

## APPENDIX S3

### ECOSPHERE

#### HABITAT PREFERENCE OF AN HERBIVORE SHAPES THE HABITAT DISTRIBUTION OF ITS HOST PLANT

NICOLAS M. ALEXANDRE, PARRIS T. HUMPHREY, ANDREW D. GLOSS, JIMMY LEE, JOSEPH FRAZIER, HENRY A. AFFELDT III, AND NOAH K. WHITEMAN.

#### STATISTICAL SUPPLEMENT

1. Herbivory survey	1
1.1. Biological motivation	1
1.2. Modeling zero-inflation	2
1.3. Modeling the damage intensity process	3
1.4. Accounting for heteroscedasticity	5
1.5. Model diagnostics and comparisons	5
2. Herbivore choice tests I: Sun versus shade derived bittercress	9
2.1. Models for choice experiments	9
3. Herbivore choice tests II: Effects of light and temperature	10
3.1. 2014 Trials	10
3.2. 2015 Trials	10
3.3. Combined 2014-2015 Analysis	11
References	12

#### Contents

#### 1. HERBIVORY SURVEY

**1.1. Biological motivation.** Herbivore damage (in this case, counts of feeding punctures made by adult females [‘stipples’], leaf mines made by larvae, and eggs laid by adult females) is typically recorded as positive integer counts, and such count data often exhibit under-dispersion (zero-inflation), over-dispersion (excess variance), or both, with respect to expectations of Poisson or Poisson–Gamma mixture (i.e. negative binomial, NB) models. Our understanding of the foraging ecology of *Scaptomyza nigrata* leads us to expect that dispersion of both of these types is likely. Female flies are choosy, often visiting several leaves before making feeding punctures; females may also avoid leaves entirely, either actively or due to stochasticity in the host sampling process, and local abundances of foraging *S. nigrata* may vary across sampling locations. All three processes can affect the prevalence of damage among leaves and can lead to under-dispersion, and the extent to which this is observed may well vary among sites or habitat types.

In addition, once a host has been preliminarily accepted, we expect that variation in the intensity of feeding damage arises from factors perceived subsequent to the initiation of damage. We thus assume that separate (but potentially related) biological processes govern the host acceptance versus the host damage stages of herbivory as measured in our study, and this also may vary across plants, site or habitat type. Below, Figure S1 shows the distribution of stipple and mine counts from our herbivory survey, broken down by habitat type, and from these raw data plots we can already appreciate the strong difference between habitat types in the prevalence (0 versus  $> 0$  counts) as well as the average intensity of damage (distribution of counts  $> 0$ ).

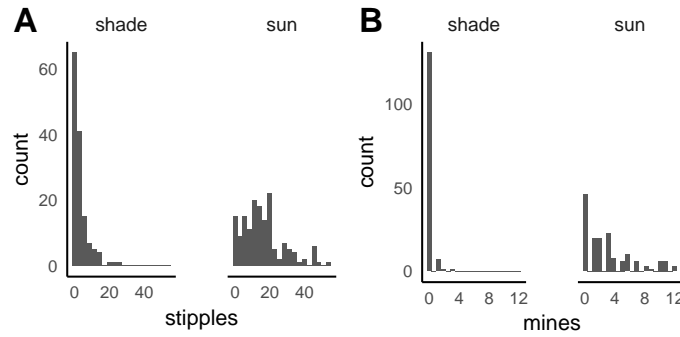


FIGURE S1. Distribution of (A) stippling and (B) leaf miner damage on sun and shade-grown bittercress across  $n = 15$  sites in a field herbivory survey.

When we compute the ratio of the variance to the mean for stipple and leaf mine counts for each site (Figure S2), we observe variation both within and between habitat types. When examining these two figures, we might suspect that a standard Poisson model may be inappropriate for accounting for the range of variation observed in these data, given that we expect the sample variance to sample mean ratio to be  $\sim 1$  under a standard Poisson process. Instead, we observe ratios  $> 1$  and distributions of ratios that differ between site types (Figure S2; i.e., heteroscedasticity). These two patterns indicate that our count data will *not* be suitably modeled without accounting for zero-inflation and over-dispersion simultaneously. Below we describe our modeling approach for handling dispersion of both types.

