Performance of a phylogenetic independent contrast method and an improved pairwise comparison under different scenarios of trait evolution after speciation and duplication

Tina Begum[1,2], Martha Liliana Serrano-Serrano[1,2], Marc Robinson-Rechavi[1,2]*

[1]Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

[2]SIB Swiss Institute of Bioinformatics, Lausanne, Switzerland

*Corresponding author E-mail: marc.robinson-rechavi@unil.ch

Short title: Methods matters in gene family evolution

Abstract

Despite the importance of gene function to evolutionary biology, the applicability of comparative methods to gene function is poorly known. A specific case which has crystalized methodological questions is the "ortholog conjecture", the hypothesis that function evolves faster after duplication (i.e., in paralogs), and conversely conserved between orthologs. Since the mode of functional evolution after duplication is not well known, we investigate under what reasonable evolutionary scenarios phylogenetic independent contrasts or pairwise comparisons can recover a putative signal of different functional evolution between orthologs and paralogs.

We investigate three different simulation models, which represent reasonable but simplified hypotheses about gene function (our "trait") evolution. These are time dependent trait acceleration, correlated changes in rates of both sequence and trait evolution, and asymmetric trait jump. For each model we tested phylogenetic independent contrasts and an improved pairwise comparison method which accounts for interactions between events and node age.

Both approaches loose power to detect the trend of functional evolution when the functional trait accelerates for a long time following duplication, with better power of phylogenetic contrasts under intermediate scenarios. Concomitant increase in evolutionary rates of sequence and of trait after duplication can lead to both an incorrect rejection of the null under null simulations of trait evolution, and a false rejection of the ortholog conjecture under ortholog conjecture simulations by phylogenetic independent contrasts. Improved pairwise comparisons are robust to this bias. Both approaches perform equally well to trace rapid shift in traits.

Considering our ignorance of gene function evolution, and the potential for bias under simple models, we recommend methodological pluralism in studying gene family evolution. Functional phylogenomics is complex and results supported by only one method should be treated with caution.

Keywords: Phylogenetic Independent Contrasts, Pairwise comparisons, Brownian model, Gene duplication, Functional evolution, Simulation

Introduction

In comparative biology, pairwise comparisons of terminal taxa of genes or of species are commonly used to detect phenotypic or morphological or character associations (Felsenstein 1985, Maddison 2000; Martins & Garland 1991; Read & Nee 1995). This pairwise approach can handle sparse data and poorly resolved phylogenies, and relies on relatively few assumptions, compared to other methods that use reconstructed ancestral states, stochastic models of evolution, or phylogenetic branch lengths (Felsenstein 1985; Read & Nee 1995; Purvis & Bromham 1997; Garland *et al.* 1992; Maddison 2000). A potential issue is that multiple pairwise comparisons of data across the same branch of a tree pseudo-replicate data, and overlook phylogenetic relatedness of species (Pagel 1994; Blomberg et al. 2003; Dunn *et al.* 2018). The phylogenetic independent contrasts (PIC) method solves these issues by taking into account the evolutionary history of species (Felsenstein 1985; Grafen 1989; Cooper *et al.* 2016; Dunn *et al.* 2018). This approach needs well-resolved phylogenies. Moreover, it is sensitive to model assumptions (e.g. known branch lengths of the phylogeny, or Brownian character evolution), and may lead to erroneous interpretations if these assumptions are violated (Garland *et al.* 1992; Maddison 2000; DÌaz-Uriarte & Garland 1998; Freckleton & Harvey 2006; Cooper *et al.* 2016). Most comparative functional genomic studies still rely on pairwise comparisons, although some do use PIC.

Applications of both approaches to test the "ortholog conjecture" model, i.e. that there are larger functional differences between paralogs (homologous genes diverging since a duplication) than between orthologs (homologous genes diverging since a speciation), have shown how methodological differences can lead to different results concerning the evolution of gene function (e.g., Nehrt *et al.* 2011; Kryuchkova-Mostacci & Robinson-Rechavi 2016; Dunn *et al.* 2018; Begum & Robinson-Rechavi 2020). Despite

68 these conflicting empirical results, the theoretical expectations of these methods under
69 different models of evolution of gene function have not been studied.

70 Two common features of most observations of gene evolution are asymmetry of rate
71 changes between phylogenetic branches, and that sequence rates are variable (i.e., there
72 is no molecular clock). Both of these can be expected to affect testing of the ortholog
73 conjecture. Functional traits evolve asymmetrically after duplication (Gu *et al.* 2005;
74 Kim & Yi 2006; Cusack & Wolfe 2007; Studer & Robinson-Rechavi 2009). At the
75 sequence level (e.g., dN/dS), there appears to be both an acceleration following gene
76 duplication, and asymmetry of this acceleration (Conant & Wagner 2003; Brunet *et al.*
77 2006; Kim & Yi 2006; Scannell & Wolfe 2008; Studer & Robinson-Rechavi 2009;
78 Panchin *et al.* 2010; Pegueroles *et al.* 2013; Pich & Kondrashov 2014; Holland *et al.*
79 2017). In yeast, Gu *et al.* (2005) reported asymmetric trait acceleration in a time-
80 dependent fashion shortly after duplication, with less dramatically acceleration in
81 sequence evolutionary rates. Trait evolutionary rates are also variable in phylogenies.
82 For example, there can be a period of rapid trait shifts or jumps according to changes
83 in the fitness landscape (e.g. ecological opportunities) following a speciation event
84 (Bokma 2008; Duchen *et al.* 2017; Landis *et al.* 2013; Simpson 1944). Such jump in
85 traits, rather than a change in the evolutionary rates, might also affect trait evolution
86 after gene duplication. The complexity of evolution after duplication led us to enquire
87 under which reasonable simple evolutionary scenarios after gene duplication we can
88 recover a signal of different functional evolution between types of events in gene trees.

89 To address this question, we consider the ortholog conjecture as a test case and we use
90 a gene trait with a bound limit of 0 to 1; we call it τ, like the tissue-specificity score
91 used in previous studies (Yanai *et al.* 2005; Kryuchkova-Mostacci & Robinson-
92 Rechavi 2016; Dunn *et al.* 2018). We compared the performances of PIC and of an

improved pairwise method using three simple simulation models with different parameters of divergence following speciations and duplications.

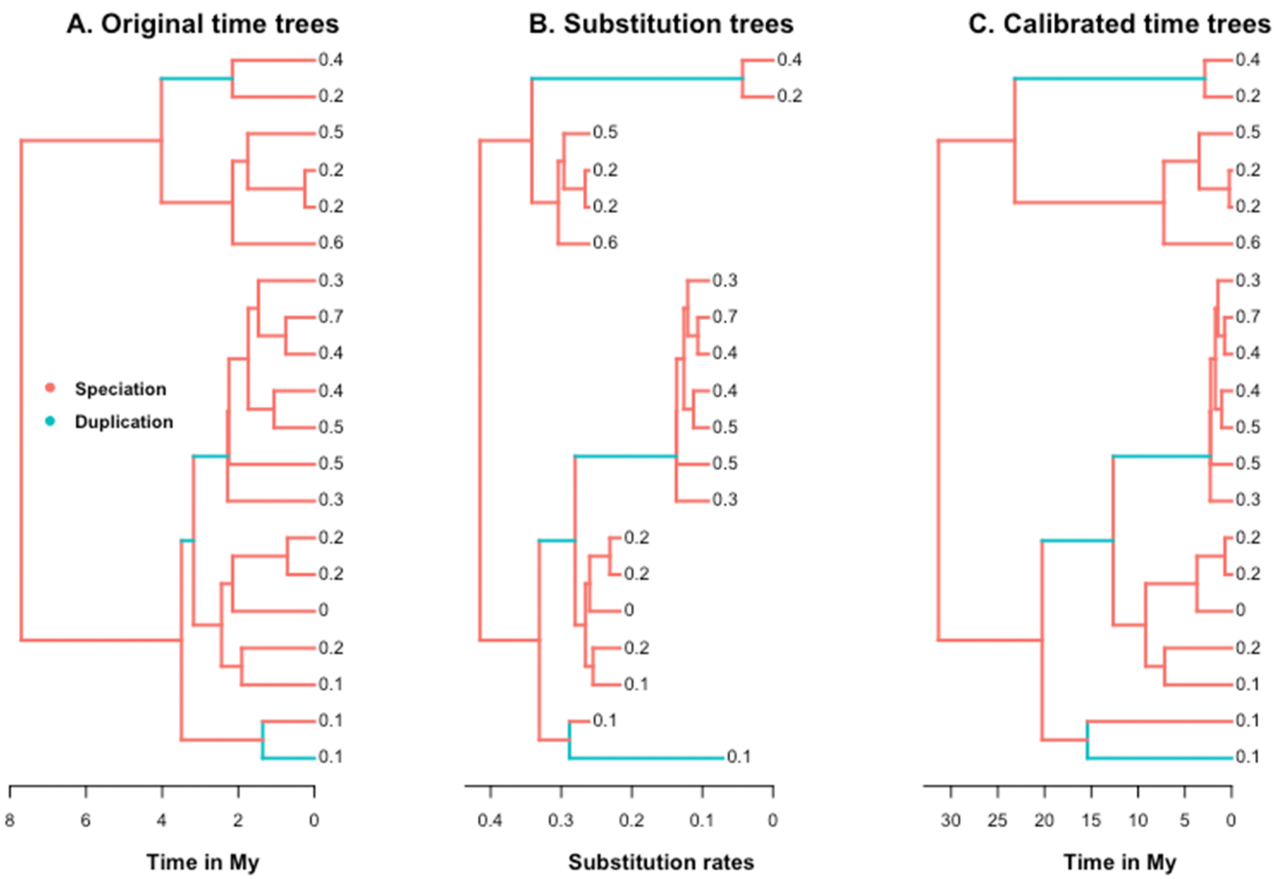
Materials and Methods

Simulated gene trees

We simulated ultrametric trees (n = 10000), each with 100 tips generated under a pure birth-death model using the TreeSim R package (Stadler 2011) with a fixed speciation rate of 0.4, and an extinction rate of 0.1. Since our study aims to mimic simple evolutionary scenarios, where standard trait evolutionary models are frequently applied on comparatively smaller phylogenies (Chira & Thomas 2016), we limited our tree size to100 tips. We call these simulated time trees as "Original time trees" (Fig 1A), with the branch lengths corresponding to units of time (i.e. Million Years – My). For consistency, we used the same set of 10000 original time trees for all further analyses.

Annotation of internal node events

We annotated 99 (number of tips - 1) internal node events ("speciation" or "duplication") so that each tree had at least one speciation and one duplication node events. We considered 3 fixed proportions of duplication events (number of duplication events / total number of internal nodes): 0.2, 0.5, and 0.8 in this study (S1 Fig). At first, we randomly annotated "duplication" events to the internal nodes based on the proportion of duplication events we considered for the tree set. We then assigned "speciation" events to the rest of the nodes of the tree. We, thereby, had 3 sets of 10000 original time trees with 3 different proportions of speciation and duplication events (S1 Fig).



117

118 **Fig 1**: An example of simulated tree models used in this study. (A) Original (true) time
119 tree, (B) substitution rates tree, and (C) pseudo time tree inferred from (B), with
120 simulated trait values at tips. We used a simulated tree with 20 tips for illustration.
121 Branches following gene duplication are asymmetrically painted using the phytools R
122 package (Revell 2012) to introduce different sequence evolutionary rates to those
123 branches. This demonstrates how the scale (i.e. the branch lengths) of the calibrated
124 tree can differ from the original time tree.

125

126 From our reanalyses of Dunn *et al.* (2018) on 8 vertebrate species (Dunn *et al.* 2018;
127 Begum & Robinson-Rechavi 2020), and from analysis of 15 vertebrate genomes
128 (unpublished), we find that the median proportion of duplications events in a range of

129 0.1 - 0.2. Therefore, we used the tree set with a proportion of 0.2 duplication events in
130 all our main text analyses to compare the performances of different approaches in
131 testing the ortholog conjecture on different simulation models. To verify that our results
132 did not depend on the number of duplication events in a tree, we reproduced results
133 using proportions of duplications of 0.5 and 0.8 (supporting materials).

