nature portfolio

Peer Review File

Gut microbiome community structure correlates with different behavioral phenotypes in the Belyaev farm-fox experiment

Corresponding Author: Dr Lara C Puetz

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The authors here use a nice and well-studied dataset of wild and domesticated foxes to investigate the potential role of the microbiome in facilitating the domestication process. I think this is a very interesting study, and overall shows compelling results and was handled carefully. I think it will make a nice contribution to the literature particularly given the lack of knowledge on the relationship between domestication and the microbiome in non-model systems. I detail some minor comments below regarding the manuscript and discussion of the results.

Minor comments:
• Batch effects (Lns 102-105) – I do think it is best not to lump data from separate batches if they were stored differently, however it would be useful to report whether the same effects were seen within the 2015 dataset as a replicate. Additionally, if the 2015 data was sufficiently different enough to warrant removal, should it have been used to generated the MAG reference catalogue? Some further clarification and justification regarding this batch is needed.
• I think the manuscript does a good job of not over-interpreting, and framing some connections as interesting without over-reaching. However, I would caution throughout interpretation at the phylum level (e.g. lines 124-135 wrt Tenericutes), only because there is such variation within phyla generally in direction and magnitude of effect across genera / families/ asv so worth making that caveat clear.
• Fig 1 – If you're dropping 2015 I think best to leave out of analyses here as well (or have a description of 2015 results in supplementary material). Additionally, in panel A – those differences do not visually appear to be significant at the $p < 0.0001$ level. Investigating the package and method used for testing these differences, the package information "DivNet takes care of all of these issues, and gives you a confidence interval that incorporates all of these challenges" suggests to me that these models incorporate additional sources of error / variation, and that this would therefore make the CI larger and not smaller. However the CI in 2A are extremely small for the Ns in each group. How do these results compare to standard testing of Shannon diversity between the groups?
• Lns 161-163 – the use of "they" in the following sentence --"In this regard, when we explored our data at the ASVs level, we see they are also depleted in the tame strain" – seems very vague to me. Does this refer to ASVs specifically within the Bacteroides genus? If so, is there a significant depletion when the data is analysed at the genus level? There is just something about the discordant taxa level tested and dicussed that seems odd here but could be addressed with slight wording clarification.
• Lns 359-363 – The relationships between the microbiome and immune system can be bidirectional which may be worth acknowledging here.
• Lines 498-509 - The methods here seem to suggest that the model was run on both batches of data controlling for collection year (?), but given the rest of the analyses were run separately due to the significant differences it would make sense to me that the DivNet is run on 2017 only.

Reviewer #3

(Remarks to the Author)
Synopsis
In this study, the authors investigated whether, and if so how, the gut microbiota of two selection lines of the Belyayev silver

fox domestication experiment differed. The goal was to establish whether selection for tameness (perhaps the most important trait for domestication) is discernible in the gut microbiota; the hypothesis being early selection for tameness might select for specific microbes that are have the potential to affect host behaviour.

First, the authors 16S-sequenced the gut microbiota of about 50 foxes from each selection line, and sequentially model each encountered taxa to see if they differentially show up in each selection line. Broadly, the relative abundance of some taxa differ between the aggressive and tame fox selection lines (Fig 1 & Fig 2). The authors then performed shotgun metagenomics to attempt to reconstruct the genomes of the bacteria found in these foxes. This was done with the goal of testing whether the gut microbiota of foxes from each selection line differed in their presence/absence and abundance of Gut-Brain Modules (GBMs) (i.e. whether the reconstructed microbial genomes have the capacity to perform metabolic functions that may affect the behaviour of the foxes). Some GBMs are indeed present in different proportions in foxes of each line (Fig 3 & Fig 4), suggesting that the gut microbiota of each selection line can potentially (and partly) explain the differences in behaviour found between tame and aggressive foxes.

This work is very interesting to the field and adds to the increasing evidence that the gut microbiota is important in shaping animal behaviour. The text is overall well-written, but does lack clarity in some sections (examples below). The statistical analyses are for the most part appropriate, but statistical reporting is lacking in some sections.

Below you can find some specific major and minor comments, that in my opinion should be addressed to improve this manuscript.

Major(-ish) concerns

While this paper is overall quite good, it does suffer from a few issues, but these are mainly contextual:

1 – While, at times, the authors are rightfully aware of the limitations of this study (mainly that it is purely correlational); at other times, thoughts that cannot be corroborated by the analyses are presented almost as fact; and there are very few (if any) instances where alternative hypothesis are considered and discussed. For instance, the text is not clear on weather both lines come from the U.S.S.R, Canada, or each line from a different place? (which would suggest founder-effects could be at play).

2 – While it is convincingly shown that the tame and aggressive groups of foxes differ in relative abundances of some microbes and that there are differences in the functional potential these microbiomes might have on host behaviour, we are never actually shown how these foxes differ in: a) broad-brush microbiome (i.e. beta diversity between selection lines); and b) in behaviour. Where these foxes also behaviourally assayed?

Minor suggestions

L29 – The first sentence of the abstract reads quite opaque. “Biological shifts” in what sense?

L45-47 – This sentence reads a bit redundant: it seems the authors are stating that “understanding the evolutionary mechanisms … of domestication” allows us to understand how animals “adapt” (i.e. evolve), to “new social environments” (i.e. human-dominated; domestic environments). Boils down to “understanding the evolution of domestication allows us to understand the evolution of domestication”. I’d recommend the authors remove this sentence entirely as the ms would be to a much clearer, stronger, start if it were to start from the second sentence.

L48 – Strictly speaking “selection for tameness” is not a “trait” rather a shift in a trait/phenotype.

L59 – “gut-derived” needs a hyphen.

L62 – “initiation of immune (…)” word missing?

L75 – “reduced fear” doesn’t need a hyphen.

L80 – While I understand that the authors mean that “the two previous points lead us to hypothesize that…”; “taken together” is not a clear segway for the sentence that follows.

L94-96 – While it is mentioned further down the ms that each sample corresponds to one individual, for clarity, I’d recommend it be mentioned somewhere in the paragraph too.

Line 108 – Do you mean the samples collected in 2017? If so, not clear.

Line 112 – Effect size, statistical test and degrees of freedom not reported; here in particular but also in other places of the ms.

Line 112 – Would also be interesting to see the Observed richness, maybe as a supplement.

Line 119 – There are other possible reasons for these differences in microbial richness between tame and aggressive foxes, e.g. founder effects. Moreover, it is unclear the mechanism though which selection for tameness/domestication would

deplete certain gut microbes?

Line 134 – I think the authors mean something along the lines of “Taken together, these findings lead us to hypothesize that…”

Line 310 – I think the authors mean “are different” rather than “differentiate” (i.e. to make different).

Line 310 – I’m not sure I follow how the gut microbiota could influence the “genetic composition” of the brain of the hosts.

Figure 1A – If the data from 2015 was excluded from most analysis, why show it here?

Figure 2A – What exactly to the dots and whiskers represent? Mean plus/minus CI?