**1.2. Modeling zero-inflation.** Zero-inflated models are mixture models composed of a binomial component that models the zero-inflated proportion of the data, and a count component (typically Poisson or NB) that is jointly fit. In such models, the extra zero count compartment results from an inflation of the probability density at  $Y = 0$  given some unmeasured latent variable that governs the exposure of sampling units to the processes that generate the actual counts [5]. When combined with the probability density function of the Poisson distribution, a zero-inflated Poisson becomes

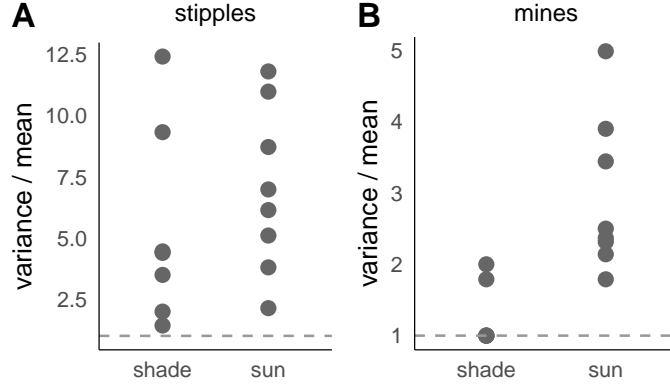


FIGURE S2. Variance/mean ratios for (A) stippling and (B) leaf miner damage on sun and shade-grown bittercress. The dashed line indicates unity.

$$(1) \quad Pr\{Y = y|\theta\} = \begin{cases} \pi_0 + (1 - \pi_0) \cdot e^{-\theta}, & \text{if } y = 0. \\ (1 - \pi_0) \cdot \frac{\theta^y}{y!} e^{-\theta}, & \text{if } y > 0. \end{cases}$$

where  $\theta$  is the Poisson mean and  $y$  is an observed count, while the latent variable  $\pi_0$  is the probability that a zero count is drawn from the under-exposure sample category rather than from the exposed category of samples. We might wish to describe the influence of habitat type on  $\pi_0$ , and to do so we construct a binomial linear model which estimates the zero-inflation coefficient  $\pi_0$  for each habitat type  $k$ , using the canonical logit link function:

$$(2) \quad \text{logit}(\pi_0^k) = \log\left(\frac{\pi_0^k}{1 - \pi_0^k}\right) = \begin{cases} \delta_0 & \text{if habitat} = \text{shade.} \\ \delta_0 + \delta_1 & \text{if habitat} = \text{sun.} \end{cases}$$

where  $\delta_0$  is the estimated coefficient for the shade habitat, taken as the reference level (i.e., the Constant, or Intercept), to which the effect of the sun habitat ( $\delta_1$ ) is added. (Note that the superscript  $k$  indicates a group-level index and is not an exponent). In our analysis, we generated maximum likelihood estimates for these ZI parameters using **R** [3] package **glmmTMB** [1, 4, 2]. In the main text (Table 1) we report the estimates of  $\pi_0^{\text{shade}}$  and  $\pi_0^{\text{sun}}$  on the probability (i.e. inverse logit) scale along with 95% (Wald type) confidence intervals.

**1.3. Modeling the damage intensity process.** Our damage intensity counts are empirically over-dispersed compared to expectations of a Poisson process (Figure S2). This arises when the variance grows faster than the mean such that variance/mean ratios exceed 1. To accommodate this fact, we model counts of herbivore damage (stipples and leaf mines) as a Poisson random variable ( $Y_i$ ) where the Poisson mean ( $\Theta_i$ ) is itself a random variable.

This gives rise to the following expression for the conditional probability of  $Y_i$  given  $\Theta_i$ , in the context of zero-inflation:

$$(3) \quad Pr\{Y_i = y | \Theta_i = \theta\} = \begin{cases} \pi_0^k + (1 - \pi_0^k) \cdot e^{-\theta} & \text{if } y = 0 \\ (1 - \pi_0^k) \cdot \frac{\theta^y}{y!} e^{-\theta} & \text{if } y > 0 \end{cases}$$

Conceptually, this means that the variance of our random variable  $Y$  includes the Poisson variance associated with each Poisson mean  $\Theta_i$  as well as additional variance in the distribution of  $\theta$  itself. We adopt the conventional probability density function for  $\theta$  as a Gamma distribution with scale parameter  $\alpha$  and rate parameter  $\beta$  [5]:

$$g(\theta) = \frac{\alpha^\beta}{\Gamma(\beta)} \theta^{\beta-1} \cdot e^{-\alpha\theta}$$

We recover the the unconditional probability of  $Y_i$  by integrating out the  $\theta$  (i.e. averaging over the distribution of Poisson means) to recover the parameterization of the negative binomial (NB) that we will use [5]:

$$(4) \quad \begin{aligned} Pr\{Y_i = y\} &= \int_0^\infty Pr\{Y_i = y | \Theta_i = \theta\} \cdot g(\theta) \cdot d\theta \\ &= \frac{\Gamma(\alpha + y)}{y! \Gamma(\alpha)} \cdot \left( \frac{\mu}{\mu + \beta} \right)^y \left( \frac{\beta}{\beta + \mu} \right)^\beta \end{aligned}$$

