23-Sept-2020

Dear Dr. Gareth Jenkins,

We thank you, as well as the Associate Editor and the reviewers for these comments that in our view have greatly improved our manuscript. Please find here our detailed response. We attached an annotated document with changes highlighted in yellow. Moreover, as suggested, we moved our supplementary material into the appendix. We hope that these improvements will make our manuscript acceptable for publication in E&E.

Best regards

Alice Balard for the authors

22-Jul-2020

Dear Dr. Balard:

Many thanks for your manuscript ECE-2020-06-00967 entitled "Coupling between tolerance and resistance differs between related <i>Eimeria</i> parasite species: implications for co-evolution with their mouse hosts" which you submitted to Ecology and Evolution.

As you will see below, comments from our Associate Editor and referees suggest a major revision before your paper can be published. Their comments should provide a clear road-map for you to revise, hopefully improving the clarity and rigour of the presentation of your work. I'd also encourage you to move your supplementary material into the main text (or as an appendix). It makes it much easier for readers to access this information, as they can download the entire paper as a single file. We have an Editorial explaining our thinking on this, if you're interested: https://onlinelibrary.wiley.com/doi/full/10.1002/ece3.2101

Once again, thank you for submitting your manuscript to Ecology and Evolution and we look forward to receiving your revisions.

Sincerely,

Dr. Gareth Jenkins

Editor in Chief, Ecology and Evolution

gjenkins@wiley.com

Associate Editor Comments to Author:

