

Dear Prof Bilde,

Thanks for your decision on our manuscript now entitled "Alteration of gut microbiota with a broad spectrum antibiotic does not impair maternal care in the European earwig" and submitted to *Peer Community in Evolutionary Biology*.

We are happy to see that you and the 3 reviewers liked our study and agreed on the fact that it provides novel insights into our general understanding of the link between gut microbiota and parental care in animals.

We have addressed all the suggestions formulated by you and the reviewers. Specifically, we 1) tuned down some of our predictions, 2) edited the methods section to clarify why we cannot test intra-individual correlations between gut microbial communities and maternal traits and explain why it was not necessary to address our main research question, 3) provided additional statistical tests to demonstrate the robustness of our non-significant effects, 4) removed the few data on offspring development and survival as they were out of the scope of this study and 5) moved contents from the supplementary material to the main text. We have also carefully considered all the other comments and incorporated changes that we believe should satisfy reviewer concerns.

We warmly thank all of you for the time you have spent assessing our manuscript and for your very detailed and constructive comments. This was of great help for improving the general quality of our study.

Below, you can find a detailed point-by-point response (in black) to the different comments (in blue) and the corresponding changes in the manuscript (in yellow).

Sincerely,

S. Van Meyel, S. Devers, S. Dupont, F. Dedeine and J. Meunier

#####  
REVIEW BY TRINE BILDE (extracted from pdf annotations)  
#####

L32. "studies"

Changed (L31).

L49. "showing"

The entire sentence has been changed (based on comments from other reviewers).

L63. this is a very strong statement, tone down to suggest the microbes could alter the investment in parental care.

We have edited the sentence accordingly: "On one hand, gut microbes could indirectly alter the investment in parental care, because parents are expected to adjust their level of care to their own condition [24] and altered gut microbial communities can lower these conditions in multiple ways (see above)." (L59-61)

L71-72. I think some moderation might be useful here, there are so many benefits to parental care that the idea that it is dived by gut microbiota is not justified. Revise to state that we know very little about whether and how the gut microbiota influences the investment in parental care.

You're right. We changed the sentence accordingly: "However, we still know very little about whether and how the gut microbiota influences the investment in parental care." (L67-68)

L193. explain the interpretation of this measure in more detail.

We have edited the sentence to provide more information on the interpretation of this measure and add a reference supporting this interpretation: "Proxies of females' physiology were the number of feces pellets produced per 24 hours (a number positively associated with their digestive/foraging activity [60]) and the gain in fresh weight between two life stages. (L197-200)

L223. these are typically referred to as taxonomic units ASV (Amplicon Sequence Variants) or OUT the term 'species' is not precise, as this level of resolution is not possible with this analysis.

Sorry for this mistake. We have corrected the text: "... which estimate the number of OTUs in the microbiome ..." (L211).

L335: "host's"

Changed (L333)

L338-349. it is a very strong prediction, given that there is no discussion of particular microbial taxonomic groups, or how they are transmitted. Phrase with more care

We have edited the sentence to tune down the strength of this prediction: "One might have expected that gut microbes directly drive the expression of parental care, as enforcing this social behaviour may allow (at least some) symbionts to reach new hosts (i.e. offspring) that are typically young (thus offering long-lived habitats), display poor immune defences (thus facilitating bacterial establishment and development [68]) and harbour only a few resident microbes (thus limiting the risk of competition within the microbiome [28]). However, our results are at odds with this prediction. This." (L336-341)

L347. "studies"

Changed (L347)

L403-404. Although not the primary focus, and in appreciation of the need to keep the text short, it would be interesting to consider what the function of the gut microbiota is, if its removal does not affect any traits?

This would be interesting, indeed. However, considering that this part is not the primary focus of our study (but should be a follow-up), that it would be highly speculative and would lead to extend the length of our (already long) manuscript, we gently ask to not address this point.

L407-408. This statement is at odds with the strongly formulated 'hypothesis' presented in the introduction, that the microbiome might 'drive' social interactions. Please reconcile these statements by being more nuanced in the introduction.

