Dear Dr. Marañón,

thank you for serving as Editor for our manuscript "Species richness and functional attributes of fish assemblages across a large-scale salinity gradient in shallow coastal areas".

Please find below our details on how the manuscript was revised to address the reviewer's comments (which are kept in italic style).

This is the second round of review for Koelher et al. I thank the authors for addressing and answering in detail the reviewers' comments. I think that the authors have done good work reviewing the manuscript and that this has helped make the manuscript more robust and streamlined. I still have a few minor comments for the authors to consider. Please note that I refer to the Lines from the manuscript with track changes, sorry for the confusion.

Main:

- Figure 2- and table 3: why the expectation of linear regression? how are the residuals looking? (any patterns?). In many cases it looks like a non-linear relationship would be more appropriate, maybe some GAMs with restricted k smoothing factors. For example, Fig 4c. SR obs does not seem to follow a linear relationship (residuals not homogeneously distributed, all observed SR at low salinity are under the regression line)

Here, we wanted to conduct univariate statistical testing on the different SR estimates (overall and for the functional groups, as well as observed, standardized and estimated SR) vs. salinity (and sometimes temperature). To address the aspect of potential non-linearity, the distributions of the variables were carefully assessed, and transformations (log10 or Yeo-Johnson, see Sect. 2.6 "Statistical analyses") applied to best meet the requirement of normally distributed residuals. The resulting residual distributions were assessed using standard residual diagnostic plots in R, assuring that no pronounced residual patterns remained.

Concerning Fig. 4c, only part of the green data points is well visible, as they are very similar to the other (red and blue) data points and overlying each other in some cases. We added a note in all respective figure legends stating that data points, lines and confidence intervals are in some cases overlying each other, and referring to the Table with regression results and to the Supplementary Table with raw data for details. The residuals of Fig. 4c did not violate the assumption of homogeneously distributed residuals.

- Incidence (e.g. L83 and throughout). The term "incidence" is not defined. Is an incidence an observation / occurrence? If so, then why put in opposition with "presence" data (e.g. L460). When reading I thought that the term was sometimes confusing, notably when calculating % incidence of undetected species. You could define what is "incidence information" the first time the word is used.

We understand that we did not in all cases distinguish data sources/formats and terms clearly enough. We separate between: data from systematic samplings, where all caught species were determined to species and the precise sampling locations were known (these were entered into the "fish incidence database"; Sect 2.2), and: data from additional sources, where only

presence/occurrence was reported. As these additional data do not stem from systematic sampling efforts they could not be used in the statistical analyses (see Sect 2.3: "These "additional data sources" were used as complementary information on SR_{obs} but could not be used in the statistical analysis since they did not have complete sampling and species incidence information.").

To accommodate this comment on remaining ambiguity in terminology in some places, we are now more specific on the term "incidence". We included an explanation/definition in the concerned Methods Section, Sect. 2.3. Further, in cases where we previously used the term "presence" in relation to the "additional data sources", we now use "species records" instead. Last, we have made sure that we use the same term (i.e. "incidence data plus additional data sources", as in Table 2) when referring to this data.

Others:

- L81: definition of coastal areas? distance from the shore? For IBTS data, which criteria is used to be defined as coastal sampling?

Coastal areas were defined in alignment with the EU Marine Strategy Framework directive (also corresponding with biological parameters for the EU Water Framework Directive), and obtained from the Water Information System of Sweden. We added the definition and source into the text. We also clarified that the same geographical criteria was used for IBTS data.

- L32-33: the link between the first and second part of the sentence does not follow a logical flow, consider separating and rephrasing.

We separated the sentences, and edited this text part.

- L130 "Geographical delineation". I might have missed it, but which geographical delineation are you referring to?

This refers to the statement above (L117/118): "Coastal areas were delineated using official national

definition". We added "of shallow coastal areas" in the pointed out place for clarity.

- L165: This is done per sub-basin?

Correct. We have edited accordingly.

- Table 2 "species incidence" and footnote a. Is this the sum of all samples' unique species richness? (Referring back to comment earlier on incidence).

This is the total number of unique species across sampling occasions. We have edited the footnote for clarity.

