**Title:** Trophic phenological mismatch: Disconnects between underlying ecological theory and climate change responses

**Authors:** Heather M. Kharouba1, Elizabeth M. Wolkovich2,3

**Affiliations:**

1Department of Biology, University of Ottawa, Canada K1N 9B4

2Harvard

3UBC

Email: heather.kharouba@uottawa.ca,

Running Title:

**Keywords:**

**Type:** Ecology Letters (Ideas & Perspectives)

**Abstract**

Many researchers hypothesize that climate change will lead to phenological mismatches with negative consequences for those interacting species and their ecological communities; yet, evidence documenting negative impacts on fitness is mixed. The most common ecological theory that underlies these studies is the Cushing match-mismatch hypothesis. It offers a testable, generally applicable hypothesis for predicting these consequences due to climate change. Here, we argue most studies do not provide strong tests of this hypothesis or define pre-climate change baselines in their study system, thus making it difficult to assess support for this major hypothesis. To accurately predict the magnitude and prevalence of mismatches due to climate change, relating empirical observations to underlying mechanisms through hypothesis testing will be required. Adjusting study designs, however, can allow more rigorous tests of this hypothesis. We highlight how these approaches could rapidly advance mechanistic understanding and thus allow robust predictions of shifts with continuing climate change.

**Introduction**

Climate change is causing phenological shifts (i.e. changes in the timing of life history events) that vary across species in different functional groups and trophic levels (Thackeray et al. 2016; Ovaskainen et al. 2013; CaraDonna et al. 2014). Such species-specific variation in response to climate change has led to changes in the relative timing of key activities (phenological synchrony) among interacting species (Kharouba et al. 2018). These changes have caused fitness consequences—often termed ‘phenological mismatch’ (Box 1)—and have influenced ecosystem-level properties in some contexts (Post and Forchhammer 2008; Plard et al. 2014; Doiron et al. 2015; Burkle et al. 2013) but not others (Vatka et al. 2011; Burthe et al. 2012). Despite many theoretical (Bewick et al. 2016; Johansson et al. 2015) and empirical studies (REF) based in single systems, we still have no general ability to predict the outcomes of shifts in phenological synchrony due to climate change.

Here, we argue that much of the difficulty in predicting the consequences of climate change-driven shifts in synchrony is due to the disconnect between ecological theory and current empirical approaches in the phenological mismatch literature. Current methodological inconsistencies across studies make it difficult to test the relevant underlying ecological theory in the context of climate change. Without better evidence, we cannot attribute variation in findings across studies to species, site, or mechanism. Without an understanding of the mechanisms underlying the well-documented patterns in phenological shifts, our ability to make accurate predictions about species’ responses, and species’ interactions, to climate change remains limited (O’Connor et al. 2012; Chmura et al. 2018).

Here, we focus on the widely-cited Cushing match-mismatch, or trophic mismatch, hypothesis (1974), the most commonly applied hypothesis to consumer-resource interactions in this literature. We show how advances could come from direct tests of the hypothesis and clear definitions of baselines, when possible. Our aim is not to put forward additional hypotheses about the context in which phenological mismatch will occur, which has been reviewed extensively elsewhere (e.g., Miller-Rushing 2010; Renner and Zohner 2018), but rather to help guide the study of phenological mismatch to develop more robust predictions.

Although the Cushing hypothesis has been applied to other types of interactions (e.g. mutualism), we limit our discussion to consumer-resource interactions (i.e. antagonistic). Below, we provide an overview of the Cushing hypothesis, review the current approach in the phenological mismatch literature and then outline the divide between the hypothesis and the empirical studies. We discuss how current approaches are impeding major progress in the field but that changes in our approach could rapidly advance our understanding and help forecast of the impacts of climate change on ecological communities, the ultimate goal of most of the phenological mismatch literature.