Note that the mean of  $g(\theta)$  is  $\alpha/\beta$  and its variance is  $\alpha/\beta^2$ . Setting  $\alpha = \beta = \phi$  allows us to re-write the mean as equal to 1 and the variance  $\sigma^2$  as  $1/\phi$ . This gives the expectation of  $Y$  for leaf  $i$  equal to  $\mu$  and the variance equal to  $\mu + \sigma^2 \mu^2$ . Thus, the 'dispersion parameter' of the NB,  $\sigma^2$  (or alternatively,  $1/\phi$ ), can be interpreted as the variance of the Gamma distribution from which the Poisson means are drawn. When  $\sigma^2 \rightarrow 0$  (and thus  $\phi \rightarrow \infty$ ) the negative binomial collapses to the Poisson, where  $E[Y_i] = \text{Var}[Y_i] = \mu$ . (Note that the  $\mu$  of the NB is not equivalent to the  $\alpha/\beta$  from the Gamma distribution).

The full expression for the unconditional probability of  $Y_i$  becomes (using the  $\sigma^2 = 1/\phi$  parameterization):

$$(5) \quad Pr\{Y_i = y\} = \begin{cases} \pi_0 + (1 - \pi_0) \cdot \left(\frac{\phi}{\phi + \mu}\right)^\phi, & \text{if } y = 0. \\ (1 - \pi_0) \cdot \frac{\Gamma(r + y)}{y! \Gamma(r)} \cdot \left(\frac{\mu_i}{\mu + \phi}\right)^y \left(\frac{\phi}{r + \mu}\right)^\phi, & \text{if } y > 0. \end{cases}$$

Thus,  $Pr\{Y_i\}$  depends on the three parameters,  $\pi_0$ ,  $\mu$ , and  $\phi$ . In our zero-inflated negative binomial (ZINB) framework, the means of the NB ( $\mu$ ) are modeled, via the log link function, as a linear function of several predictor variables:

$$(6) \quad \log(\mu_{ijl}^k) = \begin{cases} \alpha_0 + \beta x_i + \gamma_{1,j} + \gamma_{2,l} & \text{if habitat} = \text{shade.} \\ \alpha_0 + \alpha_1 + \beta x_i + \gamma_{1,j} + \gamma_{2,l} & \text{if habitat} = \text{sun.} \end{cases}$$

Here,  $\alpha_1$  indicates the fixed effect of the sun habitat type;  $\gamma_{1,j}$  represents a random effect of stem-to-stem variation; and  $\gamma_{2,l}$  represents a random effect of site-to-site variation, where  $\gamma_1 \sim \mathcal{N}(0, \tau_1^2)$  and  $\gamma_2 \sim \mathcal{N}(0, \tau_2^2)$ .  $\beta$  captures the effect of area ( $x_i$ ) of individual leaves. Again,  $k$  is a group-level superscript indexing habitat. We sampled two leaves from each of ten plants at all 15 sites, giving a total of  $n = 298$  observations (two data points were removed due to missing leaf area data). Thus, we have five total parameters to estimate, corresponding to each level of  $\alpha_i$  and the single  $\beta$  term for the stipple model, as well as the group-level standard deviations for plant stems ( $\tau_1$ ) and sites ( $\tau_2$ ).

**1.4. Accounting for heteroscedasticity.** Differences in the average variance to mean ratio (e.g. Figure S2) across habitat types indicates that certain sub-sets of the data may behave more like a pure Poisson (e.g. Figure S2B, shade) while another subset may behave more like an NB process. When data sub-sets have unequal variances, one can estimate separate residual variance terms for each group in order to improve goodness-of-fit and achieve more reliable population-level parameter estimates. In the context of an NB process, accounting for heteroscedasticity can take the form of estimating separate dispersion parameters ( $\phi_k$ ) for focal factor levels 1 through  $k$ . In this case, we utilize the flexibility of package `glmmTMB` to fit the NB dispersion parameter  $\phi$  as a linear function of habitat type  $k$ , using a log link:

$$(7) \quad \log(\phi_k) = \begin{cases} \delta_0 & \text{if habitat} = \text{shade.} \\ \delta_0 + \delta_1 & \text{if habitat} = \text{sun.} \end{cases}$$

