local adaptation and host specificity to copepod intermediate hosts by the *Schistocephalus solidus* tapeworm

Authors: Kum C. Shim[1], Christopher Peterson[1], Daniel I. Bolnick[2]*

[1]Department of Integrative Biology, University of Texas at Austin, TX.
[2]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT

*Corresponding author: daniel.bolnick@uconn.edu

Abstract:

We investigated if there was local adaptation and host specify in the tapeworm *Schistocephalus solidus* to its copepod first intermediate hosts. The tapeworm is locally adapted and host specific to its threespine stickleback second intermediate host. We exposed copepods from five lakes in Vancouver Island (BC, Canada) to local (i.e. same lake) and foreign tapeworms in a reciprocal exposure experiment. Results indicate that the tapeworm is not locally adapted to the copepods, but there was host specificity as a copepod genus was more parasitized than another genus.

Introduction:

One of the most intriguing features of parasites with complex life cycles is their ability to infect several very disparate hosts during each of their life stages (Schmid-Hempel 2011). Transmission of these parasites (especially helminths) usually involves search for, and penetration of, their intermediate hosts. These parasites then passively infect their final hosts when the intermediate hosts are predated. Thus, complex-life-cycle parasites usually lower the overall fitness of their intermediate hosts (i.e. increased predation), and cannot be as selective on infecting

final hosts (Schmid-Hempel 2011, Poulin 2007, Noble et al. 1989). Accordingly, these parasites should be more host-specific to intermediate than to final hosts (Poulin 2007, Noble et al. 1989). Increased host specificity and negative fitness effects, imply that host-parasite coevolution and local adaptation may be more likely between parasites and their intermediate hosts (Lively et al. 2004).

Moreover, in host-parasite coevolution, the species with the higher dispersal rates is predicted to locally adapt to the other (Gardon and Nuismer 2009, Greischar and Koskella 2007, Morgan et al. 2005). This theoretical result is contrary to our usual expectation that dispersal and gene flow homogenize populations and counter-act divergent selection (Lenormand 2002). But in antagonistically interacting species, gene flow (within moderation) provides genetic diversity that aids in adapting to the opposing species (Gardon and Nuismer 2009). Parasite dispersal rates are usually higher than their hosts' (Mazé-Guilmo at al. 2016, Hoeksema and Forde 2008), so parasites should be more locally adapted to their hosts than vice-versa. Moreover, hermaphroditic parasites can have higher reproductive success (i.e., it can fertilize other individual's eggs and at the same time receive sperm to fertilize its eggs) and higher dispersion rates, both of which can lead to increased local adaptation (Mazé-Guilmo et al. 2016, Hoeksema and Forde 2008).

Combining all the propositions above (i.e., hosts-specificity, negative fitness effects on their intermediate hosts, and higher dispersal rates), parasites with complex life cycles, especially those that are hermaphroditic, should be often locally adapted to their intermediate hosts.

We tested the above prediction using the hermaphroditic tapeworm *Schistosoma solidus* (Eucestoda: Pseudophyllidea) and its first intermediate hosts, freshwater cyclopoid copepods. This tapeworm is found mainly in Holarctic lakes. It has copepods and threespine sticklebacks (*Gasterosteus aculeatus*) as first and second intermediate hosts (Barber and Scharsack 2009,

Dubinina 1980) and the final hosts are warm-blooded vertebrates, usually fish-eating birds. The tapeworm reproduces sexually in the finals hosts' intestines and its eggs are dispersed with these hosts' feces, so the tapeworm has higher dispersal rates than its first two intermediate hosts which rarely disperse between even adjacent lakes (Caldera and Bolnick 2008). The tapeworm can be bred in-vitro, making it an excellent laboratory system for host-parasite studies (Barber 2013, Barber and Scharsack 2009, Smyth 1990). The tapeworm is not host specific to its final hosts, infecting several species of birds and even fish-eating mammals like otters (Hoberg et al. 1997, Dubinina 1980). However, the tapeworm is very host-specific to the stickleback (Barber 2013, Dubinina 1980). The tapeworm affects negatively the fitness of the fish (Weber et al. 2017b and references therein), and is locally adapted to this host (Hafer 1017, Kalbe et al. 2016). In laboratory infections, this tapeworm had negative fitness consequences to lab-reared *Macrocyclops albidus* copepods (Benesh 2010, Wedekind 1997); however, no work has been done on wild copepod species that are sympatric with the tapeworm to establish host-specificity and local adaptation, as has been done with stickleback.

We anticipate that this tapeworm would be similarly host specific and locally adapted to their copepod hosts as in their stickleback host. To test this hypothesis, we used reciprocal infection trials using factorial combinations of *S. solidus* tapeworms and native copepod species collected from lakes on Vancouver Island. We measured local adaptation through mainly infection rates in copepods by local (same lake) and foreign (different lake) tapeworms and by intensity (number of parasites inside hosts) in the infected copepods. To measure host specificity, we infected different copepod genera with the tapeworm and measured infection success in each genus. Results indicate that there was no local adaptation by the tapeworm in the copepods, but there was host specificity

as a specific crustacean genus had overall higher infection rates than another used in this experiment.

Material and methods:

Copepod colonies:

We used copepods from established laboratory colonies from five lakes on Vancouver Island (Boot, Echo, Gosling, Lawier, and Roberts Lakes. The coordinates for these lakes are in supplementary Table 1). These colonies were established from plankton tows collected on September 15, 2017, and June 24, 2018. Colonies were kept in five gallon buckets at 20ºC and under 16:8 hrs light:dark to simulate summer conditions in Vancouver Island until the start of the experiment on October 20 2018. We fed copepods in each bucket weekly with ~500mL of *Paramecium caudatum* and mixed rotifer cultures plus a ground protozoan pellet, both from Carolina Biological Supply Company (Burlington, NC). We also added 10-20 autoclaved wheat seeds once a month to each bucket for bacterial growth, which contributed to the copepod and paramecium diets. Before the start of the experiment, we identified each lake's copepods to species level under a dissecting scope and using the Image-Based Key to the Zooplankton of North America (Haney 2013).

The laboratory colonies for each lake only had one surviving copepod species just before the start of the experiment. These were *Macrocyclops albidus* for Boot and Lawier Lakes, *Macrocyclops fuscus* for Roberts Lake, *Acanthocyclops robustus* for Echo Lake, and *Acanthocyclops brevispinosus* for Gosling Lake. All these copepods were from the order Cyclopoida.

Tapeworm colonies:

We used tapeworm eggs from three lakes in Vancouver Island (Boot, Echo, and Gosling Lakes). Lawier and Roberts Lakes lack infected stickleback fish, so tapeworms were unavailable from these two lakes; thus, we infected copepods from five lakes with tapeworms from three. The advantage of this design is that copepods from Lawier and Roberts lakes could be highly susceptible to the tapeworm due to their lesser exposure to the parasites; thus, serving as positive controls. The tapeworm eggs were collected from laboratory crosses of randomly chosen wild tapeworms obtained from infected fish, following established methods (Weber et al. 2017b, Smyth 1990). We hoped that these randomly chosen tapeworms would reflect the tapeworm genetic diversity in each lake. These crosses were done in June – Sept. 2018, and the eggs were kept at 4^{o}C until the experiment.

Experimental set-up:

To test for tapeworm local adaptation and host specificity to copepods, we carried out a reciprocal infection experiment by exposing the copepods from each lake to local and foreign tapeworm larvae (coracidia) from three lakes (i.e., Boot, Echo, and Gosling lakes). We hatched tapeworm eggs and exposed the coracidia to copepods following published methods (Weber et al. 2017b, Smyth 1990). We used six-well plates, each well holding a different combination of copepods (n=10 individuals per well) from a lake and tapeworms (n=20 coracidia per well) from the same or a different lake (Figure 1). We used a combination of 1:2 copepod to tapeworm ratio to account for the short lifespan (~24hr) of the parasite (Dubinina 1980). We used three tapeworm families or strains per lake. We also had six to eight wells per lake with copepods unexposed to tapeworms as negative controls to measure tapeworm exposure and infection effects on host

mortality (supplementary table 2). The plates were kept in the same conditions as the copepod colonies (i.e. 20°C and 16:8hrs light:dark). We randomized the positions of the copepod-tapeworm combinations within plates, and plate locations within the incubator. We dissected each surviving copepod to ascertain infection status 17-22 days post exposure when tapeworms reached maximum size inside copepods (Dubinina 1980).

In total, we used 49 6-well plates, exposing 2,890 copepods (10 per well) with 5,780 tapeworms (20 per well, nine families in total, three per lake. See supplementary table 2). At the end of experiment, 1622 exposed and 330 control copepods survived. Exposure to tapeworms did not affect copepod survival (P value = 0.996, supplementary figure 2). The survival rate for copepods in the experiment was 56%.

Bayesian analysis (data analysis):

We used mixed-effect hurdle models to simultaneously estimate the effect of copepod and parasite origin on infection rate (prevalence) and intensity (number of worms per successfully infected copepod). Conceptually, these models combine a logistic regression on parasite presence/absence with a truncated Poisson regression on non-zero parasite counts. Our models considered tapeworm lake and its interaction with either copepod genus or lake as fixed effects; we also included an indicator for whether the tapeworm and copepod were from the same lake (i.e. "native"). Plate number and tapeworm lake were included as random effects. The full model contains all of these terms as predictors for both prevalence and incidence. We created a series of reduced models from a list of all possible combinations of predictors, excluding models that contained interactions without their main effects, copepod lake without genus, and tapeworm family without lake.

We fit all models with the *brms* package in R v. 4.0.4 (Bürkner 2018, R Core Team 2018). The predictive value of each model was determined with Bayesian stacking weights calculated by the *loo* package (Yao et al. 2018); conceptually, this is similar to AIC model weighting. A combined ensemble was created by pooling a weighted sample of each model's posterior distribution. We defined effect sizes as the standard deviation of a term's marginal effects at each posterior sample from a model where the term was included. Prior specification and other details are provided in the supplementary material section.

Results:

In this paper we tested if the tapeworm *S. solidus* is locally adapted and host specific to their copepod hosts. My results indicate that the tapeworm is not locally adapted to the copepods (figure 2), but it might be more host specific to a genus of copepods as rates of infection and intensity (number of parasites inside infected hosts) were higher in *Acanthocyclops* than *Macrocyclops* copepods (figures 3 to 5).