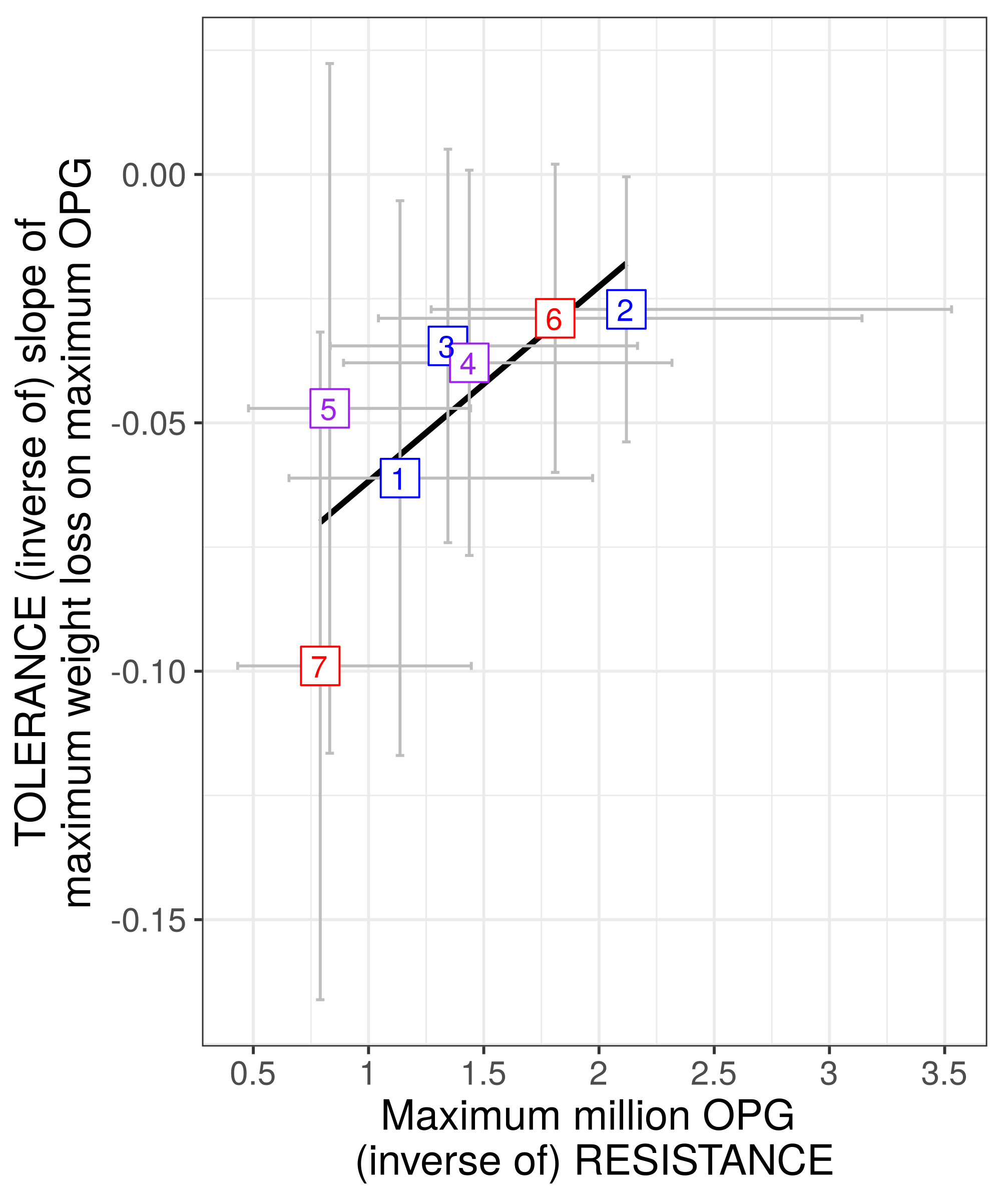
Associate Editor

Comments to the Author:

This study investigates the association between resistance and tolerance of different Eimeria species in different strains of the mouse host to determine whether there is evidence of adaptation to the host. Two reviewers found the manuscript to be on an interesting topic and the study to be largely well-executed. One reviewer had more minor comments focused on clarifying aspects of the history of the host and parasite populations, and made suggestions for other aspects for the Discussion. Another reviewer was more critical, focusing largely on aspects of the analyses, and in particular, the lack of justification for certain decisions, including the sample sizes and the methods used to analyse the data. If the authors pay close attention to these comments, particularly strengthening the rationale for their statistical approaches, this will make a nice paper.

We thank the Associate Editor for his/her comments. Following the useful reviewers comments, we modified the manuscript and provide here a detailed answer to the reviewers concerns. We hope that our manuscript is now suitable for publication in E&E.

C1. I do have one minor comment to add myself. The association between resistance and tolerance depicted in figure 5E is supposedly strong, but it looks as though it’s driven entirely by one data point (number 8). The suggestion that this is unlikely to be a statistical artefact is itself based on a marginal result, and as such I’m not sure that the association in figure 5E is especially robust.

This is a very good comment. We tested the significance of this association by removing the point 8 (corresponding to one mouse strain, PWD) and found similar results ; we did not write that information in the previous version and see now that it was missing. We added an in-text explanation to the manuscript as follows:

“This result was robust to the exclusion of the extreme point corresponding to mouse strain PWD (point 8 in figure 4D; Spearman’s =-0.93, P<0.01).” line 354-356.

For review purpose, here is the corresponding figure (we chose to not add this visualisation in the article for the sake of brevity).

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author

This is a very interesting and well explained study that addresses an interesting and important topic. The paper looks at the coupling between tolerance and resistance in different parasite species. In addition it asks whether there is any evidence for adaptation to specific parasite clones in different mouse host strains. The introduction, methods and discussion are well written and explained, and the results are mostly well presented with a few omissions.

C2. My major concerns are about the sample sizes used and the presentation of the results. First, I’m a little concerned about the ethics here. The numbers of mice used for each treatment combination is very low. It is just as unethical to subject animals to treatments that impact their welfare in numbers that are too small, as too large. Too large impacts more animals than is necessary to answer the question, too small kills animals for no real purpose as the numbers are not large enough to detect an effect. This seems to be the case with the strain by genotype effects shown in figures 3A and B. There are apparent differences between the strains that would likely be detected if more than 6 or 7 animals were used for each treatment. Why so few? This seems like an incredibly small number to determine effects in this kind of study. I would expect to see that a power analysis for this study had been carried for the ethical approval, indicating how many animals were required to show differences between groups. If the number of animals that could be used was limited by time, space or money, then the ethical option is to ask fewer questions with the same number of animals.

We thank the reviewer for raising this point, and understand this concern as ethics in animal experiment is a crucial point that we considered carefully. The mice used in the present article are wild-derived inbred strains and there was indeed a limitation in the number that could be used, as these mice are more complex to obtain than classical laboratory strains. A preliminary study was conducted with NMRI mice and stronger effects were detected (Al-khlifeh et al., bioRxiv, doi: https://doi.org/10.1101/611277), which guided our choice of using 6-7 mice per batch.