Figure 2B – I’m not sure I follow the logic of plotting the total (I’m assuming across samples?) absolute abundance vs prevalence – it is the field’s best practice to treat all microbiome datasets as compositional (i.e. relative abundances). It is also not clear how the discriminant ASVs are represented in the plot.

Figure 3 – The legend for A reads: “Six MAGs were differentially abundant or exclusive to a behavioral phenotypic group and highlighted on the tree based on the enrichment in either the aggressive (yellow) or the tame (teal) populations”; is this what the outermost layer of the tree represents? If yes, there are way more than just six MAGs annotated in such a way. If not, what is this layer representing?

Figure 3 - The legend for B reads: “Detection of GBMs (minimum 66% coverage) in the 22 differentially occurring fox MAGs highlighted in A)” (full stop missing). There seem to be way more than 22 highlighted GBMs in panel A?

Figure 4 – This might be a submission issue, and the authors completely blameless; but this figure is barely legible. I think I can see some sort of phylogenetic tree at the top of panel A? And cant at all read any of panel A’s left labels.

Methods or Supps: no mention of how the phylogenetic tree was generated.

Methods: are male and female foxes equally present and equally distributed in the dataset?

Line 418: how many sequencing runs were the samples divided into? If only one sequencing run was performed, then all good, but if more than one, dada2 has to be performed on each run separately.

Methods: did the authors look at the rarefaction curves of the 16S data? How do they inform the sampling completeness?

Reviewer #4

(Remarks to the Author)
This paper uses 16S rRNA and metagenomic profiling to test whether artificially selected tame and aggressive fox lines differ at the level of the gut microbiome. Because gut microbiomes could change as a consequence of domestication, or could be mechanistically involved in the domestication process, the study has an interesting premise. A major strength is the use of the Belyaev fox experiments, which have already revealed a remarkable syndrome of responses to strong artificial selection on behavior. The system is a textbook example of change during domestication, and therefore an excellent choice for the present work.

I have no real concerns about the main data analyses, which are basically straightforward tests for differential alpha diversity, microbial abundance, or genic abundance. I agree that the results establish clearly differentiated gut microbiomes in the two selected lines.

My primary concern is that there is a great deal of speculation woven into the presentation of results (also in the abstract, intro, and conclusions—but I think it’s most likely to mislead in the results). The authors clearly favor the idea that the gut microbiome is causal to the domestication phenotype and do a lot of selective citing of the literature to make this argument. However, there is no experimental evidence showing that manipulation of the microbiome per se (and particularly the microbes identified as differentially abundant in this study) affects other behavioral or domestication-related phenotypes, only plausibility arguments. There is also no serious discussion of alternative interpretations of their findings. For example, domesticated animals often experience changes in gut length, which could also affect the microbiome, making the differences reported here a by-product, not a cause, of the domestication syndrome.

Consequently, hypotheses that are testable in this study and find some support (e.g., lines 80-81: “we hypothesize that selection for behavioral traits in the fox domestication experiment has…led to selection on…[the] gut microbiota”) are combined with hypotheses that are not tested at all (line 81-82: “…contribut[ing] to shaping fox behaviors.”). Straightforward results presented in one sentence and based on empirical data are then combined with long discussions of “possibilities” that support the authors’ favored interpretation (this is an issue throughout the entire Results/Discussion section).

Without much more targeted experimental work and functional characterization, my view is the conclusions suggested in the study are therefore premature--even though the evidence that the foxes’ gut microbiomes differ in a number of ways (the

meat of the actual data generation and analysis) is strong.

Minor comments:
-Differential abundance analyses are reported with “a controlled false discovery rate” but it is not clear what FDR correction approach is used throughout and what the FDR threshold for significance is. This makes it hard to interpret statements about significance, and should be clarified throughout.
-In Figure 3, the legend refers to six MAGs that were differentially abundant between fox groups, which are shown, I think, in the outermost ring of Figure3A. However, there are far more than six color blocks on the outer ring, so I must be missing something—please clarify.
-The conclusions in the GBM section would be moderately strengthened if, instead of discussing case example GBMs that were differentially abundant between fox strains, there was clear evidence that neuroactive GBMs overall were more likely to be differentially abundant than expected by chance (a simple variant of a gene set enrichment approach). It’s not clear that this is the case.

Version 1:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The authors have responded to all of my comments sufficiently, and have carried out thoughtful revisions to all reviewers' comments.

Reviewer #3

(Remarks to the Author)
The authors have appropriately addressed my main concerns in the text, or have convincingly replied to them in the rebuttal letter. I support the publication of this manuscript, I just found two minor points that should be addressed before that happens:

L84 – provide reference for your red jungle fowl experiment.

L418 – “descent” instead of “decent”

Reviewer #4

(Remarks to the Author)
This is the second time I’ve seen this manuscript. I appreciate the additions the authors have made in the revision process. However, while I still find the basic analysis sound, I continue to view the interpretation and discussion as speculative. The authors mention several hypotheses that could account for differential abundance of microbial taxa and/or “gene modules” in their data set (roughly lines 118 to 130), but typically limit their discussion of individual results to a causal role of the gut microbiome in modulating behavior/brain function.

The value of the paper to readers would be further increased if the alternative hypotheses could be linked to specific, stated, differentiable predictions, but ultimately this is a decision in the hands of the authors and the editor. As a small related point, it would be helpful to differentiate when alternative hypotheses have been tested and found no support versus when they haven’t been tested (e.g., lines 125 – 126: “such morphological changes have not been identified in the experimentally domesticated foxes”—does that mean this possibility hasn’t been examined, or that it has, and is not supported?).

Title: Gut microbiome community structure correlates with different behavioral phenotypes in the Belyaev farm-fox experiment - COMMSBIO-24-3924

Referee expertise:

Referee #2: Ecology, Host-Microbe Interactions

Referee #3: Animal Microbiota and Behaviour

Referee #4: Behaviour and Evolution

Reviewers' comments:

Reviewer #2 (Remarks to the Author):

The authors here use a nice and well-studied dataset of wild and domesticated foxes to investigate the potential role of the microbiome in facilitating the domestication process. I think this is a very interesting study, and overall shows compelling results and was handled carefully. I think it will make a nice contribution to the literature particularly given the lack of knowledge on the relationship between domestication and the microbiome in non-model systems. I detail some minor comments below regarding the manuscript and discussion of the results.

We thank the reviewer for their kind words of support and nice summary of the work.

Minor comments:

1.1. Batch effects (Lns 102-105) – I do think it is best not to lump data from separate batches if they were stored differently, however it would be useful to report whether the same effects were seen within the 2015 dataset as a replicate. Additionally, if the 2015 data was sufficiently different enough to warrant removal, should it have been used to generated the MAG reference catalogue? Some further clarification and justification regarding this batch is needed.

