Editor (Violaine Laurens)

Dear authors,

First, I would like to apologize for the delay in the peer-review process. The three reviewers already solicited for the previous version of the manuscript have now all evaluated the changes made to your manuscript. All three reviewers and myself acknowledge the substantial effort made to shorten and clarify the manuscript. Most previously-raised issues have been successfully addressed. Nevertheless, some reviewers still have reservations and their comments need to be addressed. In particular, I agree with one out of the three reviewers about the section on the comparison of the landscapes and its putative conservation: this section is still a bit inconclusive and the conservation of the landscape across the comparison is not straightforward from the results shown in the manuscript. The effect of linked selection vs. background selection on these landscapes is also potentially an original result but is not explicit enough. Pending you successfully address these last comments, I would be able to recommend your manuscript for PCI evolutionary biology.

Sincerely yours,

Violaine Llaurens

#Reviewer 1 (Camille Roux)

Dear authors and editor,

I am reviewing the second version of this manuscript which has gained in clarity. My first review was already positive, it can only be that way. However, I still have some minor points to raise: It may seem like nitpicking (or maybe a mistake on my part), whether I agree with the authors that there is clearly some recent introgression, I don't see what makes this introgression "ongoing".Couldn't a past secondary contact at some time generate the observed pattern? Such an intermediate period of migration that is neither "ancestral" nor "current" would allow species to continue to diverge. This does not change the paper at all, but it is to discuss (or reinforce) the fact that there is actually introgression between species whose net divergence is 2.5%. My opinion on the matter being biased, I think that one cannot conclude between "migration (more or less) recent" and "ongoing migration" without an explicit test. Even if in one extreme these two scenarios overlap, there may be some

proportion of the current divergence that is due to an accumulation of mutations since the last contact.

Dear Camille Roux,

Thank you for this comment. We agree on this point and we have changed the text accordingly (l.138-143). To provide a bit more information here, we indeed used explicit ABC modeling with DILS between the two closest species (*P. davidiana* and *P. rotundifolia*) as part of the master project of Francesca Beclin (Beclin, 2021, University of Vienna). We found clear support for ongoing gene flow in this pair (higher support for the SC scenario when the genome-wide variation of N_e is taken into account). Regarding more divergent species pairs, which are the core of your remark I guess, we agree that we should be more cautious. Hybrids between *P. tremuloides* and *P. gradidentata,* two North American representatives of the section Populus*,* are described (*Populus x. smithii*) and present in areas where the two species overlap (upper Midwest of North America). Hybrids between these two species can be obtained by artificial crosses, even if they have a reduced fitness and limited interest for breeding, they exhibit intermediate phenotypes as compared to their parents. *P. alba* also naturally hybridizes with *P. tremula* in Europe and the hybrids (P. × *canescens*) are used in poplar breeding. Having said that, we agree that explicit demographic modeling on these species are needed to more precisely date and quantify the levels of gene flow during the divergence of these species.

There is a methodological point that I still don't fully understand. Box 1 on genomic landscapes of diversity, differentiation and divergence is very instructive on the effects of the proposed scenarios, deserving to be in a text book. However, the methodological application escapes me beyond the verbal explanation. Perhaps the test is explicitly described in a reference that I have not read. I can imagine that a significant/non-significant status + sense of relationship would be sufficient to classify scenarios 1, 2 and (3, 4). But I don't see how the authors distinguish scenarios 3 and 4 with a simple categorical approach.

The 3 and 4 scenarios are not the same, even if we expect the same series of correlations, at least regarding the direction of the correlations (+ / -). The local signatures on the genome are however not the same. In the recurrent selection scenario (scenario 3), selection leaves a consistently lower level of genetic diversity and a higher level of differentiation between two divergent populations. In contrast, in the balancing selection scenario (scenario 4), ancestral polymorphism is maintained between nascent species, resulting in elevated genetic diversity and low genetic differentiation. This is the way that we distinguished these two scenarios. Having said that, we agree that some differences remain subtle and this will require some new developments in the future.

Concerning the figures, I am still not convinced by the way populations are labelled: the letter p, next to the first 3 letters of the populations/species. This makes the pairs difficult to read, especially figure 2-A: *padepgra*, *palbpgra*, etc ... it must certainly speak to the *Populus* community, but it hinders the reading for outsiders.

