Exercises – Phylogenetics

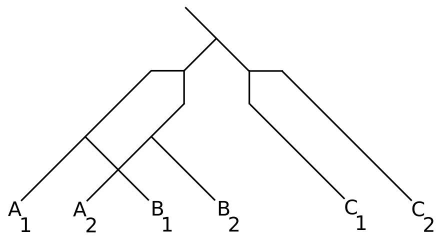
Universität Bielefeld, SS 2019
Dr. Roland Wittler, M. Sc. Tizian Schulz
https://gi.cebitec.uni-bielefeld.de/Teaching/2019summer/Phylogenetik

Exercise Sheet 2 — 18.04.2019

Due: 25.04.2019

Task 1 Gene and species trees. (2 points)

(a) Give all pairs of paralogous and orthologous genes.



(b) Find out the meaning of *xenolog*. Draw a picture and explain the term.

Task 2 Gene tree / species tree reconciliation. (3 points)

Consider the following gene tree G and species tree S.

$$\begin{aligned} G &= ((g_4,(g_5,g_6)g_3)g_2,g_7)g_1; \\ S &= ((A,B)n_2,C)n_1; \end{aligned}$$

Let g_4 be present in species A. Genes g_5 and g_6 are in B and gene g_7 is found in C.

(a) Draw G and S. Label the internal nodes and leaves. Furthermore label the leaves of G with the corresponding names of the species.

(b) Calculate $\gamma(g_1)$, $\gamma(g_2)$ und $\gamma(g_3)$. Calculate the $M(.)$ for every node (including the leaves) in G and draw the matching arrows (lecture notes: Figure 2.6) from G to S. Decide for every internal node whether it corresponds to a duplication or speciation.

(c) Draw the resulting scenario (similar to Figure 2.5 (right) in the lecture notes).
Hint The scenario might include deletions.

Task 3 Last common ancestor. (2 points)

Consider a rooted tree $G = (V, E)$. Write an algorithm in *pseudo code* that determines the *last common ancestor LCA(u,v)* for two given nodes U and V. What is the asymptotic worst case running time of your algorithm?

Hint: You can find examples for pseudo code in the lecture notes (page 13, 22/23, 28, 35, 36 etc.)
Don't get confused! *lowest* and *last/least common ancestor* are basically the same. In phylogenetics, the term *last/least common ancestor* is more common since we're looking at the evolutionary time of events. However, if you think about vertices and trees in graph theory, the term *lowest common ancestor* is used since there is no time involved in the analysis. Luckily, the abbreviations of the words are the same.