

## RESEARCH ARTICLE

# The analysis and interpretation of critical temperatures

Joel G. Kingsolver\* and James Umbanhowar

## ABSTRACT

Critical temperatures are widely used to quantify the upper and lower thermal limits of organisms. But measured critical temperatures often vary with methodological details, leading to spirited discussions about the potential consequences of stress and acclimation during the experiments. We review a model based on the simple assumption that failure rate increases with increasing temperature, independent of previous temperature exposure, water loss or metabolism during the experiment. The model predicts that mean critical thermal maximal temperature ( $CT_{max}$ ) increases non-linearly with starting temperature and ramping rate, a pattern frequently observed in empirical studies. We then develop a statistical model that estimates a failure rate function (the relationship between failure rate and current temperature) using maximum likelihood; the best model accounts for 58% of the variation in  $CT_{max}$  in an exemplary dataset for tsetse flies. We then extend the model to incorporate potential effects of stress and acclimation on the failure rate function; the results show how stress accumulation at low ramping rate may increase the failure rate and reduce observed values of  $CT_{max}$ . We also applied the model to an acclimation experiment with hornworm larvae that used a single starting temperature and ramping rate; the analyses show that increasing acclimation temperature significantly reduced the slope of the failure rate function, increasing the temperature at which failure occurred. The model directly applies to critical thermal minima, and can utilize data from both ramping and constant-temperature assays. Our model provides a new approach to analyzing and interpreting critical temperatures.

**KEY WORDS:** Critical maximum and minimum temperatures, Critical thermal maximum, Heat stress, Statistical model

## INTRODUCTION

The upper and lower thermal limits of organisms are important determinants of their ecology, distribution and evolution (Angilletta, 2009). Measurements of thermal limits are increasingly used to predict organismal responses to climate change and climatic gradients, and empirical estimates of upper and lower thermal limits are now available for a variety of ectotherms (Chown and Terblanche, 2007; Sunday et al., 2011; Buckley and Huey, 2016). However, estimates of thermal limits can depend on the experimental methods and conditions in which they are measured, so it is important to understand the factors that may influence thermal limits (Lutterschmidt and Hutchison, 1997).

Thermal limits are often quantified in terms of critical temperatures using a dynamic measurement. For example, determining the critical thermal maximum ( $CT_{max}$ ) involves briefly acclimating an individual to some starting temperature ( $T_0$ ), increasing the temperature at some

controlled or experimentally targeted ramping rate ( $r$ ), and noting the temperature (and time) at which physiological failure occurs (e.g. onset of muscle spasms, knockdown, loss of righting response). This technique provides a repeatable, rapid method for providing a quantitative measure of the thermal limit for each individual in a sample, unlike static methods that measure the binary response (e.g. fail or not fail) at a constant temperature.  $CT_{max}$  (and the critical thermal minimum,  $CT_{min}$ ) also has a simple biological interpretation as the temperature at which failure occurs. These important advantages have facilitated a burgeoning body of empirical literature on critical temperatures (Sunday et al., 2011; Buckley and Huey, 2016).

Terblanche and colleagues (2007) demonstrated how methodological details during ramping experiments can systematically affect estimates of  $CT_{max}$  and  $CT_{min}$ , a finding confirmed by subsequent studies in other organisms (Overgaard et al., 2011; Terblanche et al., 2011). For example, they found that  $CT_{max}$  increased with heating rate, and  $CT_{min}$  decreased with cooling rate. In addition, higher starting temperatures generally increased both  $CT_{max}$  and  $CT_{min}$ , at least at lower ramping rates. Ramping rate and starting temperature have substantial effects, changing mean  $CT_{max}$  and  $CT_{min}$  by 6–8°C or more (Terblanche et al., 2007). These methodological effects on estimates of critical temperatures have been interpreted in terms of the amount of time that organisms are exposed to stressful high or low temperatures during the experimental ramping assay (Overgaard et al., 2011; Rezende et al., 2011; Terblanche et al., 2011). There has been spirited discussion about the ‘best’ experimental conditions for measuring critical temperatures. For example, if rates of temperature change in natural environments are slow, then low ramping rates are most ecologically relevant for measuring critical temperatures (Terblanche et al., 2011). Conversely, slow ramping rates could cause substantial metabolic and water loss costs during the measurement period, and (in combination with low starting temperatures) allow time to develop acclimation responses, altering the physiological state of the organism and the value of  $CT_{max}$  (Overgaard et al., 2011; Rezende et al., 2011).

Several recent papers have modeled how exposure time may affect the temperature at which failure occurs in constant-temperature experiments and in ramping experiments (Santos et al., 2011; Rezende et al., 2014). These models define the ‘basal’ upper thermal limit as the temperature at which the rate of mortality or failure is  $1 \text{ min}^{-1}$  (see Fig. 1). Rezende and colleagues (2014) used a thermal death time (TDT) model to quantify how increasing exposure time will reduce the temperature at which failure occurs to below the basal thermal limit. Santos et al. (2011) used simulations to illustrate how water loss and metabolism during ramping experiments can further reduce the temperature at which failure occurs, and why slower ramping rates can cause failure at temperatures below the basal upper thermal limit.

In this study, we review the basic statistical reason why methodological details alter observed critical temperatures, and we

Department of Biology, University of North Carolina, Chapel Hill, NC 27599, USA.

\*Author for correspondence (jgking@bio.unc.edu)

 J.G.K., 0000-0002-9839-0208

Received 7 August 2017; Accepted 26 April 2018

develop a new statistical approach to estimating critical temperatures. The focus of this approach is to determine the failure rate function: the relationship between the current rate of failure and the current temperature, independent of the previous history of temperature exposure. First, we illustrate why decreasing ramping rate or starting temperature necessarily decreases estimates of  $CT_{max}$ , in ways often reported in the empirical literature. We also describe how our analyses of failure rates relate to previous methods used to model critical temperatures (Santos et al., 2011; Rezende et al., 2014). Second, we develop a new statistical model using maximum likelihood that accounts for the effects of ramping rate and starting temperature, and apply the model to the dataset of Terblanche et al. (2007). Our analyses show the model can account for much (but not all) of the variation in  $CT_{max}$ , independent of the effects of prior thermal history. Third, we extend the statistical model to incorporate the effects of other biological and experimental factors, including stress and acclimation. We apply this model to two datasets to illustrate how stress or acclimation can alter the position or slope of the failure rate function. Our models provide a new set of tools for designing experiments, analyzing data and understanding the meaning and determinants of critical temperatures.