Results from the Bayesian analysis:

The ensemble mixed-effect hurdle model contained 1.37 million posterior samples, with 4,450 different models contributing at least one sample. No single model had a stacking weight higher than 0.3%; however, both copepod and tapeworm origins contributed to over 80% of both the intensity and infection rate model components (Table 1). For both model components, copepods had the largest effect size of any term (intensity: 0.640 [0.421, 1.289]; infection rate: 0.255 [0.213, 0.364]; brackets signify 95% credible interval). Copepod effects can be decomposed into genus and lake of origin, with lake nested in genus; 61% of posterior samples with copepod

genus terms also contained copepod lake. Copepod genus effect sizes were generally smaller when lake effects were also present (intensity effect sizes: 0.712 [0.518, 0.958] without lake, 0.342 [0.011, 2.714] with lake; infection rate: 0.338 [0.303, 0.375] without lake, 0.210 [0.031, 0.395] with lake; supplementary figure 1).

Copepod by tapeworm interactions (the typical test for local adaptation) had the lowest ensemble inclusion frequencies for both the intensity and infection rate model components (Table 1), and their effect sizes when present had wide, noisy posterior distributions. The 'native' effect (indicating copepods and tapeworms from the same lake) had higher inclusion but consistently small effect sizes; we interpreted this as insufficient evidence for local adaptation. All of these effects had lower inclusion rates than the 6-well plates used for in the experiment.

We also ran mix-effect linear and GLM models in R (R Core Team) to supplement the analyses and results above. For these analyses, the best predictors for infection rate were the copepod and tapeworm lakes, and the best predictors for intensity in infected copepods was copepod lake. These results did not differ considerably from the best Bayesian mixed-effect hurdle models above, suggesting our results are robust to either choice of analytical method. For more details on the mix-effect and GLM models and results, see supplementary material.

As mentioned in the Bayesian results section, there was not enough evidence for local adaptation of the tapeworm to their copepod hosts. This can also be seen in figure 2, where infection rates by the tapeworm on local (from the same lake) and foreign (from different lakes) copepods were very similar. However, there was evidence of host specificity as copepod genus was a strong predictor in infection rate and infection intensity in the crustacean. For example, copepods from Echo and Gosling lakes (both of the genus *Acanthocyclops*) were three to six times more susceptible to infection than the other copepod genus (*Marcocyclops*) from the three

188 remaining lakes (figure 3). This was true for all tapeworm strains used (figure 4). Moreover, the
189 infected copepods from Echo and Gosling lakes (again both of the genus *Acanthocyclops*) also had
190 between 0.3 to 0.5 times more tapeworms than those (of the genus *Marcocyclops)* from the other
191 three lakes (figure 5). This accounts for the relatively high effect sizes of the copepod genus factor
192 in the Bayesian analysis.
193

194 **Discussion:**

195 We tested for local adaptation and host specificity of the tapeworm *S. solidus* from three
196 lakes in Vancouver Island to copepods from the same plus two more lakes where the tapeworm is
197 absent (Weber et al. 2017b, personal observations). Researchers argue that parasites with complex
198 life cycles should be more host-specific (Poulin 2007, Nobel et al. 1989), and that parasites with
199 higher dispersal rates should locally adapt to their hosts (Barber and Scharsack 2009, Morgan et
200 al. 2009). Thus, the *S. solidus* tapeworm, being a parasite with a complex life cycle and having
201 higher dispersal rates than their intermediate hosts (Dubinina 1980), should show local adaptation
202 and host specificity to its copepod hosts in a similar fashion to the tapeworm's second intermediate
203 host (i.e. threespine sticklebacks [Hafer 2018, Weber et al. 2017a, Kalbe et al. 2016]).
204 However, our results indicate that there was no evidence of differences between infection
205 rates by local and foreign tapeworms on the copepods (figure 2). Our experiment also shows that
206 copepods from Echo and Gosling Lakes (genus *Acanthocyclops*) were more susceptible to *S.*
207 *solidus* tapeworm infection than the ones from the other three lakes (genera *Macrocyclops*) (figure
208 3), and these copepods also had slightly more tapeworms when infected (figure 5). These infection
209 and intensity rates were very similar among the different tapeworm strains from the three lakes

used (figures 4 and 5). Thus, at least for this parasite-host system, we did not observe local adaptation by the tapeworm to the copepods.

Instead, the success of the tapeworm within a given lake depended mostly on whether a copepod genus (*Acanthocyclops*) was present. Variation in zooplankton community structure, between lakes, means that tapeworms will be locally maladapted to lakes with *Macrocyclops* spp copepods. The higher susceptibility of Echo and Gosling Lakes' copepods (of the genus *Acanthocyclops*) to the tapeworm explains why copepod and tapeworm lake variables in our models fit most the data.

To emphasize more the lack of local adaptation in our experiments, Boot and Lawier Lakes had the same species of copepods (*Macrocyclops albidus*), but both lakes' copepods had very similar infections rates by the three strains of tapeworms used (figure 4). Specifically, Boot Lake tapeworms are no more (or less) effective at infecting Boot Lake *M. albidus* than they are at infected Lawier Lake *M. albidus* (a home-versus-away criterion for local (mal)adaptation). Nor are the Boot Lake tapeworms any better (or worse) at infecting their native Boot Lake copepods, relative to tapeworms from two other lakes (a native versus immigrant criterion for local (mal)adaptation). Thus, for both lakes, the infection rate by local tapeworms was not significantly different to that of foreign tapeworms.

Local adaptation aside, our experiments show the tapeworm is clearly capable of infecting multiple copepod genera, but it is most efficient at infecting a particular genus. The copepods with the highest infection rates (those from Echo and Gosling lakes) were from the same genus (i.e. *Acanthocyclops*). This was true regardless of whether the tapeworms were taken from a lake dominated by *Acanthocyclops*, or not. Currently, *M. albidus* copepods are used for experimental infections in sticklebacks (Weber et al. 2017a and 2017b, Barber 2013, Benesh 2010, Wedekind

1997, Smyth 1990); perhaps, future work should employ *Acanthocyclop* species instead to maximize resources and time for better results. Weber et al. (Weber et al. 2017a) argued that to understand the patchiness of the tapeworm infections in stickleback populations, more data is needed on ecological processes like parasite encounter rates and abundance of suitable primary hosts (copepods). Although, the primary reason for the different stickleback infection levels in the lakes sampled was due to recent evolution of the fish's immunology (Weber et al. 2017b), copepod infectivity might still play a role. We sampled in the same lakes for this work, so we can comment on the stickleback infection levels to our copepod infection levels. The high copepod infection levels in Gosling and Echo Lakes might contribute to the high stickleback infection levels in these lakes (figs 1 in Weber et al. 2017a and 2017b). And even if wild sticklebacks in Roberts Lake lack tapeworms, our experiments here show that this lake's copepods can get infected, validating the hypothesis that this lake's fish are exposed to tapeworms in the wild but still have zero infections due to recently evolved immunological mechanisms to combat infections (Weber et al. 2017b).

Table 1: Inclusion frequencies and effect sizes of each term in the ensemble model for intensity (tapeworm count in infected copepods) and infection rate (prevalence). Effect sizes are provided as medians with 95% credible intervals and were calculated over the portion of the ensemble posterior where terms were present. Pooled terms indicate the combined effects of all copepod or tapeworm terms that were present in the model. Random effects are noted with (RE). Effect sizes for intensity and infection rate should not be compared, as they are in different units (counts and proportions, respectively).

Model Component[1]	Term[2]	Ensemble Frequency[3]	Effect Size[4]	[95% CI][5]
Intensity	Copepod (pooled)	0.849	0.640	[0.421, 1.289]
Intensity	Copepod Genus	0.849	0.621	[0.019, 2.201]
Intensity	Worm (pooled)	0.817	0.187	[0.027, 1.297]
Intensity	Worm Lake	0.817	0.182	[0.023, 1.374]
Intensity	Copepod Lake	0.522	0.503	[0.190, 2.464]
Intensity	Plate (RE)	0.500	0.095	[0.004, 0.421]
Intensity	Native	0.473	0.126	[0.005, 0.798]
Intensity	Worm Family (RE)	0.420	0.113	[0.006, 0.575]
Intensity	Genus x Worm Lake Interaction	0.354	0.616	[0.052, 4.309]
Intensity	Copepod x Worm Lake Interaction	0.208	0.721	[0.079, 7.482]
Infection Rate	Copepod (pooled)	0.871	0.255	[0.213, 0.364]
Infection Rate	Copepod Genus	0.871	0.274	[0.043, 0.385]
Infection Rate	Worm (pooled)	0.840	0.058	[0.020, 0.118]
Infection Rate	Worm Lake	0.840	0.060	[0.018, 0.128]
Infection Rate	Copepod Lake	0.593	0.143	[0.072, 0.242]
Infection Rate	Plate (RE)	0.500	0.032	[0.002, 0.078]
Infection Rate	Worm Family (RE)	0.409	0.018	[0.001, 0.060]
Infection Rate	Native	0.372	0.039	[0.002, 0.142]
Infection Rate	Copepod x Worm Lake Interaction	0.235	0.095	[0.045, 0.164]
Infection Rate	Genus x Worm Lake Interaction	0.218	0.053	[0.009, 0.175]

[1] For Model Component; Infection Rate is the Prevalence of infection in copepods by the tapeworm. Intensity is the number of tapeworms inside infected copepods.

[2]Factor indicates the model term; Copepod (pooled) and Worm (pooled) refer to the combination of whatever copepod/worm related terms were present in each model.

[3]Frequency the proportion of the model ensemble that the term appears in.

[4]Effect size of each model term accounting for the data.

[5]CI: 95% credible interval of the effect size.

