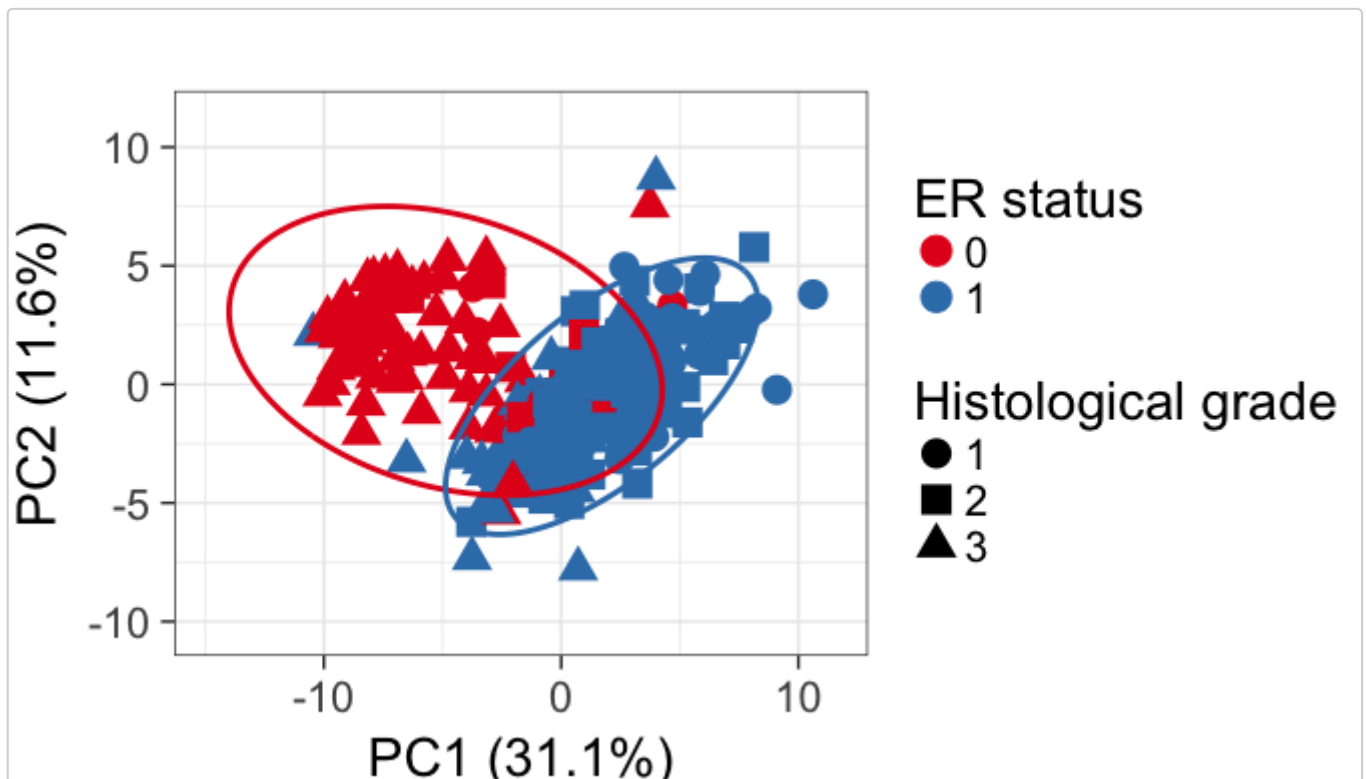# ClustVis examples

Tauno Metsalu

2018-12-03

This vignette shows some minimal examples how to generate plots using R package clustvis.