Footnote c also needs clarification.

We think this is clarified both by defining the term "incidence" in response to the reviewers second main comment, and by editing the footnote "b", since footnote c is similar but for observed and estimated incidences.

- Table 3: There are so many relationships tested that the p-value has little meaning, you could consider adjusting p-values for multiple comparisons, or delete the p-value, the R2 on its own can give support on whether there is a clear pattern or not. I am not a big fan of p-value, but would of course understand if the authors want to keep it.

These are the results from simple linear regressions, i.e. between salinity and fish SR in the different functional groups, one group at a time. Therefore, multiple comparison adjustments of the test statistics (and P-values) were not required (no multiple testing conducted). We edited the table legend for clarity in this respect, and like to keep the P-values (in addition to the R^2 values that are also included in the table).

- L275: instead of "exist" maybe "present in the sub-basins"

We edited as suggested.

- Fig 1. could you increase the height of the figure? small height gives a wrong impression that all curves have reached the asymptote.

We followed the reviewers suggestion and edited the height-width ratio of the figure for improved visualisation.

- L438: So more concretely, the IC suggests that how many species are probably occurring but not observed, 1-2 species still not identified? More?

The IC gives the proportion of species detected from the estimated "true" SR (explained in Sect. 2.3). Hence, 100% - IC gives the percentage of species statistically likely to exist but not detected in the sampling. We understand that the reviewer would like to "translate" the IC to "number of undetected species", i.e. SR_{est} minus SR_{obs}. We edited in this information.

- L445: this sentence could be improved: "which are only more rarely present"?

We agree, and edited the respective sentence.

- L451: Based on your results I would say that SR was also well described for rare species? And that few, extremely rare species might still not be listed

We generally agree. This is also stated in Sect 3.3.: "The SC exceeded 98.5% in all sub-basins (Table 2), suggesting that these undetected species were highly rare, likely representing <1.5% of incidences." The argument here is based on that SR_{obs} was similar to SR_{est} for Shannon diversity (number of frequent species) and Simpson diversity (number of highly frequent species; Table S5). In contrast, the total SR_{obs} was lower than the total SR_{est} by 1-37 species, depending on sub-basin, i.e. this number of (highly) rare species likely remained undetected. Hence, the SR of all species (including rare species) was not as well described as the SR of frequent and highly frequent species.

To accommodate this comment and clarify the argument we moved this respective paragraph higher up in the discussion. We included the information how many rare species (in numbers) were statistically estimated as undetected based on the incidence data. This discussion aspect is then immediately followed by the statement that the species lists were essentially complete when additional data sources were considered, and explicitly edited to name the rare species here again. We think that this revision gives a better flow and takes care of the raised aspect. Further, we edited in Sect. 3.5 from "rare and very rare" to "highly rare", and similarly in the concerned paragraph of the discussion.

- L462: "50-90%" where does this range comes from? Maybe should be written in the results section.

We now clarify in brackets how this ratio is calculated. The values are presented in the results, but we like to highlight this comparison in the discussion.

"are found" (= are occurring in the area) or 50-90% were "observed in the samples"? If the last one, then shouldn't it be "at least 50-90% …"?

Here, as now specified in response to the comment (see above), we compare SR_{est} with the HELCOM species numbers per sub-basin. Our SR_{est} is corrected for sample size. The SR_{obs} from HELCOM is not, and may hence be a lower bound estimate. Therefore, the here estimated ratio might be biased somewhat high, which is why we wrote "ca.". We wrote "are found" since it concerns the estimated SR, i.e. the sample-size corrected species richness.

- L516: "expected" instead of indicated? you didn't test

We agree and changed the wording as suggested.

- SC: SC, throughout the text for simplicity maybe only write "sample coverage" without the abbreviation as it is not consistently used

We think the acronym is useful, since it is commonly used, and also used repeatedly throughout the manuscript. In some places the word was spelled out for specific reasons, e.g. in the Abstract prior to definition, or in a Figure legend, but otherwise the acronym is used consistently. We like to keep it.

- L545: do you rather mean "functionally-driven" differences?

We removed this word from the sentence, as it was clear without.