## STATISTICAL MODELS AND RESULTS

### Failure rate, current temperature, and the time and temperature at failure: a simple model

Our initial assumption (which we will relax in subsequent sections) is that the instantaneous rate of 'failure' ( $M$ ) of an organism depends solely on its current temperature ( $T$ ), independent of thermal history, so there is a failure rate function  $M(T)$ . As a simple example,

suppose the failure rate increases exponentially with  $T$  for temperatures above some starting temperature  $T_0$ :

$$M(T) = ae^{b(T-T_0)}, \quad (1)$$

where  $a$  is the baseline failure rate and  $b$  is the rate parameter. Fig. 1 illustrates several failure rate functions that we will consider below, including the exponential function defined by Eqn 1.

Note that in the special case where temperature  $T$  is constant over time  $t$ , the failure rate  $M(T)$  is also constant. In this case, the relationship between the expected value of time to failure ( $t_c$ ) and (constant) temperature is known as the thermal death time (TDT) curve, which has been widely used in the microbiological and food storage literature (Tang et al., 2007). Using standard results from survival analysis (Cox and Oakes, 1984), constant rates of failure lead to expected failure times of  $M(T)^{-1}$ . Using the failure rate from Eqn 1, we arrive at an expected time to failure:

$$E(t) = \frac{e^{b(T_0-T)}}{a}. \quad (2)$$

Rezende et al. (2014) applied this approach to model heat tolerance, and similarly illustrated how under constant temperature conditions, the mean (expected) time to failure increases exponentially with decreasing temperature.

Suppose the temperature at time  $t=0$  is the starting temperature  $T_0$ , and temperature then increases linearly at ramping rate  $r$  (here we ignore potential effects of thermal inertia: see Appendix S2 of Rezende et al., 2014). As a result, the temperature  $T(t)$  at time  $t$  is simply:

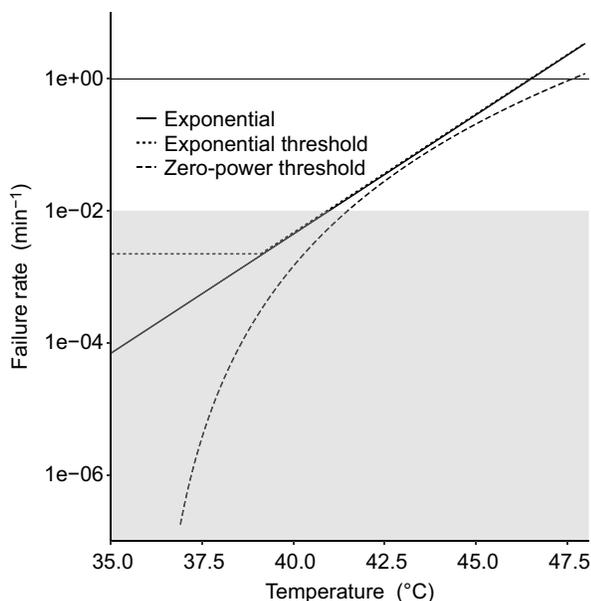
$$T(t) = T_0 + r(T - T_0). \quad (3)$$

These two equations imply that the failure rate also increases with time. Let the time at failure be  $t_c$  and the temperature at failure be  $CT_{max}$ : i.e.  $CT_{max} = T(t_c)$ . As we describe in the next section, we can combine Eqns 1 and 3 to compute the expected values of  $t_c$  and  $CT_{max}$ . A closed-form solution is not always possible, but simulations show that  $CT_{max}$  increases non-linearly with both ramping rate and starting temperature (with the reverse effects on failure time) (Fig. 2). Note that  $CT_{max}$  continues to increase with increasing ramping rate even at high ramping rate, although the rate of increase declines (see Discussion).

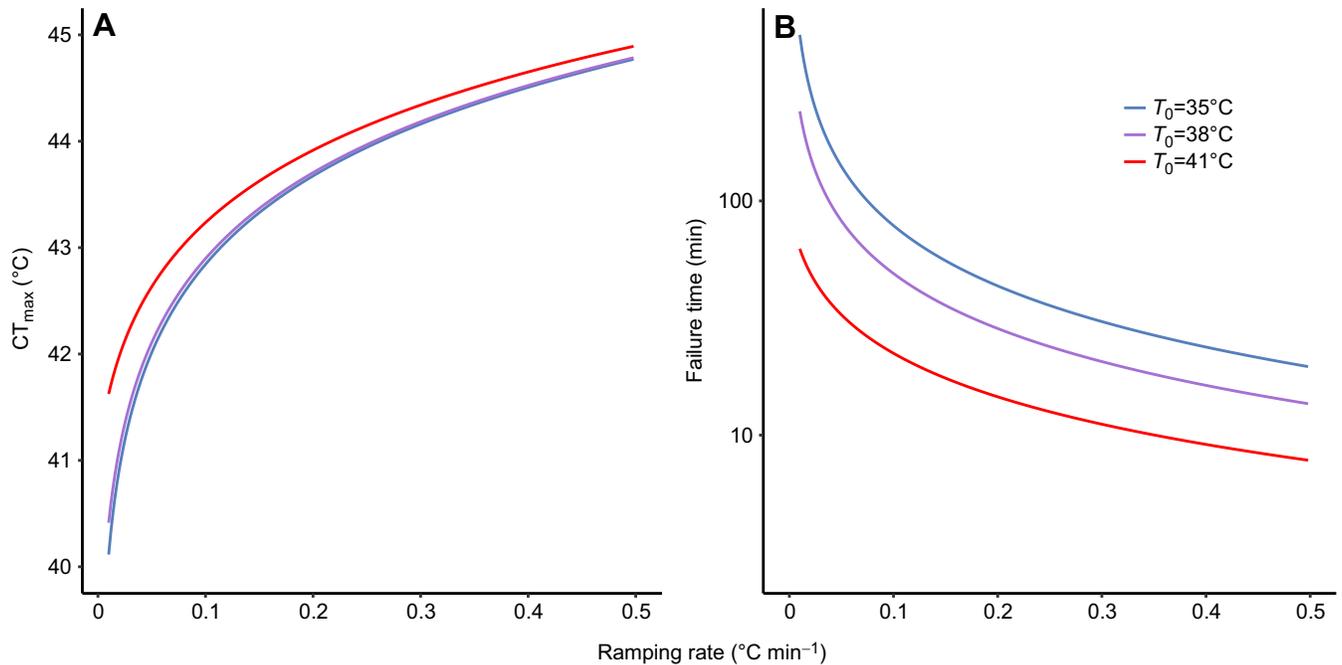
The key to understanding these effects is that expected values of failure time and the corresponding  $CT_{max}$  depend both on the current rate of failure and on the likelihood that failure did not occur at any previous time. The qualitative effects of ramping rate and starting temperature on  $CT_{max}$  will occur for any failure rate function  $M(T)$  that increases monotonically.