PCA example

```
library(clustvis)
file = "https://raw.githubusercontent.com/taunometsalu/ClustVis/master/datasets/nki.csv"
imp = importData(file)
proc = processData(imp)
pca = generatePCA(proc)
savePCA(pca, file = NA)
```
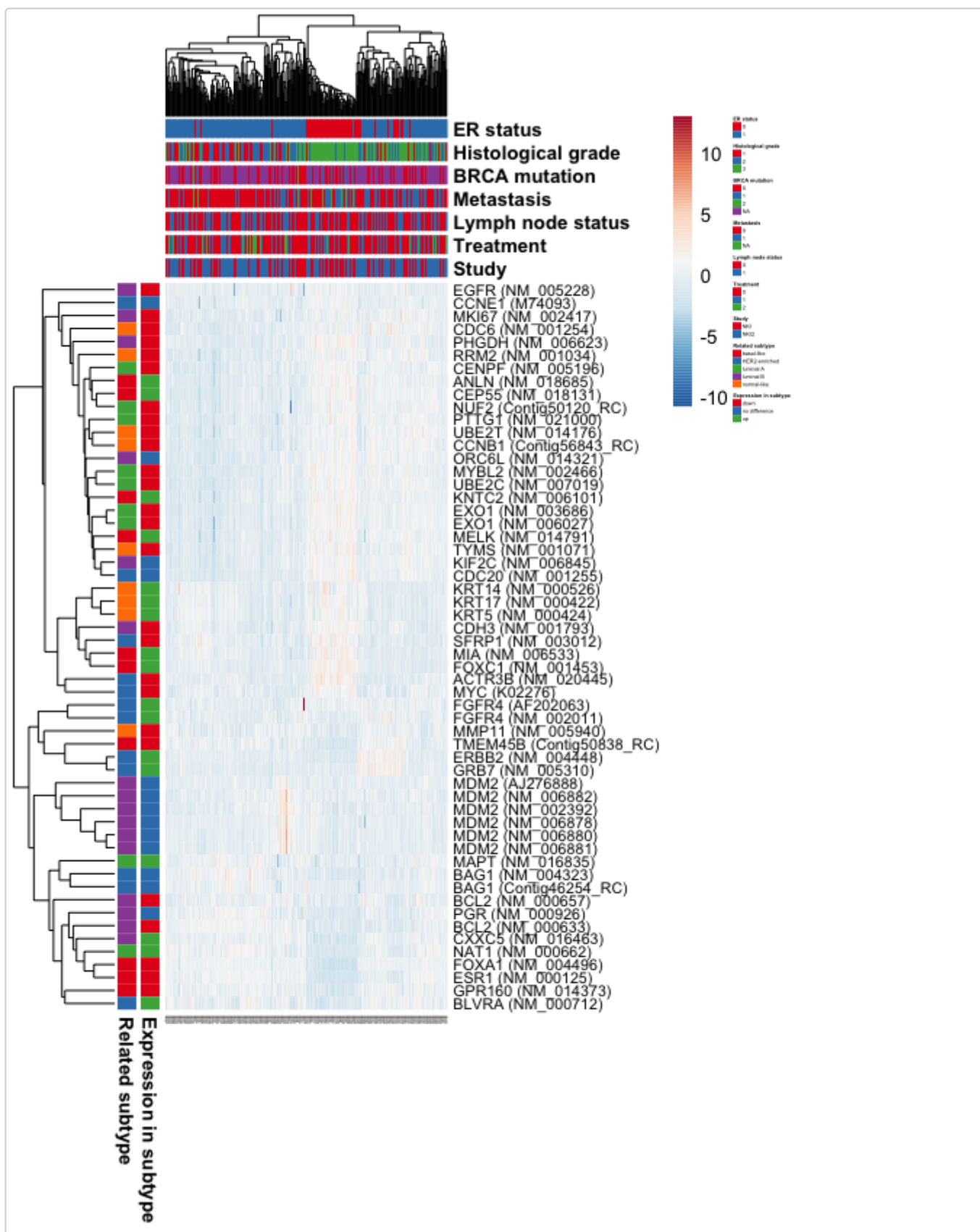

Heatmap example

```
library(clustvis)
file = "https://raw.githubusercontent.com/taunometsalu/ClustVis/master/datasets/nki.csv"
imp = importData(file)
proc = processData(imp)

hm = generateHeatmap(proc)
saveHeatmap(hm, file = NA)
```

Export example

```
library(clustvis)
file = "https://raw.githubusercontent.com/taunometsalu/ClustVis/master/datasets/nki.csv"

imp = importData(file)
proc = processData(imp)
pca = generatePCA(proc)
```

```
hm = generateHeatmap(proc)
exp = c(
  exportData(imp),
  exportData(proc),
  exportData(pca),
  exportData(hm)
)
print(str(exp))
#> List of 16
#>  $ initialMatrix   : num [1:56, 1:337] -0.306 -0.994 0.724 -0.086 -0.027 0.103 -0.396 0.123
-0.134 -0.101 ...