134 **Simulation of trait**

135 In this study, we considered a trait with a range between 0 to 1. We call it, τ like the
136 tissue-specificity score used in several previous studies of the ortholog conjecture
137 (Yanai *et al.* 2005; Kryuchkova-Mostacci & Robinson-Rechavi 2016; Dunn *et al.*
138 2018). For the sake of simplicity, we only used a Brownian model (BM) of trait
139 evolution:

140 $$dX(t) = \sigma^2 t$$

141 where $dX(t)$ quantifies the change in trait, and the Brownian variance σ^2 paremeter
142 describing the rate at which taxa diverge from each other through time t (Cavalli-sforza
143 & Edwards, 1967; Harmon *et al.* 2010; Chira & Thomas 2016). We used FastBM
144 function of the Phytools R package (Revell 2012) to simulate the trait τ on each original
145 time tree with a σ^2 value of 0.02. The same trait values were used for the corresponding
146 subtituion rate trees (described later), and for the pseudo time calibrated trees (described
147 later) for further analyses with phylogenetic method.

148 For modeling of "Time dependent trait acceleration", we wrote our own functions to
149 paint tree with different colors to simulate trait on each original time tree using a BM.
150 In this model, we used a 5 times higher trait evolutionary rate over a specific period of
151 time following gene duplication. The trait evolutionary rates do not vary between events
152 under the null hypothesis.

153 **Generation of calibrated pseudo time trees**

154 We used the NELSI R package (Ho *et al.* 2015), and modified the original
155 "simulate.rates" function to "simulate.rates_heterogeneous" function to introduce
156 different sequence evolutionary rates to the different node events. To avoid any bias in
157 our inference, we used the sequence evolutionary rates, K of 0.02, as used for
158 simulation of the trait τ (σ^2 = 0.02) in a null scenario. These trees are "Substitution
159 trees" (Fig 1B), where the branch lengths of each tree representing substitution rates.
160 To generate "Rates of sequence and trait evolution model", we considered a multiplier
161 of rate parameter (K = 0.02) to a randomly chosen daughter branch of a chronogram
162 following gene duplication.

163 An advantage of simulations is that we have the original time trees. With real data, we
164 never have these original time trees. Rather we use "Substitution trees" (Fig 1B) to
165 generate pseudo time trees, i.e. "Calibrated time trees" (Fig 1C), for hypothesis testing.
166 We used the chronos function of the APE R package (Paradis *et al.* 2004) to generate
167 calibrated time trees (Fig 1C). For such time calibration, we normally have access to
168 only a few speciation time points (focal speciation nodes) for which we have external
169 references (e.g. fossil data) (Dunn *et al.* 2018; Begum & Robinson-Rechavi 2020).
170 However, due to simulated tree structure, very few trees passed time calibration step
171 when we fix focal speciation nodes, and their ages across all trees. Hence, we used the
172 ages of all the speciation nodes of a tree to time calibrate them. The scale of the original
173 tree drastically changes after calibration if we do not fix the age of the root node. It may
174 impact the inference of phylogenetic method due to its branch length dependence.
175 Hence, we also used the age of the root of each simulated original time tree to maintain
176 the scale after time calibration. Following this method, 7711, 6752, and 8251 trees were

177 successfully time calibrated for 0.2, 0.5, and 0.8 proportion of duplication events,
178 respectively.

179 **Phylogenetic analyses**

180 For each internal node of a gene tree, we calculated the PIC of simulated trait τ from
181 the APE R package using the following formula:

182 $$PIC_n = \left(\frac{\tau_1 - \tau_2}{\sqrt{(branchlength_1 + branchlength_2)}} \right)$$

183 where PIC_n is the node of which PIC is to be computed, τ_1, and τ_2 are the trait values
184 of daughters of the corresponding node. The branch lengths of the daughters of
185 corresponding node are represented by $branchlength_1$, and $branchlength_2$, respectively
186 (Felsenstein1985; Paradis *et al.* 2004; Dunn *et al.* 2018). We then contrasted the PIC of
187 each event (speciation or duplication) using a Wilcoxon two-tailed rank test.

188 **Pairwise analyses**

189 We considered all possible pairwise comparisons for each tree to obtain Pearson
190 correlation coefficients R between τ values of events, as was done by Dunn et al.
191 (2018). Few correlation points of speciation and duplication events often make it
192 difficult to infer the statistical significance level precisely. Since most of simulated trees
193 have maximum node ages less than 20 My, we computed the correlation coefficient, R
194 of speciation and of duplication events over each 0.05 My time intervals. This gives the
195 estimates of correlation for both the events at equal time points. We then performed
196 curve fitting to find the best fit model. Previous pairwise comparisons of Kryuchkova-
197 Mostacci and Robinson-Rechavi (2016) used a linear model (i.e. lm(R ~ Event)) to
198 distinguish the effects of events on correlation coefficients. In contrast, our improved
199 pairwise comparison considers an ANCOVA analysis (i.e. aov(R ~ My + Event +

My*Event)) for a linear best fit model, or a nonlinear regression (i.e. gnm(R ~ My + I(My^2) + Event + My*Event)) or gnm(R ~ My + I(My^2) + I(My^3) + Event + My*Event)) to find out the significance level. In all tests of our improved pairwise approach, we included the interaction terms between events and node age, and compared with model without interaction term to determine the best model among them. Since none of our models supported linear fit as best fit model, we used polynomial (quadratic or cubic) regression in our improved pairwise analyses. 70% of our data is used to train the model, and then we used them to predict the value of rest of the 30% test set data. Based on the obtained root mean square error, we chose the best fit (quadratic or cubic) polynomial model for each simulation study to avoid curve over fitting problem. In addition, we considered repeated 10 fold cross validation approach to choose best polynomial model.

Statistical significance

For any method, we considered a result is significant if $P < \alpha = 0.05$. Since we used a large number of tests (n = 84) in our study, we additionally applied a Bonferroni correction where a result is considered to be significant when $P < \alpha'$, where $\alpha' = 0.05 / 84 = 5.95e^{-4}$.

Code availability

For reproducibility, all scripts are available at *https://github.com/tbegum/Method-matters-in-gene-family-evolution*.

220 # Results

221 ## Model-1: Time dependent trait acceleration model

222 Immediately after duplication, both duplicates might experience accelerated evolution
223 for a period of time due to relaxed selection (Rogozin 2014). This can be represented
224 by a simple model of acceleration of trait evolutionary rate for a given time after
225 duplication (Fig 2). After that period of acceleration, the evolutionary rates return to
226 the pre-duplication level. In this model, speciation does not have any impact on
227 evolutionary rates: a speciation during the time period of the duplication-caused
228 acceleration does not revert the gene to a "speciation" rate. To study our capacity to
229 recover evolutionary signal under this model, we simulated with different durations of
230 accelerated evolution, and different proportions of duplications to speciations in the
231 gene trees (Fig 2).

232 In this comparative study, we used an improved pairwise comparison due to the low
233 power of a simple pairwise comparison approach as used by Kryuchkova-Mostacci and
234 Robinson-Rechavi (2016) (S1 Table). First, both methods recover the proper
235 evolutionary signal under the null, with no difference of PIC between duplication and
236 speciation branches (Fig 3A), and no difference between ortholog and paralog pairs
237 (Fig 3B). Second, when there are few duplications in proportion to speciations, both
238 approaches recover the evolutionary signal, namely faster evolution of paralogs or of
239 duplication branches (Fig 3C-F, S1 Table). This is despite the acceleration lasting 50%
240 to 90% of the time since the oldest duplication. It is possible because the rarity of
241 duplications provides many speciation branches without acceleration (Fig 2). While the
242 performance of both methods at recovering the evolutionary signal decreases when the
243 proportion of duplication nodes increases, the PIC is more robust than the pairwise

approach (S1 Table). With 50% of duplication nodes, only the PIC still recovers the evolutionary trend when trait acceleration period increases. At 80% of duplications and an acceleration lasting 70% to 90% of old duplication age, neither approaches can recover the signal of the ortholog conjecture (S1 Table). Under such scenarios, most speciation branches or ortholog pairs have evolved mostly or entirely under the influence of a duplication-caused acceleration, and there is no signal to recover.

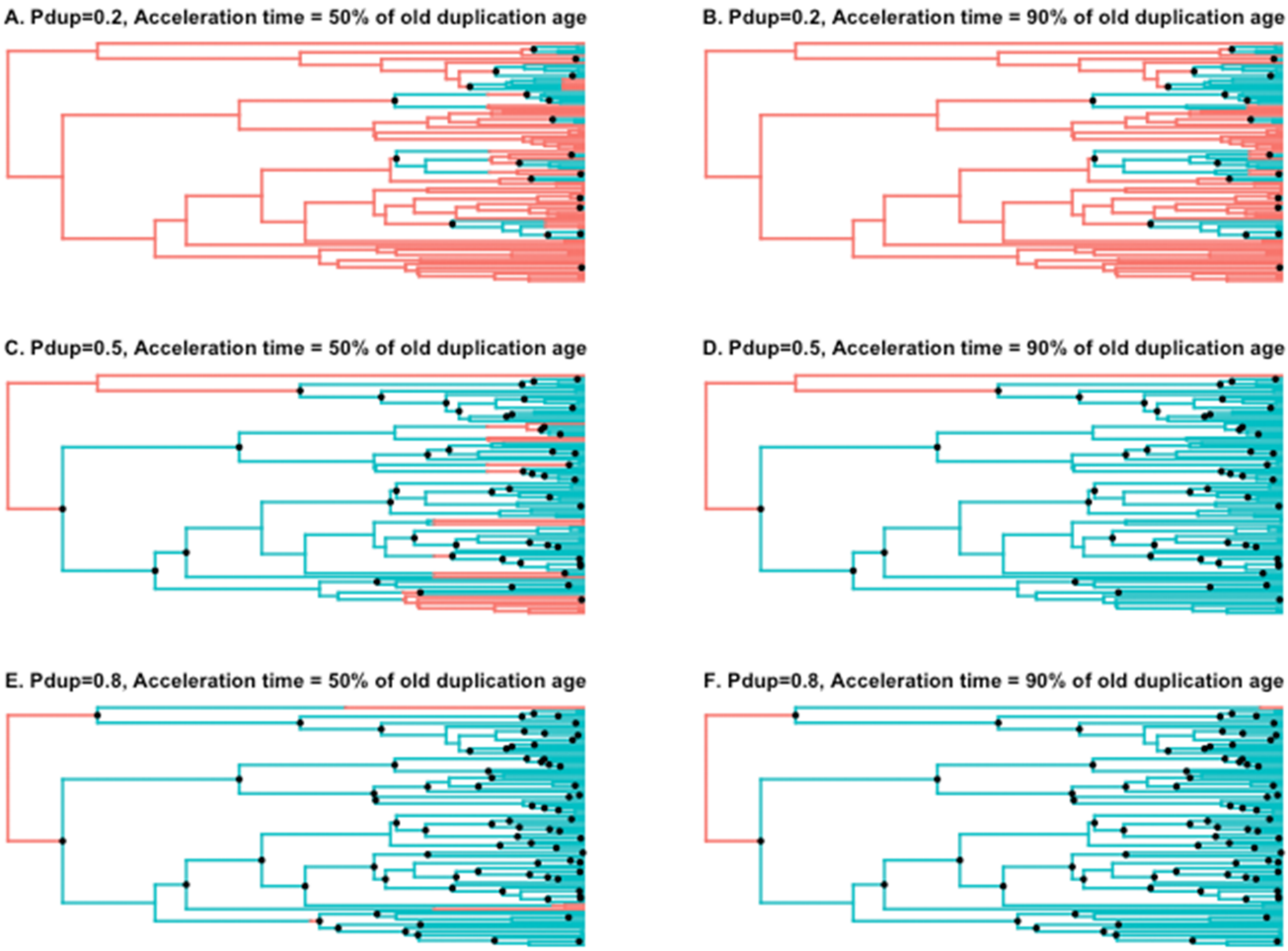


Fig 2: Example of time dependent trait acceleration model. We painted trees to demonstrate the effect of trait acceleration for 50% or 90% of the time since the oldest duplication, respectively. We show here a symmetrical model of trait acceleration, i.e.

after duplication both the duplicates experience trait accelerations. Pdup: Proportion of duplication. Duplication nodes are shown in black, while others are speciations.

