PHILOSOPHICAL TRANSACTIONS B

Research

Using museum specimens to track morphological shifts through climate change

Heidi J. MacLean[1,2], Matthew E. Nielsen[1], Joel G. Kingsolver[1] and Lauren B. Buckley[3]

[1]Department of Biology, University of North Carolina at Chapel Hill, Coker Hall 120 South Road, Chapel Hill, NC 27599, USA
[2]Department of Bioscience, Aarhus University, Ny Munkegade, 8000, Aarhus C, Denmark
[3]Department of Biology, University of Washington, 24 Kincaid Hall, Seattle, WA 98195, USA

JGK, 0000-0002-9839-0208; LBB, 0000-0003-1315-3818

Museum specimens offer a largely untapped resource for detecting morphological shifts in response to climate change. However, morphological shifts can be obscured by shifts in phenology or distribution or sampling biases. Additionally, interpreting phenotypic shifts requires distinguishing whether they result from plastic or genetic changes. Previous studies using collections have documented consistent historical size changes, but the limited studies of other morphological traits have often failed to support, or even test, hypotheses. We explore the potential of collections by investigating shifts in the functionally significant coloration of a montane butterfly, *Colias meadii*, over the past 60 years within three North American geographical regions. We find declines in ventral wing melanism, which correspond to reduced absorption of solar radiation and thus reduced risk of overheating, in two regions. However, contrary to expected responses to climate warming, we find melanism increases in the most thoroughly sampled region. Relationships among temperature, phenology and morphology vary across years and complicate the distinction between plastic and genetic responses. Differences in these relationships may account for the differing morphological shifts among regions. Our findings highlight the promise of using museum specimens to test mechanistic hypotheses for shifts in functional traits, which is essential for deciphering interacting responses to climate change.

This article is part of the theme issue 'Biological collections for understanding biodiversity in the Anthropocene'.

1. Introduction

The evidence for organismal responses to climate change predominantly consists of shifts in phenology, distribution and abundance [1]. However, some studies have shown that species' morphologies are also shifting in response to climate change. These studies are often limited to exploring shifts in body size [2,3], and whether body size shifts are driven by plasticity or evolution is often unknown. Museum specimens represent a largely untapped resource for detecting shifts in functional phenotypes beyond size [4,5]. Coupling phenotype measurements with genetic data, experimental data or models may reveal whether phenotypic shifts result from plasticity or evolution. Here, we explore the prospects for using natural history collections to detect morphological responses of populations and species to climate change via evolution or plasticity.

Multiple, often interacting, responses to climate change can limit phenotypic responses [6]. Shifts through space or time can enable organisms to track their environmental niche and reduce the need for phenotypic change. Plasticity can hinder genetic adaptation by reducing the strength of selection or it can facilitate genetic adaptation by enabling persistence through climate change [7]. Complex life cycles, such as life stages differing in thermal sensitivities, buffering capacities

THE ROYAL SOCIETY PUBLISHING

or habitats, also complicate climate change responses [8]. These interacting responses present challenges for studies using museum specimens, because the specimens generally lack ecological context. However, specimens are often the only source of historical phenotypic information and can be abundant for taxa such as insects, particularly butterflies, which have been favoured by collectors.

Supplementing morphological data from museum specimens with other types of data can clarify responses to climate change by providing insight into ecological and evolutionary context [9]. For example, quantifying levels of stress hormones in feathers addresses whether organisms face environmental stress including diet limitation [10]. An isotopic study of penguin feathers revealed diet shifts over the past century [11]. An analysis of the pollen composition on declining bee species revealed relatively low taxonomic diversity in the pollen of declining bumblebees [12]. Physiological evolution has been documented by repeating historic experiments or reviving dormant eggs [13,14], but inferring physiology directly from specimens is difficult.

Declines in body size have been a widespread response to climate change in both endotherms and ectotherms [2,3] and illustrate the potential of using natural history collections. Larger body sizes in ectotherms typically result in higher equilibrium body temperatures, but also increase thermal inertia and potentially reduce thermal stress in fluctuating environments [15]. Larger organisms incur higher overall, but lower mass-specific, metabolic demands. Owing to the exponential temperature dependence of ectotherm metabolic rates, the same amount of warming will tend to cause greater increases in metabolic rate in warm locations [16]. For example, several salamander species have decreased their body sizes consistent with the greater metabolic costs in a warmer climate [17]. Nevertheless, body size is associated with a host of life-history and other traits [18] and as a result, the adaptive significance of size reductions during climate warming is not always clear. What studies exist of morphological shifts for other traits that may have clearer adaptive significance?

We reviewed the literature to identify studies that used museum collections to document animal morphological changes other than size associated with climate change. The 16 studies meeting our criteria (electronic supplementary material, Appendix S1) focused on vertebrates (primarily rodents and birds) and shifts in coloration, cranial morphology or body shape in response to temperature or precipitation. Studies of traits other than coloration rarely (4/11) made explicit hypotheses for responses to climate change and seldom (1/11) found strong responses (electronic supplementary material, table S1). On the other hand, coloration studies consistently and often accurately predicted climate change responses, aided by mechanisms such as the direct link between the degree of pigmentation and solar absorptivity of invertebrates [19,20] (table 1). All but one of the five studies of coloration predicted observed shifts in the frequency of discrete colour morphs with a known genetic basis (established by non-collection-based research either prior to or as part of the study), allowing them to attribute the observed changes to evolution.

Here we further explore the exceptional case that failed to predict shifts in coloration: our previous research on the continuous colour variation of *Colias meadii* butterflies in Colorado, USA [21]. *Colias* body temperatures are influenced by two key thermoregulatory traits: (i) increased melanism on the posterior ventral hindwings allows the butterflies to absorb more solar radiation via closed wing basking and (ii) longer setae on the thorax reduce convective heat loss [26,27]. Body temperatures influence fitness by determining the duration of flight, which is thermally restricted [26,27] and required for activities including mating, feeding and oviposition [26,28]. However, butterflies may overheat, which not only limits flight time, but can also directly reduce the number of viable eggs produced by each female [26]. Wing melanin is both heritable and increases with elevation for *Colias* species [29], providing landscape-scale support for its thermoregulatory function. Montane *Colias* also exhibit plasticity whereby increased temperature during pupal development decreases melanism [13,30]. The plasticity makes it difficult to infer whether shifts in wing coloration in response to climate change have a genetic basis. Wing melanism is a rare trait in being both observable in museum specimens and having calculable fitness consequences.

