To: Peer Community in Evolutionary Biology editor and reviewers

Dear PCI Editors,

Thank you for your appreciation of our submission. We submit here a revised version. We apologize for the time we took to answer, which is partly due to the first author starting a new job outside academic research.

This is also the reason why, although we took great care of all the comments we received in this first round, which yielded major re-writings, additional analyses and that greatly improved the article, we could not do all the extra work asked for. We did everything necessary to make the article correct and clear, but we could not explore new dimensions such as the identifiability study requested by one of the reviewers.
!!!!Vincent: replace by: We took great care of all the comments we received in this first round, which yielded major re-writings, additional analyses and that greatly improved the article. However, for a lack of time and workforce, we could not explore new dimensions such as the identifiability study suggested by one of the reviewers.!!!!

Our major revisions are:

- We completely rewrote section 2 in order to clarify what is computed and how in 2-level reconciliation
- We added a description of the 3-level inference in algorithmic style to clarify all the steps and refer to them in the text and complexity issues

Besides this, we achieved a number of minor revisions detailed below, along with answers to reviewers.

We hope that this work can be recommended with its current boundaries and limits, that we acknowledge but are unable to largely expand by now.

Below we report detailed answers to the editor and reviewers, along with descriptions of the changes in the article, compared with the first version.

Respectfully,

The authors : Hugo Menet, Alexia Nguyen Trung, Vincent Daubin, Eric Tannier

---

Round #1

by Emmanuelle Jousselin, 08 Nov 2022 15:35
Decision concerning your submission

Dear authors,

Your preprint has now been reviewed by two experts, and I have also reviewed it myself. You will see that both reviewers are very positive about your manuscript and suggest that implementing a model of evolution of

three nested levels (i.e. accounting for the coevolution of hosts, symbionts and their genes) represents a very useful contribution to the field of reconciliation.
They also acknowledged that the model is well implemented, its presentation is thorough and the paper clearly presented.
Being myself a user of reconciliation methods (and not a developer), I found the paper very well written and found the figures very informative and able to deliver the principle of the methods. I also think that such methods are needed to address reconciliation in three-way associations.
The two reviewers have nevertheless several suggestions to improve the study (see reviews). Important ones (which might require some significant inputs) concern:
-a clarification/confusion throughout the text between most likely reconciliation and maximum likelihood;

Anwser: We did our best to clarify this in the text. There are two possible versions of the dynamic programming algorithm of the 2-level reconciliation. One indeed computes, as the reviewer writes it, in a forward manner, the likelihood of the most likely scenario, and in a backward manner, the most likely scenario itself. The other one, which is the one that we implemented and expanded, computes, in a forward manner, the likelihood of the model (summing over all possible scenarios), and in a backward manner, outputs a sample of scenarios according to their likelihood. The most usual implementation is the second one, because it allows to compute the likelihood and obtain a sample of reconciliations. We adopted this one in the 3-level model. In the text we completely rewrote section 2, adding dynamic programming equations, in order to make that clear for readers.

-a test of the method using data simulated under the model developed;

Anwser: We unfortunately did not achieve this recommendation. We acknowledge that there are identifiability issues in the model, and that it would be interesting to address them by simulating under the model itself. The DTL 2-level model has known identifiability issues, as shown by published results. In consequence it is probable that we face similar ones, with more non-identifiability results because we have more parameters. However we have encountered the following obstacles:

- Identifiability is a problem per se, in addition to all the studies we present. It is highly interesting but would deserve a thoughtful additional simulation campaign, and add a significant amount of information in the text which contains already several different messages.
- We do not have the necessary human resources to achieve such a job correctly. We feel that the absence of identifiability study, despite it

would certainly increase the quality of the paper, does not affect the correctness and interest of what remains.

In consequence we propose this new version without the simulation campaign with the same model. We have added a warning in that sense in the text, which does not change, if we understand this issue well enough, our main results. We trust the editor to choose if the revised article is consistent enough for a PCI recommendation.