Santos et al. (2011) similarly demonstrated how several mechanisms can drive methodological details to produce different estimates of  $CT_{max}$ . In particular, they used detailed physiologically based models to produce simulations that model the temporally varying probability of survival for individuals. These simulations demonstrated that water loss and decreasing tolerance to high temperature can lead to negatively biased estimates of critical thermal temperatures in slow ramping experiments. Parameters and functional forms in these models were estimated individually based on an extensive knowledge of the biology of the target organisms. In contrast, the approach we develop here provides a statistical method for fitting unknown parameters and functional forms.

Our simple example above illustrates a statistical reason why  $CT_{max}$  and failure time depend on the methodological conditions used in dynamic assays of critical temperatures. In the next section, we develop a more general set of statistical models for analyzing data on critical temperatures that account for this effect.



**Fig. 1. Failure rate (log scale) as function of temperature for three failure rate models: exponential, exponential threshold and zero-power threshold.** See Table 1 for the equations for these models. The shaded region indicates failure rates below  $10^{-2} \text{ min}^{-1}$ : in this region, the expected time to failure is greater than 100 min (longer than the duration of most critical temperature experiments). The horizontal line indicates a failure rate of  $1 \text{ min}^{-1}$ , representing the 'basal' critical thermal maximum ( $CT_{max}$ ) parameter as defined by Santos et al. (2011) and Rezende et al. (2014). The best model for the tsetse fly data is the zero-power threshold model, with parameter estimates: baseline failure rate  $a=2.82 \times 10^{-7}$ , failure rate parameter  $b=6.13$ , threshold temperature  $T_c=36.0^\circ\text{C}$ . See 'Statistical models and results' for details.



**Fig. 2.** Expected values of  $CT_{\max}$  and failure time as a function of ramping rate for three different starting temperatures ( $T_0$ ). (A)  $CT_{\max}$  and (B) failure time at 35, 38 and 41°C. Predictions are based on an exponential failure rate function (Fig. 1).

### The general model framework

Here, we derive a general formulation to fit temperature-dependent failure rates using ‘survival’ (in the current context, ‘non-failure’) data (Cox and Oakes, 1984). The fundamental measure of a survival analysis is the time between events,  $t$ . In the context of temperature-dependent survival data, this will be the time until failure of the subject. We will describe how to use standard techniques of survival analysis to derive the probability density function (pdf) of the survival time, and then use maximum likelihood estimation to find the best-fitting parameters for a failure rate function.

The heart of a parametric survival analysis is the hazard function, which is equivalent to our failure rate function,  $M(t)$  (Cox and Oakes, 1984). Distributions of survival times from this function can be derived with the assumption that all failure events are independent from each other and determined solely by the hazard function. These assumptions describe a Poisson process which has easily determined statistical properties for the distribution of events given a rate process. The pdf for survival times given a failure rate function is:

$$f(t) = M(t)e^{-\int_0^t M(x)dx}. \quad (4)$$

In the case that the failure rate is an exponential (Eqn 1), we can use Eqn 4 to generate a pdf for the survival times dependent on the initial temperature and ramping rate. Using Eqn 2, we obtain the hazard function:

$$M(t) = ae^{b(T_0+rt)}. \quad (5)$$

Substituting Eqn 5 into Eqn 4 and integrating, we get the pdf for survival times where mortality is an exponential function of temperature.

$$f(t) = ae^{b(T_0+rt)}e^{-(ae^{bT_0}/r)(e^{rt}-1)}. \quad (6)$$

Note that this particular distribution is known as the Gompertz distribution and is commonly used in survival analysis. Not all mortality or failure functions will necessarily lead to well-studied probability distributions (Cox and Oakes, 1984).

The next step is to generate the best-fit parameters for a particular failure rate function. We choose to do this via maximum likelihood estimation (MLE). Suppose we have a set of  $n$  failure times,  $t_{c1}, t_{c2}, \dots, t_{cn}$ , from an experiment that are independent and identically distributed, i.e. have the failure rate function. The joint distribution function for this will be:

$$f(t_{c1}, t_{c2}, \dots, t_{cn} | \theta) = \prod_{i=1}^n f(t_{ci} | \theta). \quad (7)$$

Here,  $\theta$  is vector of parameters. The likelihood is then defined as:

$$L(\theta | t_{c1}, t_{c2}, \dots, t_{cn}) = f(t_{c1}, t_{c2}, \dots, t_{cn} | \theta) = \prod_{i=1}^n f(t_{ci} | \theta). \quad (8)$$

Typically, log likelihoods are used because they are easier to compute, so the log likelihood,  $l(\theta)$  is:

$$l(\theta) = \sum_{i=1}^n f(t_{ci} | \theta). \quad (9)$$

We then use computational methods to maximize Eqn 8 over values of the parameter set. For the exponential failure rate function, this parameter set is the two parameters  $a$  and  $b$  (Eqn 1). Unless otherwise stated, numerical optimization was completed using the `bbmle` package in the R software program using the Nelder–Mead algorithm within function `mle2` (Wickham, 2009; <https://CRAN.R-project.org/package=bbmle>).

### Choice of failure functions

The relationship between temperature and failure may vary depending on the organism and aspect of ‘failure’ considered (Cossins and Bowler, 1987; Tang et al., 2007); here, we used the strategy of assessing the fit of a set of plausible functions (Table 1). There are three basic components to each of the functions that we fit. The first is the functional form for changes in failure rate with increasing temperature. Here, we considered exponential, power and

**Table 1. Failure rate functions fitted to the  $CT_{max}$  data from Terblanche et al. (2007)**

Model	Equation	AIC	No. of parameters
Constant	$M(t)=m$	1041.2	1
Constant threshold	$T < T_c: M(T)=0$ $T > T_c: M(T)=m$	928.5	2
Exponential	$M(T) = ae^{b(T-T_0)}$	769.6	2
Exponential threshold	$T < T_c: M(T)=0$ $T > T_c: M(T) = ae^{b(T-T_c)}$	768.4	3
Zero exponential	$M(T) = ae^{b(T-T_0)} - a$	769.55	2
Zero exponential threshold	$T < T_c: M(T)=0$ $T > T_c: M(T) = ae^{b(T-T_c)} - a$	770.9	3
Power	$M(t)=c+a(T-T_0)^b$	782.9	3
Zero power	$M(t)=a(T-T_0)^b$	780.9	2
<b>Zero-power threshold</b>	$T < T_c: M(T)=0$ $T > T_c: M(T)=a(T-T_c)^b$	<b>766.6</b>	<b>3</b>

$a$ , baseline failure rate;  $b$ , failure rate parameter;  $c$ , initial mortality rate;  $m$ , constant mortality rate;  $M$ , rate of failure;  $t$ , time;  $T$ , temperature;  $T_0$ , starting temperature;  $T_c$ , threshold temperature.

