

Dear PCI editor and reviewers,

We are grateful to the 3 reviewers and the editor for their thorough and objective assessment of our manuscript. First of all, we deeply apologize for the long delay between the decision and the resubmission of this new version. Second, we share the view that the previous version was far too long and we have worked on a shorter and more focused version. Third, and even more broadly, thanks to the help of this first round of feedback from PCI, we have considerably reworked on this manuscript, including the (number of) figures presented in the main text, and done most suggested analyses. As a result of the redrafting, all reviewers' comments cannot be directly associated with the new version. We apologize for the inconvenience. Please find below our point-by-point response to the reviewer comments.

Sincerely,

Huiying Shang & Thibault Leroy, on behalf of all the authors

Editor (Violaine Laurens)

Dear authors,

I found that your manuscript the genomic landscapes of differentiation across *Populus* speciation continuum brings original and news insights on the genomic processes involved in population differentiation. The quality of the datasets and of the analyses carried out have impressed the three reviewers and myself. Nevertheless, the manuscript is quite long and is sometimes difficult to read. Importantly the take-home messages are not straightforward to understand, preventing my recommendation at this point.

We agree with all these points. The quantity of analyses and the complexity of the interpretation are challenges for this kind of work. We should remain cautious with our interpretations and this situation contributed to making our take-home messages less visible. In this version, we have completely reshaped the main text, which is shortened in length, and, more broadly, tried to provide a more straightforward version.

The three reviewers have listed specific points that should be addressed, and I would be happy to evaluate again an improved version of the manuscript, clarifying the results and the

main conclusions drawn from your study. The number of analyses and figures reported is very large, and you may consider sorting out the most important ones, and move the less important ones in the supplementary section.

In this new submission, the ms now includes a box and a new version of Figure 1. We also reduced the number of figures in the ms and the length of the text thanks to the help of three supplementary notes added to the Supporting Information. This also helped us to provide additional information for two main analyses suggested by the reviewers (the detection of selective sweeps and the computation of ABBA-BABA statistics).

Overall, you may especially aim at reaching a larger audience of evolutionary biologists interested in speciation, including readers lacking background in population genomics.

Thank you, we appreciate your support and hope that this new version will be more accessible to a large readership.

Sincerely yours,

Violaine Llaurens

Specific comments:

Line 33 : You may briefly define what you mean here by 'genomic landscape'.

=> Done (l. 32-22): "looking at the variation of nucleotide diversity and differentiation across the genome (hereafter genomic landscapes)"

Line 44 and 47: Again providing a hint about what 'linked selection ' and 'background selection' mean here, and how these selective forces might influence genomic divergence between species seems necessary.

=> Now the abstract only contains this sentence "we find support for signatures of selection at linked sites (hereafter linked selection) - the interaction between natural selection and genetic linkage - in shaping these genomic landscapes" (l.42-45)

Line 48: It is unclear to me how you can infer that the divergent sorting happened before speciation ?

=> We changed the sentence to "...standing genetic variation" (l. 45-56) to make it clear. In our study, we found long shared haplotypes between species, i.e., between *P. davidiana* and *P. rotundifolia* ($F_{ST} < 0.25$). Besides, we found significantly positive correlations between D_{XY} and F_{ST} , which indicates that the genomic islands of differentiation were derived from divergent sorting of ancient haplotypes or restricted gene flow. However, we noticed the correlations of these two parameters are highly positively correlated, whatever the level of gene flow. Therefore, we considered that ancient polymorphism played a role in generating the heterogeneous landscape of differentiation.

Line 49-51: this conclusion is very vague and a bit circular (studying the genomic patterns informs on genomic landscape of differentiation?).

=> Sorry, we agree, a part of the sentence was missing here. "better understand the evolutionary forces shaping the". Our last sentence of the abstract now has a wider scope (l. 46-48): "*We highlight the importance of investigating genomic patterns on multiple species across a divergence gradient and discuss prospects to better understand the evolutionary forces shaping the genomic landscapes of diversity and differentiation.*"

Line 61-63: this sentence is too vague, what can of 'overall drivers' are you referring to?