-some estimation (or at least discussion) of how often the model could lead to time inconsistent scenarios.

Anwser: We have computed statistics of time inconsistency in simulated data, and added a discussion on this point, based on the literature on time consistency in DTL 2-level models. As expected time inconsistency is significantly reduced when 3-level reconciliations are computed instead of 2-level scenarios, which is an additional argument for our work. We highlight in addition that time consistency for one scenario does not mean time consistency for a dataset containing several gene families, limiting the extant of a method computing time consistent reconciliation scenarios.

I have a few minor comments in addition to the reviewers comments and suggestions:
-concerning the use of the term coevolution: in evolutionary ecology coevolution refers to “reciprocal adaptation in interacting species”, please, if you can, use the terms cophylogeny, codivergence, cospeciation rather than coevolution that refers explicitly to the adaptive process of species.

Anwser: In the revised version we have removed the term coevolution whenever it explicitly qualifies reconciliation, replacing it by “inter-dependencies”, “cophylogeny”, or “similarities in diversification pattern”, according to the context. Coevolution is still used in the article, in general sentences where it refers to the adaptive process and not the comparison of phylogenies.

-p2 line 50, I am not sure that reconciliation methods (using DTL model) have been implemented in Biogeographic analyses: when historical biogeography was first developed with approaches such as Brooks parsimony analyses, the method was applied to both the history of species interactions and biogeography. But biogeographic reconstructions such as the DEC approach now widely used (ref 45) are not similar to reconciliation: there is not “input tree” for geographic areas and the method is more a sophisticated method of ancestral trait reconstruction than a reconciliation. Ref 28 is a review on reconciliation methods and a comment on how they should take into account the biogeography of species rather than an example of how reconciliation methods are applied to reconstructing the biogeographic histories of species.

Anwser: This is right, and we replaced the sentence by “and biogeography has been imagined as one possible level.” We replaced ref 28 by Ronquist (1997) which is closer to this purpose.

-in the repository : please make sure you make the output of the analyses on test datasets available (can you produce the reconciliation on Cinara aphids with Third-Kind in a jpeg or pdf format in sup mat)

Anwser: We updated the data repository, adding all the results and three figures showing reconciliations with Thirdkind. We added a supplementary material pdf file to gather the figures and their legends.

I hope you find these comments helpful and I look forward to handling your revised preprint.

Please address all reviewers’ comments in your response.

Kind regards,

Emmanuelle Jousselin

Reviews

Reviewed by Vincent Berry, 01 Nov 2022 18:28

Overall comment:

The paper considers the co-evolution of hosts, symbionts and genes within symbionts for which it proposes a three-level probabilistic model of evolution. An important contribution is the proposal of two methods to estimate reconciliations and evolutionary event rates in this probabilistic framework. The authors also propose a method to infer the symbiont phylogeny through amalgamation from gene trees and a host tree and a test to check whether considering three entangled levels is worthwhile. Reconciliation inference methods are evaluated both from simulated and real data.
Overall, this paper presents an invaluable step forward on three level reconciliations, a framework jointly considering, e.g., a host tree, a symbiont tree and gene trees (or, e.g., a species tree, a gene tree and a domain tree) linked together through evolutionary events. This problem has been previously considered in a parsimony setting by Stolzer et al, then by Li and Bansal, and in a probabilistic setting by Muhammad et al. The latter focus on inferring gene and domain trees under a Duplication-Loss (DL) model, while the focus of the current paper is on inferring reconciliations, event rates and places of these events in the reconciliations, in a probabilistic model allowing not only duplications and losses but also transfers (DTL model). The practical evaluation is convincing. In particular (but not only!) a documented transfer is only inferred if taking into account the three-level picture. This shows that the model and sampling process presented here are not only elegant but also needed in practice. Moreover reported complexities, confirmed by reported running times, show that the method is computationally efficient.