Nevertheless, non-significant results should be carefully interpreted. As pointed out by the reviewer, an ethical option is to ask fewer questions. At the light of this comment, as well as the comment C9, we chose to remove the section on host adaptation, as it appears that it is not completely conclusive (see full explanation on comment C9); moreover this was a side-aspect of the article, and its removal - we are convinced now - makes the results presented stronger and more focused.

Regarding the coupling of resistance and tolerance, as the core aspect of our manuscript, the relevant unit of replication is the mouse strain rather than the mouse individual. For this reason, we are confident that our data allows us to answer this core question.

C3. Second, some of the results are oddly presented with no justification. In some cases data is logged, in others not. Main effects are not presented, data are presented in a way that doesn’t match the analyses etc. in one case the analysis presented does not test the stated hypothesis. I have detailed these queries below.

We thank the reviewer for this comment and corrected our manuscript accordingly, see answers to comments below for detailed corrections.

C4. A more minor comment is around the framing. There is some discussion of hybrid zones and the role of parasites. Hybrid mice are used but there is no analysis or discussion of whether hybrids are different to the parental genotypes, more like one than the other, what this might mean for hybrids in the field? This seems like an odd omission.

In this article, we used wild-derived inbred mice that often breed badly. Hybrids were used to increase the number of strains, and therefore the statistical power, but also to investigate the difference of resistance and tolerance in hybrids compared to pure *M. m. domesticus* and *M. m. musculus* in a controlled environment, which is the scope of another article in preparation. As this is out of the scope of the present manuscript, we chose to not focus on this aspect . We however understand that this aspect needs to be mentioned, therefore we added the following :

line 123-127 : “Hybrids between *M. m. domesticus* and *M. m. musculus* are used in the present study solely to increase statistical power for comparisons among strains (such as resistance-tolerance correlations). In future analyses of a hybrid effect (Balard et al. 2020) could investigate tolerance and resistance employing a larger panel of such hybrid strains allowing statistical analysis of an outbreeding effect.”

C5. Lines 209-213: use log OPG, any evidence of curvature in the fits?

We used log OPG only for visualisation, but modelled using (negative binomial) glms on untransformed data. For clarity we now plot untransformed data in all figures (see Figure 3A & 4A).

C6. Line 226: why? Surely if the interaction term is not significant then this is not statistically justified as it inflates the type 1 error rate?

We thank the reviewer for this comment. This was an important addition, which we implemented and added to the manuscript as follows:

”For each of our models that showed a significant interaction term, we also asked within each parasite isolate if the response differed between mouse groups using likelihood ratio tests (G) as described above. In the case of a non-significant interaction term, we performed post-hoc tests corrected for multiple testing (Tukey Honest Significant Differences (HSD)) to compare within all pairwise comparisons between groups (parasite isolate-mouse strain).” lines 230-235.

We reorganised and modified the results accordingly (modifications highlighted, see Results section 2.1). We also accordingly removed the part (in text and in plots) about impact on weight loss, as the interaction factor for this proxy is not significant. This does not change our main results and rationale, but is more stringent.

