Understanding uncertainty in temperature effects on vector-borne disease: a Bayesian approach

Leah R. Johnson,[1,2,14] Tal Ben-Horin,[3,4] Kevin D. Lafferty,[5,6] Amy McNally,[7] Erin Mordecai,[3,8] Krijn P. Paaijmans,[9] Samraat Pawar,[1,10] and Sadie J. Ryan[11,12,13]

[1]*Ecology and Evolution, University of Chicago, Chicago, Illinois 60637 USA*
[2]*Integrative Biology, University of South Florida, Tampa, Florida 33620 USA*
[3]*Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, California 93106 USA*
[4]*Marine and Coastal Sciences, Rutgers University, Piscataway Township, New Jersey 08854 USA*
[5]*Western Ecological Research Center, U.S. Geological Survey, Sacramento, California 95819 USA*
[6]*Marine Science Institute, University of California, Santa Barbara, California 93106 USA*
[7]*Geography Department, University of California, Santa Barbara, California 93106 USA*
[8]*Biology, University of North Carolina, Chapel Hill, North Carolina 27599 USA*
[9]*Barcelona Centre for International Health Research, Universitat de Barcelona, Barcelona, Spain*
[10]*Department of Life Sciences, Imperial College London, Silwood Park, Ascot, Berkshire, United Kingdom*
[11]*State University of New York College of Environmental Science and Forestry (SUNY-ESF), Syracuse, New York 13210 USA*
[12]*State University of New York Upstate Medical University (SUNY-UMU), Syracuse, New York 13210 USA*
[13]*School of Life Sciences College of Agriculture, Engineering, and Science, University of KwaZulu-Natal, Durban, South Africa*

Abstract. Extrinsic environmental factors influence the distribution and population dynamics of many organisms, including insects that are of concern for human health and agriculture. This is particularly true for vector-borne infectious diseases like malaria, which is a major source of morbidity and mortality in humans. Understanding the mechanistic links between environment and population processes for these diseases is key to predicting the consequences of climate change on transmission and for developing effective interventions. An important measure of the intensity of disease transmission is the reproductive number R_0. However, understanding the mechanisms linking R_0 and temperature, an environmental factor driving disease risk, can be challenging because the data available for parameterization are often poor. To address this, we show how a Bayesian approach can help identify critical uncertainties in components of R_0 and how this uncertainty is propagated into the estimate of R_0. Most notably, we find that different parameters dominate the uncertainty at different temperature regimes: bite rate from 15°C to 25°C; fecundity across all temperatures, but especially ~25–32°C; mortality from 20°C to 30°C; parasite development rate at ~15–16°C and again at ~33–35°C. Focusing empirical studies on these parameters and corresponding temperature ranges would be the most efficient way to improve estimates of R_0. While we focus on malaria, our methods apply to improving process-based models more generally, including epidemiological, physiological niche, and species distribution models.

Key words: Anopheles gambiae; basic reproductive number; Bayesian statistics; climate envelope; malaria; Plasmodium falciparum; sensitivity analysis; thermal physiology; uncertainty analysis.

Introduction

Malaria is a vector-borne disease that is a major source of illness and mortality in humans, especially in developing countries. Like many vector-borne diseases, the dynamics of malaria are greatly influenced by extrinsic environmental factors, such as temperature and rainfall. As climate changes over time, the distribution of both epidemic and endemic malaria will likely change as well, presenting new challenges for control. A better understanding of how the dynamics of malaria depend on environmental factors will be vital for understanding and planning for shifts in malaria incidence.

Various approaches have been used to try to understand the question of how environmental change is likely to impact the prevalence and distribution of malaria (reviewed in Guerra [2007] and Johnson et al. [2014]). Many of these models can be classified as niche or species distribution models, and they seek to link climate factors to observations of the prevalence of vectors, parasites, or disease occurrence. For mechanistic versions of these models (in contrast to geographical correlation models), it is necessary to understand how the vital rates of all players in disease transmission respond to the environment. Temperature strongly influences vital rates, particularly in ectotherms, and its effects can be measured under laboratory conditions. Despite this basic premise and our reasonable knowledge about thermal physiology, data on responses of

vital rates to temperature are not widely available. Even for species that have been well studied, like malarial parasites and their mosquito vectors (see Plate 1), the quality and quantity of the data are uneven across traits and temperatures. The paucity of data compromises the quality of model predictions, such as the range of temperatures that are conducive to disease transmission. Moreover, the sensitivity of model predictions to errors in empirical estimates is not well known. We develop methods for estimating sensitivity of model outputs to model inputs, focusing on the effect of temperature on malaria transmission.

An important and simple measure of transmission, the reproductive number, R_0, is often used in disease studies, as it is related to both how quickly a disease can spread in a naïve population and to the level of prevalence (the proportion of individuals that have been infected) for endemic diseases (Keeling and Rohani 2008). Recently, approaches have been developed to model how R_0 depends on temperature by incorporating thermal responses of traits underlying R_0, such as mosquito and parasite development rates (Mordecai et al. 2013, Molnár et al. 2013). For malaria, the method involves the use of laboratory data collected on the temperature dependence of all components of R_0 that depend upon parasite or vector physiology; temperature response curves fitted to each component are incorporated into the R_0 equation to find the overall thermal dependence of transmission. Including these physiologically based thermal responses produces predictions of transmission that are more in line with observed incidence patterns than previous models that do not incorporate these detailed physiological responses, even without accounting for rainfall (Mordecai et al. 2013). Although many other factors, such as key control measures, may better predict malaria morbidity and mortality, this kind of relatively simple modeling is a promising first step in prioritizing global health policy to respond to broad changes in the spread and intensification of infectious diseases (Altizer et al. 2013).

Empirical research on the factors or components that determine R_0 is costly. Thus, it is important to direct future research towards aspects that will result in the greatest reduction in uncertainty in R_0 overall, and thus improve our predictions of changes in future transmission the most. Currently, the laboratory data necessary to understand the temperature dependence of the components that determine the response of R_0 to temperature are often limited. Available data leave substantial uncertainty about the relationship between each component of R_0 and temperature, especially at the temperatures that are marginal for transmission. Thus, we anticipate considerable uncertainty in how R_0 varies with temperature. By better understanding all of the sources of uncertainty, we can prioritize laboratory studies more efficiently and design effective intervention strategies (Elderd et al. 2006, Merl et al. 2009). Mordecai et al. (2013) addressed this issue to some degree by performing a sensitivity analysis by perturbing the R_0 components with respect to temperature. However, this kind of simple, single-parameter local sensitivity analysis does not allow a full understanding of either the uncertainty in components or in R_0 overall. Further, additional data on components of R_0 for closely related species or less well controlled experiments are often available. These additional data, even if not ideal for fitting the final models directly, can be informative.

