**COMMENTS TO AUTHORS:**

I now had three reviewers provide comments about the manuscript that has been submitted to PCI Evolutionary Biology entitled ‘Evolution at two-time frames shape structural variants and population structure of European plaice (Pleuronectes platessa)’. This manuscript explores the population structure and variation in two structural variants (SVs) in the European plaice in the North and Baltic Sea. Previous work identified these two SVs on chromosome 19 and 21 and this study further explores this variation by incorporating additional sampling locations and attempting to date the SVs. The paper also investigates whether the SVs could have been introduced through introgression from another species that is known to hybridize with plaice. Overall, I think the paper is interesting, and this was also shared by the reviewers. Given the increasing awareness about the importance of SVs to population structure, I think this paper would be of interest to many readers.

However the reviewers have also highlighted a number of issues that need attention and have provided detailed and constructive comments below on how the manuscript could be improved. The major areas in need of attention are:

1. The general framework needs to be broadened, particularly in the Introduction, to include general details about why the authors investigate the population structure of plaice. Right now, the authors mostly focus in SVs in their Introduction.
2. All of the reviewers found that the Methods lacked critical detail and explanations (e.g. lack of details about sampling and dataset sizes). Please go over the comments that the reviewers have made on a point by point basis and clarify this section.
3. Reviewer 3 added some thoughts about the dating of the SVs, and possible problems with it. Further, it would also be good for the context and interpretation if the authors could provide a bit more detail about the genes that are located in the SVs, and to qualify statements like ‘...many of these were involved in ion transport’ (how many?).
4. The reviewers felt that some of the statements were too vague, and were rather descriptive and lacked quantitative support, and I agree with that. I suggest the authors go over the manuscript again and qualify some of these (many-say how many, most of the times-how often?).

Sincerely,

Maren Wellenreuther

We went through the extensive comments of the three reviewers and found that they raised interesting and fair points. Thus, we have now modified the manuscript with several changes, including the following major revisions:

1- We changed the first paragraph of the introduction, which was focusing too much on the structural variants, to make it broader in order to include the entire framework of the manuscript (basic population genomic and contribution of the SVs).
2- We provided more details in the method section about the geographic sampling and the number of SNPs analyzed, and included some formal tests that were missing for the discussion (e.g. correlation between pairwise Fst genome-wide and pairwise Fst for the SVs to support the fact that they are not correlated).
3- We included a more formal test to check for gene enrichment within the SVs in order to discuss the potential functional implication of the SVs. Previously, we highlighted the role of
ion transport, but this new analysis allowed us to identify strong enrichment in immune
related genes. We discuss it briefly, in order to avoid unnecessary speculation mentioned by
reviewer 2.

4- We also agree with reviewer 3 that the age of the split between Iceland and the European
shelf can provide interesting information about the age of the SVs. However, despite the
numerous biases that can exist in using the phylogeny to age SVs, we also think it provided
additional relevant information. Thus, we have decided to discuss the phylogeny and the
demographic inference together, in order to support the hypothesis that these SVs are at least
older than the age of the gradients in which they are currently segregating.

5- We reorganized the discussion and provided statistical support to clarify the vague sentences
of the previous version. As suggested by reviewer 1, we removed the paragraphs on the
maintenance of the SVs in the population that was too speculative. In addition, we have
included a paragraph about how the plaice SVs seem to fit with the model of evolution at two
time-frames at the end of our discussion, as suggested by reviewer 2.

All the reviewers also raised some concerns about the data filtration based on Hardy-Weinberg
equilibrium.

We understand this comment, however, as we first described the plaice SVs using the same pipeline
in Le Moan et al, 2019 (https://www.biorxiv.org/content/10.1101/662569v1),
we have decided to retain it here. Also, this filter only removed SNPs that deviate from HWe in more
than 60% of the populations, which is not a stringent filtration step as it only removed 37 SNPs from
the data set. These 37 SNPs were located on 17 out of the 24 chromosomes. None of the removed
SNPs were found on the chromosome carrying the first SV and only 3 were located on
the chromosome of the second SV. Consequently, we think that this HWe filtering step has limited
consequences on the SV signals that we detected.

All the reviewers also commented on the relevance of using the collinear regions of chromosome 19
that carries one of the SVs to infer a phylogeny representative of the genome wide divergence. We
understand the issues, however, we decided to keep it in the manuscript, as it shows that the
divergence is limited to what we think is a chromosomal rearrangement and that no signal is detected
in the immediate vicinity of the putative SVs. This provides relevant information about the strength
of such a barrier, that does not seem to affect at all the rest of the genome, even when we looked
within the same chromosome. Nevertheless, we have now included a sentence in the manuscript to
justify our choice as this was unanimously raised by the 3 reviewers:

“Chromosome 19 was used to infer phylogenies both within the SV and to represent a genome wide
phylogeny in order to estimate any effects that the SVs may have on collinear regions of the same
chromosome”

In addition to this, we went through all the specific comments raised by each of the reviewers and
have responded to individual comments below.
We feel that all of these modifications have greatly improved our manuscript. We were really impressed by the quality of the reviews and the relevance of the comments. Thus, we are sincerely grateful for this. The number of comments also partially explained why we took such a long time before resubmitting our manuscript. We are sincerely sorry for this delay, but we hope that you will understand.

Kind regards,

Alan

Reviews 1

Reviewed by anonymous reviewer, 2019-08-13 13:43

I read "Evolution at two-time frames shape structural variants and population structure of European plaice (Pleuronectes platessa)" by Le Moan et al. I found the MS interesting and the findings novel and relevant. I think the methods used are mostly appropriate to lead to the conclusions the authors arrive (except for some minor confusing cases). On the downside, however, I found the MS difficult to read at places, some important details missing and many descriptive results that need quantitative support. I also feel the discussion can be consolidated and/or improved in some of the sections. This paper will be a nice contribution showing the relevance of Structural Variants in the evolution of divergence and demographic change.

Below I include many comments. All of my comments are either minor details or are intended to add a more detailed explanation around the major issues I found in this MS:

1) Lack of details about sampling and dataset sizes etc.
2) Many different concepts are included without properly considering each of them in-depth or without consolidating a proposed mechanism to connect them. e.g., "edge effect", "founder effect", "IBD", "allele surfing".
3) Some inferences are rather descriptive and lack quantitative support.

Comments:

L 9-11 At this early point in the MS, it is a bit unclear why SV "provide evidence" for evolution at two-time frames. I think in general this sentence is a bit confusing.

This was rephrased:

"Changing environmental conditions can lead to population diversification through differential selection on standing genetic variation. Structural variant (SV) polymorphisms provide examples of ancient alleles that in time become associated with novel environmental gradients."

