Supplementary Figures

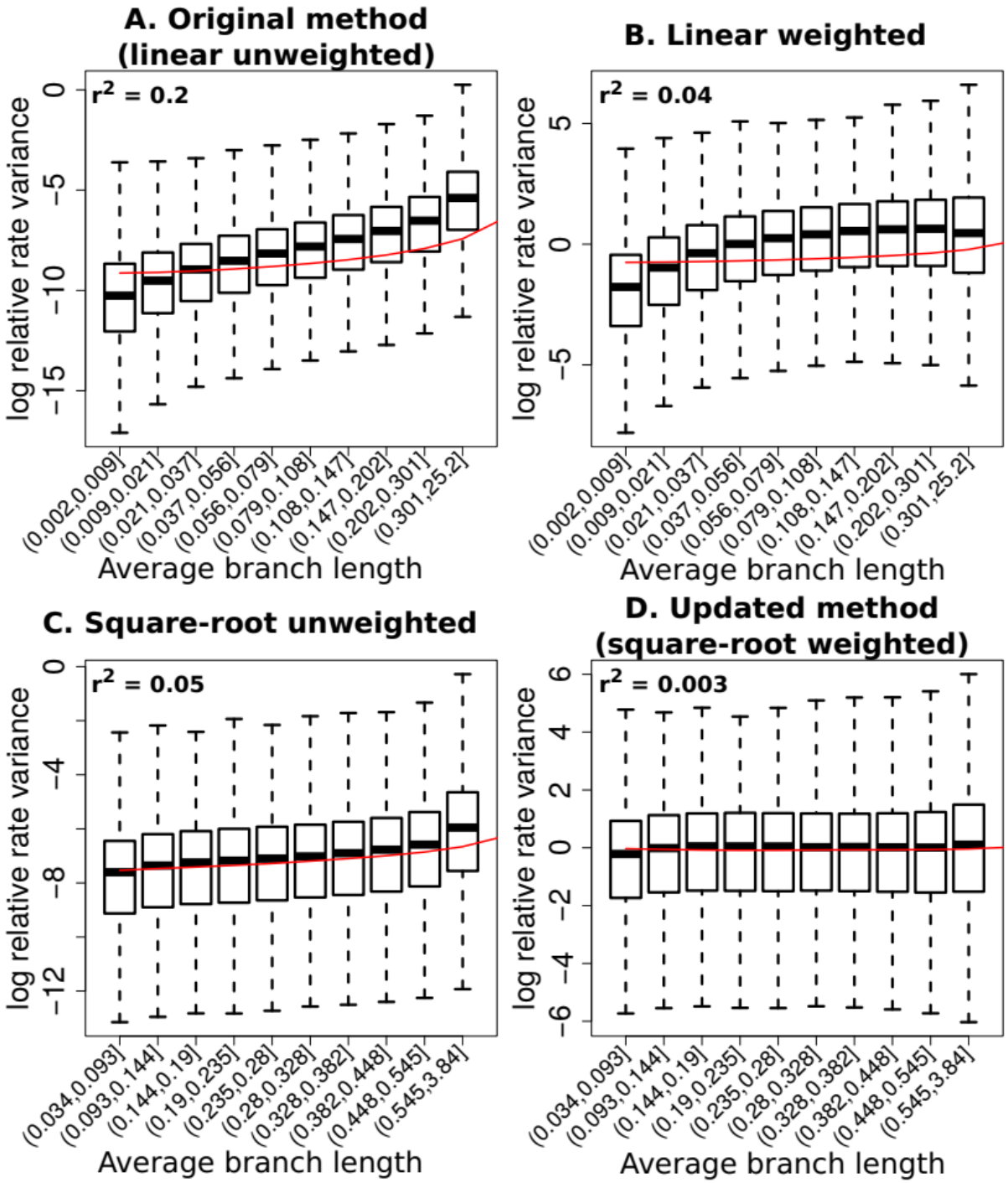


Fig S1. Comparison of mean-variance trends in relative rates computed using original, updated and intermediate methods. A corresponds to original method, D the updated method. Panels B and C reflect methods that are intermediate to the updated method, with no transformation (B), and no weighted regression (C).

