

10. Januar 2007

Abgabe: 17. Januar 2007

Aufgabe 1: Schreiben Sie ein Programm, das für das Felsenstein-84-Substitutionmodell nach Eingabe von zwei Sequenzen den Maximum-Likelihood-Schätzer für die zeitliche Distanz der beiden Sequenzen, die Nukleotidwahrscheinlichkeiten und für das Verhältnis zwischen den Transitions- und Transversionsraten ausgibt. Erproben Sie Ihr Programm an simulierten und echten Daten.

Aufgabe 2: Simulieren Sie jeweils 10 Sequenzpaare im Felsenstein-84-Modell und im Jukes-Cantor-Modell. Verwenden Sie dabei sehr verschiedene Parameterwerte und Sequenzlängen. Lassen Sie dann AIC, BIC und den Likelihood-Quotiententest für jedes Sequenzpaar raten, mit welchem der beiden Substitutionsmodelle es erzeugt wurde.

Aufgabe 3: Schreiben Sie ein Programm, das nach Eingabe eines Baumes, dessen Blätter mit homologen DNA-Sequenzen beschriftet sind, die Likelihood des Baumes berechnet. Kombinieren Sie dabei ein Substitutionmodell Ihrer Wahl mit dem Modell der gammaverteilten Ratenfaktoren. Die Anzahl der Diskretisierungsschritte und der Formparameter der Gammaverteilung sowie die Parameter des Substitutionmodells sollen vom Benutzer eingegeben werden. Erproben Sie Ihr Programm an simulierten und/oder echten Daten.