Thank you very much for the suggestion. We have changed all the labels accordingly and now use the full species names.

In conclusion, the paper is a rigorous study of the evolutionary forces shaping intra/interspecific genomic patterns, with data still impressive. I will only ask to add in the material and method a link to the reference genome used for the mapping to avoid any surprises if a rather different version is published.

#Reviewer 2 (Steven van Belleghem)

The authors fully resolved my earlier comments and did an amazing job at improving the structure and presentation of their manuscript. I think their result are of value to a broad audience interested in adaptation, divergence and population genomics. I only have a few small further remarks.

Dear Steven van Belleghem,
Thank you for your comments and support throughout the reviewing process.

Box 1 is a very helpful way of presenting the expected patterns! I think, perhaps, it would be helpful to clarify that the identification of the patterns may be dependent on the (early) timing of the sampling relative to the time of divergence. For example, I would expect the increased Fst relative to background in scenario 2 to show up only in a certain time window of divergence. If populations have been separated too long, this signal would get lost. In scenario 3, I believe Dxy should only be reduced if selection happened before the divergence event.

We agree that this is an important point. Thank you for this suggestion. It remains difficult to add this component directly to the Box, but we indicate in the figure caption "*It is important to note that the contrast is also dependent on the divergence time (not shown here), with more distinct F_{ST} peaks above the genomic background in the least divergent pairs*". More broadly, in this new version, we worked again on the introduction in order to more explicitly and rapidly introduce our key point, which is the progression through time. Thanks to this change, we hope that our introduction has increased clarity.

Figure 2b is a brilliant addition.

As before, it is great to see the presentation of how the relationship between these statistics changes along the divergence gradient in Figure 4 and how these changes in relationships match with theoretical predictions.

Supplementary Note 3 regarding the fd analysis does not seem to be referenced or mentioned in the main text.

Thank you. All supplementary notes are now cited in the main text.

#Reviewer 3 (anonymous reviewer)

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, recurrent selection and balancing selection). They conclude that, in Populus, allopatric speciation fits most of the 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.

The authors seem to have considered reviewers comments on a previous draft but it is difficult to tell given that they did not provide a response letter of any kind.

It seems that the manuscript has been extensively rewritten and is now clearer.

Dear reviewer,

Thank you for this nice summary of our work. We are sorry to read that our previous point-by-point reply did not reach you. We can only hope that you will not experience the same issue for this new round of reviewing.

I however still have a couple of concerns with the current version of the manuscript.

First, the distinction between model 3 and 4 involves more quantitative testing than qualitative testing, as the signs of the correlation between statistics are similar but just of a different magnitude. I therefore did not understand how the authors manage to distinguish between these two scenarios looking strictly at genome wide correlations?

Thank you for this question, which is in line with a comment from the reviewer #1 (see also above). We agree that the correlations are expected to be the same, at least are expected to generate a series of correlations with the same directions. According to us, it remains difficult to use the slopes in order to distinguish between the two scenarios and to classify them empirically based on the correlation alone. We used another strategy by considering the local genomic signatures for F_{ST} and D_{XY}, which are different under these two models. This is the strategy we used in our study. We made this point explicit in the figure legend.

Second, the authors use the term 'genomic landscapes' to describe the variation of many statistics along the genome. I find it a source of confusion personally.

We agree that this term can sometimes be confusing at first sight, since it could be confused by some population geneticists with landscape genetics/landscape genomics. However, the "genomic landscapes" term is classically used in this field today, in particular on the papers that have inspired our work (e.g. Han et al. 2017; Burri et al. 2017, etc), so we have considered better to keep this terminology to remain consistent with the literature.

Third, the authors recurrently mention that these landscapes are heterogeneous, but never illustrate it or test it. I think a test could be done to verify if the windows have higher, lower, or similar levels of divergence (or any other statistics) than the rest of the genome and see if it a different distribution than what can be expected by chance (using simulation data with an allopatric model possibly). I feel it is quite important given the aim of the paper is to understand why there might be heterogeneity in these landscapes.

