Dear Dr. Clark,

We have received the reports from our advisors on your manuscript, "Behavioral differences between ant species determine the ecological consequences of facultative food-for-protection mutualism", which you submitted to Journal of Insect Behavior.

Based on the advice received, we feel that your manuscript could be reconsidered for publication should you be prepared to incorporate major revisions. When preparing your revised manuscript, you are asked to carefully consider the reviewer comments which can be found below, and submit a list of responses to the comments. You are kindly requested to also check the website for possible reviewer attachment(s). We also request that you edit the references to conform with Journal of Insect Behavior standards (currently they do not) and please remove all DOI.

Please make sure to submit your editable source files (i.e. Word, Tex)

In order to submit your revised manuscript, please access the following web site:

Your username is: [rclark@wesleyan.edu](mailto:rclark@wesleyan.edu)

If you forgot your password, you can click the 'Send Login Details' link on the EM Login page at <https://urldefense.proofpoint.com/v2/url?u=https-3A__joir.editorialmanager.com_&d=DwIFAw&c=C3yme8gMkxg_ihJNXS06ZyWk4EJm8LdrrvxQb-Je7sw&r=2JDpwQMgJHV9A2B3ABaTcn0CONcfcs_UO7poy50ewbc&m=Pc_qua18Cu_OTVx_DgXnkTvpZqVKPSAv0MWJTlMijTo&s=4pLba0PjP4ee-KY_PHf93k7iKQ1Kq0A_8NFIfKUSbgY&e>=.

We look forward to receiving your revised manuscript before 30 Oct 2018.

With kind regards,

Drs. Jeremy Allison and Ring Cardé

Editor in Chief

Journal of Insect Behavior

*Dear Editor,*

*Thank you for your feedback and for the response from reviewers.*

*The reviewers bring up valid criticisms and we have made an attempt to address each of them. Luckily, both reviewers provided detailed feedback on how to improve the manuscript, for which we are thankful. We believe that a clarified methods section, reorganization and rewritten introduction and discussion, and a large number of incremental changes have substantially improved manuscript. We have outlined our changes in response to each comment below.*

COMMENTS FOR THE AUTHOR:

Reviewer #1: Review: JOIR-D-18-00009

This manuscript reports on experiments designed to inform readers about facultative mutualisms involving ants. The authors used two species of ants, known to be sap feeders, to explicitly test four hypothesis, all listed in the introduction. Ants and sap feeders were collected in the field and aggression was tested using dyadic encounters in small arenas. I commend the authors' intent to understand how behavioral interactions can affect a community of animals.

This paper covers an interesting topic but reads to me as if it is a bit under-cooked; i.e., the message of the paper is a confusing and I have questions about how the design was planned and described in the paper.

Honestly, the manuscript reads as if it is in draft form and could use a few more rounds of peer-review in in laboratory meeting before re-submission. I think that some thought about how to organize the information would greatly benefit the message and it is nearly impossible to properly assess the results because I have had a very difficult time understanding the experimental design, especially the aggression assay, and how the stated hypothesis related to the experiments conducted.

I provide my major concerns below. I understand that my comments may be annoying, but I encourage the authors to give serious thought to how to revise this manuscript before resubmission to JIB, or any other venue.

*We appreciate the constructive criticism and have made attempts to address both the minor and major comments outlined by reviewer #1. We have made major revisions to update the methods as well as the introduction and discussion. We believe most of the confusion came from unclear presentation of the methods, which, as the reviewer states, made it difficult to evaluate the manuscript in its entirety. We hope that these changes address the major comments.*

These cover my major concerns about the paper and I hope they prove informative to the authors:

1. Lines 81 - 105: I find that the four hypotheses stated in the introduction are not very well justified by the background material presented. Are these new hypotheses driven by the authors and their study system? And, how they informed by the literature?

*The hypotheses and their predictions come from other papers, as well as a series of natural history observations made while working in this system. We have changed the introduction to reflect this at lines 74-102 (this section has largely been rewritten and/or reorganized).*

a. The "superior protector hypothesis" states that one ant species engaged in mutualism with a sap-feeder will exhibit higher aggression compared to the rest of the community. This is a very broadly stated, and therefore not a very testable, hypothesis. Please define the "community". And, does the design of the relevant experiments (e.g., 3) test this hypothesis? I do not believe that testing interactions between two ants tells us much about "the rest of the community".