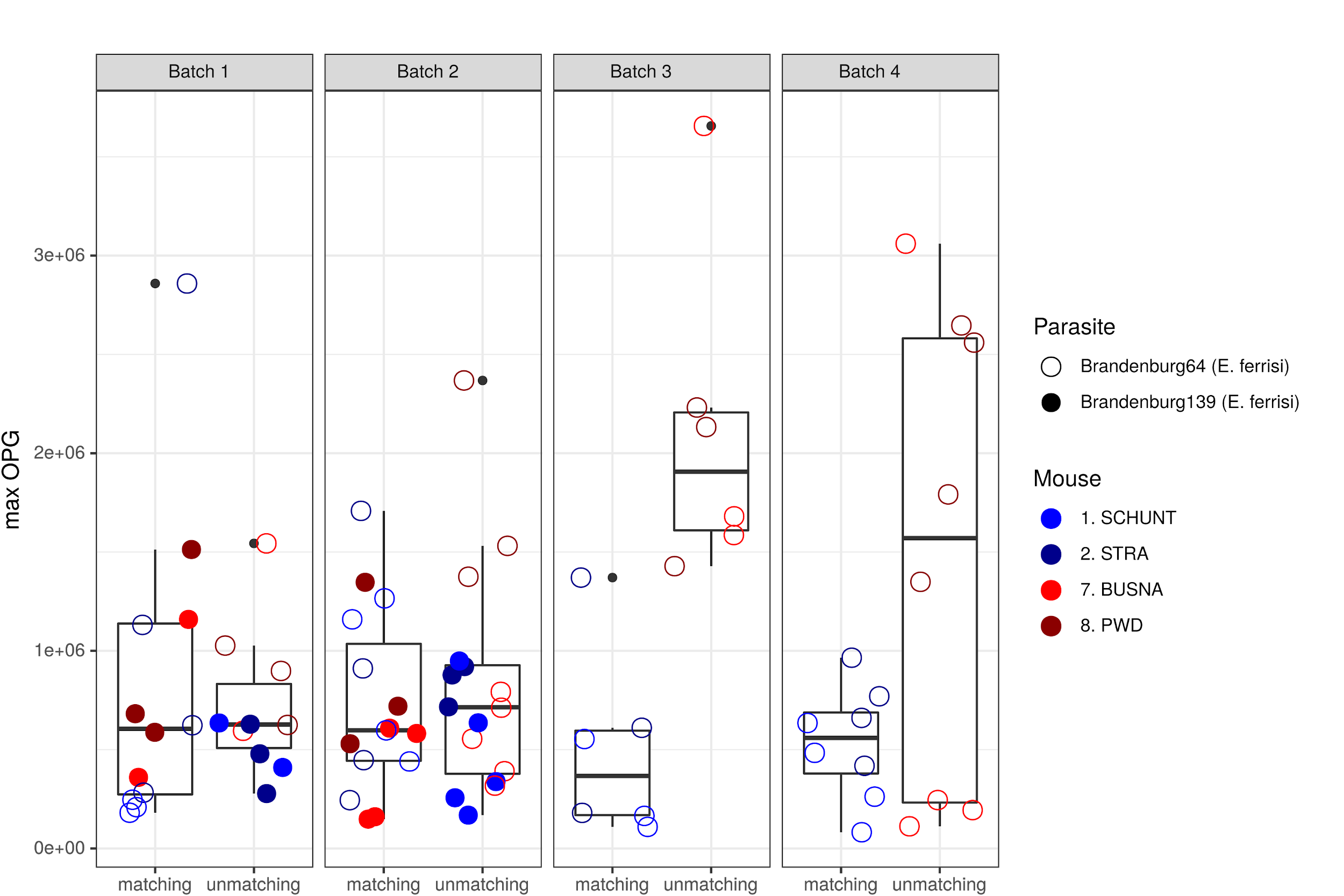
C7. Lines 258-262: why not log OPG and use pearsons?

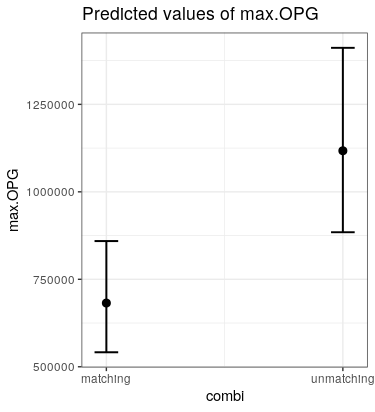
As a non-parametric test, Spearman’s rank correlation test is less sensitive to outliers than Pearson’s test, and also more stringent which is important for us to avoid false positive correlations. As asked by the Associate Editor, we tested also this correlation removing the point of mouse strain 8 (PWD) (see comment C1)

C8. Lines 294-296: what was the main effect of parasite isolate?

We chose to present only significant effects. We understand that this is confusing, and now mention the non-significant terms (main effects) throughout the results section.

C9. Lines 302-305: I’m not sure you’ve specifically tested for this? To test whether strains are better (or even worse) adapted to their own parasite (the parasite might be winning this co-evolutionary battle) you would need to pool the domesticus and musculus and ask whether ‘matched’ i.e. eastern w/ eastern and western w/ western was different ‘unmatched’ i.e. EW and WE. I don’t think they will be, looking at the data, but you still need to specifically test for it.