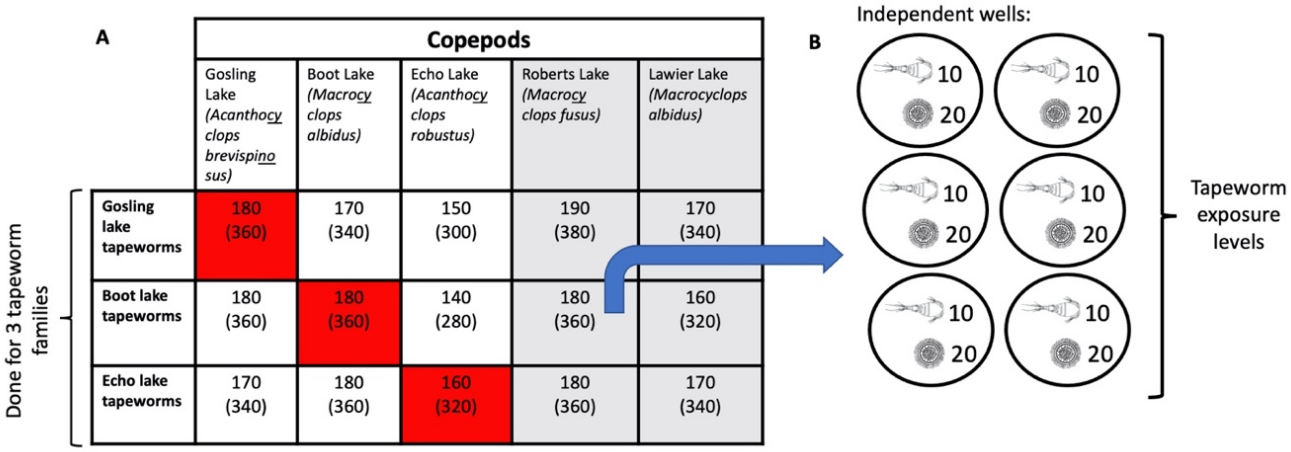
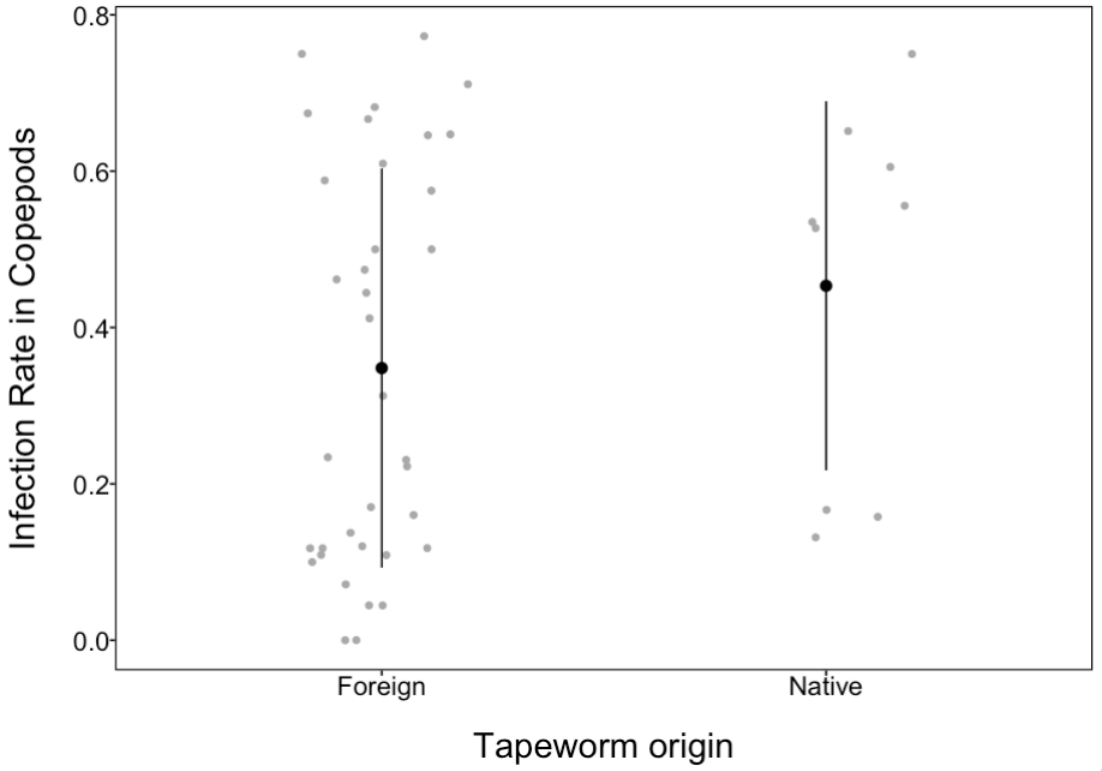


Figure 1: Graphical representation of the experiment setup. **A)** The combinations of the tapeworm *Schistocephalus solidus* by copepod exposures, using three tapeworm families per lake; red squares indicate tapeworms exposed to sympatric copepods. Roberts and Lawier Lakes are shaded in grey representing control lakes where the tapeworm is lacking in threespine sticklebacks. The numbers inside each square represent total numbers of copepods and tapeworms used (the latter in parenthesis). Names of the copepod species used are below each lake's names. **B)** A diagram of how each tapeworm family was exposed to each lake's copepods (in this example Boot Lake tapeworms to Roberts Lake copepods): in six different wells from different 6-well plates, each with 10 copepods exposed to 20 tapeworm larvae. All well positions for all exposures in panel A were randomized in the 6-well plates, and the position for each 6-well plates were also randomized in the experimental room.

280



281

282 **Figure 2.** Overall, infection rates on copepods by local or native tapeworms (i.e. where the *S.*
283 *solidus* tapeworms are from the same lakes as the copepods) are very similar to that of foreign
284 tapeworms.

285

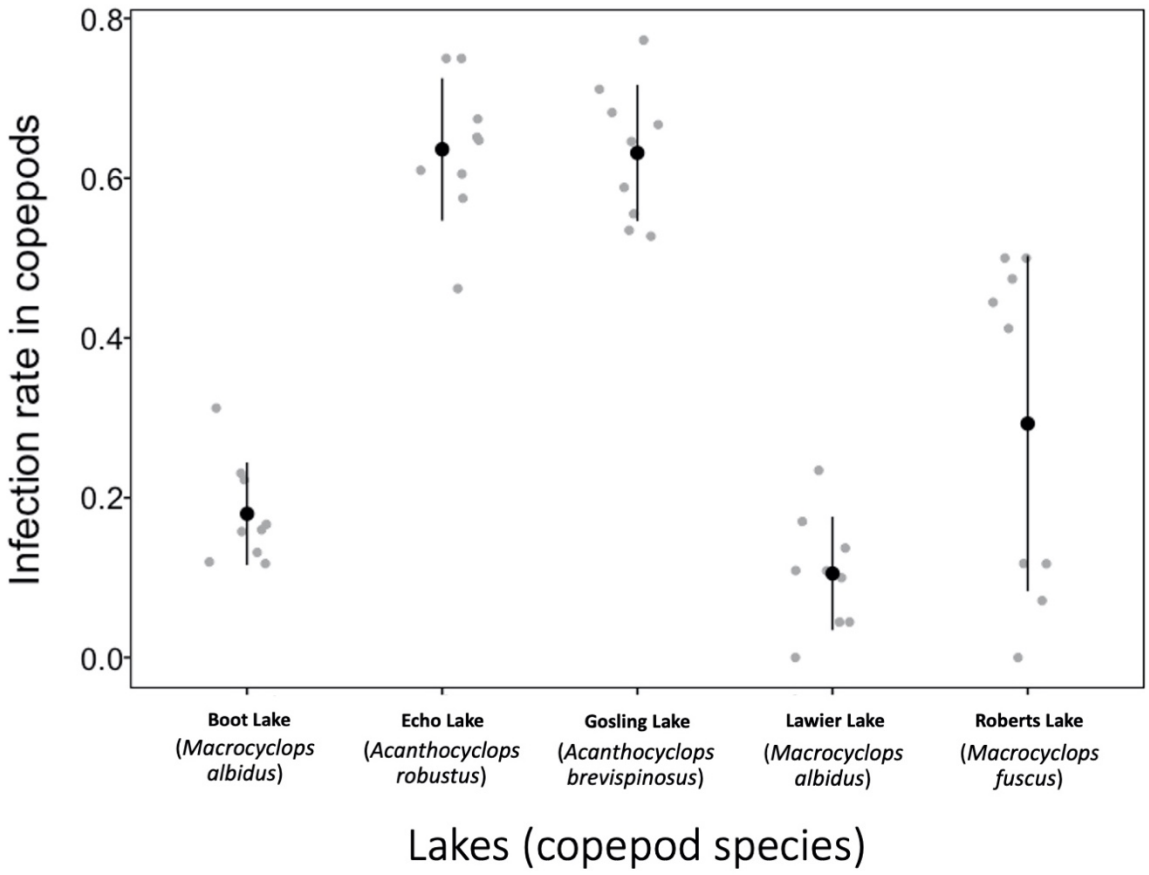


Figure 3: Infection results indicate that copepods from Echo and Gosling Lakes were three to six times more susceptible to infection by the tapeworm than those of the other three lakes. The scientific names of the copepods from each lake are in parenthesis under the lake names. The tapeworm is not found in Lawier and Roberts lakes (at least from stickleback fish surveys).

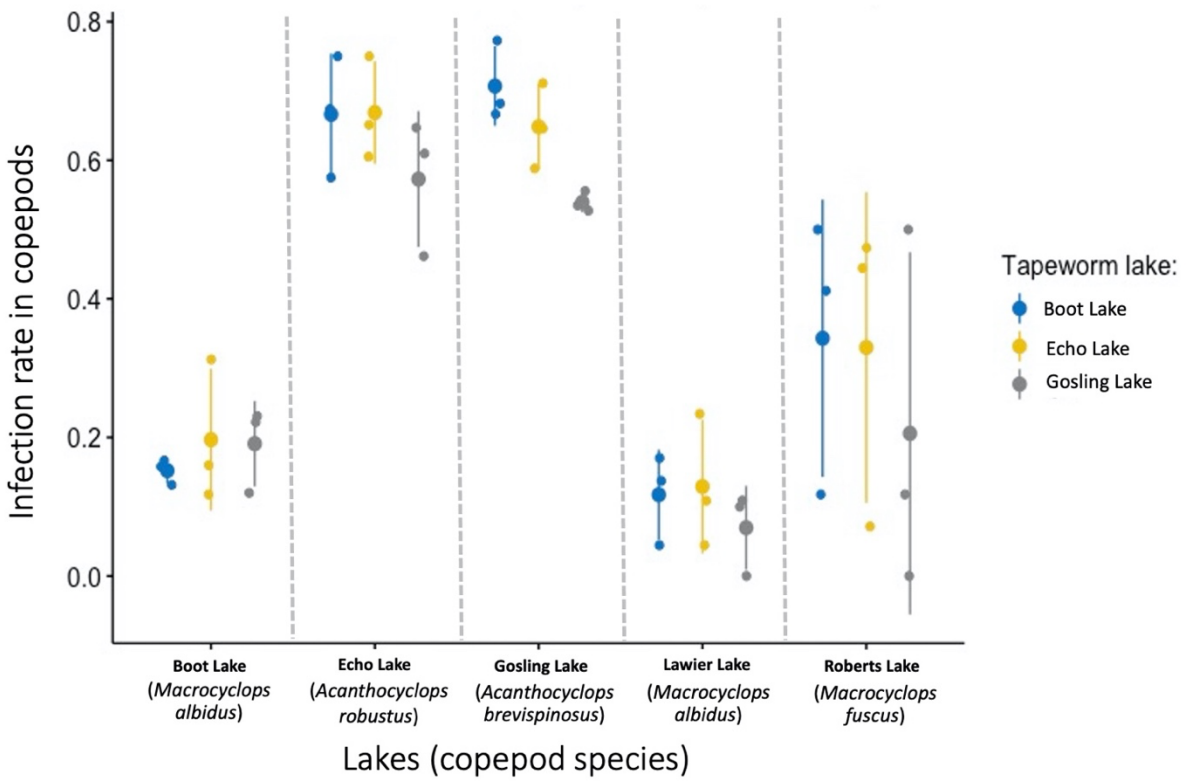


Figure 4. The infection rates in copepods were similar among all the tapeworm families or strains from the three lakes used. Again, copepods from Echo and Gosling Lakes were three to six times more susceptible to infection than those of the other three lakes (see figure 2). The scientific names of the copepods from each lake are in parenthesis under the lake names. The tapeworm is not found in Lawier and Roberts lakes (at least from stickleback fish surveys).