**1.5. Model diagnostics and comparisons.** Below we illustrate our modeling approach by presenting a justification for using NB over Poisson models, as well as our inclusion of factor-specific zero-inflation and NB dispersion parameters, focusing on the stipple counts from the herbivory survey as an example. We do so by calculating several test statistics that help us judge whether the model construction produces simulated counts which

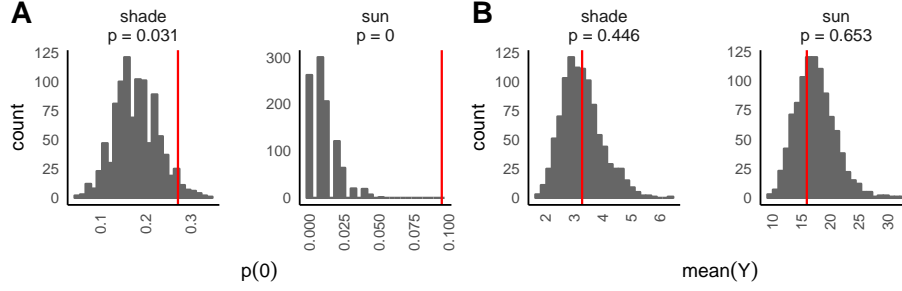


FIGURE S3. Predictive model checks of (A) proportion of zero counts  $p(0)$  and (B) mean( $Y$ ) for Poisson-only model of stipple counts. Observed values are indicated by red vertical lines.

adequately reflect the underlying but unknown count generating processes that generated the observed data in our studies. From each model fitted under maximum likelihood, we generated  $n = 1000$  predicted distributions of count data and calculated the proportion of simulated draws more extreme than the observed value. This serves as an empirical  $p$ -value calculated as  $p_Z = \frac{1}{R} \sum_{r=1}^R I(\tilde{Z}_r \geq Z_{obs})$ , where  $I(\cdot)$  is the indicator function that equals 1 if the condition is true and 0 otherwise,  $\tilde{Z}_r$  is the test statistic of the simulated data for replicate  $r$ , and  $Z_{obs}$  is the observed value of the same test statistic. If  $p_Z < 0.05$  or  $> 0.95$ , we are inclined to think that the model provides a poor representation of the underlying data generating process through the lens of a given test statistic. Our test statistics are (1) the proportion of zeros in the data  $p(0)$ , (2) mean( $Y$ ), and (3) max( $Y$ ). These were chosen as ways to characterize goodness-of-fit of the estimated data-generating process at the extremes of the data distributions as well as around its central tendency.

**1.5.1. ZI justification.** To justify our use of a ZINB model for the herbivory survey data, we first constructed Poisson-only models and compared them to global or habitat-specific ZI-Poisson models. (For the present illustration of our approach, we fit all models *without* the leaf area term to improve convergence of the Poisson model with habitat-specific ZI terms; models *with* this term are reported in the main text, and all direct comparisons can be found in the associated R code [<https://github.com/phumph/habitat>]). Model checking (Figure S3) reveals that the proportion of zeros under the fitted Poisson-only model simulations is not representative of the observed data ( $p_{Z_0} = 0.03$  for shade sites,  $p_{Z_0} \leq 10^{-3}$  for sun sites). In other words, the model simply fails to accurately represent the prevalence of damage across the dataset. When introducing a global zero-inflation term  $\pi_0$ , model fit greatly improves at the low end, as  $p_{Z_0}$  increases to 0.85 for sun sites (Figure S4), and the overall goodness-of-fit improves substantially when judged by AICc ( $\Delta\text{AICc} = -129$ ; Table S1). But we are still doing a poor job at capturing  $p(0)$  in shade habitats, suggesting that habitat-specific estimates of  $\pi_0^k$  may help. Indeed, modeling habitat-specific ZI improves the goodness-of-fit in shade habitats specifically ( $p_{Z_0} \sim 0.40$  for both habitat types; Figure S5). Including an additional model parameter is also statistically favored overall ( $\Delta\text{AICc} = -1.8$ ; Table S1).