L 22 - What global distribution? If the plaice is not distributed globally, or at least its global distribution is not assessed here...

We have changed “global” to “northern”

L 70-71 This argument needs to be developed a bit further and/or a reference included.

The sentence was: These co-adaptations are expected to arise continuously de novo after the SV associates with an environmental barrier. We could have referred to "reinforcement" here but we
decided to remove this sentence from the current manuscript as it is not directly related to our current framework.

L 83 This needs a reference, a combo later used in the text would work well here (e.g. Jones et al., 2012; Morales et al., 2018)
Done

L 118 - "to examine multiple hydrographic gradients" It is not very clear which are these gradients.
Would be good to mention them here.
Done

L 118-123 Given that the evolutionary history / temporal framework of SV's is the highlighted aspect of the MS, I'm surprised this is not included in the goals here

We rephrased the main goals of the manuscript:
In the present study, our overall objective was to increase our understanding of the origin of the SVs and their effects in contemporary populations of the European plaice. Specifically, the main goals of the study were to: i) re-assess the population structure in European plaice from northern Europe and from Iceland with the use of a population genomics approach, ii) evaluate the contribution of SVs to population structure, iii) test for a potential flounder origin of SVs; iv) estimate the age of the SVs in plaice with a phylogenetic approach; and v) provide relevant data to understand the extent to which selection is involved in maintaining the allelic clines observed at the SV. This work thus provides increased insight into the relative roles of environmental gradients, demographic history, hybridization and genomic structural changes in the evolution of a widespread and highly abundant species.

L 145-146 "Most of THE northern limit of the plaice distribution was covered with this sampling design"
What about the baltic distribution? How much of the Baltic distribution is included in this sampling?
All of the Baltic Sea distribution. From what we know, the plaice is not found after Bornholm nor in the north of the Baltic Sea.

L 184 Why a multiple-testing correction was not used for Hardy-Weinberg?
We did the test per population and we only removed SNPs with a significant deviation from Hardy-Weinberg in more than 60% of the populations. At the end, this filtration is not so stringent (we removed 37 SNPs out of 28000). It is a kind of correction for multiple testing since we removed only the SNPs that deviate from HWe in more than 4 population out of the 7.

L 191 - 195 I suggest to add the final sample sizes for all these different dataset
We have included a table with detailed sampling information to the method section. Moreover, we provide more details about the size of the different dataset throughout the entire manuscript.
L 211 - 222 I'm a bit confused behind the logic of "examine the demographic histories associated with the major population breaks identified in the overall dataset" aren't the largest breaks Icel, Norws and Katte? And then the most interesting to include Bals to date the Baltic colonisation? Maybe I'm being a bit dense and cannot fully understand what the authors are trying to test here. Also, the results of the demographic modelling seem to have very little weight in the discussion later on, so it feels a bit forced. In any case, I would like to see cartoons of the different demographic models tested in the SI.

By major breaks, we meant between the continental shelf and outside the continental shelf. We have rephrased this part to:

"We used an approximation-of-diffusion approach, as implemented in the software from δaði (Gutenkunst et al., 2010), to examine the demographic histories associated with the major population breaks that separate the populations sampled along continental shelf from the Icelandic population (also described in Hoarau et al., 2002). Specifically, we assessed the demographic history of the divergence of Iceland and its closest continental shelf population...."

A cartoon of the model can be found in the supplementary material figure S9.

Additionally, the demographic history of the North Sea and the Baltic Sea was already inferred in our other paper (Le Moan et al, 2019: Beyond parallel evolution: when several species colonize the same environmental gradient, available in bioRxive.) Here, it seems that the Baltic Sea has evolved in face of continuous gene flow from the North Sea (IM model).

L 232 Why is this a haplotype allele frequency? Are not these markers for the entire chromosome? Thus, unlikely they represent a single haplotype?

We understand this comment, and therefore we changed the word “haplotype” by the word “haplogroup”, which is more accurate.

L 248 I don't understand why the genomic architecture is invoked here? "The genomic architecture of differentiation was examined using SNP specific FST values"

This comes back in line 334. I think this is not about the genomic architecture of the SVs, given that SVs are a feature of the genomic architecture. This refers more to the genetic or spatial structure of SVs or something along those lines. I suspect the authors are misusing the definition of genomic architecture here.

We agree with the reviewer and we have changed “genomic architecture of the structural variants” into “Genomic analyses of the structural variants”

L 255 Was LD calculated pairwise across the entire chromosome? With a sliding-window? Please clarify.

Done, it was calculated pairwise.

L 277-281 How these coordinates were defined? By eye? Please explain.

The answer to this question was written in lines 281-283. We have now moved it to the part about the gene content of the SVs (line 265 – 269) and added more details about how we defined the limits of the SVs.
Also why the authors did not use the other chromosomes to represent the genome-wide estimates? How likely/unlikely it is that these regions within the same chromosome are subject to some linked effects? E.g. How well the boundaries of the SV are defined and how long is the LD decay?

See general comments

L 316-320 I do not fully understand what the authors are trying to say here. I find it difficult to appreciate a correlation within the blue dots with only 3 comparisons, also I cannot see the different effects between panel a and b of Fig. 2

We agree with the reviewer that the blue color is not needed here, it just highlight the most distant pairwise comparisons. Therefore, we have removed it from the presentation. Nevertheless, the two graphs show that when including the SVs results in a noisier IBD pattern, which can then be improved when the SVs are removed ($r^2$ improvement of 0.4, 0.56 in Figure 2a and 0.96 in 2b). This noise is clearly shown also in Figure 3a-b.

Methods: there are some details that are missing in the methods when specific things are presented e.g. How many SNPs the authors started with, how many were filtered in each step and how many they end-up with. Also the final sample sizes of the different "datasets". Or how many genes were found where they say "and all genes with more than 80% mapping were recorded". Among other things. Would be good for the authors to double check the methods section and add all these small but important specific details. Particularly how many samples and markers go into each of the analyses.

We agree with the reviewer that our manuscript lacked details about the different datasets. Therefore, we have added this information to all of the relevant parts. We have also included a supplementary table summarizing the number of SNPs obtained after each filtration step.

Figures: the style for presenting panels as "(a)" or "A" and the way they are presented in the legends (e.g. before or after their corresponding explanation) varies between figures. Please consolidate in a single style.

Corrected

L 356 "the SV21 FST was elevated only in pairwise comparisons including Nors/Katte ... (Fig. 3B)" But this Figure does not show comparisons with Katte

Corrected by referring to the supplementary table S3, high value of Fst are highlighted in bold

L 359-360 "The genome(-)wide differentiation outside... " Ideally this would be backed-up qualitatively, I think mean and sd Fst of SV and collinear will suffice.