This is a very good comment, and we thank the reviewer for suggesting this test. Following this advice, we performed a linear regression (with a negative binomial distribution) of maximum OPG as a response of the variable taking values “matching” or “unmatching”. We found a statistically significant effect of this variable, entirely unexpected for ourselves. Nevertheless, plotting the data in the boxplots below, it appears that this could be mainly driven by the batch 3, in which only the Eastern parasite isolate was used. The number of mice and the unbalanced design may be insufficient to conclude that this effect is indeed significant. We, now, however, also lack confidence to trust the negative results of our previous tests. For this reason, as pointed out in comment C2, we chose to focus on the main question (coupling in *E. ferrisi* and *E. falciformis*) and remove our analysis of host adaptation as statistical power and design can neither exclude false negative nor false positives stringently enough.



C10. Lines 334-336: this is confusing, you found a positive correlation between OPG and tolerance, which indicates a negative correlation between tolerance and resistance. However, again, this fit is terrible, try log OPG. Figure 3: What do the error bars represent? Is this raw data or predictions from the model? If the latter, for A and B, given no significant interaction, why plot the parasite strains separately? Why is OPG plotted on a log scale for A but not for C?

It appears that there may be a confusion: we did not fit a statistical model at this point, but used a non parametric correlation (Spearman) that would not be affected by a change of scaling. The line is not the result of a fit but was shown for illustration only. This comment is not relevant anymore in the current manuscript, as figure 3 was removed (see comment C9). Finally, we homogenized all figures and removed the log scale that was confusing.

C11. Figure 4: As for 3, why is OPG plotted on a log scale for A but not for C,D or E?

As mentioned in comment C5, log scale was only for visualisation, we now plot everything on the untransformed scale.

Reviewer: 2

Comments to the Author

This article reports the results of careful experiments on infection resistance and tolerance in pure and crossed wild-derived mouse strains and wild-derived Eimeria isolates. The experimental subjects are well linked to a natural context in terms of the Mus musculus x Mus mus hybrid zone in central Europe. The rationale for the study is clearly set out and compelling, and the experiments are clearly described and analyzed. The authors find a certain pattern of significant results and interpret these in a reasonable and interesting way, leading to the introduction of an interesting idea about whether correlation between tolerance and resistance affects the tendency of host-parasite systems to undergo antagonistic coevolution. Broadly I found the article clear and I feel usefully enlightened and glad that I have read it – I am sure many other readers will as well.

The data are what they are, and there may be some limitations in terms of the inference drawn from the pattern of results. The negative correlation between resistance and tolerance proxies in *Eimeria falciformis* infection is convincing up to a point, but there is also the slight concern about autocorrelation (although, to be fair, the authors are very explicit about this). C12. Also, there is a lack of a direct statistical test for a difference in the relationship between tolerance and resistance in the two *Eimeria* species infections. The authors depend on qualitatively comparing coupling or uncoupling based on the pattern of test results within the different compartments of the study – and this could be affected by statistical power issues within compartments.

We thank the reviewer for pointing out this, and added such a direct test. We did find a significant difference between both parasites for the relationship between relative weight loss and OPG, and this was not significant for the relationship between tolerance and OPG, likely due to the fact that this is based on eight points per parasite (the replication level being the mouse group). We added text in Material and Methods and Results as follows:

lines 259-269: “After testing the resistance-tolerance coupling separately in both parasites, we tested the statistical difference in the relationship between (1) health-effect and resistance and (2) tolerance and resistance in the two Eimeria species infections. To this aim, we used the mean values predicted by our three models (see 4.2) for each eight mouse groups to perform first a linear regression of the mean predicted relative weight loss as a response of the mean predicted OPG, parasite isolate and their interaction, and second a linear regression of the mean predicted tolerance value as a response of the mean predicted OPG, parasite isolate and their interaction. The significance of the marginal contribution of each parameter to the full model was assessed by removing each parameter from the full model, and the difference between full and reduced model was assessed using likelihood ratio tests (G).”

