Interpreting empirical estimates of experimentally derived physiological and biological thermal limits in ectotherms

<table>
<thead>
<tr>
<th>Journal:</th>
<th>Canadian Journal of Zoology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manuscript ID</td>
<td>cjz-2018-0276.R3</td>
</tr>
<tr>
<td>Manuscript Type:</td>
<td>Review</td>
</tr>
<tr>
<td>Date Submitted by the Author:</td>
<td>10-Nov-2019</td>
</tr>
<tr>
<td>Complete List of Authors:</td>
<td>Bates, Amanda; Memorial University of Newfoundland, Ocean Sciences Morley, Simon; British Antarctic Survey</td>
</tr>
<tr>
<td>Is your manuscript invited for consideration in a Special Issue?:</td>
<td>Canadian Society of Zoologists 2018</td>
</tr>
<tr>
<td>Keyword:</td>
<td>physiological experiments, ecological metrics, critical limits, lethal limits, performance</td>
</tr>
</tbody>
</table>
Interpreting empirical estimates of experimentally derived physiological and biological thermal limits in ectotherms

Bates, A.E., Morley, S.A.

1. This review is one of a series of invited papers arising from the symposium “Aquatic environments and organisms in transition” that was co-sponsored by the Canadian Society of Zoologists and the Canadian Journal of Zoology and held during the Annual Meeting of the Canadian Society of Zoologists at Memorial University of Newfoundland, St. John’s, Newfoundland, 7–11 May 2018.

2. Department of Ocean Sciences, Memorial University of Newfoundland, St. John’s A1C 5S7, Canada

Abstract

Whole organism function is underpinned by physiological and biological processes, which respond to temperature over a range of time scales. Given that environmental temperature controls biological rates within ectotherms, different experimental protocols are needed to assess the ability of organisms to withstand extreme weather events versus gradual temperature change. Here we emphasize the importance of time in shaping ecological and evolutionary processes, and as an experimental parameter that is key when interpreting physiology studies reporting thermal limits. We discuss how acute and chronic thermal performance is underpinned by mechanisms operating at different time scales - resistance, acclimation and adaptation. We offer definitions of common physiological and biological temperature metrics and identify challenges inherent to compiling the wealth of historical temperature limit data now available into meta-analytic frameworks. We use a case study, data across temperate fishes, to highlight that false positives may occur when differences in the thermal tolerances of species are in fact due to experimental protocols. We further illustrate that false negatives can arise if researchers fail to recognize differences in thermal limits of species emerging from macrophysiological approaches that are due to biological mechanisms. We strongly advocate for the careful design, interpretation and reporting of experimental results to ensure that conclusions arising from data synthesis efforts are grounded in theory.

Keywords

physiological experiments, ecological metrics, critical limits, lethal limits, performance
Introduction

Throughout the history of thermal biology there has been a strong focus on understanding the thermal limits of critical whole organism functions, and the mechanisms underlying these limits. Physiology is typically measured at the level of individuals of a population and the outcomes are compared across geographic and temporal scales to infer fundamental patterns that constrain the evolution of thermal tolerance. A number of theoretical frameworks to explain patterns in thermal tolerance over large geographic scales have been developed (summarised in Gaston et al. 2008; Chown and Gaston 2008; Gaston et al. 2009), such as the latitudinal variation in thermal tolerance breadth (Janzen 1967) and acclimation capacity (Vernberg 1962). Physiological approaches now have strong societal value due to the requirement to identify the populations, species and communities that will win and lose as our climate changes (Somero 2010).

It is key to identify which controlled experiments of physiological and biological limits are most appropriate, as different mechanisms are likely to be more important over different time scales (Peck 2011). Intuitively, assessments of the ability to withstand acute events of minutes to days (extreme events and weather) will require different experiments than those designed to assess the capacity of organisms to withstand chronic exposure to long-term change (climate change and decadal oscillations). While many frameworks have been developed to explain the mechanisms underlying physiological and biological limits, explicit incorporation of time is often neglected (e.g., Pörtner 2002; Verbeck et al. 2016).
The dimension of time is fundamental for the design of experiments. Yet some large-scale analyses that synthesize thermal limits have ignored experimental duration and rate of experimental temperature change as key parameters for inclusion as covariates (e.g., Sunday et al. 2011; Hasnain et al. 2013), in spite of a wealth of literature highlighting these specific issues. Even when covariates on acclimation time and rates of heating and cooling are included, consideration of the rate of temperature change in relation to specific mechanisms which underpin classic response metrics (such as critical temperatures for activity or survival) often remain unexplored. For instance rates of heating may vary by many orders of magnitude but are combined into a uniform response curve, e.g., 0.0007 to 13.2 °C per minute: Comte and Olden 2017. Moreover, the literature is riddled with experiments that expose animals to short-term temperature treatments, and then infer sensitivity of species hundreds of years into the future, including our own work (e.g., Beers and Sidell 2011; Morley et al. 2012). As a consequence, key theoretical and methodological issues are embedded within the eco-physiological literature.