Done

L 360 "Several SNPs" How many? Quantitative support needed...

Done
The decrease $F_{st}$ / increase $pi$ inference really relies on a visual pattern that to me is not immediately obvious. This also rests support to the statement in L498-499. Some way of quantifying this pattern would add support to this statement.

I think this section tends to be very descriptive and need to be backed-up with more quantitative data when possible.

We agree with the reviewer, but with the genomic coverage we have, this is all we can do. The fit on the 95% upper quantile clearly shows the drop of $F_{st}$ in the center of each structural variant. Moreover, loci with $F_{st}$ of 1 tend to be localized on the breaking point of the inversion (but there are only few of these loci). We can not completely rule out at this stage that this pattern is due to the random sampling of the ddRAD approach. However, the pattern seems quite consistent across both SVs. We hope to go deeper in the analysis of this pattern in future studies with increased genomic coverage.

We removed all the parts referring to this pattern, since it does not provide any important information for the purpose of the manuscript, but it will be something that we would be interested in investigating more carefully in future work with increasing coverage.

Just a minor note, Figure 3D The quality of the image is not enough to appreciate the differences between red and purple in the triangular LD plots. Either include a higher quality figure in next version or maybe change the colour palette a little bit.

We have the PDF of the figure, all figures will be resubmitted in high quality pdf format.

L 408-410 I do not understand what the authors mean by "decoupled from the species’ geographic distribution". The “decoupled” has been rephrased to “not correlated” and we have included a Mantel test to test the correlation between the $F_{st}$ genomewide (without the SV) and the $F_{st}$ for each SV to add statistical support to this claim.

L 416 and L 459 "to our knowledge is one of the clearest" and "is the first example described " I personally consider this type of statements unnecessary, but if the authors feel are needed they definitely need at least some support from the literature. We have toned this statement down.

L 446-460 This section is very difficult to follow because it jumps between different demographic/evolutionary models without much structure or connection. Namely, "edge effects", "founder effects" and IBD. These have obvious connections to each other and to the observed pattern of genetic diversity, but I feel the discussion here does not do a good job at making these connections. I found the edge effect particularly confusing as at some point the author seems to imply that edge effect and founder effect are the same things?

We refer to “flounder effect” when we speak about the Icelandic population (which is an island population living in isolation from the rest of the continental shelf population).

Then, we refer to the “edge effect” when we speak about the Baltic Sea and the Barents Sea which are, to our knowledge, are both on the edge of the plaice distribution range.
We have rephrased the paragraph about the IBD to include the connection between edge effect and IBD.

L 462 "The two large SVs were polymorphic in most of the sites studied" I'm interested in entertaining the idea that the SV's are not in fact polymorphic but that individuals from different groups were combined in a single site. For example, there is not enough information in the Methods section to know how the sampling was conducted and where these individuals come from. Is it possible that individuals come from different micro-habitats or that there some cryptic variation in such a way that individuals could be grouped into some kind of ecotype with alternative SV's? i.e. the SV's are not really polymorphic.

This, of course, does not invalidate that SV's are polymorphic to the site level. Feel free to ignore this if it does not make sense given what is know about the ecology of the plaice.

At this stage, we cannot exclude that the structure we detect along the gradient between the North Sea and Baltic Sea also occur at more fine scale somewhere else in Europe (for example, between coastal and offshore locations, as found in the European anchovies or in the European Sea horse). However, with our current sampling design, it is difficult to evaluate this hypothesis.

L 474 - 477 This edge effect is not super obvious by only looking at the pie-charts in Fig.3. I suggest that authors to add a figure supporting this statement, e.g. correlation of frequency with latitude or something like that.

We agree with the reviewer in that our sampling design is probably not ideally suited to thoroughly evaluate this hypothesis. However, we still mention it as one hypothesis to be further evaluated in future work. In addition, the latitude does not cover the position of the Baltic Sea on the edge.

L 477-478 "the ancestral haplotypes for SV19 disappeared" This is one more example of how the lack of detail about sampling does not allow to evaluate this type of statements. E.g. I cannot see if the sample size is lower/enough in this region, potentially preventing the authors to not see the less common haplotype.

The sampling sizes were detailed in the first table of the supplementary material. We have now moved it to the main manuscript text. We have also included a column to mention the number of analyzed individuals after data filtration.

L 491-494 It would be good to give an example of what the authors refer to here. A good citation would be Faria et al. 2019; Molecular Ecology; 28: 6, 1375-1393 where they show how to find inversions in non-model organisms with the type of approaches you talk about in this sentence.

Ok, done.

L 539 - 540 I do not see where the evidence for 2.5 times more net divergence is coming from? I might just have missed it, though.

It is the ratio between the long and the short branch in the phylogenetic tree: 0.14/0.06=2.33 and 0.12/0.05=2.4. This has been clarified in the discussion.
I think the section "The evolution and maintenance of the structural variants" could be summarised a bit further given that it contains a lot of speculation. It is a nice discussion, it just feels a little bit too long.

We agree with the reviewer and therefore we have condensed and moved this interesting part of the paragraph to the paragraph “Implication of structural variants for population structure of European plaice: “ line 586-592

L 584-585 I'm not sure what the authors talk about here, the first and the second? what? SV's? This sentence is rather unclear.

We have added a full paragraph entitled “Structural variants promoting evolution at two-time frames:” that we hope will clarify our conclusions.

L 588 Why waiting the entire MS to propose the SV's are inversions? It is kind of irrelevant at this point and quite unexpected... I think the author should either not compromise and call them SV's all throughout, or to bring this on earlier and justify it.

We now bring this up to the beginning of the discussion, after the first paragraph summarizing the main results

L 589 "decoupled from geography" does not make sense as they might have a different geographic distribution, but they cannot be "decoupled" from geography.

We have changed the sentence:
The structural variants are likely two large chromosomal inversions with high variation in allele frequencies across population that leads to population structure different from the isolation-by-distance described genome-wide.

The "Conclusions" section is very little of conclusions and a lot of future directions.

We have change the title of the section to “Conclusion and perspective”
This paper presents an analysis of population structure in the European plaice, using RAD-seq data and concentrating especially on the contribution of two previously detected large structural variants. The population genetic data suggest a genomic background of isolation by distance (strong correlation between Fst and geographic distance), from which the structural variants deviate. They make the case that the structural variants are likely to be ancestral polymorphisms, possibly maintained by balancing selection.

In my opinion, the main merits of the manuscript are that it uses standard methods in a sensible way, and that it is well written, especially the Introduction and Discussion sections.

Thanks!