We have followed this suggestion. In particular, we have edited the sentence presenting this hypothesis in the introduction to provide more nuanced predictions: "On one hand, gut microbes could indirectly alter the investment in parental care, because parents are expected to adjust their level of care to their own condition [24] and altered gut microbial communities can lower these conditions in multiple ways (see above). On the other hand, gut microbes could serve as a direct promoter of parental care because, by enforcing the expression of care in adult hosts, parental

gut microbes could maximize their chances to reach novel hosts [25–30](but see [31])” (L59-64).

#####  
**REVIEW BY NADIA AUBIN-HORTH**  
#####

Van Meyel and colleagues aim to quantify if modifying the gut microbiota of females in a species exhibiting maternal care will affect directly or indirectly their level of care to their offspring. To assess if this effect (if any) is direct or through effects on other parts of the phenotype, they also quantified physiological and reproductive traits in these females. To obtain females with altered gut microbiota, they used a treatment with a broad-spectrum antibiotic (rifampicin) known to alter gut microbiota in other insect species at this dose. They first confirmed that indeed this treatment affects the structure, composition, and diversity of the microbiota in earwig females. They then showed that none of the 6 maternal care behaviours were significantly changed in females treated with the antibiotic rifampicin, but affected 3 life-history traits.

This ms presents a clear lack in our knowledge of the association between gut microbiota and maternal care behaviour and a clear objective. The method is straightforward for a non-specialist of the microbiota like myself and the quantification of the different parental behaviours is adequate. The measurements of life-history traits in females is also based on standard protocols. The authors had planned a paired design to analyse effects of treatment on behaviour and life history traits, with siblings in the control and treated groups to be compared directly, but mortality prevented them to use this approach, and they acknowledge clearly that they replaced it by general average comparisons (and they provide the paired analysis on a smaller sample in supplementary material). They also used the appropriate adjustment of p-values to account for multiple testing. The main text contains enough explanations and the supplementary material is exhaustive and well-made.

The results are clear and well presented. The figures are clear and I think presenting individual samples on the box-plot is an excellent idea, as it helps the reader appreciate biological variation within a treatment. The treatment had very little effects on the phenotypes studied, but it does not seem to be the result of a lack of statistical power. The authors also discuss the possibility that while the antibiotic treatment did affect the gut microbiota, it may have not targeted the microbes (fungi, protist...) that may be the driver of behaviour (if any). The data is available on public repositories.

This manuscript is well-written, clear and presents significant new knowledge about this question, using a well-designed experimental study.

Thanks a lot for these very positive comments!

Minor comments

I would like to see the variables presented in figure 4 included in table 1 instead of table S1

Excellent point. We have added 4 additional traits in **table 1** to include the variables presented in the different figures of the manuscript.

#####  
**REVIEW BY GABRIELLE DAVIDSON**  
#####

This is a timely and valuable study investigating the role of the gut microbiota and host behaviour in the European earwig. The authors manipulated the gut microbiome by administering antibiotics

and tested whether this treatment affected various measures of parental care and a suite of other physiological and offspring traits. While the antibiotic treatment affected beta diversity, there was no effect on maternal care. The treatment affected the number of faeces the female produced, and her young were lighter. Overall this is a well written and designed study, and I found the Discussion particularly thorough and in line with the results.

Thanks a lot.

However, I wonder why the authors chose not to include any correlations between the microbiome and the maternal traits, and to look at links between lower-level taxonomic abundance and parental care. If the authors choose not to add this analysis, they should highlight this as a potential follow-up to confirm no links between gut microbiota and maternal care (even in the absence of any experimental treatment).

We apologize for the misunderstanding, which likely stems from unclarity in our experimental design. Our experimental setup does not allow to draw any correlation between the microbial community and the maternal traits because these two parameters have been measured on different individuals. In particular, the measurements of gut microbial diversity has been done in females dissected before oviposition (i.e. in females for which we have no measurement of maternal traits) and on females dissected at egg hatching (i.e. in a few females for which we only have a few measurements of pre-hatching maternal traits).

We have edited the method section to emphasize the use of different set of females: "To test whether and how rifampicin treatment altered the earwigs' gut microbial communities, we extracted the gut of 10 females per treatment (n total = 20) on the day we observed the first oviposition of their 1st clutch (i.e. about 2 months after being fed with or without rifampicin), and 10 rifampicin- and 8 water-treated females one day after their 1st clutch eggs have hatched (i.e. about 1 month later; Figure S1)." (L142-146).