*We have made this hypothesis more specific in the introduction. Importantly, the comparison being made isn’t the rest of the community, it is between another dominant member of that ant community.*

*Lines 74-81 now read:*

“In the **superior protector hypothesis**, a dominant ant species will monopolize mutualistic interactions (Grover et al 2007), and this same ant species are responsible for intensified competition (Wilder et al. 2013) and increased predation (Kaplan and Eubanks 2005). This hypothesis predicts that ecological dominance in ant communities is a significant predictor of a superior protector species, including either behavioral dominance (Grover et al. 2007) or numerical dominance (Kaplan and Eubanks 2005). Consequently, this study focuses on one putatively behaviorally dominant ant species as well as one numerically dominant ant species.”

*These two species are far and above the most abundance taxa in this system. This is now reflected in the text at line 115:*

“Our experiments include *Camponotus chromaoides* and *Formica neogagates,* two abundant ant species in upland mixed deciduous forests in New England, USA (Ellison et al. 2012) that engage in mutualisms with sap-feeders, (Clark and Singer 2018), and are actively engaged in competition (personal observations, Fig S1).”

2. I would encourage the authors to state other alternatives to these hypotheses and discuss alternative explanations in the discussion section. I feel as if the author worked very hard to collect a large data set that could be more creatively analyzed, possibly by using a model comparison approach (e.g., Akaike's information criterion (AIC)). The results and discussion read as if hypothesis a is not supported then hypothesis b must be supported which is not entirely true.

*First, did not mean to imply that if hypothesis a is not supported then hypothesis b is supported. These hypotheses are neither mutually exclusive nor does lack of support for one mean support for the other mechanism. We have changed the text to reflect this. As for AIC, we believe this is a fine suggestion, but AIC would be useful in a more complicated study with a larger number of predictor variables. First, the current models we employed are based on a priori predictions made before the study was run, and AIC would be more appropriate if this were an exploratory analysis or larger dataset. Second, we have already included all the potential predictors with the mixed model approach (ant species \* interactor \* honeydew + site\_random\_effect). We have tried additional combinations of factors, the results are the same even if interaction terms are dropped, and even in main effects are tested independently with one-way-ANOVA, etc. In other words, there aren’t other creative analyses to do (as much as we wish there were). AIC is not likely to yield any additional insight into the data generated by these experiments.*

*It is our hope to not make the analysis anymore complicated. We believe the current interpretation of the statistics is best: First, in multiple assays (honeydew baits and treehoppers), Camponotus was more aggressive than Formica. Second, no matter which combination of baits, species, or interactors we tried, honeydew did not induce aggression. We have changed results section to make these points more straightforward.*

*For example, lines 265-268 now read:*

“In experiment 2 we observed no evidence for any of the induced behavior hypotheses. Ants were not more likely to attack prey or competitors if they had recently been collected near baits (Fig. S2, Table S1, Binomial GLMM, *P* = 0.973), and this result did not differ between either *C. chromaoides* or *F. neogagates* (Table S1, Binomial GLMM, *P =* 0.75).”

3. Lines 123 - 139: Were the behavioral measurements collected from video in a blinded fashion? Who collected the data and did they have an investment in the outcome or prior knowledge about the treatments and hypotheses? How was observation bias account for? Or, was it at all? We know from van Wilgenburg and Elgar (2013, Plos One) that confirmation bias occurs in studies in which non-blinded measures of nestmate recognition behavior by ants were measured. I understand it is difficult to find an observer blinded to caterpillar versus and ant, let's say, but there are methods to circumvent this problem, for example, by hiring an observation panel with no knowledge about the study.

*This line of questions is important to the point of the paper. We believe one of the problems here is that we did not adequately describe the methods nor was the introduction framed to show how rare a field experiment with a diet manipulation is. Unlike most behavioral studies with ants and hemipterans, observations were completed in the field with ants collected directly from host plants immediately (within seconds). In other words, there was no travel time, no collection of ants to be brought to the lab, which significantly alters ant behavior. This took considerable effort, but the hope is that it captures ant behavior in a much more realistic manner than laboratory studies. Importantly, we noticed that in the field, ants are much more docile than we anticipated since most comparable studies are in the lab and/or with invasive ants, hence the low frequency of attacks. However, despite the low frequency, we have seen that these ant species have strong effects on native caterpillars in other predator exclusion experiments. While we admit that we did not have a naïve observer in our design, we were testing multiple hypotheses in each experiment and the results did not support the most interesting hypotheses, so confirmation bias impact should be minimal.*

