Charles Valentine  
Homework 10  
11/13/2016

Note:  
*All code is in script hw08.R --- I have made use of cat and print statements to display information easily!*

**Problem 1**

1. The R command for preprocessing the raw microarray data is:

INPUT

eset = expresso(yeast.raw,

bgcorrect.method="mas",

normalize.method="quantiles",

pmcorrect.method="pmonly",

summary.method="medianpolish")

1. The mean expression values for the first five genes across all samples in the E-MEXP-1551 microarray data is:

OUTPUT

8.9361

5.666

5.6505

11.3809

9.7525

1. The number of genes in the preprocessed samples is.

OUTPUT

10928

The number of samples in the preprocessed samples is:

OUTPUT

30

**Problem 2**

1. The annotation package for the yeast data in Problem 1 is:

OUTPUT

yeast2

1. The number of GO numbers related to Molecular Function (MF) in the 1769308\_at gene:

OUTPUT

7

1. The number of GO parents to the GO numbers in Problem 2b are (excluding NA):

OUTPUT

7

1. The number of GO children to the GO numbers in Problem 2b are (excluding NA):

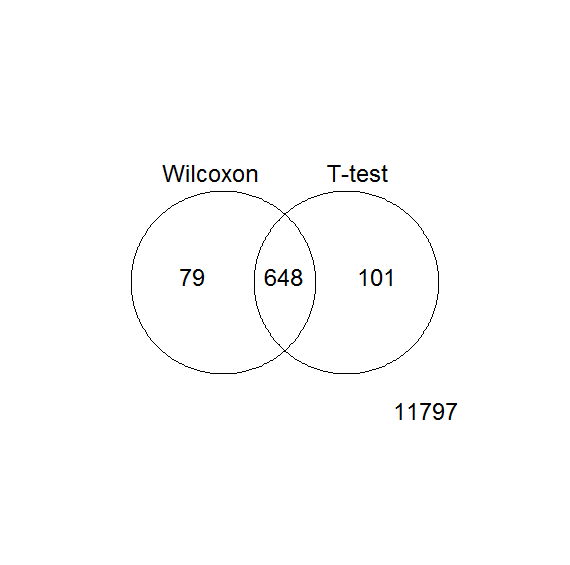
OUTPUT

422

**Problem 3**

1. See code
2. The Venn diagram for the genes that are expressed differently in stages B2 and B3 of ALL patients using the Wilcoxon exact test and the Welch t-test (p-value < 0.001) are presented here:

OUTPUT



1. The genes that pass the Wilcoxon exact test filter are:

OUTPUT

727

1. The genes that pass the Welch t-test filter are:

OUTPUT

648

1. The genes that pass both the filters in Problem 3b that are also oncogenes are:

OUTPUT

0

**Problem 4**

1. See code
2. The top five genes with nonzero means in the B3 group from a selection of persons with B-cell leukemia in the stages B1, B2, and B3 are:

OUTPUT

logFC AveExpr t P.Value adj.P.Val B

AFFX-hum\_alu\_at 13.61 13.53 355.6 5.059e-127 6.387e-123 270.8

32466\_at 12.71 12.71 316.7 4.247e-123 2.681e-119 263.9

31962\_at 13.05 13.09 307.1 4.695e-122 1.976e-118 262.0

32748\_at 12.15 12.12 302.8 1.407e-121 4.406e-118 261.2

35278\_at 12.52 12.48 302.0 1.745e-121 4.406e-118 261.0

1. The number of differentially expressed genes found using two contrasts and an analysis of variance to test the null hypothesis of equal group means are:

OUTPUT

314

The top five genes in this analysis that express differently among the three groups are:

OUTPUT

logFC AveExpr t P.Value adj.P.Val B

35991\_at -0.5964 4.142 -6.439 8.977e-09 0.0001133 9.707

36711\_at 2.3665 7.694 6.252 2.007e-08 0.0001267 8.970

33873\_at -0.5708 7.206 -6.065 4.439e-08 0.0001868 8.243

35614\_at -1.7249 5.570 -5.903 8.760e-08 0.0002765 7.620

37750\_at -0.3812 4.530 -5.823 1.226e-07 0.0003047 7.312