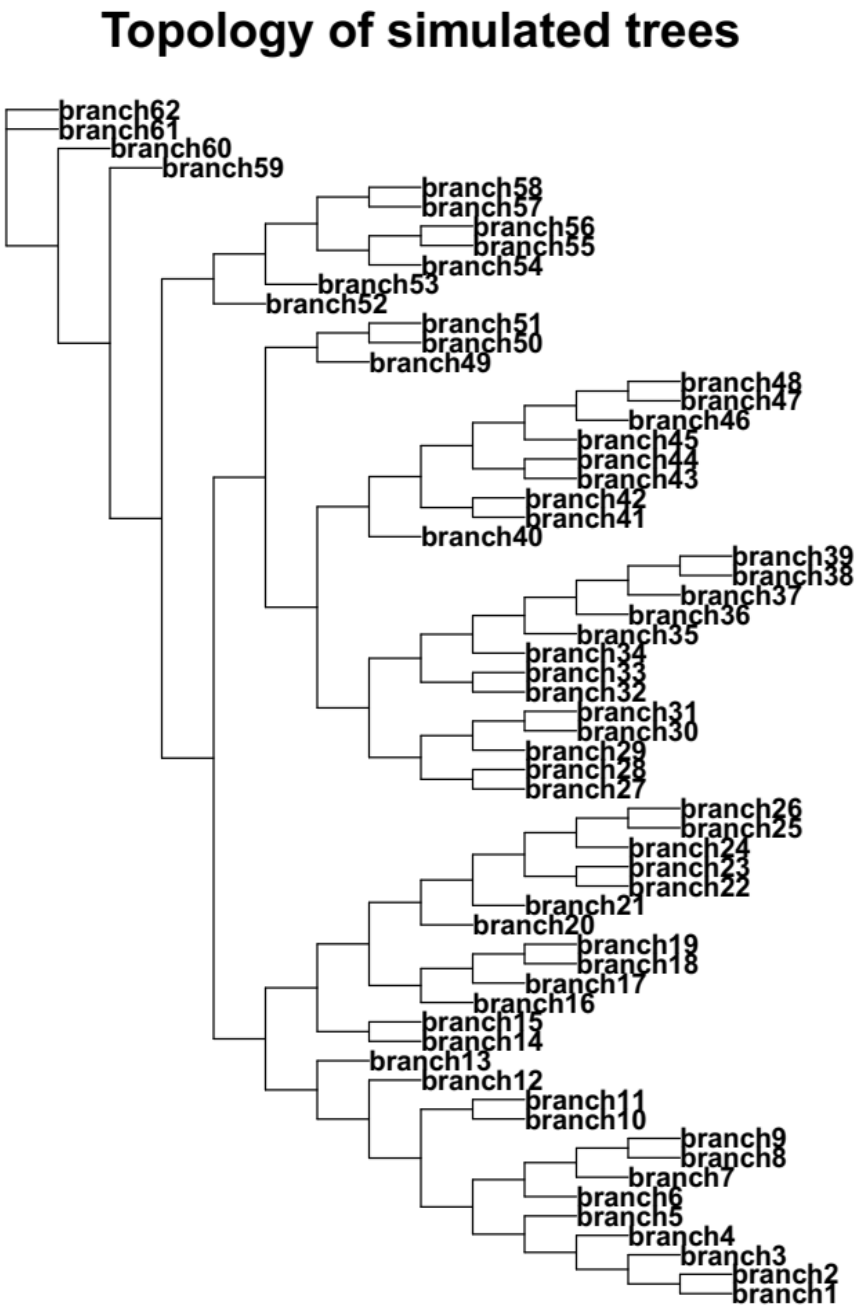


Fig S2. Topology describing the relationship between branches of simulated trees. The topology is constructed based on the relationships of 62 mammalian species as reported in the UCSC genome browser.