General comments/questions to authors:
- Current models such as the one presented here allow gene transfers between symbionts. There are a number of documented cases where gene transfers occur between hosts and symbionts.
Allowing such transfers in models would necessitate to trace also species genes (not only symbiont genes), but modeling this might not be as hard as adding a fourth level to the reconciliation model, as genes in hosts and symbionts are in essence evolving from the same ancestral genes... To what extent could your model be extended to integrate this kind of events?

Anwser: Indeed the reviewer is right, the model virtually allows genes for which part of the evolutionary history, possibly including leaves, belongs to the host. The way to implement it would be to construct a phylogeny identical to the host tree, along with a fixed reconciliation between the two

trees. We did not mention this possibility in the first version of the manuscript because we did not have any application, so it is not fully implemented. In particular a fixed reconciliation has to replace our H/S reconciliation step, which is trivial because the two trees would be identical, but not implemented. We added in the revised version a sentence about this possibility. Even if we don't use it, it can be useful for a reader to understand his fact.

- You consider undated trees and somewhere you remark that indeed this might lead to time inconsistent scenarios sometimes. Though in the experimental part of the paper you never come back to this point. Could you state the percentage of unfeasible scenarios you get in the sampling process both on simulated and biological data? I think this is a point that deserves some stats in the experiments.

Anwser: We computed the number of time inconsistencies in simulated data, comparing 2-level and 3-level reconciliations. We reported in the revised text the statistics: 35% of the 2-level scenarios presented inconsistencies, while 15% of the 3-level reconciliations did. This is expected: using the host tree (even if it is in its undated version) reduces time inconsistencies because the host tree contains timing information (ancestors live before their descendants). However the quantification of this effect argues for the use of 3-level methods when possible. We did not report time inconsistency figures of biological data because we had nothing to compare them to. We added a paragraph on time inconsistency in the revised version.

- Also, from an unfeasible (time inconsistent) scenario, do you have a method that brings to a feasible scenario? Would that be possible without losing too much in likelihood? Please say something about that in the paper (or that this point remains to be investigated).

Anwser: We added a subsection that summarises what we know on this subject: (1) inconsistency in reconciliations arises not only in a single scenario for one gene family, but comparing the scenarios from several gene families (2) it is NP-hard to find time consistent scenarios and to our knowledge there is no satisfactory solution to this problem. Solutions implemented for example in Notung or Eucalypt consist in sampling solutions and subsampling time consistent ones. This regularly leads to infinite computations and in the case it leads to solutions, it does not solve inconsistency between families. Post-processings like MaxTic solve inconsistencies between families but do not produce time consistent scenarios. The only way we see to produce time consistent scenarios is to use dated reconciliation. Using undated reconciliation has this inherent drawback, balanced by its simplicity.

- It might seem the Monte Carlo approach you propose is from the importance sampling family. Could you state that explicitly (and give reference) or indicate in which aspect it differs from this family of other MC methods?

Anwser: The idea is indeed close to importance sampling: it consists, just like for importance sampling, in estimating a likelihood by a Monte Carlo sampling a variable in a distribution that is different from the probability distribution of the variable itself. However, importance sampling is more specific: it consists in reducing the variance of a an estimator by sampling according to a biased distribution and correcting the bias in the estimator formula. As we don't exactly fit this definition, we chose not to link our method and this method family.

Detailed comments:

I lacked time to make several reading passes over the paper and to read in detail some recent related works, which might explain the relatively large number of comments / questions I mention below. But overall, I'm confident with the fact that the paper is a useful addition to the known theory and practice in the reconciliation field.

P1, L16: « hosts, symbionts and symbiont genes » would be more precise than « and their genes ». There are other places in the paper were you might be more explicit

Anwser: We acknowledge that there is an ambiguity in this formulation. In this sentence, it is said that the model "can account" for the evolution of hosts, symbionts and their genes. It is not untrue to state that the model can account for the evolution of host genes, as discussed earlier in this letter, so we decided to maintain this formulation. We checked in the remaining of the manuscript that no ambiguous formulation could lead to a wrong perception of the capacities of the model and its implementation.