Here we draw on the experience that we have gained through conducting thermal limit experiments across hundreds of species (e.g., Bates et al. 2010; Morley et al. 2014a), in combination with arguments that have been made before (e.g., Helmuth et al. 2010; Peck 2011; Vinagre et al. 2015; Jutfelt et al. 2018). We focus on the temporal context of interpreting and designing laboratory experiments, primarily for aquatic invertebrates and fishes, and highlight the importance of the rate and duration of temperature change in shaping fundamental ecological and evolutionary processes.
First, we discuss how acute and chronic thermal limits will determine the capacity of organisms to resist, adjust and adapt to environmental extremes, as well as to long-term change. Second, we highlight why the duration of experiments and rate of temperature change is a fundamental experimental parameter that needs to be carefully considered when interpreting the results of thermal limit studies. Third, we define key metrics of thermal limits, and pinpoint where the temporal duration of classical experimental protocols are important in interpreting the ecological significance of different thermal limits. We synthesize data from Canadian freshwater fish (originally compiled in Hasnain et al. 2010) to illustrate the variability in different metrics of physiological and biological limits, based on experiments presenting both acute and chronic temperature exposures. We explore how inferences from these data can be classified as false positives or false negatives. Last, we offer suggestions on best practices for macrophysiological studies aiming to compare thermal limits across diverse taxa.

**Environmental Temperature Change and Temperature Limits: Resistance, Acclimation and Adjustment**

Physiological and biological responses to increasing temperature will vary depending on the rate and duration over which the warming occurs. At the very fastest rates of warming (seconds) physiological systems will not be able to respond in time to heating and rapid death will occur. This type of rapid heating does not often occur in nature and is limited to the interface of volcanic activity (e.g., hydrothermal systems), or following catastrophic events (e.g., large asteroid impact).
Biologically relevant rapid heating on scales of minutes to days, such as exposure to high air temperatures or solar heating (Helmuth 1998), migration through the thermocline (Elder and Seibel 2015), and heated thermal effluents (Verones et al. 2010), will heat organisms at a rate where physiological resistance pathways can provide short-term protection. It is resistance pathways, such as the production of heat shock proteins (Clark et al. 2008) or the switch to anaerobic metabolism (Pörtner 2002), that maintain physiological pathways and prevent cell death during acute temperature events. Short-term resistance to temperature can also be increased through changes in behavior, such as cessation of feeding or digestion, which can reduce the exposure to temperature or allow energy to be reallocated into resistance mechanisms (Morley et al. 2014b). Evaporative cooling can also slow the rate of heating, particularly in the intertidal during low tide (Lowell 1984; McMahon 1990; Helmuth 1998). However, the protection afforded by these resistance mechanisms is time limited and comes with a significant energetic cost (e.g., Williams et al. 2016). The energetic costs of resistance mechanisms are temporary, and essentially bridge the period before acclimation is complete, or until the end of the acute environmental perturbation.

If the rate of warming is slow enough (i.e., hours to months) then “acclimation” can occur. Acclimation refers to the changes in biochemical pathways and molecules that allow for a new stable physiological state in response to treatment conditions. Acclimation experiments measure an integration of energetic processes that return, for example, the temperature at which allocation of energy to reproduction is maximised given particular environmental conditions. Experiments are necessarily performed over longer time scales of weeks to months to record measurable
changes, in particular for species from colder climates (Peck et al. 2010; Morley et al. 2019). When multiple factors are altered along with temperature, acclimation capacity can be used to assess whether the health and condition of an animal has declined in response to long-term sub-lethal stressors, thus lowering its resistance to further warming (Folt et al. 1999). For example, under combined temperature and salinity stress, the growth rate of adult grass shrimp was reduced due to a change in the balance between energy gained through feeding and energy loss due to changes in metabolic rate (Vernberg and Piyatiratitivorakul 1998).