lines 315-329: “When performing a linear regression of the mean predicted relative weight loss for each eight mouse groups as a response of the mean predicted OPG, parasite isolate and their interaction, we found that the mean number of OPG varies with the relative weight loss (LRT: G=10, df=2, P<0.01), differs between both parasites (LRT: G=8.9, df=2, P=0.012), and more importantly we found a significant interaction term (LRT: G=8.3, df=1, P<0.01). This means that the relationship between mean health-effect and mean resistance differs between the two *Eimeria* species infections. Then, we performed a linear regression of the mean predicted tolerance for each eight mouse groups as a response of the mean predicted OPG, parasite isolate and their interaction. In this case we found that the mean number of OPG varies along with tolerance (LRT: G=8.5, df=2, P=0.01) but does not statistically differ between both parasites (LRT: G=1.1, df=2, P=0.57), and the interaction term was not found significant (LRT: G=0.03, df=1, P=0.86). In this respect, the correlation between resistance and tolerance was not found to significantly differ between both parasites. Following these results, we looked at the coupling of resistance and tolerance within each of the two isolates“

Moreover, we modified the discussion and the title to avoid misleading the reader by claiming that there would be a statistically significant difference in these relationships. We nevertheless argue that our approach on coupling allows pointing out possible differences between these two parasites.

C13. Some of the co-evolutionary discussions assume that the resistance and tolerance phenotypes are underpinned by genetic variation. But what if environment or genotype x environment effects – which could be confounded with strain or isolate - are important? Furthermore, it is likely that laboratory animals will often adopt phenotypes (including immunological and infection phenotypes) very unlike those in the wild. If they have been bred or passaged in the laboratory they may also have undergone unrepresentative genetic changes. For this reason, more details should be given on the husbandry / passage history of the host strains and parasite isolates, although, in general, the form of the study has been very clearly represented.

We added the following husbandry informations following this comment:

“Previous to the experiment, the isolates had been passaged respectively 3 and 4 times in NMRI laboratory mice.” lines 114-115

“These four strains were obtained from xxx generations of brother–sister mating” lines 133-134. Ask Jarda for the exact number of generations of each mouse strain.

Notwithstanding, I think that, overall, this article is detailed and relevant work that will be interesting for many readers.

Specific points:

C14. 71 “Eventually, ….” This doesn’t seem to be the right word, implying a timeline – but the citation is an early one.

We removed ”eventually”

C15. 50-52. But arguably the determinants go beyond metabolic costs and involve the disruption of function and ultimately fitness. These arguments, and those above, don’t seem to quite paint a full picture.

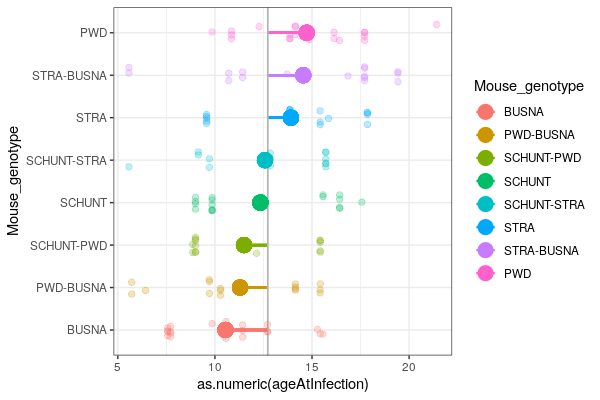
We modified as follows:

“From a mechanistic perspective tolerance alleviates direct or indirect damage (e.g. excessive immune response underlying resistance against parasites, called immunopathology; Graham et al. 2005) caused by parasites (Råberg et al. 2009). Tolerance mechanisms include modulation of inflammatory response (Ayres & Schneider 2012), tissue repair (stress response, damage repair and cellular regeneration mechanisms; Soares et al. 2017), and compensation of parasite-induced damage by increase of reproductive effort (Baucom & Roode 2011). Even in the absence of parasite infection, the maintenance of tolerance mechanisms can be detrimental to other functions, ultimately affecting host fitness (Stove et al. 2000; Råberg et al. 2009). The resulting ~~metabolic~~costs of resistance and tolerance determine the optimal (steady state and infection inducible) extent and of both immune defences (Sheldon & Verhulst 1996).” lines 43-54

C16. 138 This is quite high variation in mouse age, given the life-span of the lab mouse.