I don't have any major criticisms, but several minor comments and questions about framing, some details that are missing from the Methods section, and some of the Results that could be more clearly presented.

Minor comments

Title and conclusions (lines 583-590): The framing of the paper as about the "evolution at two-time frames" (by the way, shouldn't it be "at two time-frames"?) strikes me at a somewhat odd choice, since it's not that clear to me what this model entails in terms of testable predictions, and how it fits with the data. This becomes especially clear in the Conclusions section, that highlights this model to the exclusion of other results, but in very general terms. It is almost a necessity that any widespread polymorphism first is established in a population, then spreads, but what else does the model predict that fits, or doesn't fit, with the plaice data? I acknowledge that this may just be an issue of how the model is presented, or a matter of personal taste.

We added a full paragraph to further develop how the plaice inversion seems to fit to the model of evolution at two time-frames at the end of the discussion entitle Structural variants promoting evolution at two-time frames:

Throughout the paper: I would advice against using bespoke abbreviations. "SV" for "structural variant" might be acceptable (though my preference would be to spell that out too), but I see no reason to for example abbreviate "North Sea" to "Nors". This saves little space and impairs readability. I would also encourage spelling out isolation by distance, since the abbreviation "IBD" unfortunately has two meanings in population genetics (even if it should be obvious from context that you aren't talking about identity by descent).

Ok, we have only kept SV, LD, and kya/Mya for the abbreviations

Lines 38-39: The sentence says that because structural variants can harbour several genes, they may have functional consequences. This does not follow, so I guess the sentence is trying to say something else?

Our main point here was simply to say that, in comparison to mutations associated with single base pairs, SV often affect several genes, which make them likely to have functional consequences.
Lines 76-81: The idea of "evolution at two time-frames" could be fleshed out in more detail.

See earlier response

Line 98: This is the first mention of the European plaice in the introduction. I recognize that this might be a matter of personal style, but in my opinion, the study system and the question to be addressed would benefit from being introduced much earlier.

We understand the comment, but we feel that it will break the flow if we mention the plaice earlier in the introduction.

Line 114: The notion that structural variants explain most of the differentiation could be made more precise. How much, and what differentiation?

We have incorporated more precision from the previous paper where we identified these SVs: In a recent study of the genomic basis underlying the colonization of the Baltic Sea, we identified two large polymorphic SVs (Le Moan et al., 2019) carrying a strong signal of selection between the North Sea and the Baltic Sea plaices (66% of the top 0.1% outlier loci localized within two SVs).

Line 160: Does "DNA extractions were standardized" mean that samples were diluted?

Yes, we have changed the wording.

Line 182: What does it mean that a SNP is "present" in 80% of individuals? That 80% of individual had genotype calls for that SNP, or something else?

Yes, we have changed the word “present” to “genotyped”

Line 184-185: Hardy-Weinberg equilibrium filtering of SNPs is problematic even in the best cases, and this threshold (p < 0.05, presumably uncorrected) is very low.

How many variants were removed because of this, and is this really appropriate?

We used HWe filtration to remove SNPs that show a significant departure of HWe in 60% of the populations. At the end, we removed a very low number of SNPs (34 in total), likely mostly associated with potential paralogous SNPs that are polymorphic within one of the paralogous and monomorphic in the other, which can lead to strong departures from HWe.

The HWe filtering would be highly problematic if the plaice SVs were associated with important heterosis effects or unfit hybrids, but this is definitively not the case here. In addition, we identified the SVs with the same pipeline in our previous study, and we have preferred to keep procedures identical in order to facilitate comparisons of the studies.

Lines 187-191: Could you explain in detail how and why "size selection was slightly shifted"? This sentence is not clear to me. What happened, why did it happen, and how did you detect it (could the evidence be shown in a supplementary figure)?

This is a “library prep” effect. When we prepared the first set of libraries, an inset size between 350 and 450 bp was targeted, but for the last library, the size selection was between 300-400bp. With
ddRAD, this is quite problematic and we therefore also discuss the implications thoroughly in the manuscript.

Line 199: How did you choose which SNP to keep in each bin, at random? Given the strength of LD, is the 1 kbp limit enough to remove the effect of LD?

Yes, we only kept one random SNP over bins of 1kb for the population structure analyses. This clearly not enough to remove the SV signals, which is why we repeated the analyses with and without the chromosomes carrying the SVs. However, it is sufficient for the rest of the genome (at least with the ddRAD).

Here is a plot showing the LD decay with and without the SVs on 100kb with (left) and without (right) the SVs. The red line shows the 1kb.

![LD Decay Plot](image)

We have included this figure in the supplementary material.

Line 207: The Mantel test is never discussed again. Is this an omission?

The Mantel test is not specifically mentioned but the results are shown in brackets, which comprise all the correlation and p values from line 349 to line 363. We have included “(Mantel test: r = -0.56, p-value < 0.01)” in the first bracket only to clarify this.

Line 243-256 and elsewhere: Is "genomic architecture" really a good term for diversity and differentiation?

Thanks for pointing this out. We have rephrased genomic architecture to “genomic analyses of the structural variants”

Line 258: Could you give reference genome coordinates of the extracted sequences?

The coordinates are given in lines 277-281 (lines 265-269 in the revision).

Line 262: What does 80% mapping mean precisely?

We did this part again using the genome annotation directly and not through blast. Thus, we re-wrote the entire method section about the gene content of the SVs.
Lines 267-269 and 284-286: This procedure could be better explained. I also worry that the selection of random alleles might mean the creation of haplotypes that are not even present in the population. Is there a risk of getting the sequence wrong? Would statistical phasing help?

We understand this concern. However, by focusing only on the homozygote individuals, this protocol will not really affect the estimate of divergence between the two alleles of the SV, since we work on individuals that are necessarily “pure” for one or the other major allele of the SV. However, the statistical phasing would provide valuable information regarding the potential presence of substructure within each of the haplogroups of the structural variants. It will be something to dig into in future work (it is mentioned briefly in the conclusion/perspective paragraph).

Lines 295-296: Is a strict molecular clock using human mutation rate really the best possible way to date this polymorphism? Are there no mutation rate estimates from fish available?

We used the mutation rate used in the sea bass population genomic for similar analyses (c.f. Tine et al, 2014, Nature communication 5:5770 doi: 10.1038/ncomms6770). We refer to this publication in our manuscript.

Lines 307-311: The observed heterozygosity may be a useful descriptive statistic, but its importance here is never really explained. Why are these results presented here?

We have included some details in the paragraph about the plaice. However, the off-shelf population shows a higher genetic diversity than most of the population from the northern distributional range of the plaice (Hoarau et al., 2002), which could suggest that this population has a more ancient origin than what is currently thought.