Akaike's information criterion (AIC) values and the number of parameters for each model are also included. The best model is in bold. Use of the Bayesian information criterion (BIC) gave qualitatively similar results.

constant functions. The second is whether or not we allowed a discontinuity in the failure function. This discontinuity would represent a threshold temperature  $T_c$  where the failure rate changes from a low, constant failure rate to an increasing function of temperature (Cossins and Bowler, 1987; Tang et al., 2007; Santos et al., 2011). If such a discontinuity exists, we fit the threshold temperature as one of the parameters. If not, we assume that failure rates increase from the lowest starting temperature. The third is, if a discontinuity in failure rates exists, we either assume zero failure below the threshold temperature or we fit a constant failure rate at temperatures below the threshold temperature. For comparison, we also consider two other models: a constant model, in which failure rate is constant and independent of temperature (a null model), and a constant-threshold model, in which failure rate is zero below some threshold temperature  $T_c$  and is constant for temperatures above  $T_c$  (Table 1). Note that the constant-threshold model directly reflects the idea of critical temperature as a 'threshold' temperature above which failure quickly occurs regardless of previous thermal exposure (Sunday et al., 2011; Buckley and Huey, 2016).

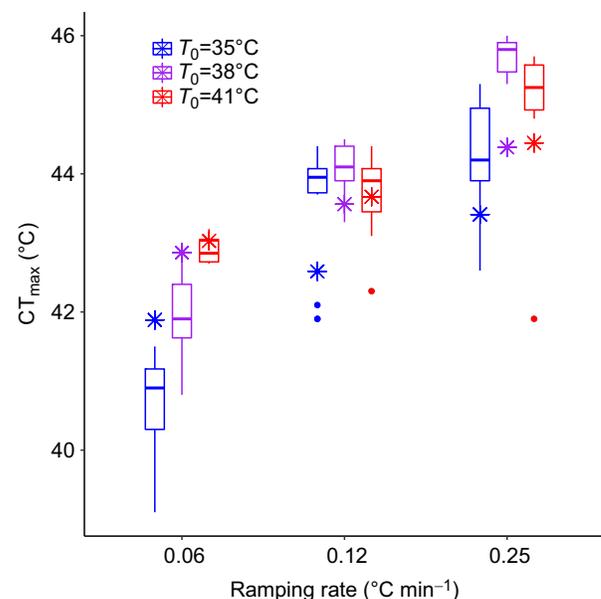
### Applying the model

We used data on  $CT_{max}$  of adult tsetse flies (*Glossina pallidipes*) from the experiments of Terblanche et al. (2007); John Terblanche generously provided the original data on time to failure from this study. The experiment included a full-factorial design with three starting temperatures and three ramping rates, and time to failure was measured for each individual. We fitted each of the failure rate models (Table 1) and used Akaike's information criterion (AIC) to select the best model.

The zero-power threshold model had the lowest AIC. The exponential threshold and exponential models also provided reasonable fits (Table 1). Use of the Bayesian information criterion (BIC) rather than AIC yielded qualitatively similar results. Based on the parameters estimated for these models, the predicted failure rates for these functions are quite similar except at temperatures below 38.5°C (Fig. 1), where failure rates are very low ( $<10^{-3} \text{ min}^{-1}$ ): in these non-stressful conditions, the expected time to failure is  $>1000$  min, far outside the range of the data (and the duration of the experiment). These results support the notion of a

threshold temperature for heat stress (Santos et al., 2011). By contrast, the constant and constant-threshold models had much higher AIC values (Table 1), and provided very poor fits to the data. For the constant-threshold model, the estimated threshold temperature is  $T_c=39^\circ\text{C}$ : in fact, the estimated  $T_c$  for this model will always be the minimum value of  $CT_{max}$  in the dataset under consideration, making this model unsuitable. This suggests that the interpretation of  $CT_{max}$  as a 'threshold' temperature (Sunday et al., 2011; Buckley and Huey, 2016) is problematic for these data.

The best-fit (zero-power threshold) model explains 58% of the total variance in  $CT_{max}$  in the dataset. The model correctly predicts that  $CT_{max}$  increases with increasing ramping rate and starting



**Fig. 3. Observed and predicted effects of ramping rate and starting temperature ( $T_0$ ) on  $CT_{max}$  for tsetse flies.** The different  $T_0$  (35, 38 and  $41^\circ\text{C}$ ) are represented by different colors. Boxplots are data from Terblanche et al. (2007); star symbols represent predictions from the best model for the data, the zero-power threshold model (see Table 1 and Fig. 1).

temperature (Fig. 3): much of the apparent effects of these methodological details can be directly accounted for in terms of the failure rate function. However, the predicted values from the model fail to account for all effects of ramping rate and starting temperature on mean  $CT_{max}$ . For example, at lower starting temperatures, predicted  $CT_{max}$  is greater than the mean  $CT_{max}$  at the slowest ramping rate, but the reverse is true at the higher ramping rate (Fig. 3). This interaction suggests that ramping rate and starting temperature may also have biological effects that alter the relationship between temperature and failure. In the next section, we extend our model to incorporate these potential effects.

### Extending the model: incorporating stress and acclimation

The models considered above (Eqns 1–9, Table 1) assume that failure rate is independent of prior thermal history: the current rate of failure depends only on the current temperature. However, abundant evidence shows that previous exposure to heat or cold (and other environmental conditions) can alter thermal tolerance, as a result of stress and acclimation (Lutterschmidt and Hutchison, 1997; Hofmann, 1999; Sørensen et al., 2003; Bowler, 2005). We will call these ‘time-dependent effects’, in which biological responses to current temperature depend on the pattern and duration of previous temperature exposure (Schulte et al., 2011; Kingsolver et al., 2015; Kingsolver and Woods, 2015). A natural way of incorporating time-dependent effects into our model is to allow the parameters of the failure rate function to depend on prior thermal history – for example, to depend on ramping rate during the experiment.