4. Lines 123-139: I have read this section several times and I would not be able to replicate the behavioral measure based upon how it is described in the text. This section should be re-drafted with a clearer explanation of the experimental design. Perhaps moving Fig S5 to the text would be helpful, but I am confused about other issues:

1. Basically, I can understand how an ant was collected. But, was a beetle, competing ant, or caterpillar in the arena when the ant dropped in? Some information about the order events in is missing in this section. Also, how was the sequence of treatments randomized, or were they?
2. For example, what is a "competing ant"? What evidence exists that the two focal species chosen are competitors?

*The sequences of events were not randomized, but done in the same order. Since the behavioral trials did not use the same ant repeatedly (new ants were collected for each trial) randomizing order or trials was not necessary (e.g. each was fully independent). We changed the methods section considerably to address this comment, for example, Lines 128-136 now read:*

Behavioral assays examined one-on-one interactions of ants in fluon-lined containers with an individual ladybird beetle, competing ant, or waxworm caterpillar (“interactors”). Ant workers were collected in the field from plant species that also host phloem-feeding Hemipterans (sap-feeders). A container was held under foliage of the host plant, which was gently tapped causing the ants to drop. Interactors were already present in the container and the assay started when the experimental ant reached the container. The experimental manipulations included a) the type of interactor in the fluon-lined container and b) the source of the experimental ant (branches with sap-feeders, baits, or controls). Attack behaviors were recorded once ants were placed into assay containers (Fig. S2).

1. And, why are waxworm caterpillars used as a treatment? Except for lines 109-111 (which should be placed in the introduction) there is no reason provided as to why this should be a treatment, given the information provided. Any caterpillar is useful as a fill-in for native members of the community?

*We have edited this section to be more careful in the decision to use waxworms. The idea here is that waxworm is a ‘live protein bait’ since it is used as a feeder insect for ant colonies just like the honeydew proxy bait is used as an active source of carbohydrates and water for ant colonies reared in the lab. Waxworm caterpillar were not intended to be stand-in or replicas of caterpillars in the field, but instead we used waxworms because they had key qualities relevant to ant behavior (e.g. being palatable prey).*

*Lines 164-169 now read:*

“We used *A. grisella* as a palatable prey because captive colonies of ants are often reared on this food source (Hood et al. 2003). This assay tests ant responses to a standardized prey item known to function as a protein-rich food for various ant species. The design of this assay enables a test of the prediction that ant workers are more likely to collect protein-rich food a compensatory response to carbon-rich honeydew provided by sap-feeders (e.g. Ness et al. 2009).”

d. Why was the time of an encounter recorded and how was it used in the analysis? This does not seem relevant. A binary response was measured, correct? 1. An ant showed attack behavior. Or, 2. It did not.

*We removed this from this sentence since it was not relevant to the analysis presented in this manuscript.*

e. Define what is an "encounter" (line 133)?

*Sentence added which defines “encounter”. Lines 138-141 now read:*

“We used this single encounter approach (e.g. one-on-one interactions) fitting with natural history observations in previous field seasons; *Formica* and *Campouts* species forage singly, and competitive encounters and predation events are typically between a single worker ant and another interactor (personal observations).”

f. Lines 129-132: I ask the authors to please explicitly define "attack behavior". Why does it "include" the listed aggressive behaviors? And, why were other behaviors presumably not included? As written, I am suspect of the measurements made. What instructions were given to the observers of the data? What was the sample unit? A fight? An ant? Was a full factorial design employed? What were the sample sizes used?

*This section has been reorganized. The sample unit, scoring of events, number of ants used, and factorial design varied is outlined under the header for each experiment below this paragraph.*

*Lines 141-148 now read:*

“In these single-encounter trials, we counted aggressive behavior between ants as biting or tangling between two ants. We counted aggression towards prey when we observed biting and/or carrying behavior. Intermediate levels of aggression, brief encounters (under one second) were not counted as “presence of aggressive behavior” in our analysis. Once any encounter occurred (aggressive or otherwise) the assay was ended, and the ant was removed from the container. Encounters not scored as aggressive included brief contact between an ant and its interactor, such as touching of antennae, legs, or mandibles, which did not lead to further behaviors outlined above.”

g. Lines 134-137: The logic of why a single encounter approach was used does not make sense to me, as written. Why is it more conservative? And, could the authors provide some additional information as to why "rapid aggressive behaviors" were preferable to slow responses, I guess? Returning to my first point, if confirmation bias was not accounted for in the design, this single encounter method is highly suspect to me because observers may unintentionally miss aggressive behaviors when they were not expected to occur (van Wilgenburg and Elgar, 2013).