We use a Bayesian approach (Clark 2007) to understand the full range of uncertainty in the thermal response of malarial R_0. The focus of a Bayesian analysis is the posterior distribution, i.e., the probability that the parameters have some value given the data. This is obtained by combining a likelihood (the probability of observing the data given parameters with particular values) and a prior distribution (the assumed probability that the parameters have some values independent of the observed data) using Bayes rule. A full discussion of the Bayesian approach can be found elsewhere (e.g., Clark 2007). A Bayesian approach allows us to incorporate prior knowledge about the various components of R_0, for instance, by using data from related species in the inference procedure. This is especially useful in applications that rely heavily on sparse data, such as the one explored here.

We are interested in two primary aspects of the relationship between transmission and temperature. (1) Which temperatures prevent transmission? (2) Which temperatures promote transmission? Earlier work on temperature and disease transmission in general, and for malaria in particular, has produced mixed results, in part because the impact of temperature on preventing transmission (as opposed to promoting it) is often ignored (Hay et al. 2002, Gething et al. 2010, Rohr et al. 2011, Siraj et al. 2014). We use a Bayesian approach to explore the uncertainty and sensitivity of these two transmission outcomes, prevention and promotion, to mosquito and parasite traits.

We begin by introducing the R_0 model and its components, the potential thermal responses for all its components, and the available data on these thermal responses. We next introduce the data model, initial uninformative priors, and our overall methodology. We then step through a series of uncertainty and sensitivity analyses, together with the results for each analysis. This is followed by a discussion of how the approach taken compares to more classical analyses and the implications of the results.

Data, Models, and Methods

The standard model of malaria transmission by a vector is the Ross-Macdonald model (Macdonald 1952) from which the reproductive number R_0 is derived. R_0 determines the dynamical threshold for disease transmission and is defined as the average number of secondary infections caused by a single infected individual in an entirely susceptible population. It specifies the

TABLE 1. Component parameters of R_0 and their definitions.

Parameter	Definition	Functional form
a	bite rate (1/gonotrophic cycle length)	Brière
MDR	mosquito development rate	Brière
p_{EA}	egg to adult survival	quadratic
EFD	fecundity (eggs·female^{-1}·day^{-1})	quadratic
μ	mosquito mortality rate	quadratic
bc	vector competence	quadratic (and Brière)
PDR	parasite development rate (1/EIP)	Brière

Note: EIP is the extrinsic incubation period of the parasite.

relationships of parameters in the model that are required for an infection to spread within a population ($R_0 > 1$) as opposed to dying out ($R_0 < 1$). The most widely used formulation for malarial R_0 (Dietz 1993) is

$$R_0 = \sqrt{\frac{M}{Nr} \times \frac{a^2 \mathrm{bc} \exp(-\mu/\mathrm{PDR})}{\mu}} \qquad (1)$$

where M is the density of mosquitoes, a is the bite rate, bc is vector competence, μ is the mortality rate of adult mosquitoes, PDR is the parasite development rate (1/EIP, the extrinsic incubation period of the parasite), N is the human density, and r is the human recovery rate. Most of these model components are directly measurable or are closely related to quantities or traits that can be observed (Mordecai et al. 2013). Following Mordecai et al. (2013), we assume that the expected mosquito density is given by

$$M = \frac{\mathrm{EFD} p_{EA} \mathrm{MDR}}{\mu^2} \qquad (2)$$

where EFD is number of eggs produced per female per day, p_{EA} is the probability that an egg will hatch and the larvae will survive to the adult stage, and MDR is the mosquito development rate. The parameters that jointly define R_0 and M are summarized in Table 1, and we refer to these as components of R_0.

Virtually all physiological traits in ectotherms exhibit unimodal temperature responses, i.e., they have an optimal temperature at which the trait is maximized and declines on either side (e.g., Angilletta 2009, Dell et al. 2011, Amarasekare and Savage 2012). However, the exact functional form of the unimodal response is still under debate, especially because it is known to vary with the type of trait (Dell et al. 2011, Mordecai et al. 2013). Therefore, as in Mordecai et al. (2013), we determined the appropriate thermal response model for each component trait by fitting candidate functional forms: quadratic for symmetric responses and Brière for asymmetric (see *Data, models, and methods: Approach, likelihoods, and priors: Likelihoods* and Fig. 1). These were chosen as they are among the simplest functional forms that exhibit the desired unimodal behavior.

All analyses were conducted in R (R Development Core Team 2008) with Markov chain Monte Carlo (MCMC) implemented in rjags/JAGS (Just Another

Gibbs Sampler, [Plummer 2003, 2013]). Computer code for all analyses are included in the Appendices.

Data

We use two sets of data in our analysis: the main data set that contains the focal data for the thermal responses for the components that make up R_0 (Table 1) and a prior data set used to elicit priors for our Bayesian analysis (both sets included in the Appendices; for sources see Appendix A: Table 1). Ideally, our main data set would exclusively comprise laboratory data on *Plasmodium falciparum*, the causative agent of the majority of tropical malaria, and its primary vector *Anopheles gambiae*, held at constant temperature for all components. These data were available for only three traits: mosquito development rate (MDR), egg to adult survival (p_{EA}), and adult mosquito mortality rate (μ). For other traits ideal data are unavailable (Mordecai et al. 2013), and instead we used data from related species collected under appropriate laboratory conditions. More specifically, we prioritized data collected in the laboratory at constant temperatures. For mosquito and parasite traits, we prioritized based on the relative to efficacy of transmission and severity of disease in humans. Thus, for mosquito traits, we prioritized data for *An. gambiae*, followed by other anophelene species, and finally for *Aedes* species. For parasite traits, *P. falciparum* was prioritized, followed by *P. vivax*. When possible, only a single mosquito or parasite species was used for an individual trait in the main data. For our prior data set, our conditions for inclusion were more flexible. Although data on related species held in the lab at constant temperature were preferred, we also allowed more distantly related species or less controlled (variable temperature) experiments. These data are distinct from the main data set and were used, along with expert opinion, to elicit informative priors for the parameters of the unimodal temperature responses for each R_0 component.