***Overview of the main ecological theory***

The most common ecological theory that underlies phenological mismatch studies (Appendix) is the Cushing match-mismatch hypothesis. This hypothesis predicts a concave down curve between consumer fitness and relative timing between the consumer and its resource (1974; Figure 1). While this curve has been applied across many ecosystems (CITES), the theory originally emerged from the marine fisheries literature as a way to explain the variation in population recruitment of fish stocks. Based on life-history theory, it postulates that there should be selective pressure for the consumer to temporally ‘match’ the peak of its energetic phase with the peak of resource availability (Figure 1). Given this strong selective pressure, if there is any change to the relative timing of the interaction, there will be a decrease in fitness for the consumer (i.e., a mismatch), thus producing the now-often-shown curve (Figure 1). For example, if a consumer emerges too early in the spring and the resource has not yet emerged, the consumer will risk starvation. At the curve’s limits, consumer fitness should fall to zero when the change in relative timing is sufficiently large (Figure 1). For example, the early-emerging consumer can only survive so many days without food.

This hypothesis is based on two important assumptions. First, the resource is the major controller on the consumer (i.e., strong interaction strength, bottom-up control; Cury et al, 2003). Second, the consumer and resource must present a certain degree of seasonality (e.g., vegetation only being available for part of the year; Durant et al. 2005; Durant et al. 2007), limiting the optimal period for growth and reproduction. If neither of these assumptions is met, then fitness consequences due to changes in the relative timing of the interaction will be either weak or non-existent.

***Disconnect between theory and empirical studies***

In its original state, the hypothesis has been debated, contested and criticized, particularly in the marine literature (Durant et al. 2007, Leggett and DeBlois 1994\*). In part because, although a relatively simple hypothesis, it is inherently difficult to test in the field, an assertion even Cushing himself made. When explicitly tested, there has been unequivocal support for it (ref in Durant et al. 2007: Gotceitas et al. 1996; Nooker et al. 2005). However, when not explicitly tested, the shape and strength of the relationship between consumer fitness and the relative timing of the interaction varies greatly (e.g., AO001, HMK003, HMK025, HMK054).

While others have suggested that this is because of data limitations and the model’s implication of complex multitrophic dynamics (Kerby chapter, Durant et al. 2007), we argue that there are key methodological that make it difficult to determine whether this hypothesis is widely supported in the context of climate change. Below, we introduce the current objectives of the phenological mismatch literature, and then discuss how studies are not rigorously testing the Cushing hypothesis or defining pre-climate change baselines.

To put our argument in context, we reviewed the phenological mismatch literature. We examined 41 observational studies that evaluated the effects of a change in relative timing on the fitness or a performance (hereafter referred to as performance) of the consumer (see Appendix for details). These studies largely focused on: i) documenting how climate change is affecting the timing of a trophic interaction and how those changes have affected the consumer’s fitness; and ii) advancing the theory necessary to predict long-term demographic changes due to changes in synchrony. The ultimate goal of these studies is to predict the impacts of climate change on ecological communities (hereafter called ‘climate change’ studies). Our database also included studies whose aim was to understand the underlying processes related to timing (e.g., coevolution, life history trade offs, food web dynamics) that drive consumer or resource dynamics (hereafter called ‘fundamental studies’), independently of climate change. Combined, these studies have improved our understanding of the importance of the relative timing of an interaction for consumer fitness.

*i) Testing fundamental theory*

The Cushing hypothesis offers a testable, generally applicable hypothesis for predicting the magnitude and direction of demographic changes in response to climate-change driven shifts in synchrony (Figure 2). To date, much research in the biological impacts of climate change literature has focused on the direct relationships between organisms and the environment (Cleland et al. 2007 (TREE); Korner?; Pau et al. 2011?) rather than testing theory (O’Connor et al. 2012). However, progress on the Cushing hypothesis requires tests of a diversity of ecological and evolutionary theory. This represents the major challenge of the hypothesis and—we argue—may be why support for it has been so mixed.

Many mechanisms can produce the Cushing curve, and they most likely vary across systems, space and time. For example, mechanisms arising from life-history theory (Box 1) can produce the Cushing curve—as consumers maximize their fitness through ideal timing with their primary resource—this forms the theoretical basis for the original hypothesis. The Cushing hypothesis, however, is also connected to food web theory (Box 1), since it is about the timing of a consumer to its food resource. However, studies deviate in what forces they hypothesize control the peak in the food resource (a critical component of the Cushing curve, see our Figure 1). For example, in aquatic systems—where top-down forces are generally more common compared to terrestrial systems—many studies suggest that the resource peak is controlled by release from, or predation by, a consumer (CITES?). This is a very different hypothesis from others that suggest seasonality in the environment produces the resource peak (CITES). In terrestrial systems, the relative contributions of bottom-up vs. top-down effects on consumer dynamics remains poorly understood (e.g. Boggs and Inouye).