The 2015 dataset was mainly used as a pilot study for proof of concept and to direct future experimental design. One outcome from this initial study indicated that we needed to increase the number of samples by 5X for adequate statistical power, as indicated by a power analysis on the 2015 - 16S sequence data. Because we only sampled 10 individuals per selection line (aggressive = 10 and tame = 10) in 2015, this dataset alone was under-sampled for robust statistical analyses. Originally, we had thought to combine the 2015 and 2017 dataset in our statistical analyses and control for

collection year on dispersion in our statistical model. However, fresh frozen (-20°C) samples, in the absence of a preservation buffer, can dramatically alter the abundance (although not necessarily the presence/absence) of the microbiota present within a given sample (Song *et al*., 2016; Li *et al*., 2023), as was indicated in our results for the 2015 fecal collections (see L105-108). We therefore opted to remove the 2015 data set altogether from our statistical analyses for the above two technical reasons.

However, we feel it is still valid to include the 2015 shotgun metagenomic data to generate the MAG reference catalogue. Typically sequencing more individuals leads to better recovery of MAGs (e.g. increased capture of microbial diversity, improved binning and assembly, better recovery of rare and low-abundant taxa). Therefore, with the main goal of curating a collection of host-associated microbial genomes present within the gut of farmed foxes; by including the 2015 sequence data, we likely generated a better representation of the overall gut microbiota present in the farmed foxes.

1.2. I think the manuscript does a good job of not over-interpreting, and framing some connections as interesting without over-reaching. However, I would caution throughout interpretation at the phylum level (e.g. lines 124-135 wrt Tenericutes), only because there is such variation within phyla generally in direction and magnitude of effect across genera / families/ asv so worth making that caveat clear.

We thank the reviewer for the suggestion and have reworded the text to point out that we are discussing some of our differences at broad taxonomic levels. (See lines L141-144). It is nonetheless intriguing that certain taxa within the Tenericutes phulym, such as *Anaeroplasmataceae,* similarly positively correlated with aggression in rodents. And further, that, of the very few gut bacteria identified to date that appear to show heritability, several Tenericutes are among these heritable bacteria.

1.3. Fig 1 – If you're dropping 2015 I think best to leave out of analyses here as well (or have a description of 2015 results in supplementary material).

We thank our reviewers for this suggestion and have now removed the 2015 results from Fig 1A.

Additionally, in panel A – those differences do not visually appear to be significant at the $p < 0.0001$ level. Investigating the package and method used for testing these differences, the package information "DivNet takes care of all of these issues, and gives you a confidence interval that incorporates all of these challenges" suggests to me that these models incorporate additional sources of error / variation, and that this would therefore make the CI larger and not smaller. However the CI in 2A are extremely small for the Ns in each group. How do these results compare to standard testing of Shannon diversity between the groups?

One of the strengths of DivNet is that we can incorporate information about the experimental design as a covariate (e.g. behavior in our case) in the model for estimating diversity. In other words, DivNet uses covariate information to share strength

across samples and obtain a mean Shannon diversity estimate of the group of samples/ecological community of interest. Detailed comparisons of DivNet with other popular methods of estimating diversity can be found in the DivNet publication (Willis, A. D. & Martin, B. D. Estimating diversity in networked ecological communities. Biostatistics (2020) doi:10.1093/biostatistics/kxaa015.)

1.4. Lns 161-163 – the use of “they” in the following sentence --“In this regard, when we explored our data at the ASVs level, we see they are also depleted in the tame strain” – seems very vague to me. Does this refer to ASVs specifically within the Bacteroides genus? If so, is there a significant depletion when the data is analysed at the genus level? There is just something about the discordant taxa level tested and dicussed that seems odd here but could be addressed with slight wording clarification.

We have now adjusted the text to clarify at which taxonomic level the results were being assessed at (L172-173; L178). In the specific sentence above, we do mean that the tame population were depleted of many ASVs within the *Bacteroides* genus (Figure 2 A), despite *Bacteroides* not being significantly differentially abundant at the genus level. Thanks for pointing this out.

1.5. Lns 359-363 – The relationships between the microbiome and immune system can be bidirectional which may be worth acknowledging here.

Thank you for your suggestion. We have now included the bi-directional relationship in our conclusion (L383-386).

1.6. Lines 498-509 - The methods here seem to suggest that the model was run on both batches of data controlling for collection year (?), but given the rest of the analyses were run separately due to the significant differences it would make sense to me that the DivNet is run on 2017 only.

We apologize for the confusion. We originally combined the 2015 and 2017 dataset in our statistical analyses and controlled for collection year on dispersion in our statistical model. However, due to the strong batch effects described above, we removed the 2015 data set from our statistical analyses and did not amend the text. It has now been corrected. (L529-531). Thank you for catching this.

Reviewer #3 (Remarks to the Author):

Synopsis
In this study, the authors investigated whether, and if so how, the gut microbiota of two selection lines of the Belyayev silver fox domestication experiment differed. The goal was to establish whether selection for tameness (perhaps the most important trait for

domestication) is discernible in the gut microbiota; the hypothesis being early selection for tameness might select for specific microbes that are have the potential to affect host behaviour.

First, the authors 16S-sequenced the gut microbiota of about 50 foxes from each selection line, and sequentially model each encountered taxa to see if they differentially show up in each selection line. Broadly, the relative abundance of some taxa differ between the aggressive and tame fox selection lines (Fig 1 & Fig 2). The authors then performed shotgun metagenomics to attempt to reconstruct the genomes of the bacteria found in these foxes. This was done with the goal of testing whether the gut microbiota of foxes from each selection line differed in their presence/absence and abundance of Gut-Brain Modules (GBMs) (i.e. whether the reconstructed microbial genomes have the capacity to perform metabolic functions that may affect the behaviour of the foxes). Some GBMs are indeed present in different proportions in foxes of each line (Fig 3 & Fig 4), suggesting that the gut microbiota of each selection line can potentially (and partly) explain the differences in behaviour found between tame and aggressive foxes.

This work is very interesting to the field and adds to the increasing evidence that the gut microbiota is important in shaping animal behaviour. The text is overall well-written, but does lack clarity in some sections (examples below). The statistical analyses are for the most part appropriate, but statistical reporting is lacking in some sections.

We thank the reviewer for their kind words of support and nice summary of the work.

Below you can find some specific major and minor comments, that in my opinion should be addressed to improve this manuscript.

Major(-ish) concerns

While this paper is overall quite good, it does suffer from a few issues, but these are mainly contextual:

2.1 – While, at times, the authors are rightfully aware of the limitations of this study (mainly that it is purely correlational); at other times, thoughts that cannot be corroborated by the analyses are presented almost as fact; and there are very few (if any) instances where alternative hypothesis are considered and discussed. For instance, the text is not clear on weather both lines come from the U.S.S.R, Canada, or each line from a different place? (which would suggest founder-effects could be at play).

