

Übungen zur Phylogenetik Vorlesung

Universität Bielefeld, WS 2014/2015, Dr. Roland Wittler, Kevin Lamkiewicz

<http://wiki.techfak.uni-bielefeld.de/gi/Teaching/2014winter/Phylogenetik>

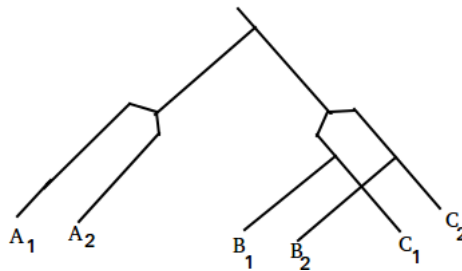
Blatt 2 vom 15.10.2014

Abgabe in einer Woche zu Beginn der Vorlesung oder vorab bei der Tutorin oder beim Veranstalter.

Aufgabe 1 Gen- und Speziesbäume.

(2 Punkte)

- (a) Gib alle Paare von paralogen und orthologen Genen im Bild an.



- (b) Informiere dich, was die Begriffe *in-paralog* und *out-paralog* bedeuten. Erkläre kurz den wesentlichen Unterschied.

Aufgabe 2 Gene tree / species tree reconciliation.

(3 Punkte)

Gegeben sei der Genbaum G und der Speziesbaum S in Newicknotation.

$$G = (g_4, ((g_6, g_7)g_3, g_5)g_2)g_1;$$

$$S = (X, (Y, Z)s_2)s_1;$$

Gen g_4 sei in Spezies X enthalten. Die Gene g_5 und g_6 sind in Y und das Gen g_7 in Z .

- (a) Zeichne G und S nebeneinander. Beschrifte jeweils die inneren Knoten und die Blätter. Zusätzlich sollen in G die Blätter mit den entsprechenden Namen der Spezies beschriftet werden.
- (b) Berechne $\gamma(g_1)$, $\gamma(g_2)$ und $\gamma(g_3)$. Gib weiterhin $M(\cdot)$ für alle Knoten (auch die Blätter) in G an und zeichne entsprechende Pfeile von G nach S in deinem Bild. Handelt es sich bei den inneren Knoten jeweils um eine Duplikation oder um eine Speziation?
- (c) Zeichne das erhaltene Szenario aus Speziesbaum und enthaltenem Genbaum, ähnlich zu dem in der Vorlesung bzw. Bild 2.5 (rechts) im Skript.

Tipp: Das Szenario enthält eine Genlöschung.

Aufgabe 3 Last common ancestor.

(2 Punkte)

Gegeben sei ein gewurzelter Baum $G = (V, E)$. Wie lässt sich für zwei gegebene Knoten u und v der *last common ancestor* (LCA) finden? Gib einen Algorithmus in *Pseudocode* an, der den $LCA(u, v)$ bestimmen kann. Welche Laufzeit hat dein Algorithmus im *worst case*?

Hinweis: Beispiele für Pseudocode lassen sich überall im Skript (Seite 13, 22, 23, 28, etc.) finden.

Nicht verwirren lassen! *lowest* und *last common ancestor* sind im Prinzip dasselbe. Es ist in der Phylogenetik üblich vom *last common ancestor* zu reden, wenn man den letzten gemeinsamen Vorfahren meint. Abstrahiert man jedoch auf Knoten in einem Baum entspricht dies dem Begriff *lowest common ancestor* (niedrigster gemeinsamer Vorfahrknoten) aus der Graphentheorie. Glücklicherweise sind die Abkürzungen der beiden Begriffe identisch.