*We added justification for the one-on-one approach to this paragraph (above). For the second point, we edited the paragraph for clarity since some of the concerns raised are probably due to a poorly articulated description of the methods on our part. To prevent bias, we were very careful to only count aggressive behaviors that were unambiguously attack behaviors. We did this with hopes to account for confirmation bias so there were no “judgment calls” for intermediate displays of aggression. All outcomes of aggression assays were scored by one person (R.E. Clark), there were multiple alternative hypotheses, and results did not support intended hypotheses that we expected to find support for.*

h. Lines 136-137: Why were different time durations used for the different experiments?

*This Information has been added. Lines 148-152 now read:*

“Trials were timed for 600 seconds in experiment 1 and 300 seconds in experiments 2 and 3. We halved the duration to increase the number of trials that could be run in a day because most interactions (47 out of 49 interactions) in experiment one happened within 300 s. Each replicate of the experiment involved a different individual ant and a newly cleaned container.”

i. Line 139: I am unclear as to why releasing a focal ant after a trial prevented recollecting the ant. Please re-state as it seems that keeping the ant in a container for the duration of all trials before returning it would be the only way to ensure no pseudoreplication occurred.

*We agree. Ants were kept in containers for the duration of all assays at each location in the field. The text now states this point explicitly.*

j. I would ask that the authors better justify using a dyadic encounter approach. What do we know about the aggressive behavior of the two ant species to support the use of this assay?

*This is a fair point and one we can address with natural history observations. While these ant species are abundant, their overall density is quite low (e.g. typically 1 ant per branch in Clark and Singer 2018). We often observed competitive interactions between these two ant species and it was typically a single individual running into another single individual (whether it is a prey item or competitor) while foraging. This information has been added in Lines 138-141, which also justifies the one-on-one approach discussed in comments above.*

k. Why were not focal ants provided with a one-on-one interaction with a nestmate ant? Because two ants or an ant and a sap feeder are constrained to an arena, shouldn't there be a control or comparator for the other treatments with a treatment that would be predicted to have low levels of aggression? In other words, how do the authors account for the effect of the treatment and assay container on ant aggression?

*We found this question confusing because we did not have an assay where ants and sap-feeders were constrained to an area. The only interactions we observed in fluon-lined containers were ant-ant competition, ant-ladybird interactions, and ant-caterpillar interactions. The only attribute we manipulated (e.g. treatment) was the source of ants. To address this, we edited the methods section to be more clear.*

*Lines 145-146 now reads: “*Once any encounter occurred (aggressive or otherwise) the assay was ended, and the ant was removed from the container.”

*Furthermore, as the reviewer suggested, we did not do an assay on nestmate recognition (e.g. two ants from the same colony) as a point of comparison for competitive aggressiveness. We assume the attack rate between two ants from the same colony would be zero.*

5. Lines 141-148: Again, not enough information is provided about the design to assess this section.

*This section has been edited and additional information in the “scoring ant behavior” subheading should provide sufficient information (Lines 126-169 in manuscript, comments addressed above for specific criticisms).*

6. Lines 149-152: I would recommend that the authors provide additional context for why this information is provided and why it was used in the model.

*This is a fair point. We added a sentence here to clarify. These host trees were chosen because they are dominant woody plants in this system, and this experiment is based on ant community surveys from previous experiments on these same 8 host-plant species.*

*Lines 182-186 now read: “*Experiment 1 examined the response of *F. neogagates* workers collected from the tree species *Acer rubrum, Betula lenta, Carya* spp., *Fagus grandifolia, Hamamelis virginiana*, *Prunus serotina* and *Quercus rubra* because they are the most common woody plants in this environment, and related work on this system includes these tree species (Singer et al. 2017, Clark et al. 2016).”