We have also followed another suggestion (see below) and added a timeline of the experiment in the supplementary document (Figure S1) emphasizing the origin of the dissected females, and added the presently suggested follow-up in the discussion: "...future studies will be required to confirm that no other members of the gut microbiota shape parental care in our study species, and to explore causal links between the presence of certain members of the microbiome and the level of maternal care expressed by its host. ." (L347-350)

Finally, I am curious about the variation of parental care traits in the cohort studied here and whether there was a sufficient range of behavioural phenotypes to detect differences. For example, looking at figure 3, Egg grooming duration and egg defence seem to have a wide range of expressions, whereas Nymph searching and Nymph defence have very little population variance. This may be something that could be added to the discussion, or in the methods to describe whether these ranges of phenotypes in the experimental cohort are typical of this species.

This is an important point. Thanks for raising it. The ranges of phenotypes measured in the present study are comparable to the values obtained in previous studies in other earwig populations using the same protocols. Our values are thus likely to be typical of this species and to reflect the natural variation in maternal care exhibited by earwig females. We have added this information in the method: "The recorded range of values of maternal care is comparable to the range of values obtained in previous studies conducted in other populations [41,42,44] and thus likely reflects the natural variation in maternal care exhibited by earwig females." (L192-194)

Abstract:

Line 24-25: You have not ruled out developmental effects/critical developmental windows of the gut microbiota on host traits. Moreover, there may be taxa that were unaffected by the antibiotic treatment that could be linked with parental care. Suggest changing to highlight that these findings are limited within the context of this specific antibiotic treatment.

This is true. We have edited the abstract accordingly: "Overall, these findings reveal that altering the gut microbiota with a large spectrum antibiotic such as rifampicin does not necessarily affect host sociality." (L21-22)

Lines 50-56: Please provide further details of the suggested mechanisms. I believe there are at least two – one whereby the microbiota correlates with host social behaviour, and one where the gut microbiota correlates with conspecific's behaviour. For example, does the microbiota contribute to pheromones, thus affecting conspecifics (rather than host) social behaviour? It would also be useful for the authors to comment on whether these relationships are correlational only, or if there is evidence of causation.

This is an excellent suggestion. We have profoundly edited this paragraph to provide this information. "Recent studies also suggest that gut microbiota can play a critical role in the sociality of their hosts by shaping the expression and nature of social interactions and/or by transforming mediators of social aggregation. For instance, family-living rats with a diet-altered gut microbiota exhibit deficient sociability and increased social avoidance [15,16]. Antibiotic-induced modifications of gut microbiota also alter the chemical signatures of social hosts and lead to higher levels of aggressiveness toward conspecifics in the leaf-cutting ant *Acromyrmex echinator* [17] and the honeybee *Apis mellifera* [18]. Finally, alteration of the gut microbiota reduces the production of aggregation pheromones in the swarm-living desert locust *Schistocerca gregaria* [19] and diminishes the presence of aggregation pheromones in feces of the gregarious German cockroach *Blattella germanica* [20]." (L45-57)

Line 76-43: How much individual variation is there in parental care in this species? Is there repeatable individual variation in parental care, or do mothers only have one brood? Is there any studies on phenotypical plasticity in this species that may be useful when considering whether there is 1) scope for variation in parental care due to environmental inputs (e.g. reaction norms), and 2) whether parental care may be intrinsically linked, either genetically or with the gut microbiome.

We have edited the introduction to add this important information: "Earwig females present important inter-individual variation in the expression of maternal care within populations [41,42], and this variation is partly inherited from the parents [43] and partly depends on environmental inputs, such as the social environment or food resources [44–46]." (L81-84)

Lines 93-97: I like this approach. Short of doing functional analyses, this is a good method for identifying/excluding co-varying factors that may explain any relationships between microbiome and maternal care.

Thanks.

Line 101: Were the subjects virgins?

