Reply Second Review Round

We would like to thank you once again for your constructive comments, which have significantly helped us enhance the clarity of our article. Below, you will find our point-by-point responses to each comment. Given the numerous and detailed nature of some reviewers feedback, we have only provided responses where necessary. No response is given if the suggestion was directly applied following the reviewers' advice. The replies to several comments on the methods section have been addressed in a general response. Among the most important changes to the manuscript, two supplementary figures and some sentences referring to them were added following comments regarding the population size.

Additionally, we have added Tanguy Postic as a new co-author for his involvement in the development of ForCEEPS, which indirectly benefited our work, and "for his help in this round of revisions – as well as for his assistance in proofreading this article.

Thank you again for your time and expertise.

Abstract

- “The main objective and originality of our work is” → “The main objective and originality of our work are”
  We think the original sentence is correct.
- L7 adaptation is confusing for several reviewers, reviewer 2 suggests "emerging solution" instead and reviewer 1 suggest “management actions”
  We insist on keeping the term "adaptation" as it is a well defined term in the literature. That term also reference the IPCC working group II. We modified the sentence to make it more explicit and added the expression "climate change adaptation and mitigation solution".

Introduction

- L18. I find the statement that questions about tree adaptation to climate change are focused on the potential for evolutionary rescue to be overly restrictive. Evolutionary rescue refers to a specific scenario where populations persist through rapid adaptation to severe environmental changes that would otherwise lead to extinction. However, in forest trees, much of the research focuses on populations that are not at risk of extinction but instead experience reduced productivity or increased mortality. These challenges, while significant, do not necessarily threaten the persistence of the population.
  We changed the text in order to avoid the equivalence between adapting to climate change and experience evolutionary rescue. L18

- L23. “evolution and species diversity can interact in complex ways” Could the authors provide references to support this claim or include a specific example to illustrate the types of complex interactions they are referring to?
  The next sentences were precisely added in our previous revision to support this claim and provide examples of the types of complex interaction we can expect. We added "For example" to make it more obvious: "For example, in species-rich communities, certain species could be prevented from adapting to a new environment because of competitive exclusion from other pre-adapted species (Johansson, 2008; Price and Kirkpatrick, 2009; Thompson and Fronhofer, 2019). This effect is even more important for rare species, for which a reduction in population size due to competition can have particularly harmful consequences on their adaptive capacity by increasing genetic drift (DeMazancourt et al., 2008; Eldijk et al., 2020). Lastly, the selection pressures exerted by interactions between species can act in the same or in the opposite directions to those imposed by an environmental change, thus accompanying or hindering species adaptation(Osmond and Mazancourt, 2013)."
- L82: “the phenotypes of the offsprings are drawn” → “the phenotypes of the offspring are drawn” (offspring is both singular and plural)
  "Offspring", "seedling" etc. are correct for singular and plural but this convention is disappearing and offspring is also very common in the scientific literature. We changed it to "offspring" but we have kept the ‘s’ for ‘seedlings’ in the plural because this form is really dominant in the literature. L82

Methods

General reply for the method section:

We have compiled the responses to the comments of the second reviewer and our recommender relating to the RBA calculations, its relationship to population sizes and the criteria for considering evolutionary rescue in the following paragraphs:

Discussions with other evolutionists (Patrice David and Luis-Miguel Chevin, pers. comm.) have led us to conclude that effective sizes are very complex to estimate in practice. The problem is that since the probability of being a parent is proportional to the basal area of an individual, the contribution to reproduction can be very asymmetrical. In addition, there is a high turnover of mature individuals, i.e. 4% of reproducing individuals are replaced each year in a mature forest (by analogy with a Moran process: Moran, 1958), and this percentage can be much higher when mortality increases (as in the climate warming phase).

The number of reproductive individuals in a given year can be extremely low (in some simulations it can even be less than 10 in fir in the year when basal area is lowest), but it can suddenly rise once a group of young trees reaches maturity 40 years after a massive mortality episode. We therefore reproduced the median minimal RBA heatmaps but with the number of individuals *present the year of the minimal basal area (not just the reproducers). We added a specific paragraph (LXXX) with some justification to why we can use the basal area as a proxy of the population size and general state of the forest, and we refer to the new supplementary figures on this subject and their captions (Fig. S.7 and S.8). In addition we also explained in this section how we computed the median minimal and final RBA.*

Regarding evolutionary rescue, the problem we faced is that it is not really possible to characterize it in a discrete way. If the basal area (and the number of trees) of a simulation does not strictly goes down to 0 (i.e. extinction), its growth will sooner or later resume during the simulation. Even if this implies that stand densities and growth rates are very close to 0. In a model considering large-scale disturbances (fires, winds), these simulations would very likely have experienced extinctions because they were limited to very low tree densities over long periods. We therefore felt that it would be misleading to consider that these simulations had experienced an evolutionary rescue. In order to eliminate these cases, we chose the arbitrary criterion of 5% of median minimal RBA, which mainly enabled us to comment quantitatively on the lower parts (low h2 and VR) of the heatmaps.