7. Lines 146-148: What is a "baited plant"? What does this mean and what are the methods? Is it assumed that if two ants, each of the a different species, were on a baited plant that they were necessarily competitors?

*Changed baited plant to “honeydew proxy bait” instead of “baited plant” to be more specific. This is now located at line 175 in the manuscript. We assumed at this close proximity ants would be competitors based on natural history observations (Lines 179).*

8. Lines 141-152: Given the opaqueness of the previous section on the bioassay, I cannot properly understand the design here. Was a F.n. and placed in a vial with either a C.c., or caterpillar? Both? It seems a bit clumsy for the authors to take a reductive, hypothesis driven approach that combines the testing of two, very different hypotheses, into one experiment.

*This paragraph has been rewritten (mainly now found in lines 174 to 181, relevant to this comment) to show that it is two assays (competition and predation, strictly speaking) and it is combined into one experiment since the source of Formica neogagates ants was shared.*

9. The title could more specifically describe the results within. It is a bit too generic, I think, and does not indicate any inferences made from the data.

*We changed the title to “*Differences in aggressive behaviors between two ant species determine the ecological consequences of a facultative food-for-protection mutualism*.”*

10. Lines 153-186: It is not stated, but I assume experiment 2 is designed to test the aggressive defense hypothesis.

*Information added to the first sentence of the paragraph – this experiment was intended to evaluate all four hypotheses outlined in the introduction. Lines 189-191 now read:*

“In this series of bioassays, we evaluated the prediction that ant aggressive behaviors would be more frequent for ants that had been recently consuming carbohydrates from mutualists to test all four hypotheses simultaneously.”

11. Line 160-161: It is stated that comparable studies used similar honeydew proxy baits, but only one citation is provided. The authors should cite the other studies.

*Petry et al. is the one study we know that is most comparable to our experiments. We changed the sentence, so it says “study” instead of “studies” in line 196.*

12. Line 166: What does it mean to "score the behavior"? Again, this relates to my previous comments about the methods narrative for the behavioral bioassay. Didn't the authors simply measure if the ants were aggressive towards ladybird beetles, etc.? I think more precise language would help clarify the message, here, and throughout the paper.

*We have edited the manuscript so it is now more precise and reflects details of the assay instead of using a shorthand term like ‘scoring behavior’ at various points. The criteria is described in lines 142-148.*

13. Lines 187 - 200: As written, this experiment seems out of place and is very poorly stated. Much background material is presented but little about the experimental design. Why do we learn, in lines 199 - 200 that experiment 2 used a complete factorial design? Shouldn't that have been made obvious in the previous section? I assume there were 6 treatment groups if there were 15 replicates and 90 trials, but I am very confused about what were the treatments.

*We* *have rewritten this paragraph to clarify the importance of the chosen design. Sentences relevant to the factorial design have been rewritten as well, and the number of replicates has been updated. Lines 201-205 now read:*

“This design provided tests of all four hypotheses to assess the relative importance of intraspecific induced behavior (e.g. behavioral shifts in *F. neogagates* or *C. chromaoides* after consuming carbohydrates) from interspecfic differences (e.g. *C. chromoaides* being more likely to attack prey than *F. neogagates* overall). Experiment 2 had 20 full replicates ((ant species \* interactor \* bait presence/absence).”

14. The discussion is largely a restatement of the results. I would appreciate that more words are devoted to providing a narrative that conceptually models and interprets the meaning of the data given what we already know in the literature.

*We reduced the focus of the opening paragraph (lines 279-301) which restates key results and tells which hypotheses were ultimately supported or not. The remaining introduction links these findings with other work on ant community ecology, ant invasion ecology, and arthropod food web interactions. We have shorted the discussion briefly to focus more on the importance of ant behaviors.*

Reviewer #3: In this manuscript, Clark & Singer propose an interesting look into ant-hemipteran mutualisms using multiple experiments that examined both inter- and intraspecific behavioral differences for two ant species, Camponotus chromaiodes and Formica neogagates. In doing so they focus on four hypotheses: (1) the superior protector hypothesis, (2) the aggressive defense hypothesis, (3) the intensified competition hypothesis, and (4) the prey acquisition hypothesis. Support was only found for the superior protector hypothesis, and no evidence was found that indicated food rewards provided by fluid feeding hemipteran induced behavioral changes. I found this paper to be well written, interesting, and my comments are primarily focused on the hypotheses/methods/results sections where I believe an increase in clarity would be helpful. I list major comments below with minor points for each section in chronological order.