We can make specific predictions about the morphological and demographic responses of *Colias* to climate change using a set of microclimate, heat balance, demographic and evolutionary models that we have developed and validated [31]. Can these phenotype-based models help understand shifts in quantitative phenotypes influenced by both genetics and plasticity? The models predict that recent warming in Colorado has increased both flight time and overheating risk for *C. meadii*, resulting in fitness declines at lower elevations and fitness increases at higher elevations in the absence of trait evolution [28]. We estimate that directional selection shifts from favouring wing lightening to favouring wing darkening with increasing elevation across *C. meadii*'s distribution but that seasonal and annual variation in climate causes the strength and direction of selection to fluctuate [32]. The models suggest that plasticity can facilitate evolution, particularly at lower elevations with long seasons, by dampening variation in selection [33]. Phenological shifts associated with changes in developmental rate can also reduce variation in selection [34].

Our initial analysis of historic *C. meadii* specimens from the southern part of the species' range (Southern Rocky Mountain (RM) region below) underscored how environmental variability and interacting responses can complicate morphological shifts [21]. Regression models indicated that the two thermoregulatory traits (wing melanism and setae length) increased in the direction that would increase body temperature despite climate warming and additional analysis suggested a phenological shift [21]. Analysis using multiple regression models suggested melanism increased significantly both over time and in cool seasons. Pupal temperature, date of collection and an interaction between pupal temperature and date of the collection also received support for model inclusion. Estimates of the temporal trend may have been influenced by interactions with these other predictors.

These previous results illustrate the potentially complex interactions among phenological shifts, plasticity and evolution in response to climate change. We examine these interactions using an expanded set of museum *C. meadii* specimens, spanning two additional geographical regions. We first test three hypothesized shifts over recent decades: (i) temperatures during the development period have increased, reflecting regional climate warming; (ii) warmer developmental temperatures have advanced adult phenology; and (iii) thermoregulatory trait values have declined to reduce the heat load associated with warming. However, these declines may be lessened if phenological shifts enable

Table 1. Summary of previous studies using natural history collections to assess coloration shifts. We include the predicted and observed responses to climate change and whether they concur. See electronic supplementary material, table S1 for additional morphological traits.

specific traits	taxa	method	time range	region	changes considered	predicted response to climate change	predicted response to climate detected?	source(s)
wing melaninsm, thorasic setal length, body size	Colias meadii (Insecta: Lepidoptera)	collection survey	1953–2012	Colorado, USA	temperature	decreased wing melanism setal length, and body size	no, increase in wing melanism, setal length, and size over time with increasing temperatures	[21]
shell melanism	Cepaea nemoralis (Gastropoda)	resampling	1967 versus 2010	the Netherlands (specific site)	temperature, increased forest cover	increased frequency of light morphs	yes, increase in frequency of one light yellow morph, other changes in morph abundance variable among patches	[22]
shell melanism	Cepaea nemoralis (Gastropoda)	resampling	various dates (1915–1962) versus 2010	the Netherlands	temperature	increased frequency of light morphs	yes, increase in lighter morphs at two of three sampled locations	[23]
plumage coloration	Strix aluco (Aves: Strigiformes)	Resampling[a]	1915–1980	Finland	snow cover	increased frequency of brown morph with reduction in snow cover	yes, increased frequency of brown morph, associated with decreases in snow cover	[24]
plumage coloration	Otus scops (Aves: Strigiformes)	collection survey	1870–2007	Italy	temperature, precipitation	increased frequency of 'dark-reddish' morph	yes, increase in 'dark-reddish' morph, explained in part by climate changes	[25]

[a]Collections only used for historical baseline in this study, changes documented by long-term observations of a specific population.
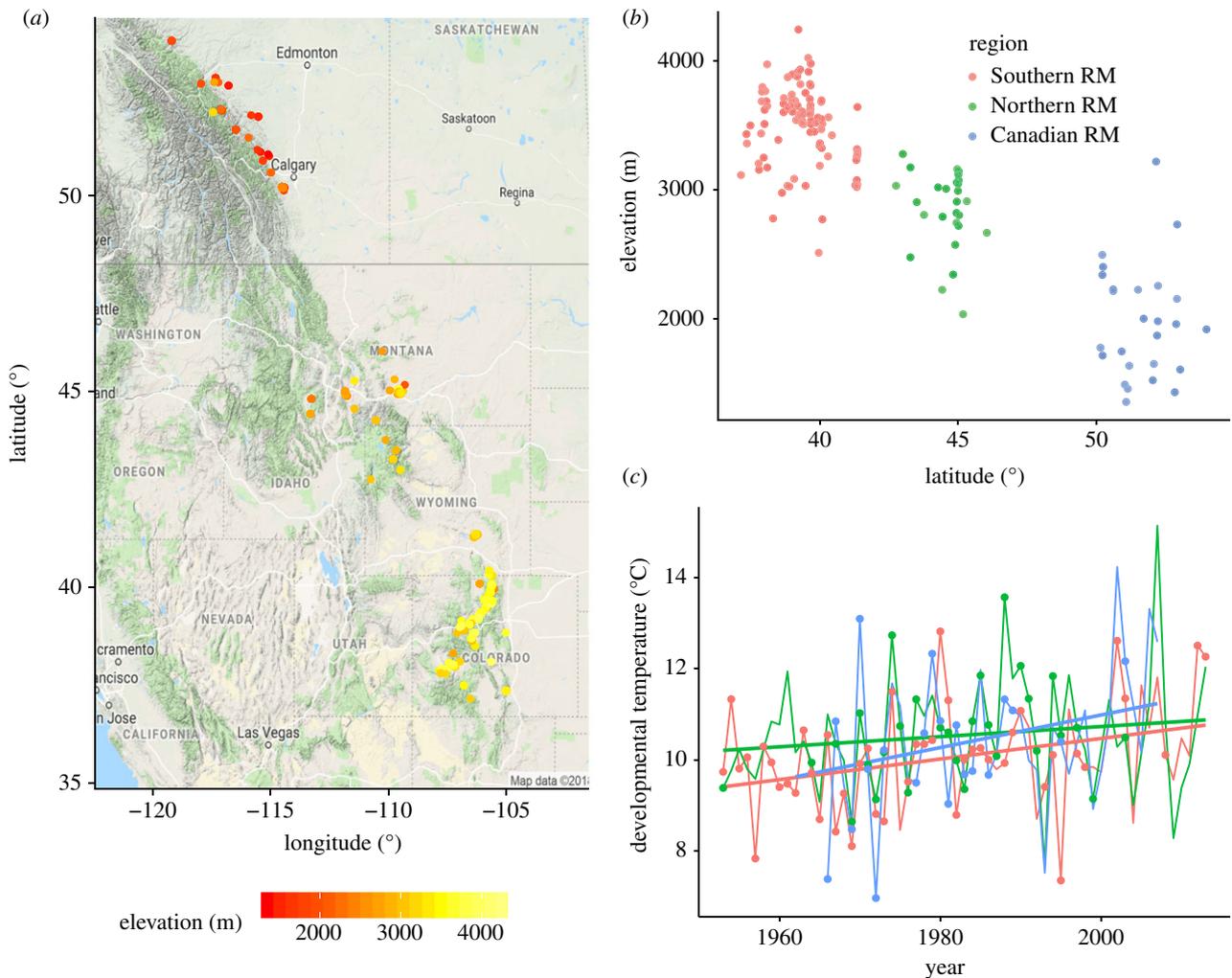
Figure 1. (a) *Colias meadii* specimens (and populations) cluster in three regions according to topography: the Southern, Northern and Canadian Rocky Mountains. (b) *C. meadii* declines in elevation with increasing latitude. (c) Temperatures during the developmental period (days 162–202) have increased in recent decades in each region (increase non-significant in Northern RM). We depict temperature data from representative weather stations (Northern RM region adjusted for elevation). Dots indicate years with specimens in each region. On the map, heat colours indicate the elevation of specimens.

