Point by point response letter to Reviewer comments

Reviewer 1.
A few decades of experiments have shown that, all other things being equal, communities with lower species diversity often have reduced ecosystem functioning and stability. Therefore, conserving and restoring species diversity is often touted as a key priority for science policy programs. However, the implications of these "biodiversity-ecosystem function" (BEF) experiments have been controversial because species do not go extinct randomly, and also because ecological stressors placed on populations may have strong consequences for ecosystem function that precede species extinction. For example, threatened species often have small or declining populations with low population genetic diversity and high inbreeding. Yet, there has been far less investigation into the ecosystem consequences of these declines.

Raffard et al present an elegant experiment where they test the consequences of changes in genotypic and functional diversity of a fish, *Phoxinus phoxinus*, on biodiversity and ecosystem functioning of fish and adjacent trophic levels. They make two important contribution to biodiversity-ecosystem functioning (BEF) literature. First, the authors seek to understand whether genetic diversity is a better proxy for differences in phenotypic expression of complementary but potentially unmeasured functional traits (e.g. body size), or whether measures of functional diversity are sufficient to capture effects of consumer variation on ecosystem functioning. Second, they use fish as their model organism. In a literature dominated by experiments reporting manipulation of plants, it is refreshing to see the ecosystem effects of a changes in larger consumer species also being documented.

They found that fish diversity consistently influenced ecosystem functioning by altering benthic invertebrate diversity, and also by altering zooplankton diversity but only when functional diversity was high. The experiment is carefully conceived and the treatments are well justified using several combinations of fish from wild populations with previously determined genotypic differences. The observed effects are likely robust as the experimental mesocosms were stocked with zooplankton and algae and natural colonization of invertebrates was also allowed. The natural colonization added an extra component of realism and stochasticity that is often missing from highly controlled experiments.

R: We really appreciate your general comments about our study, and we warmly thank the Reviewer for highlighting the strengths of our experiments and results. As you will see below, we have carefully addressed the comments you provided.

My only minor concern about the manuscript is that cryptic phenotypic diversity is emphasized as a mechanism underlying genotypic diversity. While I don't doubt this is true, it seems that there should be several more functional traits that could be easily measured in addition to body size, which is the trait predominately emphasized in the text. It would be helpful to give some concrete examples of other traits, even if they are not explicitly measured in this experiment. Otherwise, discussion of the term cryptic diversity may seem unnecessarily vague.
R: We agree that 'cryptic' functional diversity may include a lot of functional traits, and that we did not provide enough details on what could be the underlying traits. Following Reviewer's suggestion, we added details on the potential functional traits that may affect the ecological process studied in the manuscript. For instance, behavioural traits (such

as activity) are increasingly studied as strong predictors of ecosystem functioning as they are linked to individual diets, ultimately altering trophic interactions in food webs (e.g., Toscano et al. 2016, *Oecologia*). See the added section L. 74-75 and 343-351.

The following comments are made only to show that my lack of major criticisms is not due to lack of reading the manuscript. The rationale for the experiment is well defined and well-justified within the scope of broader literature. The methods are informative. The statistical analyses are appropriate and clearly described. All tables and figures are useful and carefully prepared. The discussion thoughtfully evaluates the results. Relevant literature is cited discerningly.

R: Thanks again for these positive comments!

Minor comments Line 335 typo: extend not extent
R: Following Reviewer 2 'extend' has been changed to 'expand on'.

Line 337 typo sustain(s)
R: Done (L.369)

Reviewer 2
This study experimentally explores how intraspecific genotypic and functional diversity of consumers affect ecosystem functioning at lower trophic levels. I found this especially interesting given that I have rarely come across studies that account for intraspecific diversity, let alone considering both genotypic and functional diversity, and (!) that also considers the effects of intraspecific diversity at different trophic levels. Therefore, I found this a very exciting and especially novel test of BEF theory that potentially brings to light some very important findings.

R: We thank the Reviewer for this positive comment. As you will read, we have carefully considered the specific comments you provide, which have improved the quality and clarity of the manuscript.