=> Agreed. This sentence has been removed.

Line 66: 'local adaptation AND/OR reproductive isolation

=> Done (l. 64).

Line 70: i.e. should be italics

=> Done.

Line 83: It is unclear what did you mean by 'genomic characteristic influenced by life history traits' ? Would selfing rate be an example of such life history traits? Because selfing in plant would decrease heterozygosity throughout the genomes? It is worth being more explicit here in my opinion.

=> The "life history traits" referred to the link between life history traits and genome evolution, in particular effective population size as shown by Romiguier et al. 2014 Nature. Selfing is another very relevant example. To keep the ms focused, we have removed this sentence.

The processes listed line 82-83 are nevertheless not independent from adaptive processes: for instance, variation of recombination can be linked to adaptation. Maybe the beginning of this sub-section should be rephrased.

=> This sentence has been removed. The beginning of the subsection has also been changed.

Lines 125-139 could be moved to a box containing Fig. 1

=> Thank you for this suggestion. The new ms version now includes a box, containing both these sentences and an edited version of Fig. 1, both contribute to increase the readability of the introduction.

Line 151: add parenthesis in the reference Burri (2017a)

=> Done (l. 90)

Line 155: remove 'This is due to the fact that'

=> Done

Line 147-162: could also be grouped together in a single Box summarizing the expectations and Fig. 1.

=> The new version includes a single Box. To improve readability, we have removed this subsection of the introduction since we have considered that it was too early to introduce the evolution of the landscapes through time. The line of reasoning is introduced at the end of the results before discussing the associated results.

Line 182: 'open questions' it would help the reader to focus on the main question addressed in the manuscript to ease the understanding, rather than providing such highly general statements.

=> Agreed. This sentence has been removed.

Line 170-182: Overall this section should be moved to the Material and method section, this would make the introduction shorter and more impactful.

=> In this new ms version, this sentence is at the beginning of the Materials and Methods section

Lines 201-208 : This could move to the material and method section as well.

=> This sentence has been moved to the Materials and Methods section

Lines 213-221: this seems a bit descriptive and could be removed. A brief mention on fig. S4 separating the most recently diverged species would be sufficient.

=> Agreed. We have decided to remove the PCA from the main text since the NJ tree and the Adxmiture analyses are sufficient and self-explanatory

Line 252: 'in A previous study (Shang et al. 2020)'

=> Done (l. 141v)

Line 287-289: replace 'significantly' by 'significant' (line 287 and line 289) and 'evident' by 'obvious'

=> Done

Line 291: 'indicates THAT gene density is...'

=> This sentence has been removed from the ms.

Line 302: 'To have knowledge of...' it is important to avoid vague sentences, and to guide readers into the different questions addressed in the sub-sections of the manuscript.

=> We hope that our ms is now more focused and straightforward.

Line 305: It is unclear why you used only 5 species while you have data on many other ones? Is it simply for showing a simplified figure? I guess that using the whole datasets would even strengthen the point you are making on the conserved pattern of differentiation?

=> Yes, we have used five representative species pairs among the 21 possible pairs. We tried to focus on pairs at different stages across the continuum of divergence. To provide this information more clearly, we now show some results for the 21 pairs, highlighting the five selected in Figures 2 (blue labels) and 4 (blue triangles).

You may consider reducing the number of figures in the main document, there is a large amount of data presented, I guess it would improve the manuscript by emphasizing more

some results and reducing the number of analyses presented. Some figures could be moved in the supplementary material, to help reader focus on the main results. Maybe figure 5 is slightly redundant with fig.4, and could be moved to supplementary for instance?

=> We have reduced the number of figures. The previous manuscript version contained 6 main figures, our current version contains 4 (+ the box). Regarding the previous Fig. 5 (now Fig. 3), the information shown is the correlations of the landscapes between two species pairs for the same summary statistics (e.g. F_{ST} for 5a, now 3a), while the previous Fig. 4 showed the genomic variations of F_{ST} for a given pair (Fig. 4a, now 2a). We introduce a part of redundancy in Figure 2C, but we think this is important to understand our overall line of reasoning.

