ELECTRONIC SUPPLEMENTAL MATERIAL

Hall SR, Becker CR, Duffy MA, Cáceres CE. 2011. Epidemic size determines population-level effects of fungal parasites on *Daphnia* hosts. *Oecologia*

In this Appendix, we provide supporting statistical results and figures. Table A1 provides full ANOVA tables for models used to analyze data from the experiment. Figure A1 shows quantile plots describing prevalence of infection in several different lakes. Fungal outbreaks did not start in all lakes in all years. Maximum prevalence of infection exceeded zero in 64% of the lake-year combinations (37 of 58 lake-years). Most of those 37 epidemics remained small; maximum prevalence exceeded 10% in only 16 of them (Figure A1). Furthermore, the frequency of these more severe epidemics varied among lakes. For example, although all four epidemics observed in Little Mill Lake exceeded 15% infection prevalence, the largest epidemic (>40%) occurred in Basset Lake where epidemics only exceeded 5% infection prevalence during one other year (Figure A1A). Furthermore, 80% of 186 lake-dates with some infection present showed less than 10% prevalence. In 20% of cases overall among lake-dates, fungal prevalence ranged from 10-45%, but the lakes varied considerably in shapes of these quantile curves (e.g., Warner had relatively little infection overall, Little Mill had much, and other lakes fall between; Figure A1B).

Figure A2 illustrates sample calculations used to create the severity index on the *y*-axis of Figure 1. Figure A3 shows the benefits of first using standardized regression coefficients in Figure 1. When regression slopes were estimated from each epidemic without first standardizing the data, we saw that large epidemics have seemingly small, negative effects on hosts, while small epidemics have comparatively larger positive relationships between prevalence and disease. The discrepancy between Figures 1A and A3 involves the slope (rise on the *Y*-axis [density] over run on the *X*-axis [prevalence]). Small epidemics have small 'runs' along the *X*-axis, thus have small denominators on the slope estimator; thus, even with small changes on the *Y*-axis, slope estimates can be large. The standardized slope estimates make each epidemic

comparable – the size of the epidemic is judged in terms of standard deviation units. Figure A4 presents more data on infection prevalence, epidemic severity, and timing of epidemics; this information is complementary to Figure 1. Figure A5 shows positive relationships between infection prevalence and population density at the beginning of four larger epidemics. This result echoes positive density-prevalence patterns seen in small epidemics (Figure 1). Figure A6 illustrates negative relationships between density of host *Daphnia dentifera* and egg ratios (eggs/individual) of uninfected adults in spore addition (top panel) and no spore added treatments of the mesocosm experiment (Figure A6A). These relationships are steeper in spore addition treatments, as evidenced by larger intercepts (*t*-test, $t = -2.54$, $P = 0.044$, $df = 6$) and more negative slopes (*t*-test, $t = -2.66$, $P = 0.038$, $df = 6$; Figure A6B). These data buttress the compensation argument made in the Discussion. They show that drops in host density should increase per capita birth rate of uninfected hosts.

More details on the multivariate auto-regressive (MAR) models

Essentially, the MAR approach fits multiple regressions between ln-transformed density of the different stage classes of uninfected hosts at some sampling date (lag, τ) in the future as linear functions of ln-transformed densities of I, J_U, and A_U in the present. Here, lags τ between 1 and 4 sampling dates were used since we lack a good method to select the most appropriate lag. To illustrate, we could write a linear equation for ln-transformed density of the uninfected adult stage class (A_U) at τ sampling visits in the future ($t+\tau$), $A_U(t+\tau)$, as a function of present, ln-transformed densities of this adult stage class, density of the uninfected juvenile stage class (J_U), and density of the infected host class (I), all observed at time t:

$$A_U(t+\tau) = \alpha_1 + \beta_{aa} A_U(t) + \beta_{ja} J_U(t) + \beta_{ia} I(t) + \varepsilon_1(t) \qquad (1)$$

where α_1 is an intercept and $\varepsilon_1(t)$ is a normally distributed error. In matrix form, three similarly constructed linear regressions can be written compactly as:

$$\begin{bmatrix} A_U(t+\tau) \\ J_U(t+\tau) \\ I(t+\tau) \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \beta_{aa} & \beta_{ja} & \beta_{ia} \\ \beta_{aj} & \beta_{jj} & \beta_{ij} \\ \beta_{ai} & \beta_{ji} & \beta_{ii} \end{bmatrix} \begin{bmatrix} A_U(t) \\ J_U(t) \\ I(t) \end{bmatrix} + \begin{bmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \\ \varepsilon_3(t) \end{bmatrix} \quad (2)$$