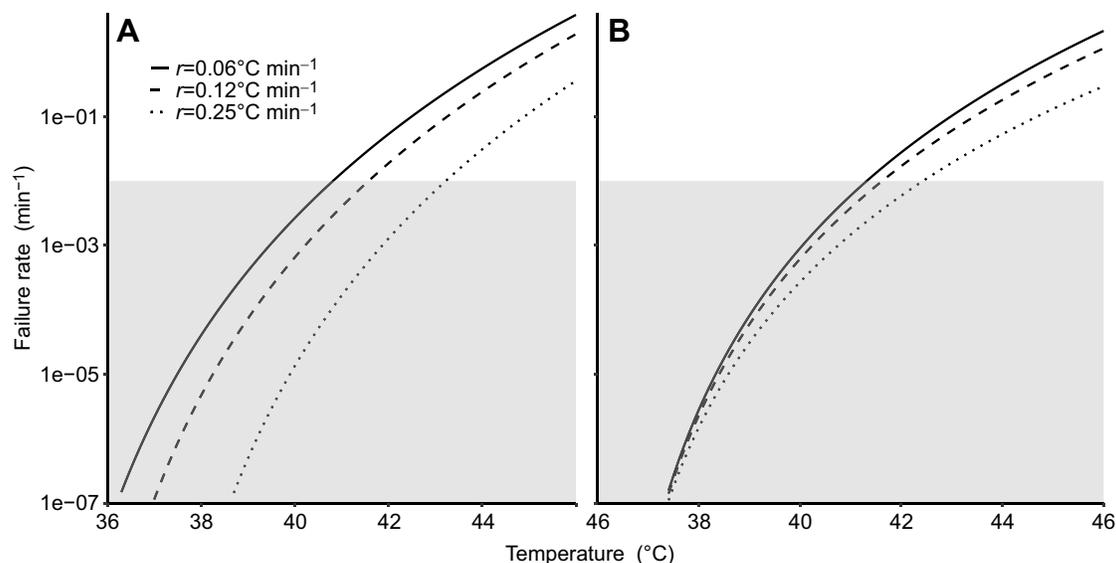
To illustrate this idea, let us consider the best model for the Terblanche et al. (2007) data, the zero-power threshold model (Table 1). The parameter  $T_c$  in this model represents the threshold temperature above which failure rate begins to accelerate (Fig. 4). Suppose that slow ramping rates increase the time an individual is exposed to warmer temperatures, depleting metabolic or water reserves and causing stress (Rezende et al., 2011, 2014; Santos et al., 2011). In this case, increasing the ramping rate will increase the parameter  $T_c$ , shifting the failure rate function to the right

(Fig. 4A). Alternatively, suppose that such stress responses increase the failure rate at higher but not lower temperatures (Fig. 4B). In this case, parameter  $b$  (the exponent of the function; Table 1) will decrease with increasing ramping rate. Acclimation and heat hardening will have opposite effects on the failure rate function: at slow ramping rates, temperatures experienced prior to failure induce heat shock or other acclimation responses that can improve tolerance to subsequent high temperatures. As a result of acclimation, increasing ramping rate will decrease parameter  $T_c$  or increase parameter  $b$ . In principle, both acclimation and stress responses could alter both  $T_c$  and  $b$  (and other model parameters); in practice, distinguishing among all possible models of interest may not be possible.

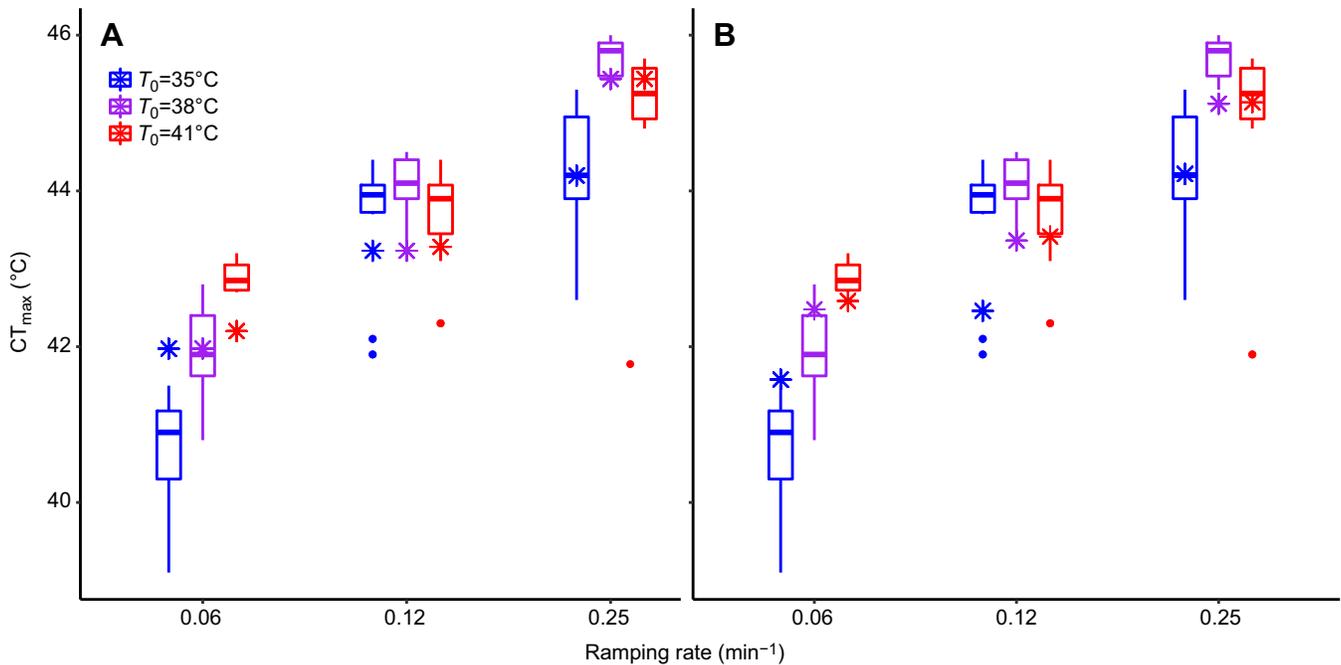
Here, we implement the simple case in which failure rate parameters  $T_c$  or  $b$  are linear functions of ramping rate, and apply these models to the tsetse fly data. Each of these models has a substantially (and significantly) lower AIC than the model in which  $T_c$  and  $b$  are constant ( $T_c$  and  $b$  constant: AIC=766.62, d.f.=3;  $T_c$  varies with ramping rate: AIC=724.02, d.f.=4, 69% of variance in  $CT_{max}$  explained;  $b$  varies with ramping rate: AIC=733.88, d.f.=4, 71% variance in  $CT_{max}$  explained), and improves the fit the models at low and high ramping rates (Fig. 5). The parameter values in the best model reveal that the threshold temperature  $T_c$  increases with increasing ramping rate. These results support the hypothesis that lower ramping rates may cause stress and decrease the temperature at which failure rate begins to accelerate (Figs 4A and 5A).

### Applying the extended model to stress and acclimation experiments

In the previous section, we quantified how ramping rate may alter failure rate functions, as a result of stress or acclimation. In most studies, only a single starting temperature and ramping rate are used, and the possible effects of experimental treatments on  $CT_{max}$  or  $CT_{min}$  are evaluated. Here, we illustrate how we can apply our model to data from such studies, and to estimate how experimental treatments alter the failure rate function.



