Raudenbush 2009). In addition, we provide likelihood ratio test statistics. Model 3 was conducted without the nested random terms, because their effect was already accounted for by model 1.

sensitivity of climate predictions to temperature threshold

Arthropod thermal requirements vary among species, and our use of a 10°C temperature threshold was an arbitrary decision. It resulted in a global median winter onset around Oct 11, which is within the range of commonly reported phenological windows and threshold values (Halkett *et al.* 2004; Waldock *et al.* 2013). To explore the sensitivity of our meta-analysis to the arbitrary threshold, we systematically varied it between 0 and 15°C, and calculated the R^2 profiles of models 1 and 2a.

Results

Overview of studies

We found 57 studies that were conducted on 45 species and represent 447 reaction norms of 402 populations (Table 1). The populations were sampled between 14.3 °N (21.0 °N if excluding populations without climate data) and 69.0 °N (Fig. 3) and stemmed from Japan (57%), Europe and the Caucasus (26%), northern America (13%) and the Asian continent (4%). For each reaction norm there were on average 6.8 day length treatments available (range 4 – 21). The variance components, as well as their sums and ratios could be most reliably calculated for nearly flat or very steep curves, whereas intermediate reaction norms were more difficult to estimate (Supp S2, see also Fig. 4C,D). Most reactions norms were highly plastic, while canalized and flat reaction norms were generally rare (Supp S2). The credible intervals of all parameters decreased with the number of day length treatments (Supp S3).

Overview of climate variables

Winter onset followed the expected latitudinal cline in the northern hemisphere (Fig. 3a), increasing nearly linearly with latitude. The day length at winter onset, however, increased exponentially with latitude (Supp S4). A linear interpolation between 21 and 69° N resulted in a day length change of 46.34 minutes per 5° latitude ($R^2 = 0.54$). The two estimates of predictability followed a different pattern, being mostly influenced by altitude and continentality (Fig. 3 b-c).

Analyses

The critical photoperiod increased linearly with latitude, by 48.45 ± 2.01 min per 5 ° N (Fig. 4A; $R^2_{total} = 0.55$, $R^2_{order} = 0$, $R^2_{species} = 0.66$, $R^2_{population} = 0.59$; LRT ratio = 364.1, $p < 0.0001$). The converted phenological reaction norm means, i.e. mean diapause timing, correlated only weakly with latitude (Fig. S5; $R^2_{total} = 0.11$, $R^2_{order} = 0$, $R^2_{species} = 0$, $R^2_{population} = 0.35$; LRT ratio = 136.9, $p < 0.0001$) and mean winter onset (Fig. 4B; $R^2_{total} = 0.17$, $R^2_{order} = 0$, $R^2_{species} = 0.08$, $R^2_{population} = 0.39$; LRT ratio = 158.9, $p < 0.0001$), as high-latitude populations failed to track the earliest winter onsets. In both analyses the resolution of the reaction norms, and hence their reliability, varied considerably, so that 40% (critical photoperiod) to 49% (mean diapause timing) of the reaction norm estimates contributed to 90% of the overall weight (reflected by point size differences in Fig. 4).

Contrary to the predictions of our framework (Fig.1), the ratio of variance components (within vs among) correlated only very weakly with day length predictability (Fig. 4C; $R^2_{total} = 0.10$, $R^2_{order} = 0$, $R^2_{species} = 0.26$, $R^2_{population} = 0.08$; LRT ratio = 33.36, $p < 0.0001$), and not at all with temperature predictability (Fig. 4D; $R^2_{total} = 0$, $R^2_{order} = 0$, $R^2_{species} = 0$, $R^2_{population} = 0$; LRT ratio = 0.14, $p = 0.71$). Residual variation in mean timing was also neither correlated with day length predictability (Fig. 5A; $R^2 = 0$; LRT ratio = 0.16, $p = 0.69$) nor with temperature predictability (Fig. 5B; $R^2 = 0$; LRT ratio = 0.02, $p = 0.90$), indicating no evidence for conservative bet-hedging.

The results were largely independent of the definition of winter onset. R^2_{total} from model 1 ranged from 0.16 to 0.18 for thresholds between 2 and 11.5 °C, and R^2_{total} from model 2a ranged from 0.06 to 0.12 between 3 and 11 °C (supp. S6).

Discussion

With this meta-analysis we tested the prediction that reaction norm shapes correlate with the environmental conditions in which they are adaptive. We concentrated on the relationship of conservative and diversified bet-hedging with arithmetic mean optimization and phenotypic plasticity, and hence tested whether the mean and the variance composition of diapause reaction norms correlate with mean winter onset and predictability of winter onset (Joschinski & Bonte 2019; Fig. 1). In addition, we correlated the inflection point (critical photoperiod) with latitude, clarifying a long-standing empirical law that predicts a shift by 60 – 90 minutes per 5° latitude (Danilevskii 1965).

Critical photoperiod and latitude

Our analysis broadly supports Danilevskii's rule. The timing of winter onset (by which we mean gradually cooling autumn temperatures) is consistently earlier at higher latitudes and altitudes (Fig. 3A), and early autumn day length increases with latitude. Taken together, these two factors cause a rapid increase of day length at winter onset with latitude, which in turn should cause a latitudinal cline in reaction norm means. The day length decline is with 46 min/5°N (Supp. S4), however, still much less than the response of insects as estimated by Danilevskii and colleagues. The early empirical observations were based on few case studies with data of relatively low resolution, while our meta-analysis integrates data from 57 high-quality studies and applies robust statistical approaches. Accordingly, we arrive at an estimate of 48 ± 2 min/5°N (Fig. 4A), a result that is in line with the estimate of 46 min/5°N from climate data. Thus, we do not only provide strong empirical evidence for Danilevskii's observation, but also support it with climate data.

Although a linear approximation of reaction norm shifts is warranted with the published data at hand (Fig. 4A) the climate data suggests that the relationship should be exponential, with the slope increasingly deviating above about 55°N. Unfortunately, very little data is available above 55°N, and the existing data is furthermore mostly confined to a single genus (*Drosophila*). Further studies at higher latitudes, especially on other orders, are hence required for a more robust understanding of reaction norm changes with latitude.

Mean timing and mean winter onset

Day length is not a selective agent and has rarely direct implications for fitness (Joschinski *et al.* 2015; but see Dunbar & Shi 2013; Joschinski *et al.* 2019) – rather the correlated temperature drop imposes selection. Changes in the day length reaction norm must hence be translated into changes in species phenology. This correlation is much weaker than the correlation of the untransformed critical day length – latitude relationship, and the slope is only 0.5, not 1 as should be expected (Fig. 4B). This low correlation is caused by two factors: First, variance is elevated at low latitudes with late winter onsets. Day length changes less over the year at low

latitudes, so random deviation from the optimal day length response (or measurement error) causes large variation in diapause timing. Secondly, the statistically more influential northern *Drosophila* species (early winter onset) deviate from the strong negative linear trend, causing an overall decline in R^2 values and slope. Bias to a single genus complicates the interpretation, but it is conceivable that the evolution of photoperiodic reaction norms is constrained at high latitude, particularly with the exponential increase of required day length with latitude in mind (Supp. S4). Whether the failure to adapt to very long day lengths at high latitudes increases extinction risks remains to be determined, as the comparably late diapausing strategy may be supplemented by increased cold tolerance and other non-diapausing strategies (e.g. Kimura 2004). Winter may thus have to be defined differently for northern species. Species- or population- specific winter onset definitions are unfortunately not available, and further assumptions on differences in cold tolerance would inflate the researchers degrees of freedom (Simmons *et al.* 2011) to an unfeasible level. Nevertheless, our results were robust to changes in the winter onset thresholds, so the observed low correlation is unlikely to be caused by our winter onset definition alone. More research is needed to determine the vulnerability of northern populations to climate change, but we demonstrate that interpretations based on photoperiodic reaction norms alone do not draw an adequate picture.

Diversified bet-hedging and phenotypic plasticity

Our analysis revealed that day length reaction norms are more variable than is commonly acknowledged. The reaction norm shapes ranged from very steep to entirely flat (Fig. 1C, D), though steep reaction norms were more common than flat ones. We suspect study bias towards species with steep reaction norms, because researchers chose species with the most reliable photoperiodic reaction. Nevertheless, we found enough flat reaction norms for a thorough analysis. The existence of reaction norm shapes with high variance within environments does not necessarily constitute bet-hedging, however. Ultimately, it needs to be demonstrated that phenotypic variance increases geometric mean fitness in a population's respective environment (Simons 2011). We took a comparative approach and correlated variance within environments with winter predictability, i.e. tested whether reaction norms are fine-tuned by local selection as predicted by theory (Joschinski & Bonte 2019). This correlation was, however, very weak (Fig. 4C). This may be partly explained by limitations of the data (many studies with only 3 or 4 populations of varying reliability), and partly by limitations of the statistical approaches (ratios other than 0 and 1 contain inherently more variance). Given the large scope of the meta-analysis, however, we conclude that diversified bet-hedging in diapause phenology is truly rare.

So far we treated phenology as a one-dimensional response to photoperiodic cues. Photoperiodic response curves are, however, also sensitive to temperature change, as the photoperiodic response shifts to longer day lengths (later timing) if temperatures are warmer (e.g. Beach 1978; Ichijo 1986; Chen *et al.* 2013). This form of plasticity is clearly adaptive provided that temperatures during diapause induction are good predictors for winter onset. Unfortunately, experiments that assess photoperiodic reaction norms under multiple temperatures and for several populations are rare, hence the inclusion of temperature plasticity was not possible in this meta-analysis. Nevertheless, genotypes that rely on temperature plasticity should exhibit less day length plasticity. Hence (indirectly) the calculated variance

ratio of the day length reaction norm should correlate negatively with temperature predictability. We did not find this relationship (Fig 4D), but the reasoning is admittedly very indirect. We therefore urge for further studies on the combined effects of day length and temperature plasticity of arthropod diapause.