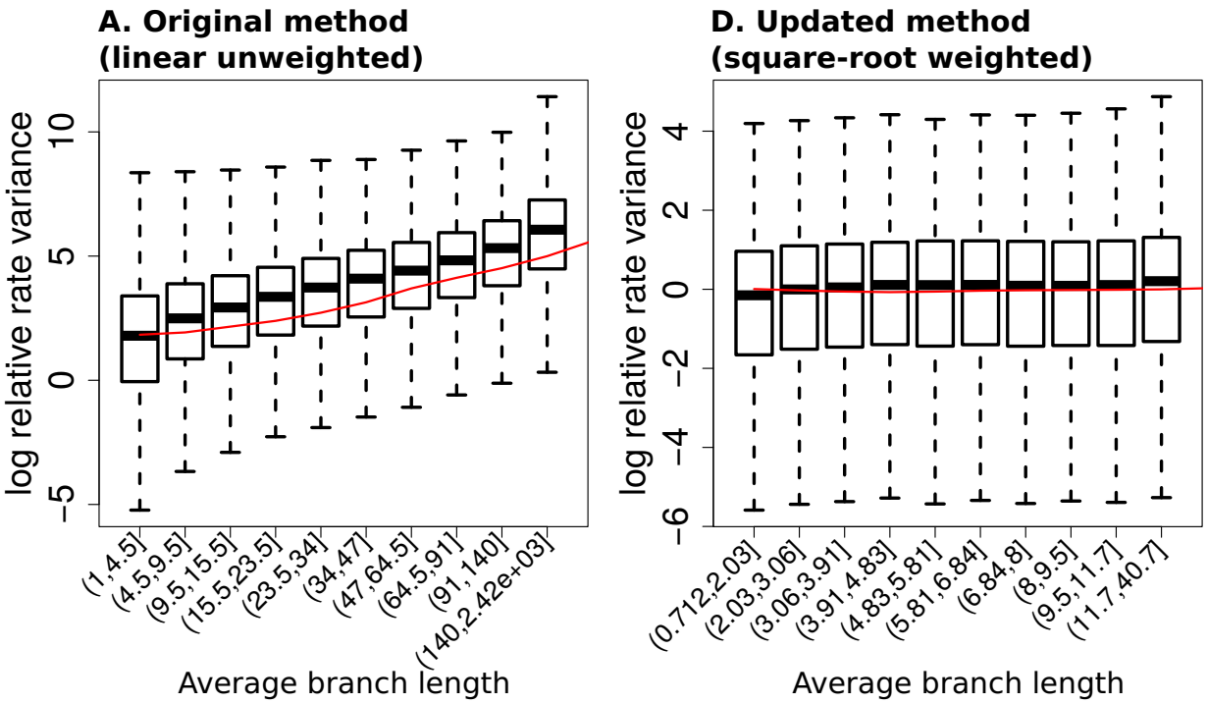


Fig S3. Mean-variance trends in relative rates on branches of simulated phylogenetic trees computed using the two methods. The original method (A) produces heteroscedastic relative rates that show a strong mean-variance trend, whereas relative rates calculated using the updated method (B) show constant variance across branches of different lengths.