P1, L20: please indicate that you first fix the H/S reconciliation before sampling S/G reconciliations.

Anwser: We added "It consists in computing reconciliation between pairs of trees, with a dependence between pairs." at this position of the text.

P2, L79: « it is possible to jointly handle three nested levels in a single computational model ... »: this sentence is overly long, please break it in two.

Anwser: we split it and got rid of the parenthesis.

P5, Fig1 is rather useful.

Anwser: thanks

P5, L169: « The inference consists in » -> « begins with » + state that this is done for all gene trees independently (as G can also sometimes denote a set of gene trees).

Anwser: Indeed G is a set of gene trees, so we corrected the text according to the two suggestions.

P5,L169-171: how do you estimate the D, T, L rates for the S/H reconciliation?

Anwser: We added the precision: we estimate the rates using an expectation maximization algorithm based on the ALE algorithm. This is also clarified by the new writing of the algorithm in the algorithmic format.

P5, L172: « it is then possible to estimate the evolutionary rates »: based on the sampled scenarios?

Anwser: Following this remark and other remarks below, we formalized and clarified the presentation of the method. The parameter estimation is included in this new presentation.

P5, L173: « among reconciliation scenarios »: of both H/S and S/G?

Anwser: Scenario sampling is also included in the new presentation of the Monte Carlo method.

P5, L175: it is a bit surprising that a probabilistic approach does not take branch length into account, usually, this is a plus of such methods over parsimony ones. For instance the probability of a non-speciation event would intuitively be considered smaller for short branches.

Anwser: The remark is correct, we cannot handle branch length at this stage, and it would certainly be valuable for the precision of transfer inference. We based our approach on ALE which does not handle branch lengths as well. There are propositions in the literature to do so, as PrimeGSR, but it is challenging to construct an ALE-like method with branch lengths.
We benefit from the other pluses of probabilistic models, like the possibility to compare likelihoods, to estimate rates, to sample solutions according to their likelihoods, to obtain a posteriori probabilities for evolutionary events, to avoid statistical inconsistency and to widen the window of reliability.

P5+6: in sections 3.2+3.5 there is an important Figure missing somewhere here that would clearly depict the steps of the inference process, with their respective input and how they coordinate together. The text in its current

state is not enough for me to be sure of the whole multi-step inference process. I'm sure readers would consider this a useful addition.

Anwser: We added a step by step algorithmic description of the Monte Carlo method in Section 3.2.

P6, L196, « In our model, gene transfer » (add a comma there)

Anwser: Done

P6, L204: « while being in the same target (?) host » (currently there is some ambiguity)

Anwser: We changed the sentencing of this paragraph, which indeed was not optimal

P7, Figure2: the clarity of this figure should be improved.

Anwser: We completely rewrote the caption of this figure, better explaining every part of the figure.

P7,L220: « among the $|S_h|$ symbiont branches present in h » (might be worth precising).

Anwser: Done

P7, equation (4): the last sum is over $k \in H$ ancestor branches of e?

Anwser: The sum is over all branches that are not ancestors of e. We made this clear in the text.

P8, Figure 3: in the middle picture on the top row: is that a transfer from the dead?

Anwser: We precised this in the caption

P9, Table 1: the legend is not self contained. State what is 2-level in particular (G/S without being aware of the H/S co-evolution?)

Anwser: We completed the legend of the Table.

P9, section 3.5: please number the inference steps and refer to these numbers whenever appropriate. Here again the lack of a figure depicting the inference process makes it hard to parse the text.

Anwser: We added a description of the inference method in algorithmic style, with number for the steps.

P9,L276: « Finally we can compute the host aware gene/symbiont

reconciliation »: not coming back to put into question the h/s reconciliation? Why not then? And if not, this seems like a contradiction with what is stated on page 12.