Lines 318-323: and Figure 2: What, exactly, is "expected under an IBD scenario", and what is it that "disappears" in panel B compared to panel A? From what I can tell, the relationship pictured in panel B is stronger than the one in A. This part of the paragraph, with multiple patterns that are lost under various removals is very hard for me to follow. It would also help to provide some context about what this is investigating. Could it be rephrased?

We have rephrased this section and have added the Mantel test to provide statistical support to our claims.

Line 321: The p-value can’t be precisely 0, can it?

It is actually 0.000001, it was rounded... We have rephrased it as p-value<0.01.

Lines 375-377: Drawing any kind of conclusion about gene function based on 900 genes is highly questionable.

We agree and this was the reason why we tried to be cautious during the discussion of gene function. However, this question seems to divide the reviewers as reviewer 3 ask more details about the genes and their potential function to discuss the selective forces acting on the SVs. We have now performed a formal gene enrichment analysis in order to obtain some statistical support to discuss relevant functions potentially affected by SVs.
We have rephrased it, it is now in lines 434-440

The entire paragraph was re-written and shortened. We have also moved it to the beginning of the discussion.

The two SVs covered nearly half of chromosomes 19 and 21 of the Japanese flounder genome, where a strong LD was maintained over 9Mbp. These large linkage blocks are expected with chromosomal rearrangements such as inversions, duplications and translocations which can formally be distinguished by use of a linkage map or genome sequencing (e.g. Faria et al., 2019), but not with the reduced representation approach used in this study. However, our data filtration steps (filtering by heterozygosity) would have resulted in the loss of duplicated regions. Moreover, assuming a high degree of synteny between European plaice and Japanese flounder genomes, the size of the LD blocks and their central position are consistent with the presence of at least two major inversions in the genome (Kirkpatrick et al., 2010).

We have rephrased the last paragraph, but as more details are provided in the introduction, we did not want to expand too much on this again.

Data availability: Will the data be made available in a standard repository? This is not mentioned in the manuscript. Depending on journal (or actually, even if the journal does not require it), a clear data availability statement would be nice.

Thanks for pointing this out. We have included this information in the manuscript.
Evolution at two-time frames shape structural variants and population structure of European plaice
(Pleuronectes platessa)

This paper explores population structure and variation in two structural variants (SVs) in European plaice. Previous work identified these two SVs on chromosome 19 and 21 and this study further explores the variation by incorporating additional sampling locations and attempting to date the SVs by comparison with other species. The paper also investigates whether the SVs could have been introduced through introgression from another species that often hybridizes with plaice. Further, the paper also uses the new data to re-assess population structure in the species and examine demographic histories to investigate the strong divergence that has been found between Icelandic populations and other European populations.

The results of the population structure and demographic analyses reveal that the divergence of Icelandic population may be explained by the possibility of a different glacial refugium (different from other European populations). This is supported by their genetic divergence, demographic modelling, and by no clear reduction in genetic diversity in the Icelandic population. In addition, the study finds support for isolation by distance particularly when excluding SVs from the analysis. The authors claim this is the strongest IBD documented for a marine fish. Although population structure is generally weak.

Overall, I think the paper is interesting. Given the importance of SVs to population structure, I think this paper would be of interest to many readers. While the study investigates an important question, I have several concerns that the authors should address. These concerns are all detailed below; however, a few major points include:

1- Flow and organization of the manuscript: The introduction provides extensive information on structural variation, but this is not the only thing that the authors are interested in. The authors also investigate overall population structure and diversity, as well as investigate demographic histories. This is not well established in the Introduction, and (as detailed below) it is not clear from the Introduction why the authors investigate demographic history with the Icelandic population. This needs better context and requires some reorganization of the Introduction and Methods to improve flow.

Thanks for pointing this out. We have rewritten the introduction to broaden it and include all aspects of our work.

2- Methods – I think there are a few cases where parts of the methods need to be clarified. Why were loci pruned for LD if the primary point of the paper is to investigate SVs, which will have many loci in high LD? And by removing loci that are out of HWE, could this remove some loci that are of interest for complex structural variation? This is important for neutral population structure, but authors don’t attempt to remove other outliers before investigating population structure. I would also suggest that authors investigate overall population structure with neutral loci (not all loci). This also relates to the
methods that were chosen here to investigate population structure (PCA). Structuring can be explained by a few important loci. It is not clear which loci contribute to the population structure overall, as there is no information on the loadings or significance of loci on PC axes. Was it many loci or just a few loci that are under selection that separate the populations?

The loci were pruned for LD only in the analyses of the overall patterns of population structure and not when we analysed the genetic diversity along the SVs. We have now clarified this throughout the manuscript. Additionally, we have included the loading plots from the PCA in the supplementary material (Figure S4) to show the contribution of the SVs to the current population structure.

3- In addition, I am not sure about the dating of the SVs. Selection and reduced recombination on SVs can quickly lead to divergent haplotypes, I am not sure if the branch length of trees will be informative for dating the inversion. My thought would be that the earliest date that can be given to the formation of these SVs is at the timing of the split between Iceland and other populations (what is this timing, it is not clearly stated in the text?). However, if there were two refugia it is possible that the derive haplotype could have evolved only in the European refugium (southern) and then have been introduced into Iceland after the last glacial maximum (through secondary contact), so it may be even more recent than the split. I would like the authors to explain how these analyses with the SVs do not violate the assumptions. If they do, then I am not sure how informative they are. Further, the trees are compared to a collinear region of chromosome 19 to represent genome-wide average. But using this chromosome may be problematic given that the SV exists on this chromosome (only 5MBP away). I would suggest using a chromosome that does not contain SVs.

These are all relevant and fair comments. Unfortunately, there is no robust procedure to take variation in selection and recombination into account. Our main point with the dating of the SVs was mainly to investigate if the SVs are potentially were older than the age of contemporary populations (e.g. in the Baltic Sea). All analyses seem to support this hypothesis. See the answer to reviewer 3 on line 959.

4- There is a very short section on functional annotation in the SV regions. But not much detail is provided and I could not find access to Supplemental File 1 (List of genes). I would like to see a bit more discussion on the genes in the SVs. Have they been found in other SVs in other species (particularly marine fish)? Did the authors perform an enrichment analysis? With over 1800 genes in the SVs, all that the authors say is that many are associated with ion transport and other functions like sexual recognition. There are >1800 genes, so how many are associated with ion transport? This section needs more details and these details may provide some hints as to the function of the SVs. And as mentioned below, it may be most interesting to examine which genes are found within the SNPs where FST and Ho are 1.