As the previous evolutionary rescue criterion of 5% minimum basal area was neither transparent nor satisfactory, we have changed it to a criterion that is now clearly explained in the new section dedicated to this subject in the methods (L374-379). Our new criteria for characterising evolutionary are 1) that the basal area is maintained at a positive value during the whole simulation and 2) that the population growth rate between the date on which the basal area is minimum and the end of the simulation is significantly different from 0. This slope is estimated according via a linear regression between the basal area and the time with a confidence threshold of 5%. A new paragraph was added in the "data analysis and software" section to describe our new methodology and Figure 2 has been modified to take account these criteria. Analysis with the new criteria did not lead us to change our conclusions.

*Finally, we added this paragraph in the "limitation" section (L514-521) to discuss the limitation of our approach regarding the small population sizes: "It is also important to mention that during the simulation, population sizes can be extremely low, especially for the most parsimonious evolutionary rescue scenarios (see main text) (Fig. S. 8). If we had taken into account large-scale disturbances such as fires and winds, it is highly likely that these populations, constrained to small sizes for long periods of time, would have become extinct. For the sake of clarity, we do not in practice predict evolutionary rescues for the most parsimonious scenarios of h^2 and VR, this approach allows us to describe and compare the "slower" evolutionary scenarios of the heatmaps of the Fig. 2."

General comments :

- **Editor:**

  - In evolutionary ecology, the widely acknowledged metric for assessing the consequences of genetic drift, the possibility of genetic rescue of the efficacy of selection is the number of individuals (ideally the effective population size), not the Relative Basal Area (RBA). I think it is crucial to better explain why RBA is retained a the criterium evolutionary rescue in this study, why this arbitrary threshold of 5% is selected, and how RBA relates to demography. Specifically, we need to understand how many (reproductive) trees correspond to the 5% RBA criterion: depending on the distribution of tree sizes, 5% of RBA can correspond to very different population sizes.
    Please refer to the general reply for the method section above.
  - Comments regarding how regeneration is managed by the model, and how is the number of new seedlings is defined are also relevant in this perspective (see comment by Reviewer 1 at L152-175)
    We've modified this section to better explain how we modified the regeneration processes, starting by explaining why we had to do so.
  - Regarding genetic correlation among traits, I think addressing this topic is beyond the scope of this manuscript, and would require considering the pairwise relationships between traits values at the individual scale, which is not done in the manuscript (if I am correct).
    We are not certain to understand this comment, is it related to the comment of reviewer 2 on line L140-141 ? In such case, please refer to our reply to this comment.
- **Reviewer 1:**
  - I don't think the creation of a seedling and its transition to adulthood is described in the main text. Some elements from the response to Reviewer 2 L107 should be included in the main text such as: "A new tree "appears" in the simulation at year n with a height greater than 1.3 meters, and is no longer considered a seedling at the year n+1.
    We added those informations in the Regeneration and Species composition Feedback section. L174-180
- **Reviewer 2:**
  - Provide precise metrics and thresholds to determine whether a population has persisted under climate change or recovered after climate change.
  - L364. The explanation of how the median minimal RBA and the median final RBA are calculated could be significantly improved, as it remains difficult to follow. Since these metrics are central to tracking the state of each species in the simulations, their explanation and interpretation would benefit from a dedicated section in the Materials & Methods. For example, sentences such as "The median minimal Relative Basal Area (RBA) (Fig. 2) can be interpreted as the population's safety margin for persistence." and "We used the final RBA to assess the population's ability to recover its pre-warming basal area by the end of the 1000-year stabilization period in a warmed climate." would be more appropriately placed in a section specifically focused on these metrics. It could

be made clearer that median minimal RBA is used to assess the short-term responses of the populations (persistence under climate change), while the median final RBA is used to assess long-term response (recovery after climate change). This information is scattered in the first paragraph of the Results section and would benefit from being consolidated into a single, focused paragraph dedicated to these two metrics.