where the α_i parameters correspond to intercepts, the ε_i parameters denote errors (i.e., normally-distributed variation not explained by the deterministic portion of the model), and the β_{nk} parameters capture the focal interaction strengths. Here, β_{nk} indicates the effect of n on k; for example, the format reads for β_{ia} in equations (1) and (2) as, "effect of infected hosts [i] on uninfected adults [a]". We present results for β_{ia} and β_{ij}, the effect of infected hosts on uninfected juveniles [j], in Figure 3. Since we fit equation (2) using conditional least squares (Ives et al. 2003, Hall et al. 2009c), parameter estimation was straightforward (i.e., it used the familiar least squares approach; see Ives et al. 2003 and Hall et al. 2009c for the details on calculation of 95% bootstrapped confidence intervals). We considered β parameters to be significant if 95% confidence intervals surrounding them did not overlap zero. Again, we estimated standardized coefficients (i.e., fit after each time series was first divided by its standard deviation) to facilitate comparison of parameters among lakes. See Hall et al. (2009c) for more details.

Table A1. Summary of results from fits of repeated measures ANOVA models to data from the mesocosm experiment. (A) ANOVA table results. (B) Slicing results. Statistically significant ($P < 0.05$) *P*-values are bolded; marginally significant ones ($P < 0.10$) are underlined.

(A). ANOVA tables

Effect	***F*-ratio[1]**	***P*-value**	***F*-ratio[1]**	***P*-value**
	Total density		*Density of D. pulicaria*	
Treatment	1.65	0.25	0.03	0.87
Day	29.3	**<0.0001**	24.8	**<0.0001**
Treatment x Day	2.82	**0.0167**	1.46	0.21
	Density of uninf. juveniles		*Egg ratio*	
Treatment	3.03	0.13	4.8	<u>0.0709</u>
Day	34.1	**<0.0001**	22.5	**<0.0001**
Treatment x Day	3.94	**0.0022**	0.87	0.54
	Density of uninf. adults		*Chlorophyll a*	
Treatment	0.54	0.49	0.12	0.74
Day	9.03	**<0.0001**	54.0	**<0.0001**
Treatment x Day	0.73	0.65	0.48	0.85

(B). Slicing results

Effect	**Day**	***F*-ratio[2]**	***P*-value**	***F*-ratio[3]**	***P*-value**
		Total Density		*Density of Uninf. Juveniles*	
Treatment x Day	1	0.74	0.4	1.42	0.25
	3	0.24	0.63	0.73	0.40
	7	0.02	0.88	0.04	0.86
	10	0.38	0.54	1.33	0.26
	14	10.13	**0.0042**	13.24	**0.0014**
	17	6.47	**0.0183**	11.15	**0.0029**
	21	0.03	0.86	0	1.00
	24	0.07	0.79	0.16	0.69

[1] Degrees of freedom (numerator, denominator): Treatment (1, 6); Day, Treatment × Day (7, 42).

[2] Degrees of Freedom: 1 for numerator, 22.4 in denominator for each day.

[3] Degrees of Freedom: 1 for numerator, 22.8 in denominator for each day.

Figure A1. Quantile plots showing the fraction of epidemics (*y*-axis) that reached a particular prevalence of infection or less (*x*-axis). Data come from nine host populations in seven years (2002-2008). (A) Maximum annual prevalence in the 37 lake-years in which *Metschnikowia* was detected in the hosts. The overall curve shows that most epidemics were small (e.g. 60% of the epidemics had a prevalence ≤ 10%). The symbols indicate different lakes. Filled symbols denote the three intensively monitored lakes shown in Figures 2 and 3. (B) All sampling dates on which *Metschnikowia* was found, grouped by lake. Examining each lake's curve indicates how frequently that lake had an epidemic of a given size.