Anwser: Indeed we are not performing any feedback loop on the h/s reconciliation. This is for computing time reasons. It would be possible to imagine such an optimization, which we did not implement.
It does not contradict the sampling procedure because the sampling of h/s reconciliations is fully performed before the s/g ones. So the sample is considered fixed and the method does not come back on it, after parsing G.

P9,L278: « on the donor-receiver symbiont couple »?

Anwser: Indeed this was ambiguous and we changed it to “on the position within symbionts of the donor-receiver couple”.

P9,L282: « we repeat all steps except the initial host/symbiont reconciliation » $->$ « we repeat steps 2 and 3»?

Anwser: We partially rewrote this section on time complexity and this issue is solved by the reformulation.

P10, L293: running times seem quite reasonable given the model sophistication, good! Is that a result of the fact that you do not entangle together the three levels but rather consider the first two levels, then the second and third level with only the knowledge for the second level of whether several symbionts belong to the same host or not (ie you mostly consider partial information from the h/s reconciliation, maybe because only this partial information is relevant to the s/g reconciliation)? Could you elaborate a bit in the paper on this.

Anwser: Yes it is correct. The inference methods are heuristics and dependencies are considered only one-way. This explains why the computing times stay bounded, even if they are still too big for large datasets. We added a sentence on this point.

P10,L307: « We consider the host/symbiont DTL parameters as fixed, i.e. estimated without knowing the data. This makes it possible to compare, based on the likelihood, our approach and a 2-level one »: ok, the inference process might lose a lot of its potential accuracy doing that?

Anwser: It is true that such a proposition is not optimal for precise inference. However it is adapted to the comparison of likelihoods with the 2-level method. So depending on the purposes, either comparative evaluation or optimization of the performances, rates can be fixed or estimated. In that case we fix the rates because we are interested in evaluating the gain of the 3-level method compared with the 2-level

method, based on the likelihood.

P12, L366: does this mean that gene trees are rather similar to one another? Or gene trees are close to symbiont trees? Could you give distance measures between input trees? (dRF or dMAST for instance)

Anwser: For the revised version we compared gene trees based on their number of leaves. The proposed comparisons are more difficult because of the differences in gene content due to duplication and transfer. Leaf numbers go from 5 to more than 100, with a variance of 80, and within a single simulation, it never happens that all gene trees have the same leaf number. So there seem to be a good diversity despite the gene trees being generated from the pruned versions of the symbiont trees.

P12,L378-380: here R(S,H) is sampled which seems to be a contradiction with what is said in section 3.5

Anwser: As answered above, the sample is fixed before doing the s/g reconciliations. We modified the text in order to make clear that h/g reconciliations can be sampled, and considered fixed. The presence of Algorithm 1 as a figure should also clarify this point.

P17, L511-512: it might be a good place to recall that the program is available on GitHub, this was only briefly mentioned in the abstract.

Anwser: Done

P19,L549-550: this is quite honest from you to recognise this.

Anwser: we hope that this remark does not reflect the scarcity of honesty in the scientific literature

Concerning the code repository: beware that some comments are not in English, for instance in main.py

Anwser: We fixed this point.

Reviewed by Catherine Matias, 06 Oct 2022 10:00

This manuscript proposes a 3-level probabilistic co-evolution model, to reconcile the phylogenies of host species with their symbionts species and the genes of the latter. The model generalizes the method called ALE (Amalgamated likelihood estimation) in its version for undated trees, that enables exploring reconciliations between 2 trees under a Duplication-Transfer-Loss model of co-evolution. When considering 3-level undated phylogenies, the new model proposed by the authors enables

considering duplications, transfers and losses of the symbionts within their host species (with probabilities θ_H that are fixed and not estimated by the model; rather these are pre-estimated through an Expectation-Maximization algorithm), together with duplications, intra-transfer and losses of the genes inside the symbiont species (probabilities θs). Here, intra-transfer means that a gene may transfer only between 2 symbionts that are within the same host species at the time of the transfer. Nonetheless, as in ALE, the method includes the possibility to use so-called ghost lineages for (indirect) transfers of genes between symbionts not present in the same host. Note that the method does not check for time feasibility, so it can output invalid reconciliations.