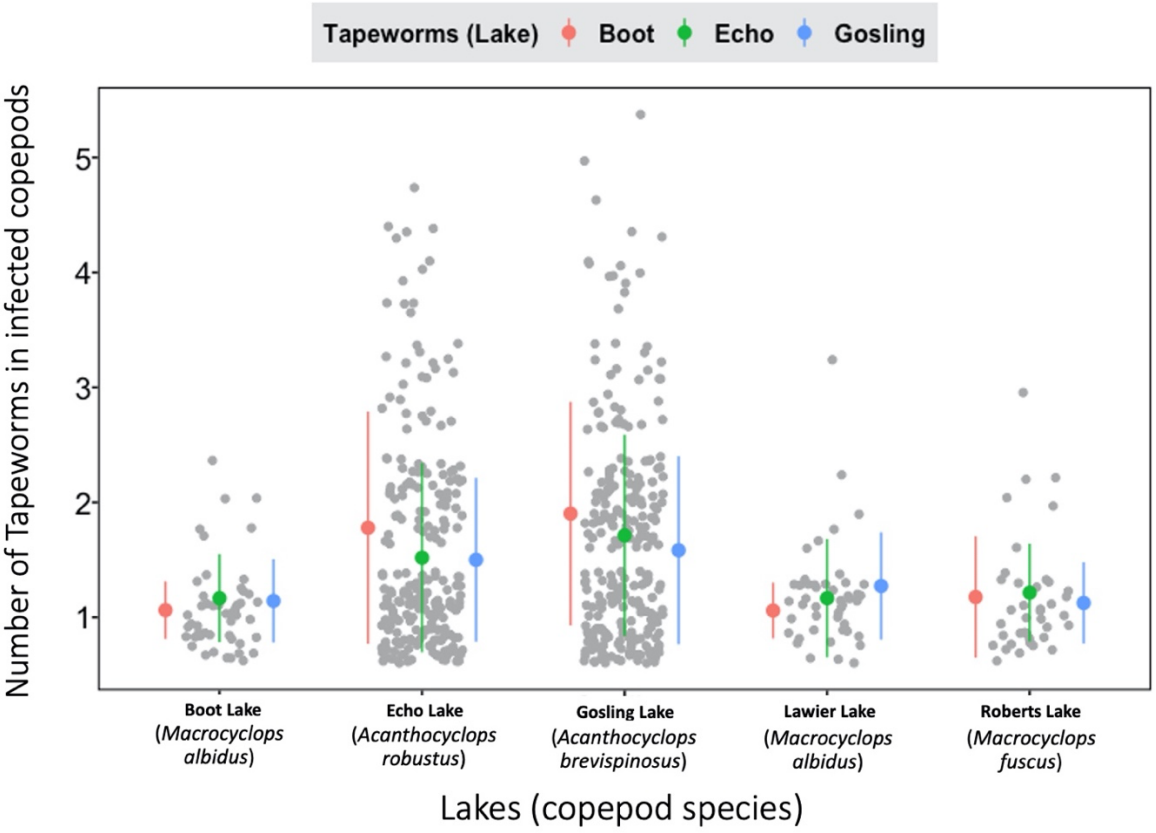


Figure 5. The infected copepods from Echo and Gosling Lakes also had slightly more parasites on average than the ones from the other three lakes. Again, the averages were very similar in the three tapeworm strains from the three lakes used. The scientific names of the copepods from each lake are in parenthesis under the lake names. The tapeworm is not found in Lawier and Roberts lakes (at least from stickleback fish surveys).

References:

Barber, I. 2013. Sticklebacks as model hosts in ecological and evolutionary parasitology. Trends in Parasitology, 29: 556-566.

Barber, I. and J. P. Scharsack. 2009. The three-spined stickleback- Schistocephalus solidus system: an experimental model for investigating host-parasite interactions in fish. Parasitology, 137: 411-424.

Benesh, D. P. 2010. What are the evolutionary constraints on larval growth in a trophically transmitted parasite? Oecologia, 162: 599-608.

Bürkner P. C. 2018. Advanced Bayesian Multilevel Modeling with the R Package brms. The R Journal. 10(1), 395-411. doi.org/10.32614/RJ-2018-017

Caldera, E.J. and Bolnick, D. I. 2008. Effects of colonization history and landscape structure on genetic variation within and among threespine stickleback (Gasterosteus aculeatus) populations in a single watershed. Evolutionary Ecology Research, 2008: 575-598.

Dubinina M. N. Tapeworms (Cestoda,Ligulidae) of the fauna of the USSR (Translated from Russian). New Delhi: Amerind Publishing Co. Pvt. Ltd., 1980.

Gandon, S., and Nuismer, S.L. 2009. Interactions between Genetic Drift, Gene Flow, and Selection Mosaics Drive Parasite Local Adaptation. The American Naturalis, 173: 212-224.

Greischar, M.A. and Koskella, B., 2007. A synthesis of experimental work on parasite local adaptation. Ecology letters, 10:418-434.

Hafer, N., 2018. Differences between populations in host manipulation by the tapeworm Schistocephalus solidus–is there local adaptation? *Parasitology*, *145*(6), pp.762-769.

Haney, J.F. et al. "An-Image-based Key to the Zooplankton of North America" version 5.0 released 2013. University of New Hampshire Center for Freshwater Biology <cfb.unh.edu> 1 Oct 2019.

Hoberg, E. P., Henny, C. J., Hedstrom, O. R. and Grove, R. A. 1997. Intestinal helminths of river otters (Lutra canadensis) from the Pacific Northwest. Journal of Parasitology, 83: 105-110.

Hoeksema, J.D. and Forde, S.E., 2008. A meta-analysis of factors affecting local adaptation between interacting species. The American Naturalist, 171(3), pp.275-290.

Kalbe, M., Eizaguirre, C., Scharsack, J.P. and Jakobsen, P.J., 2016. Reciprocal cross infection of sticklebacks with the diphyllobothriidean cestode Schistocephalus solidus reveals consistent population differences in parasite growth and host resistance. Parasites & vectors, 9(1), pp.1-12.

Lenormand T. 2002. Gene flow and the limit to natural selection. Trends in Ecology and Evolution, 17: 183-189

Lively, C.M., M.F. Dybdahl, J. Jokela, E. Osnas, L.F. Delph. 2004. Host sex and local adaptation by parasites in a snail-trematode interaction. American Naturalist, 164:S6-S18.

Mazé-Guilmo, E., Blanchet, S., McCoy, K.D. and Loot, G., 2016. Host dispersal as the driver of parasite genetic structure: a paradigm lost? Ecology Letters, 19(3), pp.336-347.

Morgan, A.D., Gandon, S., and Buckling, A. 2005. The effect of migration on local adaptation in a coevolving host–parasite system. Nature, 437:253-256.

Noble, E. R., Noble, G. A., Schad, G. A., and MacInnes, A. J. 1989. Parasitology: The Biology of Animal Parasites, 6th ed. Lea & Febiger, Philadelphia

Poulin, R., 2007. Evolutionary ecology of parasites. Princeton University Press, New Jersey.

R Core Team. 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Schmid-Hempel, P., 2011. Evolutionary parasitology. Oxford University Press, Oxford, UK.

Smyth, J. D. 1990. In vitro Cultivation of Parasitic Helminths. CRC Press, Boca Raton, FL.

Weber, J. N., M. Kalbe, K. C. Shim, N. I. Erin, N. C. Steinel, L. Ma, and D. I. Bolnick. 2017a. Resist globally, infect locally: a transcontinental test of adaptation by stickleback and their tapeworm parasite. The American Naturalist, 189: 43-57.

Weber, J. N., N. C. Steinel, K. C. Shim, and D. I. Bolnick. 2017b. Recent evolution of extreme cestode growth suppression by a vertebrate host. Proceedings of the National Academy of Sciences, 114: 6575-6581.

Wedekind, C. 1997. The infectivity, growth, and virulence of the cestode Schistocephalus solidus in its first intermediate host, the copepod Macrocyclops albidus. Parasitology, 115: 317-324.

