

Studi Kasus 6

Kevin Siswandi

2023-10-24

Pustaka

```
library(Biobase)
```

```
## Loading required package: BiocGenerics

##
## Attaching package: 'BiocGenerics'

## The following objects are masked from 'package:stats':
##
##   IQR, mad, sd, var, xtabs

## The following objects are masked from 'package:base':
##
##   anyDuplicated, aperm, append, as.data.frame, basename, cbind,
##   colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
##   get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,
##   match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,
##   Position, rank, rbind, Reduce, rownames, sapply, setdiff, sort,
##   table, tapply, union, unique, unsplit, which.max, which.min

## Welcome to Bioconductor
##
##   Vignettes contain introductory material; view with
##   'browseVignettes()'. To cite Bioconductor, see
##   'citation("Biobase)", and for packages 'citation("pkgname)".
```

Clustering

Muat data dari Montgomery dan Pickrell eSet

```
con =url("http://bowtie-bio.sourceforge.net/recount/ExpressionSets/montpick_eset.RData")
load(file=con)
close(con)
pdata=pData(montpick.eset)
edata=as.data.frame(exprs(montpick.eset))
fdata = fData(montpick.eset)
ls()
```

Lakukan clustering untuk dataset ini dengan tiga cara:

1. Tanpa perubahan apapun pada data
2. Filter semua gen dengan rowMeans di bawah 100.
3. Setelah transformasi log2 tanpa filtering.

Warnai sampel berdasarkan asal studi mereka (gunakan data fenotipe). Untuk mewarnai dapat memakai <https://www.rdocumentation.org/packages/rafalib/versions/1.0.0/topics/myplclust> sebagai referensi. Bagaimana apa yang kalian amati.

Kelompokkan sampel menggunakan k-means clustering setelah menerapkan metode log2 transform (pastikan untuk menambahkan 1). Tetapkan seed untuk hasil yang dapat direproduksi. Cobalah dengan beberapa jumlah kluster yang berbeda.

Sekarang pakai setting dua kluster di k-means, bandingkan dengan kluster yang didapatkan dengan `hclust` setelah dipotong dengan `cutree` menjadi dua kelompok.

Kluster mana yang paling cocok dengan label studi?