

Paper Review - Cheating Interactions Favor Modularity in Mutualistic Networks

Editor (Dr. Wittmann)

Comment # 1 by editor Please carefully address all the comments. I would like to encourage you to explore population size feedbacks in some form, as suggested by reviewer 1.

Response by the authors: We have now explored potential effects of population sizes on the coevolutionary dynamics as suggested by reviewer 1. We added four supplementary figures exploring this role. These results indicate that our conclusions hold after incorporating the numerical effects of the abundances of partners on the structure of selection.

Comment # 2 by editor Also, I would like to emphasize that the clarity and language of writing need to be improved. Reviewer 2 makes many relevant suggestions here but you will also have to carefully proof-read the paper yourself, getting some help if necessary, for example from a native English speaker.

Response by the authors: The authors reviewed the text for english mistakes and for better clarity of the Main Text and Supporting Information.

A few more comments of my own:

Comment # 3 by editor Abstract: Could you add a little bit of biological intuition for why cheaters have these effects?

Response by the authors: We followed your suggestions and we have now added biological intuition on how cheaters change the structure of selection operating on species in the mutualistic network.

Comment # 4 by editor L55: Please make clear that trait mismatching is favored only from the host's perspective.

Response by the authors: We have rewrite this sentence that now clearly stated that selection favors trait mismatch in hosts.

Comment # 5 by editor L115-L116: I did not follow the logic in this sentence.

Response by the authors: We now clarify that higher frequency of cheaters may foster trait disparity which, in turn, may lead to the disruption of ecological interactions.

Comment # 6 by editor As also pointed out by reviewer 2, the explanations around equation 2 and how the trait barrier was implemented was completely unclear to me, both in the main text and in the corresponding part in the supplement. I also went to the Andreazzi et al. (2017) paper and it became clearer there. However, I am not sure

you are making the same assumptions as in that paper. If I understand L158-160, it seems that if the difference $Z_j - Z_i$ is positive, then it is further amplified by adding epsilon, and vice versa for negative differences. This seems to be the opposite of what is happening in Andreazzi et al. 2017. Is this maybe a mistake?

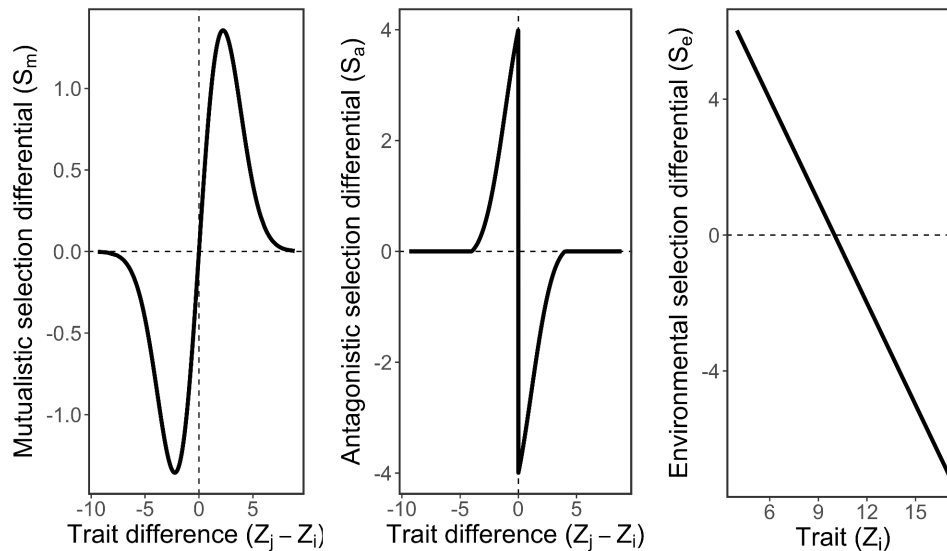
Response by the authors: We apologize for the lack of clarity and for the typo when describing the model, which in turn leads to the wrong reasoning for how selection operates in our model. Here, we are making the same assumptions that Andreazzi *et al.* 2017. Selection imposed on victims favor trait mismatching increasing. If the trait difference between cheater j and victim i is positive, selection will favor the victim trait to decrease. Otherwise, if the trait difference between i and j is negative, selection will favor the victim trait i to increase. We have now clarified the text and cited the section on the Supporting Information that explains about the model assumptions.

Comment # 7 by editor Furthermore, if all the other coefficients are positive, it seems that a positive trait difference between exploiter and victim would cause an increase in victim trait value, which doesn't make sense. Or is the direction of selection reversed by a negative v_{ij} ? Please explain all of this more clearly.