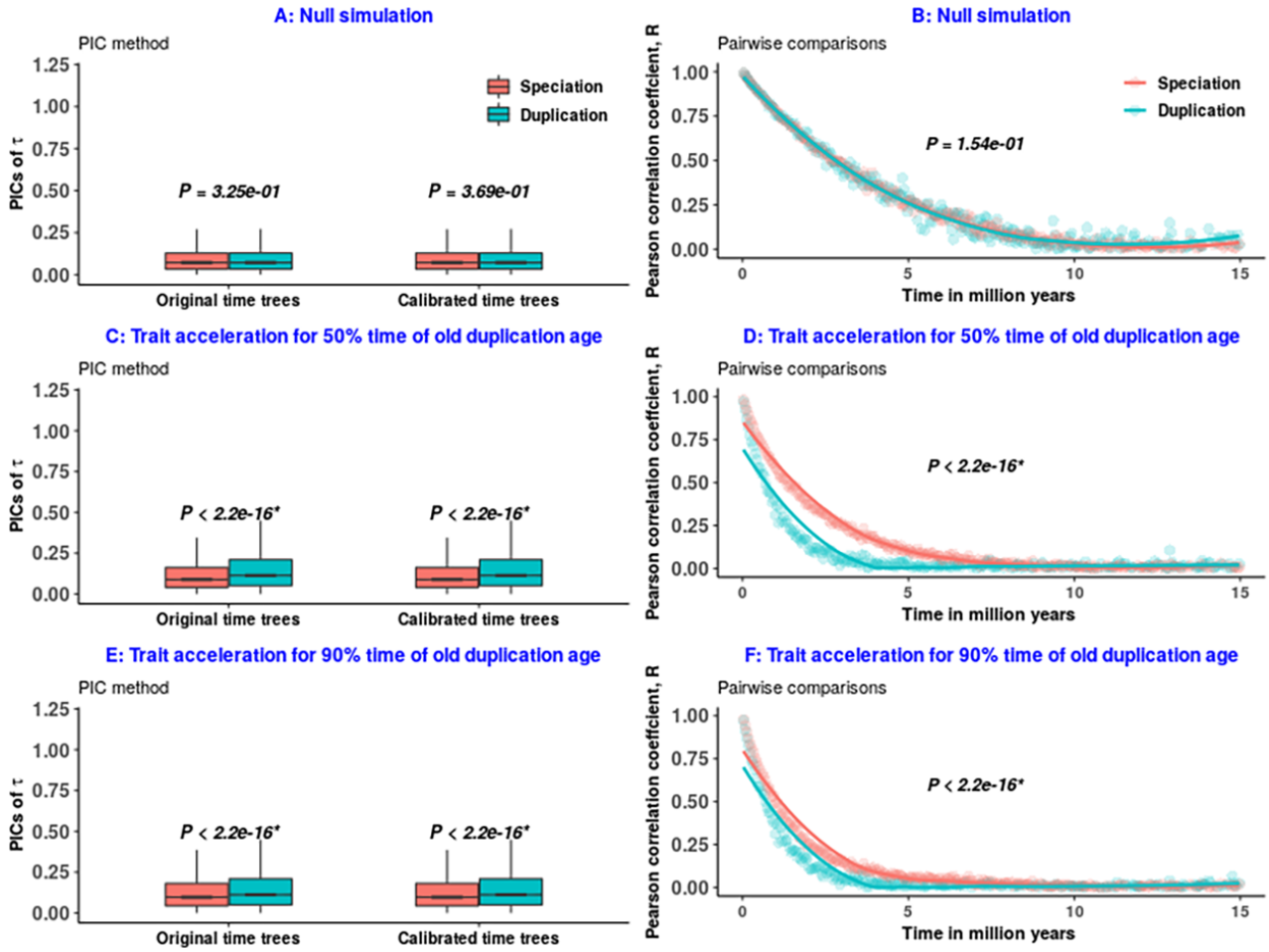


Fig 3: Performances of approaches in different time periods of trait accelerations since oldest duplication. PICs: Phylogenetic Independent Contrasts. In this symmetrical model of 0.2 proportion of duplications, we used a cubic fit polynomial in our improved pairwise approach. *: Bonferroni-corrected significant *P* values.

Model-2: Rates of sequence and trait evolution model

While there is debate over the evolutionary rate of gene function after duplication, it has been shown repeatedly that there is an acceleration of sequence evolutionary rates

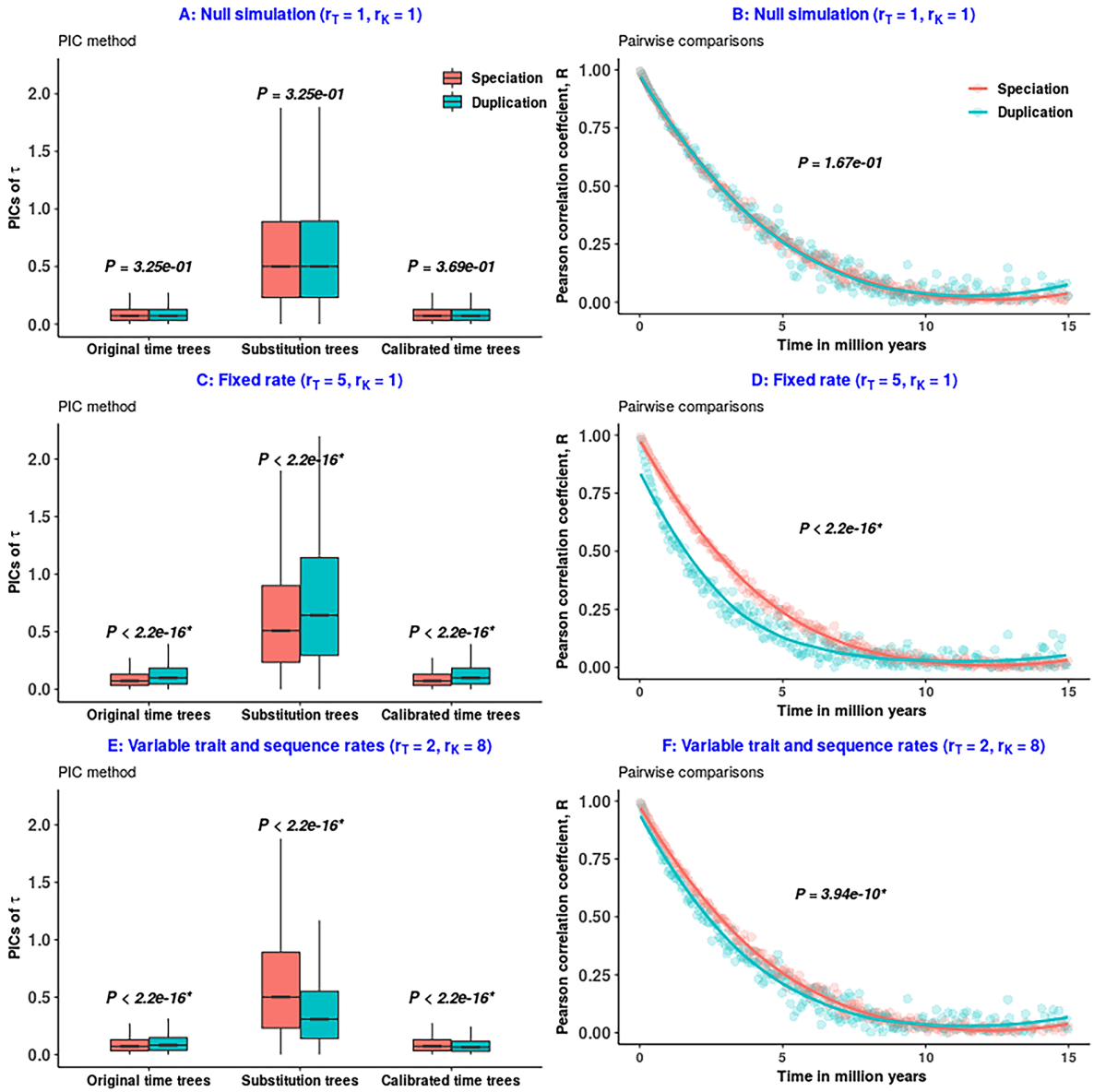
267 (Conant & Wagner 2003; Gu *et al.* 2005; Brunet *et al.* 2006; Kim & Yi 2006; Cusack
268 & Wolfe 2007; Scannell & Wolfe 2008; Studer & Robinson-Rechavi 2009; Panchin *et*
269 *al.* 2010; Pegueroles *et al.* 2013; Pich & Kondrashov 2014; Holland *et al.* 2017). We
270 explore the impact of this acceleration on the PIC and improved pairwise comparisons
271 under different scenarios (Table 1). Because many empirical studies have found
272 asymmetry of this acceleration of duplicate gene sequence evolution, we used an
273 asymmetric acceleration in our simulations.

274 Unlike the previous model, the proportion of duplication events in the trees has no
275 impact on the capacity to recover trends (Fig 4, S2-S4 Figs). An important difference
276 on the other hand is that the null is not always properly recovered. When the trait
277 evolutionary rate is not impacted by duplication, but the sequence evolutionary rate is
278 (scenario 2 of Table 1), the PIC approach wrongly rejects the null for trait evolution.
279 This is due to time tree calibration using sequence evolution as a reference. Indeed there
280 is no difference in PIC when the original time tree is used (A in S2, C in S3 and S4
281 Figs). However, the higher sequence evolutionary rates after duplication lead to higher
282 expected variance, and thus the calibrated time trees produce lower contrasts for
283 duplication than for speciation, with a significant difference in PIC (A in S2, C in S3
284 and S4 Figs). Moreover, if both the trait and the sequence evolutionary rates accelerate
285 after duplication events, but the sequence acceleration is higher (scenario 4c of Table
286 1), the PIC can again estimate lower contrasts for duplication branches with calibrated
287 time trees (Fig 4E, K in S3 and S4 Figs). This can lead to wrongly accepting the null if
288 a one-sided test is used, or rejecting the null in the wrong direction. This bias does not
289 impact the pairwise approach, under neither of these scenarios (Fig 4F, B in S2, D and
290 L in S3 and S4 Figs). Thus even though the assumption of independent data points is

291 violated, the improved pairwise approach is more robust than the PIC in this type of

292 model, since it does not depend on the branch lengths.

293



294

295 **Fig 4**: Comparisons of approaches in 3 different scenarios of Table 1. PICs:

296 Phylogenetic Independent Contrasts. We used an asymmetric model, where after

297 duplication one among the two duplicates experience acceleration. A quadratic fit

298 polynomial is used for our improved pairwise comparisons. *: Bonferroni-corrected

299 significant *P* values. A-B. Both the approaches provide testable predictions under a

300 uniform null. C-D. An ideal ortholog conjecture scenario, i.e. scenario 3 of Table 1. C.

301 5 folds higher trait evolutionary rates produce higher contrasts for duplication than
302 speciation in all 3 tree models, and rejects the null hypothesis. D. Higher correlation
303 coefficients for orthologs than for paralogs also support the conjecture in this case. E-
304 F. The ortholog conjecture scenario 4c of Table 1, where both the rates ratios varies but
305 with an excess of sequence evolutionary rates ratio than trait ratio following duplication
306 than speciation. E. 2 folds higher trait evolutionary rates following duplication supports
307 the conjecture for the original time tree. 8 folds higher sequence evolutionary rates after
308 duplication generate 8 times higher expected variances for duplications, and produce
309 lower contrasts for duplications than speciation in the substitution rates trees. This led
310 us to obtain 2/8 times lower contrasts for duplications than for speciations in the
311 calibrated time trees. Hence, we obtain a significant result but in a wrong direction by
312 rejecting the conjecture in a real ortholog conjecture scenario. F. Improved pairwise
313 comparisons still recovers the expected trend in this case.