The subjects were either mothers (mated) or juveniles (virgins). Because we have removed all measurements conducted in juveniles in this new version of the manuscript, we believe there is no more room for unclarity regarding this point. Nevertheless, we have edited the sentence to improve clarity: "[...] we investigated the effects of rifampicin on 24 other traits measured throughout the mothers' lifetime and reflecting their general physiological state, investment in future reproduction and longevity." (L94-96)

Line 109-111: Please could you expand on how cytoplasmic incompatibility arises due to “inter-familial microbiome variability”

We have briefly expanded on this: “There are only limited signs of inbreeding depression in this species [48] and sib-mating allowed us reducing the risk of poor reproductive success due to possible inter-familial cytoplasmic incompatibility between certain bacterial strains, as reported with Wolbachia and Cardinium in several arthropod species [49,50].” (L112-116)

\*adding to this comment after reading supplementary: the supplementary is a lot more detailed and easier to follow. I would suggest merging the section on Rifampicin and insect rearing in supplementary to the main text and omit it from supplementary entirely.

We have followed this suggestion and moved the section on Rifampicin and insect rearing from the supplementary material to the main text (for clarity, this change is not highlighted in the main text).

Line 125: How did you ensure the moistened sand was not a source of bacterial contamination?

The sand was not sterilised, and we thus cannot rule out that it was a potential source of bacterial contamination. However, we believe that it contamination is very unlikely and that it would not be a major issue to robustly address the main question of our study because 1) we have used this type of sand for earwig rearing over the last 10 years and always obtained a comparable level of variation in term of maternal care (showing that putative sand microbes do not homogenise the presence of earwigs gut microbes possibility associated with the expression of maternal care), 2) the use of this sand never triggered an excess of earwig mortality suggesting that it does not contain pathogenic microbes that could lethally alter earwig gut microbiota, and 3) the sand is very unlikely to contain bacteria that are specific to earwigs and thus to inhibit rifampicin-induced variation in their gut microbial communities. We have nevertheless edited the text to clarify that the sand was no sterilised: “Except when stated otherwise, individuals were always maintained in Petri dishes (diameter 9cm) lined with non-sterile moistened sand.” (L131-132)

Line 137-140: Please elaborate on how females were selected for dissection, as this sample size is quite reduced from the 74 families quoted in line 104. Also, a timeline diagram would be helpful for visualising the parental care stages and when and how many females were dissected for gut microbiome analysis. Presumably some were dissected following their first clutch, and some after their second? Or perhaps I have misunderstood why the females were allowed to have a second clutch.

We apologize again for the lack of clarity. A total of 20 females were dissected just before the production of their 1<sup>st</sup> clutch eggs and 20 other females were dissected on the hatching day of their 1<sup>st</sup> clutch eggs. We allowed (the non-dissected) females to produce a second clutch to test the effect of Rifampicin on their total reproduction. We have edited the method to clarify the origin of the dissected females, added a timeline as **Figure S1**, as well as edited the introduction, the method and the figure legends to emphasize that dissection occurred before oviposition and at the hatching of the 1<sup>st</sup> clutch.

In the introduction: “Specifically, we first determined how the antibiotherapy alters the diversity and structure of the gut bacterial community of females at two periods of their life-cycle (just before the production and at the hatching of their 1<sup>st</sup> clutch eggs) by sequencing 16S rRNA gene (V3-V4 region) amplicons. We then tested the effects of rifampicin on the expression of four pre- and two post-hatching forms of maternal care toward 1<sup>st</sup> clutch eggs and nymphs, respectively. » (L87-92)

In the method: "To test whether and how rifampicin treatment altered the earwigs' gut microbial communities, we extracted the gut of 10 females per treatment (n total = 20) on the day we observed the first oviposition of their 1st clutch (i.e. about 2 months after being fed with or without rifampicin), and 10 rifampicin- and 8 water-treated females one day after their 1st clutch eggs have hatched (i.e. about 1 month later; Figure S1)." (L142-146)

In figure legends: "Guts were sampled either before oviposition or at the hatching of the 1st clutch" (Figure 1 and Figure 2)

Line 154-155: Unless there is a strict word count limit, I would advise including the bioinformatics pipeline in the main text, rather than supplementary.

We have considered this suggestion. However, our manuscript is already particularly lengthy. This is why we would like to keep the bioinformatics pipeline in the supplementary material. Note that this pipeline will be easier to read in the supplementary material, which is now much reduced.