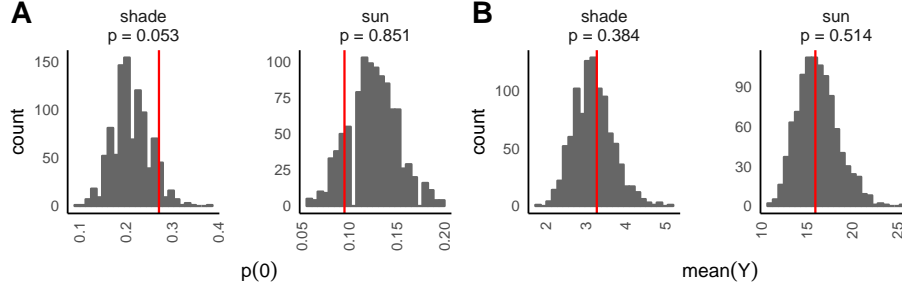


FIGURE S4. Predictive model checks of (A) proportion of zero counts  $p(0)$  and (B) mean( $Y$ ) for ZI-Poisson model of stipple counts with single global ZI parameter  $\pi_0$ . Observed values are indicated by red vertical lines.

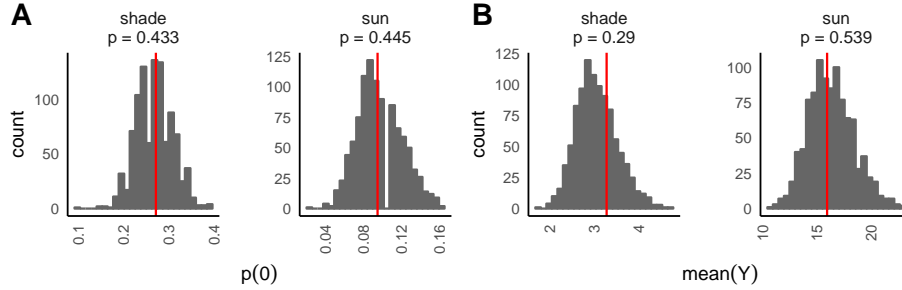


FIGURE S5. Predictive model checks of (A) proportion of zero counts  $p(0)$  and (B) mean( $Y$ ) for ZI-Poisson model of stipple counts with habitat-specific ZI parameters  $\pi_0^k$ . Observed values are indicated by red vertical lines.

When we conduct a third predictive check to see how the improved model behaves at the high end, via calculating  $\max(Y)$ , we see that the Poisson model fails to capture the difference between sites in the maximum counts observed (Figure S6).

1.5.2. *Zero-inflated negative binomial*. Model checking in this way shows that Poisson models, even when including habitat-specific ZI terms, can only take us so far in predicting the observed data across its entire range. The lack of fit of the *same data used to estimate the model* indicates lack of flexibility in the model. Because we care about accurately modeling zero counts (i.e. damage prevalence) as well as damage abundance patterns, we explore a set of NB models in order to test whether introducing an additional latent parameter (NB dispersion parameter  $\phi$ ) improves the ability of the model to capture habitat-specific patterns of herbivory.

Specifically, we compared NB-only to a series of NB models containing a single global ZI, habitat-specific ZI, as well as habitat-specific dispersion parameters with and without consideration of ZI (Table S1). Ultimately, the best model that gives both the lowest AICc (Table S1) as well as the most representative  $p_Z$  estimates is a NB model with a global

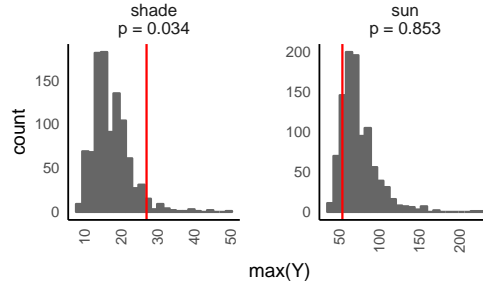


FIGURE S6. Predicted versus observed  $\max(Y)$  for ZI-Poisson model of stipple counts with habitat-specific ZI parameters  $\pi_0^k$ . Observed values are indicated by red vertical lines.

$\pi_0$  and habitat specific  $\phi^k$ . Specifically, this model performs at least as well as the two-level ZI-Poisson while also fitting better at the upper range, as indicated by the more representative range of  $\max(Y)$  (Figure S7).

This same analysis procedure was applied for the leaf miner data from the herbivory survey. These results are not displayed here but can be found in the **R** scripts in the **github** repository associated with this manuscript (<https://github.com/phumph/habitat/>).