temperature tracking. We then examine how phenology, temperature and year interact to influence thermoregulatory traits. In particular, we look for evidence that higher pupal temperatures decrease melanism via plasticity. The specimen data offer a unique opportunity to examine phenotypic shifts, but we acknowledge that we cannot control for some factors that may vary over long time scales such as habitat, host plant availability or cloud cover.

2. Methods

(a) Trait measurements

Colias meadii occurs in subalpine and alpine meadows above 2.5 km elevation throughout the Rocky Mountains (RM), from northern New Mexico, USA to southern Alberta, Canada [27]. They have a single generation each year (univoltine) and diapause over the winter as 3rd instar larvae; the adult flight season lasts four to six weeks in July and August (between days 180 and 240). Individual butterflies live only a few days and the flight season in any particular location is limited to several weeks, but the timing of the flight season can differ among locations. We use date of collection to approximate adult phenology, but we note that it may be compromised by sampling biases. We measured (methods below) the ventral hind-wing melanism, thorax setae length and wing length (proxy for

body size) for historical *C. meadii* specimens collected from 1953 to 2013 in three regions: Southern RM, Northern RM and Canadian RM (figure 1). We analyse the regions separately because they are genetically distinguished with clear evidence of restricted gene flow between all three regions [35]. Maximum-parsimony models constructed using mitochondrial DNA confirm that these three regions are genetically distinct, with the Northern RM showing signs of earlier isolation [36]. We additionally focus on a single Southern RM site that is well-sampled: Loveland Pass, Colorado, USA (39.6636°N, 105.8792°W, 3655 m).

We processed specimens from seven museums in the United States and one private collector (details in electronic supplementary material, Methods in Appendix S2). Because the collections contained relatively few females, we restricted our analyses to males (Southern RM: $N = 493$; Northern RM: $N = 205$; Canadian RM: $N = 163$). Specimens for the Southern RM are those analysed previously [21] with the addition of some contemporary field-collected specimens. The similar elevation clines between sexes [37] suggest that they respond to climatic gradients similarly and our observations should be indicative of both sexes. All specimens collected during or after 1953 (owing to the sparsity of earlier specimens and lack of climate data) were analysed except for two Colorado specimens with anomalously early collection dates (electronic supplementary material, Appendix S2).

Specimens were photographed along with black and white standards through a 100 mm macro lens in RAW format with a

5

rstb.royalsocietypublishing.org   Phil. Trans. R. Soc. B 374: 20170404

Canon Rebel XSi mounted on a copy-stand. We developed a MatLab program to assess the degree of wing melanism on the posterior ventral hind-wing within a triangulated region between the eyespot, hind-wing insertion and the wing margin [38,39]. Wing scales are pigmented by either melanin (dark) or pterin (light). We calculated a grey level value between 0 (white) and 1 (black), which reflects the proportion of melanic scales. We used a spectroreflectometer to verify grey level ($n = 60$, correlation $R^2 = 0.78$; see the electronic supplementary material, Appendix S2). A digital micrometer was used to measure the forewing of the specimen from the thoracic insertion to the apex of the wing. We measured setal length directly on the ventral thorax with an ocular micrometer on a Wild M5 microscope as the longest setae between the first and second leg.

(b) Climate data

For each region, we selected a representative weather station that occupies a high, relevant elevation and offers a near complete record. For the Southern and Northern RM, we used stations from the National Weather Service Cooperative Observed Program (COOP) at Climax, CO ($39.37°$ N, $106.18°$ W, 3514 m) and Moran 5WNW, WY ($43.86°$ N, $110.59°$ W, 2072 m), respectively (data retrieved from www.ncdc.noaa.gov). For the Canadian RM, we used the Blue Hill Lookout, Alberta station ($51.70°$ N, $115.22°$ W, 1951 m, data retrieved from http://climate.weather.gc.ca). The Moran 5WNW weather station is substantially below most *C. meadii* habitat in the Northern RM (25th, 50th and 75th percentiles: 3019, 3057 and 3144 m, respectively) and thus exhibits warmer temperatures than the other weather stations. We used an environmental adiabatic lapse rate of $6.5°C/1000$ m to adjust the data up 500 m to facilitate plotting while not influencing results. (Shifting up 1000 m resulted in colder temperatures than other sites.) We chose to use daily maximum and minimum temperature data from a representative weather station for all specimens within each region because long-term data are lacking for most high-elevation stations and, while interpolated datasets offered more complete coverage, they can perform poorly at capturing temporal variation for high-elevation sites in the Rocky Mountains [40].

We consider two temperature metrics as predictors: pupal (T_p) and developmental (T_d) temperatures. Based on previous studies in the closely related *Colias eriphyle* [21], we predicted that T_p might alter wing melanism via plasticity. We expect T_d (encompassing late larval and pupal development) to influence phenology and forewing length, a proxy for body size [41]. Pupation lasts for 14 days at $20°C$ for the closely related *C eriphyle* [13]. Slightly expanding the window to account for uncertainty in this duration, we estimated T_p as the average of mean daily temperatures 26 to 6 days prior to the collection date. We further expanded the window to include late larval development and estimated T_d as the average of mean daily temperatures 46 to 6 days prior to the mean flight date (day 208, i.e. days 162–202). We used fixed dates to estimate T_d to indicate annual conditions and avoid circularity in predicting phenology. We confirmed that our results are qualitatively robust by varying the day range used to estimate pupal and developmental temperatures. We focus on temperature, but note that shifts in cloudiness and radiation could also influence plasticity and evolution [21].