The expectation that acclimation will be beneficial has been the subject of much study (Wilson and Franklin 2002). Differential responses of individuals and populations at different times of year can provide a relative measure of capacity to acclimatise. Under identical laboratory conditions seasonal differences in the magnitude of physiological shifts in traits can indicate that a population or species undergoes temporal variation in acclimation capacity. Under seasonal warming, acclimatisation may result in a switch between biochemical pathways to produce, for example, a different isoform of an enzyme that has a higher thermal reaction norm (Angilletta et al. 2003, 2006). These new physiological states can be the result of changes in membrane fluidity (e.g., Cossins 1977) or shifts in energy allocation (for example the mobilisation of energy from stores to fuel increased costs, e.g., Das et al. 2006) that perform better under the altered temperatures. The benefit of more heat resistant pathways depends on the environmental stimulus being predictable, such as seasonal warming. This is because acclimation can take months and it is energetically costly to switch to new molecular pathways, which will be less efficient if conditions return to previous levels, or are not stable.
Presumably, there are also phylogenetic constraints on the potential for species to acclimate. For example, in the Antarctic where ocean temperatures are relatively stable and cold, acclimation may take many months (more than 3 months in marine invertebrates: Morley et al. 2019). Even so, polar invertebrates and fishes show both extremes of acclimation capacity. For example, springtails (Collembola) which live across the marine-terrestrial interface can cold harden on a daily basis as the temperature can change by more than 30°C in a 24-hour period (Peck et al. 2006).

By contrast, polar ectotherms experiencing subtidal ocean conditions are highly stenothermal, as are tropical marine ectotherms (Nguyen et al. 2011) and forest reptiles (Deutsch et al. 2008). Whereas, temperate marine ectotherms are more eurythermal, with wider thermal safety margins (Deutsch et al. 2008; Richard et al. 2012). This overarching geographic signature highlights the importance of different signals of environmental temperature variability that occur at different time scales in shaping thermal physiology (Waldock et al. 2018; Morley et al. 2019).

“Adaptation” refers to the process of genetic change resulting in resistance and acclimation pathways, and the patterns of thermal reaction norms in response to experienced climate. Thus adaptation is the process by which new pathways, with adjusted reaction norms, are expected to appear due to genetic mutation, followed by fixation of this mutation through natural selection (Sandblom et al. 2016). Genetic adaptation can take place over only a few months to many thousands of years, depending on generation time and the strength of the selective force (Thompson 1998; Rudman et al. 2017). The rates of warming in experiments leads to different variance of phenotypes, which in turn influences the heritability and evolvability of traits (Chown et al. 2010; Mitchell and Hoffman 2010). How temperature is
manipulated in experiments thus has important implications for the effect of the rate of climate change on the adaptive capacity of species in the wild.

Greater attention to how these terms are applied in the literature is required. Acclimation is often incorrectly used to refer to the period when animals are held after capture, and before the start of experiments. To avoid confusion with clearly defined biological functions, the time period over which animals are exposed to laboratory conditions prior to the initiation of experiments should more correctly be referred to as a period of adjustment from the effects of capture and transfer to laboratory conditions. Adaptation is also frequently used to describe short-term adjustment to experimental or laboratory conditions. These uses are rooted in the application of these terms in the human sciences. For instance, adaptation refers to the ability of society to change in response to conditions such as climate change (Moser and Ekstrom 2010) and in medicine, as a synonym for acclimation or acclimitization (e.g., Lenfant and Sullivan 1971). Importantly, physiological acclimation and adaptation do not typically occur at the same rate – these processes are shaped at macroecological scales by the environment (such as across latitude), and relate to differences in relative generation times within taxa, and rates of evolution. Therefore, tolerances to extreme events are not independent from either acclimatory or evolutionary capacity, as acute tolerance shifts with both (Gerken et al. 2015). Evolution will act across the diverse set of responses to shape thermal limits to the magnitude, variability and predictability of the experienced environment (Schulte et al. 2011). For example, in environments such as the south east coast of Australia, where temperature variation is unpredictable, daily variation in temperature can be as high as the annual variation in mean temperature. In this environment,
intertidal limpets (*Cellana tramoserica* (Holten, 1802)) have broad thermal windows, and because physiological tolerance limits are maximized, little capacity for further physiological adjustment (Morley et al. 2012). Similarly, high thermal tolerances are also known from species exposed to high thermal variability, such as from many terrestrial habitats due to strong diurnal variability in temperature (Sinclair et al. 2006), and at hydrothermal vents where heated effluents mix with cool ambient fluids (Bates et al. 2010).