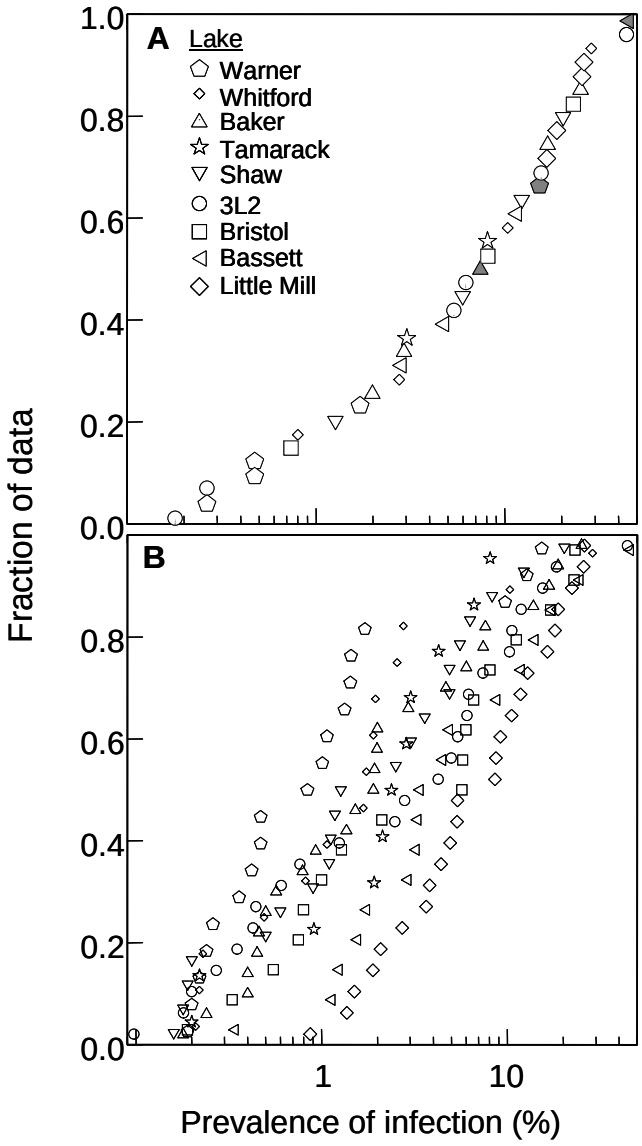

Figure A2. Sample calculations involved in Figure 1. Time series (left panels) and phase plots (right panels) of prevalence, scaled as percent infected, and ln-transformed areal density. Top panels: a small epidemic in Whitford during 2006 showed a positive relationship between density and prevalence. Bottom panels: a larger epidemic in Baker during 2002 shows a negative relationship. R^2: coefficient of determination of regression; $\hat{\beta}$: standardized regression coefficient between prevalence and density, an index of the effects of epidemics on host densities (severity).

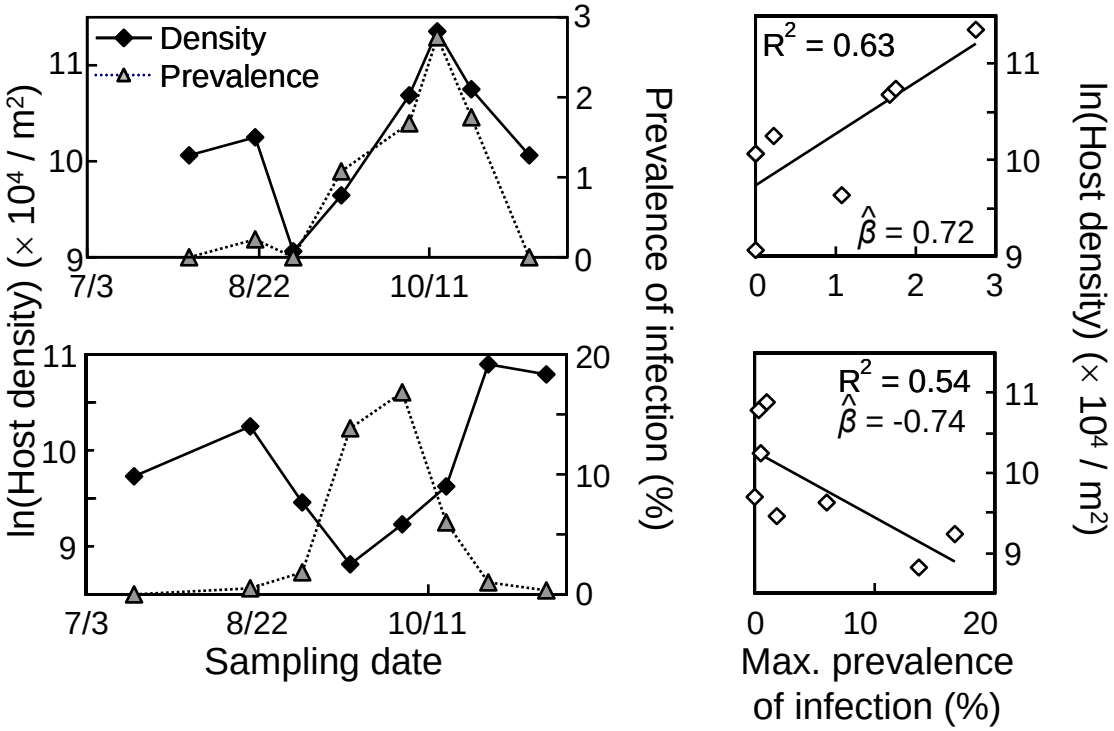

Figure A3. Analogue of Figure 1A in the text. Here, the *Y*-axis shows the regression slopes for each epidemic calculated from the raw data (i.e., neither prevalence nor density for a given epidemic were first divided by their standard deviations before fitting the linear regressions). Grey shading denotes positive unstandardized regression coefficients.