Arithmetic mean optimization and conservative bet-hedging

As alternative to diversified bet-hedging, unpredictable conditions may select for early diapause, so that the risk of fitness loss by early frost is mitigated at the cost of population growth (conservative bet-hedging, Seger & Brockmann 1987; Starrfelt & Kokko 2012). Hence we expected residual variation in mean phenology to correlate with environmental predictability, such that populations in highly unpredictable environments diapause earlier than expected based on mean winter onset. We did not find any evidence for conservative bet-hedging, neither in day length reaction norms (Fig. 5A) nor in temperature reaction norms (Fig. 5B). Empirical evidence for conservative bet-hedging is rare (Simons 2011), and indeed not expected for many life history strategies (Poethke *et al.* 2016). Our results therefore support existing evidence that conservative bet-hedging is only selected for under very specific conditions.

Evolvability in a changing climate

In this meta-analysis we inferred the potential of phenological strategies to evolve in the face of a rapidly changing climate. By correlating reaction norm shapes (Fig. 1) with their respective environments, we assessed the repertoire of strategies that arthropods applied in the past, e.g. after range expansions or human-mediated introduction to novel habitats. Two general patterns emerged:

First, the mean phenology of northern populations did not match environmental conditions, potentially due to the extreme shifts of the day length reaction norms that would be required at high latitudes. Although other mechanisms (e.g. frost tolerance) may have been co-opted at high latitudes, the discovered constraints in day length reaction norms may slow down adaptation to short days, and thus decelerate the potential for northward range expansion. There is additional concern that species- or guild-specific phenological strategies cause a phenological mismatch among interacting species (Visser & Gienapp 2019); however, because the constraints in the evolvability of day length reaction norms have a clear physical basis (exponential increase of day length at winter onset with latitude), we find it likely that phenologies converge with increasing latitude. Thus, we currently see little scope for increasing phenological mismatches. Nevertheless, biotic factors may further amplify the extinction risk imposed by changes in means and constraints in reaction norms.

Secondly, both conservative and diversified bet-hedging were rare. Finding evidence for bet-hedging is notoriously difficult (Simons 2011). First of all, bet-hedging has to be distinguished from plasticity, i.e. it has to be shown that there is random variation of a trait. We overcame this difficulty by formalizing the relationship of diversified bet-hedging and phenotypic plasticity as extremes on a continuous axis (Joschinski & Bonte 2019). This differentiation allows also studying partial plasticity and hence provides a more realistic view on the plasticity

- bet-hedging dichotomy. Another common problem is that the identification of random variation is not sufficient evidence for bet-hedging, because one needs to demonstrate that this variation also confers a fitness benefit. We clearly separated reaction norm shapes from evolutionary strategies and indeed showed that reaction norms may be flat or vary in their means, but that these properties correlate only weakly with environmental predictability – in other words, early mean phenology or variance in phenology are unlikely to represent bet-hedging strategies.

We can only speculate on the reasons for a lack of diversified bet-hedging despite considerable variance in winter onset. Potentially, many insects hedge their bets in space rather than in time. Bet-hedging via dispersal has benefits and costs that are similar to bet-hedging via dormancy (Buoro & Carlson 2014), and may thus be a viable alternative to variance in diapausing reaction norms. Similarly, some species may hedge their bets in identity, i.e. adjust their mating strategies according to environmental conditions (Fox & Rauter 2003; Gerber & Kokko 2018). These alternative bet-hedging strategies are clearly outside the scope of this review; future studies need to address their coevolution in more detail.

Conclusions

Our analysis revealed that diapause reaction norms readily evolved towards earlier diapause at high latitudes. This shift in reaction norms is however not sufficient to induce an earlier phenology, and the discrepancy will likely increase as species expand their range northwards. Moreover, bet-hedging strategies did not evolve under unpredictable conditions in the past, so we predict that species are sensitive to changes in climate predictability.

Acknowledgments

This research has benefitted from a statistical consult with Ghent University FIRE (Fostering Innovative Research based on Evidence). In addition we would like to thank Jan Baert and Thomas Hovestadt for discussion of this manuscript. JJ was financially supported by a DFG research fellowship. DB is funded by FWO project G018017N.

- Bache, S.M. & Wickham, H. (2014). *magrittr: A Forward-Pipe Operator for R*.
- Bathiany, S., Dakos, V., Scheffer, M. & Lenton, T.M. (2018). Climate models predict increasing temperature variability in poor countries. *Science Advances*, 4, eaar5809.
- Beach, R. (1978). The required day number and timely induction of diapause in geographic strains of the mosquito *Aedes atropalpus*. *Journal of Insect Physiology*, 24, 449–455.
- Becker, R.A., Wilks, A.R., Brownrigg, R., Minka, T.P. & Deckmyn, A. (2018). *maps: Draw Geographical Maps*.
- Bell, J.R., Alderson, L., Izera, D., Kruger, T., Parker, S., Pickup, J., *et al.* (2015). Long-term phenological trends, species accumulation rates, aphid traits and climate: five decades of change in migrating aphids. *Journal of Animal Ecology*, 84, 21–34.
- Bradshaw, W.E. (1976). Geography of photoperiodic response in diapausing mosquito. *Nature*, 262, 384–386.
- Bradshaw, W.E. & Holzapfel, C.M. (2001). Genetic shift in photoperiodic response correlated with global warming. *PNAS*, 98, 14509–14511.
- Bradshaw, W.E., Quebodeaux, M.C. & Holzapfel, C.M. (2003). Circadian Rhythmicity and Photoperiodism in the Pitcher-Plant Mosquito: Adaptive Response to the Photic Environment or Correlated Response to the Seasonal Environment? *The American Naturalist*, 161, 735–748.
- Buoro, M. & Carlson, S.M. (2014). Life-history syndromes: Integrating dispersal through space and time. *Ecol Lett*, 17, 756–767.
- Chen, Y.-S., Chen, C., He, H.-M., Xia, Q.-W. & Xue, F.-S. (2013). Geographic variation in diapause induction and termination of the cotton bollworm, *Helicoverpa armigera* Hübner (Lepidoptera: Noctuidae). *Journal of Insect Physiology*, 59, 855–862.
- Childs, D.Z., Rees, M., Rose, K.E., Grubb, P.J. & Ellner, S.P. (2004). Evolution of size-dependent flowering in a variable environment: construction and analysis of a stochastic integral projection model. *Proc Biol Sci*, 271, 425–434.
- Clarke, A. (2017). *Principles of Thermal Ecology: Temperature, Energy and Life*. Oxford University Press.
- Danilevskii, A.S. (1965). Photoperiodism and seasonal development of insects. *Photoperiodism and seasonal development of insects*.
- Dowle, M. & Srinivasan, A. (2017). *data.table: Extension of `data.frame`*.
- Dunbar, R.I.M. & Shi, J. (2013). Time as a constraint on the distribution of feral goats at high latitudes. *Oikos*, 122, 403–410.
- Duncan Temple Lang and the CRAN team. (2018). *RCurl: General Network (HTTP/FTP/...) Client Interface for R*.
- Fox, C.W. & Rauter, C.M. (2003). Bet-hedging and the evolution of multiple mating. *Evol Ecol Res*, 5, 273–286.
- Garnier, S. (2018). *viridis: Default Color Maps from “matplotlib.”*
- Gerber, N. & Kokko, H. (2018). Abandoning the ship using sex, dispersal or dormancy: multiple escape routes from challenging conditions. *Philos T R Soc B*, 373, 20170424.
- Gienapp, P., Lof, M., Reed, T.E., McNamara, J., Verhulst, S. & Visser, M.E. (2013). Predicting demographically sustainable rates of adaptation: can great tit breeding time keep pace with climate change? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368, 20120289.
- Gomi, T., Adachi, K., Shimizu, A., Tanimoto, K., Kawabata, E. & Takeda, M. (2009). Northerly shift in voltinism watershed in *Hyphantria cunea* (Drury) (Lepidoptera: Arctiidae) along the Japan Sea coast: Evidence of global warming? *Applied Entomology and Zoology*, 44, 357–362.
- Gomi, T. & Takeda, M. (1996). Changes in Life-History Traits in the Fall Webworm within Half a Century of Introduction to Japan. *Functional Ecology*, 10, 384–389.

- Halkett, F., Harrington, R., Hullé, M., Kindlmann, P., Menu, F., Rispe, C., *et al.* (2004). Dynamics of production of sexual forms in aphids: theoretical and experimental evidence for adaptive “coin-flipping” plasticity. *Am Nat*, 163, E112–E125.
- Hijmans, R.J. (2017). *geosphere: Spherical Trigonometry*.
- Hou, Y.-Y., Xu, L.-Z., Wu, Y., Wang, P., Shi, J.-J. & Zhai, B.-P. (2016). Geographic Variation of Diapause and Sensitive Stages of Photoperiodic Response in *Laodelphax striatellus* Fallén (Hemiptera: Delphacidae). *J Insect Sci*, 16.
- Ichijo, N. (1986). Disjunctive Cline of Critical Photoperiod in the Reproductive Diapause of *Drosophila lacertosa*. *Evolution*, 40, 418–421.
- IPCC. (2014). *Climate Change 2014: Synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Geneva, Switzerland.
- Ito, K. & Nakata, T. (2000). Geographical variation of photoperiodic response in the females of a predatory bug, *Orius sauteri* (Poppus) (Heteroptera: Anthocoridae) from northern Japan. *Applied Entomology and Zoology*, 35, 101–105.
- Jordan, R.G. & Bradshaw, W.E. (1978). Geographic Variation in the Photoperiodic Response of the Western Tree-Hole Mosquito, *Aedes sierrensis*. *Ann Entomol Soc Am*, 71, 487–490.
- Joschinski, J. & Bonte, D. (2019). *Trans-generational plasticity and bet-hedging: a common eco-evolutionary framework of utter relevance for climate change adaptation*. *EcoEvoArxiv*. Available at: [dx.doi.org/10.32942/osf.io/trg34](https://doi.org/10.32942/osf.io/trg34). Last accessed 6 April 2019.
- Joschinski, J., Hovestadt, T. & Krauss, J. (2015). Coping with shorter days: do phenology shifts constrain aphid fitness? *PeerJ*, 3, e1103.
- Joschinski, J., Kiess, T. & Krauss, J. (2019). Day length constrains the time budget of aphid predators. *Insect Science*, 26, 164–170.
- Kato, Y. (2005). Geographic variation in photoperiodic response for the induction of pupal diapause in the *Aristolochia*-feeding butterfly *Atrophaneura alcinous*. *Applied Entomology and Zoology*, 40, 347–350.
- Kimura, M.T. (1984). Geographic variation of reproductive diapause in the *Drosophila auraria* complex (Diptera: Drosophilidae). *Physiological Entomology*, 9, 425–431.
- Kimura, M.T. (1988). Adaptations to Temperate Climates and Evolution of Overwintering Strategies in the *Drosophila melanogaster* Species Group. *Evolution*, 42, 1288–1297.
- Kimura, M.T. (2004). Cold and heat tolerance of drosophilid flies with reference to their latitudinal distributions. *Oecologia*, 140, 442–449.
- Kimura, M.T., Bessho, A. & Dai, Z.-H. (1993). The influence of gene flow on latitudinal clines of photoperiodic adult diapause in the *Drosophila auraria* species-complex. *Biological Journal of the Linnean Society*, 48, 335–341.
- Kimura, M.T., Ohtsu, T., Yoshida, T., Awasaki, T. & Lin, F.-J. (1994). Climatic adaptations and distributions in the *Drosophila takahashii* species subgroup (Diptera: Drosophilidae). *Journal of Natural History*, 28, 401–409.
- Kingsolver, J.G. & Buckley, L.B. (2017). Quantifying thermal extremes and biological variation to predict evolutionary responses to changing climate. *Philos T R Soc B*, 372, 20160147.
- Koveos, D.S., Kroon, A. & Veerman, A. (1993). The Same Photoperiodic Clock May Control Induction and Maintenance of Diapause in the Spider Mite *Tetranychus urticae*. *J Biol Rhythms*, 8, 265–282.
- Kurahashi, H. & Ohtaki, T. (1989). Geographic variation in the incidence of pupal diapause in Asian and Oceanian species of the flesh fly *Boettcherisca* (Diptera: Sarcophagidae). *Physiological Entomology*, 14, 291–298.