Approach, likelihoods, and priors

We fitted the thermal response of each component of R_0 (Table 1) to independent data using a Bayesian approach to obtain the posterior distribution for the parameters that describe each response (and thus the posterior distribution of the response itself), as well as for R_0 overall. Inference in the Bayesian framework
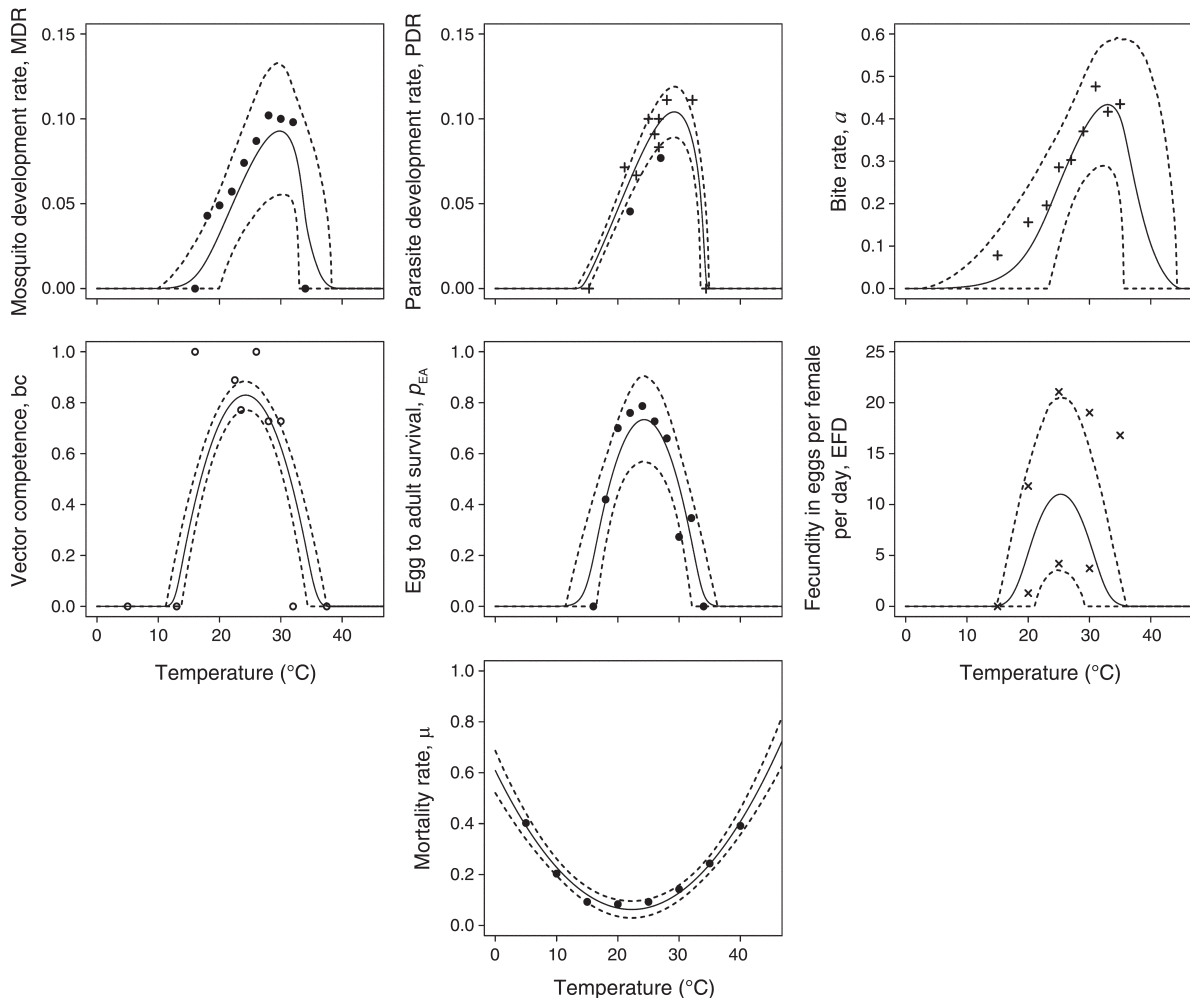
Fig. 1.   Posterior mean (solid line) and 95% highest posterior density (HPD) (dashed lines) of the thermal responses for all components of R_0, with informative priors together with the main data. Traits modeled with a Brière thermal response ($cT(T - T_0)\sqrt{(T_m - T)}$) are grouped in the top row, concave-down quadratic ($a(T - T_0)(T - T_m)$) in the middle row, and concave-up quadratic ($aT^2 + bT + c$) in the bottom row, where T is temperature, T_0 is the lower temperature limit of response and T_m is the upper limit. R_0 is the basic reproductive number, and a, b, and c are constants that determine the shape of the unimodal responses in the quadratic or Briere functions. Data symbols correspond to the species of mosquito or parasite used for the analysis. Solid dot is *Anopheles gambiae* or *Plasmodium falciparum* in *An. gambiae*; + is other anopheline species or *P. falciparum* in other anopheline species; × is *Aedes* species; and open dot is *Plasmodium vivax* in other anopheline species.

proceeds in three steps. First, a likelihood is defined for each type of data. Second, appropriate prior distributions are determined. Third, samples from the posterior distribution of the parameters, given the data, are obtained via MCMC. We used this procedure first for the prior data and for the main data, assuming uninformative priors in both cases (see *Data, Models, and Methods: Approach, likelihoods, and priors: Priors* and Appendix A: A.2). Next, the posterior distributions obtained from analyzing the prior data were used to build informative priors, and the inference procedure was repeated for the main data using these informative priors. We then compared the resulting posterior distributions obtained using the uninformative and informative priors. Further, we calculated R_0 with both

sets of results and compared them. This gave an indication of the sensitivity of the individual components and of R_0 to the choice of prior. We followed this with further sensitivity and uncertainty analyses (Appendix A: Table A.3).

Likelihoods.—We assumed functional forms for each component based on our previous work (Mordecai et al. 2013) and on the types of functional forms (unimodal and frequently asymmetric) that are typical for similar traits in other arthropod species (Dell et al. 2011). More specifically, we used either a quadratic (symmetric) or Brière (asymmetric) function, depending on the component. At any given temperature, the mean response should be determined by this functional form. Further, all model components are, by definition, greater than or
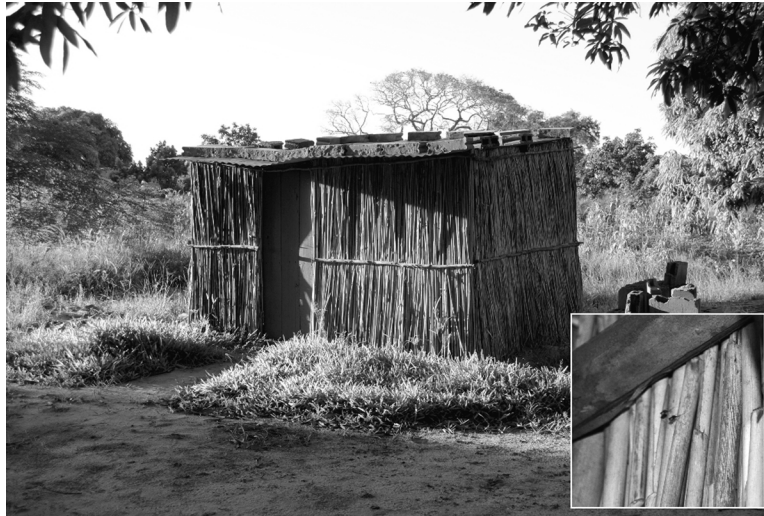
PLATE 1. Typical house in southern Mozambique, Africa, where malaria is endemic. (Inset) Blood-fed *Anopheles funestus* mosquitos rest indoors on a bamboo reed wall. Photo credits: K. Paaijmans.

