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. 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 taxanomic 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). 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.

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.

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.

Line76-43: 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.

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.

Line 101: Were the subjects virgins?

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

*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.

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

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.

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.

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 unifrac.

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

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?

Line 280: In what way is this interesting? I would omit the word interesting, unless you plan to elaborate. Perhaps this is better suited for the discussion. My interpretation is having a significant unifrac suggests phylogeny of the community structure 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.

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. 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.

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.

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.

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.

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

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

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

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

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.