Thank you for pointing this out. We have now included alternative hypotheses in the manuscript, including more emphasis on how host factors shape the GM. (e.g. L120-126; L354-359; L370-393)

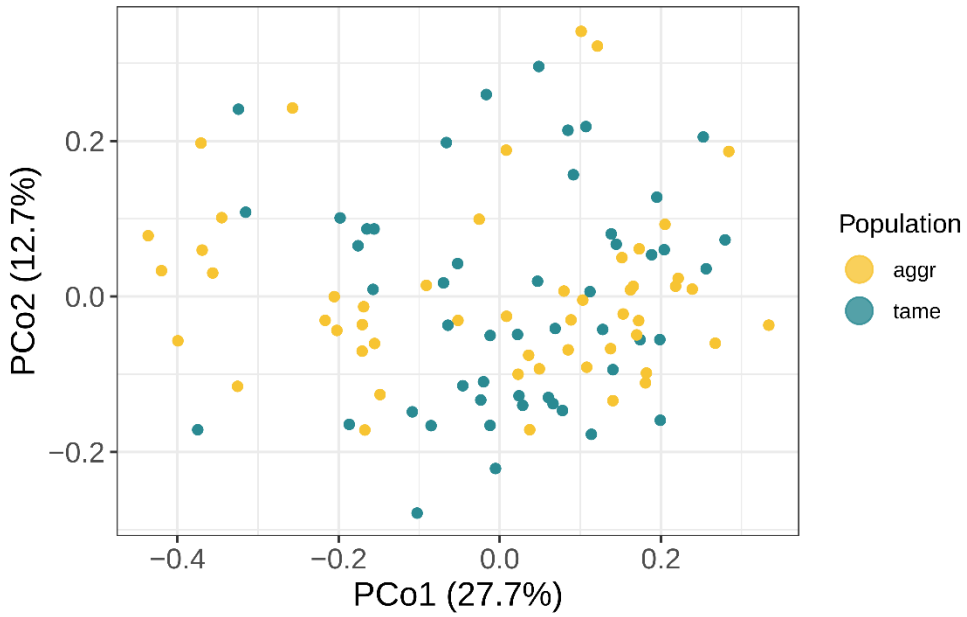
In regards to source of the experimental foxes, the tame and aggressive fox populations originated from a conventional farm-bred fox population of Canadian descent. The first successful captive-bred colony of red foxes was established in Canada in 1896. The

breeding stock started to be exported to Eurasia (including Russia) the 1910s-20s. Studies of farm-bred foxes in Russia using mtDNA and Y-chromosome markers found that all of the Russian experimental farm-bred foxes (including the tame and aggressive foxes) are descendants of Canadian foxes. The summary of these findings were recently published in the Journal of Heredity (Rando et al., 2024, https://doi.org/10.1093/jhered/esae022). We have now altered the text in the methods section of the manuscript to clarify (L417-419)

Selection of conventional farm-bred fox for tame behavior began in 1959 and for aggressive behavior in the late 1960s. Extensive genetic studies showed that the two populations are closely related (Kukekova et al, 2018, https://doi.org/10.1038/s41559-018-0611-6). Importantly, inbreeding was strenuously avoided in the selected populations since the beginning of the breeding program which also helped to avoid any bottlenecks during selection. Therefor founder-effects are likely not at play here.

2.2 – While it is convincingly shown that the tame and aggressive groups of foxes differ in relative abundances of some microbes and that there are differences in the functional potential these microbiomes might have on host behaviour, we are never actually shown how these foxes differ in: a) broad-brush microbiome (i.e. beta diversity between selection lines); and b) in behaviour. Where these foxes also behaviourally assayed?

A) The overall microbial community structure overlaps between the tame and aggressive populations of foxes as seen in the principle coordinate analysis (PCoA) plot of the Bray-Curtis dissimilarity matrices of the relative abundance of bacterial 16S ASV sequence data in the 2017 data set below. This was not overly surprising because, a) as described in the comment (2.1) above, the tame and aggressive fox populations are closely related and b) they were housed in the same or similar environments, raised in standard conditions and underwent identical treatment with minimally necessary human interaction until behavioral testing was performed. In our study, we focused on the more subtle shifts in the microbiota, that are commonly not captured by beta diversity metrics alone, and to assess the potential biological implications of these differentially abundant taxa.



B) Fox behavior was systematically scored based on their response to humans using a categorical scoring system that was well established early in the breeding program (Overview can be found in Kukekova et al., 2014, doi:https://doi.org/10.1016/B978-0-12-394586-0.00010-X). We have now included this in the methods section (L410-412)

Minor suggestions

L29 – The first sentence of the abstract reads quite opaque. “Biological shifts” in what sense?
Here, we used "biological shifts" in the context of domestication to include the significant and wide-ranging changes in biological, ecological, and evolutionary systems (e.g. genetic, morphological and behavioral changes of domesticated animals, ecosystem level impacts of replacing natural habitats with human-modified ones, evolutionary interactions, etc).

L45-47 – This sentence reads a bit redundant: it seems the authors are stating that “understanding the evolutionary mechanisms … of domestication” allows us to understand how animals “adapt” (i.e. evolve), to “new social environments” (i.e. human-dominated; domestic environments). Boils down to “understanding the evolution of domestication allows us to understand the evolution of domestication”. I’d recommend the authors remove this sentence entirely as the ms would be to a much clearer, stronger, start if it were to start from the second sentence.

Thank you for the suggestion. We agree and have removed the first sentence of the introduction. (L45-47 from original text removed).

L48 – Strictly speaking “selection for tameness” is not a “trait” rather a shift in a trait/phenotype.

Good point. The text has now been altered slightly to clarify (L47-48).

L59 – “gut-derived” needs a hyphen.
Thank you for noticing. It is now edited in the text. (L59)

L62 – “initiation of immune (…)” word missing?

No, we do indeed want to say initiation of immune activation, where the GM can activate specific immune cells and/or pathways rather than the overall activation of the immune system as a whole.

L75 – “reduced fear” doesn’t need a hyphen.
Thank you for noticing. It is now edited in the text. (L75)

L80 – While I understand that the authors mean that “the two previous points lead us to hypothesize that…”; “taken together” is not a clear segway for the sentence that follows.

We have now reworked the hypothesis in the section of the manuscript to help with flow and clarity (L77-84)

L94-96 – While it is mentioned further down the ms that each sample corresponds to one individual, for clarity, I’d recommend it be mentioned somewhere in the paragraph too.

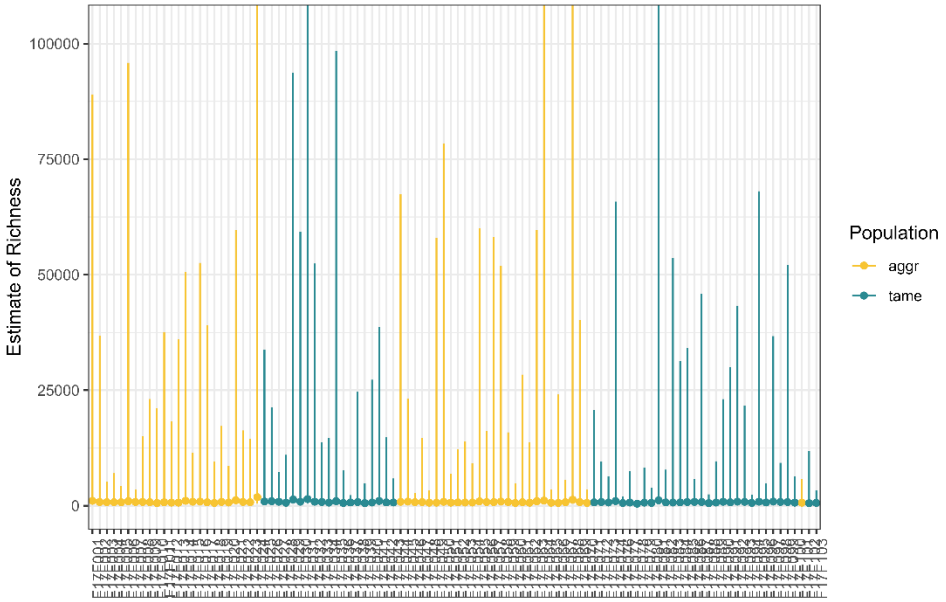
Thank you for pointing this out. We have now reworded to include that the samples were collected from 123 individual foxes (L95)

Line 108 – Do you mean the samples collected in 2017? If so, not clear.
We did indeed mean the fecal samples collected from the foxes in 2017. We have now reworded for clarity (L109-110)

Line 112 – Effect size, statistical test and degrees of freedom not reported; here in particular but also in other places of the ms.