Yao, Y., A. Vehtari, D. Simpson, and A. Gelman. 2018. Using Stacking to Average Bayesian Predictive Distributions (with Discussion). Bayesian Analysis. 13 (3): 917-1007

Supplementary material

Supplementary table 1: copepod collection dates and locality:

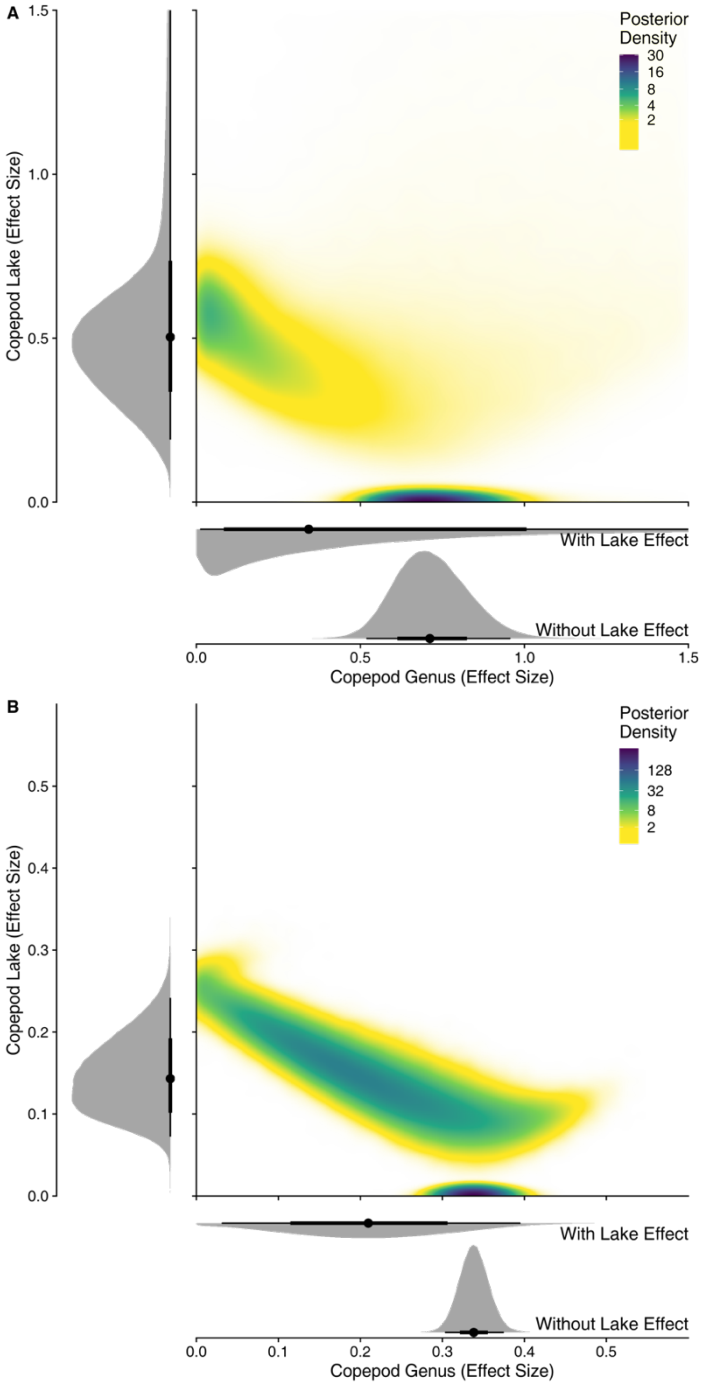
Lake	Longitude	Latitude	Collection date
Boot Lake	50.05503	-125.526	9/16/17, 6/24/18
Echo Lake	49.98765	-125.411	9/15/17, 6/24/18
Lawier Lake	50.083144	-125.515052	9/15/17, 6/24/18
Roberts Lake	50.216390	-125.544687	9/15/17, 6/24/18
Gosling Lake	50.04592	-125.501	9/16/17, 6/24/18

Supplementary table 2: Tapeworm families used for the experiments and number of exposures (i.e. wells in a 6-well plate) for each lake's copepods (reminder: each well had 10 copepods exposed to 20 tapeworms):

Tapeworm families[1]	Gosling Lake copepods	Roberts Lake copepods	Lawier Lake copepods	Echo Lake copepods	Boot Lake copepods
Control	6	6	6	8	7
Boot 11Bx1A (6/17/18)	6	6	6	5	6
Boot 2Ax2C (6/17/18)	6	6	5	4	6
Boot bulk (7/3/18)	6	6	5	5	6
Echo bulk (6/15/18)	6	6	6	5	6
Echo 3Ax1A (6/10/18)	6	6	6	5	6
Echo 27Ax31A (6/17/18)	5	6	5	6	6
Gosling 7Ax1A (6/10/18)	6	6	6	5	6
Gosling 10Ax12A (6/17/18)	6	7	5	5	5
Gosling 2 (9/22/18)	6	6	6	5	6

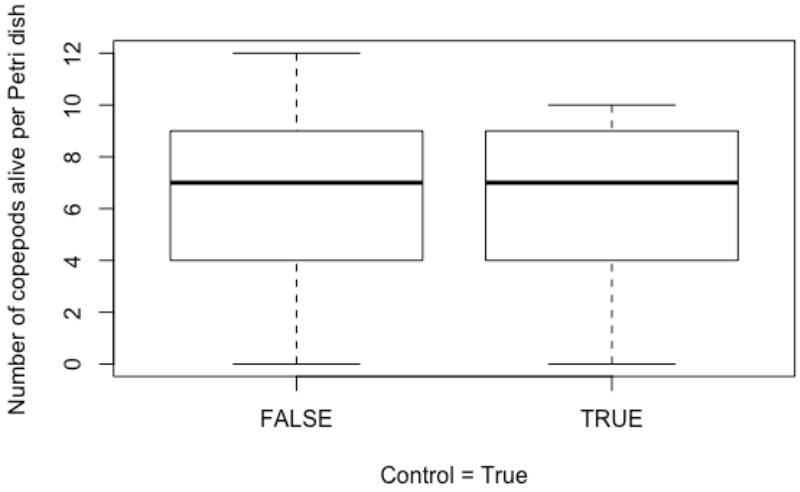
[1] Tapeworm families used in the experiment; 3 families per lake; dates in the parentheses below each family indicate the time the tapeworm eggs were harvested in the lab for the experiment. Control indicates no tapeworms were used to exposed the copepods (i.e. negative control); this is to evaluate the survivorship of the copepods in the wells during the experiment.

423



424

425 Supplementary figure **1**: Distribution of effect sizes for Copepod Genus and Copepod Lake for the
426 infection intensity (**A**) and rate (**B**) model components. Both panels include the marginal effects
427 of copepod lake (left), the marginal effects of copepod genus conditioned on whether lake was
428 included in the model (bottom), and their bivariate distributions (upper right). For both terms, the
429 genus effect increases when the lake effect declines or is absent; this is particularly notable for the
430 intensity component.



Supplementary figure **2:** The number of copepods alive after termination of experiment did not differ significantly between those exposed to the tapeworm and those that were not (i.e., control) [P value = 0.996, see more details of analysis in Supplementary mix-effect linear and GLM model analyses below]

Priors and iterations used in the mixed-effect hurdle analysis:

Each model was run for 4 chains with 1000 warmup and 1000 sampling iterations each. We checked model convergence by verifying N_eff > 1000, R-hat > 1.01, and all Hamiltonian Monte-Carlo diagnostics were acceptable.

Our prior distributions were Normal(mean = 0, sd = 6) priors for the intercepts of both model components. For incidence, we used Normal(0,1) priors for fixed effect coefficients and half-t(df = 7, mean = 0, scale = 1) priors for the standard deviation of the random effects. For the prevalence model, our priors were Normal(0, 1.5) for fixed effects and half-t(7, 0, 1.5) for the random effects. These priors were selected because they were flexible enough to allow for large effects but conservative enough to avoid spurious results.

Note: The complete R script for the mixed-effect hurdle analyses is in Christopher Peterson's GitHub (**https://github.com/Christopher-Peterson/copepod_worm_adapt**).

Supplementary mix-effect linear and GLM model analyses:

I) Analyzing if there was any difference on survival rate between copepods exposed to tapeworms and the not-exposed ones (i.e. control):

```
> model15 = glm(cop.alive ~ is_control, data = control_df)
> anova(model15, test = "LRT")
```

Analysis of Deviance Table

Model: Poisson, link: log

Response: cop.alive

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			291	2409.6	
is_control	1	0.0001767	290	2409.6	**0.9963**

Summary: not significant difference on number of copepods alive after experiment from control vs. exposed copepods (i.e. exposed to tapeworms)

> model14 = glm(cop.death.numb ~ is_control, data = control_df, family = poisson())
> anova(model14, test = "LRT")
Analysis of Deviance Table

Model: poisson, link: log

Response: cop.death.numb

Terms added sequentially (first to last)

	Df	Deviance	Resid. Df	Resid. Dev	Pr(>Chi)
NULL			291	684.6	
is_control	1	0.00056907	290	684.6	**0.981**

Summary: not significant difference on number of copepod deaths during experiment from control vs. exposed copepods (i.e. exposed to tapeworms)

II). GLM and GLMM analyses companion to the Bayesian analysis:

Using lme4 version 1.1-13 package for R

> summary(model1)
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [
glmerMod]
Family: binomial (logit)
Model 1 : infected.yes.no ~ cop.lake * worm.lake + (1 | plate)
Data: copepods

```
510
511       AIC       BIC  logLik    deviance   df.resid
512    1754.9    1841.1   -861.4    1722.9      1606
513
514    Scaled residuals:
515       Min         1Q     Median      3Q      Max
516    -1.6429    -0.4962  -0.3703    0.7090  3.7777
517
518    Random effects:
519     Groups Name        Variance   Std.Dev.
520     plate  (Intercept)  0.02598    0.1612
521    Number of obs: 1622, groups:  plate, 51
522
523    Fixed effects:
524                              Estimate Std. Error z value  Pr(>|z|)
525    (Intercept)              -1.7593     0.2789  -6.308   2.83e-10 ***
526    cop.lakeech               2.4600     0.3361   7.318 2.51e-13 ***
527    cop.lakegos               2.6596     0.3460   7.686 1.52e-14 ***
528    cop.lakelau              -0.1886     0.3737  -0.505  0.61378
529    cop.lakerob               1.0933     0.4011   2.726  0.00642 **
530    worm.lakeech              0.3491     0.3839   0.909  0.36318
531    worm.lakegos              0.3241     0.4025   0.805  0.42065
532    cop.lakeech:worm.lakeech  -0.3245     0.4721  -0.687  0.49182
533    cop.lakegos:worm.lakeech  -0.6611     0.4740  -1.395  0.16310
534    cop.lakelau:worm.lakeech  -0.3105     0.5246  -0.592  0.55388
535    cop.lakerob:worm.lakeech  -0.3525     0.5873  -0.600  0.54837
536    cop.lakeech:worm.lakegos  -0.7522     0.4870  -1.545  0.12246
537    cop.lakegos:worm.lakegos  -1.0517     0.4811  -2.186  0.02881 *
538    cop.lakelau:worm.lakegos  -0.9553     0.5728  -1.668  0.09537 .
539    cop.lakerob:worm.lakegos  -1.1049     0.6368  -1.735  0.08271 .
540    ---
541    Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
542
543    Correlation matrix not shown by default, as p = 15 > 12.
544    Use print(x, correlation=TRUE)  or
545              vcov(x)          if you need it
546
547    convergence code: 0
548    Model failed to converge with max|grad| = 0.00320527 (tol = 0.001, component 1)
549
550    > anova(model1,test="LRT")
551    Analysis of Variance Table
552                               Df   Sum Sq    Mean Sq   F value
553    cop.lake                   4   326.57     81.643   81.6425
554    worm.lake                  2    12.22      6.110    6.1102
555    cop.lake:worm.lake         8     6.20      0.775    0.7746
```