Response by the authors: We apologize for the typo in model description that lead to this confusion. We fixed the model description (see response to comment #6 by editor).

Comment # 8 by editor In addition, to facilitate understanding, I would strongly recommend to make an additional figure showing the difference in trait values for a species pair on the x axis and the resulting selection differential on the y axis, maybe with different panels for the different components as given by equation 1 and 2.

Response by the authors: Thank you for the suggestion. We have followed your advice and now in the Supporting Figure 2 the readers will find how the different components of equation 1 and 2 change with trait values. The Sup. Figure 2 is depicted below:



Sup. Figure 2. Theoretical representation of the selection differentials S from mutualisms, cheating interactions and the environmental selection based on Z_i and the trait difference between species i and j .

Comment # 9 by editor L168: I thought the S values are taking care of the slope of the adaptive landscape. So why does ϕ_i still include part of this?

Response by the authors: The parameter ϕ_{ij} is a combined parameter that includes the additive genetic variation for the trait and a scaling constant affecting the slope of the adaptive landscape and . The S values are the selection differentials which, in turn, shift the slope of the adaptive landscape.

Comment # 10 by editor L179: Can you provide some rationale for this equation?

Response by the authors: Yes, the notion here is, for species with multiple partners, the higher is relative levels of trait matching the stronger is the selection imposed by a given partner. We changed the text, improving the explanation of the equation and its basic assumptions.

Comment # 11 by editor L185-L197: Could you add one or two sentences how you used these empirical networks in your simulations?

Response by the authors: We have now inserted two sentences at the end of the “Mutualistic Networks” section explaining that we used empirical networks to parameterize the network structure in our coevolutionary model.

Comment # 12 by editor L234-236: How was that done? Please explain or refer to the respective section in the supplement?

Response by the authors: We have now referred to the Supporting Information’s section “Interaction shifts”, in which we now explain in detail the analysis. Briefly, we perform a set of simulations in which we allow outcomes shifts during the

simulations. At each timestep there is a probability g which is the probability of an outcome shift during the simulation. We vary this probability from 0.01 to 1. The results, which are qualitatively similar to the benchmark simulations, were reported in Sup. Figure 10.

Comment # 13 by editor L348-351: If I understand correctly, your analytical results are just for a single species pair. So they cannot really support the result that overall trait disparity in the network increases with the frequency of cheating interactions. Since you do not discuss these analytical results much elsewhere I think, I would suggest to leave them out.

Response by the authors: We agree. We have followed your advice and removed these results from our Main Text and Supporting Information.

Comment # 14 by editor Table 1: I didn't see a parameter A in the model description. Also, carefully check the English in this table.

Response by the authors: We apologize for this typo. We changed it for α .

Comment # 15 by editor Table 3: I am confused because the table suggests a highly significant t-test. But in L366-374, you claim that there is no clear effect.

Response by the authors: We have now removed the entire question ii) from the paper, including Table 3, following the suggestion of reviewer #2 (please see Comment # 1 by reviewer #2). Having said that, we clarify that the p-values of numerical simulations usually can be highly significant ($p < 0.001$) due to the elevated number of simulations. In this situation, the more important aspect of the simulation outcomes is how strong the effect is (e.g., measured by estimating the slope of the relationship).

Revisor 1

Comment # 1 by reviewer #1: I don't know if this a drawback in their model but the authors do not consider any feedback from species densities. In such large communities, population densities will play a major role in impacting how evolutionary dynamics are played. The authors do not consider species densities to impact the evolutionary trait dynamics of either cheaters or mutualistic species. Can the authors please justify why one major component of evolutionary dynamics are not being considered here?

Response by the authors: We agree with the reviewer that the feedback between trait evolution and population densities is a key aspect of evolutionary dynamics. Indeed, recently, our research group began to explore the trait-density feedback (Andreazzi et al. 2018. PRSB, Assis et al. 2020. Ecology Letters), emphasizing that these trait-density feedbacks may generate novel dynamics. Having said that, we follow the usual theoretical approach to ecology, focusing on the simplified models

first. Then, after a good understanding of the simplified dynamics, to explore the consequences of the complexity of trait-density feedbacks (see also our response #2 to reviewer #1). This is specially important when exploring selection imposed by ecological interactions because the link between ecological effects of these interactions often do not translate in simple ways to selective pressures (Benkman 2013. Ecology Letters). Specifically, our model assumes that the populations are in equilibrium and therefore there is no specific ecological feedback. We now explicitly state this information in the text. Additionally, we have now incorporated potential abundance-related effects in a sensitivity analysis in the text, by exploring the role of abundances in equilibrium in shaping coevolutionary dynamics (see comment #2 by reviewer #1).