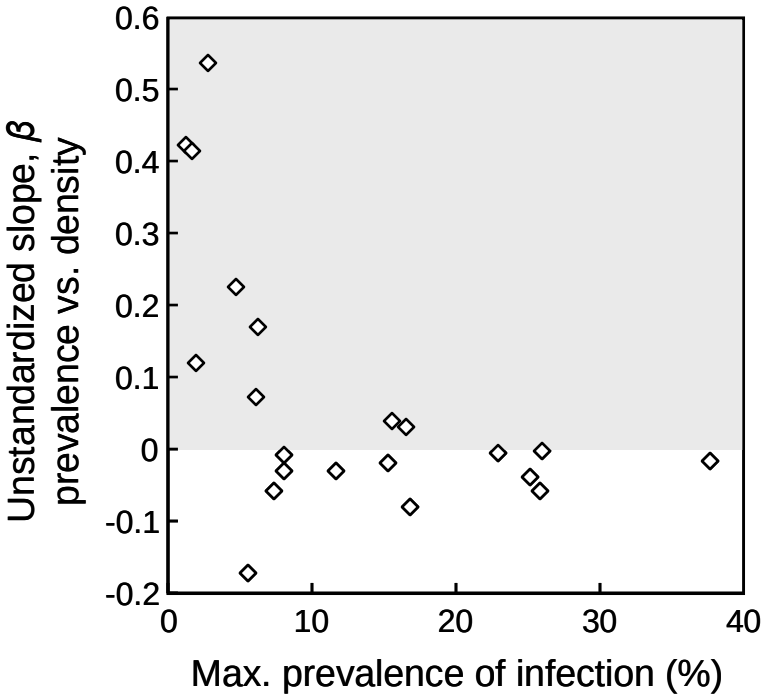

Figure A4. Complement to Figure 1. (A) The date of peak prevalence of infection (centered) correlated with severity of epidemics, indexed by standardized regression slopes ($\hat{\beta}$) fit between infection prevalence and host density (where shading denotes regions with positive $\hat{\beta}$). (B) This date of peak prevalence also correlated negatively with maximum prevalence of infection reached. Larger epidemics that more virulently depressed host density peaked earlier in the season than small/less severe epidemics. (C) Epidemics that started earlier peaked earlier.

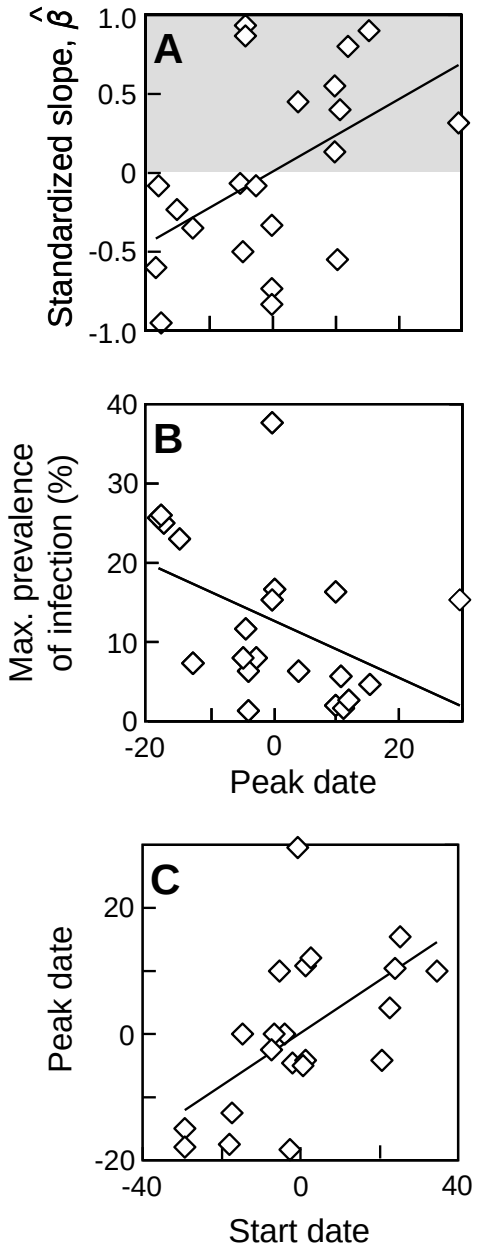

Figure A5. Prevalence increases with density during some of the early stages of several fungal (*Metschnikowia*) epidemics in populations of *Daphnia dentifera*. 2004 epidemics: (A) Bassett; (B) Bristol; (C) Warner; 2003 epidemic: (D) Baker. Arrows point in the direction of time.