- It is unclear how the minimal median RBA can be interpreted as the population safety margin for persistence. Why couldn’t the population size decrease further and then recover? A justification for this interpretation is needed. This comment aligns with Reviewer 4’s concerns from the previous review regarding the use of the median minimum RBA as a metric for evolutionary rescue.
    - Reviewer 4 (1st round of review) suggested examining tree density and diameter to provide a more comprehensive view of the forest's state. While I acknowledge that the paper and figures are already dense, I believe the justification for using median minimum RBA still needs to be strengthened. To address this, the authors could consider adding graphs in the supplementary information that depict tree density and diameter as functions of the median minimum basal area. These visualizations would aid readers in understanding how patches vary across low and high densities. Furthermore, the authors might include their justification to Reviewer 4 directly in the main text, ideally supported by some references.
    - Also, for the discussion on effective population sizes at the end of the manuscript (L510), it would be helpful to know what the tree densities are at median minimal RBA and median final RBA for each species. It would provide an idea of the minimal effective population sizes that the species go through during the simulations
      For the previous comments, please refer to the general reply for the method section above.
- I believe the structure of the methods section could be improved, as some information related to the same parameters/processes is repeated in different parts. For example, information about the three evolving traits is given both in L136-147 and in the “Evolution module” section. I suggest merging this information into a single section (preferably the “Evolution module” section) to make it easier for readers to follow. Another example is the explanation of the slow growth state, which is first introduced at L124 and then explained in more detail at L240. A final example is the explanation of the regeneration process. I think the description of how the parents of the locally regenerated seedlings are chosen (L237-240) should be included alongside the regeneration information in the section about regeneration. So I recommend that the authors reconsider the structure of the methods section, as there is still room for improvement.
  In line with the editor's recommendation, we have agreed not to insist any further on

reducing the number of repetitions that accompany the description of such complex models.

Specific comments:

- L140-141. “In order to be comprehensive with our investigations” Is this part really useful? I am not sure I understand what it means.
  The sentence has been modified so as not to present these correlations as a subject related to a scientific question on which we are testing hypotheses, which is indeed not the case. L144
- L150-151. Does “changes in the average values” refer to averages across patches? At which time step are these values extracted? I think this part could be clearer.
  When a single value is represented on a graph (a point for example), it is the average trait value of the population as stated in the text. When an aggregated estimate is represented or discussed, we specify it such as the median in Fig. 5. Please refer to the new "data analysis and software section".
- L152-175. I think the section on regeneration could still be clearer. If I understand correctly, the authors chose to adopt an approach previously used by other models that account for evolutionary processes (the first approach described in the paragraph). However, how is the number of new seedlings defined? And how is the mortality filter applied?
  - “We incorporated a feedback mechanism from the current composition of the forest (based on the relative abundance of species in relation to their BA across all patches) on the identity of colonizing seedlings each year." Does this essentially mean that the proportion of each species in the seedling pool was determined based on the proportion of the basal area of each species in the plot?
    Yes, exactly, the previous version of the text was indeed confusing. We changed it to make this point more explicit. L161-167
  - “Among the previous examples of models that account for evolutionary processes, most have adopted the first approach. However, the second approach renders regeneration independent of the current state of the simulated forest.” These two sentences are a bit confusing I think. I would suggest (if I have understood correctly): “Most models that account for evolutionary processes have adopted the first approach because the second approach renders regeneration independent of the current state of the simulated forest and therefore does not allow evolution to occur.”
    The proposition does not reflect what we meant. To be exact, the second approach does not allow interspecific competition, to play a role during the regeneration process, which is a problem for us because our objective is to study the evolution of species in a multi-specific context. Our previous formulation was indeed unclear, we tried to improve it in the regeneration and species composition feedback section. L149-159
- L183-191. Would it be possible to briefly explain how these traits are obtained and the range of expected values? For example, what would be the expected values for a drought-

tolerant species compared to a drought-intolerant one? Including the equation in the supplementary information or referencing relevant papers could help. This clarification is important because, as presented in Table 1, it is not possible to determine whether species are drought-tolerant, shade-tolerant, or exhibit high growth values without an idea of the range for these traits.
We amended the text to give more informations about the range of these parameters (i.e. mentioning minimum and maximum possible values), and better explained their actual meaning. More informations could also be found in Morin et al. 2021 (see its supporting information). L190-201