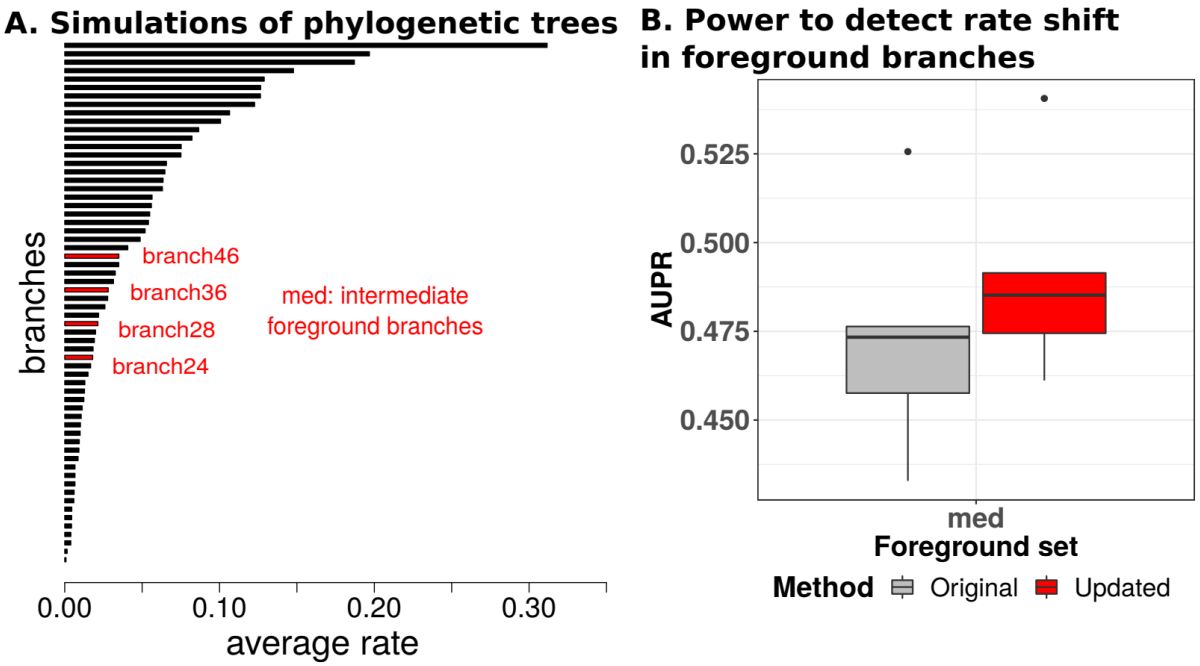
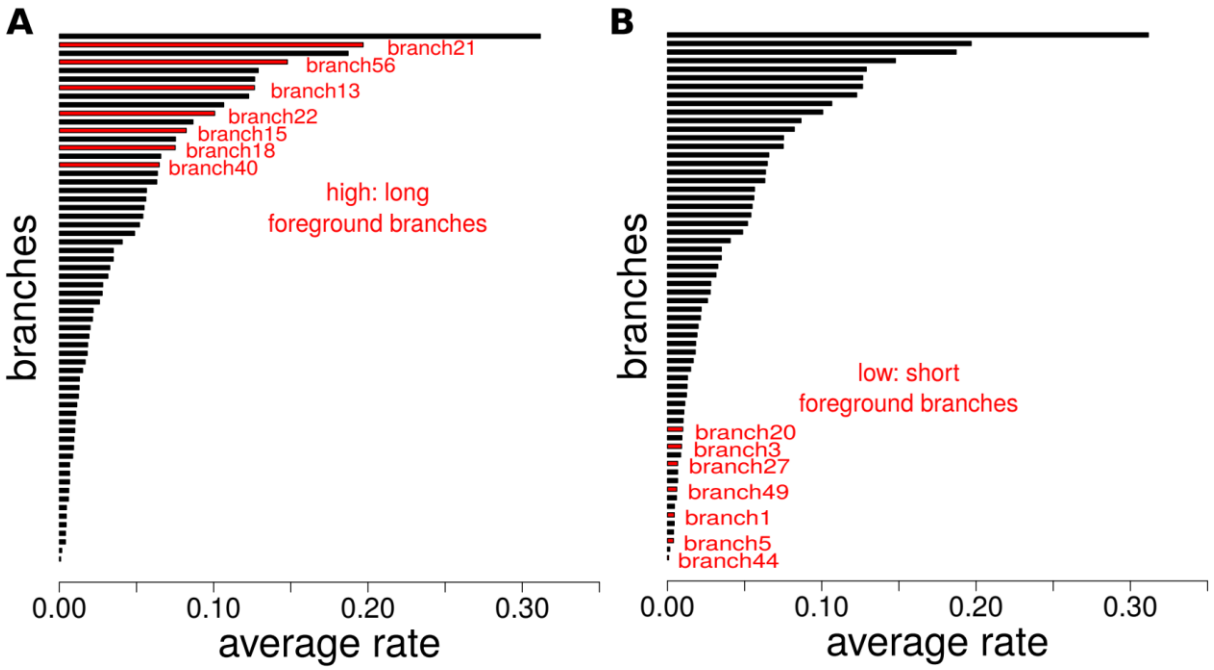


Fig S4. Simulations of phylogenetic trees with foreground branches of intermediate length A. Foreground branches (in red), used for simulating convergent acceleration on intermediately long branches. B. Comparison of power between the two methods to detect convergent rate acceleration on intermediately long foreground branches. Across five independent simulations of control trees and positive trees, we measured the area under the precision-recall curve (AUPR) to precisely detect positive trees using the foreground acceleration score. The AUPR distributions obtained using the updated method to calculate relative rates are significantly elevated compared to the original method.



