Vignette Title

Vignette Author

2017-04-05

Package

About

Function test_diff

Expression

What is it?

About dataset

In this vignette, we will work with the data set BRCA_mRNAseq_chr17, which contains information about gene expression. This data set contains per-gene read counts computed for genes from 17 chromosome for 100 patients with breast cancer.

```
load("BRCA_mRNAseq_chr17.rda")
```

```
BRCA_mRNAseq_chr17[1:5,1:5]
```

```
##                            SUBTYPE AANAT AARSD1 AATF AATK
## TCGA-A1-A0SB-01A-11R-A144-07  Normal     9   2354 2870  317
## TCGA-A1-A0SD-01A-11R-A115-07    LumA     2   1846 5656  312
## TCGA-A1-A0SE-01A-11R-A084-07    LumA    11   3391 9522  736
## TCGA-A1-A0SF-01A-11R-A144-07    LumA     0   2169 4625  169
## TCGA-A1-A0SG-01A-11R-A144-07    LumA     1   2273 3473   92
```

Nbinom test

Testa for differences between the base means of two conditions. Beside a data set containing read counts, we need a description of the samples, which we keep in a factor, whose elements correspond to different gorups.

In our example we will test for differential expression between groups with LumA breast cancer subtype and other subtypes of that cancer. We will use factor `conditions`, which consist of two values corresponds to subtype of breast cancer: LumA and other.

```
condition<-ifelse(BRCA_mRNAseq_chr17$SUBTYPE=="LumA","LumA","other")
head(condition,8)
```

```
## [1] "other" "LumA"  "LumA"  "LumA"  "LumA"  "LumA"  "other" "LumA"
```

Evaluation of expr_nbinom may take a few minutes

```
test.expr <- test_diff(BRCA_mRNAseq_chr17[,-1], condition, "nbinom")
```

As a result we obtain a data frame contains among others: base mean, means for each group and p-value of test.

```r
head(test.expr)
```

```
##        id    exp.mean exp.mean.gr1 exp.mean.g2   exp.fold exp.log.fold
## 1   AANAT    3.455436     3.901537    2.952385 0.7567235   -0.40216187
## 2  AARSD1 2779.448414  2712.105777 2855.387983 1.0528306    0.07427334
## 3    AATF 6750.269650  7042.041468 6421.250366 0.9118450   -0.13313947
## 4    AATK  352.805108   355.415829  349.861103 0.9843712   -0.02272566
## 5   ABCA5 1933.257431  1888.913529 1983.262257 1.0499487    0.07031881
## 6   ABCA6  689.547294   589.046234  802.878277 1.3630140    0.44680041
##     exp.pval  exp.padj
## 1 0.9453634 0.9920531
## 2 0.6627296 0.8992428
## 3 0.6864598 0.9111193
## 4 0.8051408 0.9585881
## 5 0.5837682 0.8489059
## 6 0.1289441 0.3857753
```

Metylation

What is it?

About dataset

Data set BRCA_methylation_chr17 contains information about methylation of CpG islands located on 17 chromosome for patienst with breast cancer.

```r
load("BRCA_methylation_chr17.rda")
head(BRCA_methylation_chr17)[1:5,1:5]
```

```
##                              SUBTYPE cg00021527 cg00031162  cg00032227
## TCGA-A1-A0SD-01A-11D-A112-05    LumA 0.03781858  0.7910348 0.006391233
## TCGA-A2-A04N-01A-11D-A112-05    LumA 0.01437552  0.7359370 0.008752293
## TCGA-A2-A04P-01A-31D-A032-05   Basal 0.01360124  0.6967802 0.009442039
## TCGA-A2-A04Q-01A-21D-A032-05   Basal 0.01525656  0.5341244 0.014674247
## TCGA-A2-A04T-01A-21D-A032-05   Basal 0.01167384  0.7378100 0.012251559
##                              cg00050312
## TCGA-A1-A0SD-01A-11D-A112-05 0.02356901
## TCGA-A2-A04N-01A-11D-A112-05 0.02770303
## TCGA-A2-A04P-01A-31D-A032-05 0.01402589
## TCGA-A2-A04Q-01A-21D-A032-05 0.01510194
## TCGA-A2-A04T-01A-21D-A032-05 0.01174021
```