- Kurota, H. & Shimada, M. (2003). Geographical variation in photoperiodic induction of larval diapause in the bruchid beetle, *Bruchidius dorsalis*: polymorphism in overwintering stages. *Entomologia Experimentalis et Applicata*, 107, 11–18.
- Lane, J.E., McAdam, A.G., McFarlane, S.E., Williams, C.T., Humphries, M.M., Coltman, D.W., *et al.* (2018). Phenological shifts in North American red squirrels: disentangling the roles of phenotypic plasticity and microevolution. *J. Evol. Biol.*, 31, 810–821.
- Lankinen, P. (1986). Geographical variation in circadian eclosion rhythm and photoperiodic adult diapause in *Drosophila littoralis*. *J. Comp. Physiol.*, 159, 123–142.
- Lehmann, P., Lyytinen, A., Piironen, S. & Lindström, L. (2015). Latitudinal differences in diapause related photoperiodic responses of European Colorado potato beetles (*Leptinotarsa decemlineata*). *Evol Ecol*, 29, 269–282.
- Lumme, J. & Oikarinen, A. (1977). The genetic basis of the geographically variable photoperiodic diapause in *Drosophila littoralis*. *Hereditas*, 86, 129–141.
- Lushai, G. & Harrington, R. (1996). Inheritance of photoperiodic response in the bird cherry aphid, *Rhopalosiphum padi*. *Physiological Entomology*, 21, 297–303.
- Menne, M., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., *et al.* (2012a). Global Historical Climatology Network - Daily (GHCN-Daily), Version 3. 22.
- Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E. & Houston, T.G. (2012b). An Overview of the Global Historical Climatology Network-Daily Database. *J. Atmos. Oceanic Technol.*, 29, 897–910.
- Monceau, K., Bonnard, O. & Thiéry, D. (2014). *Vespa velutina*: a new invasive predator of honeybees in Europe. *J Pest Sci*, 87, 1–16.
- Moritz, S. (2018). *ImputeTS: Time Series Missing Value Imputation*.
- Murata, Y., Ideo, S., Watada, M., Mitsui, H. & Kimura, M.T. (2013). Genetic and physiological variation among sexual and parthenogenetic populations of *Asobara japonica* (Hymenoptera: Braconidae), a larval parasitoid of drosophilid flies. *EJE*, 106, 171–178.
- Musolin, D.L., Tougou, D. & Fujisaki, K. (2011). Photoperiodic response in the subtropical and warm-temperate zone populations of the southern green stink bug *Nezara viridula*: why does it not fit the common latitudinal trend? *Physiological Entomology*, 36, 379–384.
- Nakao, S. (2011). Geographical variation of photoperiodic wing form determination and genetic background of reproductive diapause in arrhenotokous populations of *Thrips nigropilosus* Uzel (Thysanoptera: Thripidae) in Japan. *Appl Entomol Zool*, 46, 111–116.
- Nechols, J.R., Tauber, M.J. & Tauber, C.A. (1987). Geographical variability in ecophysiological traits controlling dormancy in *Chrysopa oculata* (Neuroptera: Chrysopidae). *Journal of Insect Physiology*, 33, 627–633.
- Noda, H. (Shimane-ken A.E.S. (1992). Geographic variation of nymphal diapause in the small brown planthopper [*Laodelphax striatellus*] in Japan. *JARQ (Japan)*.
- Paolucci, S., Zande, L. van de & Beukeboom, L.W. (2013). Adaptive latitudinal cline of photoperiodic diapause induction in the parasitoid *Nasonia vitripennis* in Europe. *Journal of Evolutionary Biology*, 26, 705–718.
- Parmesan, C. & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421, 37.
- Pegoraro, M., Zonato, V., Tyler, E.R., Fedele, G., Kyriacou, C.P. & Tauber, E. (2017). Geographical analysis of diapause inducibility in European *Drosophila melanogaster* populations. *Journal of Insect Physiology*, 98, 238–244.
- Plummer, M. (2018). *rjags: Bayesian Graphical Models using MCMC*.

- Poethke, H.J., Hovestadt, T. & Mitesser, O. (2016). The evolution of optimal emergence times: bet hedging and the quest for an ideal free temporal distribution of individuals. *Oikos*, 125, 1647–1656.
- Pullin, A.S. (1986). Effect of photoperiod and temperature on the life-cycle of different populations of the peacock butterfly *Inachis io*. *Entomologia Experimentalis et Applicata*, 41, 237–242.
- Qureshi, M.-H., Murai, T., Yoshida, H. & Tsumuki, H. (2000). Populational variation in diapause-induction and -termination of *Helicoverpa armigera* (Lepidoptera: Noctuidae). *Applied Entomology and Zoology*, 35, 357–360.
- R Core Team. (n.d.). *R: A language and environment for statistical computing*.
- Raudenbush, S.W. (2009). Analyzing effect sizes: Random effects models. In: *The handbook of research synthesis and meta-analysis*. Russell Sage Foundation, New York, USA, pp. 295–315.
- Reznik, S.Y., Dolgovskaya, M.Y., Ovchinnikov, A.N. & Belyakova, N.A. (2015). Weak photoperiodic response facilitates the biological invasion of the harlequin ladybird *Harmonia axyridis* (Pallas) (Coleoptera: Coccinellidae). *Journal of Applied Entomology*, 139, 241–249.
- Riihimaa, A., Kimura, M.T., Lumme, J. & Lakovaara, S. (1996). Geographical variation in the larval diapause of *Chymomyza costata* (Diptera; Drosophilidae). *Hereditas*, 124, 151–164.
- Rinker, T.W. (2017). *textreadr: Read Text Documents into R*. Buffalo, New York, USA.
- Rohatgi, A. (2017). *WebPlotDigitizer*. Austin, Texas, USA.
- Ryan, S.F., Valella, P., Thivierge, G., Aardema, M.L. & Scriber, J.M. (2018). The role of latitudinal, genetic and temperature variation in the induction of diapause of *Papilio glaucus* (Lepidoptera: Papilionidae). *Insect Science*, 25, 328–336.
- Sadakiyo, S. & Ishihara, M. (2011). Rapid seasonal adaptation of an alien bruchid after introduction: geographic variation in life cycle synchronization and critical photoperiod for diapause induction. *Entomologia Experimentalis et Applicata*, 140, 69–76.
- Schroeder, M. & Dalin, P. (2017). Differences in photoperiod-induced diapause plasticity among different populations of the bark beetle *Ips typographus* and its predator *Thanasimus formicarius*. *Agricultural and Forest Entomology*, 19, 146–153.
- Seger, J. & Brockmann, H.J. (1987). What is bet-hedging? In: *Oxford surveys in evolutionary biology* (4). Oxford University Press, Oxford, UK, pp. 182–211.
- Shimizu, K. & Fujisaki, K. (2002). Sexual differences in diapause induction of the cotton bollworm, *Helicoverpa armigera* (Hb.) (Lepidoptera: Noctuidae). *Applied Entomology and Zoology*, 37, 527–533.
- Shimizu, K. & Fujisaki, K. (2006). Geographic variation in diapause induction under constant and changing conditions in *Helicoverpa armigera*. *Entomologia Experimentalis et Applicata*, 121, 253–260.
- Shimizu, T. & Kawasaki, K. (2001). Geographic variability in diapause response of Japanese *Orius* species. *Entomologia Experimentalis et Applicata*, 98, 303–316.
- Shintani, Y. & Ishikawa, Y. (1999). Transition of Diapause Attributes in the Hybrid Zone of the Two Morphological Types of *Psacotha hilaris* (Coleoptera: Cerambycidae). *Environ Entomol*, 28, 690–695.
- Shintani, Y., Tatsuki, S. & Ishikawa, Y. (1996). Geographic Variation of Photoperiodic Response in Larval Development of the Yellow-Spotted Longicorn Beetle, *Psacotha hilaris* (PASCOE) (Coleoptera: Cerambycidae). *Applied Entomology and Zoology*, 31, 495–504.