314

315 Model-3: Asymmetric trait jump model

316 Finally, we investigated our ability to recover gene evolutionary patterns under a jump
317 model rather than changes in continuous rates. The idea is that, similarly to rapid trait
318 jumps when organisms transition into new adaptive zones (Simpson 1944; Duchen *et*
319 *al.* 2017), paralogs might undergo rapid trait jumps when they appear, e.g. because of
320 a new chromosomal environment in the case of asymmetric duplication. We used a
321 Brownian motion model with asymmetric jumps in trait (Fig 5). Jumps can follow either
322 speciation or duplication events, and different gene evolution models correspond to
323 different probabilities of jumps after the two types of events. We used a simple model
324 where there can only be one or zero jump per branch. Given proportions of speciation

and duplication branches are randomly chosen, and one of their daughter branches (randomly selected) experiences a rapid jump in trait, then returns to progressive trait evolution along the branch; there is only one progressive evolutionary rate. Under the null, equal proportions of speciation and duplication branches (30%) were randomly chosen. Reassuringly, the null was rejected neither by PIC (A in S5 Fig) nor by the pairwise approach (B in S5 Fig). For the ortholog conjecture, we simulated more asymmetric jumps in trait following duplications than speciations (50% *vs.* 20%). More jumps in trait should introduce more divergence in trait, and this should be detected as support for the ortholog conjecture. Indeed, both approaches did support the ortholog conjecture in this case (C and D in S5 Fig). This was robust to trees with different proportions of duplication events (S6 and S7 Figs).

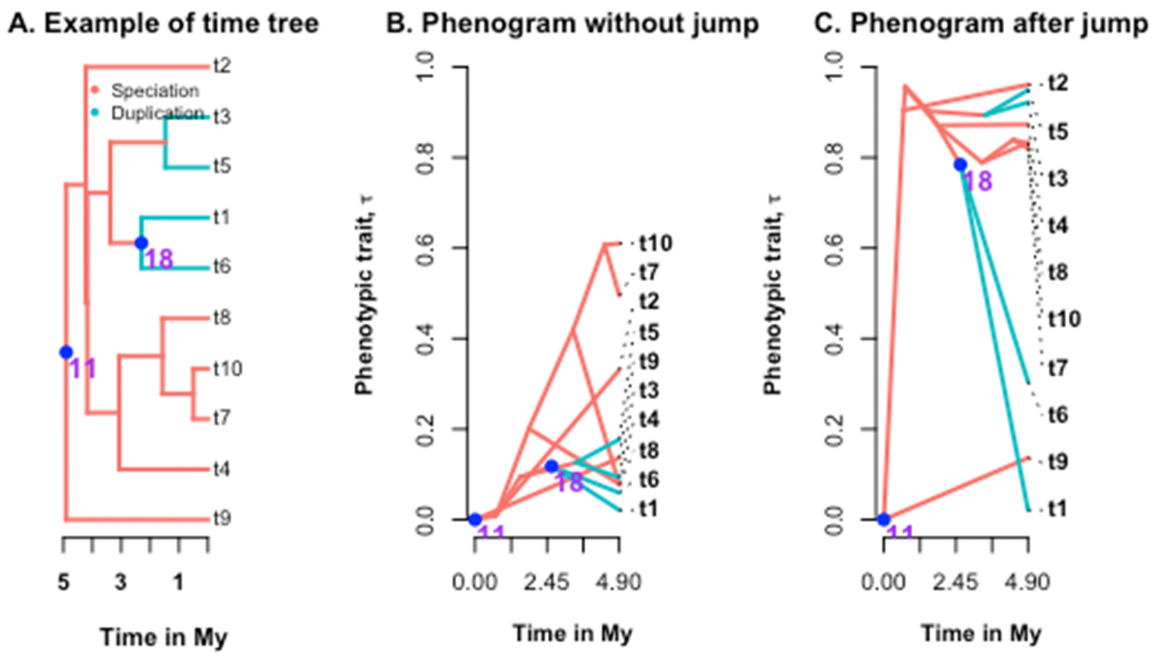


Fig 5: Working principle of an asymmetric trait jump model. A. We used a tree with 10 tips (t1-t10) with a proportion of duplications of 0.2 for illustration. This means that

the tree has 7 speciation and 2 duplication nodes. To represent an ortholog conjecture scenario, we considered that 20% of the speciation nodes, and 50% of the duplication nodes experienced instantaneous jumps in trait between 0 to 1. We chose one of the two branches randomly following the blue colored nodes with numbers, where the trait jump took place. B. The phenogram shows the projection of the phylogenetic tree in a space defined by trait τ of the both events (on the y axis) and evolutionary time (on the x) before the jump in trait. C. The phenogram shows how the trait value changed over the time after experiencing trait jumps. For example, node 11 experienced a trait jump. Hence, the trait value remained unchanged for t9 but alters for the other branch following a jump. This in consequence changed the values of all the following descendants over the time. Hence, the trait value of 18 in C differs from B. For another jump in node 18, t6 changed but not t1.

352 ## Discussion

353 A good method to detect patterns of gene functional evolution should not reject the null
354 hypothesis when data are simulated under the null, should successfully reject the null
355 when simulated under an alternative hypothesis, and should recover the right direction
356 of pattern difference. Here, our null hypothesis is that gene function evolves
357 independently of duplication, and our alternative is the "ortholog conjecture", that
358 function evolves faster after duplication. While the most naive form of this ortholog
359 conjecture is one evolutionary rate on duplication branches, and another speciation
360 branches, we explored three more complex models of trait evolution.

361 First, while the jump model is very different from this basic model, both methods tested
362 (pairwise and PIC) performed well under it (S5-S7 Figs). Thus, testing the ortholog
363 conjecture seems robust to change being sudden or gradual. Second, when gradual
364 change extends in time to both speciation and duplication branches, methods lose power
365 to detect the correct pattern. This loss increases with the duration of acceleration and
366 with the proportion of branches affected. This time dependent trait acceleration model
367 represents a reasonable assumption of gene evolution, as there is no strong reason to
368 expect paralogs to stop evolving like paralogs because there was a speciation. Indeed
369 duplication affects each gene's direct environment, whereas speciation can have at most
370 a very indirect effect. The PIC was more powerful than the pairwise approach under
371 this model (S1 Table). Thirdly and most worryingly, when both sequence and trait
372 evolutionary rates are affected by duplication, PIC can both incorrectly reject the null
373 under null simulations of trait evolution, and report the opposite direction of change
374 under ortholog conjecture simulations (Fig 4E, A in S2, C and K in S3 and S4 Figs).

375 While all of the simulation models used are quite simple, they represent realistic
376 scenarios of gene function evolution. The mechanisms of trait divergence following
377 gene duplication are still poorly understood. For example, we do not have direct
378 evidence for an asymmetric trait jump model. The purpose of our simulation study was
379 to identify scenarios where the PIC and the pairwise comparisons may produce distinct
380 signals of functional evolution between orthologs and paralogs. Although our simple
381 simulations used the same mechanism of trait evolution for all the duplicates of a tree,
382 a mix of different models for duplicates in different parts or at different times in a same
383 gene tree cannot be excluded.

384 Pairwise comparisons of traits between genes has been criticized in studying functional
385 evolution (Dunn *et al.* 2018). Behavior under a null model can help to assess the
386 presence of a systematic bias in the approach. In all different models, we found that our
387 improved pairwise comparisons gave the expected results under the null. The simple
388 pairwise comparison of Kryuchkova-Mostacci and Robinson-Rechavi (2016) also
389 produced the expected signal under the null (S1 Table), although it lacked power under
390 the alternative model. The largest difference between pairwise and PIC was the failure
391 of PIC when both rates of sequence and of trait are affected by duplication (S2-S4 Figs).
392 This is worrying since acceleration of sequence evolutionary rates is common following
393 gene duplications (Gu *et al.* 2005; Scannell & Wolfe 2008; Panchin *et al.* 2010; Assis
394 & Bachtrog 2013; Pegueroles *et al.* 2013; Kryuchkova-Mostacci & Robinson-Rechavi
395 2016; Jiang & Assis 2017; Lafond *et al.* 2018). Thus the scenario of these simulations
396 is expected to be widespread in real data. In this case, higher sequence evolutionary
397 rates for duplicates lead to higher expected variances, and thus produce lower PIC for
398 duplication events than speciation events. These lower PIC are in contrast to the null
399 expectation of no difference.

400 Our simulations also show that in the most probable ortholog conjecture scenario,
401 where both the trait and the sequence evolutionary rates accelerate after duplication
402 events (Gu *et al.* 2005), but the ratio of sequence to the trait acceleration is higher
403 (scenario 4c of Table 1), the PIC method fails to detect the signal of ortholog conjecture
404 (Fig 4E, K in S3 and S4 Figs). In this case, improved pairwise comparisons seem to be
405 a better choice due to its branch lengths independence.

406 The only model for which the ratio of duplication to speciation events impacts the
407 results is with time-dependent trait acceleration. When there are many duplications, and
408 the duplication-induced acceleration lasts for 70% to 90% of old duplication age, most
409 speciation branches also evolve under the influence of trait acceleration like duplicates.
410 Hence, both the approaches fail to recover the signal of the ortholog conjecture in a real
411 ortholog conjecture scenario (S1 Table). Under intermediate scenarios, PIC allows to
412 better recover the signal of the ortholog conjecture than pairwise comparisons.

413 Overall the main message of this study is that our ignorance of gene function evolution
414 makes it difficult to choose one method as the gold standard. Each method has its own
415 limitations, and we call for methodological pluralism, at least in the present state of our
416 knowledge. Relying on a single approach to interpret result can be problematic in many
417 cases of comparative functional genomics study involving complex gene family
418 evolution. These results apply in principle not only to the ortholog conjecture, but to
419 any other cases of gene trait evolution (e.g., horizontal gene transfer), where evolution
420 might be gradual or by jumps, affect a more or less large proportion of branches of the
421 gene tree, and be confounded with changes in sequence evolutionary rates. In
422 conclusion, functional phylogenomics is complex and results supported by only one
423 method should be treated with caution.

Acknowledgements

We sincerely thank Julien Wollbrett, Elsa Guillot, Sara Fonseca Costa, and all the members of the Robinson-Rechavi and Salamin groups for their help and useful discussions. Parts of the computations were performed at the Vital-IT (*http://www.vital-it.ch*) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics, as well as using the Wally cluster of the University of Lausanne.

Author contributions

TB, MLSS, and MRR conceived the ideas; TB and MLSS designed the original methodology; TB and MRR refined the methodology; TB and MLSS wrote the code for analysis with input from MRR; TB analyzed the data; TB and MRR led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Table 1: Hypothetical scenarios of the rates of sequence and trait evolution model.

Sequence evolutionary rates (K) → / Trait evolutionary rates (T) ↓	$K_{Speciation} = K_{Duplication}$	$K_{Duplication} > K_{Speciation}$		
$T_{Speciation} = T_{Duplication}$ **OC false**	**1: Null simulation** ($r_T = 1$, $r_K = 1$)	**2: Fixed trait** ($r_T = 1$, $r_K > 1$)		
$T_{Duplication} > T_{Speciation}$ **OC true**	**3: Fixed rate** ($r_T > 1$, $r_K = 1$)	**4: Variable trait and sequence rates** ($r_T > 1$, $r_K > 1$)		
		(a) $r_T > r_K$	**(b) $r_T = r_K$**	**(c) $r_T < r_K$**

r_T: Ratio of trait evolutionary rates, and r_K: ratio of sequence evolutionary rates between speciation and duplication. Ratio as 1 refers to no change, while ratio > 1 refers to higher rates following gene duplication than speciation in the corresponding rate parameter(s). 1: Both the rate parameters remain same under a null, 2. The trait evolutionary rate is not affected by duplication, but the sequence evolutionary rate is in

the fixed trait scenario. 1-2: Null scenarios due to no change in r_T. 3: Under a fixed rate, the trait evolutionary rate varies following duplication, but the sequence evolutionary rate does not. 4. Correlated changes in both the rate parameters with three different possibilities like (a)-(c). 3-4: Representations of the ortholog conjecture scenarios due to the increase in r_T following duplications.

452 **Table 2**: Summary table for 3 simulation models used in this study.