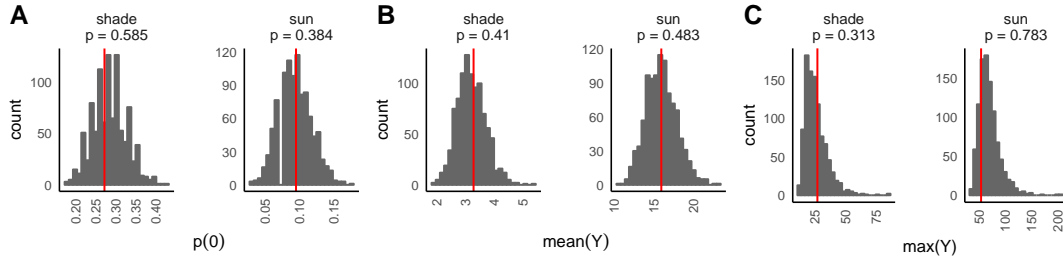


FIGURE S7. Predictive model checks for best-fitting ZINB model of stipple counts, with habitat-specific dispersion parameters  $\phi^k$  and global ZI parameter  $\pi_0$ . Observed values are indicated by red vertical lines.

TABLE S1. Goodness-of-fit statistics for Poisson and NB models with and without accounting for ZI and dispersion.

family	$\Delta\text{AICc}$	df	$\pi_0^*$	NB $\phi^*$	$p(0)^{\text{shade}}$	$p(0)^{\text{sun}}$	$\text{mean}(Y)^{\text{shade}}$	$\text{mean}(Y)^{\text{sun}}$	$\text{max}(Y)^{\text{shade}}$	$\text{max}(Y)^{\text{sun}}$
NB	0	7	1	2	0.59	0.38	0.41	0.48	0.31	0.78
NB	1.8	8	2	2	0.47	0.43	0.44	0.47	0.36	0.77
NB	19.3	7	2	1	0.44	0.45	0.34	0.56	0.09	0.93
NB	20.6	6	1	1	0.09	0.80	0.39	0.51	0.08	0.94
NB	63.3	5	0	1	0.07	< 0.001	0.46	0.62	0.32	0.99
Poisson	72.5	6	2	–	0.43	0.45	0.29	0.54	0.03	0.85
Poisson	74.3	5	1	–	0.05	0.85	0.38	0.51	0.05	0.90
Poisson	204.0	4	0	–	0.03	< 0.001	0.45	0.65	0.20	0.94

\* 1 = single global parameter; 2 = habitat-specific parameters; 0 = not estimated; – = not applicable  
df = degrees of freedom

## 2. HERBIVORE CHOICE TESTS I: SUN VERSUS SHADE DERIVED BITTERCRESS

**2.1. Models for choice experiments.** Using NB generalized linear mixed models (GLMMs), we model variation in stipple and egg counts on plants arising from source habitat (sun versus shade) and leaf attributes (as fixed effects), as well as so-called structural random effects of plant ID nested within cage ID (i.e. replicate) to capture experimental design constraints that determine the level of independence among data points. We do not consider Poisson models in all subsequent analyses on the basis of far worse model fits observed for stipple and leaf mine data, above.

We also disfavored ZI models in principle because we assumed that *S. nigrita* females had sufficient time to sample all available plants, and that variation in the intensity (rather than prevalence) of herbivory would drive most of our patterns. We did, however, consider whether including source habitat-specific estimates of  $\phi_k$  empirically improved model fit, via model checking and AICc (see Results and Table 1 of main text).

Our mixed model for stipple and eggs counts, for both the whole-plant and detached leaf assays, takes the following form:

$$(8) \quad \log(\mu_{ijk}) = \begin{cases} \alpha_0 + \beta x_i + \gamma_{(jk)} + \gamma_k & \text{if habitat = shade.} \\ \alpha_0 + \alpha_1 + \beta x_i + \gamma_{(jk)} + \gamma_k & \text{if habitat = sun.} \end{cases}$$

In this model,  $\mu_{ijk}$  is the estimate for each leaf  $i$ ,  $\alpha_0$  and  $\alpha_1$  are as in Eqn. 6, and  $\beta$  is leaf position along stem (low to high) for the whole-plant model or leaf area ( $\text{mm}^2$ ) for the detached leaf assay; finally,  $\gamma_k$  represents a random effect of cage ID ( $k$ ) and  $\gamma_{(jk)}$  represents a random effect for plant ID  $j$  nested within each level of cage ID  $k$ ;  $\gamma_{(jk)} \sim N(\gamma_k, \tau_1^2)$  and  $\gamma_k \sim N(0, \tau_2^2)$ . Model comparisons and goodness-of-fit tests for analyses associated with this assay can be found in the **R** notebooks in our [github](#) repository listed above.

### 3. HERBIVORE CHOICE TESTS II: EFFECTS OF LIGHT AND TEMPERATURE

The NB model structures for this set of choice experiments were similar to those described above but with an expanded random effects structure. As for the previous herbivore choice models, all models described below are NB-only GLMMs. The choice trials in 2014 and 2015 were conducted slightly differently, which means that each year’s data calls for slightly different random effects. We present the analyses of each dataset separately, then jointly.