(c) Data analysis

We used maximum-likelihood spatial autoregressive (SAR) models [42,43] with bootstrapping to account for sampling biases and trait similarity owing to geographical proximity (function spautolm in R package spdep; see the electronic supplementary material, Appendix S2). We used an SAR error model with a 40 km threshold distance. We used a Moran's *I*-test on the model residuals to confirm that the SAR models fully accounted for spatial autocorrelation [43]. We compared the SAR model output

to linear models and found no qualitative differences in the direction of the effects. The SAR model equations were either response ∼ predictor or response ∼ predictor$_1$ × predictor$_2$.

We accounted for uneven collection intensities by restricting our analysis to 15 individuals per site per year. We randomly sampled specimens each of the 50 times we repeated the analysis and averaged across the iterations. Our inclusion of many collectors (by visiting numerous collections; electronic supplementary material, Appendix S2) was designed to further limit spatial and (within and across the season) temporal biases. We prioritized sampling specimens in the Southern RM because it was the location of our contemporary empirical research.

3. Results

We sampled *C. meadii* specimens spread across the three regions. Sampling was more extensive in the Southern RM than in the Northern or Canadian RM (figure 1), particularly in more recent years. Specimen elevation declines with latitude, at least in part owing to declines in the elevation of suitable habitat. We find that mean daily temperatures during the developmental period (days 162–202) have increased in recent decades (1953–2013) in all regions (Southern RM: slope $\pm$ s.e.: 0.027 ± 0.0040, $z = 6.7$, $p < 10^{-9}$; Northern RM: slope $\pm$ s.e.: 0.053 ± 0.0065, $z = 8.2$, $p < 10^{-13}$; Canadian RM: slope $\pm$ s.e.: 0.095 ± 0.010, $z = 9.2$, $p < 10^{-17}$; figure 2*a*). However, when we look at this trend using all years, not just those with observations, this trend is only significant for the Southern and Canadian RM (Southern RM: slope $\pm$ s.e.: 0.023 ± 0.0078, $t = 2.9$, $p < 0.01$; Northern RM: slope $\pm$ s.e.: 0.011 ± 0.0091, $t = 1.2$, $p = 0.23$; Canadian RM: slope $\pm$ s.e.: 0.036 ± 0.017, $t = 2.1$, $p < 0.05$; figure 1).

We detect shifts in wing melanization across recent decades corresponding to increases in development temperature (figure 2*b*). The direction of shifts, however, has varied across regions: wings have darkened in the Southern RM (slope $\pm$ s.e.: 0.0013 ± 0.00026, $z = 5.1$, $p < 10^{-5}$), but lightened in the Northern (slope $\pm$ s.e.: -0.0026 ± 0.00072, $z = -3.7$, $p < 0.001$) and the Canadian RM (slope $\pm$ s.e.: -0.0017 ± 0.00048, $z = -3.6$, $p < 0.001$).

We next investigate whether phenological or phenotypic plasticity can account for the shifts in melanization. We detect suggestive, but non-significant, advancements in the phenology of adult flight across time in the Northern and Canadian RM (Southern RM: slope $\pm$ s.e.: 0.021 ± 0.046, $z = 0.5$, $p = 0.64$; Northern RM: slope $\pm$ s.e.: 0.13 ± 0.064, $z = 2.1$, $p = 0.09$; Canadian RM: slope $\pm$ s.e.: -0.14 ± 0.073, $z = -1.97$, $p = 0.054$; figure 2*c*). When analysing phenology as a function of developmental temperatures rather than time, we find that phenology advances significantly with warmer developmental temperatures in all regions (figure 3*a*). Multiple regression models support shifts in phenology with both developmental temperatures and time (table 2).

Statistical evidence for plasticity in wing melanism is ambiguous (table 2). We hypothesized that increases in pupal temperature (or later phenology as a proxy for warmer temperatures in the seasonal environments) would decrease wing melanism via plasticity (as in the closely related *C. eriphyle*). In the Southern RM, we find evidence for delayed phenology in later years increasing wing melanism (table 2). Although this observation is contrary to the hypothesized plastic response, it could result from exposure to cool late-season temperatures owing to the extended
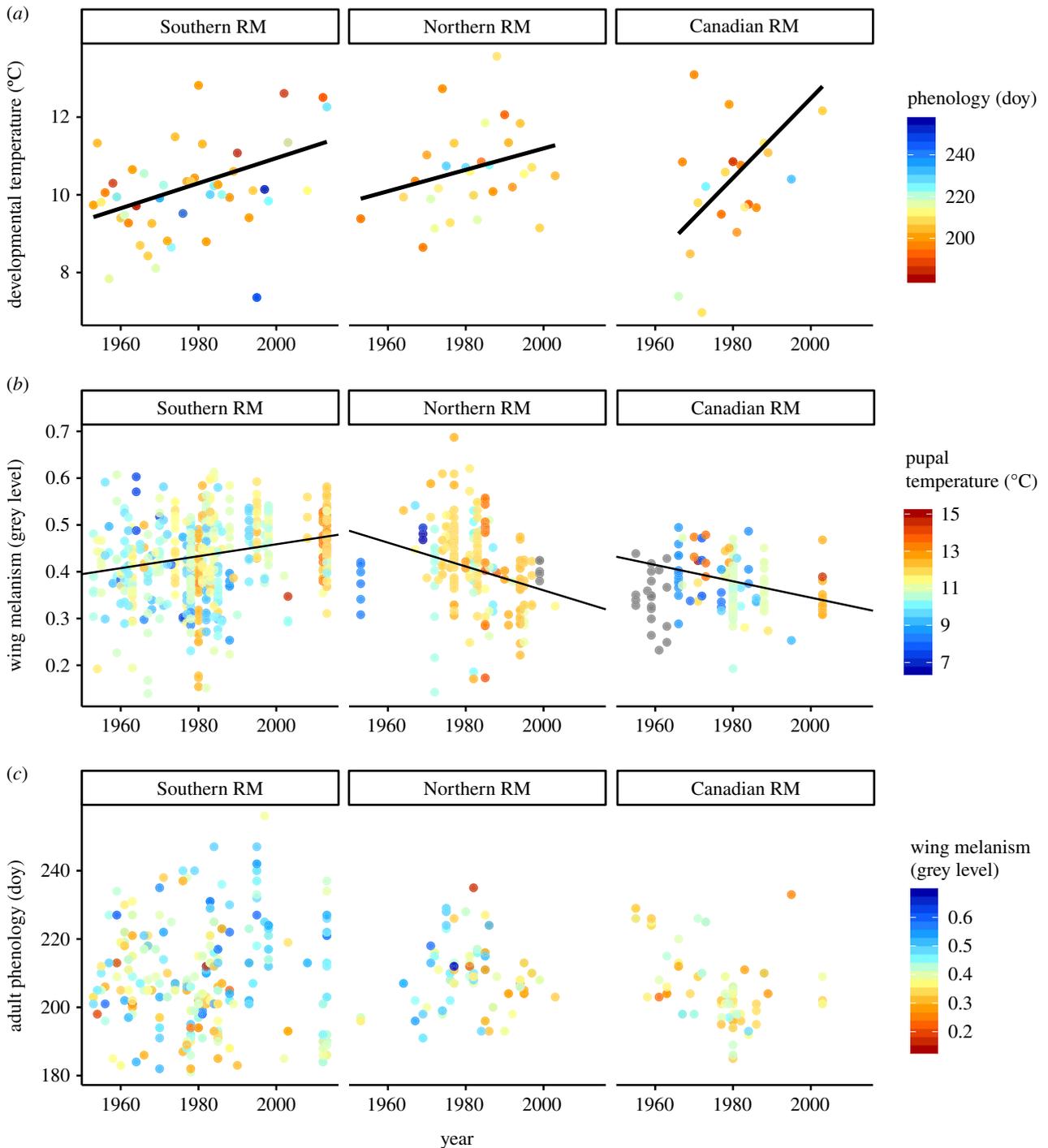
Figure 2. Across the three regions: (a) predicted development temperatures (mean daily temperature across days 162–202) have been increasing over recent decades. We only include years corresponding to specimens within each region. (b) Shifts in wing melanism (grey level) across recent decades vary among regions. (c) *C. meadii* phenology (doy: day of year) has not shifted significantly. Colours indicate interactions with additional variables (*a*: mean phenology cross specimens, *b*: pupal temperatures, *c*: wing melanism; grey indicates no temperature data. Black lines depict significant ($p < 0.05$) regressions estimated from a spatial model with bootstrapping.