- L185. The authors explained in their response to the reviewers: "A new tree "appears" in the simulation at year n with a height greater than 1.3 meters, and is no longer considered a seedling at the year n+1. However, a tree can only reproduce once it reaches its species-specific age of maturity L185." If I understand correctly, seedlings are simulated in each patch at a given year n based on local (and potentially migration) regeneration. These seedlings are directly assigned a height greater than 1.3m and, by the following year, are no longer considered seedlings. Am I correct?
    - I think this information should be included in the main text.
      This information is now included in the regeneration and species composition feedback section. L175
    - Then, it follows that seedling parameters are applied only for one year per individual. What is the justification for this? Is this step primarily aimed at removing seedlings that are not light-tolerant enough?
      We've deleted a sentence that could lead to confusion and added a sentence that makes it clear that the seedling parameters relate only to the regeneration process. The parameters for seedlings are not the same as those for adults, but with different values for younger trees. L176-180
- L202. In which cases the initial state of a population is derived from previous simulations?
  See the response to the comment on L313.
- L212. I find the definition of VR unclear. When defining a fraction, it has to be specified what is being divided by what. In this case, I suggest clarifying that VR represents the fraction of intraspecific trait variability relative to interspecific trait variability observed in the species pools parameterized in ForCEEPS. Relative to this, VR is referred to as "Trait variability" in some parts of the manuscript (e.g., L530 and in the figures), which I believe is not accurate.
  We have clarified that the VR is the intraspecific standard deviation divided by the interspecific variation in the text. Regarding the comment on "trait variability", we refer to our response to the comment on lines L220-221.
- L218. I would suggest mentioning that h^2 and VR vary between 0.1 and 0.4 depending on the simulations, and providing some references to support the biological realism of these values.
  This information is given in the "simulation plan section" (L323), which we feel is more

appropriate. We have nevertheless added references to support the fact that the range of h^2 *and* VR *we are exploring covers "realistic" values. L329*

- l.218-219: If VR is the same for all species in an evolutionary scenario, how can it be that in an assisted migration scenario with a total of 10 species, VR can be greater than 0.1
  We improved the description of VR and how it is calculated please refer to other comments.
- L220-221. ΣE and ΣP remain fixed across the simulations if I understand well. But what about ΣG? If ΣG changes, how? If ΣG changes, how does it change? And if it does not change, then h^2 cannot change either, as $h^2 = \Sigma G / \Sigma P$. I think it would be helpful to specify more clearly which parameters change across the simulations and, if so, how they change. I find it very hard to understand.
  - This comment follows the one of Reviewer 3 ("As h^2 and sigma_E are fixed, how do you control for sigma_G after local and migration regeneration?"), to which the authors' response is incomplete in my opinion. I could not understand from their reply how ΣG and h^2 vary throughout the simulation.
    This is a common point of misunderstanding about quantitative genetic models. I am often confronted with it at conferences, and other evolutionist colleagues working on QTL models encounter the same difficulties in explaining this problem. Heritability and genetic variance are initially determined by the modeler, but they cannot be properly "fixed". I tried to improve the explanation of this issue at the end the "Evolution phase" section I hope it clarifies our methodology because I cannot think of a better way to explain it. (L274-281)
- L226. Does this mean that the trait correlation matrix is assumed to be the same at both the intraspecific and interspecific levels? If so, I think this should be stated more clearly
  In the sentence "Trait correlation values were derived from interspecific correlations of proxy traits for $DrTol$, $ShTol$, *and* g *found in literature and databases" two lines above, we changed the word "derived" for "obtained" which express a more direct link. (L230)*
- L280-285. This section is somewhat confusing because the main text states that patches evolve over 2100 years, but in the figures, the timeline starts at year 1000 and ends at year 3000. Could this discrepancy be due to baseline simulations conducted prior to the evolutionary scenarios? If so, this should be explicitly clarified in the text.
  We cannot find in the text where we would be suggesting that populations evolve over 2100 years. They evolve over the entire duration of the simulation, i.e. from the year 1000 to the year 3000 (2000 years). By convention, we have started counting years from 1000, because the evolution simulations were preceded by 1000 years of baseline simulation (to generate forest inventories for the initial state), and moreover, this allows the warming period to coincide with the years 2000 to 2100. We did not expect this convention to cause so much confusion. The final version of the text now clarifies these points.
  - For example, at L358: " we computed the difference between the DrTolsp3000 and the average DrTolspy over a reference historical period (year 1500 to 2000)." Does it mean that DrTolsp3000 is the final average trait value of the simulation? Additionally, isn't the

  historical period supposed to be from 1900 to 2000 instead of 1500 to 2000?
  We agree that this section was very confusing, First, the reference period chosen spans from 1500 to 2000, with the simulation years 1000 to 1500 removed to allow time for local adaptation. This information should have been in the main text. Then, there was a mistake: the "historical period" is a repetition of climate data from 1969 (the first year of the CHELSA dataset) to 2000. Finally, the year 3000 is indeed the end of the simulation, which should be clearer now. L395