That said, I also have a number of concerns about the level of detail provided in the methods and also discussion, as well as with the analytical approach, in particular with regards to the development of hypotheses and how they have been tested (or perhaps not) in the path models. In particular, clearer justification is needed for the hypothesized relationships. for example, why do the authors expect main effects of genotypic and functional richness on zooplankton abundance and richness, but only an interactive effect on zooplankton diversity (Fig S3b)? This is assuming that the diagram in Fig S3 indicates an interaction between genotypic and functional richness, though that is also not explained anywhere. In that vein, why would the authors expect an interaction between genotypic and functional richness in (b) but not in (a)?

R: Interactive effects were initially included in all simple models (i.e., to test hypotheses related to potential interaction identified in Fig. 1), and they were removed when they were not significant so as not to bias effects of the simple terms. The procedure is actually very similar to that used for simple models (Table 1); interaction terms were initially included and then eliminated when not significant. This is now stated line 266.

It would also be helpful to see a list of the 'independence claims' (the basis set) used to construct the path model, to ensure this matches up with the hypothetical models presented in Fig S3.

R: The list of independent claims has now been added in Table S2, as required.

Furthermore, there is no explanation of the 'mechanisms' tested in these path models; that is, the modulating effects of invertebrate and zooplankton abundance/diversity on decomposition and algae. Therefore, it is still unclear exactly what hypotheses the authors are trying to test, which is critical for confirmatory analyses such as these.

R: We agree that biological hypotheses in a causal path-analysis are necessary to understand mechanisms. Therefore, we added the rationale and details on our hypotheses in the Method section (L.262-266).

In addition, the path models shown in Figure 3 appear to have weighted arrows, but it is not explained how the arrows have been weighted. Are these standardised coefficient weightings? Or are they unstandardized coefficients? This is also crucial for interpreting these path models.

R: The arrows were weighted in accordance with the p-values from the models, this is now detailed in the caption (L.651). We did not include coefficients to weight the arrows because some of our explicative variables are factors (e.g., functional and genotypic richness) and in that case it is not possible to extract coefficient estimates as it is for continuous variables. Another option would be to weight the arrows with F-values (which should not change much the figure); if you think this is more relevant (or if you have another suggestion) we would be pleased to modify the figure accordingly.

In addition, It was not at all clear to me what level of variation in genotypic richness there was between the "low" and "high" genotypic richness treatments, although there are details given in the supplement on variation in body mass and functional richness. Was genotypic variation actually quantified, or just assumed given the geospatial separation among sites where the minnows were collected? I see this might be available information from another study, but this should be included in this paper as it is rather fundamental to the experimental design of this study.

R: The low and high genotypic richness treatments differed in the number of genotypes they include. The low genotypic richness treatment included 2 genotypes (i.e., 2 isolated populations) while the high genotypic richness treatment included 4 genotypes (i.e, 4 isolated populations). We did not quantify genotypic richness (e.g., allelic richness from microsatellite markers) from the experimental populations, but we indeed choose them from previous studies (Fourtune et al. 2018, *BioRxiv*; Raffard et al. 2019, *Ecol. Evol.,* Prunier et al. 2019 *BioRxiv*) to maximize both functional and genotypic richness. The populations that we choose have been genotyped at 17 microsatellite markers in previous studies (Fourtune et al. 2018, *BioRxiv*; Raffard et al. 2019, *Ecol. Evol.,* Prunier et al. 2019 *BioRxiv*). From these data, we measured that population differentiation was high (mean pairwise Fst = 0.133, range = 0.029-0.320). Also, the number of alleles within populations varied from 5.47 to 10.17 in average across loci. Populations used for experiments were hence different genetically one from each other and displayed different allelic diversities. As a result, if we use this data to quantify the potential allelic richness for the two different types of treatment (2 vs 4 genotypes), we found that the experimental populations from the low genotypic treatment are expected to have in average 11.01 alleles, whereas experimental populations from the high genotypic treatment are expected to have in average 14.23 alleles (t-test<0.001). As this diversity

was not directly measured in the tank, we did not include this later information in the MS. Nonetheless, we now provide more information about the genetic make-up of the populations we used to assemble the experimental populations (L.144).**

Specific comments:
Line 38 - 39: Strange wording "The loss of genotypic richness was similarly prejudicial than the loss of functional richness". I'm not quite sure what the authors mean by this - please clarify.
R: This statement has been modified (L.39).