It is important to note that, after the first round of revisions, we added a new panel (current figure 2b) illustrating this heterogeneity in a specific example (chromosome 1 of the *P. davidiana*/*P. rotundifolia* pair). It is important to note that this example is representative of the heterogeneity, even if the amplitude of the contrast is the best (*P. davidiana*/*P. rotundifolia* is the most recent pair, see also comment from the reviewer 2 above). According to us, introducing such a test could generate a quite circular reasoning, since we indeed used the tails of the empirical distributions for F_{ST} and D_{XY} to identify the windows of interest.

Fourth, the authors mention repeatedly that these genomic landscapes are conserved across the speciation continuum based on correlations of statistics across the genome. However, the correlation coefficients that the authors find are significant but low for most statistics, implying that these landscape are somewhat similar but also somewhat different. I think the authors could do a better job at quantifying the amount of overlap across species. Also, some correlation coefficients decrease with time, indicating that they are actually not conserved.

Thank you very much for your suggestions. Yes, we agree that the correlation coefficients vary between species, so the word "conserved" should probably be used more cautiously. We tried to work in that direction. Having said that, the data are particularly massive and we can easily imagine a lot of potential new axes of progress, including of course working at the intersection of the detected genomic windows across species. But it is important to note that our manuscript is particularly long and our investigation has already been considerably extended thanks to the two rounds of reviewing (*e.g.* explicit sweep detection, introgression etc). For instance, our current SI version includes 3 supplementary notes, 24 supplementary figures and 4 supplementary tables. At this stage, we can only encourage the reanalysis of our publicly available data to take the analysis further and to answer additional research questions.

Finally, I do not grasp the novelty of the current study. The message I get is 'Heterogeneous landscapes emerge because of different evolutionary process'. I feel that a deeper comparison of the authors' results with past literature in the Discussion section would clarify this point.

Thank you for this comment. I think our objective is not to claim that our results are incredibly novel. Nonetheless, we can still consider that the number of studies investigating the evolution of the genomic landscapes through time is particularly limited in the literature. It requires having whole genome sequence data from a lot of species within the same clade, which is not that common. We have tried to make this point more explicit in the new version of the introduction. The results shown in Figure 4 are probably the ones that have the main added value, but some others are probably of interest for a large readership.

I personally felt that the manuscript lacks clarity on the points I mentioned above and therefore do not recommand publication of the manuscript in its current form.

We hope that this new version of the ms meets your expectations.

Specific comments:

lines 54-55: Too vague. Strictly speaking we understand this now: Mutation, selection, gene flow and drift.

Thank you for this suggestion, this sentence has been made clearer (l.25-l.36).

line 62: What are 'hotspots' of elevated genetic differentiation?

We have changed 'hotspots' to 'regions' to make the sentence better understood (l. 37).

lines 72-74: what does 'highly heterogeneous genomic landscapes' mean?

We have rewritten the sentence. '...*in shaping the variation of the differentiation levels across the genome and identified reproductive isolation genes in regions of high differentiation*' (l. 49-52).

lines 225-227: What else does shape it then? "*correlations of π were in general higher than those of ρ, indicating that not only recombination rate variation shapes nucleotide diversity*" Several explanations are possible for this result, but given that the two main components of diversity is N_e and μ (at equilibrium $\theta = \pi = 4N_e\mu$), the genome-wide variation in effective population sizes and/or mutation rates along the genome could probably shape this pattern.

line 240: I would argue that the genomic patterns are 'somewhat' conserved here.

Rather than indicating "the conserved genomic patterns", our sentence is now a bit more nuanced: "the relatively conserved genomic patterns" (l. 253).

Other expectations:

-Data are available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad (pay) or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.

-Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, code) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or code must be carefully described so that they can be reused.

-Details on experimental procedures are available to readers in the text or as appendices.

-Authors have no financial conflict of interest relating to the article. The article must contain a Conflict of interest disclosure paragraph before the reference section containing this sentence: The authors of this preprint declare that they have no financial conflict of interest with the content of this article. If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: Violaine LLAURENS is a recommender at PCIEvolBiol.

The raw data and the scripts are publicly available. The zenodo repository has been updated and this new version provides more information to the reader. The authors of the preprint have no conflict of interest of any kind with the content of this article, including financial of course. I have changed the sentence to be more general than just associated with the financial conflict. No authors are currently PCI recommenders.