Comment #2 by reviewer #1: If it's to simplify the model by not tracking population dynamics, one could still include and consider equilibrium population densities in their model:

$$f_{ij}(t) = \frac{a_{ik} \exp\left(-\alpha(z_i - z_j)^2\right) N_k^*}{1 + h \sum a_{ik} \exp\left(-\alpha(z_i - z_j)^2\right) N_k^*}$$

Where, h is the handling time and N_k^* is the equilibrium density of the mutualistic partner. I am not asking the authors to use a functional response formulation but I feel that not including any kind of population density in their model will leave one important facet that would contribute to how much cheaters or mutualistic partners could influence evolutionary dynamics. This particular functional response can be incorporated (in case the authors decide to include it) in equation S1 where then some kind of ecological dynamics (say equilibrium species density) can be at play in influencing evolutionary dynamics.

Response by the authors: We follow the reviewer suggestions of incorporating the abundances at the equilibrium. Having said that, we opted for a type-I response instead of a type-II response because (1) it allows an easier standardization of f_{ij} , since the summation of f_{ij} in j needs to sum up to 1 and (2) do not incorporate additional non-linearity that could make the results hard to detangle. Our results suggest that incorporating abundance at the equilibrium, although increasing variability in the outcomes of the evolutionary dynamics, does not lead to qualitative distinct results (see Supporting Information).

Comment # 2 by reviewer #1 A simple formulation of such ecological dynamics with their current model could be done as follows:

$$\Delta Z_i = \frac{h_i^2 \sigma_i^2 \delta}{\delta Z} \ln W_i$$

Now expanding $W_i^{(t)} = R_i^{1 - \frac{f(N_i)}{K_i}} \exp(-(Z_i - \theta)^2) \bar{B}_i$ Where B_i (captures the cumulative impact of mutualistic and or cheating pairwise interaction on species i . Expanding \bar{B}_i

$$\bar{B}_i = \exp \left(\sum \frac{a_{ij} \exp(-\alpha(z_i(t) - z_j(t))^2) N_j}{(1 + h a_{ij} \exp(-\alpha(z_i(t) - z_j(t))^2) N_j)} \right)$$

which captures what q_{ij} in the previous note that I mentioned and also introduces type 2 functional response curve in evolutionary trait dynamics and $R_i^{1 - \frac{f(N_i)}{K_i}}$ introduces density dependence with $R!$ being the reproductive rate of the species i and a_{ij} is the adjacency matrix.

With this population dynamics (with no interspecific competition at play) can simply be written as $N(t) = N(t-1) * W(t)$. (there might be some typos and mistakes but the model then would also introduce ecological dynamics in the picture). I would completely understand if the authors don't go towards this direction but it is something that the authors might want to consider in their future research. Similar eco-evolutionary dynamics in mutualistic networks can also be modelled following the formulation in this recent publication at Ecology Letters by Baruah (2021) "The impact of individual variation on abrupt collapses in mutualistic networks." If the authors do not want to use such a formulation, I would completely understand but would suggest that they include some kind of explanation that justifies their original model.

Response by the authors: We thank the reviewer for this suggestion and we will consider this approach in the future when exploring eco-evolutionary dynamics with cheaters and mutualistic partners. Having said that, we feel that it would be too much to address in a single paper and we opted to incorporate abundances in the equilibrium as a first approximation (see also our responses to the comments #1 and 2 of the reviewer #1).

Comment # 3 by reviewer #1 Increase in the frequency of cheaters in mutualistic networks increases trait disparity. This is a nice result from their model. In such co-evolutionary dynamics as frequency of cheaters increases, naturally more trait mismatch increases given the gaussian trait function that they defined in q_{ij} . However, I would argue that even with increases in frequency of cheating interactions, trait disparity would not increase as much as the authors have reported. This is again particularly linked to the earlier point that I have raised. Including ecological dynamics will probably further decrease trait disparity as eco-evolutionary feedback will be at play. As trait mismatch increase with increasing

cheating interactions, population densities of cheaters should decrease and which consequently will drive those cheater species to further decrease trait disparity. I wonder what the authors would have to say to this. Could the authors give a reasoning whether such high trait disparity will be achieved when species ecological dynamics are at play?

