

## 13. Übung „Angewandte Bioinformatik mit Perl und R“

1. Laden Sie wieder `golub`-Daten.
  - (a) Filtern Sie erneut die Gene mit den Kriterien (a), (b) und (c) aus der Aufgabe 3 der letzten Serie. Benutzen Sie die Funktion `hclust` um die Proben der Golubdaten mit diesen Genen zu clustern. Verwenden Sie unterschiedliche Distanzfunktionen (z.B. aus dem package `bioDist`) und vergleichen Sie die Ergebnisse (jeweils visuell).
  - (b) Wiederholen Sie dies für unterschiedliche Teilmengen signifikanter Gene und wiederum unterschiedliche Distanzen. Vergleichen Sie wiederum die Ergebnisse, was beeinflusst die Ergebnisse am stärksten.
  - (c) Verwenden Sie nun das k-means Verfahren (Funktion `kmeans`) für das Clustern. Vergleichen Sie erneut die Ergebnisse.
  - (d) Testen Sie einige weitere der genannten Clusterverfahren.
2. Versuchen Sie nun Genen zu clustern. Überlegen Sie sinnvolle Einschränkungen.