The diversity of ecological theory that can produce the basic shape of the Cushing curve means that different researchers may collect very different data, depending on what mechanism they suspect underlies the curve in their particular system or what their main objective is. Researchers studying the hypothesis from the lens of life history theory require per capita estimates of fitness, including measurements of fecundity and mortality and how they vary across ontogeny; such measurements allow researchers to assess fitness consequences at the level of selection (i.e., the individual) and to directly link any fitness changes to timing. Studies addressing the Cushing hypothesis from the perspective of food web theory require more equivalent data on the consumer and resource but do not require per capita fitness consequences for the consumer.

Strong tests of the Cushing hypothesis include tests of multiple mechanisms. For this performance metrics would ideally be assessed for both the consumer and resource, and at the level of the individual when possible, but such data are extremely rare in this literature. Of the studies we examined with per-capita performance data on the consumer (hereafter referred to as ‘life history studies’), none had equivalent measures of performance on the resource (Table 1). The majority of these studies (14/25) measured performance at the community level for the resource (e.g. biomass; Table 1) and many had no measure at all for the resource (7/25; Table 1). As expected, studies that measured population- or community-level data on the consumer (hereafter referred to as *food-web* studies) collected more equivalent data on the resource and consumer. The majority of these studies also collected population-level (7/17) or community level data for the resource (7/17) (Table 1). Across all life history and food-web studies, only one study measured per-capita performance on the resource.

Part of these discrepancies may stem from differing generation times—studies that can easily observe a resource curve shaped by predation often focus on organisms with generation times on the scale of days to week. For example, aquatic studies that focus on phytoplankton as the resource may observe many generations in one summer while a terrestrial study focused on caterpillars would generally observe a comparably smaller number generations (1-2(?), depending on latitude and species for caterpillars). When populations turnover too quickly to track individuals, researchers struggle to collect robust data on per capita fitness. The fundamental difference in the scale of generation times pervades aquatic/terrestrial comparisons (Gruner et al. 2008, Borer et al. 2005) and may drive the difference in the objectives of aquatic vs. terrestrial studies in the phenological mismatch literature. The majority of aquatic studies in our review approached the Cushing hypothesis from the perspective of food-web theory (14/20), whereas terrestrial studies approached it from life-history theory (20/24; Table 2).

A further divide across studies may come from the size, and population size of the consumer, properties correlated with generation time: it can be difficult to collect per capita fitness data when consumers are too small and numerous. Of the life history studies, the majority of consumers were birds (15/25) and mammals (6/25; Table 2). Comparatively, invertebrates were measured at the population or community level (Table 2). *Not sure how to end this paragraph!!*

This places the phenological mismatch studies on a continuum: at one end, aquatic studies focus generally on food web mechanisms, collect equivalent data on the consumer and resource but no per-capita fitness data and at the other end, terrestrial studies focus on life history theory from the lens of the consumer, collect individual-level data on the timing and fitness of the consumer, with much less information on the resource. The fundamental problem with these approaches is that researchers test only one piece of the much larger field of mechanisms that could underlie the Cushing curve. Though it is highly possible that both food web and life history theory together explain many of the consumer-resource systems studied, data limitations make it hard to assess both hypotheses at once. Consequently, most studies do not actually provide strong tests of the Cushing hypothesis, making it difficult to refute the hypothesis if no evidence is found. If they find their mechanisms explain little of the variation they observe, they rarely have the data to test alternative hypotheses. Without strong support for their hypothesis (i.e., high explained variation and clear underlying links), the mechanism underlying the hypothesis [again, may need to change hypothesis to curve] will be uncertain.

This uncertainty can be caused by two reasons: (1) hypotheses that are not mutually exclusive (i.e. different mechanisms have the same predictions); or (2) you get different predictions when multiple mechanisms are considered. For example, researchers have pointed out how a larger perspective on life history, such as including other sources of mortality beyond those related to resource access or alternative resource types (Samplonius et al. 2016), can lead to alternative predictions than the synchrony predicted by the Cushing curve (CITES, Singer & Parmesan, Johannsson & Jonzen). More generally, this is an issue that permeates ecology (REF).