Simulation model	**Relevant parameters**	**Performance of PIC method**	**Performance of improved pairwise comparisons**
Time dependent trait acceleration	Trait acceleration time, and heterogeneous rates of trait evolution between events	Performs well for trees with few, and intermediate duplication events. It lacks power to detect the actual trend when trait accelerates for a long period since old duplication for trees with many duplicates.	Performs well for trees with few duplications (proportion 0.2) but lacks power in most cases for trees with proportion of duplication events of 0.5, and of 0.8.
Rates of sequence and trait evolution	Rates of sequence evolution, and traits evolutionary rates of events	Fails to show the actual trend of functional evolution under a null, and under certain ortholog conjecture scenario	Recovers the actual trends in all hypothetical conditions
Asymmetric trait jump model	Branch of jump and random change in trait between 0-1	Performs well in all hypothetical scenarios	Recovers the actual signals in all hypothetical conditions

453
454

455

456 ## References

457 Assis, R. & Bachtrog, D. (2013). Neofunctionalization of young duplicate genes in
458 drosophila. *Proc Natl Acad Sci U S A*, **110**, 17409–14. Retrieved from
459 *https://www.ncbi.nlm.nih.gov/pubmed/24101476*

460 Begum, T. & Robinson-Rechavi, M. (2020). Special care is needed in applying
461 phylogenetic comparative methods to gene trees with speciation and duplication nodes.
462 *bioRxiv*. Retrieved from *https://www.biorxiv.org/content/10.1101/719336v2.full*

463 Blomberg, S.P., Garland, T. & Ives, A.R. (2003). Testing for phylogenetic signal in
464 comparative data: Behavioral traits are more labile. *Evolution* **57**:717–745. Retrieved
465 from *https://www.ncbi.nlm.nih.gov/pubmed/12778543*

466 Bokma, F. (2008). Detection of 'punctuated equilibrium' by bayesian estimation of
467 speciation and extinction rates, ancestral character states, and rates of anagenetic and
468 cladogenetic evolution on a molecular phylogeny. *Evolution*, **62**, 2718–26. Retrieved
469 from *https://www.ncbi.nlm.nih.gov/pubmed/18752617*

470 Brunet, F.G., Roest Crollius, H., Paris, M., Aury, J.M., Gibert, P., Jaillon, O., Laudet,
471 V. & Robinson-Rechavi, M. (2006). Gene loss and evolutionary rates following whole-
472 genome duplication in teleost fishes. *Mol Biol Evol*, **23**, 1808–16. Retrieved from
473 *https://www.ncbi.nlm.nih.gov/pubmed/16809621*

474 Cavalli-Sforza, L.L. & Edwards, A.W.F. (1967). Phylogenetic analysis. Models and
475 estimation procedures. Am J Hum Genet, **19**, 233–257.

Chira, A.M. & Thomas, G.H. (2016). The impact of rate heterogeneity on inference of phylogenetic models of trait evolution. *J Evol Biol*, **29**, 2502–2518. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/27653965*

Conant, G.C. & Wagner, A. (2003). Asymmetric sequence divergence of duplicate genes. *Genome Res*, **13**, 2052–8. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/12952876*

Cooper, N., Thomas, G.H. & FitzJohn, R.G. (2016). Shedding light on the 'dark side' of phylogenetic comparative methods. *Methods Ecol Evol*, **7**, 693–699. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/27499839*

Cusack, B.P. & Wolfe, K.H. (2007). Not born equal: Increased rate asymmetry in relocated and retrotransposed rodent gene duplicates. *Mol Biol Evol*, **24**, 679–86. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/17179139*

DÌaz-Uriarte, R. & Garland, T. (1998). Effects of branch length errors on the performance of phylogenetically independent contrasts. *Syst Biol*, **47**, 654–72. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/12066309*

Duchen, P., Leuenberger, C., Szil·gyi, S.M., Harmon, L., Eastman, J., Schweizer, M. & Wegmann, D. (2017). Inference of evolutionary jumps in large phylogenies using lèvy processes. *Syst Biol*, **66**, 950–963. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/28204787*

Dunn, C.W., Zapata, F., Munro, C., Siebert, S. & Hejnol, A. (2018). Pairwise comparisons across species are problematic when analyzing functional genomic data. *Proc Natl Acad Sci U S A*, **115**, E409–E417. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/29301966*

499 Felsenstein, J. (1985). Phylogenies and the comparative method. *The American*
500 *Naturalist*, **125**, 1–15. Retrieved from *https://www.jstor.org/stable/2461605*

501 Freckleton, R.P. & Harvey, P.H. (2006). Detecting non-brownian trait evolution in
502 adaptive radiations. *PLoS Biol*, **4**, e373. Retrieved from
503 *https://www.ncbi.nlm.nih.gov/pubmed/17090217*

504 GabaldÛn, T. & Koonin, E.V. (2013). Functional and evolutionary implications of gene
505 orthology. *Nat Rev Genet*, **14**, 360–6. Retrieved from
506 *https://www.ncbi.nlm.nih.gov/pubmed/23552219*

507 Garland, T.J., Harvey, P. & Ives, A. (1992). Procedure for the analysis of comparative
508 data using phylogenetically independent contrasts. *Systematic Biology*, **41**, 18–32.

509 Grafen, A. (1989). The phylogenetic regression. *Philos Trans R Soc Lond B Biol Sci*,
510 **326**, 119–57. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/2575770*

511 Gu, X., Zhang, Z. & Huang, W. (2005). Rapid evolution of expression and regulatory
512 divergences after yeast gene duplication. *Proc Natl Acad Sci U S A*, **102**, 707–12.
513 Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/15647348*

514 Hallin, J. & Landry, C.R. (2019). Regulation plays a multifaceted role in the retention
515 of gene duplicates. *PLoS Biol*, **17**, e3000519. Retrieved from
516 *https://www.ncbi.nlm.nih.gov/pubmed/31756186*

517 Harmon, L.J., Losos, J.B., Jonathan Davies, T., Gillespie, R.G., Gittleman, J.L., Bryan
518 Jennings, W., Kozak, K.H., McPeek, M.A., Moreno-Roark, F., Near, T.J., Purvis, A.,
519 Ricklefs, R.E., Schluter, D., Schulte Ii, J.A., Seehausen, O., Sidlauskas, B.L., Torres-
520 Carvajal, O., Weir, J.T. & Mooers, A. (2010). Early bursts of body size and shape

evolution are rare in comparative data. *Evolution*, **64**, 2385–96. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/20455932

Ho, S.Y., DuchÍne, S. & DuchÍne, D. (2015). Simulating and detecting autocorrelation of molecular evolutionary rates among lineages. *Mol Ecol Resour*, **15**, 688–96. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25155426

Holland, P.W., MarlÈtaz, F., Maeso, I., Dunwell, T.L. & Paps, J. (2017). New genes from old: Asymmetric divergence of gene duplicates and the evolution of development. *Philos Trans R Soc Lond B Biol Sci*, **372**. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/27994121

Huerta-Cepas, J., Dopazo, J., Huynen, M.A. & GabaldÛn, T. (2011). Evidence for short-time divergence and long-time conservation of tissue-specific expression after gene duplication. *Brief Bioinform*, **12**, 442–8. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/21515902

Jiang, X. & Assis, R. (2017). Natural selection drives rapid functional evolution of young drosophila duplicate genes. *Mol Biol Evol*, **34**, 3089–3098. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28961791

Kim, S.H. & Yi, S.V. (2006). Correlated asymmetry of sequence and functional divergence between duplicate proteins of saccharomyces cerevisiae. *Mol Biol Evol*, **23**, 1068–75. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/16510556

Kryuchkova-Mostacci, N. & Robinson-Rechavi, M. (2016). Tissue-specificity of gene expression diverges slowly between orthologs, and rapidly between paralogs. *PLoS Comput Biol*, **12**, e1005274. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/28030541

544 Lafond, M., Meghdari Miardan, M. & Sankoff, D. (2018). Accurate prediction of
545 orthologs in the presence of divergence after duplication. *Bioinformatics*, **34**, i366–
546 i375. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/29950018*

547 Landis, M.J., Schraiber, J.G. & Liang, M. (2013). Phylogenetic analysis using lèvy
548 processes: Finding jumps in the evolution of continuous traits. *Syst Biol*, **62**, 193–204.
549 Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/23034385*

550 Maddison, W.P. (2000). Testing character correlation using pairwise comparisons on a
551 phylogeny. *J Theor Biol*, **202**, 195–204. Retrieved from
552 *https://www.ncbi.nlm.nih.gov/pubmed/10660474*

553 Martins, E.P. & Garland, T. (1991). Phylogenetic analyses of the correlated evolution
554 of continuous characters: A simulation study. *Evolution*, **45**, 534–557. Retrieved from
555 *https://www.ncbi.nlm.nih.gov/pubmed/28568838*

556 Nehrt, N.L., Clark, W.T., Radivojac, P. & Hahn, M.W. (2011). Testing the ortholog
557 conjecture with comparative functional genomic data from mammals. *PLoS Comput*
558 *Biol*, **7**, e1002073. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/21695233*

559 Pagel, M. (1994). Detecting correlated evolution on phylogenies: A general method for
560 the comparative analyses of discrete characters. *Proceedings of the Royal Society B:*
561 *Biological Sciences*, **255**, 37–45.

562 Panchin, A.Y., Gelfand, M.S., Ramensky, V.E. & Artamonova, I.I. (2010). Asymmetric
563 and non-uniform evolution of recently duplicated human genes. *Biol Direct*, **5**, 54.
564 Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/20825637*

565 Paradis, E., Claude, J. & Strimmer, K. (2004). APE: Analyses of phylogenetics and
566 evolution in R language. *Bioinformatics*, **20**, 289–90. Retrieved from
567 *https://www.ncbi.nlm.nih.gov/pubmed/14734327*

568 Pegueroles, C., Laurie, S. & Alb‡, M.M. (2013). Accelerated evolution after gene
569 duplication: A time-dependent process affecting just one copy. *Mol Biol Evol*, **30**,
570 1830–42. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/23625888*

571 Pich I RosellÛ, O. & Kondrashov, F.A. (2014). Long-term asymmetrical acceleration
572 of protein evolution after gene duplication. *Genome Biol Evol*, **6**, 1949–55. Retrieved
573 from *https://www.ncbi.nlm.nih.gov/pubmed/25070510*

574 Purvis, A. & Bromham, L. (1997). Estimating the transition/transversion ratio from
575 independent pairwise comparisons with an assumed phylogeny. *J Mol Evol*, **44**, 112–9.
576 Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/9010143*

577 Read, A. & Nee, S. (1995). Inference from binary comparative data. *Journal of*
578 *Theoretical Biology*, **173**, 99–108.

579 Revell, L.J. (2012). Phytools: An r package for phylogenetic comparative biology (and
580 other things). *Methods in Ecology and Evolution*, **3**, 217–223.

581 Rogozin, I.B. (2014). Complexity of gene expression evolution after duplication:
582 Protein dosage rebalancing. *Genet Res Int*, **2014**, 516508. Retrieved from
583 *https://www.ncbi.nlm.nih.gov/pubmed/25197576*

584 Scannell, D.R. & Wolfe, K.H. (2008). A burst of protein sequence evolution and a
585 prolonged period of asymmetric evolution follow gene duplication in yeast. *Genome*
586 *Res*, **18**, 137–47. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/18025270*

Simpson, G. (1944). *Tempo and mode in evolution. a wartime book.* New York: Columbia University Press.

Stadler, T. (2011). Simulating trees with a fixed number of extant species. *Syst Biol*, **60**, 676–84. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/21482552*

Studer, R.A. & Robinson-Rechavi, M. (2009). How confident can we be that orthologs are similar, but paralogs differ? *Trends Genet*, **25**, 210–6. Retrieved from *https://www.ncbi.nlm.nih.gov/pubmed/19368988*

Yanai, I., Benjamin, H., Shmoish, M., Chalifa-Caspi, V., Shklar, M., Ophir, R., Bar-Even, A., Horn-Saban, S., Safran, M., Domany, E., Lancet, D. & Shmueli, O. (2005). Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics*, **21**, 650–659. Retrieved from *<Go to ISI>://WOS:000227241200012*

Supporting Information

S1 Table: Performances of the PIC and pairwise (PW) approaches for time dependent trait acceleration model.