**Influence of Experimental Duration and Rate of Temperature Change on Thermal Limit Estimates**

How temperature is manipulated and how animals are held in the laboratory influence the estimates of thermal limit endpoints achieved, and their interpretation. The most critical aspect of physiological experiments may be the duration of temperature exposure, and the rate of temperature change (e.g., Mora and Maya 2006; Terblanche et al. 2007; Peck et al. 2009). For example, a widely utilised metric, termed Critical Thermal (or Temperature) Maximum (*CT*$_{\text{max}}$, Table 1), is highly repeatable (Morgan et al. 2018), and can be directly compared across diverse taxa. *CT*$_{\text{max}}$ provides a relative measure of whole animal physiological condition that is dependent on the how quickly temperature is ramped up or down (e.g., Peck et al. 2009), as illustrated in the case study presented in Figure 1. The temperature limits (*CT*$_{\text{max}}$) of species decrease with slower warming rates in laboratory experiments that were otherwise executed using the same protocols (Fig. 1).

Lower ramping rates, therefore, lead to much lower estimates of *CT*$_{\text{max}}$, where the fundamental mechanisms setting critical thermal limits differ (Fig. 1). For example, if
experimental temperatures are altered by degrees per second or minute, then the assessment will be of an organism’s acute ability to cope with rapid and potentially short-lived extreme events (i.e., Resistance, Fig. 1). If experimental temperatures are altered by degrees per day or month, or longer, then acclimatory (Acclimation, Fig. 1), and ultimately evolutionary limits, are likely being tested (Adaptation, Fig. 1). Yet experiments are rarely conducted at durations longer than months due to practical constraints (Sandblom et al. 2016; Morley et al. 2019). In long duration experiments resistance and acclimation mechanisms will be important modifiers of patterns of adaptive potential (however, see Rezende et al. 2011). Indeed, much of the understanding gained on evolutionary shaping of temperature limits comes from studies that span across latitude and quantify patterns of resistance and acclimation (e.g., Sunday et al. 2011; Comte and Olden 2017).

Confusion in the literature will be decreased through careful consideration of whether resistance, acclimation and adaptation mechanisms are likely underpinning any observed shifts in thermal tolerance, as determined by the rate of temperature change and whether variability in thermal regime is incorporated in the experimental design. Researchers often leap to the conclusion that low acute tolerance indicates a poor ability to cope with climate change and visa-versa. Instead, experiments conducted at rapid rates of warming are conceptually matched to a framework of resistance and acclimation to extreme temperature events (which are increasing in both frequency and magnitude, Frolicher et al. 2018). Indeed, acute ramping experiments can be considered as analogous to an extreme temperature exposure in the wild. Metrics of critical performance (Table 1) assessed using rapid protocols may therefore be key for predicting which populations and species will be most
sensitive to heat waves and cold spells, thus predicting strong biological responses such as geographic range contractions (Wernberg et al. 2016). A key direction for future research is to understand if species that are able to survive extreme events have used so much energy to cope with these extremes that their fitness is compromised (McKenzie et al. 2016). Moreover, accumulating evidence suggests that laboratory conditions where aquaria temperatures are stable lead to different physiological responses in comparison to designs which include thermal variability during adjustment periods or during treatment conditions (Gallant et al. 2017; Morash et al. 2018), highlighting a second key area of research that will advance the field.

**Metrics of Physiological and Biological Thermal Limits**

Different thermal limit endpoints have developed over hundreds of years with inconsistent terminology (Table 1 presents examples and terminology). As introduced above, depending on experimental design, thermal limits represent resistance, acclimation and evolutionary adaptation and can therefore be easily misrepresented. Acute thermal limits are usually measured using lethal (e.g., UILT and LILT; Table 1) or loss of equilibrium (e.g., CT_{max}) measures at rates of warming over minutes to hours (Table 1). Critical and pejus temperatures, and the onset of spasms (also defined in Table 1), are also typically tested at rapid rates of warming, but not always (e.g., Nguyen et al. 2011). Indeed, acclimation and acclimatisation are often quantified through a change in acute thermal limits (Morley et al. 2019). For example, a shift in CT_{max} to a higher thermal limit in response to a warmer incubation temperature indicates that beneficial acclimation of acute thermal limits has occurred. Acclimation capacity can also be quantified by measuring thermal optima...
(e.g., growth and reproduction) and temperature preference using static protocols or a temperature choice apparatus (Table 1). Similarly, tolerances and optima have been shown to change across generations, from parents to offspring, termed transgenerational plasticity (e.g., Donelson et al. 2012; Diamond et al. 2017; Morley et al. 2017; Donelson et al. 2018). The degree to which this selection of more tolerant phenotypes becomes fixed into the population is still under investigation, yet transgenerational plasticity provides a relatively responsive mechanism through which tolerance limits can shift and represents a rapidly advancing field (Donelson et al. 2018).