Line 112 – Would also be interesting to see the Observed richness, maybe as a supplement.

We did indeed estimate richness and perform hypothesis testing using the breakaway package v.4.8.4. Please see the Figure below. The estimated richness at the ASV level was significantly higher in the aggressive compared to the tame population (aggr=723±14; tame=676±18, p-adj=0.01).



Line 119 – There are other possible reasons for these differences in microbial richness between tame and aggressive foxes, e.g. founder effects. Moreover, it is unclear the mechanism though which selection for tameness/domestication would deplete certain gut microbes?

We thank the reviewer for bringing this up and we now include alternative hypotheses, including possible founder effects, into how these factors may drive differences in the GM diversities of the farmed foxes (L120-126). We also discuss further on both points below.

1) Although the tame and aggressive selection lines were initiated at different times, they were drawn from the same population and are closely related (please see the comment 2.1.2 above). Further, inbreeding was strenuously avoided, therefore founder effects, are likely not at play in the farmed foxes. In an ideal scenario, we would have had access to, and thus been able to include, GM data from the original outbred and P0 parental generations for both fox populations, and further, followed and profiled the GM of individuals across generations

throughout the entire selection process. Had this been possible, we would have been able to get a better grasp of what the initial GM communities of each population looked like, when and how quickly compositional changes occur within the GM of animals during early domestication, and how these changes may reflect causal links to the behavioral shifts detected throughout the process. Unfortunately, at the time the experimental study began on the foxes, there was no plan to assess the GM community, and the idea of the microbiota-gut-brain-axis did not exist (or was just at its infancy), and the animals are long since gone.

2) In regard to the second point above, the exact mechanisms by which domestication may alter the gut microbiota remain unknown, however, the GM of domesticated animals consistently differs from that of their wild counterparts. Decreased GM diversity is often associated with domesticated animals (Metcalf et al., 2017; Prabhu et al., 2020; Fu et al., 2021; Bowerman et al., 2021), however this pattern is species dependent (Reese et al., 2021), and captivity does not necessarily induce directional changes in the GM diversity and composition in vertebrates (Alberdi et al., 2021). To date, most of these studies have identified that ecological shifts during domestication, specifically changes in diet, can have major impacts on reshaping the gut microbiome, in terms of community composition and functional potential (Reese et al., 2021; Rampelli et al., 2021; Metcalf et al., 2017; Alessandri et al., 2019). Further, domestic and captive animals tend to be enriched in human-associated gut microbes compared to their wild counterparts, suggesting that increased contact with humans or even their ecological convergence during domestication, contribute to the acquisition of certain human-associated taxa throughout the process 124,125. (Alberdi et al., 2021; Reese et al., 2021; Alessandri et al., 2019; Ellis et al., 2013).

   Unlike many of these studies however, the two populations of experimental foxes from our work were housed in similar controlled environments and fed the same diets at a given time point, with minimal human exposure until behavioral testing was performed. As such, environment and diet alone cannot explain how certain microbes were depleted in the tame population. Alternatively, the behavioral selection process itself may have inadvertently led to the selection of (or in this case against) specific bacterial gut microbiota community structures that may potentially contribute to shaping fox behaviors.

Line 134 – I think the authors mean something along the lines of “Taken together, these findings lead us to hypothesize that…”
This section of the text has now been reworded for clarity and to additionally incorporate the comments on alternative hypotheses from all reviewers. (L120-130).

Line 310 – I think the authors mean “are different” rather than “differentiate” (i.e. to make different).

Indeed we meant to say “are different” and have amended it in the text. (L325)

Line 310 – I’m not sure I follow how the gut microbiota could influence the “genetic composition” of the brain of the hosts.

Good point. We have now reworded the text and provide examples and references of how the GM can alter genetic signatures in the brain of their host (L327).

Figure 1A – If the data from 2015 was excluded from most analysis, why show it here?

We thank our reviewers for this suggestion and have now removed the 2015 results from Fig 1A.

Figure 2A – What exactly to the dots and whiskers represent? Mean plus/minus CI?

The dots represent the estimated model coefficients for abundance (effect sizes) and the errors bars (whiskers) represent the 95% confidence intervals (i.e. the uncertainty in the estimated coefficient. We have now included this in the figure legend as well. (L211-212)

Figure 2B – I’m not sure I follow the logic of plotting the total (I’m assuming across samples?) absolute abundance vs prevalence – it is the field’s best practice to treat all microbiome datasets as compositional (i.e. relative abundances). It is also not clear how the discriminant ASVs are represented in the plot.

The figure is mainly trying to show that many of the differentially abundant ASVs occurred in abundant and prevalent taxa identified in the gut microbiota of the farmed foxes. In the plot we grouped together the discriminant and differentially abundant taxa for ease of visualization.

Figure 3 – The legend for A reads: “Six MAGs were differentially abundant or exclusive to a behavioral phenotypic group and highlighted on the tree based on the enrichment in either the aggressive (yellow) or the tame (teal) populations”; is this what the outermost layer of the tree represents? If yes, there are way more than just six MAGs annotated in such a way. If not, what is this layer representing?

Thank you for bringing this to our attention, there was a typo in the figure legend and it has now been adjusted to state the correct number of differentially abundant MAGs (L255).

Figure 3 - The legend for B reads: “Detection of GBMs (minimum 66% coverage) in the 22 differentially occurring fox MAGs highlighted in A)” (full stop missing). There seem to be way more than 22 highlighted GBMs in panel A?

Thank you for bringing this to our attention. Fig 3A was an earlier version that included the output from the 2015 and 2017 data sets. We have now removed the output from the 2015 analyses and amended the figure.

Figure 4 – This might be a submission issue, and the authors completely blameless; but this figure is barely legible. I think I can see some sort of phylogenetic tree at the top of panel A? And cant at all read any of panel A’s left labels.

We agree that the details in the figure were difficult to decipher and have both adjusted and moved this figure to the supplementary information document.

Methods or Supps: no mention of how the phylogenetic tree was generated.

Thank you for bringing this to our attention. The phylogenetic tree was generated with FastTree 2.1.10 and visualized in anvi’o. We have now included this in the methods section (L502-503).