equal to zero. Thus, we chose to use a truncated normal distribution with mean parameter (usually denoted by μ) given by the appropriate functional form (i.e., Brière or quadratic), as the likelihood for the data for most of the components of R_0. For all of the components we examined, the lower truncation limit was at zero. Most components can take any value greater or equal to zero. Thus, for most of our data $Y_i \mid 0 < Y_i < b \overset{\text{iid}}{\sim} N(\mu = f(T_i), \sigma^2)$, where Y_i is the data point, b is the upper truncation limit (either 0 or ∞), $f(T_i)$ is the Brière or quadratic temperature response, iid indicates that the data are independently and identically distributed, and σ^2 is the variance of the normal distribution. However, two components, the vector competence (bc) and the egg to adult survival (p_{EA}), are probabilities and are thus constrained to be between zero and one. For these components, we would ideally have the actual numbers of successes and total numbers of observations so that the more appropriate binomial model could be used. These data were indeed available for vector competence and so were modeled with a binomial likelihood, i.e., $Y_i \overset{\text{iid}}{\sim} \text{Bin}(n, p = f(T_i))$, where n is the number of total observations, of which Y were successes, and the probability, p, of a success at a particular temperature, $f(T_i)$, is either Brière or quadratic. For egg to adult survival the raw data were not available, so we used a normal distribution truncated at zero and one to model the proportion of eggs that successfully mature to the adult stage. This choice keeps calculations simple, allows straightforward implementation of biologically based priors, and has shape properties that are more appropriate for these data than alternatives such as a beta distribution.

Priors.—We began by defining a set of default priors for all parameters that are chosen to be relatively uninformative. That is, these priors were designed to constrain parameter values to be biologically reasonable, but to otherwise provide wide, reasonably even support across potential parameter values. In particular, for our default priors, we assumed that the maximum temperature at which a unimodal, hump-shaped component goes to zero is 45°C and the minimum temperature should be 0°C, as these temperatures are generally lethal to mosquitoes. This upper limit is slightly higher than some observed upper lethal limits, which are closer to 40°C (Bayoh 2001, Bayoh and Lindsay 2003, Lardeux et al. 2008). We chose the higher conservative limit to allow a broader range of temperatures for which the data could inform the posterior distributions. Each of the concave-down (or hump-shaped) curves have a parameter that describes the temperatures at which trait goes to zero, notated as T_0 and T_m for the lower (0) and upper (m) limits, respectively. Since we required $T_0 < T_m$, we specified nonoverlapping priors for these parameters. For the concave-up quadratic, we chose priors that limited the quadratic curves to those that are concave up and in the appropriate quadrant. We set the priors on other parameters (including the precision parameter, $\tau = 1/\sigma$, in the normal distribution) to be diffuse, i.e., to have wide support. Details can be found in Appendix A: Tables A.2 and A.3. In all cases, we examined the sensitivity of the posterior distributions to the priors.

Uncertainty and Sensitivity Analyses and Results

Our uncertainty and sensitivity analyses consisted of multiple parts. First, we addressed sources of uncertainty in our analysis to understand the expected response of R_0 and its components to temperature and the range of responses that are supported by data. This is similar to global sensitivity analysis for the components and R_0. Our measure of uncertainty for each analysis is the 95%

highest probability density (HPD) interval which gives the range of a parameter or response corresponding to a central area containing 95% of the probability. Second, we compared how the uncertainty in R_0 overall depends on the uncertainty in its components using a variant of local sensitivity analysis and comparing the results to those obtained for the global-style analysis. Third, we addressed how sensitive R_0 is to temperature and to its components, as well as the uncertainty in these relationships. Here, sensitivity is the amount by which R_0 changes when temperature changes and is given by the derivative of R_0 with respect to T. As with R_0 itself, the uncertainty in this sensitivity is expressed in terms of the 95% HPD intervals. How the Bayesian and classical approaches to these analyses compare is addressed further in the discussion.

Uncertainty in the components of R_0

How uncertain are the responses of the mosquito and parasite traits to temperature, and how does this depend on the prior information included in the analysis?

To answer this question, we made three sets of comparisons. First, we qualitatively examined the impact of adjusting the default priors for each parameter on the inference for individual components (with both the main and prior data). Second, we compared the posterior distributions for individual components of the main data set obtained with default and informative priors. In this case, the informative priors are generated from the posterior distributions obtained from fitting the prior data to elicit informative priors. Third, we examined the impact of using an alternative functional response for the vector competence term, for which there is both relatively little data and little a priori support for a particular functional form.

The impact of including the informative priors varied between components, in some cases decreasing uncertainty (smaller HPD intervals around the mean, such as bc, p_{EA}, EFD, PDR), sometimes increasing uncertainty (larger HPD intervals around the mean, such as a, MDR), or having little impact on the posterior (no change in HPD intervals, μ; see Appendix B: B.1). Across components, modifying priors on the lower and upper temperature limits of the responses (T_0 and T_m, respectively) had the greatest impact on the posterior distributions of the temperature responses overall. The upper limit of 45°C on T_m for the default and some of the informative priors was important for components for which no high temperature data were available, such as the bite rate, a. Full results for each component, including inferences with both types of priors and a comparison of the marginal posterior distributions of parameters with their priors, are included in Appendix B.

In Fig. 1, we show the posterior mean and 95% HPD interval around this mean (summarizing the extent of our uncertainty around this response) for all the components when informative priors were used. Some interesting patterns emerged when we compared across components. First, for all components modeled with a Brière function, the lower temperature limit (the temperature below which the trait is zero) was less certain than the upper temperature limit (although this difference was small for PDR). This was partly due to the nature of the functional form, i.e., it goes to zero more quickly at high temperatures than it does at low temperatures. However, it also reflected that there were often fewer data available across lower temperatures than upper temperatures in the main and prior data together. This pattern of uncertainty at the limits was not found for the concave-down quadratic responses. Instead, in some cases, the upper limit was less certain than the lower. This indicates that the temperature resolution for experiments needed to pin down the responses may depend on the type of response (asymmetric vs. symmetric) that a trait exhibits.