In addition to variation across experimental protocols, numerous different and related metrics are used in thermal physiological experiments whereby the definitions and statistical procedures estimating different thermal limits also vary. For instance, \( CT_{\text{max}} \) has multiple definitions (Lutterschmidt and Hutchison 1997). Similarly, the terminology describing “lethal” thermal endpoints is equally as confused with many similar abbreviations indicating very different experimental procedures and statistical approaches (Kilgour and McCauley 1986). For example, the time that an experimental population can resist exposure to heat or cold (static method, LT50 in Table 1), versus a ramping design where the temperature at which a particular proportion of the population succumbs to death (dynamic method: UTL50 in Table 1, Mora and Maya 2006) - both rely on a lethal endpoint. Moreover, similar abbreviations are used for different metrics. For example, Venables et al. 1978 used “TL50” for the time to 50% mortality. Bennett and Judd 1992 used low temperature static experiments and describe the exposure time producing 50% mortality as “TL50”. Compton et al. 2007 apply logit linear regression to find the point where 50%
of bivalves died, labelling this endpoint the upper (and lower) thermal limit (ULT and LLT). Korhonen and Lagerspetz (1996) observed shrimp at a high temperature (33°C) until its pleopods stopped moving and called this LT = lethal time (after Lagerspetz and Bowler 1993). Korhonen and Lagerspetz (1996) also used a ramping rate of 12°C per minute to estimate the “LTMax”. The percentage of an experimental population that reaches a pre-defined endpoint can also distinguish an end-point, such as the time until 50 or 100% mortality occurs (LT50 vs. LT100, Table 1). There are also a number of different ways to calculate the population level upper lethal temperature limit (LT50). For example, the statistical mid-point, can be parameterized as either the mean or median of mortality. Alternatively, probit analysis can be used to fit a logistic regression to identify a mid-point. All of these approaches will result in different estimates for a particular temperature threshold – thus, features of the study design strongly influence the endpoint obtained.

While many different methodological and statistical approaches have been used to measure temperature limits, physiological performance is typically conceived as a maximum within a narrow temperature range where optimal temperatures for many processes center around a specific value (Beitinger and Fitzpatrick 1979). Yet whether the optimal temperatures for different physiological and biological processes converge upon the same optimum is unknown for most species. Indeed, as environmental temperatures approach critical limits, it is expected that loss of whole-organism function unfolds as a hierarchy of response losses across different processes and pathways (Pörtner 2002). The optimal temperature for growth may be very different in comparison to other physiological or ecological processes, such as spawning temperature (Hasnain et al. 2010). This loss may not act solely through
direct limitation, but can also occur through the reallocation of resources from one process to another. There is often a choice to allocate energy away from reproduction, as temperature increases, and towards essential functions such as the production of molecular chaperones (the heat shock proteins) that help maintain enzyme function, and therefore, sustain life in the short-term (Clark et al. 2008). It may be unrealistic to expect similar estimates of thermal limits across different levels of biological organization, from the failure of physiological pathways (e.g., the multiple performance-multiple optima paradigm: Clark et al. 2013), through organ systems, to loss of a whole animal function such as locomotion, highlighting a further direction for future research.

**Case Study Illustrating Challenges in Interpreting Data Syntheses of Thermal Limits**

Here we illustrate both confounding issues and insights that can be gained through synthesizing thermal limit data collected using different methodologies. We use a case study that combined different temperature metrics across 84 species of freshwater Canadian fishes (originally reported in Hasnain et al. 2010). We selected species found living in temperate freshwater systems that were represented by response metric data for at least two different temperature metrics (the mean number of metrics compared across each species was 4). In this compilation, experimental designs included both dynamic and static temperature treatments, depending on the thermal limit being assessed (Fig. 2: methods are described in the figure legend text). As is the case for many such studies, the experimental features underpinning the different tests for each species were not factored in to species-specific estimates for each metric (Hasnain et al. 2010). Hence teasing out trends
due to phylogenetic history and life-history traits, or identifying species-specific responses, must be accompanied with careful integration and consideration of the experiments that underpin estimates of temperature optima and tolerance. Any differences between species or taxonomic groups may be due to systematic differences in experimental design, thus returning a false positive. For example, in Figure 2, $CT_{\text{max}}$ is measured using dynamic heating, but is combined with UILT, which was measured using static heating protocols (ULT in this case was assessed after a certain time following exposure to a particular temperature, and design varied across studies). Yet overall, the upper limit for mortality (ULT) is lower than the temperature for loss of equilibrium ($CT_{\text{max}}$). To understand the mechanisms explaining the difference in UILT and $CT_{\text{max}}$ values, the various experimental protocols (e.g., selection of a dynamic versus static protocol, temperature ramping ramp, and duration of treatment exposure) would need to be considered to attribute which resistance and acclimation responses were likely at play, or even better, standardized experiments to address this specific question.