Response by the authors: We agree with the reasoning of the reviewer and indeed our model qualitatively captures this dynamics by selection, always favoring trait matching of cheaters with particular hosts. We think the comment by the reviewer illustrates that we failed in clarifying an important point in our analyses, that what we are measuring is the network-level patterns of trait mismatching and trait distributions. These results do not imply that, at the level of each pairwise interaction, there are high levels of trait mismatching. Indeed, cheaters can evolve, in our model, to become really adapted to particular victims. However, at the network level, because many victims are evolving extreme phenotypes due to the presence of cheaters, they are also creating higher levels of trait mismatching with mutualistic partners. In this sense, cheaters add to the environmental selection in the model, imposing levels of trait disparity (and mismatch) that are higher than scenarios in which only selection favoring trait mismatching by mutualisms is occurring (Guimarães et al. 2011. Ecology Letters). We emphasize, however, that the increase in trait disparity promoted by cheaters is limited by the conflicting selection imposed by the environment and by selection imposed by mutualistic partners favoring trait matching.

Comment # 4 by reviewer #1 The only way the authors have considered additive genetic variation to impact evolutionary dynamics is through the trait evolution equation mentioned in supplementary information equation S1. I would argue that such variation is also important for trait matching or mis-matching. With higher amount of variation there is more chances of trait matching and dynamically over time that would impact how evolutionary dynamics would pan out. With higher trait variation, strength in mutualistic interactions will further increase which then should further positively increase trait matching. Similarly, in the case of interactions where species cheat, higher variance would lead to higher amount of trait matching in general. The authors then could realistically formulate their trait barrier equation as something like this $|Z_i - Z_j| > f(\sigma_i, \sigma_j)$ where $f(\sigma_i, \sigma_j)$ could be an additive function of trait variances of the two species in general. If not, could the authors elaborate more on this as what this threshold is and what it could be actually dependent on?

Response by the authors: Please note that by being part of the composed parameter controlling the response to selection (equation 5 and equation S1), the additive genetic variance generates the effect of the reviewer points out, that is, the higher the genetic additive variance, the higher is the evolutionary change, allowing the

mutualistic partner to faster achieve trait matching and the higher is the likelihood of the cheater to overcome the trait threshold (epsilon). In the current model, which is a mean-field model that we track changes of mean phenotypes, we would like to point out that simpler effects of trait variance, e.g., $\epsilon = \epsilon / (\text{additive genetic variance})$, would imply in simple scaling of the effects of epsilon values. Having said that, we totally agree with the reviewer that intraspecific trait variance could directly affect the role of epsilon in shaping dynamics. Also, intraspecific trait variance could also affect the role of trait matching in shaping interaction strengths, as now known for theoretical ecological dynamics (see the studies by Daniel Bolnick on the topics). We agree that this topic deserves future research and indeed one of us pointed out in a recent review that this connection between intraspecific and interspecific network patterns is one of the cutting-edge problems in the study of ecological interactions (Guimarães 2020. AREES). However, we think that, for actually address this problem with the depth its deserves, we should explore not only trait variance per se, but the intraspecific trait distributions in detail, since different shapes of trait distribution may have important consequences for multiple of evolutionary processes, from speciation to the role of competition in promoting trait disparity. So, we think that this type of analysis would deserve a separate manuscript and we are now just beginning to explore this problem in our lab.

Comment # 5 by reviewer #1 The authors did a great job in the discussion. The discussion is very well written with carefully elaborating in detail each of the results of the questions that they put forward in the introduction. Also, their result of how cheating interactions could promote modularity at the cost of reduced nestedness in mutualistic network is another result that was probably not known before. A recent model on adaptive niche-based interaction indicated the emergence of nestedness as well as modularity in mutualistic networks Cai et al 2020 Nature comms. I suggest the authors cite this paper too and discuss on how their result could be different/similar to this particular study.

Response by the authors: We thank the reviewer for the suggestion we have now the reference in the discussion and we discuss how future work could integrate the evolutionary effects of cheaters to the study of network robustness.

Comment # 6 by reviewer #1 Nestedness could be beneficial in certain aspects but has been shown to lead to sudden species collapses (Lever et al 2014, Baruah 2021). I would like to have the authors dedicating a few sentences on how their results i.e., the emergence of trait disparity, modularity and lower nestedness could impact stability of biodiversity. The discussion currently is devoid of such and would greatly improve from this.

Response by the authors: We have followed the reviewer's advice and we have now added a discussion synthesizing our results with the literature of stability and robustness of mutualistic networks.