A consequence of not collecting per-capita fitness data on the resource is that our mechanistic understanding ends at the consumer level. This will make it difficult to make predictions of the broader consequences of climate change-driven phenological mismatches for the interaction. We will not be able to incorporate feedbacks between consumer and resource or determine the broader implications of phenological mismatches for the community.

*ii) Pre-climate change baseline*

The concept of pre-climate change baseline (i.e., defining the range of natural variation) is rarely, if ever, considered by phenological mismatch studies; yet, establishing one in a system has important implications for the Cushing hypothesis in both climate change and fundamental studies. Of the studies we considered, only 23% (9/40) of the studies began before the early 1980s, the same time as pronounced recent climate change began (REF). Even in this subset of studies, only X % had data from more than 2 years. This is also an issue with the phenological synchrony literature (Box 1), which does not necessarily consider the ecological consequences of shifts in synchrony (Kharouba et al. 2018). Even those studies with time-series that extend into decades before the 1980s do not explicitly define a pre-climate change baseline in synchrony (but see AO001, HMK031, HMK002).

There are two implications of not establishing a pre-climate change baseline (Figure 2). First, these studies are implicitly assuming conditions of stationarity (i.e., a constant underlying probability distribution) before climate change occurred (Wolkovich et al. 2014). This is problematic when climate change has led to non-stationarity to different extents in different systems (Wolkovich et al. 2014). Without a defined pre-climate change baseline in the system, we do not know what the pre-climate change baseline *should* be and thus whether and to what magnitude we should predict a divergence in synchrony due to climate change (Figure 2a). It is then further difficult to determine how much of the observed change in the interaction can be attributed to climate change, some other driver (for example, nutrient enrichment in lakes which also started well before 1980 (REF)), versus natural variation. In essence, knowledge of the pre-climate change baseline may rule out climate change and identify other drivers, which is key for the prioritization of conservation management decisions for threatened species.

Second, these studies confound a pre-climate change baseline with the peak of Cushing curve: that is, they assume that fitness was highest before climate change because the consumer’s timing was well matched to the resource’s (Figure 2b; Singer and Parmesan 2010). The resulting prediction is that climate change will necessarily change the relative timing of the interaction (i.e. non-stationarity) leading to a decline in the consumer’s fitness. However, without knowledge of the pre-climate change baseline, it is difficult to predict the direction and magnitude of fitness consequences due to climate change-driven shifts in synchrony (Figure 2c). An alternative hypothesis put forward by Singer and Parmesan (2010) suggests that in some contexts, ‘asynchrony’ (i.e. when the most energetically demanding phase of the consumer is not lined up with the peak resource availability) may be the pre-climate change baseline. They suggest that this can arise due to a pre-existing life history strategy (e.g., maximizing fecundity over mortality). In systems where asynchrony might be the baseline state, our ability to anticipate the effects of climate change will be even more challenging (Figure 2bc) as these interactions might become less (i.e. fitness increases) or even more mismatched (i.e. fitness decreases) with climate change. Theoretically, interactions might become less mismatched if the relative timing of the interaction is at either limit of the curve (e.g. bet-hedging in variable environments (Danforth 1999)) and if the phenology of the consumer and resource respond at different rates to climate change (Figure 2c).

Establishing a pre-climate change baseline also has implications for fundamental studies that want to understand the underlying processes of consumer-resource dynamics. Recognizing that these dynamics are occurring in the context of climate change, our current reality is key. The pre-climate change baseline determines whether the system should currently be in equilibrium or is only transient. For example, life-history trade-offs often end up in some form of equilibrium but if climate change has pushed the system away from the baseline, then the system might currently be under transient dynamics rather than at equilibrium (Figure 2). This is a similar condition for co-evolution, it predicts an arms-race but that arms-race varies a lot under stationary climate versus non stationary climate (ref). To this end, an understanding of the system’s pre-climate change baseline state is important for providing context and designing studies.