While some hypotheses cannot be addressed through data synthesis approaches, others can. Systematic bias across large numbers of species is unlikely (a fundamental premise of macroecology), and so the overall modelled estimates for the different metrics are likely representing real biological patterns (Fig. 2). Ignoring strong patterns translates to a false negative. Here we find a large (10s of degrees Celsius) mismatch between temperature metrics describing optimal versus lethal endpoints. Our statistical modelling approach (see Table 2 and Fig. 2) accounts for species identity and provides an estimate of the average across species where multiple temperature metrics were included for each species. On average, optimal
temperatures for spawning were lower than for egg development, which were both lower than for growth. These metrics were obtained under constant experimental conditions, and therefore, capture acclimation and adaptation responses, and support that different thermal optima may be important at different life history stages (e.g., Clark et al. 2013).

Temperature preference (the temperature which individuals migrate to and select when offered a choice in experimental aquaria) was also, on average, 1.9°C lower than the temperature that was optimal for growth. This finding goes against the “final preferendum” hypothesis, which predicts that preferred temperatures and optimal growth will align (Reynolds and Casterlin 1979). We can also conclude from this analysis that if the evolution of temperature preference is truly optimised to maximise fitness, it must be matched to some other function than reproduction (as both OS and OE, on average, were relatively cool, Fig. 2).

Further, we find that metrics of upper thermal limits are at much higher temperatures than metrics of preference or performance for these species (on average, >29°C versus <20°C, Fig. 2). Our analysis suggests that studies that use metrics such as \( \text{CT}_{\text{max}} \) (i.e., critical loss of function following rapid warming) to predict species sensitivity thresholds with long-term warming will be inaccurate (although this metric may be useful to assess the impact of acute exposures, e.g., Morley et al. 2019). Sensitivity to long-term warming, may instead, be better predicted by looking at responses, such as spawning, which are highly temperature sensitive and are limited at thresholds occurring at much lower temperatures.
Best Practices

We present a number of arguments identifying why both the biological (metric selection) and methodological (experimental design) context must be carefully considered when combining thermal limit data. Failure to appropriately combine and interpret different thermal limits and tolerance metrics may misrepresent meaningful biological processes. This is of particular concern, given that thermal limit data sets combining global efforts are increasingly emerging, such as a recent compilation coined as GlobTherm (Bennett et al. 2018). A subset of 262 aquatic species from GlobTherm (fishes: Actinopteri and Chondrichthyes, invertebrates: Asciidacea, Bivalvia, Branchiopoda, Bivalvia, Gastropoda, Malacostraca, Ostracoda) returns metrics reported as CTmax, LT50 and LT100 with ramping rates ranging from 0 to 1°C per minute. While a “pre-treatment” temperature is reported in the study metadata, the duration at which animals were held in the lab prior to experimentation is not included. It is likely that many studies did not include this information within the methodological details. Moreover, the reported metrics are often not directly comparable. For example, LT100 from Bates et al. 2010 were achieved with a temperature ramping protocol that allowed a heating and cooling period for recovery, a design that cannot be directly compared to the results from other studies with linear ramping rates – yet within the GlobTherm metadata, there is no way for users to identify the confounding methodological issues in the original study designs.

We caution that while syntheses of different response data can provide useful insights, data are now available to researchers who may not have the physiological background to identify the subtleties between the metrics, but are nevertheless incorporating these data into analytical approaches. Failure to appropriately interpret
patterns from combining different thermal metrics may lead to inaccurate predictions about species vulnerabilities and distribution constraints. Yet, rather than dismissing particular metrics in favour of others or arguing if data synthesis approaches can provide meaningful insights, we advocate for careful experimental design, comprehensive reporting of experimental methods, and the responsible re-use and interpretation of metrics.