This is indeed rather high, but mainly driven by few individuals, and more importantly not clustered by mouse strain: indeed, for all mouse strains, the means are very close. Here is a plot of the age for each mouse strain, and the distance of the group mean to the overall mean (grey line). For the sake of brevity we decided not to include the plot in the article, but we added this information lines 136-138:

“Age of the mice at the time of infection ranged between 5.6 and 21.4 weeks, with the mean for each eight mouse strains ranging between 10.5 and 14.7 weeks.”



C17. 169. If these were pinworms (can anything be said about their identity, beyond just “nematodes”?), then I think that the authors are fully justified in just going ahead and assessing the Eimeria results regardless. Other types of nematode might be more of an issue, but pinworms are what is typically found in many laboratory colonies.

We identified classically found *Syphacia* sp. and *Aspiculuris* sp. eggs and added the information line 170: “(*Syphacia* sp. and *Aspiculuris* sp.)”

C18. 195-199 Although generally the acid test for what distribution should be used for model analyses rests on the assessment of model residuals, rather than lumped raw data.

We agree, checked this, and add now:

“We confirmed the fit of our models by assessing the uniformity of the distribution of model residuals” line 198-199

C19. 217 “either”?

Thanks for spotting this typo, we removed it.

C20. 291-222. Consider refining the wording of this sentence.

We replaced the previous sentence “For tolerance, we performed a linear regression with null intercept (as each mouse was controlled against itself at start of the experiment, before losing weight or shedding parasite), modelling relative weight loss as a response of maximum OPG interacting with mouse group, parasite isolate and their interaction.” by

“Tolerance was assessed by modelling relative weight loss as a response of maximum OPG interacting with mouse group, parasite isolate and the interaction of the two latter. As each mouse was controlled against itself at the start of the experiment, before losing weight or shedding parasites, we performed a linear regression with null intercept.” lines 222-226

C21. 226. models

We corrected, thanks

C22. 272 with the

We corrected, thanks

C23. 298. “Eventually” doesn’t seem to be the right word.

This entire section was removed (see comment C9).

C24. 298-302. This is getting a bit confusing or is a lapse of explanation. The wording of this sentence implies a linear regression between mouse strain and parasite isolate (two factors), but this doesn’t seem make sense, or at least the meaning is not fully clear. Suggest re-wording.

We changed and rephrased, see comment of reviewer 1 (C9)

C25. 318 Eventually?

We replaced by « moreover »

C26. 363 literature

We added “(Fineblum and Rausher 1995)” line 384

C27. 365 tolerance mathematically

We corrected, thanks

C28. 366 meaningful

We corrected, thanks

C29. 367. Why stepwise? This implies one comes before the other – but this isn’t necessarily so.

We corrected for “double” line 387

C30. 368-370. Check spellings and wording of this sentence.

We modified as follows: “To limit the possibility of statistical artifact, our approach did not only consist in calculating correlations between resistance and tolerance, but also in testing differences in resistance, impact on health and tolerance.” line 388-391

C31. 372-375. This needs to be explained better.

We modified the original sentence “*We additionally excluded the possibility of positive correlation between mean health-effect and mean resistance of each host strains, which could indicate some host strains having few parasites-few effects on health, and others more parasites-more effects on health: this configuration would limit the possibility of detecting an actual resistance-tolerance trade-off.*” as follows:

“Of note, a positive correlation between mean health-effect and mean resistance of each host strains could indicate some host strains having few parasites-few effects on health, and others more parasites-more effects on health; This configuration would limit the possibility of detecting an actual resistance-tolerance trade-off by lack of a full range of resistance values. For this reason, our approach consisted in testing the "coupling" between resistance and tolerance, that is (1) a genotype-level correlation between tolerance and resistance additionally supported by (2) the absence of positive correlation between health-effect and resistance. We argue that this additional step increases the confidence in the presence of a biologically meaningful negative correlation between resistance and tolerance, likely implying a trade-off.” line 391-401

C32. Fig. 2 legend. Host relative weight? What hosts are these data for?

These data are for all host groups pooled. We added “host relative weight” and precision : “All mouse groups are pooled for each parasite isolate.”

C33. Fig. S2.3 legend line 46 …in E. falciformis?

We corrected, thanks (in what is now Appendix 3)