Methods: are male and female foxes equally present and equally distributed in the dataset?

In the 2017 data set, 20% of the individuals sampled were females and were equally distributed between the two selection lines (details on each individual fox sampled can be found in the original Supplementary Table 1a).

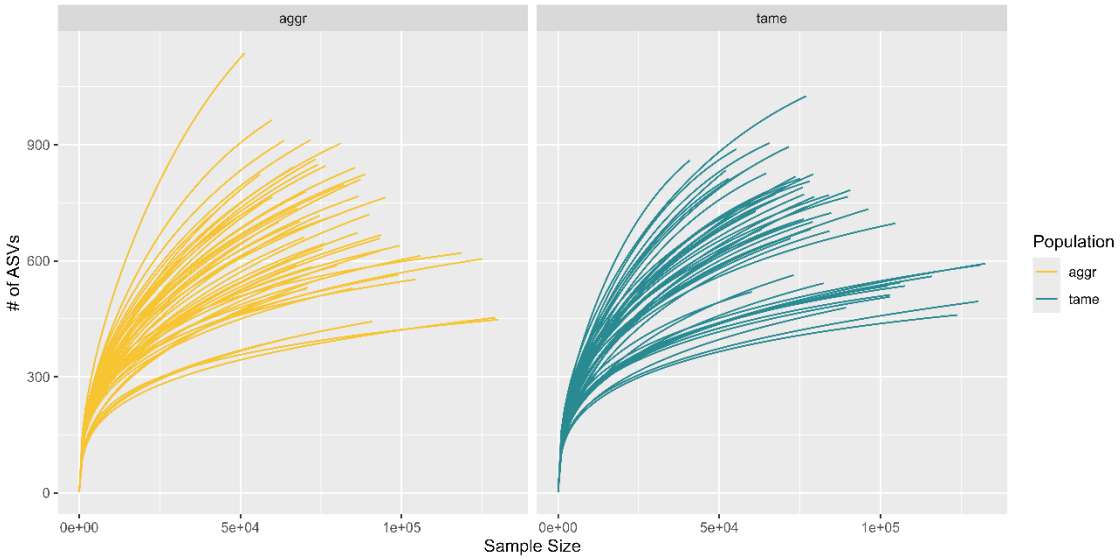
Line 418: how many sequencing runs were the samples divided into? If only one sequencing run was performed, then all good, but if more than one, dada2 has to be performed on each run separately.

All samples from the 2017 dataset were ran on the same Illumina MiSeq sequencing run to produce the 16S sequence data and dada2 on the 2017 data set was ran independently from the 2015 data set.

Methods: did the authors look at the rarefaction curves of the 16S data? How do they inform the sampling completeness?

We appreciate the reviewer’s suggestion. Yes, we examined rarefaction curves for all samples (see Figure below). The rarefaction curves indicate that the sequencing depth was generally sufficient to capture the majority of the microbial diversity in each sample. However, some samples show steeper curves that do not completely level off, which may indicate slightly lower completeness for those specific individuals. Overall, the

rarefaction analysis supports that the sampling effort was adequate for comparative diversity analyses between the tame and aggressive populations, especially when combined with the statistical packages we used herein that account for uneven and incomplete sequencing depth (i.e. Divnet and corncob).



Reviewer #4 (Remarks to the Author):

This paper uses 16S rRNA and metagenomic profiling to test whether artificially selected tame and aggressive fox lines differ at the level of the gut microbiome. Because gut microbiomes could change as a consequence of domestication, or could be mechanistically involved in the domestication process, the study has an interesting premise. A major strength is the use of the Belyaev fox experiments, which have already revealed a remarkable syndrome of responses to strong artificial selection on behavior. The system is a textbook example of change during domestication, and therefore an excellent choice for the present work.

I have no real concerns about the main data analyses, which are basically straightforward tests for differential alpha diversity, microbial abundance, or genic abundance. I agree that the results establish clearly differentiated gut microbiomes in the two selected lines.

3.1 My primary concern is that there is a great deal of speculation woven into the

presentation of results (also in the abstract, intro, and conclusions—but I think it's most likely to mislead in the results).

We thank the reviewer for bringing up their concerns and to address some of these concerns, we now include alternative hypotheses, such as how physiological and morphological differences between the fox populations, and possible founder effects, may influence the GM of farmed foxes (L120-126; L354-359), alongside our previous discussion on how host genetic mechanisms may be driving the gut microbiome community trends (L370-393).

3.2 The authors clearly favor the idea that the gut microbiome is causal to the domestication phenotype and do a lot of selective citing of the literature to make this argument. However, there is no experimental evidence showing that manipulation of the microbiome per se (and particularly the microbes identified as differentially abundant in this study) affects other behavioral or domestication-related phenotypes, only plausibility arguments. There is also no serious discussion of alternative interpretations of their findings. For example, domesticated animals often experience changes in gut length, which could also affect the microbiome, making the differences reported here a by-product, not a cause, of the domestication syndrome.

Consequently, hypotheses that are testable in this study and find some support (e.g., lines 80-81: "we hypothesize that selection for behavioral traits in the fox domestication experiment has…led to selection on…[the] gut microbiota") are combined with hypotheses that are not tested at all (line 81-82: "…contribut[ing] to shaping fox behaviors."). Straightforward results presented in one sentence and based on empirical data are then combined with long discussions of "possibilities" that support the authors' favored interpretation (this is an issue throughout the entire Results/Discussion section).
Thank you for pointing this out. We have now reworked the hypothesis in the section of the manuscript (L77-84)

Without much more targeted experimental work and functional characterization, my view is the conclusions suggested in the study are therefore premature--even though the evidence that the foxes' gut microbiomes differ in a number of ways (the meat of the actual data generation and analysis) is strong.

In our exploratory study, one of the main goals was to identify if differences in the microbial community composition exist between the two selection lines and if yes, what were some of the implications that these differences may have with regards to behavioral selection and early domestication in silver foxes. We know from previously published studies that gut bacteria are strongly implicated in brain development, host

behaviour, and cognition (as highlighted within the manuscript), therefore, it is possible that when selecting specifically for behavior in the host, we are unintentionally selecting for certain host associated bacteria that may contribute to the reduced fear-like and aggressive behaviors in the foxes. Exact mechanisms through which the gut microbiome contribute to the gut-brain-axis, and hence behavior, remain elusive. However, extensive research has associated certain gut bacteria with host behaviors, including fear, with the neuroactive compounds that are produced and/or processed in the gut by bacteria. The solid foundations of this research led us to our main hypothesis in the manuscript.

In this regard, by characterizing how gut microbiome variation and host behaviors co-vary during early animal domestication, we provide an important first step that can be followed up with manipulative approaches capable of generating further knowledge into how the GM contributes to behavioral plasticity during early domestication (which we stress in the conclusion section of the manuscript). We also acknowledge that, although intriguing, any role for the microbiota in the evolution of domestic behaviors in animals does not displace other contributing factors, such as host genetics, but rather adds an additional layer to our understanding of how such behaviors arise.