Rather than advocating for standard protocols for testing different scales of thermal tolerance, we encourage the design of experiments that are explicitly matched to hypotheses that directly link to physiological mechanisms. To advance our knowledge of the factors driving thermal physiology, experiments need to be designed to mimic daily, monthly and seasonal environmental variability, for different physiological rates and generation times of species found across different latitudes. For example, changes of 1°C per hour may be ecologically relevant for testing resistance mechanisms in the tropics, whereas 1°C per day might be more appropriate in polar oceans. Similarly, 1°C per month will test adaptive capacity of tropical species which have fast generation times, while for ectotherms in the Antarctic, 1°C change in 10 years is likely more ecologically relevant (while logistically challenging possible through coordinating across different graduate student projects, for instance). End-points such as loss of equilibrium, which is equivalent to “functional mortality”, is a consistent metric for comparison across latitudes that integrates mechanisms occurring at the molecular, cellular and tissue scales. The result will be a more common “currency” for direct comparison of temperature sensitivity across populations, species and communities. Such attempts will enhance
our ability to incorporate physiological data into predictions of vulnerability at different scales of biological organization.

We expect that experiments designed to disentangle the different dimensions of physiological tolerance, and how physiology in turns translates into ecological performance, will be key in moving the field of eco-physiology forwards. We predict experiments that include interactions between the duration, magnitude, intensity and frequency of temperature exposure, coupled with the appropriate thermal limits, will reveal unexpected physiological and biological responses with important conservation implications (McKenzie et al. 2016). While time consuming and difficult to achieve, there needs to be a combination of both short- and long-term experiments whose results are carefully interpreted to understand which tolerance mechanisms are most important under different environmental variability scenarios.

Acknowledgements
Salary support to AEB was through the Canada Research Chairs Programme. SAM is funded through Natural Environment Research Council funding to the British Antarctic Survey. We appreciate comments on a draft version from K. Gamperl and J. Chu.

References


Applied Research and Development Branch, Ontario Ministry of Natural Resources.


FIGURE LEGENDS

Figure 1. Case study highlighting thermal limits ($CT_{max}$) for experiments conducted using a similar design. Species exposed to different rates of warming in laboratory experiments ranging from $1^\circ C$ increases over minutes to months. Different timescales represent resistance, acclimation and adaptation mechanisms, with a paucity of data for longer time durations of warming (due to the difficulty in holding animals in experimental systems for long time periods). Individual lines are the regression slopes for species from Singapore: Nguyen et al. (2011), Antarctic Peninsula: Peck et al. (2009), and Ascension Island, New Zealand and McMurdo Sound: Morley et al. (2014a). Regressions show the slope of the decline in $CT_{max}$ for each species at different warming rates, where slower ramping rates in these laboratory experiments lead to much lower estimates of $CT_{max}$ that are likely underpinned by fundamentally different mechanisms.

Figure 2. Case study showing box-and-whisker plots of the thermal limits (y-axis) for key ecological and physiological metrics across species of Canadian freshwater fishes. The gray line is the median across all experimental data for each metric type with the box and whiskers representing the 5th, 25th, 75th, and 95th quantiles. Solid dots are the modelled coefficient estimates after accounting for the taxonomic structure of the data with multiple metrics reported for the same species (reported in Table 2). Data were extracted for analysis from Hasnain et al. (2010). The number of species included to estimate the temperature for each metric differs due to data availability, with 84 species included in total. Optimal spawning temperature (OS);
optimal temperature for egg development (OE); optimal growth temperature (OGT); temperature preference (Preferred); upper incipient lethal temperature (UILT) and critical thermal maximum ($CT_{max}$). Note that OS, OE, OGT, and UILT were measured using static temperature exposure, while $CT_{max}$ was through dynamic heating.
Figure 1. Case study highlighting thermal limits (CTmax) for experiments conducted using a similar design. Species exposed to different rates of warming in laboratory experiments ranging from 1°C increases over minutes to months. Different timescales represent resistance, acclimation and adaptation mechanisms, with a paucity of data for longer time durations of warming (due to the difficulty in holding animals in experimental systems for long time periods). Individual lines are the regression slopes for species from Singapore: Nguyen et al. (2011), Antarctic Peninsula: Peck et al. (2009), and Ascension Island, New Zealand and McMurdo Sound: Morley et al. (2014a). Regressions show the slope of the decline in CTmax for each species at different warming rates, where slower ramping rates in these laboratory experiments lead to much lower estimates of CTmax that are likely underpinned by fundamentally different mechanisms.
Figure 2. Case study showing box-and-whisker plots of the thermal limits (y-axis) for key ecological and physiological metrics across species of Canadian freshwater fishes. The gray line is the median across all experimental data for each metric type with the box and whiskers representing the 5th, 25th, 75th, and 95th quantiles. Solid dots are the modelled coefficient estimates after accounting for the taxonomic structure of the data with multiple metrics reported for the same species (reported in Table 2). Data were extracted for analysis from Hasnain et al. (2010). The number of species included to estimate the temperature for each metric differs due to data availability, with 84 species included in total. Optimal spawning temperature (OS); optimal temperature for egg development (OE); optimal growth temperature (OGT); temperature preference (Preferred); upper incipient lethal temperature (ULT); critical thermal maximum (CTmax). Note that OS, OE, OGT, and UILT were measured using static temperature exposure, while CTmax was through dynamic heating.
TABLES