flight season. Also contrary to the hypothesis for plasticity, our two predictor analyses estimate that melanism increases with increasing pupal temperatures in the Southern and Northern RM and later phenology in the Northern RM (table 2). Underlying levels of melanism are significant inter-actions among pupal temperature and year (Northern RM) and pupal temperature and phenology (Northern and Canadian RM, table 2). These interactions highlight the challenge of deciphering plastic and evolutionary determinants of wing melanism.

The increases in wing melanism in the Southern RM and decreases in wing melanism in the Northern and Canadian

RM remain when controlling for these interactions. We examine the residuals of the model using pupal temperature, phenology and their interaction to predict wing melanism (figure 3*b*). The observations of increasing wing melanism over time in the Southern RM (slope ± s.e.: 0.0010 ± 0.00026, $z = 4.0$, $p < 0.001$) and decreasing wing melanism in the Northern (slope ± s.e.: -0.0027 ± 0.00071, $z = -3.9$, $p < 0.001$) and Canadian RM (slope ± s.e.: -0.0015 ± 0.00049, $z = -3.2$, $p < 0.01$) persist when attempting to control for plasticity by accounting for these factors.

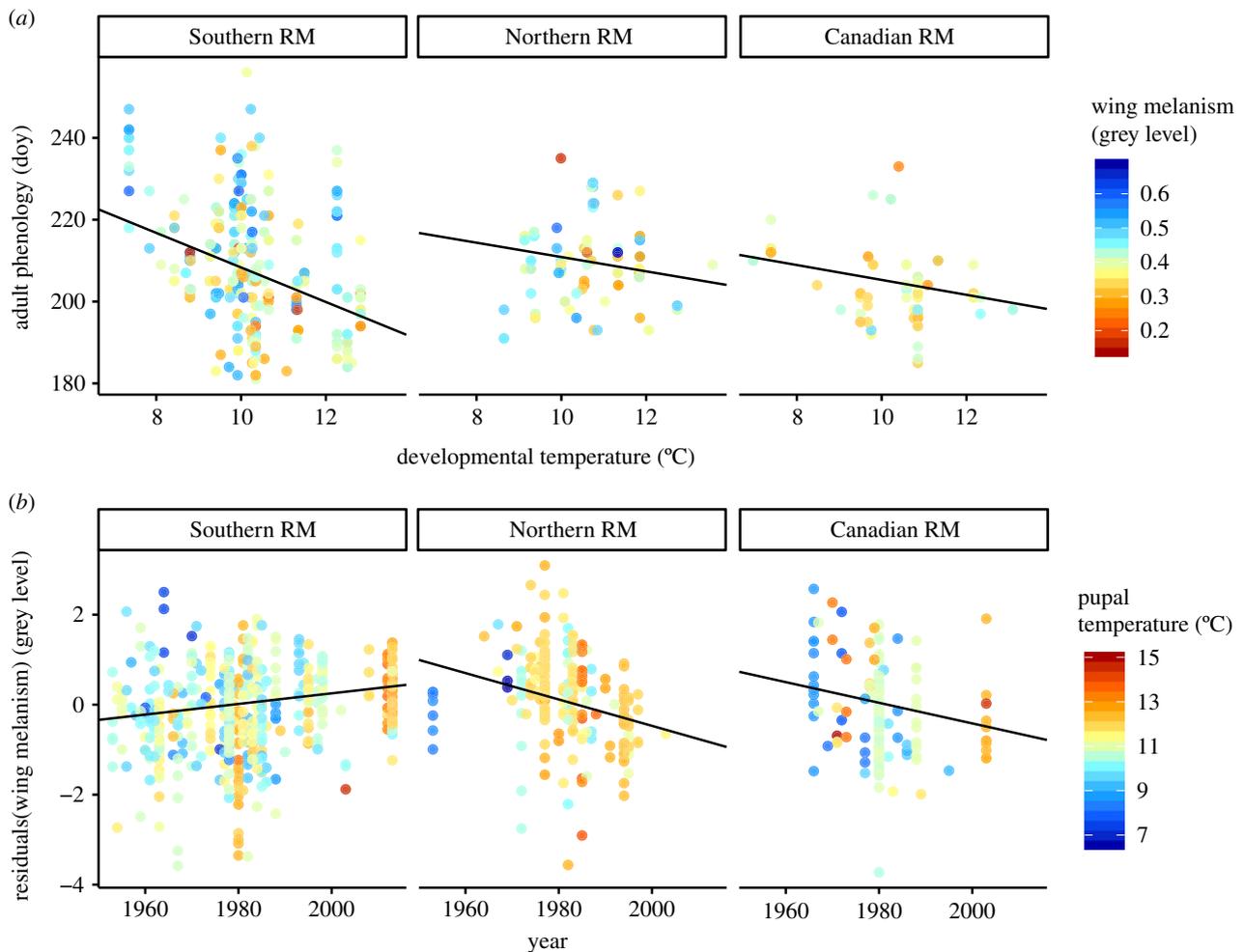We also examine phenotypes across time at a single, well-sampled site: Loveland Pass, CO (electronic supplementary