C

Number of foreground branches	Long foreground branches set	Short foreground branches set
4	21, 56, 13, 22	44, 5, 1, 49
5	21, 56, 13, 22, 15	44, 5, 1, 49, 27
6	21, 56, 13, 22, 15, 18	44, 5, 1, 49, 27, 3
7	21, 56, 13, 22, 15, 18, 40	44, 5, 1, 49, 27, 3, 20

Fig S5. Simulations of phylogenetic trees with varying numbers of foreground branches. Red branches correspond to the seven foreground branches chosen for simulating trees showing convergent rate acceleration on long (A) and short (B) foreground branches. The foreground branch sets of simulated trees used for comparing the power to detect foreground acceleration across different numbers of long and short branches are given in (C).

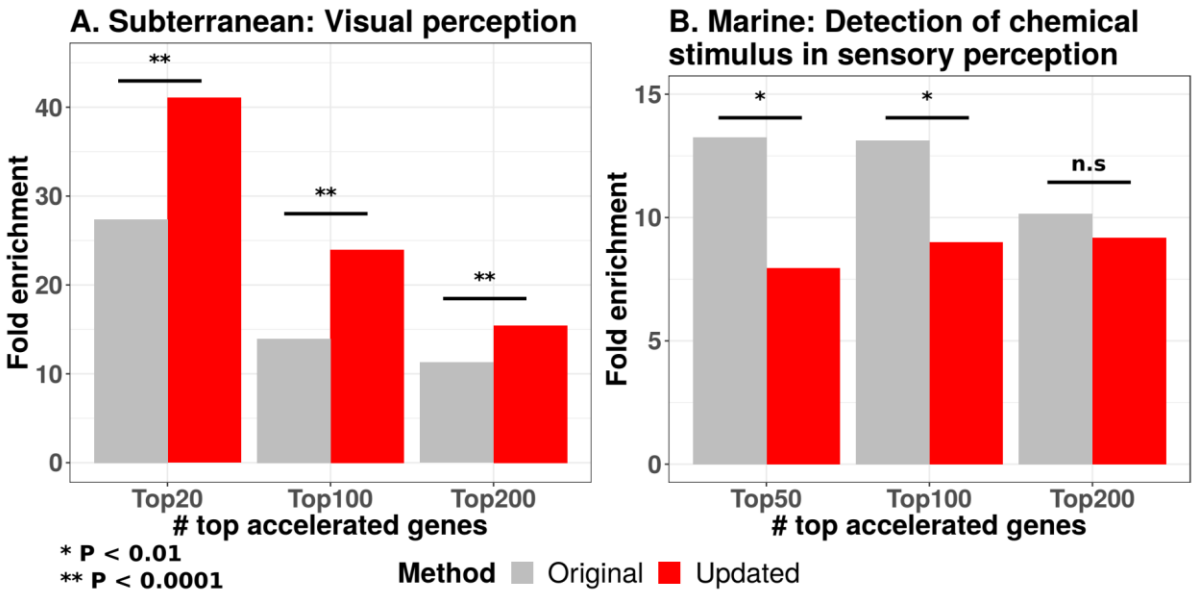


Figure S6: Comparison of fold enrichments of top enriched terms associated with two convergent phenotypes, detected across methods. A. The barplot compares the fold enrichment for the visual perception GO term across top subterranean accelerated genes discovered by the original and the updated method. Across three different cutoffs for the number of top genes, the enrichment was consistently higher for genes discovered by the updated method (p-value reported in hypergeometric test). Visual perception was the top significant term reported across all the subterranean-accelerated gene sets analyzed. B. The same analysis was repeated with the top enriched term in marine-accelerated genes, namely Detection of chemical stimulus in sensory perception. The fold enrichments were significantly higher across the Top50, and 100 genes respectively. We chose the marine accelerated Top50 instead of Top20, as no terms were enriched across either method in the Top20 gene list.