Seminar III: R/Bioconductor

August-December 2009

Leonardo Collado Torres

Bachelor in Genomic Sciences (LCG),
UNAM, Cuernavaca, Mexico

lcollado@lcg.unam.mx

http://www.lcg.unam.mx/~lcollado/index_en.php

August 14, 2009

Assistants: Alejandro Reyes areyes@lcg.unam.mx, José Reyes jreyes@lcg.unam.mx and Víctor Moreno jmoreno@lcg.unam.mx

Note: Questions through the forum please. Those who are not from the sixth LCG generation send us an email so we can register you on the forum.

Abstract

The following exercises are meant to review how to use R, create some basic plots and use the apply function family.

1 Review

1. Why does the following expression show a warning? This is part of what rule?

```
> c(2, 3) + c(4, 5, 7)
```

2. For all the prime numbers between 1 and 10, calculate its square root. What is the sum, median and mean?

2 Plots

- Read the following file into R: `ftp://ftp.ebi.ac.uk/pub/databases/genome_reviews/gr2species_phage.txt`[1] and make the following plots with your username on the title. Check whether using a log10 scale on the y axis helps.

  1. Sort the genome sizes (column 2) and plot them in a line with increasing values.
  2. Plot a histogram with a density line for the same data.
  3. Plot a boxplot for the differences between contigous sorted genomes. Meaning, 2nd smallest - smallest, 3rd smallest - 2nd smallest, etc.[2]
  4. Make a barplot showing the 10 biggest genomes. Include the names[3] on the x axis and every bar has to have a different color and/or density.[4]

3 `Apply` functions

1. What is the mean genome size for every type of replicon (column 4)? You have an atomic vector and a factor so use ...

2. Create a matrix `mat` with 10 rows and 10 columns and 100 random uniform values from 1 to 10. Create your own function and apply it to every row so that every row will now sum 1 in your new matrix `mat2`.

3. Using the same matrix `mat`, make the matrix `mat3` with row sums equal to 1 using built in R matrix functions. `mat2` and `mat3` should be the same.

[1] Look for the useful function for this case
[2] You might want to use `apropos` searching for diff...
[3] They have to be redable
[4] The `which` function might be useful.