Thanks for pointing this out. We have tried to improve this part by adding functional enrichment analyses and expanding the discussion, but we have also kept it relatively short in order to avoid unnecessary speculation, as mentioned by reviewer 2. In addition to functional enrichment analyses, we also found a few heat shock protein genes on the first SV, which are already discussed as candidate genes for local adaptation in the context of the Baltic Sea in other marine fishes (sprat, flounder and cod). We were also able to find a relatively convincing signal of immune related gene
enrichment, which could be a reason why these SVs seem to be associated with various environmental gradients.

More detailed comments are provided here below. Authors should also be careful to correct grammatical errors throughout the text. In addition, there are many areas throughout the text where authors should be more specific, instead of saying “many” or “some” or “several”, please indicate how many there actually are (see comments below for examples).

Abstract – The abstract requires more details as many parts of the study are not presented in the Abstract. The findings of the Icelandic population representing a potential different glacial refugium seem important in the text, but I don’t see it mentioned in the Abstract? Additionally, I understand that the range of dates for the SVs is 550-220 kya, but the Abstract only indicates that they evolved around 220 kya. Also, there is no mention of the tests for introgression.

We have modified the abstract to make it clearer. We decided not to highlight the specific timing here because of all the potential bias that can be introduced when aging SVs. We also decided not to highlight the introgression in the abstract, as we only rule out the flounder as a source of the origin of the SVs, but introgression may still have occurred from another species.

Line 13-15 – Indicate that this is known from previous work.

Done

Line 14 – and “shows” strong genetic differentiation...

Done

Introduction – There is a lot of information on SVs here, which provides a nice summary of a lot of the literature. However, the Introduction misses an important part of the study. The introduction focuses solely on SVs, which is not the only thing that the paper investigates. There is no mention of the Icelandic population and analyses of demographic history. The inclusion of analyses on population structure and demographic history need to be better outlined in the Introduction. It is not clear from the Introduction why the dadi analysis (demographic history) was done with Northern Europe and Iceland. This needs more background and better context in the Introduction. In the study, the authors suggest a previously unknown refugium may have existed in Iceland, but it is not clear why this question was even investigated. There should be more discussion about other aspects such as population structure and not just a focus on SVs.

We have included a broader paragraph at the beginning of the introduction and a more specific statement about Iceland in the last paragraph about the plaice.

Line 34-35 – Is this always true? If an inversion is introduced through secondary contact/hybridization, couldn’t both homozygotes be present initially? Maybe indicate “de novo” or “initially”, or reword for clarity.

They must have occurred at some point at an initial frequency of 1/2Ne, even in a different refuge

Line 40 – “SVs are likely to evolve incompatible alleles”. Why? Explain.
The lack of recombination and therefore the independent co-evolution.

Line 43-44 – I’m not sure if I understand what is meant by “become trapped in environmental gradients”. Also, physical barriers to gene flow would imply that populations are NOT fully connected? Please clarify this section.

“Trapped” is a common word used in the literature (by Barton for example).

Nevertheless, these 3 sentences were removed from the new version of the manuscript, as they were not essential.

Line 49 – Is “evolving” necessary here? Consider rewording this part.

We have changed “Structural variants are also important for evolving and maintaining locally adapted populations in the face of gene flow” to “Structural variants are also important features promoting the evolution and the maintenance locally adapted populations in the face of gene flow” that is further developed in the rest of the paragraph.

Line 54 – Consider changing to “maladapted” instead of “unadapted”

Done

Line 56 – Fix wording here.

We have removed this sentence with the reorganization of the introduction

Lines 61-63 – Not sure why allele frequency clines are relevant in this sentence? Why does having mutations in SVs make clines more likely? Not clear.

Adaptive mutations in a SV are more likely to resist swamping. We have rewritten this paragraph.

Line 70-71 – But this can only occur if genes with functional relevance are found within that region of the chromosome/SV?

Yes, we agree with reviewer 3. However, we have decided to remove this sentence (“these co-adaptation are expected to arrive continuously de-novo after the SV associates with an environmental barrier”) since it is not essential in the context of our study.

Line 106-108 – Provide references for these biological characteristics.

Done

Line 108 – What type of markers/how many? Microsatellites?

Done, it is 6 microsatellite loci

Line 115 – in European plaice specifically? Or in all four species? This was not clear. I think more information on this study in European plaice is needed here (Lines 112-118).

In the plaice only. We wrote more information about our previous study to make this paragraph clearer.
For the “larger geographical scale (line 118)” indicate where these new samples are from to provide better context.

This information is more detailed in the next paragraph in the methods section. We did not want to repeat ourselves because we think that the paper is long enough in the present state.

Line 119 – There is no mention of the Icelandic population here and the examination of demographic histories. We have included more details about Iceland in the last paragraphs:

However, the off-shelf population shows a higher genetic diversity than most of the population from the northern distributional range of the plaice (Hoarau et al., 2002), which could suggest that this population has a more ancient origin than what is currently thought.

Methods

Line 138 – Samples were collected during spawning. What do we know about the distribution of these fish outside of the spawning season? I am wondering if they remain close to these spawning grounds or are they highly migratory?

It’s difficult to answer this question for marine species like the plaice which are not studied a lot. The few existing tagging data suggest some site fidelity for spawning grounds but movement to feeding grounds up to 250 km have already been recorded. See https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2003.2473 for more details.

Line 182 - What does present mean here? Does it mean that it was genotyped in >80% individuals (missing in <20%)? Or something else?

Yes, we have changed “present” to “genotyped”

Line 185 – I wonder if removing SNPs that are out of HWE could remove potentially informative SNPs associated with complex structural variation. Can the authors comment on this?

See the answer in the general comment address to all three reviewers

Line 196 – Population structure should also be explored for neutral markers. I suggest performing analysis with all loci, neutral loci, and SV dataset(s). Or perhaps a neutral, outlier, and SV dataset. Although there are many ways to identify outliers, it might be easiest to use a method such as PCAdapt and remove loci that are outliers on the first couple of axes that explain a lot of the variation. It is difficult to know whether the patterns observed in Figure 1b are due to potential adaptive differences among populations or due to neutral structure. In a PCA differences can be driven by a few important loci. If such is the case, it would also be nice to see which part of the genome are responsible for these differences observed here (Fig. 1b). Other types of analyses for population structure may also be helpful, such as STRUCTURE (or ADMIXTURE).

We understand this comment, but we feel that including a neutral vs outlier analysis would change, and potentially confuse, the focus of our manuscript significantly. Thus, we have decided to keep the focus on the contrasts between SV and collinear regions of the genome rather than basic neutral vs outlier. Nevertheless, we have now included a PCA as well as the loading plot without the structural
variants in the supplementary material to illustrate the contribution of the SVs to the structure detected in the overall PCA in Figure 1a.