Anwser: We fully agree with this remark, but note that even with a check to time feasibility, methods can output invalid reconciliations. Indeed checking time feasibility on one gene family does not really solve the time consistency problem in general, because two time feasible reconciliations for two different genes can exclude each other because they are mutually time inconsistent. Solving consistency for many gene families means dating the species tree and this is a hard problem. Our philosophy has then been to allow a degree of incoherence. This may have the advantage of pointing at conflicting scenarios, testing trees. For example, in the revised version, we could check that 3-level reconciliations are less often inconsistent and 2-level reconciliations. The level of inconsistency is then transformed into a criterion for method evaluation, which would be impossible if the methods would output time consistent scenarios. We now discuss in the revised version of the text the time consistency problem. We added a dedicated subsection in Section 2 and a paragraph in the description of the simulations.

The authors develop an algorithm for approximating the likelihood of any dataset (trees of hosts, symbionts and their genes) and inferring the parameters of the gene/symbiont co-evolution, relying on two versions of their method (Monte Carlo approximation with samples of reconciliations from the symbiont tree to the host tree; or sequential approach that relies on the most likely reconciliation from the symbiont tree to the host tree). When the symbiont tree is unknown, they also propose an option to infer it by amalgamation. In practice, the method is applied with many gene families (thus many gene trees).

Simulations under an external model are proposed, and the authors compare the 2 versions of their method (sequential and Monte-Carlo based) with a 2-level reconciliation of genes tree in their symbionts tree. Performance is measured with respect to the capacity of the 3 methods to recover gene transfers between correct symbiont donor and symbiont recipient (precision and recall are weighted wrt estimated probability of each transfer). The difference between the likelihoods of symbiont/gene

reconciliations in the 3-level approach and in the 2-level one is used as a measure of host/symbiont co-evolution. Finally, the method is illustrated on 2 datasets: a Cinara aphids enterobacteria system and Helicobacter pylori within humans.

This is an important contribution to the 3-level reconciliation problem. The remarks below should help the authors clarify some points.

Major remarks

1. There is a confusion in the text between most likely reconciliation and maximum likelihood.
I detail the problematic points below.
• When describing the 2-level reconciliation model (line 137 and below), the authors write: "We do not have to enumerate all scenarios to compute that sum, because we can compute this likelihood using dynamic programming, considering matching all couples of gene and species sub-trees, starting from the leaves, and enumerating all possible events to get each match." This is not correct. Dynamic programming is a way to compute, for any parameter value $\theta_S = (p_S^S, p_S^D, p_S^T, p_S^L)$ the quantity

$$\max_{r_{G,S} \in R_{G,S}} \mathbb{P}_{\theta_S}(G, S, r_{G,S}) \tag{1}$$

but this quantity is different from the model likelihood, that equals the sum over all possible reconciliations

$$\mathbb{P}_{\theta_S}(G, S) = \sum_{r_{G,S} \in R_{G,S}} \mathbb{P}_{\theta_S}(G, S, r_{G,S}) \tag{2}$$

Dynamic programming algorithm constructs a table of all possible successive events from the leaves to the root, together with pointers that indicate at each stage the most likely event (for a fixed parameter value θ_S). To obtain the exact likelihood of the data, one should enumerate all possible paths (i.e. reconciliations) within that table and sum the corresponding probabilities; while backtracking in this table only outputs the most likely path (i.e. reconciliation). So if I understood correctly, at this stage (of the reconciliation between G and S) rather than sampling reconciliation scenarios, the authors compute the most likely one, say $\hat{r}_{G,S}$ that realizes the maximum in Eq (1) (for any parameter value θ_S), thanks to dynamic programming. While the chosen strategy makes sense, it's nonetheless different from a maximum likelihood one, where one would estimate θ_S by considering the argmax over θ_S of Eq (2).