**3.1. 2014 Trials.** The 2014 field and lab trials were conducted in two temperature environments simultaneously, with one cage held in each. This gives a structure of temperature environment ( $\gamma_{(lk)}$ ,  $n = 2$ ) nested within trial ( $\gamma_l$ ,  $n = 6$  each for the 2014 field and lab assays). Nested within temperature environment is cage, but since we have only a single level of cage for each temperature environment per trial, this level is irrelevant for the 2014 dataset. However, we include side-of-cage ( $\gamma_{(kj)}$ ,  $n = 2$ ; left or right, arbitrarily) to control for pseudo-replication at the level of the main treatment effect (i.e. Light environment, light versus dark), which was applied with randomization to the sides of each cage. Each random effect is  $\sim N(0, \sigma_i^2)$ . The full model is given below using the above notation for the random effects. The number of independent data points in the 2014 field and lab trials is 24 each (2 sides per cage, 1 cage per trial, 2 temperature settings per trial, and 6 trials).

$$(9) \quad \log(\mu_{ijkl}) = \begin{cases} \alpha_0 + \alpha_{1,i} + \alpha_{2,i} + \alpha_{12} + \beta x_i + \gamma_l + \gamma_{(lk)} + \gamma_{(kj)} & \text{if Light = light, Temp = warm} \\ \alpha_0 + \alpha_{1,i} + \alpha_{2,i} + \beta x_i + \gamma_l + \gamma_{(lk)} + \gamma_{(kj)} & \text{otherwise.} \end{cases}$$

We set the first levels of coefficients for the fixed effects of Light ( $\alpha_{1,1} = \text{dark}$ ) and Temperature ( $\alpha_{2,1} = \text{cool}$ ) equal to zero since these two states are incorporated into the Constant ( $\alpha_0$ , i.e. the reference level of the model). Our experiment was designed to allow us to fit an interaction term  $\alpha_{12}$  which estimates how much the effect of Light is impacted by the Temperature environment (thus  $\alpha_{12}$  has a single level). We fit models with the same structure for field and laboratory trials, and we consider condition-specific  $\phi$  estimates when conducting model comparisons. Specifically, we estimated whether  $\phi$  specific to light condition and/or temperature condition improved model fit by accounting for differential distributions of overdispersion analogous to Eq. 7 (with an additive [on the log scale] impact of temperature treatment).

**3.2. 2015 Trials.** In the 2015 trials, we used bittercress leaves collected from both sun and shade habitats, and our randomization scheme ensured equal representation of sun- and shade-derived leaves across all treatments. We did this to test whether *S. nigrita* would exhibit preference for shade-derived bittercress in the context of our temperature and light manipulations. In our analysis, we included leaf source (sun v. shade) as an additional fixed factor. Our model structure was thus

(10)

$$\log(\mu_{ijkl}) = \begin{cases} \alpha_0 + \alpha_{1,i} + \alpha_{2,i} + \alpha_{12} + \alpha_{3,i} + \beta x_i + \gamma_l + \gamma_{(lk)} + \gamma_{(kj)} & \text{if Light = light, Temp = warm} \\ \alpha_0 + \alpha_{1,i} + \alpha_{2,i} + \alpha_{3,i} + \beta x_i + \gamma_l + \gamma_{(lk)} + \gamma_{(kj)} & \text{otherwise.} \end{cases}$$

All coefficients are as in the 2014 trials above, except that the term  $\alpha_{3,i}$  is for source habitat, and the first level ( $\alpha_{3,1}$  = shade) was set to 0 and was thus incorporated into the intercept term  $\alpha_0$ . Thus, this model estimates only a single coefficient for the effect of habitat.

The experimental procedures also differed slightly in 2015. We placed two cages in one environmental chamber which was set to a single temperature per trial, and individual trials were conducted at different temperatures sequentially. This implies a structure of cage ( $\gamma_{kl}$ ,  $n = 2$ ) nested within trial ( $\gamma_l$ ,  $n = 5$ ), along with the side-of-cage nested within cage ( $\gamma_{(kj)}$ ,  $n = 2$ ). The number of independent data points in the 2015 trials is thus 40 (2 sides per cage, 2 cages per room, 10 trials [5 in each temperature regime]). While the model structure is the same as in Eqn. 8, the meaning of the random effects (and their coding in the design matrix) are slightly different. Below we discuss how we reconciled the random effects to analyze both years' data together.