Line 55: The "iBEF" abbreviation does not really work if plural, as it stands for "intraspecific biodiversity-ecosystem functioning", without "relationship". This should rather be written as "iBEF relationships". Please check throughout manuscript.
R: Done (e.g, L. 50, 79, 96, 330).

Line 64 - 65: How does that approach not "provide mechanisms"? Which mechanisms? The concepts here are quite vague and need some more development.
R: Actually, genetic diversity *per se* does not affect ecosystem functioning. The effects are necessarily supported by functional traits, which are responsible for interactions between individuals and the environment. Therefore, studying solely the ecological effects of genetic diversity does not provide mechanisms of effects on ecological processes. We amended the sentence to make it clearer (L.64-66).

Line 76: Change "and a better mechanistic understanding" to "and could provide a better mechanistic understanding".
R: Done (L.79)

Line 78 - 79: One might argue that biodiversity effects can be just as strong with changes in lower trophic levels, because of bottom-up effects that could resonate through to higher trophic levels.
R: It is true that biodiversity is important in all trophic levels. Yet, several key studies suggested that biodiversity effects can be particularly strong in higher trophic levels (Griffin et al 2013, Ecology; Duffy 2002 Oikos; and Duffy 2003 Ecology Letters). References have been added and we modified the sentence to moderate the statement (L.80).

Line 91: I am not sure about the terminology of "essential" ecosystem functions. What would be considered a "nonessential" ecosystem function?
R: We wanted to highlight the importance of ecosystem functions such as decomposition rate and primary production, which are the target of a number of ecological studies. Yet, you are right, it is very subjective to qualify ecosystem functions as essential and non-essential, so we have removed this statement.

Line 93 - 94: "Predatory and consumer species" - this is redundant and simpler to just say "consumer species".
R: Done (L.95)

Line 109: "At the opposite" - change to e.g. "Alternatively".
R: Done (L.111)

Lines 111 - 114: Why would the relationship differ in the ways presented in Fig 1? Are these hypothetical scenarios based on any previous research? Some more development of these hypotheses is needed to give some conceptual foundation underlying these expectations.

R: These scenarios were primarily developed conceptually. Since our study is relatively new in comparing genotypic and function richness, these scenarios were not theoretically grounded. As suggested, we added details on these predictions (L.113).

Line 114 - 115: Perhaps this is true, but until this point there is no explanation of these mechanisms. What are they, and how would they modulate the effects of genotypic and functional richness on ecosystem functioning? Also, change "bases" to "basis".

R: These mechanisms were essentially direct trophic effects and subsequent trophic cascades. This is now explained in the manuscript (L.122).

Line 116: "affected directly populations' biomass production" - change to "directly affected population biomass production"

R: Done (L.125)

Line 117: I think the authors need to more clearly define what they mean by 'ecosystem function'. Biomass production is typically referred to as an ecosystem function in the BEF literature so this could be rather confusing if a clear definition is not provided earlier in the introduction.

R: An ecosystem function relates a process (i.e., energy or material flux) that generally integrates multiple species (e.g., biomass production of a community of species). Here, biomass production was assessed at the population level, and was therefore considered as a property of the minnow' population and not an ecosystem function *per se*. We understand the possible confusion, therefore 'population biomass production' was replaced by 'population performance' on line 125.

Line 118: Can the authors give more information on the study system here? E.g., freshwater or marine?

R: Some information on the study system is available on line 104-108 of the Introduction ('…freshwater fish, the European minnow (*Phoxinus phoxinus*)'; and 'pond mesocosms' 106 and 104, respectively). While we understand that information on the study system is important, we choose not to add further details in the Introduction so as not to overload the text, and since information is available in the Method section (L. 131). If crucial information is missing, please let us know, and we will add it.

Line 169: Was it four or six genotypes?

R: The asymptote of the relationship was actually reached in-between four and six genotypes. We changed the text to 'approximately four genotypes' (L.183)

Line 187: "1 - nb. survival fish/nb. introduced fish" - what does this mean and what is it for? This requires some more explanation.

R: This formula assesses the mortality rate (invert of the survival rate). We have changed the format to make it clearer (L.201).

Line 226 - 227: More detail is needed here, and perhaps show results prior to outlier removal so it is possible to assess why the authors felt this should be removed.