For most of the components explored, either the data gave a strong indication of whether a symmetric or asymmetric response was appropriate (Mordecai et al. 2013) or there were biophysical reasons why we expected a response to be asymmetric or symmetric a priori (Angilletta 2009, Dell et al. 2011). However, vector competence (a compound trait) was ambiguous. Thus, we fit both a quadratic and Brière function for this component. Both fit quite well, and thus the impact of fits using both functional forms on the uncertainty in R_0 was addressed in the subsequent analysis.

Overall uncertainty in R_0

How uncertain is the response of the basic respoductive rate, R_0, to temperature (due to uncertainty in all components), and how does this depend on the prior information included in the analysis?

To answer this question we made three comparisons. First, we compared the posterior distributions of R_0 under default and informative priors for the components, looking at the overall uncertainty (95% HPD interval) of the full response curve when all components were allowed to vary according to their posterior distributions. Second, we examined the HPD intervals of three important summaries of R_0: minimum (low temperature transmission limit), maximum (high temperature transmission limit), and peak (temperature at maximal transmission) R_0. Third, we examined the impact of the two functional responses for the vector competence term on the posterior distribution of R_0. This analysis shows the overall uncertainty around which temperatures prevent transmission (low and high temperature transmission limits) and which temperatures promote transmission (peak temperature).

In Fig. 2, we show the posterior mean of the temperature dependence of R_0 and 95% HPD intervals of the temperature response of R_0 when both informative and uninformative priors are used. All curves are scaled to the maximum value of the mean $R_0(T)$ curve. These are generated using posterior samples from all components and so indicate the overall uncertainty in
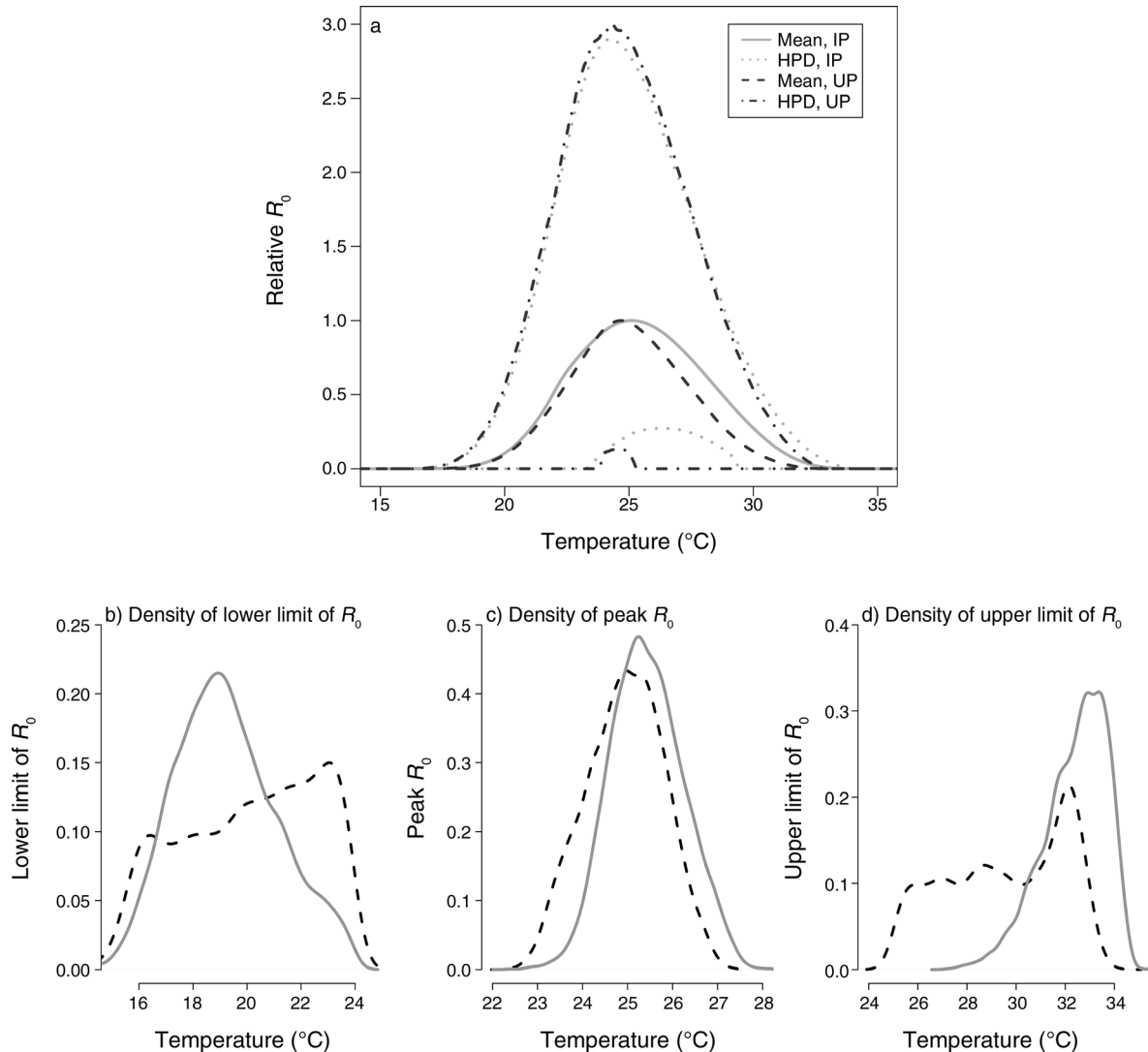
Fig. 2. (a) Relative R_0 (R_0 divided by the maximum value of the posterior mean) assuming a quadratic function for vector competence, where UP means uninformative priors, IP means informative priors, and HPD means highest probability density. Smoothed posterior distributions of the (b) lower temperature limit of R_0, (c) peak temperature of R_0, and (d) upper temperature limit of R_0 all assuming a quadratic function for vector competence. Case with uninformative prior is shown as a dashed line and with informative prior as a solid line.

the response curve due to uncertainty in all components, simultaneously. Notice that the mean R_0 curves obtained using default and informative priors are very similar. Further, the upper 95% HPD intervals were nearly identical. However, the lower HPD intervals of R_0, especially at higher temperatures, differ considerably as the additional prior information allowed us to pin down the high temperature transmission limit more precisely. This can be seen more clearly by looking at the posterior distributions of the upper and lower temperature limits of R_0 and the distribution of the temperature at peak R_0 (Fig. 2). With the default priors, almost the full range of possibilities for the lower (from 0°C to 24°C) and upper (from 25°C to 45°C) limits were equally represented. Adding in prior information indicates

support for a slightly lower temperature limit to malaria transmission, while the upper limit is at a slightly higher temperature than was predicted with default priors. In other words, the climate envelope where transmission may be possible is slightly larger than would be inferred without prior data. Further, our estimates of the temperatures that can exclude malaria, particularly at the upper end, are more precise. However, the prediction of the temperature of peak transmission, which corresponds to temperatures at which malaria is expected to be most severe and difficult to control, was robust.