#>   ..- attr(*, "dimnames")=List of 2
#>   .. ..$ : chr [1:56] "SFRP1 (NM_003012)" "TMEM45B (Contig50838_RC)" "PGR (NM_000926)" "MDM2
(NM_002392)" ...
#>   .. ..$ : chr [1:337] "NKI_4" "NKI_6" "NKI_7" "NKI_8" ...
#>  $ initialAnnoRow  :'data.frame':    56 obs. of  2 variables:
#>   ..$ Related subtype     : chr [1:56] "HER2-enriched" "basal-like" "Luminal B" "Luminal B"
...
#>   ..$ Expression in subtype: chr [1:56] "down" "down" "no difference" "no difference" ...
#>  $ initialAnnoCol  :'data.frame':    337 obs. of  7 variables:
#>   ..$ ER status        : chr [1:337] "1" "1" "0" "0" ...
#>   ..$ Histological grade: chr [1:337] "3" "2" "1" "3" ...
#>   ..$ BRCA mutation    : chr [1:337] "0" "0" "0" "0" ...
#>   ..$ Metastasis       : chr [1:337] "0" "0" "0" "0" ...
#>   ..$ Lymph node status : chr [1:337] "0" "0" "0" "0" ...
#>   ..$ Treatment        : chr [1:337] "0" "0" "0" "0" ...
#>   ..$ Study            : chr [1:337] "NKI" "NKI" "NKI" "NKI" ...
#>  $ processedMatrix : num [1:56, 1:337] -0.996 -2.151 2.433 -0.5 -0.25 ...
#>   ..- attr(*, "dimnames")=List of 2
#>   .. ..$ : chr [1:56] "SFRP1 (NM_003012)" "TMEM45B (Contig50838_RC)" "PGR (NM_000926)" "MDM2
(NM_002392)" ...
#>   .. ..$ : chr [1:337] "NKI_4" "NKI_6" "NKI_7" "NKI_8" ...
#>  $ processedAnnoRow:'data.frame':    56 obs. of  2 variables:
#>   ..$ Related subtype     : chr [1:56] "HER2-enriched" "basal-like" "Luminal B" "Luminal B"
...
#>   ..$ Expression in subtype: chr [1:56] "down" "down" "no difference" "no difference" ...
#>  $ processedAnnoCol:'data.frame':    337 obs. of  7 variables:
#>   ..$ ER status        : chr [1:337] "1" "1" "0" "0" ...
#>   ..$ Histological grade: chr [1:337] "3" "2" "1" "3" ...
#>   ..$ BRCA mutation    : chr [1:337] "0" "0" "0" "0" ...
#>   ..$ Metastasis       : chr [1:337] "0" "0" "0" "0" ...
#>   ..$ Lymph node status : chr [1:337] "0" "0" "0" "0" ...
#>   ..$ Treatment        : chr [1:337] "0" "0" "0" "0" ...
#>   ..$ Study            : chr [1:337] "NKI" "NKI" "NKI" "NKI" ...