- Shroyer, D.A. & Craig, G.B. (1983). Egg Diapause in *Aedes Triseriatus* (Diptera: Culicidae): Geographic Variation in Photoperiodic Response and Factors Influencing Diapause Termination. *J Med Entomol*, 20, 601–607.
- Simmons, J.P., Nelson, L.D. & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychol Sci*, 22, 1359–1366.
- Simons, A.M. (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proc Biol Sci*, 278, 1601–1609.
- So, P.-M. & Takafuji, A. (1992). Local variation in diapause characteristics of *Tetranychus urticae* Koch (Acarina: Tetranychidae). *Oecologia*, 90, 270–275.
- Starrfelt, J. & Kokko, H. (2012). Bet-hedging--a triple trade-off between means, variances and correlations. *Biol Rev Camb Philos Soc*, 87, 742–755.
- Suwa, A. & Gotoh, T. (2006). Geographic variation in diapause induction and mode of diapause inheritance in *Tetranychus pueraricola*. *Journal of Applied Entomology*, 130, 329–335.
- Takeda, M. & Chippendale, G.M. (1982). Phenological adaptations of a colonizing insect: The southwestern corn borer, *Diatraea grandiosella*. *Oecologia*, 53, 386–393.
- Tufto, J. (2015). Genetic evolution, plasticity, and bet-hedging as adaptive responses to temporally autocorrelated fluctuating selection: A quantitative genetic model. *Evolution*, 69, 2034–2049.
- Tyukmaeva, V.I., Salminen, T.S., Kankare, M., Knott, K.E. & Hoikkala, A. (2011). Adaptation to a seasonally varying environment: a strong latitudinal cline in reproductive diapause combined with high gene flow in *Drosophila montana*. *Ecology and Evolution*, 1, 160–168.
- Urban, M.C. (2015). Accelerating extinction risk from climate change. *Science*, 348, 571–573.
- Urbanski, J., Mogi, M., O'Donnell, D., DeCotiis, M., Toma, T. & Armbruster, P. (2012). Rapid Adaptive Evolution of Photoperiodic Response during Invasion and Range Expansion across a Climatic Gradient. *The American Naturalist*, 179, 490–500.
- Vasseur, D.A., DeLong, J.P., Gilbert, B., Greig, H.S., Harley, C.D.G., McCann, K.S., *et al.* (2014). Increased temperature variation poses a greater risk to species than climate warming. *Proceedings of the Royal Society of London B: Biological Sciences*, 281, 20132612.
- Vasseur, D.A. & Yodzis, P. (2004). The Color of Environmental Noise. *Ecology*, 85, 1146–1152.
- Vaz Nunes, M., Koveos, D.S. & Veerman, A. (1990). Geographical Variation in Photoperiodic Induction of Diapause in the Spider Mite (*Tetranychus urticae*): A Causal Relation between Critical Nightlength and Circadian Period? *J Biol Rhythms*, 5, 47–57.
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36, 1–48.
- Visser, M.E. (2008). Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society of London B: Biological Sciences*, 275, 649–659.
- Visser, M.E. & Gienapp, P. (2019). Evolutionary and demographic consequences of phenological mismatches. *Nature Ecology & Evolution*, 3, 879.
- Waldock, J., Chandra, N.L., Lelieveld, J., Proestos, Y., Michael, E., Christophides, G., *et al.* (2013). The role of environmental variables on *Aedes albopictus* biology and chikungunya epidemiology. *Pathog Glob Health*, 107, 224–241.

- Wang, X.-P., Yang, Q.-S., Dalin, P., Zhou, X.-M., Luo, Z.-W. & Lei, C.-L. (2012). Geographic variation in photoperiodic diapause induction and diapause intensity in *Sericanus montelus* (Lepidoptera: Papilionidae). *Insect Science*, 19, 295–302.
- Wickahm, H. (2017). *tidyverse: Easily Install and Load the “Tidyverse”*.
- Wipking, W. (1988). Repeated larval diapause and diapause-free development in geographic strains of the burnet moth *Zygaena trifolii* Esp. (Insecta, Lepidoptera). *Oecologia*, 77, 557–564.
- Yoshida, T. & Kimura, T.M. (1994). Relation of the Circadian System to the Photoperiodic Clock in *Drosophila triauraria* (Diptera: Drosophilidae) : An Approach from Analysis of Geographic Variation. *Applied Entomology and Zoology*, 29, 499–505.
- Yoshio, M. & Ishii, M. (1998). Geographical variation of pupal diapause in the great mormon butterfly, *Papilio memnon* L. (Lepidoptera : Papilionidae), in Western Japan. *Applied Entomology and Zoology*, 33, 281–288.
- 健一橋本, 和子飯島 & 賢一小川. (2008). モンシロチョウ *Pieris rapae crucivora* Boisduval (チョウ目 : シロチョウ科) の蛹休眠を誘起する光周反応の地理的変異. 日本応用動物昆虫学会誌, 52, 201–206.
- 哲雄後藤 & 徳純真梶. (1981). 日本産ナミハダニ *Tetranychus urticae* KOCH の休眠誘起の臨界日長とその地理的変異. 日本応用動物昆虫学会誌, 25, 113–118.
- 幸雄桑名. (1986). ブチヒゲヤナギドクガの光周反応と地理的起源. 日本応用動物昆虫学会誌, 30, 173–178.
- 武氏家. (1985). キンモンホソガの休眠に関する研究. 日本応用動物昆虫学会誌, 29, 198–202.
- 顕次藤本, 民人櫻井 & 史郎中尾. (2014). アカメガシワクダアザミウマ雌成虫における休眠性の地理的変異. 日本応用動物昆虫学会誌, 58, 47–54.

Order	Genus	Species	Popula- tions	Reaction norms	Photo- periods	Region	Reference
Coleoptera	<i>Acanthoscelides</i>	<i>pallidipennis</i>	3	3	5	Japan	(Sadakiyo & Ishihara 2011)
	<i>Bruchidius</i>	<i>dorsalis</i>	3	3	5	Japan	(Kurota & Shimada 2003)
		<i>dorsalis</i>	3	3	7	Japan	(Kurota & Shimada 2003)
	<i>Harmonia</i>	<i>axyridis</i>	4	4	5	Europe, Asia	(Reznik <i>et al.</i> 2015)
	<i>Ips</i>	<i>typographus</i>	4	4	5	Europe	(Schroeder & Dalin 2017)
	<i>Leptinotarsa</i>	<i>decemlineata</i>	5	6	6	Europe	(Lehmann <i>et al.</i> 2015)
	<i>Psacothaea</i>	<i>hilaris</i>	6	6	5	Japan	(Shintani <i>et al.</i> 1996)
		<i>hilaris</i>	8	11	4	Japan	(Shintani & Ishikawa 1999)
Diptera	<i>Aedes</i>	<i>albopictus</i>	21	21	12	US, Japan	(Urbanski <i>et al.</i> 2012)
		<i>atropalpus</i>	3	3	5-7	US	(Beach 1978)
		<i>sierrensis</i>	5	5	4-7	US	(Jordan & Bradshaw 1978)
		<i>triseriatus</i>	8	8	10	US	(Shroyer & Craig 1983)
	<i>Boettcherisca</i>	<i>peregrina</i>	6	6	8	Japan	(Kurahashi & Ohtaki 1989)
	<i>Chymomyza</i>	<i>costata</i>	8	14	6-8	Europe, Japan	(Riihimaa <i>et al.</i> 1996)
	<i>Drosophila</i>	<i>auraria</i>	8	8	4-5	Japan	(Kimura <i>et al.</i> 1993)
		<i>auraria</i>	7	7	4-7	Japan	(Kimura 1984)
		<i>biauraria</i>	11	11	4-5	Japan	Kimura <i>et al.</i> 1993)
		<i>biauraria</i>	4	4	5	Japan	(Kimura 1988)
		<i>lacertosa</i>	8	8	4-7	Japan	(Ichijo 1986)
		<i>littoralis</i>	8	8	7-11	Europe	(Lumme & Oikarinen 1977)
		<i>littoralis</i>	11	18	5-9	Europe	(Lankinen 1986)
		<i>melanogaster</i>	6	6	6	Europe	(Pegoraro <i>et al.</i> 2017)
		<i>montana</i>	6	24	4-6	Europe	(Tyukmaeva <i>et al.</i> 2011)
		<i>subauraria</i>	8	8	5-6	Japan	Kimura <i>et al.</i> 1993)
		<i>subauraria</i>	4	4	5-7	Japan	(Kimura 1984)
		<i>takahashii</i>	5	5	4	Japan	(Kimura <i>et al.</i> 1994)
		<i>triauraria</i>	3	3	7-11	Japan	(Yoshida & Kimura 1994)
		<i>triauraria</i>	10	10	4-5	Japan	Kimura <i>et al.</i> 1993)
<i>triauraria</i>		4	4	5-6	Japan	(Kimura 1984)	