Figure 3. Across the three regions: (*a*) adult phenology advances with increases in developmental temperatures. We depict (colour) how phenology and developmental temperatures correspond to wing melanism; doy, day of year. (*b*) Observations of increasing melanism in the Southern RM region and decreasing melanism in the Northern and Canadian RM regions over time persist when controlling for plasticity. We control for plasticity by analysing the residuals of the relationship with phenology, pupal temperature and their interaction across the three regions. As in figure 2, we indicate pupal temperatures associated with specimens (colour). Black lines depict significant ($p < 0.05$) regressions estimated from a spatial model with bootstrapping.

material, figure S1). Examining a single site minimizes potential confounding factors but requires analysing a relatively low number of specimens that are clustered temporally in sampling events. Neither phenology (slope ± s.e.: 0.24 ± 0.13, $t = 1.9$, $p = 0.08$) nor wing melanism (slope ± s.e.: 0.00060 ± 0.00049, $t = 1.2$, $p = 0.26$) shift significantly across time (electronic supplementary material, figure S1).

We focus our analysis on wing melanism, but additionally analysed shifts in two additional phenotypic traits related to thermoregulation (electronic supplementary material, figure S2). Forewing length, a proxy for overall body size, has increased over recent decades in the Southern RM (slope ± s.e.: 0.048 ± 0.0052, $z = 9.3$, $p < 10^{-14}$) but has not shifted significantly in the Northern (slope ± s.e.: -0.013 ± 0.011, $z = -1.2$, $p = 0.26$) or the Canadian RM (slope ± s.e.: -0.027 ± 0.022, $z = -1.1$, $p = 0.25$). The length of setae on the thorax, which decreases heat loss, has increased over recent decades in the Southern RM (slope ± s.e.: 0.024 ± 0.0018, $z = 13.5$, $p < 10^{-14}$) but has not exhibited shifts in the less thoroughly sampled regions (Northern RM: slope ± s.e.: 0.0062 ± 0.0033, $z = 1.9$, $p = 0.08$; Canadian RM: slope ± s.e.: 0.0031 ± 0.0080, $z = 0.39$, $p = 0.70$) (electronic supplementary material, figure S2).

4. Discussion

Our study of *Colias* butterflies and review of previous studies highlight some of the difficulties inherent to detecting often complex and interacting responses to climate change using collections. Although specimens represent an unparalleled resource for detecting phenotypic shifts, interpreting the drivers of these shifts and their underlying mechanisms is challenging. Even in the well-studied *Colias* system, parsing out the influence of phenological, plastic and evolutionary responses is elusive. Sampling biases exacerbate this challenge. Nevertheless, recognizing the multiple, interacting responses to climate change and developing approaches for partitioning their contributions are essential to improving our ability to forecast responses to climate change.

Testing predictions of our phenotype-based models provides insight into the drivers of the observed changes in melanism—an increase in the Southern RM but decreases in the Northern and Canadian RM—over recent decades. Despite warming temperatures, we do not observe clear phenological shifts across time. However, we do find evidence that phenology advances in years with warm temperatures during the developmental period (figure 3*a*). These interactions among phenology, temperature and year complicate

8

rstb.royalsocietypublishing.org   Phil. Trans. R. Soc. B 374: 20170404

Table 2. Results of SAR models with bootstrapping examining the correlations underlying phenotypes. Predictor variables were centred and scaled for comparability. T_d and T_p are developmental and pupal temperatures (°C), respectively. Phenology indicates day of year and melanism is quantified as a grey level. Significant predictors and interactions are indicated by asterisks. Region codes: SRM, Southern RM; NRM, Northern RM; CRM, Canadian RM.

response	region	predictor 1				predictor 2					interaction			
		estimate	s.e.	z	p-value		estimate	s.e.	z	p-value	estimate	s.e.	z	p-value
phenology	SRM	T_d −0.49	0.05	−10.0	***	year	0.24	0.04	5.6	***	−0.06	0.04	−1.5	
	NRM	−0.22	0.06	−3.5	***		0.08	0.08	1.0		−0.01	0.06	−0.1	
	CRM	−0.19	0.17	−1.1			−0.07	0.14	−0.6		0.11	0.09	1.3	
melanism	SRM	phenology 0.22	0.04	4.8	***	year	0.26	0.04	6.5	***	0.11	0.04	2.8	*
	NRM	0.18	0.08	2.3	*		−0.38	0.07	−5.1	***	−0.12	0.08	−1.5	
	CRM	−0.07	0.09	−0.8			−0.15	0.09	−1.7		0.15	0.07	2.1	
melanism	SRM	T_p 0.01	0.05	0.3		year	0.20	0.05	3.9	***	0.08	0.04	1.8	
	NRM	0.19	0.09	2.2	*		−0.57	0.08	−7.4	***	−0.17	0.04	−3.9	***
	CRM	0.19	0.10	2.1			−0.38	0.10	−4.0	***	0.06	0.06	0.9	
melanism	SRM	T_{pupal} 0.16	0.05	3.0	**	phenology	0.22	0.05	4.9	***	0.06	0.05	1.4	
	NRM	0.15	0.09	1.7			0.17	0.08	2.0		0.15	0.06	2.4	*
	CRM	−0.06	0.08	−0.8			−0.03	0.09	−0.4		0.20	0.06	3.2	**

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

attributing the morphological shifts to plasticity or evolution. This attribution challenge is exemplified by the previous [21] (using multiple regression models) and current analyses for the Southern RM suggesting somewhat different responses to temperature and phenology.

Portions of the current analyses are contrary to the hypothesis that increases in pupal temperature decrease melanism via plasticity. The limited evidence of plasticity aligns with model predictions that plasticity may be less advantageous (and perhaps less prevalent) for high-elevation *Colias*, such as *C. meadii*, with short flight seasons [31]. Additionally, morphological shifts persist when attempting to control for plasticity. This limited evidence of plasticity in wing melanism suggests a role for genetic evolution in morphological shifts and shows how, more broadly, controlling for plasticity in studies of museum specimens can help disentangle plastic and genetic shifts.