mapping to genes using function `map_to_gene`

```r
BRCA_methylation_chr17_gen <- map_to_gene(BRCA_methylation_chr17[,-1])
head(BRCA_methylation_chr17_gen)[1:5,1:5]
```

```
##                                 AANAT    AARSD1       AATF      AATK
## TCGA-A1-A0SD-01A-11D-A112-05 0.7148533 0.8625816 0.24294092 0.7835302
## TCGA-A2-A04N-01A-11D-A112-05 0.5850106 0.8355825 0.21367129 0.8466190
## TCGA-A2-A04P-01A-31D-A032-05 0.4495537 0.8786166 0.03277413 0.3417919
## TCGA-A2-A04Q-01A-21D-A032-05 0.7120650 0.8819490 0.03460160 0.7264985
## TCGA-A2-A04T-01A-21D-A032-05 0.6010397 0.7739978 0.02501599 0.6276399
##                                    ABC1
## TCGA-A1-A0SD-01A-11D-A112-05 0.01401806
```

```
## TCGA-A2-A04N-01A-11D-A112-05 0.01275865
## TCGA-A2-A04P-01A-31D-A032-05 0.01455092
## TCGA-A2-A04Q-01A-21D-A032-05 0.01283647
## TCGA-A2-A04T-01A-21D-A032-05 0.01278928
```

ttest

```r
condition <- ifelse(BRCA_methylation_chr17$SUBTYPE=="LumA","LumA", "other")
test.mety <- test_diff(BRCA_methylation_chr17_gen, condition)
```

As a result we obtain a data frame . . .

```r
head(test.mety)
```

```
##         met.log.fold  met.mean met.t.stat      met.pval      met.padj
## ICAM2    -0.15151320 0.3330801  -8.916763 3.754116e-17 3.063359e-14
## RILP     -0.05073691 0.3341447  -7.638110 2.575168e-13 1.050668e-10
## PIPOX     0.11505558 0.3647812   7.165992 5.360053e-12 1.196885e-09
## TNFSF12  -0.13412855 0.2485025  -7.151613 5.867083e-12 1.196885e-09
## CD7       0.09822690 0.8127112   6.986460 1.641919e-11 2.679612e-09
## KSR1      0.19973400 0.5549808   6.950127 2.054467e-11 2.794075e-09
##           met.B      id
## ICAM2   27.37828   ICAM2
## RILP    18.64868    RILP
## PIPOX   15.65860   PIPOX
## TNFSF12 15.56967 TNFSF12
## CD7     14.55763     CD7
## KSR1    14.33731    KSR1
```

Combining data

We can combine results for both tests, for expression and methylation, to easily use functions designed to
visualize those data.

```r
test.both <- full_data(test.expr, test.mety)
head(test.both)
```

```
##        id    exp.mean exp.mean.gr1 exp.mean.g2  exp.fold exp.log.fold
## 1   AANAT    3.455436     3.901537    2.952385 0.7567235  -0.40216187
## 2  AARSD1 2779.448414  2712.105777 2855.387983 1.0528306   0.07427334
## 3    AATF 6750.269650  7042.041468 6421.250366 0.9118450  -0.13313947
## 4    AATK  352.805108   355.415829  349.861103 0.9843712  -0.02272566
## 5   ABCA5 1933.257431  1888.913529 1983.262257 1.0499487   0.07031881
## 6   ABCA6  689.547294   589.046234  802.878277 1.3630140   0.44680041
##     exp.pval  exp.padj met.log.fold   met.mean met.t.stat      met.pval
## 1 0.9453634 0.9920531   0.07237890 0.70420013  4.4829647 1.026904e-05
## 2 0.6627296 0.8992428   0.01114861 0.75430861  0.6669291 5.052981e-01
## 3 0.6864598 0.9111193   0.01360521 0.12700442  1.2459842 2.136847e-01
## 4 0.8051408 0.9585881   0.06619897 0.69837465  4.4078014 1.427292e-05
## 5 0.5837682 0.8489059   0.00510806 0.08834311  0.4746493 6.353608e-01
## 6 0.1289441 0.3857753  -0.03402781 0.24872043 -2.0227305 4.393248e-02
##       met.padj     met.B
## 1 0.0001862119  1.559446
```

```
## 2 0.6620225949 -7.986421
## 3 0.3944948472 -7.431520
## 4 0.0002283667  1.244016
## 5 0.7590840893 -8.096320
## 6 0.1417004023 -6.171584
```