Line 199 – If you are looking for structural variants, shouldn't you not prune for LD? This may remove loci that are linked due to SV?

This was confusing in the original version. We only pruned for physical linkage for the overall population structure analyses and the PCA as these “individual-based” analyses are sensitive to it and to missing data, which can be problematic with ddRAD as missing data are not randomly distributed. See the answer to reviewer 1 line 448

Line 211-222 – It was not clear from the Introduction why this analysis is being conducted. It needs to be clear how this is related to SVs (the focus of the paper).

We have made the introduction broader and have included specific statements about Iceland in the paragraph about plaice population structure

Line 238 – How many loci had HO and FST of 1 and for which inversions? These loci may be especially important for understanding the functional differences between genotypes. The genes that these SNPs are located within may be especially informative.

This information has now been added in the results section.

Line 253-255- It is not clear to me what was done here to calculate average/smoothed FST?

We used the loess regression using ggplot on the value of the upper 5% quantile. We have clarified it in the text.

Line 280 – Why did you choose a collinear region of chromosome 19 to represent genome-wide average? Why not choose a different chromosome that does not contain any SVs, as the dynamics of collinear regions on chromosome 19 could still be influenced by the presence of the SV (which is only 5Mbp away). This does not seem like the best approach.

We understand this comment, which was also highlighted by reviewer 1. However, as we cannot detect any structure even within the same chromosome, it provide interesting additional information about the strength of barriers to gene flow, which seems to be limited to the rearranged parts of the chromosome.

Line 295 – These inversions are likely not neutral (or at least that is what this study suggests). I have often wondered how to appropriately date inversions. I don’t see how these analyses can be applied to inversions without violating many assumptions. Please explain.

See our response to line 953

Results
Line 317 – Typo (need space between ‘the’ and ‘Transition’)

Fixed
Line 333 – Provide information about timing in the text.
Line 334-342 – It seems confusing to go back to Figure 1 here after this population structure has already been discussed above (Lines 302-307). This is an area where better organization could improve the manuscript.

**The figure was moved in the result part (line 346)**

As mentioned, the demography history (above; Lines 327-333) seems out of place in the middle of this. This is where providing better context and describing all objectives in the Introduction would help with the flow of the manuscript.

We have moved this paragraph up, after the presentation of the FST between Iceland and the continental shelf, in order to justify the removal of Iceland when analyzing the IBD.

Line 336-337 – Manhattan plot or data to support this statement?

**Manhattan plot with the egeinvalue has been added in Figure S4 of the manuscript.**

Figure 3 (D1/2) – What does R2 indicate in this figure? The mean R2 for the locus with all other loci on the chromosome?

**Yes, it has been added to the text. We calculated the pairwise LD between any pair of loci along chromosomes 19 and 21.**
The genome-wide differentiation outside the SVs was lower (e.g. mean F_{ST} North Sea vs Baltic Sea=0.004, sd=0.031) than inside the SVs (mean F_{ST} SV19=0.203, sd=0.177 & mean F_{ST} SV21=0.167, sd=0.142; Figure 3b and Figure S8).

Line 360 – Be specific in the text. “Several SNPs” -> How many?

Added:
The individuals from haplogroups 1 and 3 were differentially fixed for eight and 30 SNPs within SV19 and SV21, respectively

Line 362-363 – Not clear where the reduction in FST is in the figure? Do you mean the huge gap in SV21 (where there is low LD), which likely due to differences in position along the chromosome compared to the reference genome of flounder? Or another region? I also don’t see a specific region of low FST for SV19? Moreover, I don’t see a concordant increase in π. This needs to be better labeled in the figures.

Most reviewers pointed to this part of the manuscript as a weak part. Here, it was referred to the slight drop in Fst in the middle of both SVs, with an absence of Fst of 1 between haplogroup 1 and 3 (many with Fst of 0.98 – 0.95). However, this signal could also be due to random sampling of the ddRAD method. We removed this part from the manuscript, but we would definitely be interested in looking into this pattern in future work.

Line 376-377 – There are over 1800 genes here, how did you determine that ion transport was important? Looked for over enrichment? Also I could not find access to Supplementary File 1?

We included a gene enrichment analyses.

Line 379-390 – Again, I am not sure that using chromosome 19 is the best option for the genome-wide estimates.

See general answer to all three reviewers

Lines 391-400 – How many SNPs were used here? It seems clear from the tree that this is not a case of ancient introgression (Line 398-400)

Discussion

Line 411 – derived “form”

Done

Line 412 – edge “of” the plaice distribution...
Line 417 – How much greater is this IBD relative to other marine fish? What is the magnitude of difference?

We have toned this part down.

Line 417-420 – Was the geographic distribution of previous studies similar to this one? It seems surprising that if this is the strongest IBD detected for a marine fish that all types of markers should be able to resolve this pattern.

Earlier work had a higher geographical coverage but only 6 microsatellite loci. The results from previous work is no surprise, given the high Ne of this species and the low degree of differentiation (<0.015). There are several examples of marine species that were thought to be genetically panmictic with few microsatellite, but where clear population structure has been identified thanks to the increased number of markers.

Line 433- Are there any mitochondrial data from this species that would support a different glacial refugium? Or is such a scenario observed in other marine species? This would be useful to know. In addition, a scenario of secondary contact was most likely. How does secondary contact fit into this history? I don’t see it mentioned in the Discussion here.

Previous studies on the plaice mitochondria have shown quite a lot of differences between Iceland and the continental shelf samples: (https://www.sciencedirect.com/science/article/abs/pii/S1385110103001321).

However, it is not clear from previous work if the Icelandic population shows private mitochondrial haplotypes. Consequently, we have preferred not to extend the discussion on the Icelandic sample in the current manuscript.

Line 449 – References for these time frames?

We have now included an entire paragraph about this

Line 421-460 – This section seems a bit out of place. The entire Introduction focuses primarily on structural variation, so the population structure doesn’t flow well with the general goals of the paper (as discussed above).

We have added a paragraph to the introduction in order to be make it broader.

Line 462 – I’m not sure that it was shown that the large SVs were responsible for the ‘main population differences’. They explain individual differences, but without their inclusion you see more IBD? It seems based on FST they are important for differentiating the North Sea and Baltic, but other comparisons don’t seem to show as high FST for both SVs? Based on the PCA, it is not clear what genomic regions are driving population differences. The PCA should be done without the SVs at the very least (and also with outliers removed).