Revisor 2

Comment # 1 by reviewer #2: I think the paper would benefit from removing the parts on the effect of cheater centrality on trait disparity and clustering. To me, the aspect of centrality is more of a detail that distracts the reader from the other more general findings. Also, the authors conclude that there is essentially no effect of cheater centrality. I do not recommend to entirely drop this result (since such negative findings are also important), but I suggest moving it to the supplementary document (and just refer to it in the main text).

Response by the authors: We agree with the reviewer that the question *ii*) of our paper is, at the end, a technical detail. We remove the analysis of cheater's centrality from the main text, moving it to the supporting information and referring to this result in our main text.

Comment # 2 by reviewer #2: Throughout the manuscript and particularly in the introduction and the discussion the writing must be improved. In my review, I focused mainly on methodological and conceptual aspects of the paper. Hence, I commented only occasionally on issues related to writing but I suggest a comprehensive editing of the manuscript to improve both grammar and clarity (e.g., rewriting and/ or splitting of long and cumbersome sentences).

Response by the authors: We review the entire manuscript, seeking to improve both grammar and clarity, as suggested by the reviewer.

Comment # 3 by reviewer #2: L43: Yoder & Nuismer, 2010 also suggests that the outcome of co-evolution can depend on the interaction type.

Response by the authors: We added the reference in the first paragraph of the main text.

Comment # 4 by reviewer #2: L46: This sentence does not read well. I would change "matching the flower depth that it pollinates" to "matching the depth of the flower it pollinates".

Response by the authors: Done.

Comment # 5 by reviewer #2: L64: I don't understand the second part of the sentence (starting from "highlighting the..."). Please rewrite.

Response by the authors: Done.

Comment # 6 by reviewer #2: L71: "act as antagonisms" → "act as antagonists", "as mutualisms" → "as mutualists"

Response by the authors: Done.

Comment # 7 by reviewer #2: L74: “exploring” → “exploiting”

Response by the authors: Done.

Comment # 8 by reviewer #2: L90: “a rich-species communities” → “species-rich communities”

Response by the authors: Done.

Comment # 9 by reviewer #2: L92: “to each other” → “with each other”

Response by the authors: Done

Comment # 10 by reviewer #2: L93: The sentence starting on this line is very long and cumbersome. Please rewrite.

Response by the authors: We agree with the reviewer. However, this sentence was removed of the manuscript because we have followed the reviewer’s advice and removed the analysis of the centrality of the cheaters from the main text (see our response to comment 1 by the reviewer #2).

Comment # 11 by reviewer #2: L100: Wechsler & Bascompte, 2021 can be cited earlier as a study where the effect of cheating on co-evolution in mutualistic communities has been investigated (basically a similar question than yours but using a different class of models). The current referencing implies that the study deals with pair-wise interactions and the role of cheater centrality - which is not the case.

Response by the authors: We agree with the reviewer and we have now cited Wechsler & Bascompte, 2021 in the excerpt of the text in which we discuss the importance of considering multiple interactions simultaneously for coevolution.

Comment # 12 by reviewer #2: L106: How about just writing: “How does the frequency of cheating interactions affect co-evolutionary dynamics”?

Response by the authors: Done.

Comment # 13 by reviewer #2: L115: I don’t really understand the mechanism you are explaining here (in the sentence starting with “By assuming ...”). Consequentially, it’s not clear why you expect modularity to increase with the frequency of cheating interactions (the last sentence of the paragraph is also quite cumbersome and hard to understand). Also, you are talking only about your expectations with regard to modularity. In Figure 1b, you also indicate your expectations on how cheating affects nestedness. Why do you expect an opposite trend there?

Response by the authors: We apologize for the lack of clarity in the previous version of the manuscript. We expect that a higher frequency of cheating interactions lead to higher arms race dynamics which increases trait disparity. If we assume that trait match is essential to the occurrence of an interaction, higher trait disparity would

lead to a sparse network, in which we observe less interactions compared to a community without cheaters. By disrupting interactions, trait disparity may generate clusters of species with similar traits, leading to modularity. We have now clarified this reasoning in the manuscript.

Comment # 14 by reviewer #2: L137: I am confused about the quantity N (especially with respect to its use in Equations 1, 2 and consequentially 5). My understanding is the following: To calculate S_m , the sum in Equation 1 should run over the set of all species with a positive effect on species i (hence, all species j with $a_{ij} = 1$). Likewise, to calculate S_a , the sum in Equation 2 should run over the set of all species with a negative effect on i (hence, all species j with $a_{ij} = -1$). If my understanding is correct, you need to change the notation (since both in Equation 1 and 2 you write \sum_j^N). Maybe use something like M_i to refer to the set of all species with a beneficial effect on species i and C_i for the set of all species which exploit species i . Alternatively, you may add a condition in the sum. Something along the lines of $\sum_{j, a_{ij}=1}^N$, where N is the total number of species. The second notation may be preferable since in Equation 7 and in the supplementary document (section “Network description”) you use N as the total number of species in the community.