Plots

log-log p-value

```
p_values_plot(test.both)
```
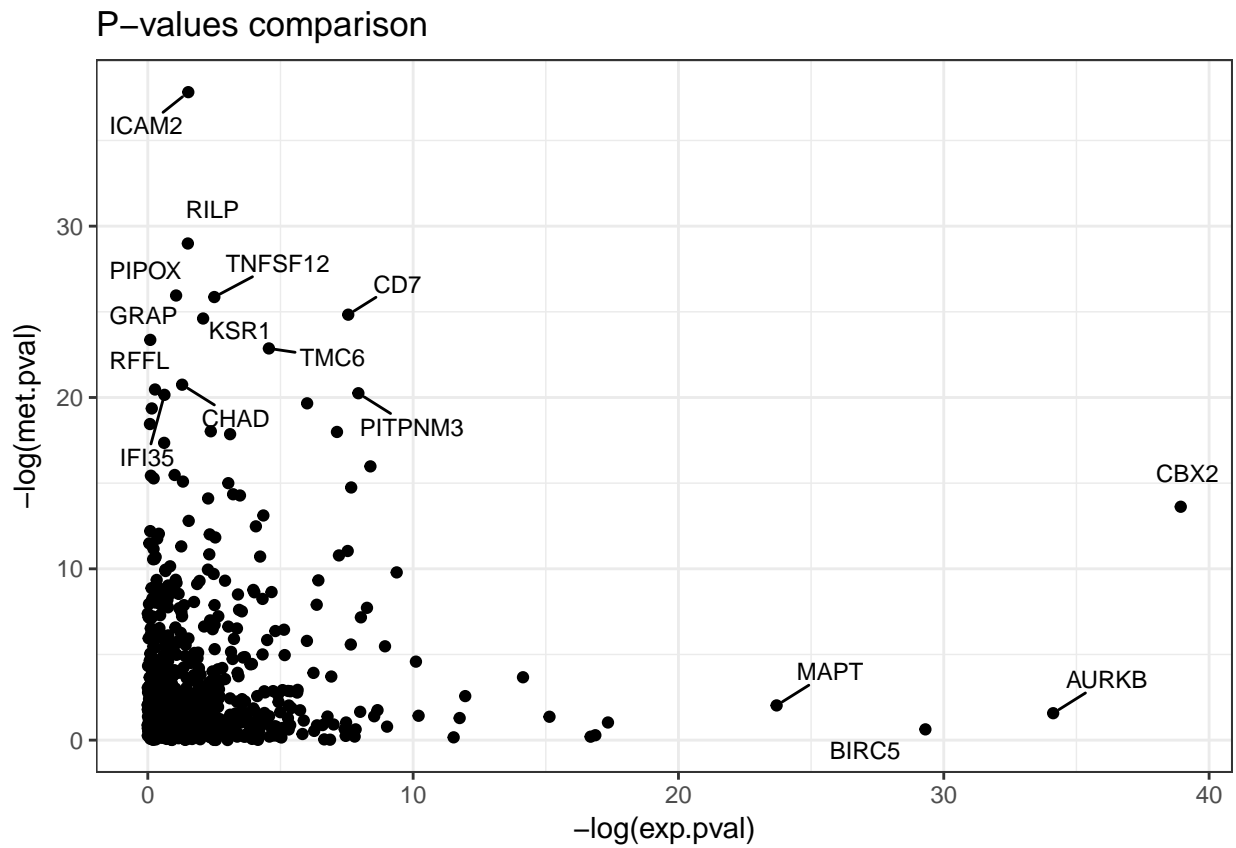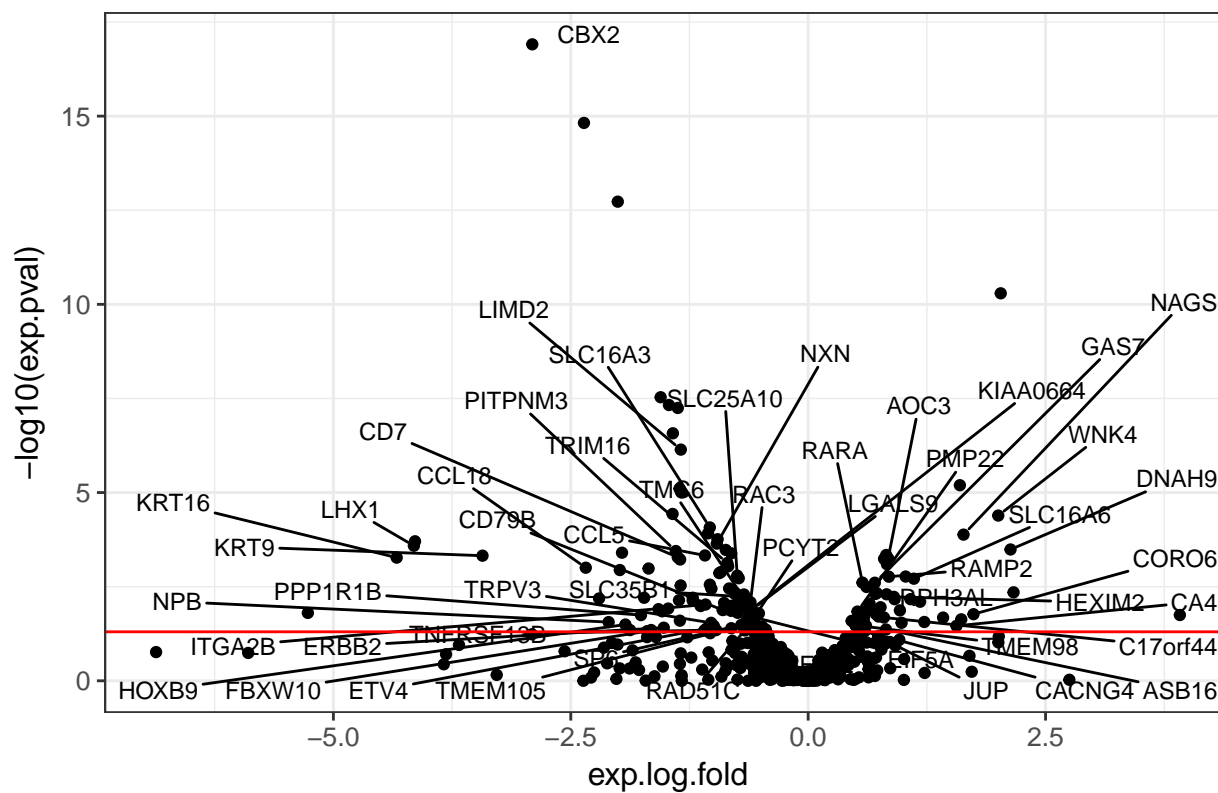