2) Doing another model without the interaction of copepod lakes and worm lakes:

```
> model2 <- glmer(infected.yes.no ~ cop.lake + worm.lake + (1|plate),
+                 data=copepods, family="binomial") #model 2 does not have an interaction

> #comparing models 1 and 2:
> anova(model1, model2)
Data: copepods
Models:
model2: infected.yes.no ~ cop.lake + worm.lake + (1 | plate)
model1: infected.yes.no ~ cop.lake * worm.lake + (1 | plate)

       Df  AIC    BIC    logLik   deviance  Chisq Chi Df Pr(>Chisq)
model2  8 1744.9 1788.0 -864.46  1728.9
model1 16 1754.9 1841.1 -861.43  1722.9     6.0436     8     0.6423
```

model2 without the interaction is slightly better! Though not significantly

3) model 6, testing for GLM without taking into account the random variable (no random effect)

```
> model6 <- glm(infected.yes.no ~ cop.lake + worm.lake + cop.lake*worm.lake, + data = copepods,family="binomial")
> summary(model6)

Call:
glm(formula = infected.yes.no ~ cop.lake + worm.lake + cop.lake *
    worm.lake, family = "binomial", data = copepods)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5645  -0.6639  -0.5187   0.8962   2. 2917

Coefficients:
                           Estimate Std. Error z value  Pr(>|z|)
(Intercept)                -1.7272     0.2713   -6.366  1.94e-10 ***
cop.lakeech                 2.4319     0.3292   7.388  1.49e-13 ***
cop.lakegos                 2.6027     0.3302   7.882  3.21e-15 ***
cop.lakelau                -0.2107     0.3704  -0.569  0.56938
cop.lakerob                 1.0912     0.3982   2.740  0.00614 **
worm.lakeech                0.3271     0.3780   0.865  0.38677
worm.lakegos                0.2921     0.3952   0.739  0.45973
cop.lakeech:worm.lakeech   -0.3263     0.4636  -0.704  0.48161
cop.lakegos:worm.lakeech   -0.6018     0.4568  -1.318  0.18765
cop.lakelau:worm.lakeech   -0.2863     0.5199  -0.551  0.58185
```

```
cop.lakerob:worm.lakeech  -0.3843     0.5788  -0.664  0.50670
cop.lakeech:worm.lakegos  -0.7143     0.4761  -1.500  0.13356
cop.lakegos:worm.lakegos  -1.0181     0.4707  -2.163  0.03056 *
cop.lakelau:worm.lakegos  -0.9051     0.5637  -1.606  0.10834
cop.lakerob:worm.lakegos  -1.0732     0.6296  -1.705  0.08828 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2167.1 on 1621  degrees of freedom
Residual deviance: 1723.3 on 1607  degrees of freedom
AIC: 1753.3

Number of Fisher Scoring iterations: 5

> anova(model6,test="LRT") #"LRT" is to get P-values here to check quickly for significance for
the variables in the model
Analysis of Deviance Table

Model: binomial, link: logit

Response: infected.yes.no

Terms added sequentially (first to last)


                         Df  Deviance Resid.  Df  Resid. Dev  Pr(>Chi)
NULL                                          1621   2167.1
cop.lake                 4      424.73        1617   1742.4    < 2.2e-16 ***
worm.lake                2       13.35        1615   1729.0      0.001265 **
cop.lake:worm.lake       8        5.71        1607   1723.3       0.6 79743
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Note: From the results above, there’re very significant effects of copepod lake and worm lake, but no interaction between them.

4) testing for effect of plates: (summary answer after running the model below: plate does not matter)

```
> model7 <- glm(infected.yes.no ~ cop.lake + worm.lake + cop.lake*worm.lake + plate,
+  data=copepods,family="binomial")
> anova(model6, model7) #does not give AIC comparisons, so this kind of useless
Analysis of Deviance Table

Model 1: infected.yes.no ~ cop.lake + worm.lake + cop.lake * worm.lake
```

```
Model 2: infected.yes.no ~ cop.lake + worm.lake + cop.lake * worm.lake +
    plate
  Resid.  Df     Resid. Dev     Df Deviance
1          1607    1723.3
2          1557    1658.6 50     64.754

> summary(model7)

Call:
glm(formula = infected.yes.no ~ cop.lake + worm.lake + cop.lake *
    worm.lake + plate, family = "binomial", data = copepods)

Deviance Residuals:
    Min       1Q        Median       3Q      Max
-2.1391  -0.7205  -0.4181      0.8657   2.7258

Coefficients:
                          Estimate   Std. Error   z value  Pr(>|z|)
(Intercept)            -3.400215   1.128677  -3.013   0.00259 **
cop.lakeech             2.648247   0.376858   7.027  2.11e-12 ***
cop.lakegos             3.271272   0.396900   8.242  < 2e-16 ***
cop.lakelau             0.004914   0.393591   0.012  0.99004
cop.lakerob             1.103968   0.423390   2.607  0.00912 **
worm.lakeech            0.547379   0.427087   1.282  0.19996
worm.lakegos            0.650357   0.444572   1.463  0.14350
platep1                1.423063   1.203531   1.182  0.23704
platep10               1.183440   1.187952   0.996  0.31915
platep11              -0.330903   1.234766  -0.268  0.78871
platep12               1.690147   1.195298   1.414  0.15736
platep13               1.281547   1.180913   1.085  0.27783
platep14               2.257173   1.174100   1.922  0.05455 .
platep15               1.166814   1.219485   0.957  0.33866
platep16               0.387203   1.170179   0.331  0.74073
platep17               1.225190   1.179514   1.039  0.29893
platep18               2.003874   1.230927   1.628  0.10354
platep19               0.919460   1.176703   0.781  0.43457
platep2                1.260738   1.188751   1.061  0.28889
platep20               0.943129   1.180947   0.799  0.42451
platep21               2.117863   1.184146   1.789  0.07369 .
platep22               0.870342   1.179558   0.738  0.46060
platep23               2.369486   1.170017   2.025  0.04285 *
platep24               1.636683   1.212048   1.350  0.17691
platep25               1.688014   1.198811   1.408  0.15911
platep26               1.536905   1.196642   1.284  0.19902
platep27               1.158677   1.194184   0.970  0.33191
platep28               1.765654   1.176215   1.501  0.13332
```

694 platep29 1.696665 1.231557 1.378 0.16831
695 platep3 1.741903 1.187559 1.467 0.14243
696 platep30 0.459434 1.208949 0.380 0.70393
697 platep31 0.996894 1.182753 0.843 0.39931
698 platep32 1.586515 1.188720 1.335 0.18199
699 platep33 1.394902 1.237889 1.127 0.25981
700 platep34 1.438833 1.154879 1.246 0.21281
701 platep35 0.365781 1.220012 0.300 0.76432
702 platep36 1.320014 1.189100 1.110 0.26696
703 platep37 1.164217 1.172606 0.993 0.32079
704 platep38 1.558819 1.183279 1.317 0.18771
705 platep39 1.482254 1.204915 1.230 0.21863
706 platep4 2.309692 1.206378 1.915 0.05555 .
707 platep40 1.950804 1.181498 1.651 0.09871 .
708 platep41 1.305689 1.223933 1.067 0.28606
709 platep42 Estimate Std. Error z value Pr(>|z|)
710 (Intercept) -3.3798150 1.1267879 -3.000 0.00270 **
711 cop.lakeech 2.6629963 0.3773947 7.056 1.71e-12 ***
712 cop.lakegos 3.2849437 0.3973952 8.266 < 2e-16 ***
713 cop.lakelau -0.0526042 0.3916659 -0.134 0.89316
714 cop.lakerob 1.1102455 0.4235788 2.621 0.00876 **
715 worm.lakeech 0.5649020 0.4274367 1.322 0.18630
716 worm.lakegos 0.6713093 0.4449921 1.509 0.13140
717 platep1 1.4001004 1.2023205 1.164 0.24422
718 platep10 1.1421783 1.1859406 0.963 0.33550
719 platep11 -0.3649309 1.2333211 -0.296 0.76731
720 platep12 1.6596767 1.1936453 1.390 0.16440
721 platep13 1.2550194 1.1795068 1.064 0.28732
722 platep14 2.2249946 1.1725085 1.898 0.05774 .
723 platep15 1.1187589 1.2173953 0.919 0.35811
724 platep16 0.3611196 1.1689989 0.309 0.75739
725 platep17 1.1901914 1.1777752 1.011 0.31224
726 platep18 1.9470174 1.2285226 1.585 0.11300
727 platep19 0.9021327 1.1759706 0.767 0.44300
728 platep2 1.2408435 1.1877500 1.045 0.29616
729 platep20 0.9219882 1.1799472 0.781 0.43458
730 platep21 2.1044216 1.1832326 1.779 0.07532 .
731 platep22 0.8400550 1.1781515 0.713 0.47583
732 platep23 2.3557240 1.1686849 2.016 0.04383 *
733 platep24 1.5905510 1.2096493 1.315 0.18855
734 platep25 1.6344116 1.1964111 1.366 0.17191
735 platep26 1.4913152 1.1944169 1.249 0.21182
736 platep27 1.1385838 1.1933126 0.954 0.34001
737 platep28 1.7516352 1.1750112 1.491 0.13603
738 platep29 1.6434851 1.2292294 1.337 0.18122
739 platep3 1.6953214 1.1854924 1.430 0.15270