As mentioned by one of the reviewers, the three last sections of the manuscript are quite difficult to understand and are probably the most important of the manuscript. It is hard for the reader to get the exact information gained by the different analyses, how much they are redundant or provide independent evidences of the same phenomenon.

=> We have reshaped the end of the manuscript and hope that this section is now easier to understand.

Line 414-416: this conclusion is surprising: while pointing out the limit of your analyses is necessary, highlighting that the scenario tested are naïve prevents the reader to get the general message of your study.

=> We also agree here. We now rather discuss the other factors contributing to the signal, in addition to the prime role of background selection.

Line 424: you may remove the emphasis; the dataset is impressive, there is no need claiming superiority. Replace by 'we provided a relevant and ambitious case-study'.

=> Sorry for the misunderstanding. Our initial sentence was "one of the most ambitious". We would like to clarify that this sentence was not to claim superiority, but rather to call for more large studies in terms of sampling and genomic resources. We have decided to edit this sentence accordingly.

Line 426: support THE role of linked selection

=> This sentence is no longer in the ms.

Line 430 : 'play a role' is vague and participate to the lack of clarity on the main findings of the study.

=> We have changed the sentence to "*The observed positive correlations between F_{ST} and D_{XY} in all species pairs indicate that shared ancient polymorphism must also play a very important role*" (l. 298-300).

Line 433-436: I am not sure this sentence is useful, it also provides a bit circular argument.

=> The last sentences of our conclusion have been completely changed.

Reviewer 1 (Camille Roux)

In the manuscript "Conserved genomic landscapes of differentiation across *Populus* speciation continuum" by Shang et al. the authors study the evolution of genomic patterns of polymorphism and divergence along a divergence gradient. Personally, I will speak of a divergence gradient and not of a speciation continuum, as I am not yet convinced that there can be a continuum of a process that is itself continuous.

=> We are fully in line with the reviewer 1 that this term, albeit widely used among evolutionary biologists, is misleading. Following this view, we have decided to use "divergence gradient" and "continuum of divergence" rather than the "speciation continuum".

For this, the authors analyze a still rare and very high quality dataset composed of 201 whole poplar genomes sequenced entirely from 8 different species, with an average coverage varying from 21X to 32X. This yields >30million SNPs for a 500Mb genome, and thus, allows for robust quantitative analyses. Overall, I am impressed with the data acquisition work and the underlying analyses. Despite the revolution in sequencing methods that began more than 10 years ago, biological models studied in genomics in such detail can be counted on the fingers of one hand. First, the study is an important step in advancing our knowledge of the *Populus* model, with results that are consistent with each other. It is impressive how the observed genomic patterns (π , Tajima'sD) are suggested by the SMC++ analysis, making the interpretations consistent. The rarity of introgression is however very surprising with here only two lineages connected by gene flow, while the IBD analysis could suggest more exchanges between species. This point is in my opinion little discussed by the authors and would deserve more development.

=> We should indicate here that we should be cautious on the number of lineages that are connected by gene flow. Several other species could probably exchange gene flow and more work in this system in that respect. In our previous version, this conclusion was mostly based on the treemix analyses and this result is consistent with the parapatric distributions of the species. Additional work done on ABBA-BABA tests suggest that gene flow could be more widespread among all our pairs. We have edited the text in that direction. In the future, demographic inferences of past and ongoing gene flow among species of the complex should be performed.

This study is also interesting beyond the *Populus* model as it is one of the pioneering studies describing the effects of the divergence process on genome-wide molecular patterns (Figure 6). I find these patterns important to empirically illustrate expectations and to aid future interpretations. However, it is not clear to me how such patterns will be used, especially to discriminate different demo-genomic scenarios (figure 1). Can't we imagine, in a future analysis, model comparisons to interpret these patterns more finely? [Background selection versus selective sweep versus background + sweep] x [isolation versus heterogeneous migration] and then quantify the parameters of the different forces? The authors have everything to perform such an analysis (recombination map + patterns along a divergence gradient), and a future study on this would be impactful in speciation genomics.