- L309. I do not understand the goal of using different VR values in the simulations. What hypothesis is being tested with this approach? Would it not be possible to use empirical values for VR? Is it particularly challenging to find values of intraspecific variation for the traits considered? I find this part unclear and think it would benefit from more explanation
  As stated in the introduction (L100), our first objective is to "Identify the evolutionary rescue conditions of beech and fir [...]". This is why we are exploring these different conditions for h^2 *and* VR*. The choice of exploring values between 0.1 and 0.4 for these two parameters is detailed in the "simulation plan" section (L325). We hope that the new wording of this section will make it more obvious.*
- L313. I understand that baseline simulations were run first to obtain the initial forest (shown in Figure S.6.). However, I do not understand what the starting points (i.e., the initial forests) are for questions 2 and 3.
  As mentioned in the first sentence of the "simulation plan" section: "In order to obtain mature, age-structured forests in a pseudo-equilibrium state to serve as the initial state for all our later simulations, we carried out 50 replicates of what we call "baseline simulations". This phrasing does not exclude the question 2 and 3 of our simulation plan.
- L314. Which parameters are used to assess persistence and recovery? Which parameters are tracked to evaluate the "evolutionary trajectories of traits"? Are the trait values themselves used? I would suggest referring to the specific metrics rather than their interpretation.
  Please refer to the grouped reply above.
- L330. Same comment as in my previous review: I believe more context and justification for the choice of these eight introduced species is needed here. In particular, it seems that they were chosen in part due to their higher drought tolerance, it has to be explicitly mentioned.
  The reasons why we chose these species are now mentioned in the text (L356), supported by the new legend in Table 1.
    - Moreover, the process of regeneration and migration remains unclear to me. For example, if M=0.1, does this mean that 100% of the local regeneration occurs (but how many seedlings are generated per patch? This is still unclear)? Then, in addition to the locally regenerated seedlings, the number of seedlings from migration (i.e., from other species) corresponds to 10% of the locally regenerated seedlings, correct? And how is the proportion of seedlings from the different introduced species determined.
      The new description of the regeneration process should clarify that point. L169-174

- L348. This is the first time that populations are mentioned in the methods, what is a population in the simulations?
  As classically done in ecology, a population is defined as a group of individuals capable of reproducing among themselves.
- L. 351-354: could you explain how age at maturation affects the evolutionary rate
  We have added a sentence to better explain the role of the AM_s *parameter at the end of the section describing regeneration (L178). We also corrected the typo in the equation of the Haldane's rates.*
- Eqn 3: in the sum, is the first term multiplied or divided by AM_s? if I am correct, the first term is usually multiplied by 1/nb.gen with nb.gen the number of generations and nb.gen=1/AM_s
  We did indeed copy this equation incorrectly into our latex document, and we thank the reviewers for pointing this out to us. We would like to reassure them that the calculation was correct in the data analysis scripts.
- L353. “depending on its age at maturation (AMs)” To what does “its” refer? And what is the age of maturation? In Bone and Farres (2001), the evolutionary rates are calculated based on the number of generations instead of the age at maturation, so why not using the number of generations?
  Following the corrected equation (7), we are indeed dividing the standardized trait change by the number of generations ($(y2 - y2)/AM_s$*).*
- L354. “Exact H0 values are provided along with their 95% confidence intervals for all replicates for the h2 and VR = 0.3 evolutionary scenarios as this setting led to the most realistic evolutionary rates (Bone and Farres, 2001).” What do the authors mean by “Exact”? Is this term necessary?
  "Exact" was removed.
    - Then, I do not understand why the scenario with h2 and VR = 0.3 gives more realistic evolutionary rates - is this conclusion based on the results of the present study, or does it stem from prior findings, such as those in Bone and Farres (2001)? If it is the latter, it would be helpful to briefly explain why this is the case. And then, does this imply that the other evolutionary scenarios examined in the study are unrealistic? I am a bit confused here. Why not write Equation 3 right after the text describing the equation ? (i.e., L352-353).
      We added "according to field study and experimentation" (L391) to the sentence to clarify that point.
- L356-358 : The text lines 356-358 leaves some doubt as to whether the same formula was used for both short-term and long-term changes.
  The sentence was modified to : "Haldanes rate measures the "short-term" evolutionary responses of populations. To quantify the "long-term" evolution of the $DrTol$ *trait in species* s *and in population* p *i.e.* $\Delta DrTol_{s,p}$*, we computed the difference between the* $\overline{DrTol_{s,p,3000}}$ *and the average* $\overline{DrTol_{s,p,y}}$ *over a reference historical period." L391*

- L356. A naive question: why can’t Haldanes be used to assess the long-term evolutionary response?
  As I understand it, because I haven't found any sources describing exactly the justifications and hypothesis behind the computations of Haldane's rates (the original book in which Haldane proposed this unit of measurement is not available). Haldane's calculation of the rate of evolution is in fact, a slope, so it makes the implicit assumption of a linear trend in the evolution of the trait, which is valid as long as the force of selection is constant (during the warming phase). However, once the stabilisation phase is reached and the force of selection weakens, the genetic drift becomes more important and the evolutionary trajectories no longer have any reason to be linear, especially as we are expecting an inflection in the trajectory of the trait which will produce - if we represent the value of the average trait on a graph as a function of time - a logistic curve and not a straight line. So the linear selection phase, the progressive decline phase and the static phase cannot all be described using one slope calculation.