Overall:

1) For the superior protector hypothesis, I felt that the predictions were quite vague in reference to the study organisms. While not stated in the introduction, would you predict for one or the other species to be a better protector based on their natural history/observations/information in the literature/traits? In the abstract (line 18), the authors make the prediction that C. chromaiodes would be a better protector. In the discussion, there is mention of numerical differences (line 285) as a mechanism for dominance. I think explicitly stating which species you think will be a better competitor, and why, would greatly increase the power of this hypothesis. For example, since F. neogagates has larger colonies than C. chromaiodes, they may need more resources to meet the queen/brood's demands and thus are more likely to protect valuable carbohydrate sources like honeydew from fluid feeding hemipteran. This would also indicate why an initial focus was placed on F. neogagates

in experiment 1. This might require introducing the two species in the introduction before these hypotheses or alternatively suggesting mechanisms why you might predict one species to be better than the other besides behavioral observations of invasive ants which can be both aggressive and numerous.

2) The attack rates on caterpillars and ladybeetles by both ant species was incredibly low (Figs. 2 & 3). I wonder if the novelty of an exotic ladybeetle may be one potential confounding variable. Additionally, the Hood et al. 2003 paper referencing the palatability of lesser wax moth caterpillars focuses on the red imported fire ant, a notoriously omnivorous invasive ground dwelling ant. I am not sure how comparable that species is with large formicines which may not be as aggressive and/or behaviorally quite different.

*We agree that this is a bit of a stretch. The focus was originally just to look at preference of protein rich prey, which this reviewer pointed out below for lines 225-226 below. Given the use of waxworms in Hood et al. 2003, and since waxworms are quite different than native caterpillars and the formicines in this study are distinct from fire ants, we edited this section (same comments and information as referred to in response to reviewer 1).*

3) While the discussion nicely covers the first hypothesis, for which the authors found support, there was a lack of discussion about the other three hypotheses besides brief mention (e.g. line 298). I think adding some ideas why no support was found for these three different ideas would improve the manuscript and, as the authors suggest, encourage other researchers to look at behavior in different ways.

*The discussion has been reorganized so there is a paragraph which attempts to explain the lack of support for these three hypotheses and what it means given other observations in the literature (Lines 312-322).*

Minor Comments:

Abstract:

Line 16: And throughout the manuscript. Check spelling of Camponotus chromaiodes.

*This is the correct spelling.*

Introduction:

Line 94: Intensified competition hypothesis - I feel the initial statement that "individual ant workers would be more likely to aggressively attack competing ants" is vague. In the following sentences, the authors expand on this idea but it may be worth just adding something simple to the end of line 95 like "in the presence of mutualists".

*This section has been rewritten. Lines 96-100 now reads:*

“In the **intensified competition hypothesis**, induced aggression is particularly relevant to ant competition such that ant workers are more likely to display aggressive competition behaviors towards other ants when in the presence of mutualists (Grover et al. 2007, Kay et al. 2010). This hypothesis would be supported if we see higher rates of interspecific aggression in ants recently engaged in mutualism.”

Methods:

Line 117: Minor, but the first listed figure is "Fig S5". Might help the organization if figures are listed in order of presentation. This appears to be a trend throughout the methods and results.

Line 129: Same as above.

*Change made. S1-S5 reordered to fit the order they come up in the manuscript.*

Line 141: Based on the experiments and results, wouldn't this be interspecific competition differences? See major comment above, but I was curious why F. neogagates was being primarily focused on for experiment 1. This may be one location where adding some additional predictions to the hypotheses in the introduction would clarify why certain behavioral assays were being performed.

*Change made. We added this information on why F. neogagates was chosen first.*

Line 146: Are these waxworms the same as used in experiment 2? If so, recommend moving their description from lines 178-186 to here.

*We added information on all trials to the “scoring ant behavior” subheading which precedes all three experiments. It is now found at line 162.*

Line 147: Baited plants are introduced here, but the description of baits are listed in experiment 2 (lines 157-159). Were these the same baits? And if so, would recommend moving the bait description up to this section.

*Changed to “trees with honeydew proxy baits” in line 175.*

Experiment #2: How many total behavioral observations were done? Asking for consistency as the number of observations is listed in experiment #1 (line 151) and experiment #3 (line 200).