Minor comments:
-Differential abundance analyses are reported with “a controlled false discovery rate” but it is not clear what FDR correction approach is used throughout and what the FDR threshold for significance is. This makes it hard to interpret statements about significance, and should be clarified throughout.

Thank you for bringing this to our attention. The false discovery rate was estimated using the Benjamini–Hochberg procedure. The appropriate reference was previously included however we have now directly stated the procedure throughout the text (L209-211 & L536)

-In Figure 3, the legend refers to six MAGs that were differentially abundant between fox groups, which are shown, I think, in the outermost ring of Figure3A. However, there are far more than six color blocks on the outer ring, so I must be missing something—please clarify.

We apologize for the confusion and thank you for bringing this to our attention, there was a typo in the figure legend and it has now been adjusted to state the correct number of differentially abundant MAGs (L255). Please see the response to reviewer #2’s comment on Figure 3 above.

The conclusions in the GBM section would be moderately strengthened if, instead of

discussing case example GBMs that were differentially abundant between fox strains, there was clear evidence that neuroactive GBMs overall were more likely to be differentially abundant than expected by chance (a simple variant of a gene set enrichment approach). It's not clear that this is the case.

We thank the reviewer for this thoughtful comment. Our approach in this study was hypothesis-driven and the case-by-case discussion of differentially abundant GBMs (e.g., those related to serotonin, glutamate, and GABA metabolism) was intentional and biologically motivated. These neurochemical signalling systems have been repeatedly implicated in behavioral phenotypes associated with domesticated animals, including foxes, and their corresponding genomics regions (e.g. genes encoding neurotransmitter receptors and transporters) have been identified as targets for selection during animal domestication. By examining the microbial capacity to contribute to these specific pathways, we aimed to explore a potential microbial complement/supplement to the molecular and genetic mechanisms already known and/or implicated in the host.

While a broader gene set enrichment analysis of GBMs is a valid and interesting approach, it was outside the scope of our hypothesis-driven design. Nevertheless, we agree that future work could benefit from incorporating such an analysis to provide a more comprehensive view of how neuroactive capacity compares to the overall functional landscape of the microbiome.

Referee expertise:

Referee #2: Ecology, Host-Microbe Interactions

Referee #3: Animal Microbiota and Behaviour

Referee #4: Behaviour and Evolution

Reviewers' comments:

Reviewer #2 (Remarks to the Author):

The authors have responded to all of my comments sufficiently, and have carried out thoughtful revisions to all reviewers' comments.

Reviewer #3 (Remarks to the Author):

The authors have appropriately addressed my main concerns in the text, or have convincingly replied to them in the rebuttal letter. I support the publication of this manuscript, I just found two minor points that should be addressed before that happens:

L84 – provide reference for your red jungle fowl experiment.

We thank the reviewer for pointing this out and have now included the appropriate reference in the manuscript. (L84)

L418 – "descent" instead of "decent"

We thank the reviewer for noticing this and have made the appropriate edit in the manuscript. (L418).

Reviewer #4 (Remarks to the Author):

This is the second time I've seen this manuscript. I appreciate the additions the authors have made in the revision process. However, while I still find the basic analysis sound, I continue to view the interpretation and discussion as speculative. The authors mention several hypotheses that could account for differential abundance of microbial taxa and/or "gene modules" in their data set (roughly lines 118 to 130), but typically limit their discussion of individual results to a causal role of the gut microbiome in modulating behavior/brain function.

We thank the reviewer for their continued engagement with our manuscript and for acknowledging the improvements made during revision. We understand the concern

regarding speculative interpretation; however, our hypotheses were informed by a well-established body of literature demonstrating links between the gut microbiota and behavioral phenotypes, including fear- and aggression-related responses. To further clarify the basis of our hypotheses, we highlight key findings from this literature and outline how they informed our conceptual framework and study design.

Some of the most compelling studies on the role of the GM in regulating fear and aggressive behavior in their hosts come from germ-free (GF) and antibiotic treated animal models. GF animals were developed to examine the consequences of a complete absence of microbial signaling in the gut on host physiology and behavior [1]. Remarkably, findings from these experimental studies have demonstrated that the presence of a functional and diverse gut microbiota affects synaptic plasticity (important for fear learning and memory) [2,3], fear memory retention [4] and extinction [2], and aggressive behaviors [5–8]. Further, the absence or depletion of the GM was accompanied by altered epigenetic and gene expression profiles in the brain associated with these behaviors [2,4–6,9,10], modified blood brain barrier integrity [6], altered immune responses [6,10], and levels of neuroactive metabolites in the brain [2,5,7] compared to control animals. Recolonization of the GM, or probiotic supplementation, could restore normal fear learning [2] and aggressive behaviors [6,7] that persisted into adulthood, however, this only occurred if the transplant was administered during early stages of development. Together these findings indicate that microbiota-derived signals at an early age influence neurodevelopment, learning-related plasticity, and behavior, with potential long-lasting affects later in life.

The GM can also induce transgenerational effects on host behavior and epigenetics. Maternally treated rodents with antibiotics, for example, can produce offspring that display long-lasting and significant shifts in the community composition of the GM, which was further accompanied by increased levels of aggression compared to controls [6,11]. Further, early life exposure to environmental stress can impair fear extinction in rats (i.e. cannot appropriately update their fear response to a once-frightening stimulus when it is no longer threatening), and the same high-fear phenotype was passed down to offspring, despite having never been exposed to the stress themselves [12–14]. Follow-up probiotic treatment not only restored normal fear memory phenotypes in the stressed-treated rats but also reversed transgenerational effects [14]. Recent findings have also identified that the GM metabolism of choline can reduce its bioavailability and downstream metabolites to the host, which in turn altered methyl donor availability, epigenetic regulation and behavior both in parents and offspring [15].

Collectively, these findings provide evidence that the gut microbiome is an important modulator of complex behavioral phenotypes in animals and may affect the ways in which they interact and respond adaptively to changing environments (e.g. animals shifting from a wild to domestic environment). Many studies indicate a biochemical basis for the role of the GM in host behavioral plasticity, which, in turn may affect immune

system responses, brain physiology, epigenetic programming or even gene expression profiles in the brain associated with such behaviors. The gut microbiota also play a key role during neonatal neurodevelopment in animals in which microbiota-derived signals induce learning-related plasticity, including fear extinction learning, that persist into adulthood. Further, both parental microbiome and response to environmental stressors are thought to interact in ways that impact the behavior, gut microbiota and epigenetics of their offspring, all of which have interesting implications for host ecology and evolution, including rapid adaptation via behavioral plasticity. However, the eco-evolutionary relevance of the microbiome-gut-brain axis outside of laboratory animals remains largely unexplored.

One of the main goals of our study was to identify if differences in the microbial community composition exist between the two selection lines of silver foxes and if so, to explore the potential implications of these differences to the emergence of the behavioral phenotypes observed in the experimental foxes. A comprehensive investigation of all possible mechanisms underlying how these gut microbiota shifts may have arisen was beyond the scope of the present work. Instead, our objective was to examine how these differences may contribute to the behavioral plasticity thought to initiate the early process of animal domestication. Consequently, we designed our study to test these literature-informed hypotheses through correlative analyses rather than causal inference. As noted in the manuscript (L354–L358), our approach was intentionally exploratory and not intended to assign causality.