R: It is true that outlier removal is an important step in statistical analyses. The tank that was removed was a highly influential point. We added a figure in the

supplementary showing results from a Cook's distance analysis (Fig S3). As shown, the tank number 13 was influential in all the variables, so we felt it should be removed so as not to bias the results. As you will see, some tanks (15 & 21) were outliers for some variables and in that case we did not remove the tanks from the analyses. But tank 13 was an outlier for all variables, which drives our choice to remove. We have no explanation for that; it is relatively well known that with mesocosms some strange (erratic) trajectories can occur.

Line 242: "and effect sizes were averaged" - can the authors please explain how the averages were calculated? Specifically, which values were averaged (average across which effect size values)?
R: Effect sizes were first measured for each variable response (n = 7) and for each genotypic or functional richness treatment separately. They were then averaged across variable responses for genotypic richness and functional richness separately. We have modified the text accordingly (L.254-257).

Line 255: Strictly speaking, I'm not really sure the term 'null models' should be used to describe these. They are simplified models that exclude the relationship between both genotypic/functional richness and decomposition/algae stock. However, they are not the simplest possible models.
R: Following this comment, we changed 'null model' for 'simplified model' throughout the text.

Line 277: "At the opposite" - change to e.g. "In contrast".
R: Done (L.301)

Lines 290 - 291: As opposed to? This is written as though an alternative was tested, but that the authors found the strongest support for a trophic cascade. However, the way the analysis is designed, this is the only possible outcome that could have been found. If the path models also allowed for direct effects of minnow richness on ecosystem functions, significant direct effects might indicate that there are other factors at play, other than indirect/cascading effects on ecosystem functioning. Of course, it could also mean that other measures of invertebrate/zooplankton community structure are important (aside from abundance or diversity), that were not quantified in this study. This seems important to include in the analyses in order to confidently make statements such as this.

Line 296 - 297: Again, this statement is unfortunately not supported by the path models, as it appears that direct effects of intraspecific diversity were not included in the full models. Thus, the results could also potentially arise from correlations of invertebrate and zooplankton abundance/diversity with other unmeasured (but potentially more important) variables.
R: Actually, the C-statistic gives support to the indirect effects. Indeed, if p-values > 0.05, the paths adequately fit the data. Here, in both path-analyses p-values > 0.05, indicating that the causal pathways were adequately reproduced (including the indirect effects). Nonetheless, we agree that our model comparison (vs. null/simplified model) was not adequate. Therefore, following your suggestion, we now compared the basic model with an alternative model including direct effects. As you can see, the results are consistent and support indirect effects of genotypic and functional richness on ecosystem functions (Table 2).

Line 306: "iBEFs" - change to "iBEF relationships".

R:Done (L.330)

Line 335: "extent" - change to e.g. "expand on".
R: Done (L.367)

Line 337: "diversities" - change to "diversity".
R: Done (L.369)

Line 345: "diversity benthic invertebrates" - change to "benthic invertebrate diversity".
R: Done (L.377)

Line 357 - 358: It is not very clear how this relates to, or explains, the results presented in the preceding sentence.
R: We think that the primary explanation for the positive relationship between functional richness and invertebrate abundance is that functionally rich populations might consume less resources due to stronger competition. Yet, it was probably not the case since functionally rich population display higher biomass production than functionally poor populations. The sentence has been modified to make it clearer (L.386-389).

Line 360 - 362: I would still be interested to hear the authors' thoughts on why enhanced intraspecific diversity of fish caused an increase in invertebrates and zooplankton. Until now, there is no clear explanations provided that could be tested in future experiments.
R: We are not sure to understand this comment. If the Reviewer means invertebrates and zooplankton abundance, or diversity, or both. Nevertheless, regarding invertebrate diversity we suggest that genotypic richness enhance resource partitioning, releasing the predation pressure on more taxonomic group, increasing the diversity of invertebrate (L.378-381). Regarding the effect of functional richness on invertebrate and zooplankton abundance the mechanisms are less clear. We propose that increasing functional richness may have constrained some individuals to switch on alternative resources (such as periphyton or detritus) decreasing the overall predation pressure on invertebrates and zooplankton, hence increasing abundance (L.389-393). Yet, these explanations remain hypothetical and we prefer to remain cautious. We also propose that measuring diet (e.g. using gut content of stable isotope analysis) may be useful in confirming those explanations (L.393-397).

Line 364: "basis" - change to "base".
R: Done (L.399)