Results & Discussion

General comments:

- **Reviewer 2:**
  - How can the final RBA exceed the RBA of the reference period ? I would expect that trees are optimally adapted to their environment, with maximal basal area during the reference period. This outcome is understandable when other species are introduced or in scenarios with gene flow, but it is unclear how this could occur, for example, in a monospecific patch of fir.
    As a general rule in evolutionary biology, organisms are not expected to be perfectly ‘optimally adapted’ to their environment. On the contrary, in our framework, where the initial parameters of the species were derived from calibrations carried out throughout the whole of Europe, we expect these trait values to be initially maladapted to the climates of south-eastern France (see first sentences of the "limitation" section). Finally, we have not capped the value for the g_s *parameter, but only a negative correlation with the* $DrTol$ *parameter, which can only slow down its evolution. The parameter* g_s *can theoretically evolve towards infinitely positive values. This is not a problem given the timescales we are working in, but it could be problematic in the longer term.*

Specific comments:

- L366. I do not see how variation in outputs due to variation in VR provides insight into evolutionary processes. Higher VR does not necessarily indicate higher genetic variation, right? If h2 is fixed, then higher VR would lead to higher environmental variation, correct? Therefore, I am unsure how this can be interpreted as shedding light on evolutionary

processes. Could the authors clarify this? Additionally, for h2, I believe it would be more accurate to discuss the relevance of accounting for "genetic parameters" rather than "evolutionary processes," as h2 does not describe a process, but rather a genetic quantity.
"Higher VR does not necessarily indicate higher genetic variation, right?", Yes it does, please refer to equation (2) and (3). It has been clarified in the "initialisation phase" section. L221

- L368-369. "cases with lower h2 and VR values" Which values are considered high and which are considered low? Please be more precise.
  Lower refer to lower value of h^2 *and* VR *i.e. 0.1 is lower than 0.2.*
- L407. "In beech-fir mixtures, evolution favored the dominant species" is confusing in my opinion; I initially understood that beech was favored simply because it was dominant. I would suggest rephrasing it with something like: "Beech better recovered than fir in the beech-fir mixtures [...]".
  We believe that the original sentences (L448-453) clarify this point.
  - Maybe I missed the information, but is there a comparison between the initial proportion of beech and fir versus the final proportion? Did the proportion change over the course of the simulations? Figure 2C shows this, but it would be useful to provide the initial and final proportions. Additionally, it seems that in Figure 2C, the proportion of fir is slowly increasing after the warm period, and the simulation stops before any stabilization occurs. Could fir have recovered its initial proportion if the simulation had continued?
    It is possible indeed, but we had to stop the simulations at some point. We have run some simulations in a longer term (+500 yr), and it appears that this process is very context dependent. Therefore we prefer not to draw too quick conclusions, but it is right that this questions could deserve to be tackled in a specific study.
- L413. Looking at Figure 5B, it seems that both species evolved towards greater drought tolerance from the beginning of the simulations, even before the warm period. It appears that beech benefited from being more drought-tolerant at the start, as its DrTol values remain higher than fir's over the course of the simulation. Perhaps the authors could elaborate on this in the discussion, as this seems to be an interesting point.
  We thank the reviewer for pointing this out. In fact, this point is already mentioned in the supporting information (table S.3 caption). Yet, elaborating on this point seems too speculative in our opinion, and we feel that there is no room for deepening this point, especially as this would probably require specific simulations. As the article is already complex enough, we prefer not to go that way.
- L426. But recovery was generally higher for higher h2 and VR (even though the maximal final RBA was not for VR = h2 = 0.4).
  We did not claim the opposite, this is why we used the word "fastest" not "faster" to refer to the VR = h2 = 0.4 scenario.
- L427-428. Can we really say that values were lower in the assisted gene flow treatment? It seems that their g mean value is very close to the one of the beech-fir mixture treatment,

  and that quantile intervals are very wide.
  This is exact, we forgot to removed g *in this sentence. L466*