	***P* values from proportion of duplications = 0.2**			***P* values from proportion of duplications = 0.5**			***P* values from proportion of duplications = 0.8**		
Approach	PIC	PW	PW_KMRR	PIC	PW	PW_KMRR	PIC	PW	PW_KMRR
Null hypothesis	$3.69e^{-01}$	$1.54e^{-01}$	$6.12e^{-01}$	$6.78e^{-01}$	$3.50e^{-01}$	$7.44e^{-01}$	$3.71e^{-01}$	$2.15e^{-01}$	$4.73e^{-01}$
Trait acceleration for 50% time of old duplication age	$< 2.2e^{-16}$*	$< 2.2e^{-16}$*	$1.11e^{-4}$*	$< 2.2e^{-16}$*	$3.74e^{-05}$*	$2.46e^{-01}$	$4.03e^{-12}$*	$7.33e^{-01}$	$6.05e^{-01}$
Trait acceleration for 60% time of old duplication age	$< 2.2e^{-16}$*	$< 2.2e^{-16}$*	$4.62e^{-4}$*	$< 2.2e^{-16}$*	$3.22e^{-03}$	$4.20e^{-01}$	$1.06e^{-04}$*	$9.44e^{-01}$	$6.92e^{-01}$
Trait acceleration for 70% time of	$< 2.2e^{-16}$*	$< 2.2e^{-16}$*	$2.04e^{-3}$	$< 2.2e^{-16}$*	$2.22e^{-01}$	$5.89e^{-01}$	$1.17e^{-01}$	$9.28e^{-01}$	$7.49e^{-01}$

old duplication age									
Trait acceleration for 80% time of old duplication age	$< 2.2e^{-16}$*	$< 2.2e^{-16}$*	$6.00e^{-3}$	$< 2.2e^{-16}$*	$4.1e^{-01}$	$7.1e^{-01}$	$9.06e^{-01}$	$8.62e^{-01}$	$7.8e^{-01}$
Trait acceleration for 90% time of old duplication age	$< 2.2e^{-16}$*	$< 2.2e^{-16}$*	$1.89e^{-2}$	$2.91e^{-11}$*	$5.69e^{-01}$	$7.97e^{-01}$	$6.39e^{-01}$	$8.17e^{-01}$	$8.00e^{-01}$

Note: PIC: Phylogenetic Independent Contrasts method. PW: Improved pairwise comparisons. PW_KMRR: pairwise comparisons as used by Kryuchkova-Mostacci and Robinson-Rechavi (2016). *P* values of PIC method are shown for calibrated time trees. Due to independence of branch lengths, *P* values of pairwise comparisons are estimated based on original time trees. *: Bonferroni-corrected significant *P* values.

607 # Supplementary figure legends

608 **S1 Fig**: We used same simulated tree with 100 tips to illustrate 3 different sets (A-C).
609 The branch lengths of the trees are in million years, My. For A-C, the proportion of
610 duplications are 0.2, 0.5, and 0.8, respectively. This means that out of 99 internal nodes
611 20 nodes, 50 nodes, and 80 nodes are duplication events, respectively in A-C. Rest of
612 the nodes are annotated as speciation nodes.

613 **Fig S2**: Comparisons of approaches for trees with proportion of duplications of 0.2.
614 PICs: phylogenetic independent contrasts. *: Bonferroni-corrected significant *P* values.
615 We used an asymmetric model, where after duplication one among the two duplicates
616 experience acceleration. A quadratic polynomial is used in our improved pairwise
617 approach. A-B. Scenario 2 of Table 1, A. Due to no variation in trait evolutionary rates
618 following speciation and duplication events, the original time trees recover an expected
619 signal of no difference in contrasts. An excess of sequence evolutionary rates following
620 duplication than speciation produces lower PICs for duplications compared to
621 speciations for the substitution rate trees, and for the calibrated time trees. This leads to
622 a false rejection of the null in a real null condition for the calibrated trees. B. The
623 improved pairwise approach shows the expected trend by showing no difference in
624 correlation coefficients between events in this case. C-F. The ortholog conjecture
625 scenario 4a and 4b of Table 1. When the trait evolutionary rates ratio is higher than (C-
626 D), or equals to (E-F) the sequence evolutionary rates ratio of events, higher changes
627 in trait rates following duplications make both the approaches to recover the expected
628 signals in support of the conjecture.

629 **Fig S3**: Performances of the PIC *vs.* improved pairwise approaches in all hypothetical
630 conditions of Table 1. These plots are generated for trees with intermediate duplication

events, i.e. for 0.5. Results are for asymmetric model. PICs: Phylogenetic independent contrasts. *: Bonferroni-corrected significant *P* values. We used a quadratic fit in our improved pairwise approach.

Fig S4: Performances of the PIC *vs.* improved pairwise methods in 6 hypothetical conditions of Table 1. We used trees with proportion of duplication events of 0.8. Results are for asymmetric model. PICs: Phylogenetic independent contrasts. *: Bonferroni-corrected significant *P* values. We used a quadratic fit curve in our improved pairwise comparisons.

Fig S5: Comparisons of the approaches for an asymmetric trait jump model. PICs: Phylogenetic independent contrasts. We here trees with few duplication events, i.e. of 0.2. A quadratic fit curve is used for the improved pairwise approach. *: Bonferroni-corrected significant *P* values. For a null uniform model (A-B), equal proportions, i.e. 30% of the speciation and duplication nodes experienced jump in trait. For the ortholog conjecture scenario, we used higher proportions of trait jumps following duplication nodes than speciation nodes, i.e. 50% and 20% in this case (C-D). Both the approaches recovered the expected signals in all the conditions.

Fig S6: Performances of the PIC *vs.* improved pairwise comparisons for an asymmetric trait jump model. PICs: Phylogenetic independent contrasts. We here used a proportion of duplication events of 0.5. We used quadratic fit curve in our improved pairwise comparisons. *: Bonferroni-corrected significant *P* values. (A-B) For a null uniform model, 30% of the nodes experienced trait jump following speciation, and duplication events. PICs or correlations of both events are drawn from the same distribution in such case. For the ortholog conjecture simulation (C-D), 20% of the speciation and 50% of the duplication nodes experienced jumps in trait. Both approaches recover expected signals in all cases.

656 **Fig S7**: PICs vs. improved pairwise comparisons under an asymmetric trait jump
657 model. PICs: Phylogenetic independent contrasts. We here used proportion of
658 duplication events of 0.8. *: Bonferroni-corrected significant *P* values. A quadratic fit
659 curve is used for the improved pairwise comparisons. 30% of the speciation and
660 duplication nodes experienced jumps in trait under a null (A-B), while 20% of the
661 speciation and 50% of the duplication nodes experienced jumps under an ortholog
662 conjecture simulation (C-D). Both the approaches recovered expected signals in all
663 cases.

664

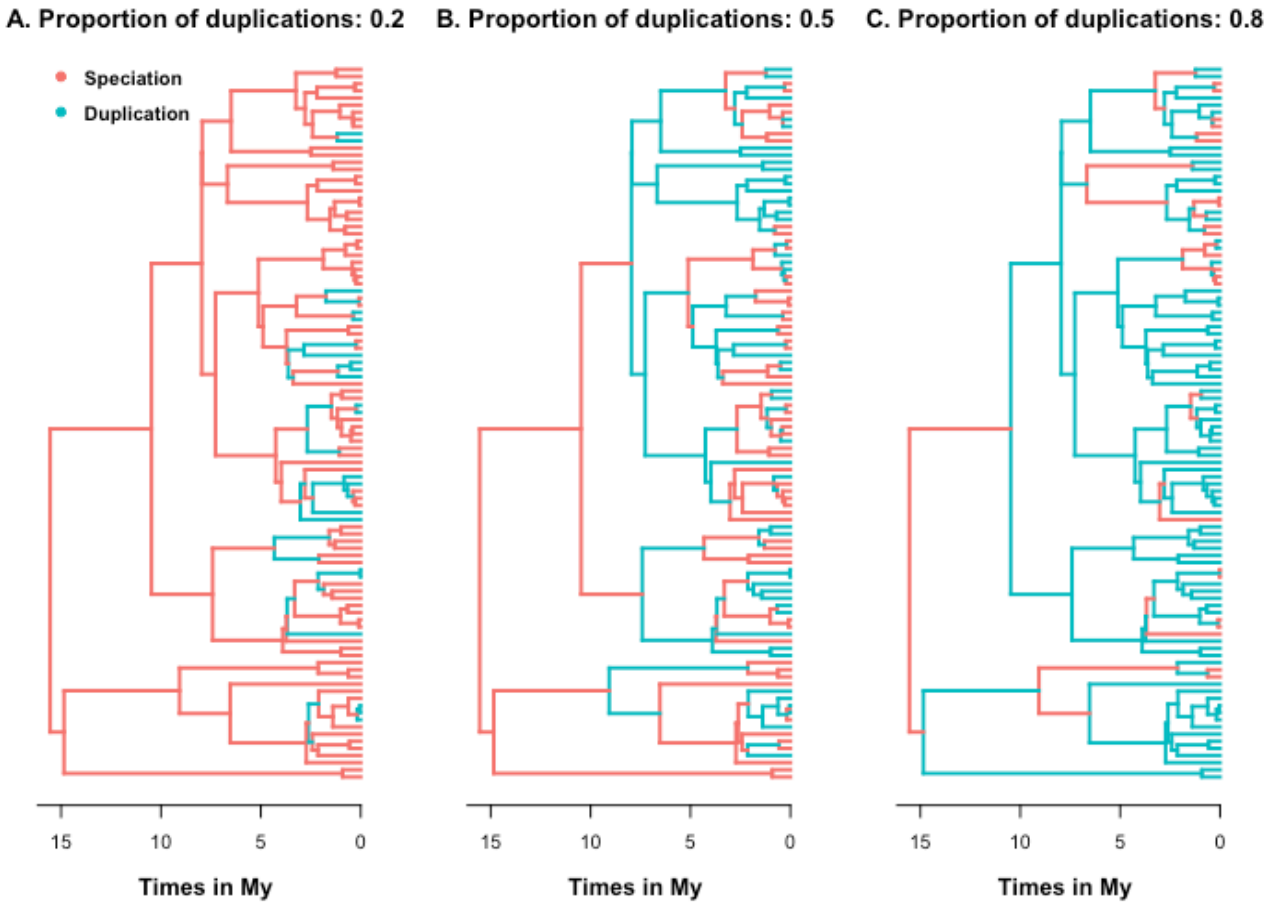
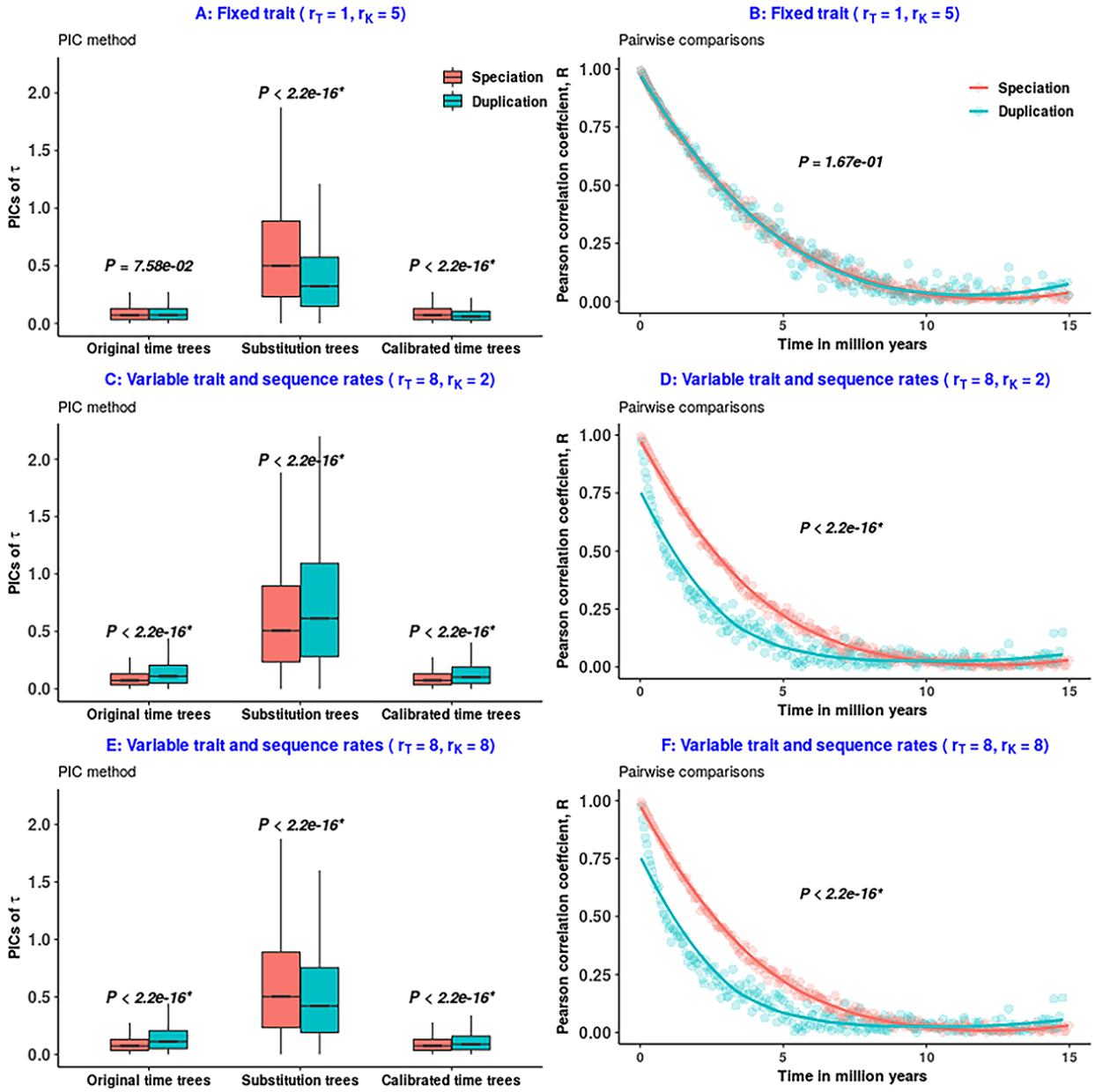