Hemiptera	<i>Wyeomyia</i>	<i>smithii</i>	16	16	16-21	US	(Bradshaw <i>et al.</i> 2003)
	<i>Laodelphax</i>	<i>striatellus</i>	3	3	5-6	Asia	(Hou <i>et al.</i> 2016)
	<i>Nezara</i>	<i>viridula</i>	5	5	5-8	Japan	(Musolin <i>et al.</i> 2011)
	<i>Orius</i>	<i>sauteri</i>	5	5	6-8	Japan	(Shimizu & Kawasaki 2001)
Homoptera		<i>sauteri</i>	8	8	6-8	Japan	(Ito & Nakata 2000)
	<i>Rhopalosiphum</i>	<i>padi</i>	3	3	11	Europe	(Lushai & Harrington 1996)
	<i>Laodelphax</i>	<i>striatellus</i>	8	8	6-8	Japan	(Noda 1992)
	<i>Asobara</i>	<i>japonica</i>	9	9	5	Japan	(Murata <i>et al.</i> 2013)
Hymenoptera	<i>Nasonia</i>	<i>vitripennis</i>	7	7	8	Europe	(Paolucci <i>et al.</i> 2013)
Lepidoptera	<i>Atrophaneura</i>	<i>alcinous</i>	6	6	5	Japan	(Kato 2005)
	<i>Diatraea</i>	<i>grandiosella</i>	3	3	6	US	(Takeda & Chippendale 1982)
	<i>Helicoverpa</i>	<i>armigera</i>	5	5	6	Asia	(Chen <i>et al.</i> 2013)
		<i>armigera</i>	3	3	6	Japan	(Qureshi <i>et al.</i> 2000)
		<i>armigera</i>	3	3	4-5	Japan	(Shimizu & Fujisaki 2006)
		<i>armigera</i>	3	3	5	Japan	(Shimizu & Fujisaki 2002)
	<i>Hyphantria</i>	<i>cunea</i>	4	4	4-5	Japan	(Gomi & Takeda 1996)
		<i>cunea</i>	3	3	4-6	Japan	(Gomi <i>et al.</i> 2009)
	<i>Inachis</i>	<i>io</i>	3	3	9	Europe	(Pullin 1986)
	<i>Leucoma</i>	<i>candida</i>	5	5	4-5	Japan	(幸雄 1986)
	<i>Papilio</i>	<i>glaucus</i>	3	3	8-11	US	(Ryan <i>et al.</i> 2018)
		<i>memnon</i>	4	4	8	Japan	(Yoshio & Ishii 1998)
	<i>Phyllonorycter</i>	<i>ringoniella</i>	5	5	4	Japan	(武 1985)
	<i>Pieris</i>	<i>rapae</i>	7	7	5-8	Japan	(健一 <i>et al.</i> 2008)
	<i>Sericinus</i>	<i>montelus</i>	6	6	9	Asia	(Wang <i>et al.</i> 2012)
	<i>Zygaena</i>	<i>trifolii</i>	5	5	8-11	Europe	(Wipking 1988)
Neuroptera	<i>Chrysopa</i>	<i>oculata</i>	9	9	4-6	US	(Nechols <i>et al.</i> 1987)
Thysanoptera	<i>Haplothrips</i>	<i>brevitubus</i>	3	13	6	Japan	(顕次 <i>et al.</i> 2014)
	<i>Thrips</i>	<i>nigropilosus</i>	6	6	4-6	Japan	(Nakao 2011)
Trombidiformes	<i>Tetranychus</i>	<i>pueraricola</i>	33	33	5	Japan	(Suwa & Gotoh 2006)
		<i>urticae</i>	10	10	7-9	Europe	(Vaz Nunes <i>et al.</i> 1990)
		<i>urticae</i>	5	5	6-8	Japan	(So & Takafuji 1992)
		<i>urticae</i>	8	8	7-12	Europe	(Koveos <i>et al.</i> 1993)
		<i>urticae</i>	6	6	5-11	Japan	(哲雄 & 徳純 1981)

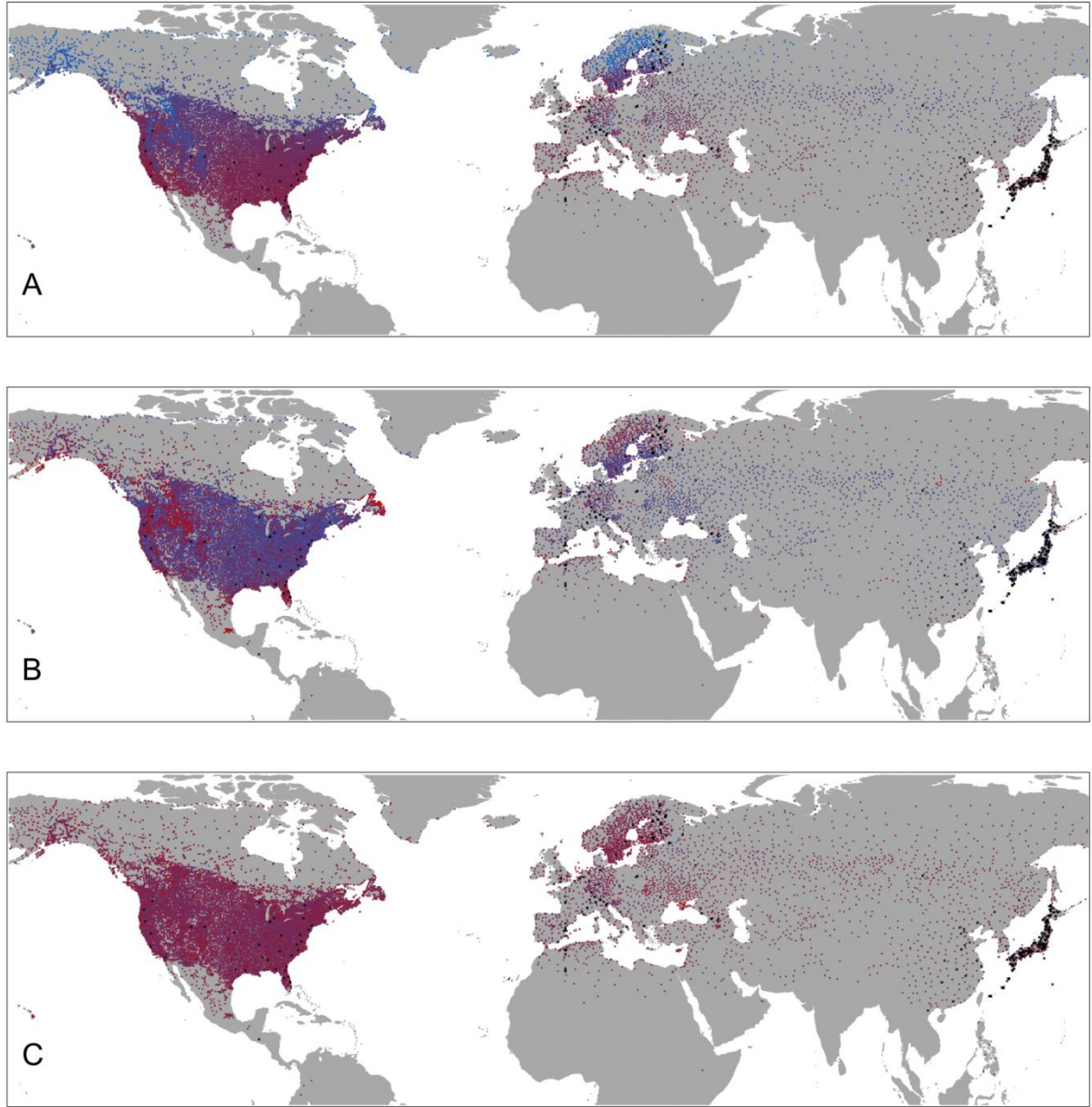


Fig. 3.: Winter onset calculation based on GHCN-daily climate data. A: Mean winter onset, **B:** Standard deviation in winter onset (day length predictability). $Sd \geq 30$ received same color as $sd = 30$; **C:** Color of noise. Capped at $\beta=2$. **Black:** sampling locations of empirical studies.

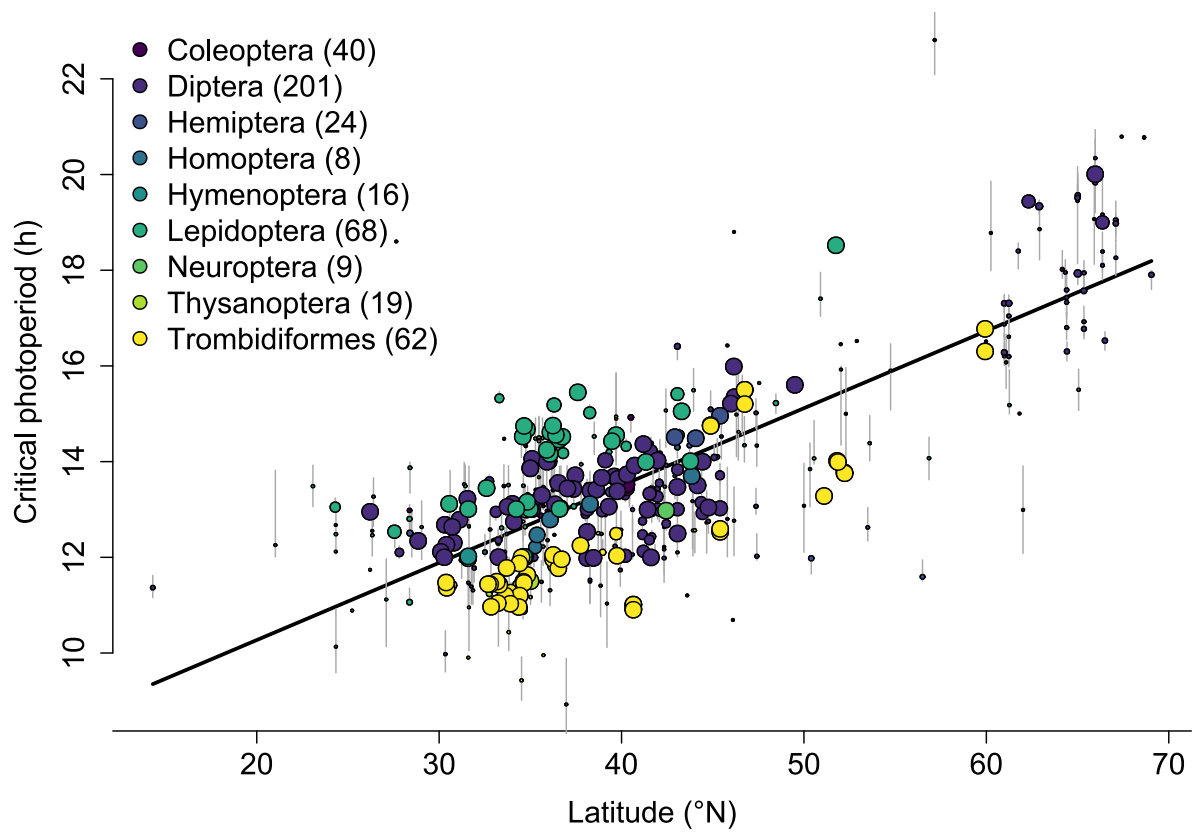


Fig. 4A. Critical photoperiod from primary studies versus latitude. Pointsizes are proportional to the inverse variances. credible intervals of the estimates are displayed as grey lines, credible intervals > 2h are suppressed. Black line: main effect, slope = 48.36 min/5°N