*This has been updated in lines 204 and 221.*

Line 163: See comment above about figure order. "Fig. S3".

*Change made in Figure numbering.*

Line 199: See comment above about figure order. "Fig. S4".

*Change made in Figure numbering.*

Lines 213-214: Was there a particular reason to include the interaction terms in your model? Specifically, were these done to test a particular prediction in your hypotheses? Would be good to clarify why particular combinations were used.

*This information is added at lines 201-205 – it is necessary to test the superior protector hypothesis which examines different in ant behavior between species in response to honeydew baits and/or with interactors.*

Results:

Line 222: See comment above about line 141. Inter- versus intraspecific competition.

*Change made, reference to the type of competition has been double-checked throughout the manuscript.*

Line 223: "of" seems out of place at the beginning of the sentence.

*Sentence rewritten to be shorter and more to the point. It now says: “This finding demonstrates that F. neogagates is behaviorally subdominant to C. chromaoides in this community” at line 251.*

Lines 225-226: Ants in the canopy are often nitrogen limited as they regularly forage on carbohydrate resources. I wonder if you did not find support for the prey acquisition hypothesis because of this. Just a thought.

*We added this information to the experimental design to address why waxworms were used in response to reviewer 1 also. In particular, we added the following sentence at the end of this paragraph from lines 167-169 to address this:*

“The design of this assay enables a test of the prediction that ant workers are more likely to collect protein-rich food a compensatory response to carbon-rich honeydew provided by sap-feeders (e.g. Ness et al. 2009).”

Line 233: Suggest changing "dropping" to "drop".

*Change made (at line 254 now).*

Line 237: Suggest changing "protector trait hypothesis" to the same hypothesis name in the introduction. Superior protector hypothesis?

*Change made (now in line 290).*

Line 252: Is Fig. 3 supposed to be Fig. 4 here?

*Yes. Change made (now in line 272).*

Line 256: There is no Table S3 and it was unclear how Fig. S3 (photograph of Camponotus workers feeding on a bait) relates to the data or statistics in this statement.

*This now refers to the correct figure (S2). Since the test is just a single comparison (two ant species) we report the GLMM results in the text.*

*These lines now read: “For both C. chromaoides and F. neogagates, ants taken in proximity to sap-feeders were not more likely to attack interactors compared to control ants (Fig. S2, Binomial GLMM, P = 0.999).”*

Discussion:

Lines 265-267: See major comment above. This statement suggests that C. chromaiodes is a superior protector of their mutualist partner yet both ant species had low attack rates on prey (caterpillar) and predators (ladybeetles) in Figs. 2 & 3.

*The comparison of importance is between F. neogagates and C. chromoaides, but it is indeed the case that the attack rate is still <20% for C. chromaoides. We changed the implication that this has “strong ecological impacts” in the discussion, but instead point out that C. chromaoides is more aggressive overall compared to F. neogagates.*

Line 301: I believe Crematogaster opuntiae is now a junior synonym of Crematogaster vermiculata (Morgan and Mackay 2017)

*In this current version, we just refer to Crematogaster genus instead of species given change in taxonomy.*

Figures:

Figure 1: On lines 238-239, it mentions that 21% of trials C. chromaiodes attacked compared to 4% of trials for F. neogagates. Just wanted to make sure these are the same values being reported in Fig. 1 as it appears there is a much lower attack rate on average (~15%). Also, on line 527 I believe this is referencing Experiment 2.

*These estimates were in error (likely based on raw values or an old R script) so we reran the GLMM to computed estimated marginal means from analyses used to generate Figure 1. The attack rates were 15.8% and 4.8% for C. chromaoides and F. neogagates, respectively. The text has been updated.*

Figure 2: I believe this is referencing Experiment 2 not 1 on line 532.

*This is correct. Change made so Fig. 2 refers to experiment 2.*

Figure 4: Same comment for Figure 1. Check % of trials and mean values reported in figures (e.g. 14% of trials in text and mean >0.2 prob of attack in Figure 4).

*Change made following similar suggestion in Figure 1 above. The text now reads 21.2% and 4.2% for C. chromaoides and F. neogagates respectively.*

References:

Morgan, C. and Mackay, W. P. 2017. The North America acrobat ants of the hyperdiverse genus Crematogaster.