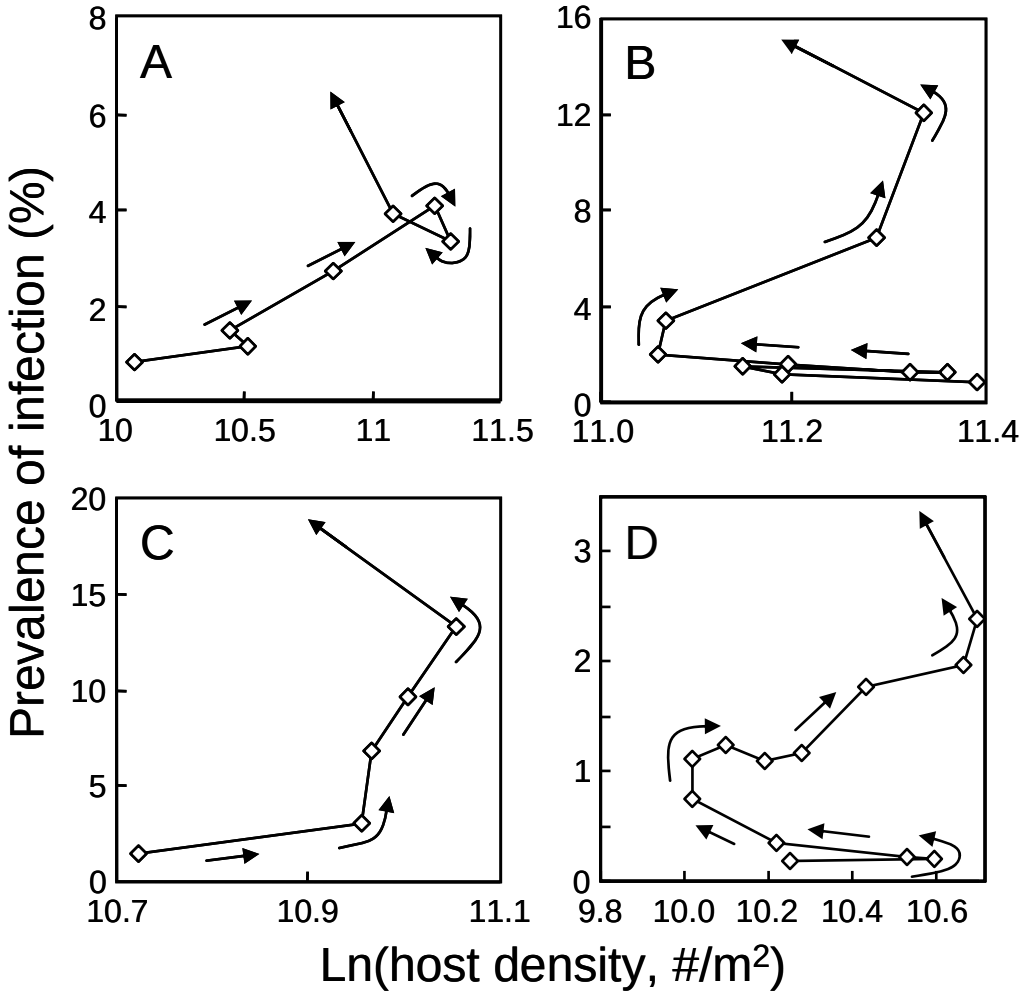

Figure A6. Results from the mecoscosm experiment showing negative density-dependence of egg ratios. (A) In both spore addition (top panel, dark symbols) and control (lower panel, light symbols) treatments, we found negative relationships between present host density (ln transformed) and egg ratio one sampling visit (3 days) ahead. Each symbol shape corresponds to a different replicate mesocosm. (B) This negative relationship between density and the key fecundity index, egg ratio, is seen in the intercept and slope summaries. The relationship is even steeper (higher intercept, steeper negative slope) in the spore addition treatments (grey bars).