- L432. Could the authors elaborate further on this interpretation? Does this imply that the gene flow from the beech population of the Montagne Noire was better pre-adapted to a warmer climate compared to the fir population? Would it be possible to compare the DrTol values of the fir and beech populations from the Montagne Noire to support this point?
  The reason why we ran simulations in the Montagne Noire was precisely to obtain genetics values from population better adapted to warmer climate. We described this treatment as follows in the "simulation plan" section : "This latter was identical to the beech-fir mixtures treatment, with the difference that we added immigration of pre-adapted beech and fir individuals during the 100-year warming period." Following the first round of review, we also added a table with the trait values of the imported individuals from the Montagne Noire (Table S.4.) to show their level of preadaptation.
- L435. The effect of trade-offs among traits on the final RBA of each species does not seem entirely convincing. Could the authors provide more robust arguments, supported by quantitative parameters (e.g., is there really a meaningful difference in g values between the beech-fir mixture and assisted gene flow treatments?), to support their claim that these trade-offs significantly impact the final RBA?
  This is indeed not true regarding the g *parameter. However, we think that the rest of the demonstration is convincing enough regarding the* $ShTol$ *and* $DrTol$ *coevolution (see L473). We modified that sentence to restrain our interpretation to only these two traits.*
- L442-446. So I am not sure to understand; did the proportion of basal area of fir vs beech change between the reference period and the end of the simulation? Maybe it would be useful in this treatment to compare the absolute basal area of the two species before and after the warming period, as the relative basal area does not seem very useful for the interpretation in this treatment.
  In the h2 and VR = 0.4 scenario, if fir has only lost 43% of its historical basal area over the course of the simulation and beech 86% (Fig. 2.D), the proportion of fir to beech basal area has necessarily changed in favor of fir yes. We informed the reader's of the problem that may be posed by examining this situation in relative terms in the main text, and figure 2.C shows accurately what is happening in absolute terms. We do not think that this question deserves a new figure especially as our article might already contains too many figures and sub-figures according to the standards of many journals.
- l.447: “beech is expected to grow more slowly due to its lower gs”: is this statement based on initial values presented in Table 1?
  Yes and this asymmetry remains true along the course of the simulations (Fig. 5.B.). We added the necessary references to justify this claim. L486
- L478-479. This is not clear to me why the authors consider that the “choice of stabilizing the climate after 100 years of warming [...] limits conclusions on any long-term effects.”. I don't believe this is the case, as this study effectively assesses the long-term effects of a warming

climate period on the proportion of each species.
We agree with the reviewer, and we think that our study brings relevant elements towards possible future of these forests. Yet, our arbitrary choice (and probably unrealistic) in the climate dynamics after 2100 led us mention this limitation to our work. Several researchers to whom we presented our work felt that this was a major limitation due to the fact that the climate has no particular reason to stabilize in 2100, and that my conclusions might have been very different if I had tested different 'post-warming' climatic scenarios such as cooling, continued warming, etc. In addition, there is the issue of evolutionary lag, which is detailed in the following sentences: as the climate continues to warm, recovery is impossible because the evolution of traits lags behind the values of the traits that would allow tree populations to recover. In our opinion, it is very important to mention this point because it has direct consequences for the way in which we perceive the threat of climate change. This means that, no matter how much the climate warms, it is always important to halt the trend to allow species to adapt to the new conditions. We leave it to the recommander to decide whether or not to delete this part of the text from the limitation section.

- L498-499. I may be mistaken, but such opposing selective forces seem to accurately capture what happens in real forests, which I believe is a strength of the study.
  We agree with this point, the only problem is that only opposing selective forces have been implemented, see the part of the discussion concerning competition for water. We added the word 'purely' L546 to the sentence to clarify it. In addition, we know that there is a trade-off between growth traits and drought stress at the interspecific level, but we don't yet know whether such a trade-off exists at the intraspecific level. Some researchers claim that we have no good reason to expect them at present. The decision to implement them or not has actually been debated with our collegues.

Perspectives

- L513. "Another reason to improve the integration of migration into models is to explore its demographic impact on populations. In our case, we have deliberately limited these impacts to highlight the genetic consequences of the arrival of pre-adapted individuals." Why not recording also the number of individuals over the course of the simulations as suggested by Reviewer #4? How did the authors limit the demography? I do not follow the rationale of those sentences.
  This point was added as a response to a comment from the first round of reviews. The sentence was changed to be more explicit: "In our case, we have deliberately limited these impacts by keeping the migration parameter low (M = 0.1) to highlight only the genetic consequences of the arrival of pre-adapted individuals. Nevertheless, assisted gene flows are also advocated for the demographic support they offer (\cite{aitkenAssistedGeneFlow2013}), ideally, higher migration rates should be tested."