**3.3. Combined 2014-2015 Analysis.** We reconcile the slight distinctions between 2014 and 2015 datasets by including trial-year as a composite random effect, now re-coded to reflect each experiment conducted in a given room (compared to a trial containing two separate temperature settings, as in the 2014-only analysis). Nested within the new trial factor is cage, which has  $n = 1$  for 2014 and  $n = 2$  for 2015; side-of-cage is modeled in the same way as above. This model structure is now identical to Eqn. 10.

For the combined analysis, we drop the term for plant source habitat ( $\alpha_3$ ) since the 2014 trials were not designed to examine plant source habitat. Additionally, we model both a continuous and discrete versions of the Temperature factor: the discrete model is the same as Eqn. 8, while in the continuous form,  $\alpha_{2,j}$  is replaced with  $\beta_2 y_j$ , where  $y_j$  is the temperature measured at the level of cage for each trial separately (Appendix B, Fig. 2C). Additionally, the fixed interaction term  $\alpha_{12}$  is now replaced with an interaction modeled by the expression  $\beta_3 y_j(\alpha_{1,i})$ , which captures how the effect of light ( $\alpha_{1,i}$ ) changes as a function of cage temp ( $y_i$ ); since  $\alpha_{1,i}$  has one level (the first level is set to 0), estimating this interaction term  $\beta_3$  adds only a single parameter to the model. The full model is thus:

$$(11) \quad \log(\mu_{ijkl}) = \alpha_0 + \alpha_{1,i} + \alpha_{2,i} + \beta_1 x_i + (\beta_2 + \beta_3 \alpha_{1,i}) y_j + \gamma_l + \gamma_{(lk)} + \gamma_{(kj)}$$

Because the full model with the continuous temperature interaction term ( $\beta_3$ ) exhibited such a poor fit to the data as judged by AIC (see below), we report the coefficient estimates for the nested model containing only the additive effect of the continuous temperature term ( $\beta_2$ ; Table S2).

TABLE S2. Coefficient estimates for all light–temp choice experiments.

	Stipples				Eggs	
	Field (2014)	Lab (2014)	Lab (2015)	Both years (1)	Both years (2)	Lab (2015)
<i>Fixed effects</i>						
$\alpha_0$ (Constant)	−2.038*** (0.603)	−2.514** (0.870)	0.364 (0.885)	−1.147 (0.661)	−4.123 (2.162)	−4.870*** (0.812)
$\beta_1$ leaf width (mm)	0.068*** (0.012)	0.063* (0.026)	−0.003 (0.022)	0.033 (0.018)	0.031 (0.017)	0.054** (0.019)
$\alpha_{1,}$ [light]		1.792* (0.838)	2.160*** (0.515)	2.019*** (0.470)	2.321*** (0.339)	3.598*** (0.537)
$\alpha_2$ [warm]		0.399 (0.885)	−0.100 (0.880)	0.109 (0.667)		1.960** (0.650)
$\alpha_{12}$ [light:warm]		1.123 (1.172)	0.293 (0.718)	0.636 (0.654)		−1.994** (0.628)
$\alpha_3$ [sun]			−0.291 (0.701)			−0.002 (0.19)
$\beta_2$ [temp]					0.149 (0.102)	
<i>Random effects</i>						
$\tau_{kj}^2$ [trial/cage/side]	0.559	1.397	0.720	1.102	1.129	0.152
$\tau_{jk}^2$ [trial/cage]	0.000	0.099	2.451	1.687	1.587	0.170
$\tau_l^2$ [trial]	0.204	0.000	0.000	0.000	0.063	0.163
$\tau_m^2$ [year]				0.036	0.000	
AIC	1152.4	874.4	2125.7	2995.8	2993.4	861.2
Log Likelihood	−567.2	−428.2	−1053.9	−1487.9	−1487.7	−421.6
Num. obs.	240	240	398	638	638	398

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

## REFERENCES

- [1] Simon Jackman. *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory, Stanford University*. Department of Political Science, Stanford University, Stanford, California, 2015.
- [2] Tom Loeys, Beatrijs Moerkerke, Olivia De Smet, and Ann Buysse. The analysis of zero-inflated count data: beyond zero-inflated poisson regression. *Br J Math Stat Psychol*, 65(1):163–80, Feb 2012.
- [3] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017.
- [4] Achim Zeileis, Christian Kleiber, and Simon Jackman. Regression models for count data in r. *Journal of Statistical Software, Articles*, 27(8):1–25, 2008.
- [5] A. Zuur, E.N. Ieno, N. Walker, A.A. Saveliev, and G.M. Smith. *Mixed Effects Models and Extensions in Ecology with R*. Statistics for Biology and Health. Springer New York, 2009.