volcano

Volcan plot for expression

```
volcano_plot(test.both, type="expression")
```
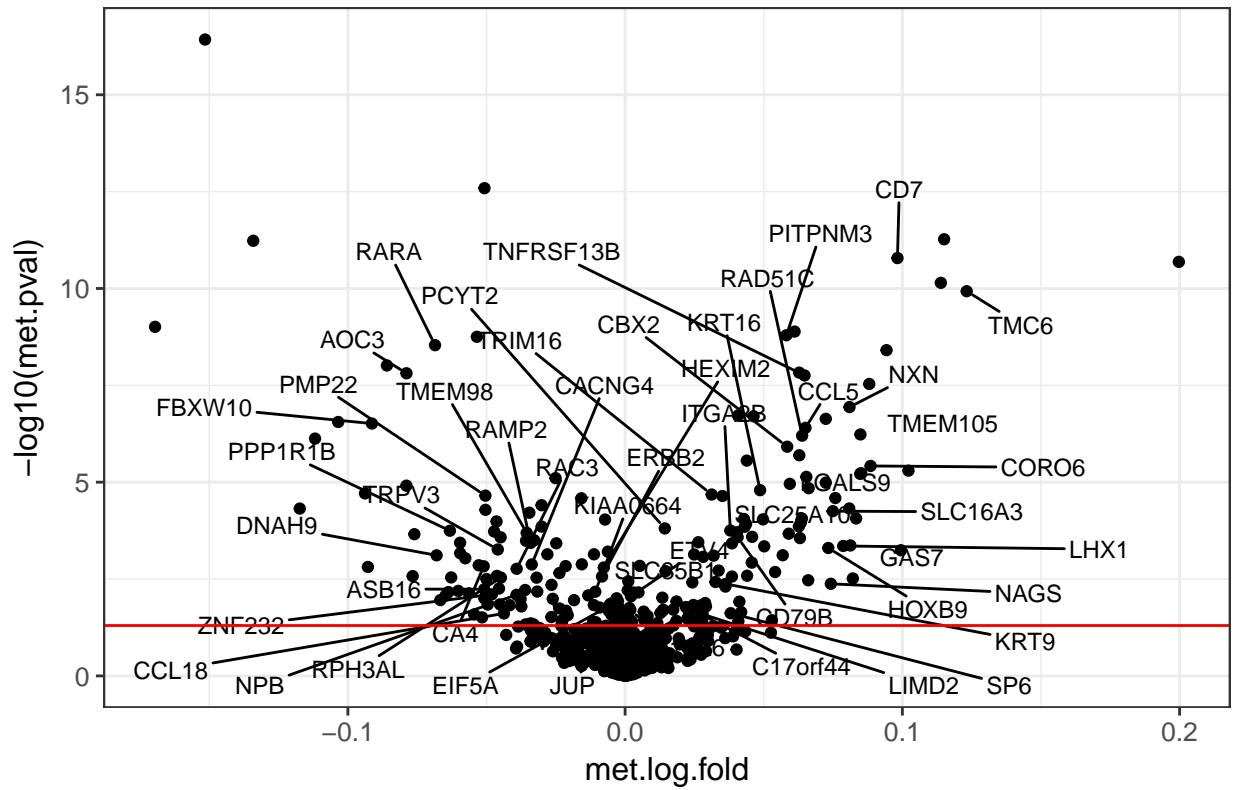
Volcano plot of expression

volcano plot for methylation

```
volcano_plot(test.both, type="methylation")
```

Volcano plot of methylation

metyl "road" + boxplots for expr