Line 226-229: Because you are using so many different indices of diversity, I think this warrants further explanation of what the difference between Jaccard & B-C are, as well as weighted and unweighted unifracs.

Very good point. We have edited the text to clarify the differences between these indices: "Diversity between the gut microbial communities (beta-diversity) was assessed using 4 measures of community similarity: 1) Jaccard index, which does not consider phylogeny of OTUs but takes into account their presence/absence; 2) Bray Curtis dissimilarity, which does not consider the phylogeny but considers the number of reads assigned to an OTU (i.e. its abundance); 3) UniFrac index, which considers phylogeny but not abundance; and finally 4) Weighted UniFrac index, which considers both phylogeny and abundance." (L224-229)

Line 233: What about second clutch as indicated in line 120? Or was that not a clutch used in the current study? Please clarify.

Sorry again for the lack of clarity. We did not analyse gut bacterial communities in females during the 2<sup>nd</sup> clutch (only during the 1st clutch). We have edited the text to clarify our experimental setup (see above).

Line 267: Specify whether you mean the decrease in alpha diversity was overall independent of the antibiotic treatment. Even better yet would be to just state the alpha diversity did not decrease in the antibiotic treatment compared to the control, which I think is what the authors mean?

We have edited the text accordingly: "Similarly, the  $\alpha$ -diversity did not decrease in the rifampicin treatment compared to the control (Chao1:  $F_{1,34} = 0.72$ ,  $P = 0.401$ ; ACE:  $F_{1,34} = 0.62$ ,  $P = 0.435$ ; Fisher:  $F_{1,34} = 0.59$ ,  $P = 0.447$ ; Shannon:  $F_{1,34} = 1.67$ ,  $P = 0.205$ ; Simpson:  $F_{1,34} = 0.55$ ,  $P = 0.465$ ; Figure 2), and it was independent of an interaction between female sampling stage and rifampicin treatment (all  $P > 0.525$ )." (L274-278)

Line 280: In what way is this interesting? I would omit the word interesting, unless you plan to elaborate.

We have removed the term "interesting" from this sentence.

Perhaps this is better suited for the discussion. My interpretation is having a significant is important. Having a significant weighted, but not unweighted suggests either deep phylogenetic differences between groups, or differences in relative abundance between groups. Because both BC & Jaccard are non-significant, it suggests that it is the phylogenetic differences that is the most

prominent factor that differs between the community structures. In this case, because weighted is significant, and unweighted is not, the phylogenetic differences is specific to clades that diverged in the more distant past than recent evolved nodes). This makes sense in terms of an affect of a broad-spectrum antibiotic, which probably works on conserved traits of bacteria.

We edited the discussion to add this very nice interpretation of our results. Thanks for that! "Our experiment first demonstrates that the ingestion of rifampicin by earwig females modified the composition ( $\beta$ -diversity) but not the richness ( $\alpha$ -diversity) of bacterial OTUs present in the gut. The fact that rifampicin only shapes indices of  $\beta$ -diversity controlling for phylogeny suggests that OTUs' phylogeny is the most prominent difference between the community structures present in treated versus non-treated individuals. Because weighted uniFrac is significant and (unweighted) uniFrac is not, our results then indicate that this phylogenetic difference is specific to clades that diverged in the more distant past compared to recently evolved nodes, a pattern in line with broad-spectrum antibiotics acting on conserved bacterial traits. Overall, our findings thus confirm that our treatment successfully altered gut microbial communities in earwigs (just like in other animal species [52–54])." (L319-328)

Lines 285-291: Did the authors consider looking at whether diversity correlated with any of these maternal care traits? Because a broad-spectrum antibiotic is a rather blunt tool to manipulate the gut microbiota, there could be taxa that are correlated with maternal care, but just were unaffected by the antibiotic treatment.

As detailed above, our experimental setup does not allow to look at correlation between gut microbial diversity and maternal care traits, because these were measured on different animals.

Similarly, the authors may also consider looking at relative abundance of genus-level taxa to get a better picture of what taxa were affected by the antibiotics.