#>  $ processedSize   : int [1:2, 1:4] 56 337 56 337 56 337 56 337
#>   ..- attr(*, "dimnames")=List of 2
#>   .. ..$ : chr [1:2] "Rows" "Columns"
#>   .. ..$ : chr [1:4] "Before processing" "After collapsing similar columns (if applied)" "After
removing rows and columns with NAs" "After removing constant rows and optionally columns"
#>  $ processedNaRows : chr [1:3, 1:21] "8" "2.4%" "no" "8" ...
#>   ..- attr(*, "dimnames")=List of 2
#>   .. ..$ : chr [1:3] "Count" "Percentage" "Removed"
#>   .. ..$ : chr [1:21] "MDM2 (NM_006881)" "MDM2 (AJ276888)" "MDM2 (NM_006878)" "MDM2
(NM_006880)" ...
#>  $ processedNaCols : chr [1:3, 1:18] "9" "16.1%" "no" "8" ...
```

```
#>    ..- attr(*, "dimnames")=List of 2
#>    .. ..$ : chr [1:3] "Count" "Percentage" "Removed"
#>    .. ..$ : chr [1:18] "NKI_12" "NKI_14" "NKI_339" "NKI_340" ...
#>  $ pcaScores      : num [1:337, 1:56] -1.2 -4.26 -4.35 5.89 -3.53 ...
#>    ..- attr(*, "dimnames")=List of 2
#>    .. ..$ : chr [1:337] "NKI_4" "NKI_6" "NKI_7" "NKI_8" ...
#>    .. ..$ : chr [1:56] "PC1" "PC2" "PC3" "PC4" ...
#>  $ pcaLoadings    : num [1:56, 1:56] 0.0838 -0.0994 -0.1093 -0.0396 0.1175 ...
#>    ..- attr(*, "dimnames")=List of 2
#>    .. ..$ : chr [1:56] "SFRP1 (NM_003012)" "TMEM45B (Contig50838_RC)" "PGR (NM_000926)" "MDM2
#> (NM_002392)" ...
#>    .. ..$ : chr [1:56] "PC1" "PC2" "PC3" "PC4" ...
#>  $ pcaVariance    : num [1:2, 1:56] 0.3112 0.3112 0.1165 0.4277 0.0678 ...
#>    ..- attr(*, "dimnames")=List of 2
#>    .. ..$ : chr [1:2] "Individual" "Cumulative"
#>    .. ..$ : chr [1:56] "PC1" "PC2" "PC3" "PC4" ...
#>  $ pcaCaption     : chr "Unit variance scaling is applied to rows; SVD with imputation is used
#> to calculate principal components. X and Y axis show prin"| __truncated__
#>  $ pcaMatrix      :'data.frame':   337 obs. of  5 variables:
#>   ..$ pcx  : num [1:337] 1.2 4.26 4.35 -5.89 3.53 ...
#>   ..$ pcy  : num [1:337] -2.377 0.037 2.734 -1.173 -1.365 ...
#>   ..$ color: chr [1:337] "1" "1" "0" "0" ...
#>   ..$ shape: chr [1:337] "3" "2" "1" "3" ...
#>   ..$ label: chr [1:337] "NKI_4" "NKI_6" "NKI_7" "NKI_8" ...
#>  $ hmCaption      : chr "Rows are centered; unit variance scaling is applied to rows.
#> Imputation is used for missing value estimation. Both rows and col"| __truncated__
#>  $ hmMatrix       : num [1:56, 1:337] -0.909 -0.372 -0.585 -1.904 0.172 ...
#>    ..- attr(*, "dimnames")=List of 2
#>    .. ..$ : chr [1:56] "EGFR (NM_005228)" "CCNE1 (M74093)" "MKI67 (NM_002417)" "CDC6
#> (NM_001254)" ...
#>    .. ..$ : chr [1:337] "NKI_26" "NKI_243" "NKI_235" "NKI_274" ...
#>  NULL
```