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/`

September 4, 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
With the following exercises you'll tune your skills with packages such as biomaRt that enable you to download public data sets.

1 biomaRt

1. Use `biomaRt` to access Uniprot. Get all the proteins from the virus domain. Find out which is the most frequent organism. Then find which is the most frequent EC number.

2. Follow `www.ensembl.org/info/website/tutorials/biomart_worked_ex.pdf` and reproduce it using `biomaRt`. Every time they mention "click result" you need to create a different object with the corresponding information.

3. Make any plot you want with the previous data. Don't use the base plotting functions. Meaning, use `lattice`, `plotrix`, ...