This is exactly what we did. Sorry for the misunderstanding. As described in the material and methods section, all comparisons of the gut microbiota between treated and untreated insects were conducted using the diversity and abundance of OTUs, which represent the finest taxonomic level we can use in metabarcoding studies. In order to gain in clarity, we have chosen to present results on the microbiota at the taxonomic level of the phylum in the Figure, but all statistics were performed on OTUs.

To make it clearer for readers, we have added this information in the legend of Figure 1: "These results are presented at the phylum level for clarity, whereas statistical analyses of gut microbial diversity were conducted using OTUs". (L629-631)

292-298: Here would also be useful to correlate the microbiome metrics with these traits to shed light on whether these changes in response to antibiotics were dependent or independent of the gut microbiota.

Unfortunately, our experimental setup does not allow to do that (see answers above).

Line 342-344: I agree, this is an important interpretation, that there may be microbiota that are important for parental care, but were not affected by rifampicin. I would suggest having this alternative interpretation in the abstract (see my comment above line 24-25) . You could also add here in the discussion that there may be functional redundancy, where if some bacteria that service the host are lost, others have similar functional roles that replace any lost taxa.

We have edited the abstract to emphasize the fact that our results are specific to the antibiotic we have used: "They also emphasize that not all animals have evolved a co-dependence with their microbiota and call for caution when generalizing the central role of gut microbes in a host biology." (L22-24).

We have also added the hypothesis of functional redundancy: "This may first suggest that earwig parental care is primarily shaped by microbes that are non-sensitive to rifampicin or that non-sensitive microbes can take over this function (functional redundancy). In insects, gut microbial communities do not only encompass a broad diversity of bacteria (among which some are resistant to rifampicin) but also fungi, protists and other microorganisms that could have key roles and functional redundancies in hosts biology [37,60]." (L341-346)

Line 363-364: a lack of sensitivity of the host to gut microbiome perturbations (in the context of parental care) could also be observed if there are developmental affects associated with microbiome-behaviour links. In such a case, it may be there is a critical developmental window in which microbiome manipulation is required to see a behavioural plasticity in parental care.

This is true and that is why we provided Rifampicin to females during their entire adult stage. Our results therefore suggest that if such a process is at play in earwig, it would take place during larval development.

Figures: I like the figures and showing individual data points within the boxplots.

Thanks.

Supplementary: Were there any negative controls carried through library prep/sequencing?

Several controls were carried out during the experiments, and as expected, all these controls were negative. Therefore, no library preparation or sequencing was produced. First, we checked the absence of inter-sample contamination by taking samples from the dissection tools (forceps) and the T1 buffer. Then, we extracted the DNA from these controls and verified the absence of 16S amplification by PCR. During this same amplification phase, we also placed 2 negative controls (water as template) to verify the absence of contamination during the PCR. At the sequencing platform (i.e., the GeT-PlaGe platform), and prior to sequencing, the lack of contamination was checked with a negative control during the second PCR step (water as template). Again, all these controls were as negative. We have added this information in a new section of the supplementary file entitled "Negative controls through library preparation and sequencing" (Supp file).

Supp Mat Line 84: I think Bokulich, N. A. et al. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat. Methods 10, (2013). should be cited here as the < 0.005% filter is their recommendation.

We have added this reference in the supplementary material.

Supp Mat Line 85: As OTU's at 97% similarity?

We used a different method than the classical 97% similarity threshold obtained from distance comparisons. As described in the clustering step, we used the SWARM algorithm [3], which is an aggregative method. We used two aggregation steps: a first denoising step with an aggregation distance of 1 followed by a second step with an aggregation distance of 3. The aggregation distance corresponds to the maximum number of differences between sequences in each aggregation swarm step. The denoising step allows to build very fine clusters with minimal differences, and the second clustering step was run between seeds from the first aggregation step. We have added this information in the corresponding section of the supplementary material.

Supp Mat: In general, there is a lot of overlap in the text between the main and the supplementary, which I think is unusual. I would expect only the additional details need be included, rather than an extended version of the main text.

Sorry about that. We followed this suggestion (also formulated by another reviewer) and have transferred most of the supplementary material in the main text. We have also substantially reduced the supplementary material to avoid any overlap between the two documents.