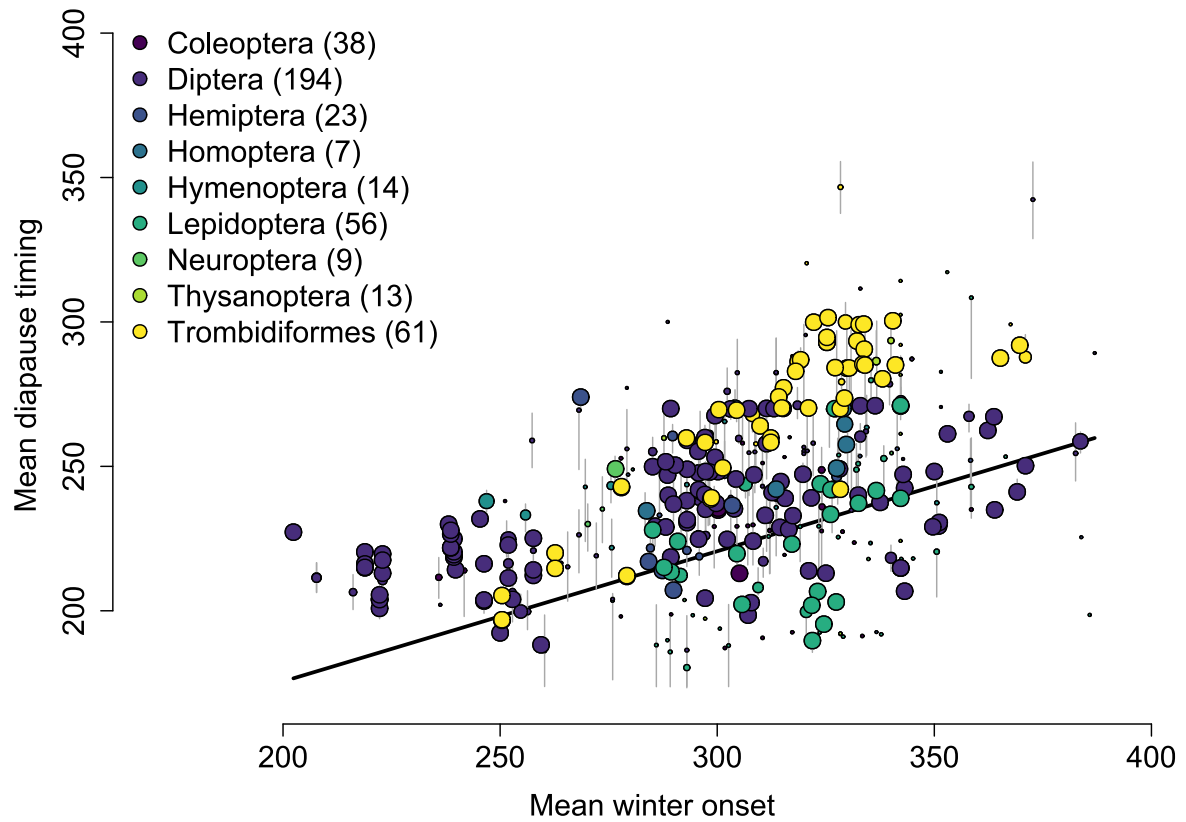


Fig. 4B: Correlation of mean diapauses timing with mean winter onset. Pointsizes are proportional to the inverse variances. credible intervals of the estimates are displayed as grey lines, credible intervals > 1 month are suppressed. Black line: main effect, slope = 0.52.

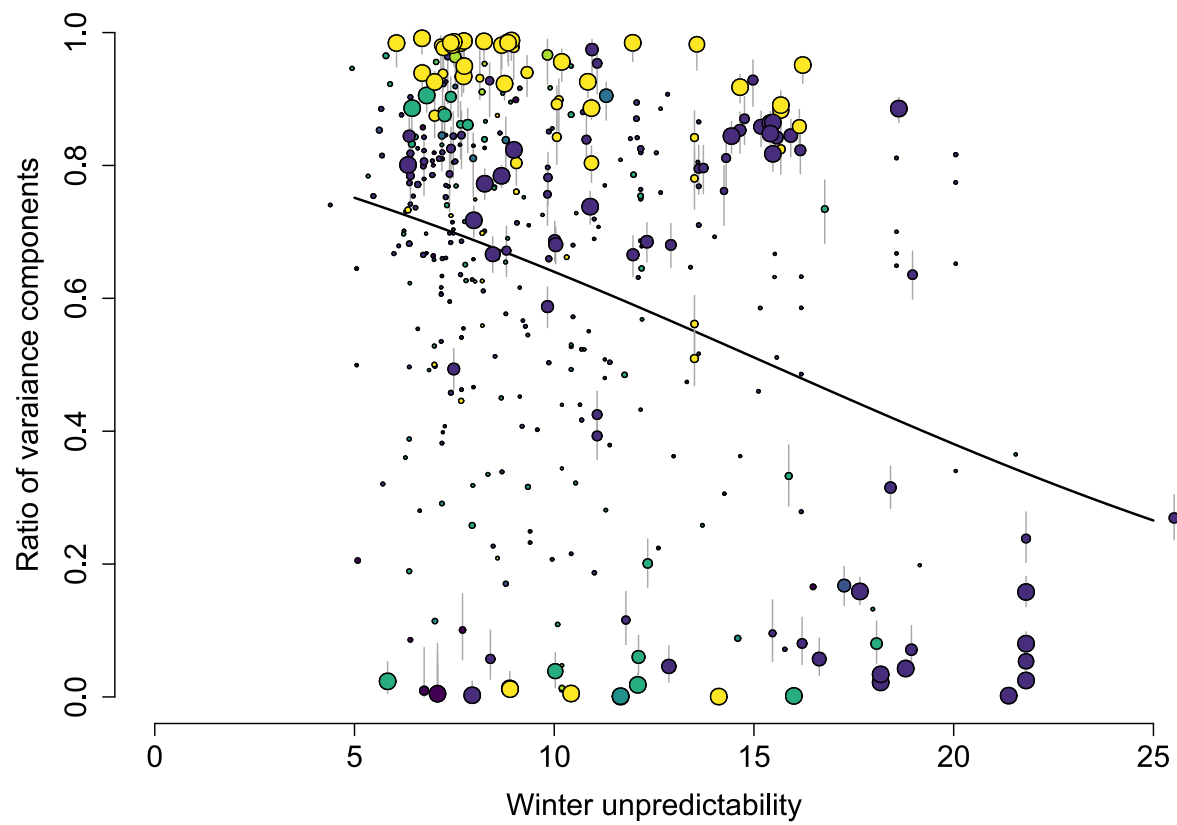


Fig.4C: Variance composition versus day length predictability. Ratio of 0 = pure bet-hedging, ratio of 1 equals pure plasticity. Credible intervals > 10 percent points are suppressed. Black line: main effect, slope = 0.47

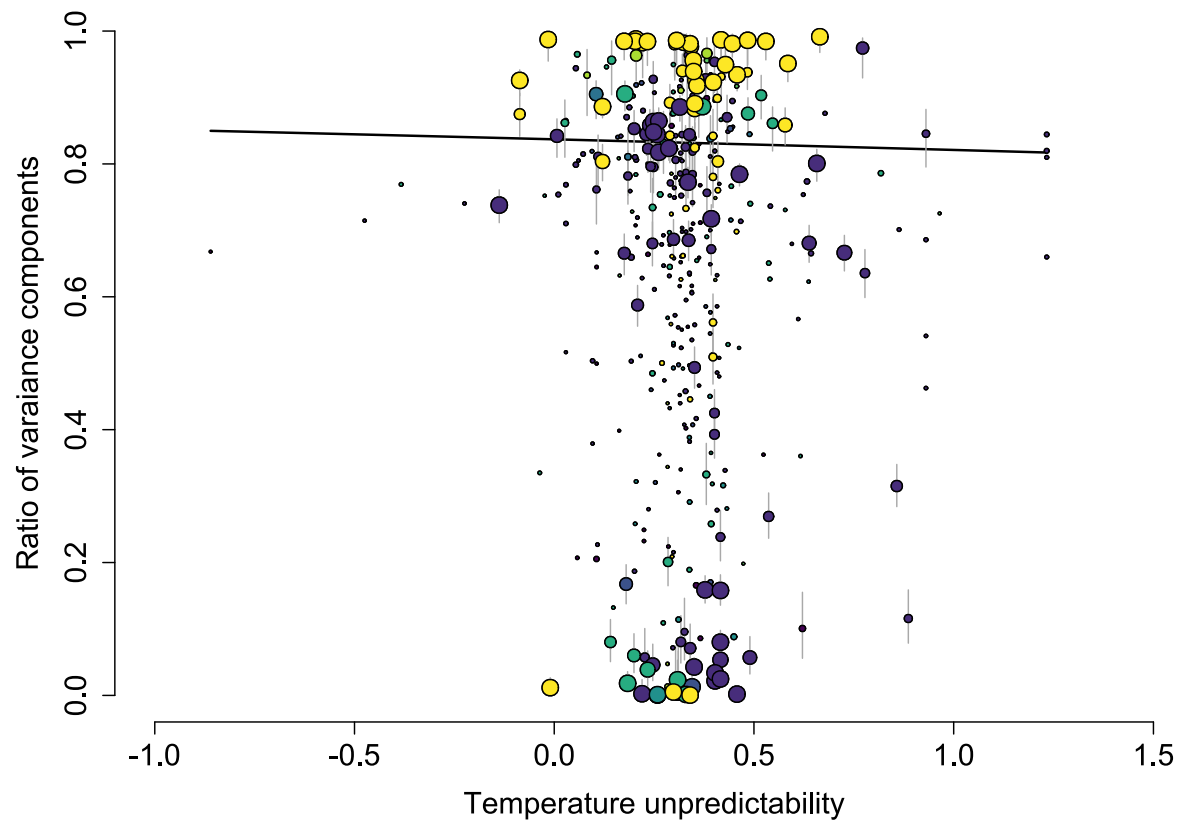


Fig.4D: Variance composition versus temperature unpredictability. Ratio of 0 = pure bet-hedging, ratio of 1 equals pure plasticity. Credible intervals > 10 percent points are suppressed. Black line: main effect (not significant)

