

In this work, the authors present a follow-up to their 2018 paper “Gene transfers can date the tree of life” in *Nature Ecology & Evolution*. This work expands upon previous efforts, demonstrating how the relative age constraints imposed by horizontal gene transfer between lineages on a tree are compatible with absolute age estimates constrained by the fossil record, for a test dataset of Cyanobacteria and Alphaproteobacteria. The authors also show that “true” relative age constraints between nodes improves the precision of age estimates for subsets of HGT types. This implementation is also added to RevBayes, so that HGT relative age constraints can be combined with other molecular dating approaches, as made available by the authors.

The authors also take a novel approach in estimating uncertainty in branch length estimations, which underpin estimates of evolutionary rates in ultrametric trees. Essentially, they generate a distribution of branch lengths from a population of trees generated from a MCMC analysis. These are then used to estimate rate parameter variances across branches. This appears to be an efficient approach as an alternative to more computationally costly methods typically employed, as stated by the authors.

This work constitutes an incremental improvement in our understanding of the utility of HGT in refining divergence time estimates. The statistical analyses and phylogenetic/bioinformatics methods appear to be sound and appropriately used. The major limitation of this work, both practically and conceptually, is the starting point of published time-trees in the case of both simulated and empirical investigations; in short, it is not surprising that many HGT constraints, be they simulated as “true” or extracted from actual HGT events, will be redundant with respect to the observed relative ages within the calibrated tree; it is also not surprising that some of the HGT constraints will improve precision of age estimates if these happen to be “active” with respect to branches of high rate variance. That being said, this proof of concept is clearly validated here, which is important. The work would be far more impactful if application of simulated or empirical HGTs to rooted phylogenies independently converged on similar relative age estimates to those that were obtained by calibrated molecular clocks alone.

Major comments:

- (1) The authors test their approach on simulated trees, reconstructed from sequences simulated under the published timetree of Betts et al. As such, their simulated trees are already going to be very close to ultrametric, even after their shuffling approach (this is acknowledged, as the authors state the tree height is the same in both cases). That is, if the relative ages of groups proposed by Betts et al are generally reasonable (being informed by fossils and evolutionary rate models), it is not surprising that the resulting simulated trees will also be largely compatible with a set of “true” transfers.

The rate transformations applied as part of the simulation are extremely modest, with the vast majority of branches only having a 10% variance in rate, with very few branches having a 20% variance. It is unclear if these rate variances were selected to fit those observed within empirical datasets (e.g., Betts et al.). If so, then this is reasonable. However, if not, then the conservative choice of rate variances seems to favor their intended result: compatibility between valid HGT constraints and timetree estimates. Regardless, since this assumption is so important to their conclusions, the authors should perform additional tests, increasing the branch rate variances and observing at what level of branch rate heterogeneity the congruence between relative HGT constraints and the molecular clock model is “broken”. In the introduction, the authors make clear that the utility of this approach is driven by the absence of reliable absolute age constraints within most microbial lineages; therefore, it would be far more convincing if the simulated studies were performed on phylograms rather than chronograms, that more reasonably reflected the information we generally know about microbial groups (in the absence of rate models or fossil constraints). Ideally, the authors would apply their simulations to both the chronogram-derived trees (as they currently do), and the rooted phylogram generated by Betts et al. without any inferred rates, to show that the sets of simulated HGTs reasonably interact with trees that do not contain any dating information within their branches.

(2) In discussing the results shown in Figure 3, the authors state that

“Results improve markedly with 5 or more constraints, with a strong effect when moving from 4 to 5 constraints, and then a slower improvement. There is no obvious feature of constraint 5 that would make it substantially more helpful than other constraints for dating.”

It appears that this “step” effect is largely determined by the order in which HGT constraints were applied. If relative age constraints have a low probability of being “active” as the relative ages of nodes they constrain are already well-resolved by the timetree, then it is expected that stepwise addition will make little change to the error in node ages, until one highly active constraint is applied. This will produce a step down in error (as is observed). Following this step, application of additional HGT constraints will generally not further increase precision, if these are consistent with the first “active” constraint encountered (in the case of the authors approach, constraint #5). In fact, HGT #5 is the first constraint to be imposed that spans the full diversity of the tree, away from the root. This in itself may be a meaningful observation. The authors should shuffle the order in which HGTs are applied in this stepwise approach, or jackknife using different subsets of HGT constraints (much like is often done with fossil calibration) in order to show that

increasing numbers of HGT constraints generally increases precision, and the result is not being driven only by one or a handful of particularly active constraints.

- (3) The results of Figure 5 are not particularly surprising from a traditional divergence time estimate perspective; for proximal age constraints, the uncertainty in the age estimates of each node will be more likely to overlap, and thus be consistent with the presence or absence of the constraint; only distal HGT constraints have the potential to “violate” the rate models strong enough to result in narrowing the age distributions on the constrained nodes. This seems trivial, and it is unclear why the authors have presented this analysis, as it does not further their case for the other hypotheses they test.
- (4) Analysis of Empirical data: The authors applied HGT-based age constraints to cyanobacteria and archaea in order to show that these constraints improved the precision of age estimates independently of the fossil calibrations. However, the list of applied HGTs, what groups they constrained, and how they were obtained is lacking from the manuscript or SI (or, at least in a form that is not recognizable). These constraints are a critical part of their test and should be very clearly documented. If it is the case that these HGT constraints are similarly “simulated” , this should be made much more clear in the methodology.
- (5) Discussion:

“We further found that constraints between nodes of similar ages were less useful than constraints between nodes of differing ages. This is encouraging since it should be easier to find transfers between nodes whose ages differ widely (distal transfers) than between nodes with similar ages, because large age differences give more time for transfers to occur and to leave a detectable footprint in extant genomes.”

This encouragement seems to be misplaced. While it is certainly true that HGTs with donor-recipient ages that vary widely are easier to detect and likely more abundant, these would seem to consist largely of HGTs where the recipient is clearly much younger than the donor clade (e.g., “forward in time” transfers, such as one genus of bacteria being the HGT recipient from within a different order of bacteria), and thus, provide little or no additional information, as branch lengths alone will clearly resolve the recipient as being much younger than the donor. Rather, it seems the most active HGTs will be ones where the relative age distributions for the donor and recipient nodes are highly uncertain, and the HGT constrains the relative ages in a way that conflicts with the rate information—that is, pushes the recipient to

be younger than the donor, when the evolutionary rate models would be most consistent with the inverse. Thus, “active” HGTs in the way described by the authors may actually be quite rare. A more nuanced discussion of this point should be provided.

Minor comments:

(1) the manuscript contains many grammatical errors and appears to have been hastily drafted. It should be carefully copy-edited.

(2) Introduction: The authors provide a useful discussion of previous work using HGT to constrain the relative ages of groups in a tree:

“Recently it has been shown that gene transfers could help date species trees, because they contain information on the chronological order of speciation nodes (Szöllosi et al. 2012; Davín et al. 2018).”

They should also include references to Wolfe & Fournier 2018 and Magnabosco et al., 2018, which also show the application of the same concept with similar methodologies.

(3) Figure 1 is of low quality and should be re-drafted.

(4) Betts et al. should not be cited for the ete3 library references.

(5) In the “Constraints improve dating accuracy” section, the Maximum A Posteriori tree is incorrectly cited as Figure 2A.