#####  
**REVIEW BY AN ANONYMOUS REVIEWER**  
#####

A highly influential study by Buffington in 2016 (cit 17) showed that in mice, maternal diet can affect the gut microbial composition in offspring, and altered compositions result in different social behaviors of offspring. The Buffington study has motivated research in other species, and this research tends to show correlations between social behaviors and gut microbiota. The present study uses antibiotic (and control) treatments to test if the expression of female parental care (a key form of social behavior) is modulated by gut microbiota in earwigs. A large number of care behaviors are quantified, but the authors do not detect an effect of different microbiota on care behaviors.

Studies like these, because they show absence of certain patterns, can be very important to prevent the scientific community from extrapolating isolated findings in one animal to animals in general. However, because it is inherently more challenging to show the absence than the presence of an effect, I suggest the authors give more weight towards assessing the robustness of their finding. Specifically, I suggest evaluating two points: 1) what is the power to detect behavioral differences between treatment groups if there were any differences? To conduct power estimates, one would of course require using expected effect sizes (which are unknown), but the authors could use previous behavioral studies to formulate an educated guess for a reasonable effect size range.

We agree with this. We are fully aware that showing an absence of effects is generally challenging and requires a high statistical power. This is why we used a median number of 36 replicates per treatment (L191), a number that is larger than what is typically used in behavioural studies, including studies about maternal care in earwigs (e.g. Diehl & Meunier, 2018; Kramer et al., 2015; Wong & Kölliker, 2012). Our high numbers and the associated high statistical power of our results is acknowledged by Nadia Aubin-Horth, one of the other reviewers.

Having said that, we agree that this might not be easy to interpret for readers that are less familiar with behavioural studies. Therefore, we have added two additional information (Table 1): the effect sizes of each result and the expected N, which is the sample size that would have been required to detect a significant effect with these effect sizes (now detailed in L251-253). These additional values (presented in Table 1) reveal that it would have required an immense number of replicates (more than 500'000 individuals for certain traits) to change our non-significant results into significant results. We believe that it convincingly demonstrates the strength of our non-significant results regarding the measured traits, and emphasize that our conclusion about the absence of effect of rifampicin on maternal behaviours is biologically robust.

2) is the detected shift in microbial communities upon treatment significant enough to expect consequences? (or could it be that, for some reason, the microbial gut communities were, in biological/functional terms, only mildly altered by the treatment?). Again, this would require

evaluation of the biological meaning (“effect size”) of the statistically significant community differences.

The different reasons why rifampicin alterations of gut microbiota do not translate into alteration of maternal traits is the topic of one entire paragraph in the discussion (which has been extended regarding comments from other the reviewers, e.g. about functional redundancies, L341-371). Nevertheless, we would like to add that we gently disagree on the fact that calculating effect sizes is appropriate to address this question. For instance, if only one OTU would be both affected by antibiotics and responsible for maternal care, then we would have a very low effect size (indicative of low statistical significance) on the test exploring the effect of antibiotics on gut microbiota, whereas the effect of Rifampicin on the presence of this specific OTUs would be biologically highly significant. Whereas the use of effect sizes (and expected replicate number) is relevant for univariate approaches (here maternal traits), we believe that it is less efficient for multivariate data sets (here OTUs) and thus not appropriate to address this specific question.

In this respect, I was wondering whether the summary of the 161 unique OTUs into 6 phyla was pertinent, or whether we might lose too much information here. For example, are there any OTUs among the 161 that are completely lost upon antibiotic treatment, as was the case in the mouse study I believe? Or did the treatments mostly affect the relative frequencies of different phyla as appears to be the case from Figure 6 – from this figure it appears that the transient stop in food intake in the control group has in fact a similarly strong effect on microbiota as the antibiotic treatment, which may suggest that perhaps antibiotic treatment did not cause the warranted shift in microbiota communities?

This is an excellent point. Thanks for that. We have added the results of a novel exploration of the data set about the prevalence of the 161 OTUs, which reveals that “The prevalence (i.e. frequency) of these 161 OTUs among the 38 tested females ranged from 0.211 to 1.000 (Table S3). The vast majority of OTUs were found in at least one female in each experimental modality (Table S3), indicating that our rifampicin treatment did not eliminate specific OTUs.” (L265-268). We have also provided the corresponding data in table S1.