=> This is definitely a great idea and at this stage we can only encourage further work in that direction. Our investigations seem too empirical at this stage to directly go in this direction. A large simulation background will be probably needed in order to be more quantitative and to precisely train a detection method. After years spent on this data, we can make the following very clear statement: the patterns are very difficult to interpret! It is one main issue we are

facing since the beginning of this work, which also explains the delay between the submission of the 1st and 2nd version of the ms. We therefore agree that the proposed framework and the development of a dedicated method could be impactful in speciation genomics. Regarding the programming skills, and particularly for machine learning, such a work does not seem to be something we could manage ourselves. But we clearly support this direction and the availability of our empirical data could facilitate these future investigations.

My various comments are mainly about presentation, especially in the introduction. There is a whole literature already discussing the distinction between background selection and selective sweep to explain genomic variation in molecular diversity. It is thus surprising to disconnect interspecific analyses as performed here from what we have learned with intraspecific studies. In particular, the recent studies of Peter Keightley finely quantifying the relative contributions of these forces in genomic patterns of (intra) diversity. So I think some of the points in the introduction might consider a bit more the conclusions drawn from intraspecific diversity analyses. Still on the introduction, it seems to me to be written too much for population genomists and may seem obscure to people outside the discipline. I thus note in my remarks below various points that could benefit from a little pedagogical effort.

=> Thank you for this suggestion. We have completely reshaped the introduction and hope this version will be more intelligible for non-specialists. It should also be noted that we made a particular pedagogical effort in the abstract.

In conclusion, the authors provide a detailed picture of divergence in poplar, with a pioneering analysis of such a process along a divergence gradient. The comparison of scenarios to explain the described patterns is so far still verbal, but could be the main subject of a future statistical study. The work done so far is impressive and I can only recommend this paper as a reference for future analyses of divergence in a given clade at such a fine scale.

=> We fully agree with the reviewer's point of view. Our approach here remains empirical and the models mostly verbal. Detailed empirical studies are important to pave the way for future statistical studies and we hope that our study will contribute to move forward in this direction.

Reviewer 2 (Steven van Belleghem)

Huiying Shang and colleagues present a detailed study on the genomic divergence landscapes between pairs of *Populus* tree species. They use more than two hundred whole-genome resequenced samples from 8 species to study genetic divergence along a continuum of speciation. By correlating measures of nucleotide diversity, relative divergence, absolute divergence, and recombination rate they set out a number of hypothesis that each well define how these relationships are expected to change as divergence increases between species. Doing this, they identify the common drivers of heterogeneity in genomic

divergence to be background selection, with contributions of ancestral polymorphisms, gene flow and potentially selection.

I think the authors did an excellent job at introducing these complex concepts and outlining their hypothesis that can be neatly dissected with their data. The number of possible expected relationships between population genomic statistics is high and I was truly impressed how well the authors did at explaining their expectations and being able to follow how their results fit to them. I think the analysis and results are solid, but please allow me to give a few remarks that I hope could further increase the accessibility of this manuscript to a broad public.

First, I very much like Figure 1. Panel A shows expected correlations within or between species, whereas panel B shows how these relationships may change over time. In the results, however, results relating to panel A are presented at the end and it would have been more logical to me to have presented those results first (regarding that section, I would also suggest including the figures in the main text, rather than the supplement, because a quantification of each hypothesis is really unique! I think doing this could also improve understanding how much each of the different factors contributed to the divergence landscape.). Next, In Fig4 and Fig6 I would suggest improving the link with figure 1, by showing the results in the same order as in Fig1 and clearly pointing the reader to how they match Fig1 and what it tells about the drivers of genomic divergence.

=> We agree. Thank you for the suggestion. We also tried to simplify Figure 1, which is now included in Box 1. Box 1 now only contains the previous Fig. 1A. To avoid information overload in the introduction, we have decided to drop the information previously shown in Fig. 1B. We only introduce this expectation at the end of the results when this information is needed to interpret the results. We hope that the new version is more straightforward to understand.