A. Proportion of duplications: 0.2
B. Proportion of duplications: 0.5
C. Proportion of duplications: 0.8
Speciation
Duplication
15
10
5
0
Times in My

Fig S1

670



671

672

673 **Fig S2**

674

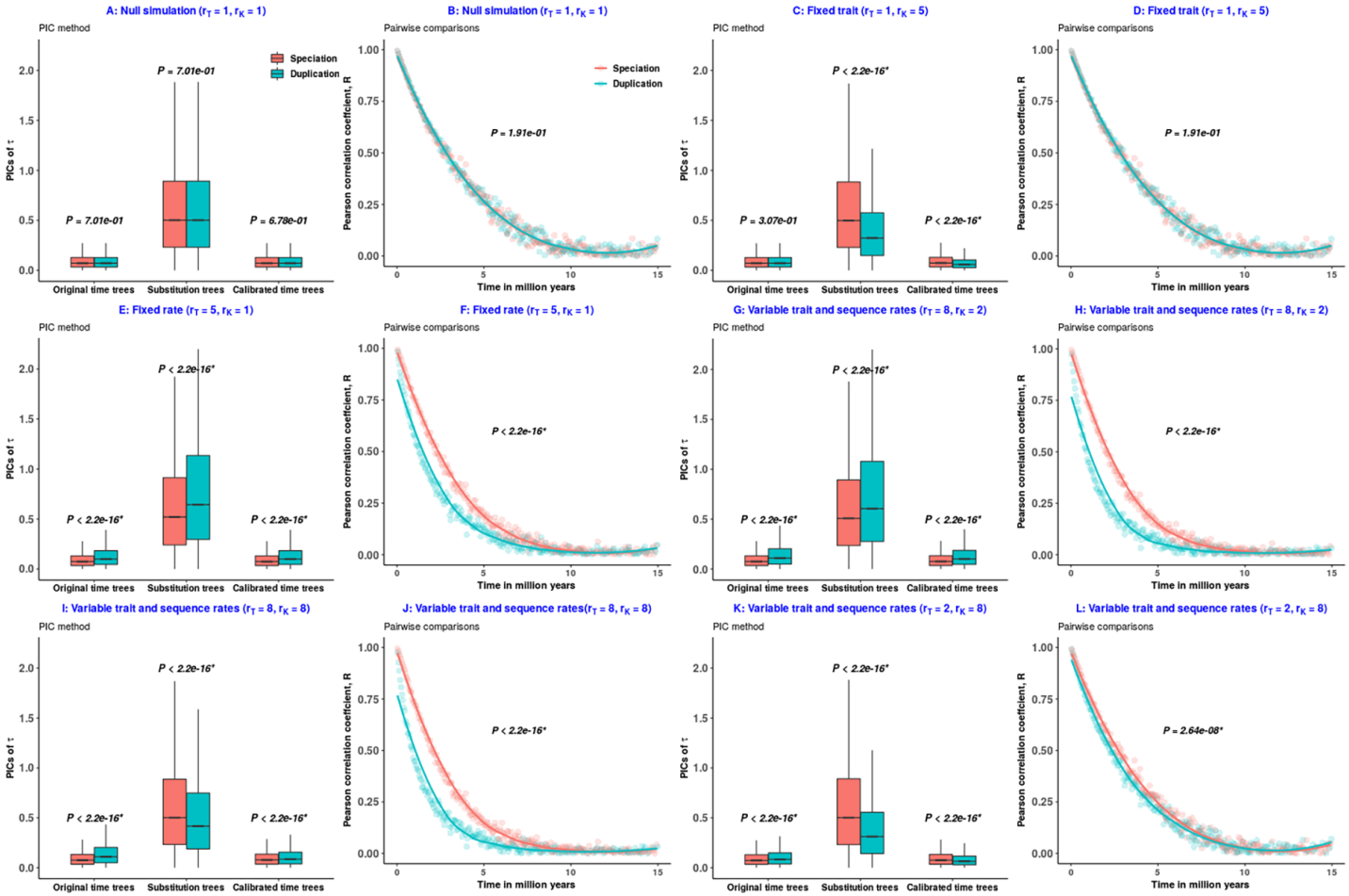

A: Null simulation (r_T = 1, r_K = 1)
B: Null simulation (r_T = 1, r_K = 1)
C: Fixed trait (r_T = 1, r_K = 5)
D: Fixed trait (r_T = 1, r_K = 5)
E: Fixed rate (r_T = 5, r_K = 1)
F: Fixed rate (r_T = 5, r_K = 1)
G: Variable trait and sequence rates (r_T = 8, r_K = 2)
H: Variable trait and sequence rates (r_T = 8, r_K = 2)
I: Variable trait and sequence rates (r_T = 8, r_K = 8)
J: Variable trait and sequence rates(r_T = 8, r_K = 8)
K: Variable trait and sequence rates (r_T = 2, r_K = 8)
L: Variable trait and sequence rates (r_T = 2, r_K = 8)
PIC method
Pairwise comparisons
Speciation
Duplication
PICs of τ
Pearson correlation coefficient, R
Time in million years
Original time trees
Substitution trees
Calibrated time trees
P = 7.01e-01
P = 6.78e-01
P = 1.91e-01
P = 3.07e-01
P < 2.2e-16*
P = 2.64e-08*

Fig S3

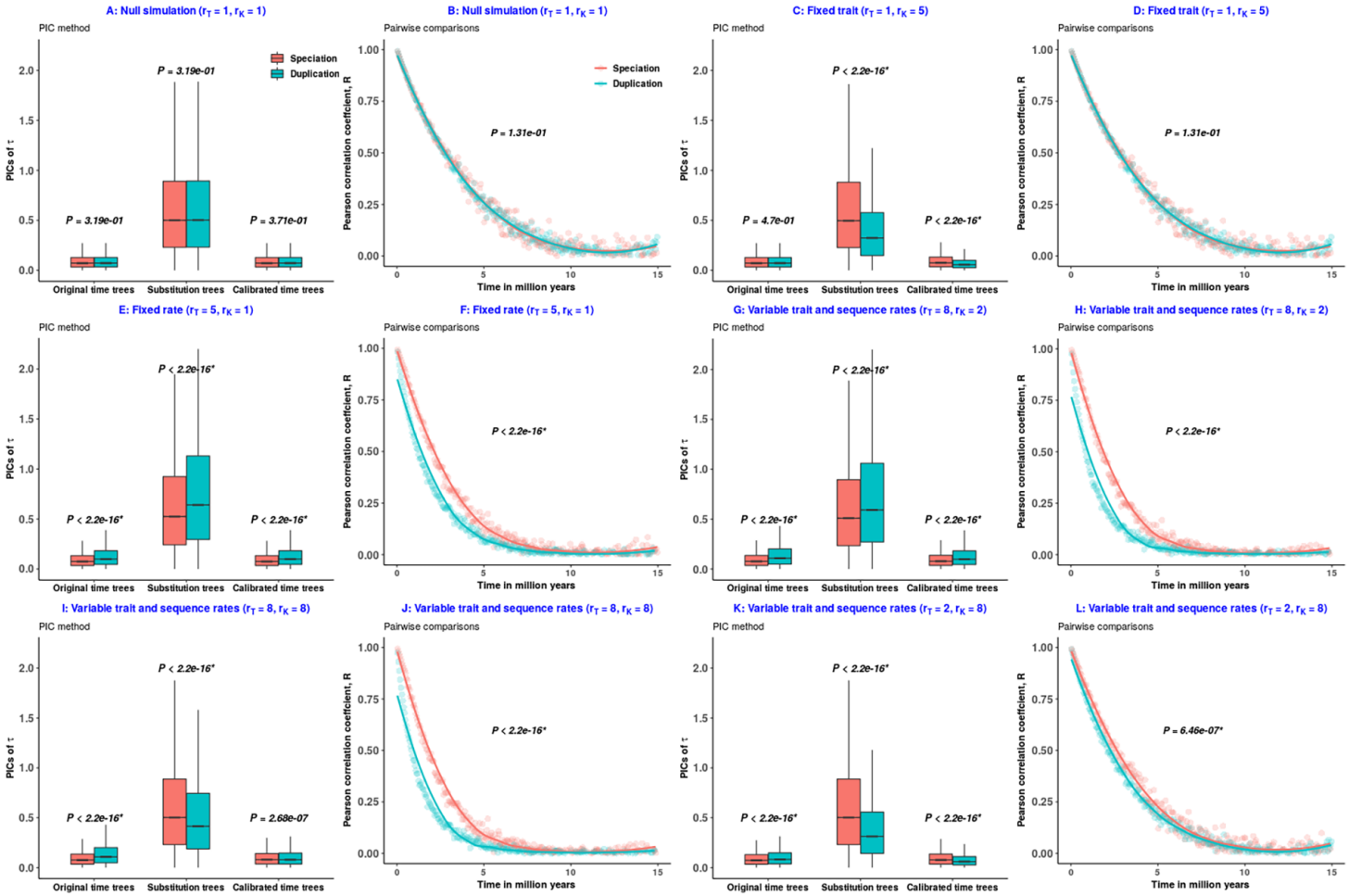

A: Null simulation (r_T = 1, r_K = 1)
B: Null simulation (r_T = 1, r_K = 1)
C: Fixed trait (r_T = 1, r_K = 5)
D: Fixed trait (r_T = 1, r_K = 5)
E: Fixed rate (r_T = 5, r_K = 1)
F: Fixed rate (r_T = 5, r_K = 1)
G: Variable trait and sequence rates (r_T = 8, r_K = 2)
H: Variable trait and sequence rates (r_T = 8, r_K = 2)
I: Variable trait and sequence rates (r_T = 8, r_K = 8)
J: Variable trait and sequence rates (r_T = 8, r_K = 8)
K: Variable trait and sequence rates (r_T = 2, r_K = 8)
L: Variable trait and sequence rates (r_T = 2, r_K = 8)
PIC method
Pairwise comparisons
Speciation
Duplication
PICs of τ
Pearson correlation coefficient, R
Time in million years
Original time trees
Substitution trees
Calibrated time trees
P = 3.19e-01
P = 3.71e-01
P = 1.31e-01
P = 4.7e-01
P < 2.2e-16*
P = 2.68e-07
P = 6.46e-07*