Table 1. Definitions of metrics used to compare thermal tolerance between individuals, populations and species. The temporal scale at which the different metrics are measured influences the thermal capacity (resistance, acclimation, and adaptation) being tested (developed from Peck 2011) – thus these metrics are not directly comparable across populations and species if the methodological timing of exposure in experiments varies (e.g., dynamic versus static protocol, or different rates of temperature change). An asterix (*) indicates metrics shown in Figure 2.

<table>
<thead>
<tr>
<th>Physiological and Biological Metrics of Performance</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Critical thermal maximum or minimum (CT&lt;sub&gt;max&lt;/sub&gt; or CT&lt;sub&gt;min&lt;/sub&gt;)</td>
<td>Temperature at which an individual loses equilibrium, or ability to perform a function, e.g. cardiac function, feeding, muscular activity</td>
</tr>
<tr>
<td>*Upper or lower incipient lethal temperature (UILT or LILT), sometimes IULT and ILLT</td>
<td>Temperature at which mortality occurs in a predefined % of individuals in a population during a specified time interval</td>
</tr>
<tr>
<td>Onset of spasms</td>
<td>Temperature at which the onset of uncontrolled muscular contractions occurs</td>
</tr>
<tr>
<td>Upper or lower lethal temperature (ULT or LLT)</td>
<td>Temperature at which a predefined % of mortality occurs, e.g., ULT50 or ULT100</td>
</tr>
<tr>
<td>LT50 LT100</td>
<td>The time at which 50 or 100% mortality occurs</td>
</tr>
<tr>
<td>Critical temperature limits</td>
<td>Temperature beyond which aerobic metabolism is no longer sufficient, leading to time limited survival</td>
</tr>
<tr>
<td>Pejus limits</td>
<td>Upper and lower boundaries of the optimal thermal range</td>
</tr>
<tr>
<td>*Optimum growth temperature (OGT)</td>
<td>Temperature at which growth is maximized</td>
</tr>
<tr>
<td>*Optimum temperature for reproduction</td>
<td>Temperature at which reproduction (spawning output, development, survival, etc.) is maximized</td>
</tr>
<tr>
<td>*Temperature preference</td>
<td>Chosen temperature within a gradient - often equated to optimum temperature</td>
</tr>
</tbody>
</table>
Table 2. Estimated temperatures for different biological and physiological limits in 84 species of Canadian freshwater fishes (compiled in Hasnain et al. 2010). Our goal was to estimate the ranking of the different temperature limits across diverse species. A linear mixed effects model (family/genus/species was included as nested random effects) to control for non-random variation in temperature limits due to taxonomic identity (a proxy for phylogeny), fitted using restricted estimation likelihood in the R package “nlme”(Pinheiro et al. 2016). Lower confidence interval (lci) and upper confidence interval (uci). See Fig. 2 for further details including the definitions of the thermal limits.

<table>
<thead>
<tr>
<th></th>
<th>lci (5%)</th>
<th>estimate</th>
<th>uci (95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OS</td>
<td>12.6</td>
<td>14.8</td>
<td>17.1</td>
</tr>
<tr>
<td>OE</td>
<td>14.2</td>
<td>16.5</td>
<td>18.7</td>
</tr>
<tr>
<td>Preferred</td>
<td>17.9</td>
<td>20.0</td>
<td>22.2</td>
</tr>
<tr>
<td>OGT</td>
<td>19.7</td>
<td>21.9</td>
<td>24.1</td>
</tr>
<tr>
<td>UILT</td>
<td>27.2</td>
<td>29.5</td>
<td>31.7</td>
</tr>
<tr>
<td>CT_{max}</td>
<td>29.1</td>
<td>31.3</td>
<td>33.5</td>
</tr>
</tbody>
</table>