As mentioned in the previous section, the most appropriate functional response to describe the temperature dependence of vector competence (bc, a compound trait) was ambiguous. Since both functional

forms fit the available data well, we examined how using each impacted the posterior inferences for R_0. To do this, we calculated R_0 using first the posterior samples for the Brière response for bc and then the quadratic. All other components of R_0 were allowed to take all possible values of their posterior distributions. Thus, our comparison shows how the uncertainty in the functional form for bc impacts the overall uncertainty in R_0 given the full uncertainty in the other parameters. The choice of the functional form for bc had little impact on the posterior distribution of $R_0(T)$ except at the high temperature limit (Appendix B). As with other components examined, the vector competence fit with a Brière function exhibited reduced uncertainty in the upper limit compared to the quadratic fit, and as a result, decreased the uncertainty in R_0 at this limit. Since there is no a priori reason to prefer one or the other of these, in all further analyses we assumed the quadratic fit, as this resulted in the most uncertainty in R_0 at the upper limit and was thus a more conservative choice.

Uncertainty in R_0 and its sensitivity to temperature, by component

Which mosquito and parasite traits drive the uncertainty in R_0 across temperatures? To answer this question, we used a variation of a traditional sensitivity analysis. We set all but a focal component to its posterior mean. Then, the posterior samples of the focal component were used to calculate the width of the 95% HPD at each temperature due to only the variation in this single component. We then normalized this to the width of the 95% HPD for the full posterior of R_0 (i.e., when all components were allowed to vary) to approximate the proportion of the uncertainty each component contributed to the full uncertainty in R_0.

In Fig. 3, we show the amount of uncertainty in R_0 due to a single component compared to the uncertainty overall. The adult mosquito mortality rate μ dominated at intermediate temperatures, i.e., the region in which temperature promotes transmission, because $R_0 \propto 1/\mu^3$ (by definition from Eqs. 1 and 2), and at intermediate temperatures the rate of adult mosquito mortality is low. More surprising was the relatively narrow range of temperatures where it dominated: μ dominated in the middle third of the transmission range and primarily determines the height and location of peak R_0. This explains why the informative priors had so little impact on the upper HPD interval for R_0 (Fig. 2): the posterior of μ was not impacted by additional prior information, and this component determined the height of the curve. In contrast, the uncertainty on the lower and upper limits for transmission was driven by bite rate and fecundity, respectively. In particular, uncertainty in the bite rate (a) drove uncertainty in the lower temperature boundary, while the fecundity (EFD) was most important at the upper temperature limit. These two components were also the most important sources of uncertainty, after μ, at intermediate temperatures. Other

components, such as vector competence (bc), contributed very little to the overall uncertainty despite the fact that the data on these components were sparse. This was primarily because parameters like this are probabilities, bounded between zero and one, and only impact R_0 proportionally. Other parameters have the potential to vary over orders of magnitude, and thus swamp the impact of these parameters over most of the range of R_0. Only when the magnitudes of the other parameters are small, such as the temperature extremes, can these contribute significantly to the overall uncertainty.

Which mosquito and parasite traits determine the sensitivity of R_0 to temperature and the uncertainty in the sensitivity across temperatures? We examined how sensitive R_0 is to changes in temperature, that is, how various components contribute to the change in R_0 with temperature, as measured by the derivative of R_0 with respect to temperature. We especially focused on the uncertainty of the sensitivity, measured by the HPD interval around the mean sensitivity, and which components drive the uncertainty. We focused on these as they indicate the data that can best help improve our understanding of what determines the shape of R_0 across temperatures. We started with a standard sensitivity analysis, calculating the derivative of R_0 with respect to temperature overall and for each component separately, $(dR_0/dT)_\theta$, where θ denotes a parameter in the set of those listed in Table 1. Thus $(dR_0/dT)_\theta$ is the change in R_0 with respect to temperature due to the temperature response of a single parameter alone (see Appendix A: Table A.5 for equations). As with the previous analysis, we then set all components, save a focal component to its mean, and used the posterior samples from that focal component to obtain the marginal posterior distribution for the sensitivity. We repeated this for all components, then examined the median sensitivity and the width of the 95% HPD intervals for the component-wise sensitivity, $(dR_0/dT)_\theta$, relative to the overall sensitivity, $(dR_0/dT)_\theta$ (calculated using the posterior samples from all the components at once) to see which components were driving the sensitivity of R_0 to temperature and the uncertainty in the sensitivity.

In Fig. 3, we show the uncertainty in the sensitivity of R_0 to temperature by each component scaled by the overall uncertainty in the sensitivity of R_0 to temperature. For instance, the median sensitivities, by component, indicate that at low temperatures R_0 was very sensitive to the bite rate and at high temperatures (even the very highest) the mortality rate drove the response of R_0 to temperature (Appendix C). However, at very high and very low temperatures, other components besides bite rate and mortality began to be important as well, and these other components were less certain. Thus, at the temperature extremes determining where transmission is not possible, the uncertainty in how sensitive R_0 is to temperature was driven by other components, such
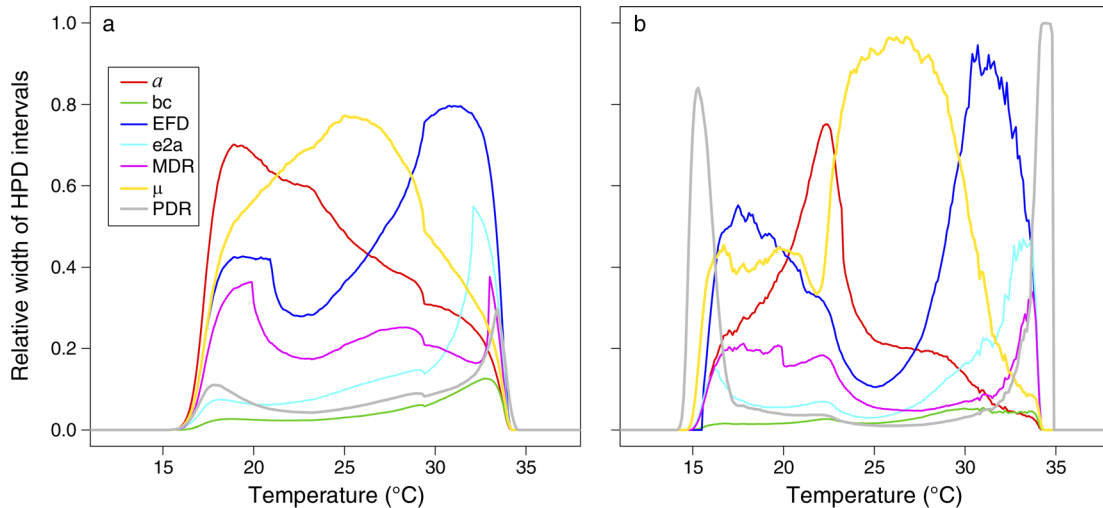
FIG. 3. (a) Relative width of the 95% HPD intervals due to uncertainty in each component, compared to uncertainty in R_0 overall. Each curve was obtained as followed. For each component, R_0 was calculated for the thinned posterior samples of that component, with all other components set to its posterior mean. Then the width of the inner 95% HPD was calculated at each temperature. This was then normalized to the width of the HPD of the full posterior distribution of R_0 at each temperature. (b) Relative width of the 95% HPD in $(dR_0/dT)_\theta$ scaled by the width of the 95% HPD for $(dR_0/dT)_\theta$ at each temperature, calculated as in panel (a). In both, a quadratic response for vector competence (bc) was used. The term e2a refers to egg to adult survival (p_{EA} elsewhere in the text).