Shifts toward darker wings in the high-elevation Southern RM but lighter wings in the lower elevation Northern and Canadian RM may indicate differential drivers among biogeographic regions, perhaps reflecting habitat differences. Climate conditions may be less extreme in the lower elevation *Colias* habitat in the Northern and Canadian RM, but our limited weather station data prevent a thorough assessment. The shifts are consistent with model predictions of selection for wing lightening at low elevation and wing darkening at high elevations [32]. Darker coloration may enable high-elevation populations to capitalize on warming [28], particularly if warming has extended the flight season into the cooler early and late seasons [21]. In recent years in the Southern RM, seasonal temperatures have reached higher mid-season maxima before falling off sharply (electronic supplementary material, figure S3). However, our ability to detect phenological shifts may be compromised by biases in the time of year of collection. Fluctuations in the strength and direction of selection owing to environmental variation also influence melanism trends in the alpine sites [32]. Given the cool temperatures in the alpine sites, selection for wing darkening to increase flight time in cold years may be stronger and more regular than selection to avoid overheating in warm years. Examining a single, relatively well-sampled location in the Southern RM (Loveland Pass) does not clarify these phenological or phenotypic shifts, likely owing to uneven temporal coverage.

Examining multiple phenotypes provides an integrated view of organismal responses. In the Southern RM (consistent with previous observations [21]), we find that both wing melanism and setae length increase, which would tend to increase body temperatures. The response of two phenotypes contrary to expected simple responses to climate warming suggests the need to further consider explanations such as those discussed above. We do not, however, detect shifts in setae length in the Northern or Canadian RM despite decreases in wing melanism. Heat retention may be less important in the lower elevation, potentially less climatically variable, habitat in these regions. Increases in body size have occurred in the Southern RM over recent decades, again in agreement with the previous study [21], but again we

detected no change in the Northern or Canadian RM. Our observation of unexpected and inconsistent phenotypic shifts, even in the well-studied *Colias* system, demonstrates the importance of integrative consideration of multiple morphological traits in future studies.

Turning to natural history or museum collections to test predictions, and ideally models, for morphological shifts in other systems in which the function of their morphology is well understood should help clarify how plasticity and evolution interact to drive morphological shifts. With the exception of size, morphological traits related to climate change are understudied in natural history collections. The studies that do exist show strong bias in taxa (mostly insects, rodents and birds), habitats (mostly terrestrial), geographical region (mostly Europe and North America) and traits considered (mostly coloration or head and body shape; see electronic supplementary material, Appendix S1). These biases may exist in part owing to the biases of past collections and the constraints inherent in preserving specimens. One promising, but resource intensive, approach is to pair museum collections or historical surveys with contemporary resampling [44]. Efforts to digitalize specimen collections that include images (e.g. LepNet, www.lep-net.org and ButterflyNet, www.butterflynet.org) have the potential to revolutionize detecting and understanding morphological shifts.

Fully understanding the causes of phenotypic shifts and their relationship to climate change requires considerable additional ecological and evolutionary knowledge. Distinguishing the relative contributions of phenotypic plasticity and genetic evolution is facilitated by knowledge of the genetics underlying the trait. The only studies we identified where the observed morphological change could be specifically attributed to evolution were studies of colour polymorphisms whose genetic basis was established separately. Cases such as our focal study of *C. meadii*, where a quantitative trait also varies plastically, require integrative studies of the species' ecological and evolutionary context. Integrative studies including comprehensive (including genetic and biochemical) measurements of specimens are needed to decipher interacting responses to past climate change, which is essential to accurately predict future responses.

References

1. Parmesan C. 2006 Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.* **37**, 637–669. (doi:10.1146/annurev.ecolsys.37.091305.110100)

2. Gardner JL, Peters A, Kearney MR, Joseph L, Heinsohn R. 2011 Declining body size: a third

universal response to warming? *Trends Ecol. Evol.* **26**, 285–291. (doi:10.1016/j.tree.2011.03.005)

3. Sheridan JA, Bickford D. 2011 Shrinking body size as an ecological response to climate change. *Nat. Clim. Change* **1**, 401–406. (doi:10.1038/nclimate1259)

4. Holmes MW et al. 2016 Natural history collections as windows on evolutionary processes. *Mol. Ecol.* **25**, 864–881. (doi:10.1111/mec.13529)

5. Johnson KG et al. 2011 Climate change and biosphere response: unlocking the collections vault. *BioScience* **61**, 147–153. (doi:10.1525/bio.2011.61.2.10)

6. Scheffers BR et al. 2016 The broad footprint of climate change from genes to biomes to people. *Science* **354**, aaf7671. (doi:10.1126/science.aaf7671)

7. Hendry AP. 2015 Key questions on the role of phenotypic plasticity in eco-evolutionary dynamics. *J. Hered.* **10**, 25–41. (doi:10.1093/jhered/esv060)

8. Kingsolver JG, Woods HA, Buckley LB, Potter KA, MacLean HJ, Higgins JK. 2011 Complex life cycles and the responses of insects to climate change. *Integr. Comp. Biol.* **51**, 719–732. (doi:10.1093/icb/icr015)

9. Ryan SF, Deines JM, Scriber JM, Pfrender ME, Jones SE, Emrich SJ, Hellmann JJ. 2018 Climate-mediated hybrid zone movement revealed with genomics, museum collection, and simulation modeling. *Proc. Natl Acad. Sci. USA* **115**, E2284–E2291. (doi:10.1073/pnas.1714950115)

10. Bortolotti GR, Marchant T, Blas J, Cabezas S. 2009 Tracking stress: localisation, deposition and stability of corticosterone in feathers. *J. Exp. Biol.* **212**, 1477–1482. (doi:10.1242/jeb.022152)

11. Hilton GM, Thompson DR, Sagar PM, Cuthbert RJ, Cherel Y, Bury SJ. 2006 A stable isotopic investigation into the causes of decline in a sub-Antarctic predator, the rockhopper penguin *Eudyptes chrysocome. Glob. Change Biol.* **12**, 611–625. (doi:10.1111/j.1365-2486.2006.01130.x)

12. Kleijn D, Raemakers I. 2008 A retrospective analysis of pollen host plant use by stable and declining bumble bee species. *Ecology* **89**, 1811–1823. (doi:10.1890/07-1275.1)

13. Higgins JK, MacLean HJ, Buckley LB, Kingsolver JG. 2014 Geographic differences and microevolutionary changes in thermal sensitivity of butterfly larvae in response to climate. *Funct. Ecol.* **28**, 982–989. (doi:10.1111/1365-2435.12218)

14. Geerts AN et al. 2015 Rapid evolution of thermal tolerance in the water flea *Daphnia. Nat. Clim. Change* **5**, 665–668. (doi:10.1038/nclimate2628)

15. Stevenson RD. 1985 Body size and limits to the daily range of body-temperature in terrestrial ectotherms. *Am. Nat.* **125**, 102–117. (doi:10.1086/284330)

16. Dillon ME, Wang G, Huey RB. 2010 Global metabolic impacts of recent climate warming. *Nature* **467**, 704–706. (doi:10.1038/nature09407)