**Towards robust forecasting of phenological mismatch**

While the Cushing hypothesis provides a testable hypothesis for predicting phenological mismatch in response to climate change, it is uncertain how much support there is for this hypothesis. Here we have suggested two major reasons for this, both of which provide a clear path forward for progress in the field. First, most studies do not actually provide strong tests of the Cushing hypothesis because (a) studies do not collect equivalent data on the consumer and resource, and (b) multiple mechanisms are not being tested. This makes it difficult to refute the hypothesis if no evidence is found. Unless researchers are extremely clear about the mechanistic hypotheses they are testing, progress will continue to be slow. Second, pre-climate change baselines are not being defined in the majority of studies in this literature. Lack of a relevant baseline makes it is difficult to determine whether recent phenological mismatch can be attributed to climate change and it limits our ability to predict the direction and magnitude of phenological mismatch due to climate change.

To move the field of phenological mismatch forward, from documentation towards forecasting, higher quality data are needed to test fundamental hypotheses and key baselines need to be defined. Below, we highlight how new approaches researchers take can improve testing of the Cushing hypothesis. This will also provide the framework for building the required depth in evidence across studies to determine general quantitative patterns and their underlying mechanisms, an approach not currently possible given the current state of the field.

*(i) Testing fundamental theory*

By having the potential to test multiple mechanisms and define key baselines, the integration of long-term data (i.e. that start before the early 1980s) and an experimental approach in a single system represents the gold standard to move the field of phenological mismatch forward. In the absence of long-term data, the integration of observational data across a spatial gradient with an experimental approach can provide a powerful approach to test multiple hypotheses. Samplonius et al. (2016) combined descriptive and experimental data to test many confounding factors of the Cushing hypothesis in their system (e.g., ontogenetic dietary shifts, habitat quality). In some systems, experiments can allow researchers to test multiple mechanisms, for example direct from indirect effects or abiotic from phenological mismatches(Box 1; REF). For example, Berger et al. (2014) manipulated water temperature, stratification depth and the presence/absence of zooplankton in lake mesocosms to separate the direct effects of light and temperature on spring plankton dynamics from effects mediated through the consumer.

Baselines themselves can also be used to test multiple mechanisms. Researchers with long-term data can define specific hypotheses about the expected vs. observed patterns of environmental change in their systems and then attribute variation in phenological mismatch to the main driver. For example, many lakes experienced large changes in nutrient inputs in the 1960s and 1970s (e.g., George 2011; Lake Washington), before the onset of recent climate change in the early 1980s so researchers can use different breakpoints to determine where non-linearity began.

Higher temporal resolution data in some systems may help with teasing apart different mechanisms and collecting more equivalent performance data on the consumer and resource. Sampling frequency should be relative to the life history of the species of interest. For example, in aquatic systems, there is often quick turnaround between producers and consumers so it can be difficult to determine how much influence producers have on consumers and vice versa. Moreover, some zooplankton are never in complete dormancy (i.e., they remain at low densities). The temporal sequencing of consumer and resource and seasonal availability of the resource are key requirements of the Cushing hypothesis. Finally, researchers can be explicit when possible about which mechanism(s) is likely driving the curve.

*ii) Defining baselines*

Researchers with long-term data have the potential to test for support for the competing hypotheses of synchrony vs. asynchrony as the pre-climate change baseline (Figure 2). By testing for a clear peak in fitness, researchers can place themselves on the Cushing curve, where a clear peak would support the synchrony baseline hypothesis. Incorporating the most energetically expensive part of the consumer’s phenology or estimating how much the phenological phase that is measured differs from the most expensive part (e.g. egg hatching is on average 7 days before maximum chick growth occurs) will help define a peak in fitness. If there is no clear peak in fitness, then well-designed experiments (in systems where possible) that have the ability to reveal the limits of the curve (i.e. where consumer fitness reaches zero) can be used to reject the null hypothesis of an asynchrony baseline. Knowledge of when/if different resources are available and whether the interaction type changes throughout the season can also help describe the limits of the curve.