as fecundity (EFD) and the parasite development rate (PDR).

DISCUSSION

Using a Bayesian approach, we identified component traits that were the main sources of uncertainty in R_0 and in how sensitive R_0 is to temperature. Overall, uncertainty about the temperature limits on transmission was greater than the uncertainty in the optimal temperature for transmission. We found that much of the uncertainty in R_0 was due to adult mosquito mortality, μ, as one would expect given R_0's nonlinear dependence on mortality. This contribution was focused in the region of temperatures that promote transmission, where μ and its uncertainty were small. Other components determined the uncertainty near the temperature limits of R_0 in terms of the relative width of the HPD intervals. In particular, near the low temperature limit, the uncertainty in R_0 was largely due to uncertainty in the bite rate, a, whereas near the high temperature limit, the uncertainty was primarily due to uncertainty in fecundity (EFD). Fecundity also contributed a relatively large amount of uncertainty across temperatures. The uncertainty in the high temperature limit itself was determined primarily by the parasite development rate (PDR).

The most important empirical data needed to improve model certainty depends on the goals. To resolve the uncertainty in the temperature optimum, empirical studies should focus on measuring adult mosquito mortality rate from 20°C to 30°C, as well as bite rate and fecundity. By contrast, resolving uncertainty sur-

rounding the lower and upper temperature limits on transmission requires measuring: (1) bite rate from 15°C to 25°C, (2) fecundity across temperatures, but especially ~25–32°C, and (3) PDR at ~15–16°C and 33–35°C. The last is especially interesting, because PDR is not driving the overall uncertainty at any temperature. Instead, our uncertainty in how sensitive R_0 is to temperature depends on PDR at both temperature extremes. More specifically, our uncertainty in the temperature at which R_0 changes from zero is driven by PDR. This also suggests that PDR could be determining the temperature limits for malaria transmission and the ability of the parasite to evolve at the edges of its thermal limits could determine where malaria could occur. Resolving the temperature limits is particularly important given that warming is expected to expand transmission into currently or recently unsuitable highland areas (Siraj et al. 2014) and may force currently warm, suitable lowland areas above the upper transmission limit. Differences in the presumed thermal limits that inhibit transmission can impact predictions of when and where transmission is likely to occur both now and in the future (Ryan et al., *in press*). A better understanding of the uncertainty in the temperatures that inhibit transmission should help inform policy priorities as climate changes.

Further, our results provide guidance as to which components may not be as high priority for further work. For instance, our analysis indicates which components are contributing relatively little to the uncertainty in R_0, such as vector competence (bc). Although these components are necessary for transmis-

sion (incompetent vectors cannot transmit disease, for example), investment in reducing uncertainty in these components will have a comparatively small impact on our overall understanding of R_0. Thus, these components could be given a lower priority for further empirical effort, especially if these components are difficult or expensive to measure.

Our method is unusual in that it combines concepts and approaches from traditional global and local sensitivity analyses with a Bayesian method of inference. Bayesian analyses often focus on the inference and uncertainty aspects, and, for simpler models, model and variable selection. If sensitivity is addressed, it is typically the sensitivity of the posterior to the prior specification. Conversely, most sensitivity analyses would typically be conducted using parameter values selected at random or evenly over some reasonable range of parameters for a system instead of performing parameter inference as part of the procedure. This traditional approach would be most similar to performing our analysis using samples from informative priors. The choice to take the extra step and fit data before doing the uncertainty and sensitivity analyses depends on whether or not appropriate, high quality data are available for the analysis. If this is not the case, then incorporating additional data could lead to biased results. However, even moderately good data, such as we have in our example, allowed us to narrow down the regions of parameter space that are reasonable for our system by allowing the data to inform us of how the parameters that describe our thermal response curves are correlated for each component of R_0, which is not easy to obtain a priori.

On the other hand, in this example, we have been forced to use data from alternative species, even for the focal data. For instance, we used fecundity data for *Aedes albopictus* for the focal data, even though we expect these are likely to be different from any *Anopheles* species. We also used *Plasmodium vivax* data for vector competence. Thus, we may have underestimated the overall uncertainty in R_0 and the uncertainty due to these components. It is unclear how to explicitly incorporate this into the current (or any other) framework to quantify the impact of using these data. However, the types of sensitivity analyses we conducted take into account the structure of the model and give information about how much each component could potentially contribute to the overall uncertainty. We can, therefore, use these analyses to complement our intuition about how to focus research efforts. Thus, although fecundity for Anophelene species and vector competence for *P. falciparum* could be useful (since we had to use alternative species for these analyses), when combined with the rest of the analyses, we would prioritize fecundity as uncertainty in that component has a bigger impact on overall uncertainty. Further, we might expect that biological reasons may explain uncertainty in some components that contribute to R_0,

for instance because there is more individual variability in some traits or local adaptation (Sternberg and Thomas 2014). With appropriate data we could explicitly include individual- level (or population-level) variation as part of the analysis or combine data across different populations, while allowing differences in thermal response curves within the framework.

