

Description of Additional Supplementary Files

File Name: Supplementary Data 1

Description: The full list of predicted and true log-likelihood differences obtained from the example dataset for all 2,462 single-step SPR moves. The 'prune_name' and the 'rgft_name' columns represent the node name (either internal or terminal) that defines the prune and regraft actions, respectively. The 'orig_ds_ll' column represents the empirical log-likelihood of the starting tree (same for all rows). The 'll' column represents the empirical loglikelihood of each neighboring tree. The 'dif_ll' column represents the difference between the empirical log-likelihood of each neighboring tree and that of the starting tree. The 'dif_ll_pred' column represents the loglikelihood difference as predicted by the machine-learning algorithm.