Anwser: We clarified in the text (rewriting fully section 2 and adding the ALE equations) that our use of the dynamic programming algorithm, like in ALE, is a computation of the sum (Eq 2) and not the maximum (Eq 1). Indeed,

the forward step of the dynamic programming sums over all scenarios. The backward step then either samples over scenarios according to their likelihood, or computes a "marginal" most likely scenario. The latter is possibly different from quantity (1) because the maximum is chosen at each step and this does not theoretically guaranties that the product of local maxima is a global maximum. This is the difference between “joint likelihood” and “marginal likelihood” described for example by Yang (Computational molecular evolution, Oxford University press, 2006).
This procedure allows us to estimate θ_S by maximum likelihood, exactly as described by the reviewer, that is, approximating argmax over θ_S of Eq (2).

• I believe that one layer of reconciliation is missing in the equations presented in Section 3.2. As far as I understand, Eq.(1) should be modified in the following way (I also added as indexes of the probabilities the different parameters θ_S and θ_H , for more clarity)

$$\begin{aligned} P_{(\theta_S,\theta_H)}(G|S,H) &= \sum_{r_{S,H}\in R_{S,H}} P_{\theta_S}(G|S,H,r_{S,H})P_{\theta_H}(r_{S,H}|S,H) \\ &= \sum_{r_{S,H}\in R_{S,H}} \sum_{r_{G,S}\in R_{G,S}} P_{\theta_S}(G,r_{G,S}|S,H,r_{S,H})P_{\theta_H}(r_{S,H}|S,H), \\ &\simeq \sum_{r_{S,H}\in R_{S,H}} P_{\theta_S}(G,\hat{r}_{G,S}|S,H,r_{S,H})P_{\theta_H}(r_{S,H}|S,H), \end{aligned}$$

where $\hat{r}_{G,S}$ is the most likely reconciliation of G in S (for the current parameter value θ_S and the fixed reconciliation $r_{S,H}$). This first approximation makes sense since the most likely reconciliation $\hat{r}_{G,S}$ contributes to the dominant term in the sum $\sum_{r_{G,S}\in R_{G,S}}$ and one hopes the other terms are negligible. Then, if I understand correctly, a sequence $r_n \in R_{S,H}$ of reconciliations of the symbiont tree S within the host tree H is sampled and the authors make the second approximation of the likelihood through

$$P_{(\theta_S,\theta_H)}(G|S,H) \simeq \frac{1}{N}\sum_{n=1}^{N} P_{\theta_S}(G,\hat{r}_{G,S}|S,H,r_n)P_{\theta_H}(r_n|S,H). \qquad (3)$$

In any case, Eq. (2) in the manuscript is not correct and a weight $P\theta_H(r_n|S,H)$ is missing in that equation.

Anwser: We hope that part of the misunderstanding is solved thanks to the rewriting of section 2 following the previous remark. Indeed, we do not compute $\hat{r}$, the most likely reconciliation, but instead sample N reconciliations from H,S according to their likelihood.
We agree that Equation (2) would not be correct if the reconciliations of S in H were sampled uniformly in the reconciliation space. If it was the case

case we would need to weight by

$$P_{\theta_H}(r_n|S,H)/(\sum(P_{\theta_H}(r_i|S,H)))$$

. But the reconciliations are sampled according to $P_{\theta_H}(r_n|S,H)$. This is achieved using the backtracking of the reconciliation. In consequence, the frequency of appearance of one reconciliation in the sum follows $P_{\theta_H}(r_n|S,H)$. Weighting by $1/N$ only then produces an equivalent (approximated) equation as the one proposed by the reviewer.