Fig S4

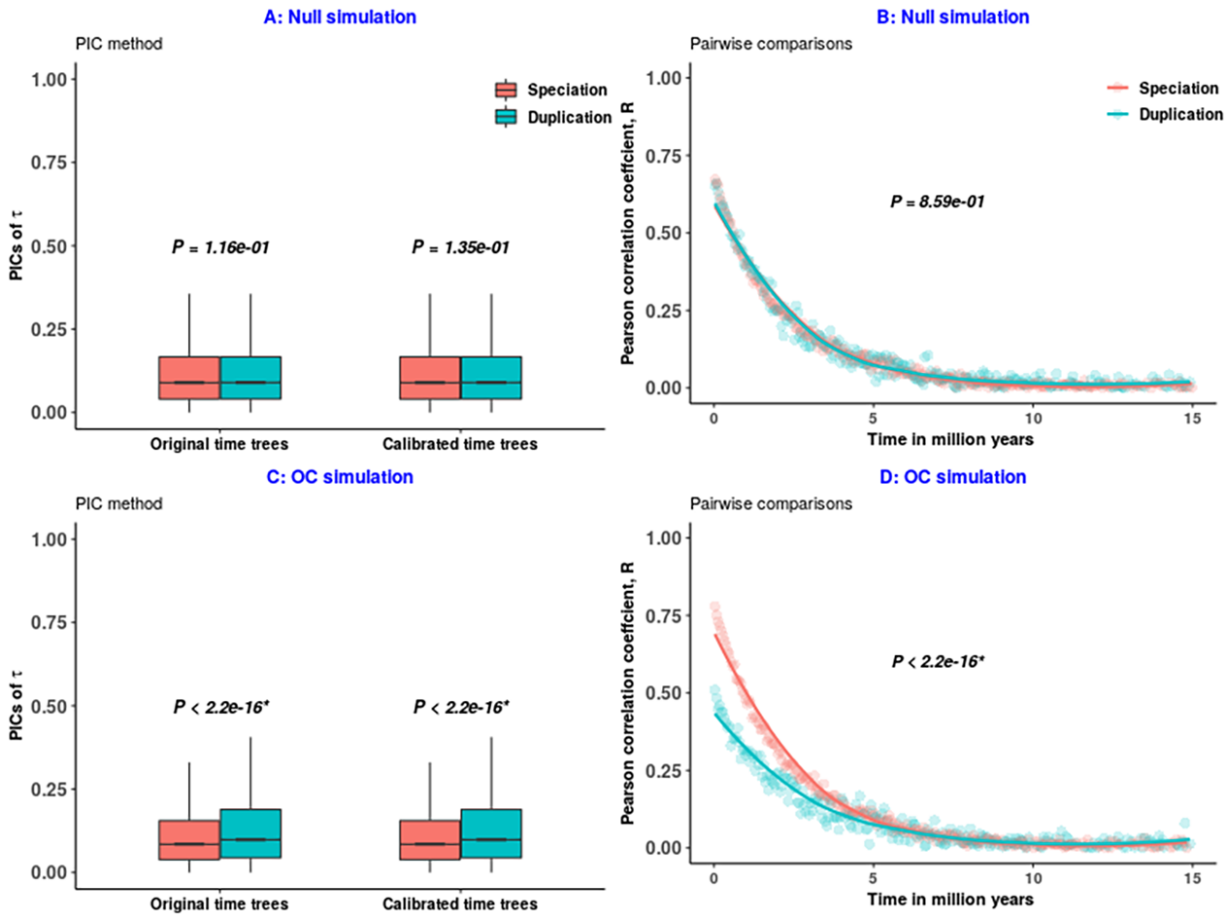

A: Null simulation
PIC method
PICs of τ
1.00
0.75
0.50
0.25
0.00
Speciation
Duplication
P = 1.16e-01
P = 1.35e-01
Original time trees
Calibrated time trees
B: Null simulation
Pairwise comparisons
Pearson correlation coefficient, R
P = 8.59e-01
0
5
10
15
Time in million years
C: OC simulation
PIC method
P < 2.2e-16*
P < 2.2e-16*
D: OC simulation
Pairwise comparisons
P < 2.2e-16*

Fig S5

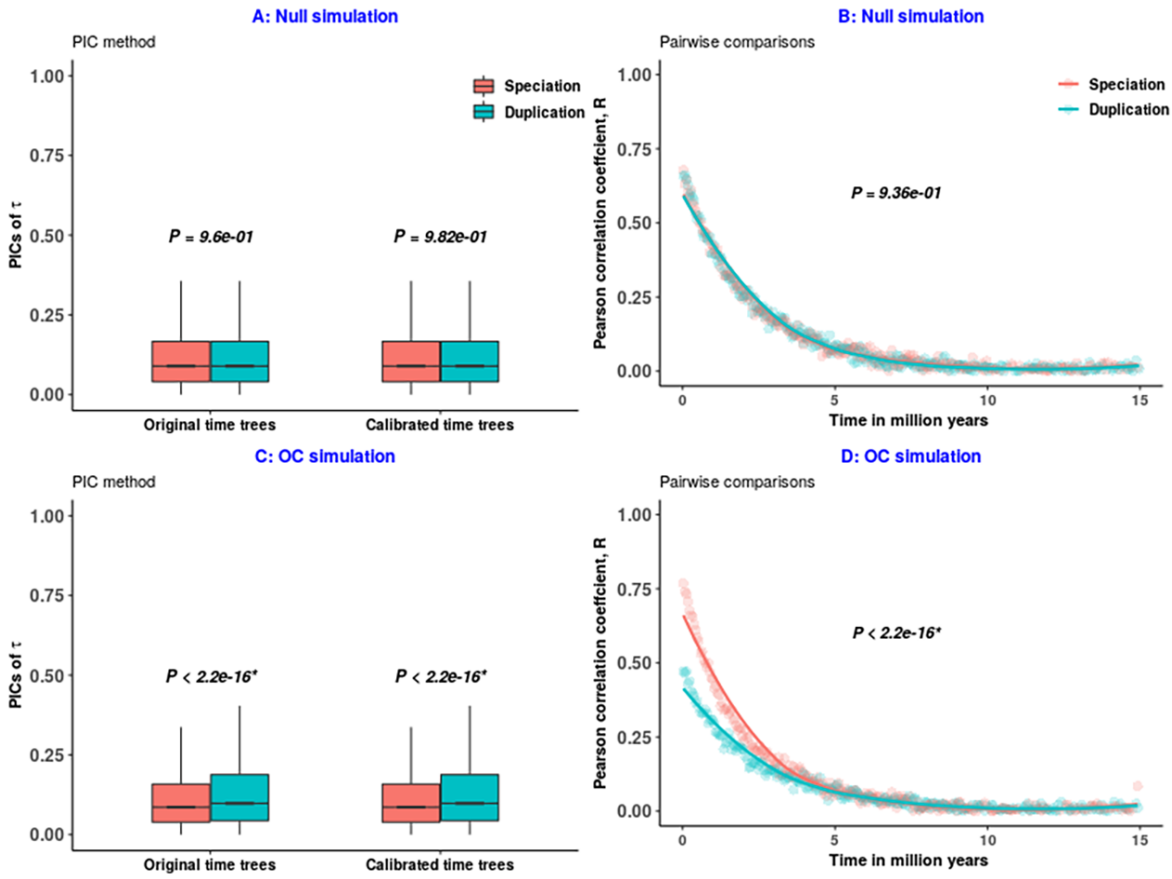

A: Null simulation
PIC method
Speciation
Duplication
PICs of τ
1.00
0.75
0.50
0.25
0.00
P = 9.6e-01
P = 9.82e-01
Original time trees
Calibrated time trees
B: Null simulation
Pairwise comparisons
Speciation
Duplication
Pearson correlation coefficient, R
P = 9.36e-01
0
5
10
15
Time in million years
C: OC simulation
PIC method
PICs of τ
P < 2.2e-16*
P < 2.2e-16*
Original time trees
Calibrated time trees
D: OC simulation
Pairwise comparisons
Pearson correlation coefficient, R
P < 2.2e-16*
Time in million years

Fig S6

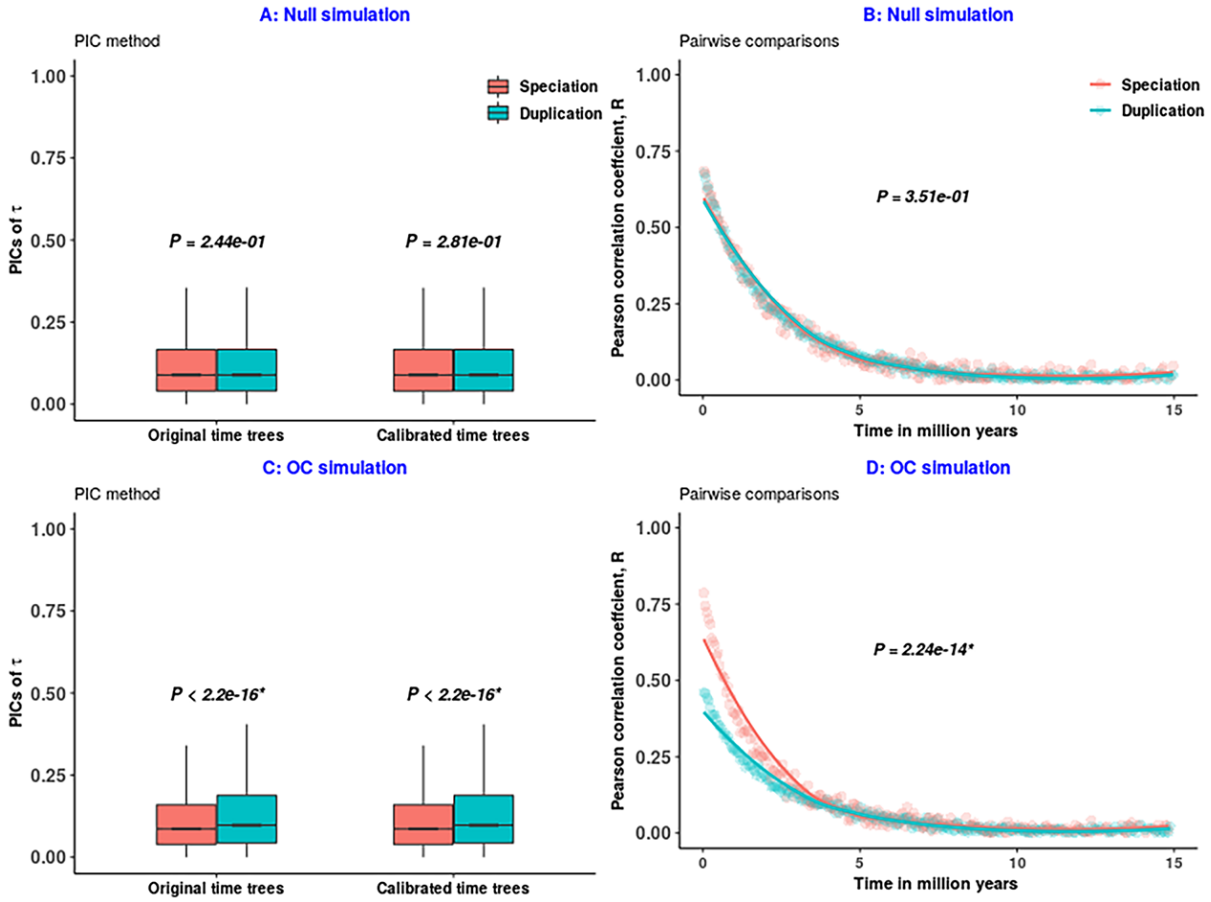

A: Null simulation
PIC method
Speciation
Duplication
PICs of τ
1.00
0.75
0.50
0.25
0.00
P = 2.44e-01
P = 2.81e-01
Original time trees
Calibrated time trees
B: Null simulation
Pairwise comparisons
Pearson correlation coefficient, R
P = 3.51e-01
0
5
10
15
Time in million years
C: OC simulation
P < 2.2e-16*
D: OC simulation
P = 2.24e-14*

Fig S7