Where long-term data are not available, there is no perfect solution to this challenge. Null modeling to simulate historical data may be an alternative (e.g., Dornelas et al. 2014; Sgardeli et al. 2016; Kharouba et al. 2018). When not all required data is available across a long time period, the integration of experimental data with observational studies could be another approach (Figure 3). For example, the performance data for an insect herbivore and its main host plant is not available beyond X number of years from a single location but researchers were able to manipulate the timing of larval emergence to determine the potential consequences of a change in the timing of the interaction on the performance of the insect herbivore (Figure 3). Combining the modelling of pre-climate change baselines based on knowledge of phenological cues using a hindcasting approach (Senner et al. 2016; Deacy et al. 2017) or process-based phenological models (Chuine and Régnière 2017) with knowledge about the Cushing curve in the system is another possibility.

*Final thoughts on forecasting*

Accurately forecasting phenological mismatch in response to climate change will involve translating short-term responses into long-term demographic responses. Beyond the points already discussed, this will require knowledge of i) the correct phenological cues for both species in the interaction (Pau et al. 2011), which are not widely known (Kharouba et al. 2018; Chmura et al. 2018); ii) the strength, and associated uncertainty, of the shape of the Cushing curve; iii) lifetime fitness for the consumer; and iv) comprehensive data on the resource. Shifts in phenological synchrony will lead to temporal changes in a consumer’s performance if the species’ performance is related to the relative timing of the interaction (i.e., the Cushing hypothesis) and there has been a directional change in synchrony over time. For site-level predictions, a combination of process-based phenological modeling (Chuine and Régnière 2017) and phenologically explicit consumer-resource models (Bewick et al. 2016) can be used. When data is available for multiple populations, hierarchical Bayesian modeling can be used, which can overcome spatial variability in phenological responses to temperature (Ibanez et al. 2010). Previous work has identified factors that will help predict *when* a mismatch is more likely to lead to a population decline but this work is largely theoretical. Given the number of potential sources of variation in phenological mismatch (site, species, mechanism etc.), comparable data across systems will be needed to make general predictions about the ecological consequences of shifts in phenological synchrony.

**Acknowledgements**

We thank Johan Ehrlen, Kjell Bolmgren and Steve Travers for interesting discussions. HMK thanks the professor writing retreats offered through the Centre for Academic Leadership at the University of Ottawa for support in writing this manuscript and *Harvard that supported travel to work with EMW*.

**References**

Gruner et al. 2008, A cross-system synthesis of consumer and nutrient resource control on producer biomass;

Borer et al. 2005, What determines the strength of a trophic cascade?

**Box 1. Glossary**

**Phenological mismatch**- between interacting species, not with abiotic environment or intraspecifc, incorporates fitness

**Phenological synchrony-** does not incorporate fitness

**Asynchrony** (singer and parmesan)

**Baseline-** a reference or benchmark to represent the conditions that fully describe functional ecosystems, Historical range of variability= broad historical envelope of possible ecosystem conditions, such as burned area, vegetation cover type area, or patch size distribution, provides a representative time series of reference conditions; Keane et al. 2009 Forest Ecology and Management; Landres et al. 1999 Ecological Applications

**Stationarity**

**Mechanism** (ecological pathways to produce the curve)

**Life-history theory**

**Food web theory**

**Tables**

Table 1. A comparison across studies of the type of performance data collected for consumer and resource.  *We define a life-history study as one that collected data at the individual level and a food-web study as one that collected data at the population or community (i.e., across species) level data.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | Resource fitness | | | |  |
|  | Life-history | Food-web | |  |
| None | Individual | Population | Community | *Totals* |
| Consumer  Fitness | Life-history | Individual | 7 | 0 | 4 | 14 | *25* |
| Food-web | Population | 2 | 1 | 7 | 7 | *17* |
| Community | 0 | 0 | 1 | 1 | *2* |
|  |  | *Totals* |  |  |  |  | *44* |

Table 2. A comparison across studies of the type of performance data collected for the consumer across systems and taxonomic group.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | System | | Taxonomic group | | | |
| Invertebrate | Vertebrate | | |
| Aquatic | Terrestrial | Fish | Bird | Mammal |
| Consumer  Fitness | Individual | 6 | 20 | 1 | 2 | 15 | 6 |
| Population | 13 | 3 | 9 | 4 | 2 | 1 |
| Community | 1 | 1 | 2 | 0 | 0 | 0 |
| *Totals* | *20* | *24* | *12* | *6* | *17* | *7* |

**Figures**

Figure 1. Simple conceptualization of the Cushing curve; with climate change predictions

Figure 2. Stationarity and change with climate change (a); then assumed max fitness, pre-climate change baseline (b); alternative baselines (c) … note this means (b) does not have the shallow curve fro Singer & Parmesan, but c would, yielding two examples of the major alternatives: (1) you’re on a different spot on the curve that max fitness before climate change and (2) the curve is different.