Although we applied our methods to a particular formulation of R_0 for malaria, the approach is appropriate for mechanistic models of other systems, including other vector-borne disease and species distribution models. To best understand the full uncertainty for a particular disease or species distribution, consideration of multiple mechanistic models and environmental drivers would allow researchers to further understand model uncertainty and the robustness of predictions to other formulations. For instance, alternative dynamical models could change the estimate of the peak and range of species distributions, as the functional relationship between various traits or vital rates and population performance could be substantially different. This is an issue that is largely ignored in the literature addressing disease incidence, species distributions, and climate envelopes. This source of model uncertainty needs to be addressed to fully quantify uncertainty in the response of populations to climate change. We suspect that for many species and diseases, including malaria, the uncertainty inherent in the individual components would often swamp model uncertainty in big picture quantities like R_0. This is especially true when data from a focal species are not available at all and further emphasizes the importance of high quality data for parameterizing these kinds of models. However, alternative formulations incorporate each component in different ways, so the conclusions about which components drive uncertainty are likely to be less robust. Consideration of multiple models can indicate what data acquisition should be prioritized, for instance if it is important across a variety of model formulations, and further, which new data would allow better discrimination between competing models. The Bayesian approach allows direct comparison of models and their uncertainty. Thus, it has the potential to be a useful tool for identifying concrete recommendations for future research to improve predictions of how factors such as climate change could impact the distribution of malaria and other vector-borne diseases.

Literature Cited

Altizer, S., R. S. Ostfeld, P. T. Johnson, S. Kutz, and C. D. Harvell. 2013. Climate change and infectious diseases: from evidence to a predictive framework. Science 341(6145):514–519.

Amarasekare, P., and V. Savage. 2012. A framework for elucidating the temperature dependence of fitness. American Naturalist 179(2):178–191.

Angilletta, M. J. 2009. Thermal adaptation: a theoretical and empirical synthesis. Oxford University Press, Oxford, UK.

Bayoh, M., and S. Lindsay. 2003. Effect of temperature on the development of the aquatic stages of *Anopheles gambiae* sensu stricto (Diptera: Culicidae). Bulletin of Entomological Research 93(5):375–382.

Bayoh, M. N. 2001. Studies on the development and survival of *Anopheles gambiae* sensu stricto at various temperatures and relative humidities. Dissertation. Durham University, Durham, UK.

Clark, J. S. 2007. Models for ecological data. Princeton University Press, Princeton, New Jersey, USA.

Dell, A. I., S. Pawar, and V. M. Savage. 2011. Systematic variation in the temperature dependence of physiological and ecological traits. Proceedings of the National Academy of Sciences USA 108(26):10591–10596.

Dietz, K. 1993. The estimation of the basic reproduction number for infectious diseases. Statistical Methods in Medical Research 2(1):23–41.

Elderd, B. D., V. M. Dukic, and G. Dwyer. 2006. Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. Proceedings of the National Academy of Sciences USA 103(42):15693–15697.

Gething, P. W., D. L. Smith, A. P. Patil, A. J. Tatem, R. W. Snow, and S. I. Hay. 2010. Climate change and the global malaria recession. Nature 465(7296):342–345.

Guerra, C. A. 2007. Mapping the contemporary global distribution limits of malaria using empirical data and expert opinion. Dissertation, University of Oxford, Oxford, UK.

Hay, S. I., J. Cox, D. J. Rogers, S. E. Randolph, D. I. Stern, G. D. Shanks, M. F. Myers, and R. W. Snow. 2002. Climate change and the resurgence of malaria in the east African highlands. Nature 415(6874):905–909.

Johnson, L. R., K. D. Lafferty, A. McNally, E. A. Mordecai, K. P. Paaijmans, S. Pawar, and S. J. Ryan. 2014. Mapping the distribution of malaria: current approaches and future directions. Pages 191–211 *in* D. Chen, B. Moulin, and J. Wu, editors. Analyzing and modeling spatial and temporal dynamics of infectious diseases. John Wiley & Sons, Hoboken, New Jersey, USA.

Keeling, M. J., and P. Rohani. 2008. Modeling infectious diseases in humans and animals. Princeton University Press, Princeton, New Jersey, USA.

Lardeux, F., R. Tejerina, V. Quispe, and T. Chavez. 2008. A physiological time analysis of the duration of the gonotrophic cycle of *Anopheles pseudopunctipennis* and its implications for malaria transmission in Bolivia. Malaria Journal 7(1):141.

Macdonald, G. 1952. The analysis of equilibrium in malaria. Tropical Diseases Bulletin 49(9):813.

Merl, D., L. R. Johnson, R. B. Gramacy, and M. Mangel. 2009. A statistical framework for the adaptive management of epidemiological interventions. PLoS ONE 4(6):e5807.

Molnár, P. K., S. J. Kutz, B. M. Hoar, and A. P. Dobson. 2013. Metabolic approaches to understanding climate change impacts on seasonal host-macroparasite dynamics. Ecology Letters 16(1):921.

Mordecai, E. A., et al. 2013. Optimal temperature for malaria transmission is dramatically lower than previously predicted. Ecology Letters 16(1):2230.

Plummer, M. 2003. Jags: a program for analysis of Bayesian graphical models using Gibbs sampling. Pages 20–22 *in* K. Hornik, F. Leish, and A. Zeilis, editors. Proceedings of the Third International Workshop on Distributed Statistical Computing (DSC 2003), Vienna, Austria, March 20–22, 2003.

Plummer, M. 2013. rjags: Bayesian graphical models using MCMC. R package version 3.10. http://cran.r-project.org/web/packages/rjags/index.html

R Development Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.r-project.org

Rohr, J. R., A. P. Dobson, P. T. Johnson, A. M. Kilpatrick, S. H. Paull, T. R. Raffel, D. Ruiz-Moreno, and M. B. Thomas. 2011. Frontiers in climate change-disease research. Trends in Ecology and Evolution 26(6):270–277.

Ryan, S. J., A. McNally, L. R. Johnson, E. A. Mordecai, T. Ben-Horin, K. P. Paaijmans, and K. D. Lafferty. *In press*. Rising suitability and declining severity of malaria transmission in Africa under climate change. arXiv:1407.7612 [q-bio.QM]

Siraj, A., M. Santos-Vega, M. Bouma, D. Yadeta, D. R. Carrascal, and M. Pascual. 2014. Altitudinal changes in malaria incidence in highlands of Ethiopia and Colombia. Science 343(6175):1154–1158.

Sternberg, E. D., and M. B. Thomas. 2014. Local adaptation to temperature and the implications for vector-borne diseases. Trends in Parasitology 30(3):115–122.

Supplemental Material

Ecological Archives

Appendices A–C are available online: http://dx.doi.org/10.1890/13-1964.1.sm