Second, I think there's a missed opportunity to better quantify the extent of gene flow between species, compared to using IBD and TreeMix. Would it be possible to calculate ABBA-BABA type of statistics (e.g. `fd` https://github.com/simonhmartin/genomics_general/blob/master/ABBABABAwindows.py) to understand what fraction of the differentiation heterogeneity results from admixture? Another set of analyses could include running SweepFinder in each population to identify common and unique selective sweeps.

=> We have done `fd`-statistics to better quantify the extent of gene flow between species. Our results detected extensive gene flow between *P. davidiana* and *P. tremula*, *P. tremula* and *P. alba*, *P. tremuloides* and *P. grandidentata* (Supporting Fig. S22). Besides, SweepFinder2 was also performed to detect regions under selection (Supporting Fig. S17-S20). We also identified the common and private selective sweep. We overall reported private regions, which is consistent with the allopatric selection scenario. Thank you for the suggestions.

Third, do reads of more distant species map equally well to the reference genome. It might be good to verify this, so that no correlation biases are created because more distant species have less reads mapping, resulting into reduced diversity values.

=> Thank you for your suggestion. We indeed observed a variation of the proportion of the genome covered by reads between 80% and 92% for different species. We now provide this

information (Supporting Table S4) but the observed difference seems relatively reasonable and therefore unlikely to introduce an important bias.

Fourth, regarding the title “Conserved genomic landscapes of differentiation”, the point of conservation is a bit lost to me (after the abstract) as most of the later figures visualize how relationships between statistics change as species diverge. I might thus suggest presenting a more direct quantification of the conservation of the differentiation landscape. I do agree that the drivers of genomic differentiation may be highly conserved.

=> We have changed the title to “*Drivers of genomic landscapes of differentiation across Populus divergence gradient*”, without introducing the conservation here.

Small comments:

Be careful in the use of ‘linked selection’ and ‘background selection’. They are not the same and I think they are sometimes used as interchangeable (e.g. in the abstract L44 and L46).

=> Thank you for your suggestion. We have checked the use of “linked selection” and “background selection” in the main text.

L87-92: Also observed in *Heliconius* (and including introgression) (Martin et al. 2019 Plos Biology; Edelman et al. 2019 Science; Van Belleghem et al. 2021 Evolution)

=> Agreed. These references are now included (l. 91-92).

L124: “different scenarios can be hypothesized regarding the extent of local gene flow...”. I found this sentence confusing because the four scenarios do not just focus on gene flow, but include balancing selection, background selection and selective sweeps. This becomes clear further in the manuscript, but I would argue to modify the sentence here.

=> Agreed. In this previous version, we tried to introduce the fact that gene flow is needed to generate heterogeneous landscapes in the presence of barriers to gene flow, but we agree that it was confusing. This sentence has been edited.

L145: “...polygenetic adaptation from standing genetic variation”. Does this include patterns arising from linked selection in the common ancestor?

=> We deleted the sentence in the manuscript.

L153: “recombination (r)”

=> Done

L154: I would suggest “across the speciation continuum”

=> Following the suggestion of reviewer 1, we now use “divergence gradient” and “divergence continuum” rather than speciation continuum.

L154: “...negative correlation of F_{ST} and r become increased with advancing differentiation...”. Is this assuming no gene flow? I imagine that with gene flow, this correlation can also become increased.

=> Yes, it is true that F_{ST} heterogeneity is expected to be associated with ρ when introgression rates increase. We have deleted the sentence to clarify and reduce the length of this paragraph.

L158: “The positive correlations between p and r should remain highly correlated with each other under background selection”. Can selective sweeps also contribute here to that relationship?

=> This is true. In the current version, we also included the results of an explicit sweep detection.

L178: *Populus trichocarpa*

=> Done

L179: *P. trichocarpa*

=> Done

L226: Fig. 2c instead of Fig. 1c?

=> Yes, sorry for this typo.

L229: Can shared haplotypes in the IBD analysis also result from incomplete lineage sorting?

=> Sure. This explains why we do not discuss the results from the closest pair (*P. davidiana* and *P. rotundifolia*) regarding gene flow, since we cannot exclude that the pattern is just due to incomplete lineage sorting. We however know that these two species can exchange genes.