```
platep30                  0.4280207  1.2074214   0.354  0.72297
platep31                  0.9763900  1.1818188   0.826  0.40870
platep32                  1.5603954  1.1872418   1.314  0.18874
platep33                  1.3499887  1.2353393   1.093  0.27448
platep34                  1.4182930  1.1539588   1.229  0.21905
platep35                  0.3351304  1.2185258   0.275  0.78329
platep36                  1.2740963  1.1869288   1.073  0.28307
platep37                  1.1311015  1.1709521   0.966  0.33406
platep38                  1.5187579  1.1813184   1.286  0.19857
platep39                  1.4487093  1.2032578   1.204  0.22859
platep4                   2.2814105  1.2049449   1.893  0.05831 .
platep40                  1.9408945  1.1805980   1.644  0.10018
platep41                  1.2763272  1.2225738   1.044  0.29650
platep42                  1.3990490  1.1997577   1.166  0.24357
platep43                  0.7191183  1.1960971   0.601  0.54769
platep44                  1.1219854  1.1798769   0.951  0.34164
platep45                  1.2439503  1.1811677   1.053  0.29227
platep46                  1.0435061  1.1943221   0.874  0.38227
platep47                  0.8521671  1.2247366   0.696  0.48656
platep48                  1.0256932  1.2030773   0.853  0.39390
platep49                  1.5725579  1.1901917   1.321  0.18641
platep5                   2.1948244  1.1726097   1.872  0.06124 .
platep50                  2.1848215  1.1983475   1.823  0.06827 .
platep6                   1.1904317  1.1870980   1.003  0.31595
platep7                   1.8411879  1.2012220   1.533  0.12533
platep8                   1.1430506  1.2024364   0.951  0.34180
platep9                   1.3928382  1.2060656   1.155  0.24815
cop.lakeech:worm.lakeech -0.1959257  0.5569642  -0.352  0.72501
cop.lakegos:worm.lakeech -1.4267929  0.5652912  -2.524  0.01160 *
cop.lakelau:worm.lakeech -0.4480476  0.5534265  -0.810  0.41818
cop.lakerob:worm.lakeech  0.0002931  0.6431409   0.000  0.99964
cop.lakeech:worm.lakegos -1.1915784  0.5580610  -2.135  0.03274 *
cop.lakegos:worm.lakegos -1.5504693  0.5540004  -2.799  0.00513 **
cop.lakelau:worm.lakegos -1.3602538  0.6179272  -2.201  0.02771 *
cop.lakerob:worm.lakegos -1.4027838  0.6821333  -2.056  0.03974 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2167.1  on 1621  degrees of freedom
Residual deviance: 1658.6  on 1557  degrees of freedom
AIC: 1788.6

Number of Fisher Scoring iterations: 5
```

```
> anova(model7, test = "LRT")
Analysis of Deviance Table

Model: binomial, link: logit

Response: infected.yes.no

Terms added sequentially (first to last)


                          Df  Deviance  Resid. Df  Resid. Dev   Pr(>Chi)
NULL                                        1621      2167.1
cop.lake                   4    424.73       1617      1742.4    < 2.2e-16 ***
worm.lake                  2     13.35       1615      1729.0     0.00 1265 **
plate                     50     54.14       1565      1674.9     0.319232
cop.lake:worm.lake  8      16.52       1557      1658.6     0.038042 *
 ---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Note: no significance in the plates, so plate does not have an effect in the infection outcome. But weird that there’s a slight significant interaction on copepod and worm lake interactions

5) testing for local adaptation

```
> model8 <- glm(infected.yes.no ~ cop.lake + worm.lake + native, + data=copepods, family="binomial")
> summary(model8)

Call:
glm(formula = infected.yes.no ~ cop.lake + worm.lake + native,
    family = "binomial", data = copepods)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5174  -0.6755  -0.5094   0.8875   2.2339

Coefficients:
                  Estimate Std. Error  z value  Pr(>|z|)
(Intercept)  -1.34434    0.18042  -7.451  9.24e-14 ***
cop.lakeech   2.11536    0.19279  10.972  < 2e-16 ***
cop.lakegos   2.08992    0.18921  11.045  < 2e-16 ***
cop.lakelau  -0.61487    0.22859  -2.690  0.00715 **
cop.lakerob   0.61262    0.25303   2.421  0.01547 *
worm.lakeech -0.01723    0.14308  -0.120  0.90415
worm.lakegos -0.44988    0.14882  -3.023  0.00250 **
nativeTRUE   -0.15688    0.14466  -1.084  0.27817
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2167.1  on 1621  degrees of freedom
Residual deviance: 1727.8  on 1614  degrees of freedom
AIC: 1743.8

Number of Fisher Scoring iterations: 4

                 Df  Deviance    Resid. Df  Resid. Dev    Pr(>Chi)
NULL                                  1621     2167.1
cop.lake         4    424.73          1617     1742.4    < 2.2e-16 ***
worm.lake        2     13.35          1615     1729.0     0.001073 **
native           1      1.18          1614     1727.8     0.278190
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Summary: no local adaptation

6) testing for effect of worm family used: (summary answer after running the model below: tapeworm family does not matter)

```
> model9 <- glm(infected.yes.no ~ cop.lake + worm.lake + cop.lake*worm.lake + worm.fam,
          + data=copepods,family="binomial")
> summary(model9)

Call:
glm(formula = infected.yes.no ~ cop.lake + worm.lake + cop.lake *
    worm.lake + worm.fam, family = "binomial", data = copepods)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5972  -0.6913  -0.4666   0.8940   2.3145

Coefficients: (2 not defined because of singularities)
                      Estimate Std. Error  z value  Pr(>|z|)
(Intercept)          -1.66090    0.30207  -5.498   3.83e-08 ***
cop.lakeech           2.43654    0.32947   7.395 1.41e-13 ***
cop.lakegos           2.60899    0.33064   7.891 3.00e-15 ***
cop.lakelau          -0.20226    0.37076  -0.546  0.58540
cop.lakerob           1.09678    0.39853   2.752  0.00592 **
worm.lakeech          0.42579    0.42204   1.009  0.31303
worm.lakegos          0.13006    0.44256   0.294  0.76885
worm.famboo2ax2c     -0.15686    0.24765  -0.633  0.52648
```

```
worm.famboobulk          -0.06496    0.24469  -0.265  0.79066
worm.famech27ax31a       -0.41496    0.25145  -1.650  0.09888 .
worm.famech3ax1a         -0.07461    0.24443  -0.305  0.76019
worm.famechbulk               NA         NA      NA       NA
worm.famg10ax12a          0.04772    0.25413   0.188  0.85104
worm.famg2                0.23203    0.25941   0.894  0.37107
worm.famg7ax1a                NA         NA      NA       NA
cop.lakeech:worm.lakeech -0.33392    0.46483  -0.718  0.47253
cop.lakegos:worm.lakeech -0.60388    0.45816  -1.318  0.18749
cop.lakelau:worm.lakeech -0.31071    0.52098  -0.596  0.55091
cop.lakerob:worm.lakeech -0.43645    0.58100  -0.751  0.45253
cop.lakeech:worm.lakegos -0.70794    0.47683  -1.485  0.13763
cop.lakegos:worm.lakegos -1.00586    0.47187  -2.132  0.03304 *
cop.lakelau:worm.lakegos -0.92188    0.56421  -1.634  0.10227
cop.lakerob:worm.lakegos -1.06786    0.63030  -1.694  0.09023 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2167.1  on 1621  degrees of freedom
Residual deviance: 1727.2 on 1599  degrees of freedom
AIC: 1773.2

Number of Fisher Scoring iterations: 13

> anova(model9, test = "LRT")
Analysis of Deviance Table

Model: binomial, link: logit

Response: infected.yes.no

Terms added sequentially (first to last)


                     Df  Deviance  Resid. Df  Resid. Dev     Pr(>Chi)
NULL                                   1621      2167.1
cop.lake              4    424.73      1617      1742.4   < 2.2e-16 ***
worm.lake             2     13.35      1615      1729.0     0.001265 **
worm.fam              8      0.74      1607      1728.3     0.999425
cop.lake:worm.lake    8      1.07      1599      1727.2     0.997795
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Summary: worm family does not matter!

Comparing all models on testing the prevalence of infection:

> AIC(model1, model2, model3,model4, model5, model6, model7, model8, model9)
Note: according to lab-mate Christopher Peterson, it is fine to do AIC comparisons between GLM and GLMM models (need to ask him for the reference).

Below are the models sorted from best to worst: (the first number after each model name is the degrees of freedom followed by the Akaike Information Criterion (AIC) value:

	df	AIC
1) model8	8	1743.843

model8 <- glm(infected.yes.no ~ cop.lake + worm.lake + native, data = copepods, family = "binomial") #testing for local adaptation

2) model2	8	1744.910

model2 <- glmer(infected.yes.no ~ cop.lake + worm.lake + (1|plate), data=copepods,family="binomial") #GLMM not testing for interaction

3) model6	15	1753.310

model6 <- glm(infected.yes.no ~ cop.lake + worm.lake + cop.lake*worm.lake, data = copepods,family="binomial") #GLM testing for interactions between cop.lake and worm.lake

4) model3	6	1753.846

model3 <- glmer(infected.yes.no ~ cop.lake + (1|plate), data = copepods, family = "binomial") #testing for copepod lake only

5) model1	16	1754.866

model1 <- glmer(infected.yes.no ~ cop.lake*worm.lake + (1|plate), data = copepods, family = "binomial") #testing for effects of cop.lake, worm.lake, and their interacions

6) model9	21	1760.933

model9 <- glm(infected.yes.no ~ cop.lake + worm.lake + cop.lake*worm.lake + worm.fam, data = copepods, family = "binomial") #testing if worm fam had an effect on the prevalence of infection (not, it didn't)

7) model7	65	1788.556

model7 <- glm(infected.yes.no ~ cop.lake + worm.lake + cop.lake*worm.lake + plate, data=copepods,family="binomial") #testing if plate had an effect on the prevalence of infection: no, it didn't

8) model4	4	2163.787

model4 <- glmer(infected.yes.no ~ worm.lake + (1|plate), data = copepods, family = "binomial") #testing for effect of worm.lake in prevalence of infection

9) model5 2 2171.093

model5 <- glmer(infected.yes.no ~ (1|plate), data = copepods, family ="binomial")

Analyzing intensity of infection (i.e. number of tapeworms per infected copepod)

Note: data is Poisson distributed

#Then using the following model:

7) Analyzing results on intensity including those not infected (i.e. number of worms >= 0)

```
> model10 = glm (numb.worm ~ cop.lake + worm.lake + cop.lake*worm.lake,
+               data=copepods,family="poisson")
> summary(model10)

Call:
glm(formula = numb.worm ~ cop.lake + worm.lake + cop.lake * worm.lake,
    family = "poisson", data = copepods)

Deviance Residuals:
      Min        1Q    Median        3Q       Max
-1.6042   -0.6794   -0.5164    0.1574   3.5375

Coefficients:
                           Estimate Std. Error z value Pr(>|z|)
(Intercept)                -1.8302     0.2425  -7.546 4.48e-14 ***
cop.lakeech                 1.9931     0.2557   7.796 6.38e-15 ***
cop.lakegos                 2.0824     0.2540   8.197 2.47e-16 ***
cop.lakelau                -0.2422     0.3382  -0.716  0.47382
cop.lakerob                 0.8747     0.3299   2.652  0.00801 **
worm.lakeech                0.3639     0.3263   1.115  0.26470
worm.lakegos                0.2461     0.3483   0.707  0.47984
cop.lakeech:worm.lakeech   -0.5104     0.3480  -1.467  0.14248
cop.lakegos:worm.lakeech   -0.5044     0.3440  -1.466  0.14260
cop.lakelau:worm.lakeech   -0.1741     0.4578  -0.380  0.70367
cop.lakerob:worm.lakeech   -0.3128     0.4640  -0.674  0.50015
cop.lakeech:worm.lakegos   -0.5809     0.3719  -1.562  0.11827
cop.lakegos:worm.lakegos   -0.6599     0.3685  -1.791  0.07336 .
cop.lakelau:worm.lakegos   -0.5585     0.4983  -1.121  0.26241
cop.lakerob:worm.lakegos   -0.8069     0.5314  -1.518  0.12892
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 2107.7  on 1621  degrees of freedom
Residual deviance: 1498.2  on 1607  degrees of freedom
```

```
AIC: 2959.8

Number of Fisher Scoring iterations: 6

> anova(model10, test = "LRT")
Analysis of Deviance Table

Model: poisson, link: log

Response: numb.worm

Terms added sequentially (first to last)


                         Df  Deviance  Resid. Df  Resid. Dev     Pr(>Chi)
NULL                                      1621      2107.7
cop.lake                  4    584.18      1617      1523.5    < 2.2e-16 ***
worm.lake                 2     19.58      1615      1504.0     5.61e-05 ***
cop.lake:worm.lake        8      5.72      1607      1498.2       0.6781
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Results: It seems that copepod lake and worm lake have significant effects on intensity, but not on their interactions

8) What if I include a fixed variable in there, let's say tapeworm family, and using "glmer" for GLMM

```
> model11 <- glmer(numb.worm ~ cop.lake*worm.lake + (1|worm.fam),
+                  data=copepods,family="poisson")
> summary(model11)
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) [glmerMod]
 Family: poisson  ( log )
Formula: numb.worm ~ cop.lake * worm.lake + (1 | worm.fam)
   Data: copepods

    AIC      BIC   logLik deviance df.resid
 2961.7   3047.9  -1464.8   2929.7     1606

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.1474 -0.4741 -0.3551  0.1667  7.4149

Random effects:
 Groups   Name        Variance Std.Dev.
 worm.fam (Intercept) 0.001736 0.04166
Number of obs: 1622, groups:  worm.fam, 10
```

```
Fixed effects:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)               -1.8330     0.2438  -7.517 5.61e-14 ***
cop.lakeech                1.9938     0.2557   7.798 6.27e-15 ***
cop.lakegos                2.0835     0.2541   8.201 2.39e-16 ***
cop.lakelau               -0.2397     0.3383  -0.709  0.47863
cop.lakerob                0.8760     0.3299   2.655  0.00793 **
worm.lakeech               0.3672     0.3282   1.119  0.26322
worm.lakegos               0.2494     0.3501   0.712  0.47618
cop.lakeech:worm.lakeech  -0.5132     0.3482  -1.474  0.14044
cop.lakegos:worm.lakeech  -0.5069     0.3442  -1.473  0.14082
cop.lakelau:worm.lakeech  -0.1784     0.4581  -0.389  0.69697
cop.lakerob:worm.lakeech  -0.3201     0.4645  -0.689  0.49076
cop.lakeech:worm.lakegos  -0.5813     0.3718  -1.563  0.11795
cop.lakegos:worm.lakegos  -0.6605     0.3685  -1.792  0.07310 .
cop.lakelau:worm.lakegos  -0.5613     0.4984  -1.126  0.26013
cop.lakerob:worm.lakegos  -0.8078     0.5315  -1.520  0.12855
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


> AIC(model10,model11)
               df      AIC
model10 15 2959.820 #glm: numb.worm ~ worm.lake + cop.lake + worm.lake*cop.lake
model11 16 2961.659 #glmer: numb.worm ~ worm.lake*cop.lake + worm.fam [fixed var]

Summary: both models seem pretty similar


9) what if I select only those copepods that got infected for the intensity analysis (as it should be)?
preva = filter (copepods, numb.worm > 0) #using “filter” in “dplyer” R package to extract infected cops from dataset.
hist(preva$numb.worm, ylab = "# copepods")
#Data is still Poisson distributed.


model12 = glm (numb.worm ~ cop.lake + worm.lake + cop.lake*worm.lake,
            data = preva,family="poisson")

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7200  -0.4976  -0.1582   0.2137   2.0344

Coefficients:
```

```
                          Estimate     Std. Error   z value   Pr(>|z|)
(Intercept)              0.060625   0.242536   0.250   0.8026
cop.lakeech              0.515466   0.255655   2.016   0.0438 *
cop.lakegos              0.582373   0.254043   2.292   0.0219 *
cop.lakelau             -0.003466   0.338200  -0.010   0.9918
cop.lakerob              0.101894   0.329884   0.309   0.7574
worm.lakeech             0.093526   0.326255   0.287   0.7744
worm.lakegos             0.072907   0.348315   0.209   0.8342
cop.lakeech:worm.lakeech -0.251881   0.348006  -0.724   0.4692
cop.lakegos:worm.lakeech -0.198414   0.344046  -0.577   0.5641
cop.lakelau:worm.lakeech  0.003466   0.457840   0.008   0.9940
cop.lakerob:worm.lakeech -0.061889   0.463968  -0.133   0.8939
cop.lakeech:worm.lakegos -0.291160   0.368767  -0.790   0.4298
cop.lakegos:worm.lakegos -0.256372   0.368523  -0.696   0.4866
cop.lakelau:worm.lakegos  0.093078   0.493503   0.189   0.8504
cop.lakerob:worm.lakegos -0.117643   0.531446  -0.221   0.8248
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 238.68  on 621  degrees of freedom
Residual deviance: 211.90  on 607  degrees of freedom
AIC: 1673.5

Number of Fisher Scoring iterations: 4

> anova(model12, test = "LRT")
Analysis of Deviance Table

Model: Poisson, link: log

Response: numb.worm

Terms added sequentially (first to last)


                         Df    Deviance   Resid. Df   Resid. Dev      Pr(>Chi)
NULL                                          621     238.68
cop.lake                 4       21.5604        617     217.12       0.0002451 ***
worm.lake                2        3.3173        615     213.80       0.1903991
cop.lake:worm.lake       8        1.9074        607     211.90       0.9837214
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Results: with only the infected copepods, it seems like now only the copepod lakes explains the results. We have the distinct hunch that it has to be either Echo Lake and/or Gosling lake's copepods who are explaining most of these results.

10) What if we include a GLMM model using worm fam as the fix variable (only for the infected copepods; this is for intensity):

Generalized linear mixed model fit by maximum likelihood (Laplace Approximation)

```
model13 <- glmer(numb.worm ~ cop.lake*worm.lake + (1|worm.fam),
                 data = preva,family="poisson")

     AIC      BIC   logLik  deviance df.resid
  1675.5   1746.4   -821.7   1643.5      606

Scaled residuals:
    Min      1Q  Median      3Q     Max
-0.6541 -0.4636 -0.1543  0.2195  2.5118

Random effects:
 Groups   Name        Variance Std.Dev.
 worm.fam (Intercept) 0        0
Number of obs: 622, groups:  worm.fam, 10

Fixed effects:
                                Estimate  Std. Error    z value  Pr(>|z|)
(Intercept)                    0.060625   0.242545   0.250   0.8026
cop.lakeech                     0.515466   0.255665   2.016   0.0438 *
cop.lakegos                     0.582373   0.254052   2.292   0.0219 *
cop.lakelau                    -0.003466   0.338206  -0.010   0.9918
cop.lakerob                     0.101894   0.329891   0.309   0.7574
worm.lakeech                      0.093526   0.326264   0.287   0.7744
worm.lakegos                      0.072907   0.348328   0.209   0.8342
cop.lakeech:worm.lakeech -0.251881   0.348015  -0.724   0.4692
cop.lakegos:worm.lakeech -0.198414   0.344055  -0.577   0.5641
cop.lakelau:worm.lakeech  0.003466   0.457846   0.008   0.9940
cop.lakerob:worm.lakeech -0.061889   0.463976  -0.133   0.8939
cop.lakeech:worm.lakegos -0.291160   0.368781  -0.790   0.4298
cop.lakegos:worm.lakegos -0.256372   0.368537  -0.696   0.4866
cop.lakelau:worm.lakegos  0.093078   0.493512   0.189   0.8504
cop.lakerob:worm.lakegos -0.117643   0.531449  -0.221   0.8248
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

 > AIC(model12,model13)
                df       AIC
model12   15   1673.494
model13   16   1675.494
```

Results: both models seem very similar

11) Local adaption on intensity of infection levels (using only infected copepods for analyses):

```
> summary(model16)

Call:
glm(formula = numb.worm ~ cop.lake + worm.lake + native, family = "poisson",
    data = preva)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7056  -0.4897  -0.1383   0.1908   2.0068

Coefficients:
                Estimate Std. Error    z value  Pr(>|z|)
(Intercept)   0.201136   0.143104   1.406  0.15987
cop.lakeech   0.354473   0.145691   2.433  0.01497 *
cop.lakegos   0.430396   0.144167   2.985  0.00283 **
cop.lakelau  -0.003657   0.195147  -0.019  0.98505
cop.lakerob   0.013441   0.202257   0.066  0.94702
worm.lakeech -0.076399   0.078612  -0.972  0.33112
worm.lakegos -0.112841   0.087777  -1.286  0.19860
nativeTRUE   -0.065931   0.081718  -0.807  0.41977
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 238.68  on 621  degrees of freedom
Residual deviance: 213.15  on 614  degrees of freedom
AIC: 1660.7

> anova(model16, test = "LRT")
Analysis of Deviance Table

Model: Poisson, link: log

Response: numb.worm

Terms added sequentially (first to last)


                Df   Deviance   Resid. Df   Resid. Dev    Pr(>Chi)
NULL                                  621     238.68
cop.lake      4     21.5604         617     217.12        0.0002451 ***
worm.lake   2      3.3173         615     213.80        0.1903991
```

native 1 0.6538 614 213.15 0.4187510

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Results: again, only copepod lake seems to account for the data

The best models for intensity of infection (using only the data from infected copepods):
Below are the models sorted from best to worst: (the first number after each model name is the degrees of freedom followed by the Akaike Information Criterion (AIC) value:

		Df	**AIC**
1.	model 16	8	1660.748
	glm (numb.worm ~ cop.lake + worm.lake + native, data = preva, family="poisson")		
2.	Model 12	15	1673.494
	glm (numb.worm ~ cop.lake + worm.lake + cop.lake*worm.lake, data = preva, family="poisson")		
3.	Model 13	16	1675.494
	glmer(numb.worm ~ cop.lake*worm.lake + (1|worm.fam), data = preva, family="poisson")		