17. Caruso NM, Sears MW, Adams DC, Lips KR. 2014 Widespread rapid reductions in body size of adult salamanders in response to climate change. *Glob.*

Change Biol. **20**, 1751–1759. (doi:10.1111/gcb.12550)

18. Brown JH, Gillooly JF, Allen AP, Savage VM, West GB. 2004 Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789. (doi:10.1890/03-9000)

19. Roulin A. 2014 Melanin-based colour polymorphism responding to climate change. *Glob. Change Biol.* **20**, 3344–3350. (doi:10.1111/gcb.12594)

20. Brakefield PM, De Jong PW. 2011 A steep cline in ladybird melanism has decayed over 25 years: a genetic response to climate change? *Heredity* **107**, 574–578. (doi:10.1038/hdy.2011.49)

21. MacLean HJ, Kingsolver JG, Buckley LB. 2016 Historical changes in thermoregulatory traits of alpine butterflies reveal complex ecological and evolutionary responses to recent climate change. *Clim. Change Responses* **3**, 13. (doi:10.1186/s40665-016-0028-x)

22. Ożgo M, Schilthuizen M. 2012 Evolutionary change in *Cepaea nemoralis* shell colour over 43 years. *Glob. Change Biol.* **18**, 74–81. (doi:10.1111/j.1365-2486.2011.02514.x)

23. Ożgo M, Liew T-S, Webster NB, Schilthuizen M. 2017 Inferring microevolution from museum collections and resampling: lessons learned from *Cepaea. PeerJ* **5**, e3938. (doi:10.7717/peerj.3938)

24. Karell P, Ahola K, Karstinen T, Valkama J, Brommer JE. 2011 Climate change drives microevolution in a wild bird. *Nat. Commun.* **2**, 208. (doi:10.1038/ncomms1213)

25. Galeotti P, Rubolini D, Sacchi R, Fasola M. 2009 Global changes and animal phenotypic responses: melanin-based plumage redness of scops owls increased with temperature and rainfall during the last century. *Biol. Lett.* **5**, 532–534. (doi:10.1098/rsbl.2009.0207)

26. Kingsolver JG. 1983 Thermoregulation and flight in colias butterflies - elevational patterns and mechanistic limitations. *Ecology* **64**, 534–545. (doi:10.2307/1939973)

27. Watt WB. 1968 Adaptive significance of pigment polymorphisms in *Colias* butterflies. I. Variation of melanin pigment in relation to thermoregulation. *Evolution* **22**, 437–458. (doi:10.1111/j.1558-5646.1968.tb03985.x)

28. Buckley LB, Kingsolver JG. 2012 The demographic impacts of shifts in climate means and extremes on alpine butterflies. *Funct. Ecol.* **26**, 969–977. (doi:10.1111/j.1365-2435.2012.01969.x)

29. Ellers J, Boggs CL. 2004 Functional ecological implications of intraspecific differences in wing melanization in *Colias* butterflies. *Biol. J. Linn. Soc.* **82**, 79–87. (doi:10.1111/j.1095-8312.2004.00319.x)

30. Hoffman RJ. 1978 Environmental uncertainty and evolution of physiological adaptation in *Colias* butterflies. *Am. Nat.* **112**, 999–1015. (doi:10.1086/283343)

31. Kingsolver JG, Buckley LB. 2017 Evolution of plasticity and adaptive responses to climate change

along climate gradients. *Proc. R. Soc. B* **284**, 20170386. (doi:10.1098/rspb.2017.0386)

32. Kingsolver JG, Buckley LB. 2015 Climate variability slows evolutionary responses of *Colias* butterflies to recent climate change. *Proc. R. Soc. B* **282**, 20142470. (doi:10.1098/rspb.2014.2470)

33. Kingsolver JG, Buckley LB. 2017 Quantifying thermal extremes and biological variation to predict evolutionary responses to changing climate. *Phil. Trans. R. Soc. B* **372**, 20160147. (doi:10.1098/rstb.2016.0147)

34. Kingsolver JG, Buckley LB. 2018 How do phenology, plasticity and evolution determine the fitness consequences of climate change for montane butterflies? *Evol. Appl.* **11**, 1231–1244. doi:10.1111/eva.12618)

35. DeChaine EG, Martin AP. 2005 Historical biogeography of two alpine butterflies in the Rocky Mountains: broad-scale concordance and local-scale discordance. *J. Biogeogr.* **32**, 1943–1956. (doi:10.1111/j.1365-2699.2005.01356.x)

36. Wheat CW, Watt WB. 2008 A mitochondrial-DNA-based phylogeny for some evolutionary-genetic model species of *Colias* butterflies (Lepidoptera, Pieridae). *Mol. Phylogenet. Evol.* **47**, 893–902. (doi:10.1016/j.ympev.2008.03.013)

37. Ellers J, Boggs CL. 2002 The evolution of wing color in *Colias* butterflies: heritability, sex linkage, and population divergence. *Evolution* **56**, 836–840. (doi:10.1111/j.0014-3820.2002.tb01394.x)

38. Stamberger JA. 2006 Adaptation to temporal scales of heterogeneity in the thermal environment. Dissertation. Stanford University.

39. Ellers J, Boggs CL. 2004 Evolutionary genetics of dorsal wing colour in *Colias* butterflies. *J. Evol. Biol.* **17**, 752–758. (doi:10.1111/j.1420-9101.2004.00736.x)

40. McGuire CR, Nufio CR, Bowers MD, Guralnick RP. 2012 Elevation-dependent temperature trends in the rocky mountain front range: changes over a 56- and 20-year record. *PLoS ONE* **7**, e44370. (doi:10.1371/journal.pone.0044370)

41. Higgins JK. 2014 *Rapid evolution and population divergence in response to environmental change in Colias butterflies*. University of North Carolina at Chapel Hill. (http://gradworks.umi.com/36/68/3668482.html).

42. Bivand R. 2005 *The Spdep package*. Comprehensive R Archive Network, Version 0.3–13.

43. Kissling WD, Carl G. 2008 Spatial autocorrelation and the selection of simultaneous autoregressive models. *Glob. Ecol. Biogeogr.* **17**, 59–71. (doi:10.1111/j.1466-8238.2007.00379.x)

44. Tingley MW, Beissinger SR. 2009 Detecting range shifts from historical species occurrences: new perspectives on old data. *Trends Ecol. Evol.* **24**, 625–633. (doi:10.1016/j.tree.2009.05.009)

45. MacLean HJ, Nielsen ME, Kingsolver JG, Buckley LB. 2018 Data from: Using museum specimens to track morphological shifts through climate change. Dryad Digital Repository. (doi:5061/dryad.f448b91)