We want to emphasize that we used rifampicin to alter gut microbial communities (which it does) and not necessarily to eliminate certain gut microbes (which it apparently does not). The conclusions of our study therefore remain unchanged and – as usual – is tightly linked to our experimental approach and the associated effect on the gut microbial communities.

I was also wondering whether the diversity analyses were in fact based on the 161 OTUs and not the 6 phyla? If so, then please clarify in the text.

We have edited the text to clarify that the diversity analyses were based on the 161 OTUs: “The structure, composition and diversity of the microbial communities were based on the 161 identified bacterial Operational Taxonomic Units (OTUs) (see results) and analysed using ...” (L217-219).

Minor comments.

I was surprised that the authors continued antibiotic treatment after egg hatching as this prevents distinguishing between direct effects of diet on juveniles from indirect effects mediated by maternal care (the focus of the study). A short explanation in the methods of the rationale for this decision would be helpful.

We continued the treatment after egg hatching because the main goal of our study was to determine the effect of Rifampicin on the expression of maternal care (and

other maternal traits), and not on the outcomes of antibiotic-alteration of maternal care in offspring.

This comment made us realise that the results on offspring development and survival were both out of the scope of our study and very difficult to interpret regarding the effect of rifampicin on juveniles (and we only poorly addressed these results in our former discussion). For these reasons, we have decided to remove these (few) measurements from our study. This removal helps to keep the study better focused (on maternal traits), makes it easier to read and follow the method section, and it does not alter the robustness of our conclusions. Moreover, it also helps in reducing the length of the manuscript. We hope that you will understand and agree with this choice.

L7 I would say increase instead of maximise

Changed.

L150 replace “appeared” by “are indicated in”

Replaced.

L187 The number of replicates per treatment or in total?

We have clarified the text: “The number of replicates for each of our measurements ranged from 9 to 59 per treatment (median = 36 per treatment; details in Tables 1 and S1).” (L190-192)

L205 I would reformulate to “whether females produced a second clutch (yes or no)” (you are looking if the event happened or not, not at the probability of it happening)

We have changed the text accordingly: “We finally assessed whether females produced a 2nd clutch (yes = 1 or no = 0) ...” (L210-211)

L236 corrected for multiple testing?

We have added this (missing) information in the main text: “To correct for multiple testing in these post-hoc analyses, the significance level was adjusted to  $\alpha = 0.0375$  using the Mean False Discovery Rate approach [64].” (L236-238). Note that the significance of our results remains unchanged.

L320 “Our experiment first demonstrates that the ingestion of rifampicin by earwig females induced stage-specific modifications in the species composition ( $\beta$ -diversity) of the gut microbiota but did not shape its species richness ( $\alpha$ -diversity).” I am not convinced by this statement as stage-specificity may be caused solely by females not feeding for some time (i.e., a difference between treatment and transient absence of treatment, and not caused by stage-specific effects of the treatment) intake.

This is a good point. Thanks. We have edited the sentence accordingly: “Our experiment first demonstrates that the ingestion of rifampicin by earwig females modified the composition ( $\beta$ -diversity) but not the richness ( $\alpha$ -diversity) of bacterial OTUs present in the gut.” (L319-321). We discussed the possible causes of this stage-specific pattern in the rest of the paragraph (L328-332).

L381. For me, the following statement is a strong overinterpretation: “In the present study, however, heavier eggs and newly hatched juveniles did not translate into higher offspring survival and improved development compared to lighter counterparts. On a proximate level, these findings suggest that rifampicin breaks the association between offspring weight and quality, either due to alteration in gut microbial communities and/or antibiotic toxicity.” I suggest a simple alternative explanation: the ad libitum food available under lab conditions simply masks the effect small eggs would have on juvenile survival/development under more natural conditions.

This is another good point. We have edited the sentence to limit speculation regarding this point: “Light eggs and newly hatched juveniles are often thought to reflect low offspring quality in insects [70], and further studies are required to confirm this association in earwigs.” (L378-380)

Table S3 (with Affiliation Table with BLAST scores) needs explanations for the different column headers.

We have added a sheet in the excel file (Table S3) to provide clearer explanations of column headers.