**Fig. 4. Predicted effects of stress on failure rate function, via the effects of ramping rate on the parameters of the zero-power threshold failure rate function.** (A) Increasing ramping rate ( $r$ ) increases the threshold temperature ( $T_c$ ), shifting the failure rate function to the right. (B) Increasing ramping rate decreases the exponent of the failure rate function ( $b$ ). In both cases, stress responses at lower ramping rates increase failure rates at higher temperatures. The shaded region indicates failure rates below  $10^{-2} \text{ min}^{-1}$ ; in this region, the expected time to failure is greater than 100 min (longer than the duration of the experiments).



**Fig. 5. Observed and predicted effects of ramping rate and starting temperature ( $T_0$ ) on  $CT_{max}$  for tsetse flies, for models in which ramping rate affects the model parameters of the failure rate function.** See Fig. 4. (A) Ramping rate ( $r$ ) influences the threshold temperature ( $T_c$ ). (B) Ramping rate influences the exponent of the failure rate (parameter  $b$ ). Boxplots are data from Terblanche et al. (2007), while star symbols represent best-fit model predictions. The best model for the data is when increasing  $r$  increases  $T_c$ , as shown in A. See ‘Statistical models and results’.

We used data on the effects of a brief (2 h) heat shock (heat hardening) pre-treatment on  $CT_{max}$  of *Manduca sexta* larvae (Kingsolver et al., 2016). All  $CT_{max}$  measurements in that study were done with a starting temperature of  $38^\circ\text{C}$  and a ramping rate of  $0.25^\circ\text{C min}^{-1}$ . Standard linear models showed that increasing heat shock temperatures significantly increased  $CT_{max}$  (Kingsolver et al., 2016). Here, we used these data to estimate how heat shock temperature affects the parameters of the failure rate function. For simplicity (and in the absence of other information), we chose an exponential failure rate function with parameters  $a$  and  $b$  (Table 1, Fig. 1). We considered models in which either  $a$  or  $b$  is a linear function of heat shock temperature. Analysis of these models showed that allowing parameter  $b$  to be a function of heat shock temperature provides a substantially and significantly better model ( $\Delta\text{AIC}=8.0$ ;  $\chi^2=9.9$ ,  $P=0.0016$ , d.f.=1); the parameter estimates revealed that the slope (exponent) of the failure rate function declines with increasing heat shock temperature (Fig. 6). As a result, the model correctly predicts that temperature at failure ( $CT_{max}$ ) increases with increasing heat shock temperature (Fig. 6). By contrast, allowing parameter  $a$  to be a function of heat shock temperature results in a worse model ( $\Delta\text{AIC}=2.0$ ). These analyses suggest that acclimation due to heat hardening increases the temperature at which failure occurs by altering the shape ( $b$ ) but not the magnitude ( $a$ ) of the failure rate function – an insight that standard linear models of  $CT_{max}$  cannot provide. Analyses of TDT time curves at constant temperatures with *Drosophila* also suggested that acclimation altered the thermal sensitivity (shape) of the failure rate (Castañeda et al., 2015).

## DISCUSSION

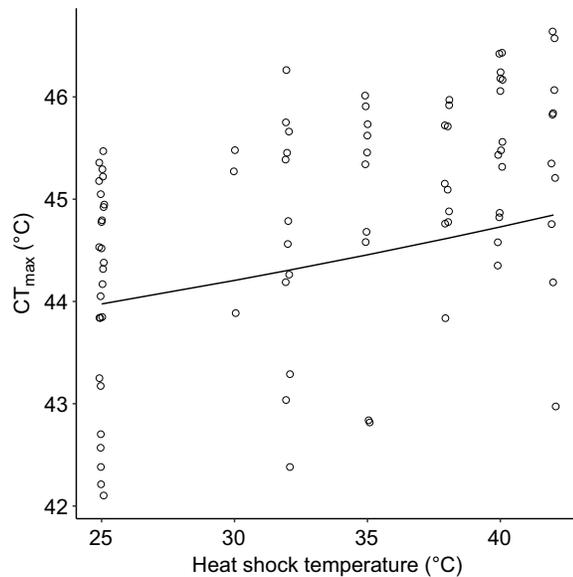
### Analyzing and interpreting critical temperatures

A central goal of our study was to explore the statistical and biological reasons why estimates of critical temperature may depend on methodological details in dynamic (ramping) experiments, and to

provide a statistical framework for quantifying these effects. The starting point for these analyses is the failure rate function, where the rate of failure at high temperature increases with increasing temperature, and is independent of the previous history of temperature exposure. (The logic is identical for low temperatures and  $CT_{min}$ .) During ramping, failure rate increases because temperature is increasing during the experiment. The time to failure (or temperature at failure,  $CT_{max}$ ) depends on the current failure rate given that failure has not yet occurred. As a consequence, the expected or mean value of  $CT_{max}$  will increase with ramping rate  $r$  and starting temperature  $T_0$  (Fig. 2). Several empirical studies have documented these predicted effects of  $r$  and  $T_0$  on  $CT_{max}$  or  $CT_{min}$  (Terblanche et al., 2007; Hoffmann, 2010) (but see Discussion below). Our simulations show that the rate of increase in  $CT_{max}$  decreases with increasing ramping rate (see also Rezende et al., 2014).

We emphasize that many of these effects are a straightforward statistical consequence of how critical temperature is operationally defined (time and temperature at failure) in ramping experiments, independent of stress, acclimation or other processes that reflect previous thermal history. As numerous authors have suggested (Hoffmann, 2010; Overgaard et al., 2011; Rezende et al., 2011, 2014; Terblanche et al., 2011), ramping rates, starting temperatures and other aspects of thermal history may indeed induce stress or acclimation responses that could alter subsequent responses to temperature (including the temperature at failure) (Lutterschmidt and Hutchison, 1997; Bowler, 2005; Loeschcke and Sørensen, 2005). For example, mean  $CT_{max}$  for codling moths was significantly greater at low ( $0.06^\circ\text{C min}^{-1}$ ) ramping rates than at higher rates, a response that cannot be accounted for by the statistical effects described here (Chidawanyika and Terblanche, 2011). A major goal of our model is to distinguish the statistical and biological effects in analyzing data on critical temperatures (see below).