Figure 3. Case study demonstrating the integration of experimental (a) and observational data (b) relating to the Cushing hypothesis in a single system- the winter moth (*Operophtera brumata*) and oak (*Quercus robur*). (a) Experimental raw data was obtained from Tikkanen and Julkunen-Tiitto (2003) and result from two experiments (green, red points). In the first experiment, the authors manipulated the number of days that neonates (i.e. early instar larvae) spent without food (green points). In the second experiment, they manipulated the emergence times of larvae. There were four cohorts, each separated by intervals of 3-5 days. All *O. brumata* eggs originated from laboratory stock originally from Turku, Finland whereas the foliage originated from trees near Banchory, NW Scotland. (b) Inter-annual variation in relative timing between median egg hatch date of *O. brumata* and the median bud opening date of *Q. robur* from 1996-2005 in the Netherlands. Horizontal error bars represent the lower and upper quartiles of the data. Raw data from the observational study was retrieved from VanAsch and Visser 2007 Figure 2. In this system, negative values along the x-axis denote where egg hatching occurred before bud opening, whereas positive values indicate egg hatching occurred after bud opening.

Notes for HK on figure 3:

(a) great tit and winter moth

* Nilsson and Kallander 2006- compared breeding phenology of great tit between coastal and inland sites where budburst differs by a week
* Visser et al. Oecologia 2006- measured caterpillar and lay dates for 20 years (no experiment)
* Van Noordwijk et al. 1995 – 1948-1972, 1975, 1980, 1982-1986 (UK)
* Visser et al. 1998- observational study 1973-1995
* Bauer et al. 2016- observational study 1961-2007 (Czeck)

(b) winter moth and oak

* Visser and Holleman (2001)- descriptive model
* Buse and Good 1996- temperature manipulation
* Bauer et al. 2016- observational study 1961-2007 (Czeck)
* Tikkanen et al. 2003- (1) effect of starvation: neonate were incubated without food for different periods of time; (2) effect of declining foliage quality: manipulated hatch times (Scotland) cohorts differed by 3-5 days; (3) time after budburst
* Tikkanen and Lyytikainen-Saarenmaa 2002- compared hatching dates between populations, and looked at effect of foliage (Finland and Sweden)
* Tikkanen and Julkunen-Tutto (2003)- neonates reared without food for different periods of time (HMK038)
* Van Dongen 1997
* VanAsch and Visser- data in Figure 2- synchrony- 1996-2005 from netherlands
* Hunter 1990- neonate larva do not tolerate starvation for long periods (2-5 days)
* Check: Crawley MJ, Akhteruzzaman M (1988) Individual variation in the phenology of oak trees and its consequences for herbivorous insects. Funct Ecol 2:409–415
* Check: Salis et a. 2017

Figure 1.



Figure 2.



Figure 3.



**Appendix**

*Literature search*

We located papers with data relating phenological data from trophic interactions to performance of the consumer and/or the resource by conducting keyword searches in ISI Web of Science published up to June 2017. Keywords included phenolog\* AND AND mismatch\* OR synchron\* AND interact\* AND (fitness\* OR performance\*). We focused on observational studies. Studies were excluded if they did not: (1) measure phenology directly (e.g. used derived measures of phenology, such as NDVI or spring temperature); (2) measure phenology of one of the species; (3) quantitatively link consumer performance to the relative timing between consumer and resource. Additionally, authors had to be explicit that the two species interacted (e.g. specifying type of interaction). To ensure a reasonable sample size and to include studies across different major biomes, we included interactions that were resolved to the family-level and below.

If more than one measure of phenology was included, we chose the one used by the authors to calculate mismatch and examine its impact on performance.

Years of study- those years with consumer and resource AND phenology and performance data (e.g. HMK051)

These studies have encompassed terrestrial, marine and freshwater ecosystems as well as large latitudinal gradient.

Based on the type of data collected for the consumer and resource, we classified these studies as life history or food web-based (Box 1).