Fig3a: Why do introgression edges go from tree edges to tree tips and not tree edge to tree edge? (Maybe I am just not familiar with the method)

=> The precise location of the events is difficult to infer. Under treemix, the precise location of the inferred introgression events and even the direction of the arrows should be interpreted with some caution.

Fig4e-f: Would it be to visualize and add the genome-wide correlation of p and r ?

=> A significantly positive correlation between average π and ρ can be found for each species pair, but the correlation becomes weaker with time (Fig. 4f).

L285: Regarding the low p but high r in *P. grandidentata*, I noticed that this species has undergone consistent population decline. Could that have affected p independently from r ?

=> We agree, it is a likely hypothesis. But we do not want to overinterpret the results for this species since we have also observed a lower mapping rate of *P. grandidentata* to the reference genome.

L318: “correlation coefficients for...”. Correlation with what in this case? I assume with divergence level, but this took me some time to understand.

=> Here, we checked the correlations between pairs of summary statistics (including F_{ST} , D_{XY} , π , and ρ) and then we reported the evolution of these correlations over time.

L341: Is “respectively” needed here?

=> This word has been removed.

L342: Again, I would suggest “across the speciation continuum”

=> To be consistent with the comments of the reviewer 1, we have now preferred the “*the Populus continuum of divergence*” here.

L352: “...expect negative correlations between F_{st} and p or r and the relationships become stronger as d_a increases”. Would there not be a limit to this, e.g., when F_{st} reaches 1?

=> Yes, but it would take a lot of time before reaching this upper limit. In our case, we are investigating already late stage of speciation for some species pairs (up to ~5mya)

L353-354: This is a difficult sentence to me. Should it read: “Besides, the positive correlations between D_{xy} and p should be highly correlated with d_a ...”?

=> This section has been completely revised.

L359-361: “... p and D_{xy} showed significantly positive correlations while the trend became weaker as divergence increased”. Is this not expected as the proportional contribution of ancestral p to D_{xy} and the current p will decrease over time? Hence, the correlation will decrease?

=> D_{xy} and π are highly correlated under BGS, because diversity is inherited from the ancestor and can be passed down over lineage splits.

Fig6: I would suggest fixing the y-axis so that all plots are readily comparable.

=> Thank you for the suggestion. All the panels from this plot now have a y axis ranging from -1 to 1.

L387: There seems something wrong with this sentence.

=> This sentence is no longer in the ms.

L405-406: It is unclear to me how support was found for reproductive barriers.

=> Even if it was not the main goal of the paper, we have tried to estimate a proportion of the detected windows exhibiting footprints that could be consistent with reproductive barriers. Our conclusion was based on this basis. But we agree that additional work will be needed in that direction.

Reviewer 3 (anonymous)

In this manuscript, Shang and colleague study a radiation of *Populus* trees, to decipher the most likely scenario explaining speciation in this group. Specifically, by sequencing individuals from 7 species, they test 4 different speciation scenarios (divergence with gene flow, allopatric speciation, background selection and balancing selection). They conclude that, in *Populus*, allopatric speciation fits most of genetic differentiation observed in the genome and is the most likely mode of speciation in the genus. Authors also detected two introgression events in the radiation.

The study system is interesting and the authors use state of the art methodologies to answer their questions.

However, the authors explain very succinctly the two last models they test for and the associations of statistics they expect to emerge if these models are true. Specifically I did not understand model 4 (balancing selection) and how this model differed from model 1. Also, the distinction between model 3 and 4 involves more a quantitative testing than a qualitative testing, as the signs of the correlations between genetic statistics are similar but just of a different magnitude. I therefore did not understand how the author did manage to distinguish these two scenarios?

=> We apologize for the lack of clarity. We have worked on a simplified version of Figure 1 (now in a dedicated Box), in order to provide an easier to understand version. To summarize briefly, in scenario 1, we can expect high F_{ST} (caused by low π) and high D_{XY} , while in scenario 4, ancestral polymorphism is maintained between nascent species, resulting in elevated π and low F_{ST} . In scenario 3, the high F_{ST} could be caused by both low π and low D_{XY} . As a consequence, our interpretation was based on the observed genomic landscapes of F_{ST} , π and D_{XY} .