The value of the paper to readers would be further increased if the alternative hypotheses could be linked to specific, stated, differentiable predictions, but ultimately this is a decision in the hands of the authors and the editor. As a small related point, it would be helpful to differentiate when alternative hypotheses have been tested and found no support versus when they haven't been tested (e.g., lines 125 – 126: "such morphological changes have not been identified in the experimentally domesticated foxes"—does that mean this possibility hasn't been examined, or that it has, and is not supported?).

We thank the reviewer for bringing forward these concerns. We have addressed several of these points in our previous revisions. For example, we expanded our discussion of how physiological and morphological differences between fox populations, as well as potential founder effects, may influence gut microbiome composition in farmed foxes (L120–126; L354–359). We also elaborated on how host genetic mechanisms may contribute to the observed gut microbiome patterns (L370–393). In the current revision, we have added further supporting analyses. Specifically, we include two new ordination plots of the gut microbial communities of foxes in the supplemental materials demonstrating no clustering among A) fox housing sheds (located on the same farm in near proximity to one another) nor B) between male (m) and female (f) foxes. These results suggest that neither

housing nor sex are driving the gut microbiome trends observed in our study (Line 120, Suppl. Fig 2).

Only recently have studies started to explore the effects of domestication on host-associated microbiomes. Most of these studies have identified that ecological shifts during domestication, specifically changes in diet, can have major impacts on reshaping the gut microbiome, in terms of community composition and functional potential [16–19]. How shifts in the GM correspond to functionality in the host during animal domestication remains elusive. However, recent findings, for example, indicate that changes in GM function may have enhanced host dietary energy harvest during the transition from a wild to domesticate diet, and thus provided an important contribution to the phenotypic plasticity required to adapt to the new environment [16,19,23]. However, in the case of our study, all foxes were fed identical diets, and housed in similar controlled environments, therefore, observed differences in the gut microbiota between the two selection lines are unlikely due to shifts in environmental factors (L118-120).

Regarding the reviewer's specific question about lines 125–126, we appreciate the opportunity to clarify this point. The original wording may have implied that gut morphological changes had been examined and found unsupported. In fact, this aspect has not yet been explored in the experimentally domesticated foxes. We have revised the sentence accordingly to explicitly state that gut morphology has yet to be examined in this system, ensuring there is no ambiguity about the status of this alternative hypothesis.

We thank the reviewer for highlighting this needed clarification.

Literature Cited

1. Cryan, J. F. *et al*. The Microbiota-Gut-Brain Axis. *Physiol. Rev.* **99**, 1877–2013 (2019).
2. Chu, C. *et al*. The microbiota regulate neuronal function and fear extinction learning. *Nature* **574**, 543–548 (2019).
3. Konstantinidis, I. *et al*. Major gene expression changes and epigenetic remodelling in Nile tilapia muscle after just one generation of domestication. *Epigenetics* **15**, 1052–1067 (2020).
4. Hoban, A. E. *et al*. The microbiome regulates amygdala-dependent fear recall. *Mol. Psychiatry* **23**, 1134–1144 (2018).
5. Jia, Y. *et al*. Gut microbiome modulates Drosophila aggression through octopamine signaling. *Nat. Commun.* **12**, 2698 (2021).
6. Leclercq, S. *et al*. Low-dose penicillin in early life induces long-term changes in murine gut microbiota, brain cytokines and behavior. *Nat. Commun.* **8**, 15062 (2017).
7. Watanabe, N. *et al*. Effect of gut microbiota early in life on aggressive behavior in mice. *Neurosci. Res.* **168**, 95–99 (2021).
8. Sylvia, K. E., Jewell, C. P., Rendon, N. M., St. John, E. A. & Demas, G. E. Sex-specific modulation of the gut microbiome and behavior in Siberian hamsters. *Brain. Behav. Immun.*

60, 51–62 (2017).

9. Bravo, J. A. *et al.* Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proc. Natl. Acad. Sci.* **108**, 16050–16055 (2011).

10. Liu, Y. *et al.* Increased persistence of avoidance behaviour and social deficits with L.rhamnosus JB-1 or selective serotonin reuptake inhibitor treatment following social defeat. *Sci. Rep.* **10**, 13485 (2020).

11. Cusick, J. A., Wellman, C. L. & Demas, G. E. Maternal stress and the maternal microbiome have sex-specific effects on offspring development and aggressive behavior in Siberian hamsters (Phodopus sungorus). *Horm. Behav.* **141**, 105146 (2022).

12. Callaghan, B. L. & Richardson, R. Early experiences and the development of emotional learning systems in rats. *Biol. Mood Anxiety Disord.* **3**, 8 (2013).

13. Cowan, C. S. M., Callaghan, B. L. & Richardson, R. Acute early-life stress results in premature emergence of adult-like fear retention and extinction relapse in infant rats. *Behav. Neurosci.* **127**, 703–711 (2013).

14. Callaghan, B. L., Cowan, C. S. M. & Richardson, R. Treating Generational Stress. *Psychol. Sci.* **27**, 1171–1180 (2016).

15. Romano, K. A. *et al.* Metabolic, Epigenetic, and Transgenerational Effects of Gut Bacterial Choline Consumption. *Cell Host Microbe* **22**, 279-290.e7 (2017).

16. Reese, A. T. *et al.* Effects of domestication on the gut microbiota parallel those of human industrialization. *Elife* **10**, e60197 (2021).

17. Rampelli, S. *et al.* Bronze Age Coprolites Unveil the Buffering Role of the Gut Microbiota in the Dietary Adaptation of Dogs to Domestication. *SSRN Electron. J.* (2021) doi:10.2139/ssrn.3850374.

18. Metcalf, J. L. *et al.* Evaluating the impact of domestication and captivity on the horse gut microbiome. *Sci. Rep.* **7**, 15497 (2017).

19. Alessandri, G. *et al.* The impact of human-facilitated selection on the gut microbiota of domesticated mammals. *FEMS Microbiol. Ecol.* **95**, (2019).

20. Prabhu, V. R., Wasimuddin, Kamalakkannan, R., Arjun, M. S. & Nagarajan, M. Consequences of Domestication on Gut Microbiome: A Comparative Study Between Wild Gaur and Domestic Mithun . *Frontiers in Microbiology* vol. 11 133 at https://www.frontiersin.org/article/10.3389/fmicb.2020.00133 (2020).

21. Fu, H. *et al.* Domestication Shapes the Community Structure and Functional Metagenomic Content of the Yak Fecal Microbiota . *Frontiers in Microbiology* vol. 12 at https://www.frontiersin.org/articles/10.3389/fmicb.2021.594075 (2021).

22. Bowerman, K. L. *et al.* Effects of laboratory domestication on the rodent gut microbiome. *ISME Commun.* **1**, 49 (2021).

23. Alberdi, A., Martin Bideguren, G. & Aizpurua, O. Diversity and compositional changes in the gut microbiota of wild and captive vertebrates: a meta-analysis. *Sci. Rep.* **11**, 22660 (2021).