The PCA without the SVs has now been included in the supplementary material. Also, we have included estimates of Fst within and outside the SVs to be more clear about our statement that the SVs were responsible for the main population differences.
There are no environmental data in this paper. I understand that the environmental break in the Baltic is well known but you need environmental data to back up this conclusion here. It would be possible to incorporate data from online databases. MARSPEC and Bio-ORACLE have salinity and temperature data that could be used here to better explore these associations. These can be accessed online and also using the ‘sdmpredictors’ package in R. This may be useful to the authors or even just to show on a map to those who are unfamiliar with this region. Further, since the fish were collected during spawning, do the environmental conditions at their capture site reflect the environmental conditions they would usually experience? Or are they migratory species? The analyses based on environmental data were performed in the other manuscript (Le Moan et al, 2019, Beyond parallel evolution: when several species colonize the same environmental gradient. BioRxve) together with 3 other species to study adaptation to the Baltic Sea environmental gradient. The allele frequencies of the structural variants were indeed correlated with the salinity gradient (together with 3 other small regions of the genome). We would prefer not to include more detailed analyses to the present manuscript as it is already relatively long.

Why not examine selection in your study? Test of selection could be performed on the genotype groups?

Cf earlier (line 751 of the reviewer comments), but mainly because the formal test for selection were done in our previous study.

Please provide examples of this phenomenon in other species/inversions? I was not aware of this pattern (which is not easy to see in the figures). In another recent study with a large inversion in a salmonid, they actually find an increase in FST in the middle of the inversion and a drop in nucleotide diversity (https://www.biorxiv.org/content/10.1101/504621v1.abstract).

We have decided to remove this part for the moment but we would be very interested in coming back to this pattern with more data to analyse.

Thanks for sharing this paper, it is very interesting. We are not sure that the pattern they observed is different from the one we described in the place. The trout inversions seem to show higher values of Fst in their breaking point. This pattern is even clearer in the trout as they have two inversions and therefore two breaking points side by side (which look like a peak of FST in the center of the LD block). The substantial drop of diversity in other breaking points of their inversions is also visible. In fact, the trout study has a better dataset to look for these rare events of recombination between the alleles of an inversion.

We have rephrased this paragraph (which is now line 484-486)

If this is true, would there not be drift occurring in these SVs? Should Iceland look different compared to the rest of the populations given the long divergence time?

It was something that we wanted to check but the shift in the insert size during the library prep made this question difficult to address. At the moment, it is briefly mentioned in the perspective.
Line 515 – “do not include effects from selection or recombination” – does this mean it is likely shorter or longer? I am not sure that you can accurately date an inversion by branch length given that inversion are much different than other parts of the genome. Selection and reduced recombination could quickly lead to very divergent haplotypes. Assuming that Hap 2 is the ancestral allele, then given that the SV19 inversion is polymorphic in Iceland and that Iceland contains Hap1 (derived) for SV21, the best date you may be able to give both inversions is only as early as the time of the split using other loci. What is the split date between Iceland and other populations (this was obtained from demographic modeling?). Demographic history also suggested secondary contact after isolation, so it could be possible that variation in the inversion was introduced into Iceland after the split and was not present before. I’d be happy to learn how these methods are appropriate and also better than just using other loci to date the split.

To answer these question in detail:

1- if there is an absence of recombination between haplotypes/alleles, it is good, because it means that aging the alleles with a basic phylogeny can work, as they behave like different species. However, recombination between the main haplotypes/alleles seems to occur (although rarely). Such events of recombination will tend to decrease the estimates of divergence because it will reduce the number of sites that are differentially fixed between the two alleles.

2- selection (positive and background) will increase the divergence in the region surrounding the mutation that are positively selected or counter selected due to the effect of selection at linked site. Nevertheless, it will not have consequences on the entire SV when they are long enough and if the alleles are frequent, because recombination still occurs in the homozygous individuals, and therefore, decreases the size of the sweeps

Thus, it is difficult to know how wrong we are because the two mechanism act in opposite directions, and the answer will depend of the balance between the two mechanisms. It should be possible to estimate the recombination rate between the divergent alleles with full genome sequencing, and then, jointly infer the divergence time and the selection forces acting on the SV under an ABC framework. This ABC analyses is something that we would like to do in the future with deeper genomic coverage. The only biases to infer selection is that the selection forces acting on the structural variant are likely variable over the time (see the Faria et al, 2019 cited in the manuscript), which can be a bit problematic to infer.

Nevertheless, in the paper, we highlighted the divergence of the ancestral allele (haplotype 2 in blue), which is likely the allele that has been the least affected by selection (i.e. it is the ancestral state that shows the highest diversity, which makes it less likely to diverge due to background selection and positive selection). This estimated time of divergence is already enough to say that it is older than the Baltic Sea, which was the main reason why we conducted included these analyses.
Then, we agree that the Icelandic population is also an interesting reference point, which we now used as additional data to confirm that the structural variants are old.

Line 522-529 – This is a really cool idea. Are there any other candidate species that could represent a possible source of introgression of this SV? Could you provide examples of some candidate for future studies?

**Maybe a species found in the western Atlantic through trans-Atlantic migration. An alternative could be from a Pacific species that can meet the European plaice somewhere in the North of Siberia. At this stage, all of these hypotheses remain extremely speculative.**

Line 55-547 – Why not test for selection in this region? Perhaps the composite likelihood ratio (CLR) method in SweeD would be appropriate here (it is more robust to variation in recombination, ascertainment, and demography than Tajima’s D). You could test it within each haplotype group (homozygotes for Hap1 and Hap2).

**We are limited by the number of markers. SweeD, Tajima’s D or other similar statistics are statistics that are calculated for sequences. It will be definitively something to do I follow-up work.**

Line 550 – I would also be interested to know if there are some life history traits that may be associated with this inversion. It seems many inversions underlie complex phenotypes (e.g., mating and migration strategies). Are there any traits that could be of interest, such as migration behaviour? Or age-at-maturity? Colour morphs? I am not sure. But it would be interesting to know about any possibilities.

**All these questions are questions we do also have 🤔. However, we just want to avoid too much speculation at this stage. Still, they certainly provides exiting perspectives for future work!**

Line 578 - Fix wording here.

**Fixed**

Line 596 – There is not really much discussion on the genes in the inversions. I think this would be interesting to explore a bit further. Were any processes over-enriched? Are any of the genes in the inversions found within inversions in other species?

**We added more analyses to answer this question. They seem enriched for immune system genes.**

Figure S6 – What does “mean FST” represent? What is the comparison that the points represent? Needs more information in caption.

**Corrected**

Figure S7 – Include something to indicate the position – or at least which end is the start of the chromosome.

**We are not sure we understand this question as the chromosomes are separated by long stretch of white.**