Rezende et al. (2014) used the relationship between (constant) temperature and the expected time to failure  $t_f$  (see Eqn 2) to suggest



**Fig. 6. Effects of experimental (pre-treatment) heat shock temperature on  $CT_{\max}$  for *Manduca sexta* larvae.** Circles represent measured values of  $CT_{\max}$ ; the solid line is the predicted relationship from the best model for the failure rate function, in which increasing heat shock temperature decreased the exponent of the exponential failure rate function (see Discussion). Data from Kingsolver et al. (2016).

a regression model for analyzing thermal limits under constant temperature  $T$ , based on TDT curves:

$$\log_{10}(t_f) = (CT_{\max} - T)/z = (CT_{\max}/z) - (T/z), \quad (10a)$$

where  $CT_{\max}$  and  $z$  are parameters. Here,  $CT_{\max}$  is defined as the temperature at which the failure rate  $M=1 \text{ min}^{-1}$  (see Fig. 1) and  $z$  is a measure of thermal sensitivity. Rezende et al. (2014) used linear regression of  $\log_{10}(t_f)$  on experimental temperature (for a set of different experimental temperatures) to estimate the values of  $CT_{\max}$  and  $z$ . Re-arranging this equation, they also proposed a model for the 'knockout' temperature  $T_{\text{ko}}$ :

$$T_{\text{ko}} = CT_{\max} - z\log_{10}(t_f). \quad (10b)$$

It is important to note that time does not 'cause' temperature at failure to 'decline' below  $CT_{\max}$ , as implied by Eqn 10b: rather, the expected time to failure simply results from the constant mortality rate. As we describe in the associated datasets and code (<https://datadryad.org/resource/doi:10.5061/dryad.3f4s88q/1>; Kingsolver and Umbanhowar, 2018), the likelihood framework we develop here can readily incorporate the special case of constant temperatures (i.e. where ramping rate  $r=0$ ), and allows data on failure terms from constant temperatures and ramping experiments to be combined within a single analysis. Indeed, using data from ramping experiments to predict (and then test) results for constant temperature experiments (or vice versa) would be very useful for evaluating the time-dependent effects of stress and acclimation.

Santos et al. (2011) also identified that the Gompertz distribution describes the distribution of survival times of organisms with mortality rates that increase with age. Additionally, they incorporated two time-varying mortality mechanisms to predict how experimental methodology alters mortality times and temperatures. One of these is the temperature stress-dependent mortality rates that we model here, while the other is higher mortality due to metabolic loss of water, which we do not model. A major difference between their approach and our approach is that we focus on a parameter-estimating method

while eschewing highly detailed physiologically based models. Our demonstration of models with parameters  $T_c$  and  $b$  varying linearly with ramping rate provides support for a broad class of mechanisms including those modeled by Santos et al. (2011). In cases where the physiology is understood well, more biologically realistic models could be incorporated into our method and parameters could be estimated to test whether those models adequately fit experiments.

We have developed and implemented a likelihood model for analyzing data for critical temperatures from ramping experiments, which estimates the parameters of the failure rate function for a set of plausible failure functions. The model can be applied to either minimum or maximum critical temperatures, and can be readily modified to consider other failure functions. Our application of the model to the exemplary dataset of Terblanche et al. (2007) showed that the model explained 58% of the variation in  $CT_{\max}$ , including most of the effects of ramping rate and starting temperature (Fig. 3). The best failure function for these data was the zero-power threshold model, which includes a threshold temperature ( $T_c$ ) below which failure rate is zero, and above which failure rate accelerates (Table 1); for these data, the estimated  $T_c=36.0^\circ\text{C}$  (Fig. 1). We emphasize that  $T_c$  is not the same as  $CT_{\max}$ , and that mean  $CT_{\max}$  is likely well above the estimated  $T_c$  of Santos et al. (2011). In addition, the change in failure rate near  $T_c$  may be quite small: different failure rate functions may be very different at temperatures near  $T_c$  (Fig. 1).

In contrast, the constant-threshold failure function provides a much worse fit to the data; and the estimated threshold temperature from this model ( $T_c=39^\circ\text{C}$ ) is well below most of the measured values of  $CT_{\max}$  (Figs 1 and 3). In fact, the estimated  $T_c$  in the constant-threshold model will always be the minimum value of  $CT_{\max}$  in the data, rendering this model unsuitable. This result suggests that the interpretation and use of critical maximum (or minimum) temperatures as a threshold temperature above (or below) which failure occurs may be problematic (Santos et al., 2011; Rezende et al., 2014). Several recent studies have combined data for critical temperatures and environmental temperatures to predict the frequency of 'extreme' environmental conditions when critical temperature thresholds are exceeded (Sunday et al., 2011; Buckley and Huey, 2016; Kingsolver and Buckley, 2017). While these studies provide an important and useful first step, the use of critical temperatures as thresholds is unlikely to provide valid quantitative predictions about the consequences of extreme temperatures (Rezende et al., 2014).

The vast majority of published studies of critical temperatures use a single ramping rate ( $r$ ) and starting temperature ( $T_0$ ). Because  $r$  and  $T_0$  systematically influence measured values of critical temperatures, combining data on critical temperatures across studies using different methodological conditions is problematic. The failure rate function provides one possible solution to this problem. Suppose we choose a particular failure rate function  $M(T)$  (Table 1), and estimate the model parameters of this function for our data. Given the parameters, one can then compute the expected  $CT_{\max}$  for different ramping rate and starting temperatures. Converting  $CT_{\max}$  from different studies to a 'standard', common methodological condition would allow more direct comparisons among studies and facilitate their appropriate use in synthetic analyses of geographic patterns and climate change responses (Sunday et al., 2011; Buckley and Huey, 2016).

### Incorporating stress, acclimation and other effects

We have shown that estimates of  $CT_{\max}$  may be influenced by  $r$  and  $T_0$  in the absence of stress or acclimation effects. But of course prior thermal history (including  $r$  and  $T_0$ ) may additionally affect thermal

limits and thus  $CT_{max}$ . We have illustrated one natural way to model these effects, by allowing the parameters of the failure rate function to depend on ramping rate. This approach allows us to evaluate predictions about how stress or acclimation influences specific aspects of the failure rate function. Our analyses of the Terblanche et al. (2007) data using this approach show that allowing the threshold temperature ( $T_c$ ) or the acceleration of the failure function ( $b$ ) to depend on ramping rate substantially improves the model (e.g. increasing the explained variance from 58% to 71%): in particular, the threshold temperature increases with faster ramping rates, suggesting that stress responses may lower thermal tolerance at slow ramping rates (Fig. 5).

More generally, an extensive body of literature documents how experimental treatments, genetic differences and other factors can alter critical temperatures. In these studies,  $CT_{max}$  or  $CT_{min}$  is typically modeled as the response variable; here, we have illustrated how experimental factors can instead be analyzed in terms of their effects on failure rate. If inferences are limited to a single set of methodological conditions, the qualitative results of these two statistical approaches are likely to be similar. One advantage of the failure rate approach is that it enables us to evaluate more specific predictions about the consequences of stress or acclimation. For example, our analysis of *M. sexta* larvae indicates that increasing heat shock temperatures increases the critical temperature by reducing the exponent of the failure rate function (Fig. 6). Modeling failure rate functions allows us to test this and other issues of interest to thermal biologists, and to account for differences in the methodological conditions used within and across studies.