Response by the authors: We have followed the reviewer’s advice and we have changed the notation of Equation 5 both in Main Text and in the Supporting Information accordingly.

Comment # 15 by reviewer #2: L148: “explored” → “exploited”

Response by the authors: Done.

Comment # 16 by reviewer #2: L152: It was difficult to understand Equation 2. In particular, the \pm notation and the corresponding explanation in the text (line 158) is confusing. In fact, I had to consult the provided reference (Andreazzi et al. 2017) to understand it. Especially, the wording “the signal of ε_{ij} will be positive [negative]” is ambiguous. From my understanding, it should be the case that: If the trait of victim i is smaller than that of exploiter j (i.e. if $Z_j - Z_i > 0$) then the victim trait gets even smaller. Likewise, if the trait of victim i is larger than that of the exploiter j (i.e., $Z_j - Z_i < 0$) then the victim trait gets even larger. Please better explain the equation and consider using another notation – maybe like in Equation 4 in Andreazzi et al. 2017.

Response by the authors: We made a typo when describing the model - indeed our model followed Andreazzi's work. We have now followed Andreazzi et al. 2017. When the trait difference is positive ($Z_j - Z_i > 0$) than the value of ε_{ij} will be negative. Contrary, if $Z_j - Z_i < 0$, ε_{ij} will be positive.

Comment # 17 by reviewer #2: L153: So, if $|Z_j^{(t)} - Z_i^{(t)}| > \varepsilon_{ij}$ there is no selection of the cheater j on the victim i . But it seems there is still selection imposed by the victim i on cheater j . Right? Hence, talking of a "trait barrier" that prevents interactions due to dissimilarity is not adequate (since there needs to be some sort of interaction for i to impose selection on j).

Response by the authors: We apologize for the lack of clarity. It is a trait barrier for an antagonistic interaction, in which one individual is benefited by the partner, whereas the other individual shows fitness loss. The interaction occurs in a different form, as a commensalistic interaction that has the potential of evolving back to an antagonistic partner if selection is strong enough to trait distances become smaller than the threshold.

Comment # 18 by reviewer #2: L167: Please indicate the range of values the parameter φ_i can take.

Response by the authors: We have used a single value for the φ_i parameter. We did so because previous work in our lab showed that φ_i just speed up the coevolutionary changes without implying in changes in the coevolutionary outcome (Guimarães et al. 2017. Nature). Having said that, we performed a new sensitivity analysis in which we show that the main patterns of trait disparity are maintained even with different values of φ_i . We have also added the theoretical ranges of φ_i in the text.

Comment # 19 by reviewer #2: L169: Please indicate the range of values the parameter γ_i can take (e.g. $\gamma_i \in [0,1]$).

Response by the authors: As with φ_i , we have used a single value for γ_i . For both parameters we write in the new main text that they assume a single value. We have performed a sensitivity analysis in which we show that the main patterns of trait disparity are maintained even with different values of γ_i . We also added the theoretical ranges of γ_i in the text.

Comment # 20 by reviewer #2: L172: I suppose that the term representing the environmental component of selection in Equation 5 should be in brackets. Hence, $\gamma_i(\theta_i - Z_i^{(t)})$ instead of $\gamma_i\theta_i - Z_i^{(t)}$.

Response by the authors: We apologize for this typo. We have changed the notation of equations 4 and 5, inserting the parentheses.

Comment # 21 by reviewer #2: L176: What are the variables a_{mij} and a_{vij} ? Are those the entries of the adjacency matrix A after transforming p connections from 1 (mutualistic) to -1 (antagonistic)? Hence, essentially $a_{mij} = 1$ and $a_{vij} = -1$? Also, what is the difference between a_{mij} and a_{ij} (as used in Equation 11)?

Response by the authors: Those are the entries of the M and V matrices. M is the matrix which denotes all the mutualistic interactions and V is the matrix which denotes all the cheating interactions. As described by the reviewer, $a_{mij} = 1$ and $a_{vij} = -1$. a_{ij} is a generic term describing an interaction. If we are using the M matrix $a_{ij} = a_{mij}$ and if we are using the V matrix $a_{ij} = a_{vij}$. We have now clarified that in the text.

Comment # 22 by reviewer #2: L180: Is the parameter α the same as the parameter A listed in Table 1? Please, clarify.

