Supplementary Materials

LS³: A method for improving phylogenomic inferences when evolutionary rates are heterogeneous among taxa

Carlos J. Rivera-Rivera[1,2] and Juan I. Montoya-Burgos[1,2]

[1]*Department of Genetics and Evolution, University of Geneva, 30 quai Ernest Ansermet, 1211 Geneva 4, Switzerland*
[2]*Institute of Genetics and Genomics in Geneva, University of Geneva Medical Center, 1 rue Michel-Servet, 1211 Geneva 4, Switzerland*

This file includes:
1. Supplementary Figures:

1. Supplementary Material Figures

Figure S1. Input topologies for producing the simulated evolution dataset followed a common backbone and 50 different arrangements of long branches in the fast-evolving lineage.

Examples of the different long-branch arrangements assigned to the taxa of Lineage C for producing the simulated data of this study. The final dataset is composed of 200 simulated gene alignments of 2,000 bp for each of the 50 cases of long branches (15 cases with two long branches, 20 cases with three long branches, and 15 cases with four long branches) for a total of 10,000.
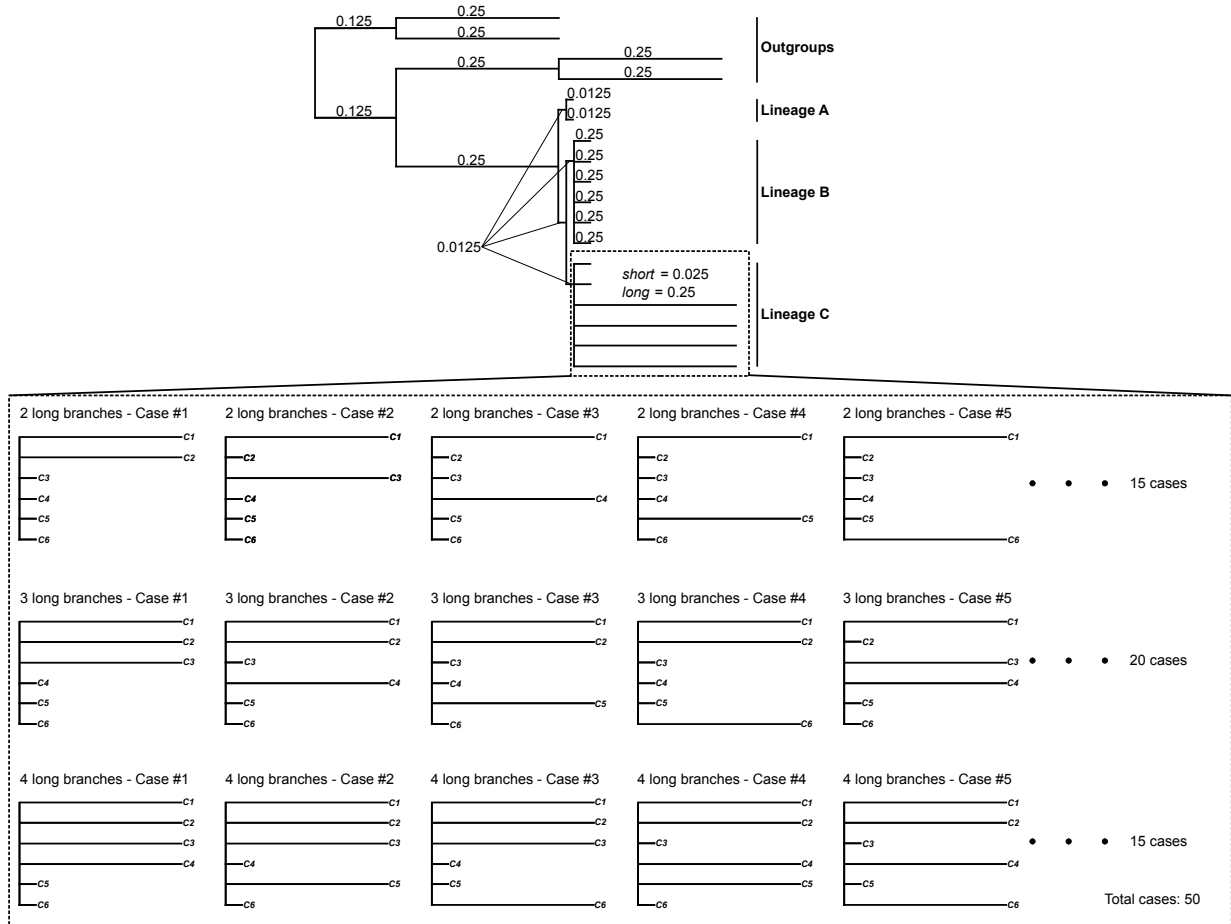
Figure S2. LBA artifact in the positionning of Glires in the mammalian tree and its resolution upon removing LS³-flagged sequences and genes.

A dataset of 57 genes leading to the non-monophyly and missplacement of Glires within the mammalian tree due to LBA was assembled, named the "only LBA" (OL) dataset. This dataset was used to assess the efficiency of the LS³ method in flagging data that could reinforce LBA artifacts. (A) Maximum Likelihood phylogeny obtained with the OL dataset showing the non-monophyly and missplacement of Glires (in red) relative to the Laurasiatheria (in blue) and the Primates (in yellow). Note the strong bootstrap support for the wrong grouping of Laurasiatheria with Primates. (B) ML phylogeny obtained after removing sequences and genes that LS³ flagged in the OL dataset (OL_{LS³}) showing the correct monophyly of Glires and their sister group relationshisp with Primates, yet with poor boostraop support. (C) ML phylogeny obtained after adding to the OL dataset two genes that produced the correct species tree, the OL+2 dataset; the boostrap support for the wrong relationshisps are reduced, but the topology is incorrect. (D) The ML phylogeny obtained with the OL_{LS³}+2 dataset, in which the LS³-flagged data has been removed, resulted in a correct phylogeny with strong boostratp supports. Only bootstrap support values lower than 100 are shown.