#### Acknowledgements

John Terblanche generously provided the original data from Terblanche et al. (2007) for our analyses. Ray Huey, John Terblanche and two anonymous reviewers provided valuable suggestions on previous versions of the manuscript. Research supported in part by NSF IOS-15-2767 to J.G.K.

#### Competing interests

The authors declare no competing or financial interests.

#### Author contributions

Conceptualization: J.G.K., J.U.; Methodology: J.G.K., J.U.; Software: J.U.; Validation: J.U.; Formal analysis: J.G.K., J.U.; Investigation: J.G.K., J.U.; Writing - original draft: J.G.K.; Writing - review & editing: J.G.K., J.U.; Visualization: J.U.; Funding acquisition: J.G.K.

#### Funding

This research was supported in part by National Science Foundation grant IOS-15-2767 to J.G.K.

#### Data availability

The datasets and R code needed to reproduce all of the results reported in this paper are available from the Dryad Digital Repository (Kingsolver and Umbanhowar, 2018): <https://datadryad.org/resource/doi:10.5061/dryad.3f4s88q/1>.

#### References

Angilletta, M. J. (2009). *Thermal Adaptation: A Theoretical and Empirical Synthesis*. Oxford: Oxford University Press.

- Bowler, K. (2005). Acclimation, heat shock and hardening. *J. Therm. Biol.* **30**, 125-130.
- Buckley, L. B. and Huey, R. B. (2016). Temperature extremes: geographic patterns, recent changes, and implications for organismal vulnerabilities. *Glob. Change Biol.* **22**, 3829-3842.
- Castañeda, L. E., Rezende, E. L. and Santos, E. S. (2015). Heat tolerance in *Drosophila subobscura* along a latitudinal gradient: Contrasting patterns between plastic and genetic responses. *Evolution* **69**, 2721-2734.
- Chidawanyika, F. and Terblanche, J. S. (2011). Rapid thermal responses and thermal tolerance in adult codling moth *Cydia pomonella* (Lepidoptera: Tortricidae). *J. Insect Physiol.* **57**, 108-177.
- Chown, S. L. and Terblanche, J. S. (2007). Physiological Diversity in Insects: Ecological and Evolutionary Contexts. *Adv. Insect Physiol.* **33**, 50-152.
- Cossins, A. R. and Bowler, K. (1987). *Temperature Biology of Animals*. New York: Chapman and Hall.
- Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*. Boca Raton, FL: Chapman and Hall.
- Hoffmann, A. A. (2010). Physiological climatic limits in *Drosophila*: patterns and implications. *J. Exp. Biol.* **213**, 870-880.
- Hofmann, G. E. (1999). Ecologically relevant variation in induction and function of heat shock proteins in marine organisms. *Am. Zool.* **39**, 889-900.
- Kingsolver, J. G. and Buckley, L. B. (2017). Quantifying thermal extremes and biological variation to predict evolutionary responses to changing and climate. *Phil. Trans. R. Soc.* **372**, pii: 20160147.
- Kingsolver, J. G. and Umbanhowar, J. (2018). Data from: The analysis and interpretation of critical temperatures. *Dryad Digital Repository*. doi:10.5061/dryad.3f4s88q.
- Kingsolver, J. G. and Woods, H. A. (2015). Beyond thermal performance curves: Modeling time-dependent effects of thermal stress on ectotherm growth rates. *Am. Nat.* **187**, 1-12.
- Kingsolver, J. G., Higgins, J. K. and Augustine, K. (2015). Fluctuating temperatures and ectotherm growth: distinguishing non-linear and time-dependent effects. *J. Exp. Biol.* **218**, 2218-2225.
- Kingsolver, J. G., MacLean, H. J., Goddin, S. B. and Augustine, K. E. (2016). Plasticity of upper thermal limits to acute and chronic temperature variation in *Manduca sexta* larvae. *J. Exp. Biol.* **219**, 1290-1294.
- Loeschcke, V. and Sørensen, J. G. (2005). Acclimation, heat shock and hardening—a response from evolutionary biology. *J. Therm. Biol.* **30**, 255-257.
- Lutterschmidt, W. I. and Hutchison, V. H. (1997). The critical thermal maximum: history and critique. *Can. J. Zool.* **75**, 1561-1574.
- Overgaard, J., Hoffmann, A. A. and Kristensen, T. N. (2011). Assessing population and environmental effects on thermal resistance in *Drosophila melanogaster* using ecologically relevant assays. *J. Therm. Biol.* **36**, 409-416.
- Rezende, E. L., Tejedo, M. and Santos, M. (2011). Estimating the adaptive potential of critical thermal limits: methodological problems and evolutionary implications. *Funct. Ecol.* **25**, 111-121.
- Rezende, E. L., Castañeda, L. E. and Santos, M. (2014). Tolerance landscapes in thermal ecology. *Funct. Ecol.* **28**, 799-809.
- Santos, M., Castañeda, L. E. and Rezende, E. L. (2011). Making sense of heat tolerance estimates in ectotherms: lessons from *Drosophila*. *Funct. Ecol.* **25**, 1169-1180.
- Schulte, P. M., Healy, T. M. and Fanguie, N. A. (2011). Thermal performance curves, phenotypic plasticity, and the time scales of temperature exposure. *Integr. Comp. Biol.* **51**, 691-702.
- Sørensen, J. G., Kristensen, T. N. and Loeschcke, V. (2003). The evolutionary and ecological role of heat shock proteins. *Ecol. Lett.* **6**, 1025-1037.
- Sunday, J. M., Bates, A. E. and Dulvy, N. K. (2011). Global analysis of thermal tolerance and latitude in ectotherms. *Proc. R. Soc. B* **278**, 1823-1830.
- Tang, J., Mitcham, E., Wang, E. and Lurie, S. (2007). *Heat Treatment for Postharvest Pest Control*. Trowbridge: Cromwell Press.
- Terblanche, J. S., Deere, J. A., Clusella-Trullas, S., Janion, C. and Chown, S. L. (2007). Critical thermal limits depend on methodological context. *Proc. R. Soc. B* **274**, 2935-2943.
- Terblanche, J. S., Hoffmann, A. A., Mitchell, K. A., Rako, L., le Roux, P. C. and Chown, S. L. (2011). Ecologically relevant measures of tolerance to potentially lethal temperatures. *J. Exp. Biol.* **214**, 3713-3725.
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.