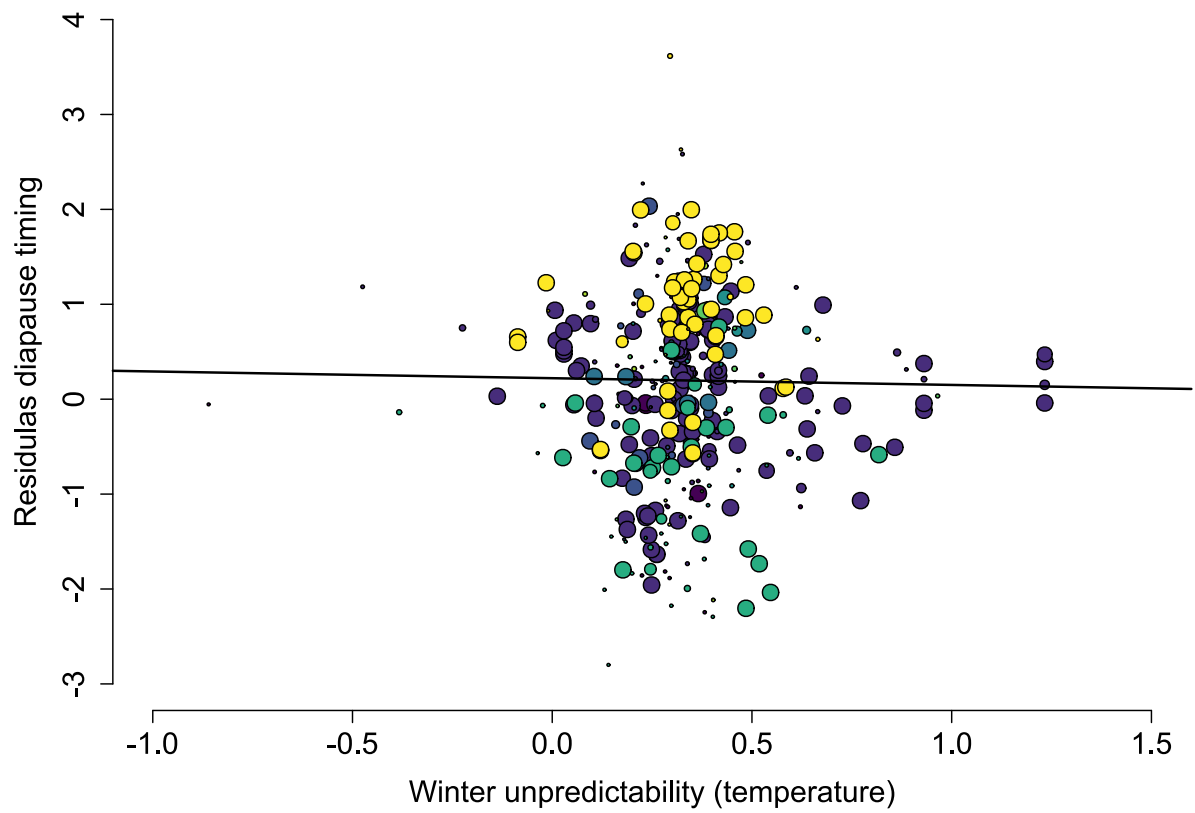
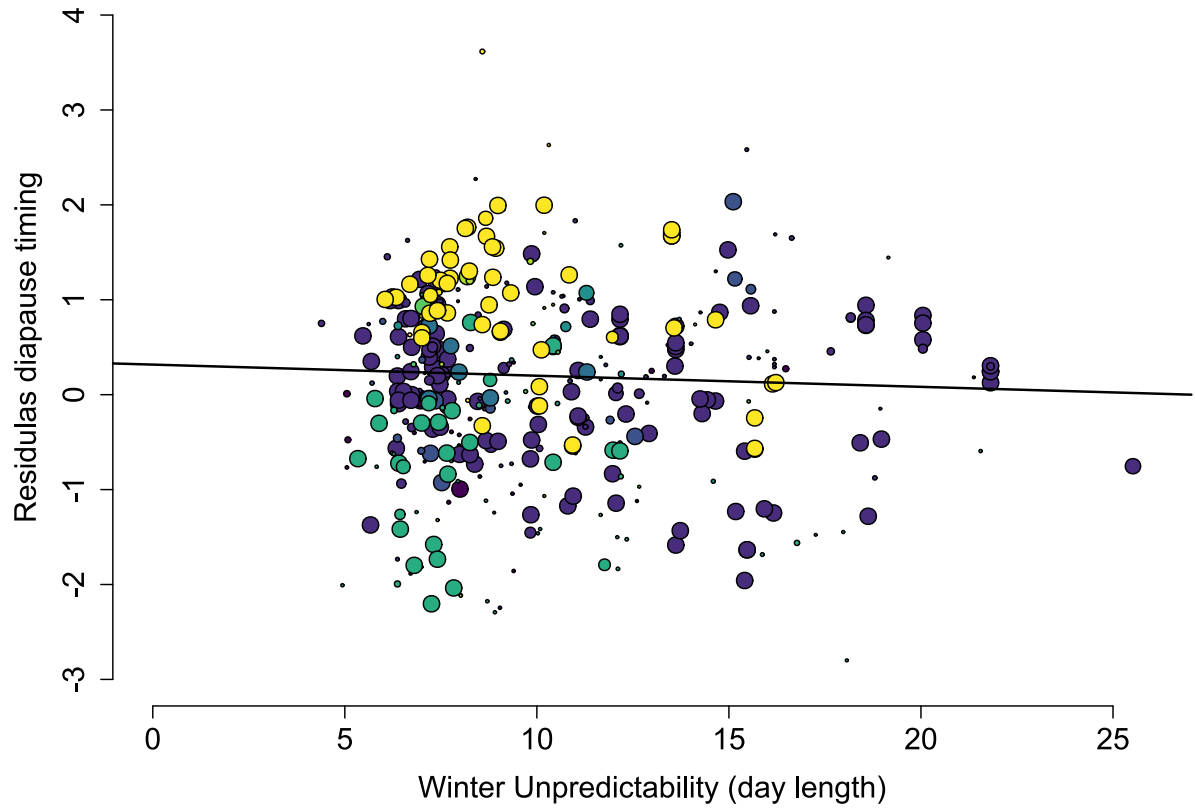
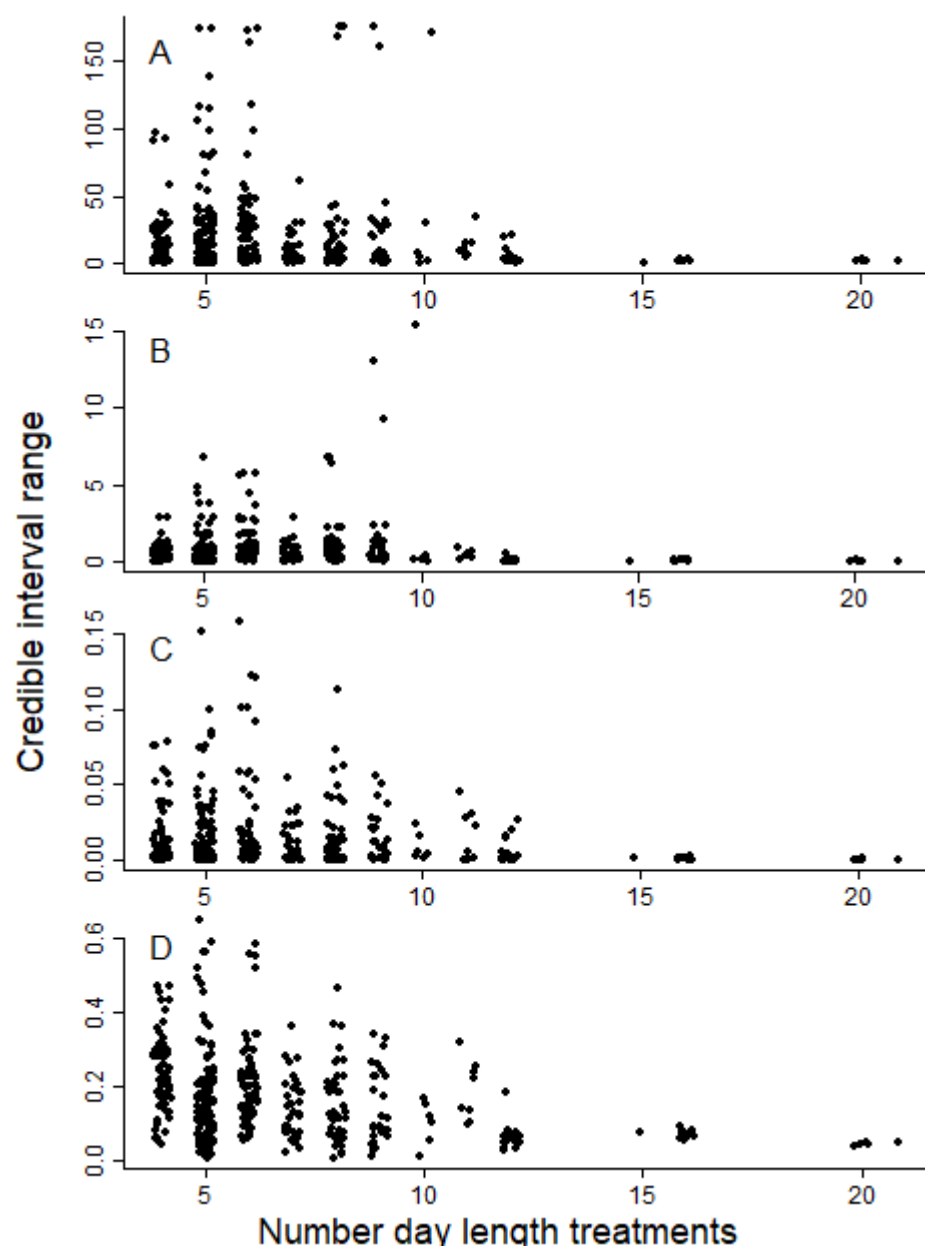
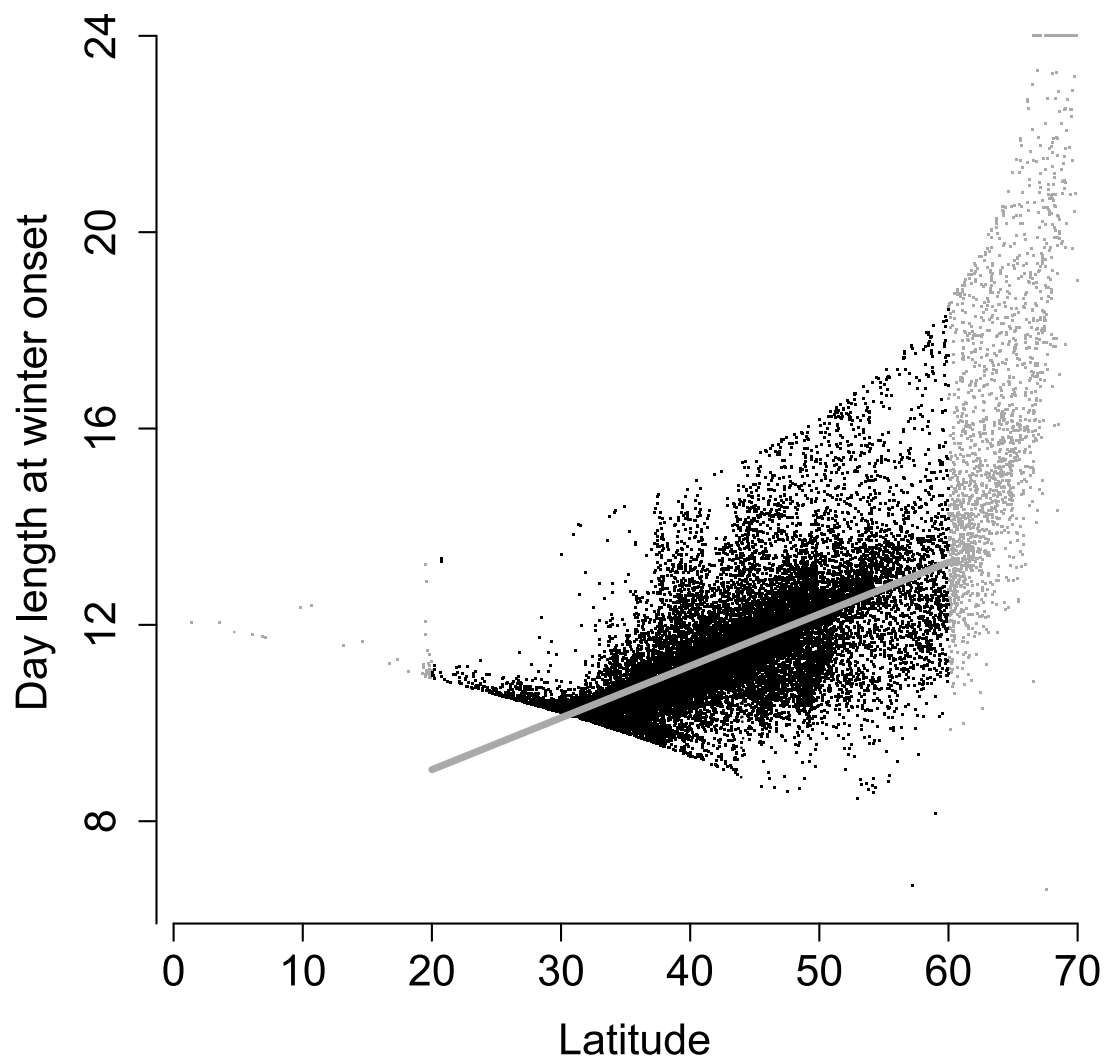