- L517-518. Since the regeneration process remains highly unclear to me, I do not understand what proportion of seedlings from other species is added in each patch and across the entire forest. Are these numbers realistic?
  We hope the explanation of the regeneration process is now clearer thanks to the improvement we made in this version of the text. Regarding the realism of the migration rate, we can't estimate what "realistic" would means for a human intervention. As a reminder, we're simulating either assisted gene flow or assisted migration, and the importance of this flow or migration rate is left to the choice of the forest manager. It's quite conceivable to try to completely replace local regeneration by migration in this framework.
- L521. "We therefore think it is urgent to explore the case where intra- and interspecific interactions are at the origin of selection pressures that accompany environmental change." I'm not sure I fully understand. Do the authors mean that it is important to better understand the interactions between selection pressures from the environment and competition from other trees?
  The text has been modified to better detail our expectations. L570-578
- l.524-526: I think it should be made clearer that in a scenario with water competition, climate and competition would select the trait in the same direction (in contrast to the current scenario)
  Same response.
- l.534-536: "without any apparent drawbacks emerging in our framework". Is this statement correct? I understood that under "assisted gene flow" the genetic correlation and competition played a more decisive role than increasing genetic variation in DrTol.
  As you can see on Fig. 2.D. recoveries in assisted gene flow treatments are consistently greater than or equal to their counterparts without assisted gene flow in all evolutionary scenarios.

Tables

- **Table 1**
  - Add the range of the the key parameters and their direction regarding tolerance (g DrTol and ShTol)
    This was improved in the related method section
  - Table 1. Full names and units of all parameters should be provided; otherwise, the table is not useful as it cannot be understood. Since there are many parameters, perhaps only those of particular interest for this study could be included here, with the others provided in the Supplementary Information.
    The table was modified accordingly.

Figures

- **Fig 1:**

  - “The initial state of any evolution simulation is a forest inventory (see methods).” Maybe mention here that the initial state is based on real forest inventory data.
    In fact, it's the opposite: we've clarified what we mean by “forest inventory”. For us, the word “inventory” applies to the actual forest inventory or to inventories generated from previous simulations, but it's not very clear. It was used in contrast to starting a simulation from virgin ground.
- **Fig 2:**
  - Is the median calculated across$Med(Min(BA1500-3000,i)/Mean(BA1500-2000,i)$ *100) replicates? If so, it has to be clearer in my opinion.*
    This calculation is now detailed in a dedicated paragraph in the "data analysis and software" section (equation (3)).*
  - Then, for figure 2B, it is written: “numerical values of median relative basal area are indicated for maximum and for minimum value above 5%”. But I am unclear about what these numbers represent. Aren’t they the minimum and maximum values for the median minimal RBA in each treatment? If so, why are the numbers in the third row (whole forest) not approximately equal to the sum of the numbers in the first and second rows (species-specific BA)? Please clarify all of this.
    Please refer to the general reply for the method section.
- **Fig 3:**
  - This figure is much improved compared to the previous version! “Median across 25 to 50 replicates of the mean relative light availability perceived by individuals of beech and fir (model output) across replicates”. This is unclear. Are the median and quantiles across replicates being represented? If so, what exactly is meant by "mean relative light availability"? Is the mean calculated over the patches, or over some other quantity?
    The expression "across replicates" was repeated twice it was a mistakes. It is as indicated, the median across replicates of the mean relative light availability perceived by individual. It is a median across replicates of a mean over individuals.
- **Fig 5:**
  - Figure 5B. Why looking at negative values for ShTol
    We've made this choice to make graphical reading more intuitive and to facilitate the parallel with the evolution of g below, but it's a rather arbitrary choice we admit.

Supporting Informations

- How do the authors interpret the fluctuations in basal area (BA) observed in the scenario with +3°C?
  This is an expected behaviour for this type of model and for real-world forests, we don't think it's necessary to comment further on this technical point in the main text or in the supporting information. Mortality often occurs in episodes when several stress factors coincide. In this case, we are certainly seeing the consequences of consecutive particularly hot and dry years

with increased forest cover, which prevents individuals from growing due to a combined stress of water and shade.

- The legend in Figure S7 is also unclear, where it states “mean individual relative light availability over time.” This doesn’t help clarify how the mean is calculated. Please clarify the legends of those two figures (Fig 3 and Fig S7).
  The legend was rephrased to: "Median across replicates (n = 50 for monospecific and beech-fir mixtures; n = 25 for assisted gene flow and assisted migration treatments) of the individual light availability (%) (i.e. averaged across all individual per species) over time for beech and fir populations under all evolutionary scenarios. The line stops when every populations goes extinct in a given evolutionary scenario."'
- Fig. S.7. Why does the blue line corresponding to the assisted migration treatment stop before the end of the simulation in some scenarios?
  The line stops when every populations goes extinct in every simulation for this evolutionary scenario. This information is now added to the figure's caption.

References

Moran, P. a. P. (1958). Random processes in genetics. *Mathematical Proceedings of the Cambridge Philosophical Society*, 54, 60–71.