I personally felt that the manuscript is poorly written in parts, making some sections unintelligible. Indeed, I did not understand a couple of result sections (Conserved genomic landscapes across the continuum of divergence, Correlated patterns of genome-wide variation across speciation continuum, and Scenarios of genomic patterns of differentiation) because of both the writing, and the many reasoning shortcuts used in these sections.

=> The manuscript has been completely reshaped and we hope that this new version will be more intelligible for a large readership.

I therefore do not recommend publication of the manuscript in its current form.

I think the above mentioned sections of the manuscript need to be rewritten. This is more of a personal preference, but the flow of the manuscript could be changed to present the main results of the manuscript (i.e. the hypothesis testing) earlier on. In that line of thinking I would put the Between species variability in demographic trajectories section in the SI, and bring from the SI Figure S11 which I feel is central to the results of this paper.

=> We have indeed removed the figure containing the demographic trajectories in order to more rapidly go to our main points regarding the genomic landscapes. This information is now provided in a dedicated supplementary note. We also changed the figure to be more explicit regarding our line of reasoning, see for instance Fig. 2b. Even if we agree that each circular drawing is informative, it would be difficult to show this result in the main text (one figure for each of the five representative species pairs, Fig. S9-S13).

Specific comments:

lines 143-145: expand reasoning.

=> One additional layer of complexity is the nature of inheritance of traits, since conserved differentiation landscapes could also be the consequence of polygenic adaptation from ancestral standing genetic variation. But to not introduce too many layers of complexity at the beginning of the introduction, this sentence has been removed from the ms.

line 148: why is this expected?

=> The negative correlation between F_{ST} and D_{XY} is expected because F_{ST} is a relative measure of divergence and is therefore sensitive to π , which is lower than the genome-wide π in regions under background selection.

Figure 2c: There is more colour than cluster number for K=5. Why is there green colour in pade and pdav individuals?

=> Thank you for mentioning this. This issue is now fixed

line 287: gene density appears from nowhere. Integrate it in the introduction.

=> This idea has been added to the introduction.

lines 289-291: Why? expand reasoning.

=> When gene density becomes greater, the diversity becomes weaker. This is due to the increased effect of background selection in gene-rich regions, which strongly reduces the local nucleotide diversity levels.

lines 305-306: How were selected the 'representative' species pair?

=> We agree with reviewer 3 that this point was not crystal clear in the previous version. The selective pairs were selected to cover different stages of the gradient. We now provide more information about the 21 species and these 5 species (blue labels in Figure 2, blue dots in Figure 4). In Figure 2a for instance, we now provide the violin plots of F_{ST} and D_{XY} for the 21 species pairs. Thank you for the suggestion.

lines 316-317: reformulate. Fst along the genome. What does 'independant' refers to for species pairs?

=> Here the sentence "independent species pairs" refers to the fact that several pairs can carry the same species which creates a lack of independence. To be more explicit, we introduced a gray background in the Fig. 3a and 3b.

lines 317-318: I do not understand this sentence.

=> We now provide more explicit information before this sentence.

lines 318-322: I do not understand this sentence.

=> We agree that this sentence was too long and therefore difficult to understand.

lines 328-330: expand reasoning.

=> Assuming that the diversity landscapes were only shaped by the recombination rate variation along the genome, we could have expected the same level of correlation. However, we reported higher correlations for π than for ρ , suggesting that recombination rate is not the sole factor contributing to the correlation of the diversity landscapes.

line 342: the speciation continuum?

=> To be consistent with the comments from the other reviewers, we no longer use this terminology.

lines 345-346: I don't understand this sentence.

=> Background selection is due to a balance between selection against deleterious mutations (which are more likely observed in gene-rich regions) and recombination. The effects of background selection are higher in gene-rich regions leading to a reduction of the local diversity. Similarly in regions of low recombination, the footprints of background selection extend across larger regions and therefore can have a particularly pronounced impact across the window used for the computation (or even larger).