Response by the authors: It was a typo. We have changed for α .

Comment # 23 by reviewer #2: L183: What is the range of values the different model parameters can take? It is nice that Table 1 provides an overview with the baseline values (i.e., the values used in the simulations) but it would be helpful to add another column indicating the values these parameters can in principle take (see also comments L167 and L169). Also, did you explore the sensitivity of your results to changes in the parameters (especially ε , b , ϕ , and γ)? For the mentioned parameters, a sensitivity analysis should be performed and it should be reported how broad the conditions are under which the presented results hold.

Response by the authors: We have followed the reviewer's advice and we have now added the range of values that all our model parameters can take. We have now performed the sensitivity analysis of all parameters listed by the reviewer (see the Supporting Information). We found that although variation in parameter values had an impact and the changes in modularity, nestedness, and number of clusters, the results are qualitatively similar to those from the baseline simulations reported in the main text.

Comment # 24 by reviewer #2: L186: You are considering bipartite networks. However, instead of representing them as incidence matrices, where one group of species is represented by the rows and the other by the columns, you use adjacency matrices. This is fine but it would help to illustrate such a matrix graphically. For that, you could reference a modified version of the matrix in Figure 1a (modified because in that example all animals happen to interact with all plants which is not

representative for your networks). It also seems that you imply (in Figure 1a) that the rows and columns are grouped by the two types of species such that the upper left and lower right blocks contain only 0's while 1's occur only in the lower left and upper right blocks. If so, please make this explicit. Please, also indicate the values the entries of the matrix can take. I assume 0, 1 and -1 (after introducing cheating connections).

Response by the authors: The comment by the reviewer made us realize that the figure was not working as a conceptual figure that, by definition, should be simple. We substitute the matrices by a network. We hope that this change will make our model approach simpler to follow.

Comment # 25 by reviewer #2: L203: The sentence starting with "To reduce computational time, .." does not sound correct. It's clear what you mean but the sentence should be rewritten. Particularly, the last part "ending the simulation" seems misplaced.

Response by the authors: Done.

Comment # 26 by reviewer #2: L227: In the adjacency matrix shown in Figure 1a), the rows and columns are grouped by means of the two species types (e.g., animals and plants). Hence, non-zero entries (i.e. interactions) occur only in the lower left and upper right blocks of the matrix. The entries in the lower left block denote the effects of the animals on the plants and those in the upper right block the effects of the plants on the animals. I am wondering if the rows and columns of the matrices used to parameterize your simulations (i.e., the real mutualistic networks) are also grouped in that manner? If yes, then "only applying the transformation from positive to negative effects [...] in the lower triangle" restricts cheating to only one group of species (in Figure 1a the animals). In case my suspicion is correct, give a justification for your choice and make explicit that only one group can cheat. Also, how would the results change if this assumption is relaxed?

Response by the authors: As with the comment #24 by the reviewer 4, this comment by the reviewer made us realize that the figure was not working as a conceptual figure. We substitute the matrices by a network, that is a more straightforward illustration of our modeling approach.

Comment # 27 by reviewer #2: L248: Please add a brief explanation on how the clustering works. Also, provide a rational of why clusters are actually considered (the justification could also go to the introduction). To me, such clusters in phenotype space are very much related to diversification /speciation. See for example: Doebeli & Ispolatov, Science, 2010 or Doebeli & Dieckmann, Am. Nat., 2000.

Response by the authors: In short, the algorithm to detect clustering sorts trait values in groups, creating clusters which minimizes the intra-cluster variation in

trait values, maximizing the inter-cluster trait distances. In intraspecific models and analyses, this approach may be useful to find phenotypic groups and record the emergence of incipient species, as indicated by the reviewer. Because, we are measuring distances among mean species traits, in our case, clusters indicate groups of species that show similar trait values, indicating complementarity between partners (trait matching) and convergence among species at the same group (e.g., among plants in the same cluster).

Comment # 28 by reviewer #2: L274: Is \bar{C} the average value of the degree centrality of the species in the set of species where i belongs to or the average value of all species in the community? Similarly, is σ computed only based on the set species i belongs to or all species?

Response by the authors: \bar{C} is the average value across all species in the network. Having said that, of C_i that was calculated as: $C_i = \frac{k_i}{N_0}$ where k_i is the number of interactions of species i and N_0 is the quantity of species in the opposite set of i . Thus, it controls for the effects of asymmetries between species richness between sets when estimating the degree centrality of the species. We computed σ across all species in the network.