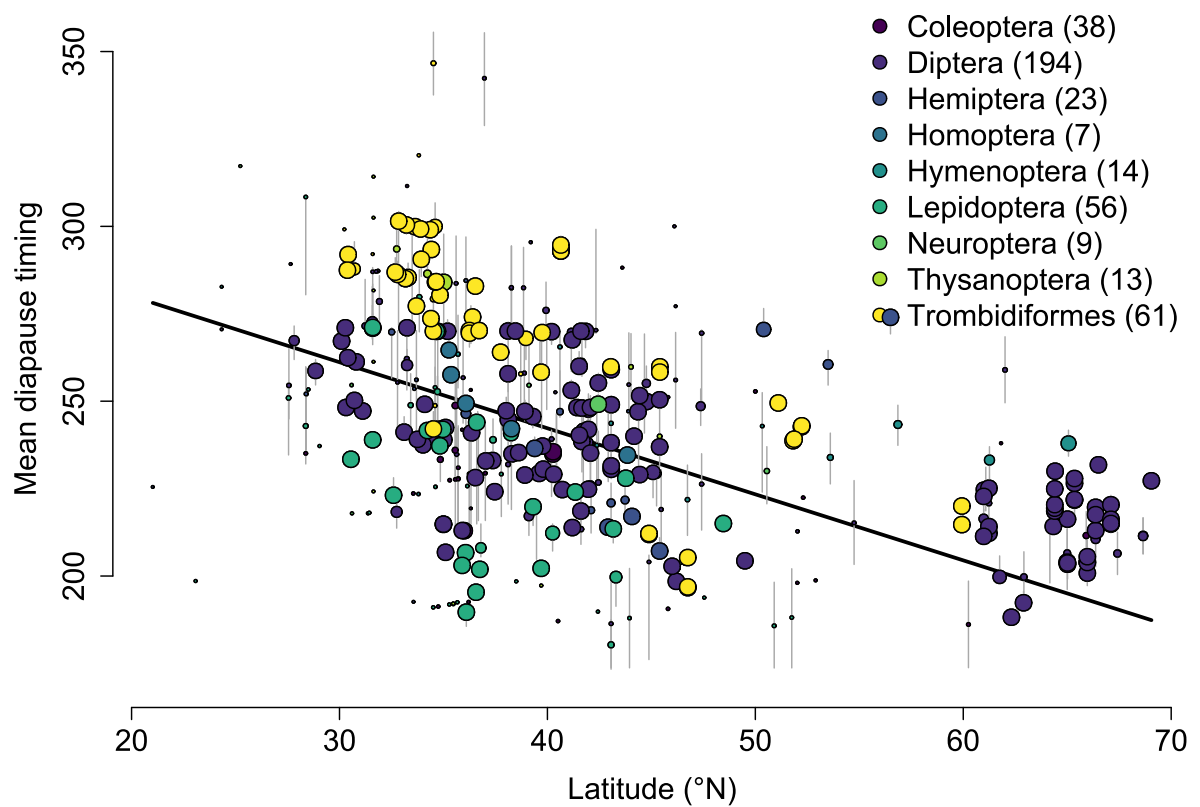
Fig. 5: Residual deviation from diapause timing against winter predictability (conservative bet-hedging). A: day length predictability; B: temperature predictability.



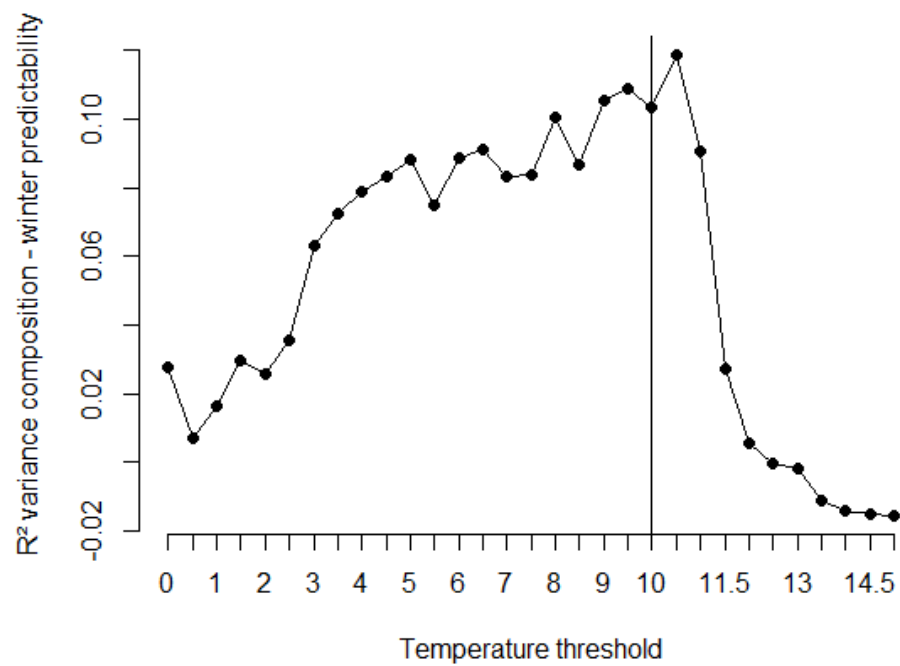
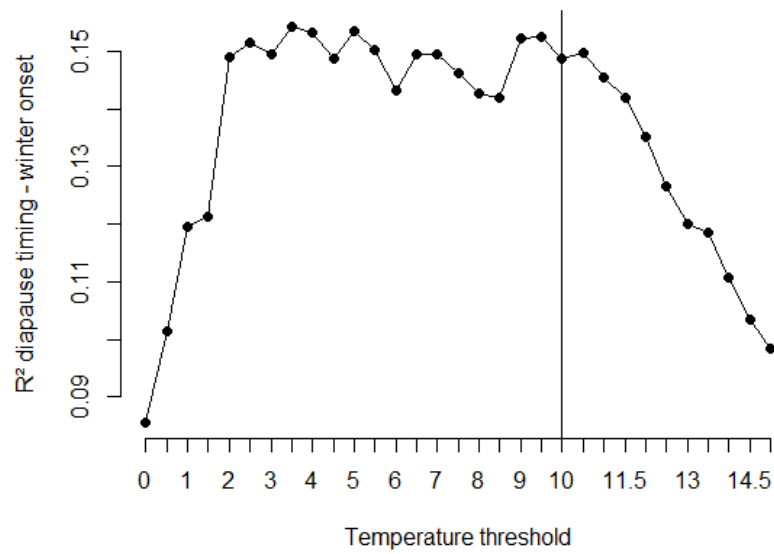
Supp. S3. Funnel plots of mean timing (A), critical day length (B), responsiveness (C) and variance composition (D).



Supp S4: Day length at winter onset vs latitude. Grey line: linear prediction between 20 and 60°N, grey points = points outside this prediction.



Supp. S5. Correlation of mean diapause timing with latitude. Pointsizes are proportional to the inverse variances. credible intervals of the estimates are displayed as grey lines, credible intervals > 1 month are suppressed. Black line: main effect



Supp S6: Sensitivity of the meta-analysis to threshold choice. The meta-analysis was repeated for parameter choices between 0 and 15. Panel A shows R^2_{total} for model 1 (Mean diapause timing vs. mean winter onset), panel B for model 2a) (variance composition vs. day length predictability).