To summarize, I understood that (in the first version of the algorithm, the sequential one being different) the authors sample a reconciliation $r_n \in R_{S,H}$; compute its probability $P_{\theta_H}(r_n|S,H)$ (thanks to a dynamic programming table); then they find the most likely reconciliation $\hat{r}_{G,S} \in R_{G,S}$ (thanks to a second dynamic programming table) that maximizes the probability $P_{\theta_S}(G,\hat{r}_{G,S}|S,H,r_n)$, together with the corresponding maximum value of that probability. (Note that this most likely reconciliation depends on the parameter θ_S and on the sampled reconciliation r_n). By doing this for many sampled reconciliations r_n, the authors finally compute the approximation in the right-hand side of Eq. (3). This quantity may be computed for a fixed parameter value (θ_S,θ_H) and the authors search for its maximum wrt (θ_S,θ_H). (In fact, they will pre-estimate θ_H with the Expectation-Maximization approach implemented in ALE; and then output mean a posteriori values for θ_S by sampling reconciliations of the gene/symbiont trees).

Anwser: We tried to make clearer in the text that we do not sample a reconciliation and then compute its probability, but sample with a frequency proportional to its probability, directly following the dynamic programming backtracking. We do not compute $\hat{r}$ the most likely reconciliation, and do not maximize the probabilities.

• In the sequential version of their method (Section 3.4), I understand that the authors now consider the following approximation

$$P_{(\theta_S,\theta_H)}(G|S,H) \simeq P_{\theta_S}(G,\hat{r}_{G,S}|S,H,\hat{r}_{S,H})P_{\theta_H}(\hat{r}_{S,H}|S,H),$$

where $\hat{r}_{S,H}$ is the most likely reconciliation between the symbiont tree and the host tree. If this is indeed the case, it could be useful to write it down.

Anwser: As judiciously suggested, we added in the text the equation of the sequential version, which is
$P_{(\theta_S,\theta_H)}(G|S,H) \simeq P_{\theta_S}(G|S,H,\hat{r}_{S,H})P_{\theta_H}(\hat{r}_{S,H}|S,H),$
where $\hat{r}_{S,H}$ is the reconciliation scenario maximizing the marginal likelihood. Note that we do not compute $\hat{r}_{G,S}$ as explained in the answer above, but the likelihood of the model.

2. The authors choose not to produce simulations under their own model (Line 165). While it's interesting to use an external model as they did, that does not replace the simulations under the true model, to evaluate both the estimation procedure and potential identification issues in the model. Indeed, as the reconciliation models become more and more elaborate, the issue of knowing what portion of information about the past co-evolutionary events remains as a signal in the data is crucial. This can only be assessed through scenarios under the model at stake.

Anwser: We recognize here a lack in our study, though we are unable to fill it currently. We added in the text a short discussion on this point, acknowledging that identifiability issues certainly arise in this framework and it would be important to know them. We miss working strengths to assess them here. We think that it does not affect the interest, honesty and extent of our work, but its completeness (But who knows a really complete work?)

Minor remarks - Line 278: "In consequence we cannot use the efficient computation trick used for uniform rates." Please give a reference for that trick.

Anwser: We added two references for this trick.

- Line 287: and below: should be made explicit that the times are given for the sequential version.

Anwser: Done

- Line 323: "We did that by adding the symbiont tree as a possible host tree". That sentence suggests that many host trees can be input in the method. However, I think this has not been said before. Please clarify this point.

Anwser: We did that by adding the symbiont tree as an additional host tree, as the reconciliation algorithm can accommodate multiple trees on separate sets of leaves, and matching the symbiont leaves with no host to themselves. C'est un essai, je ne suis pas certain que ce soit compréhensible.

Typos - Line 218, $P(e \to h)$ should be $P_T(e \to h)$.

Anwser: done

- Line 413, "*the 1.0 model*" what does that mean?

Anwser: We changed this sentence to "It happens for almost all instances in the simulation dataset with with no intra transfers (inter transfer rates of

1.0)"