Comment # 29 by reviewer #2: L311: Why are you focusing only on the loss of interactions? I don't see an obvious reason for this choice. Please justify it. Also, it seems straightforward to test scenarios where new interactions can evolve. Please investigate such a scenario to see whether and how it affects your results on the effect of cheating on network structure.

Response by the authors: We have followed the reviewer advice and we performed additional simulations in which we allow a fraction of interactions to reconnect due to trait similarity. This variation of the model did not change the patterns of change in modularity and nestedness we reported in the baseline model in the main text. We have added this analysis to the Supporting Information.

Comment # 30 by reviewer #2: L319: The sentence starting on this line does not read well. Please rewrite.

Response by the authors: Done.

Comment # 31 by reviewer #2: L346: Concerning the histograms at the bottom of Figure 2. I like that you present these examples. They help to get an idea of the actual system state at equilibrium. I however don't fully understand the explanation in the figure caption. In particular, I don't understand what you mean by: "where the left histograms show the increased trait disparity, and the right histograms shows the trait group formation.". What are the left and right histograms? Those with $p = 0.01$

and $p = 0.9$, respectively? Why are there two histograms per considered value of p ? Please clarify. Also, it would be good to reference the histograms using a separate letter as a figure label (e.g. Figure 1g).

Response by the authors: Following the reviewer's comment we realized that the histograms were neither helpful nor needed to illustrate our predictions and we removed it from the manuscript.

Comment # 32 by reviewer #2: L347: "not dependent on the structure of the network" is too strong. There may well be structural features of networks that can modulate this effect. It's just that there is no effect between the three different types of networks you considered.

Response by the authors: Totally agree. We have changed the text to "the effect did not change between networks of the different mutualism types included here"

Comment # 33 by reviewer #2: L379: It seems that the highest values of modularity (the lowest values of nestedness) are reached for intermediate levels of cheating (Figure 4). To me, it's interesting that the relationship between p and the number of clusters (Figure 2d-f) has a similar unimodal shape. Did you explore a potential connection between the number of clusters and network structures (i.e., modularity and nestedness)? One may expect a connection since, in the third part, traits determine interactions through the use of the b_{ij} parameters.

Response by the authors: We used the third part of the manuscript to actually explore the problem mentioned by the reviewer. We have now also pointed out the reasoning behind the connection between the changes in trait distributions and the reorganization of network structure. With a low frequency of cheating interactions, the dominance of mutualistic selection can drive a global trait convergence, resulting in only one or a few large trait clusters. Because under our approach traits modify the binary structure of networks, modularity will be low in this situation and nestedness will increase. With an intermediate frequency of cheating interactions, these interactions may displace the traits of certain species (through directional selection), breaking large trait clusters (and, therefore, modules) into many small clusters. These small clusters probably consist of mutualistic partners with convergent trait values. With a high frequency of cheating interactions, cheaters that impose a similar directional selection will create a few large trait clusters containing their victims. Thus, in this last scenario, the number of clusters is small even though trait disparity is very high.

Comment # 34 by reviewer #2: L380: From Figure 5, it seems clear that co-evolution with cheating leads to higher modularity compared to a scenario where connections

are removed randomly. For nestedness, however, the difference is not so clear (especially not for seed dispersal interactions). To assess the significance of these results it would be good to calculate a p-value. You could do so by creating for each co-evolved network (i.e. simulation) a large number of randomized networks (i.e., networks where the same number of links as in the co-evolved network are randomly removed). A one sided p-value can then be calculated as the fraction of randomizations with a larger increase (decrease) in modularity (nestedness) than the co-evolved network.

Response by the authors: As discussed in Comment #15 by the editor, statistical analysis should be done with caution in simulation data due to the possibility of generating extremely low p-values due to the high number of simulations performed, as discussed by Hartig et al. 2011. Ecol Lett. We opted to focus on interpreting the trends generated by the model, as suggested by Hartig et al. 2011. Ecol Lett. and usually done in theoretical studies in other disciplines, such as Physics.

Comment # 35 by reviewer #2: L384: “In mutualisms that the empirical ...” → “In mutualisms where the empirical ...”

Response by the authors: Done.

Comment # 36 by reviewer #2: L416: “are be” → “are”

Response by the authors: Done.

Comment # 37 by reviewer #2: L423: Can you be a bit more specific about the 1:1 proportion of positive and negative effects reported in Dodds, 1997 and how this relates to your results? Also, how exactly could your theory be tested empirically?

Response by the authors: We rewrote this excerpt of the discussion to address the points raised by the reviewer. We note, however, that we kept that brief due to space limitations.