Asterias: an example of using R in a web-based bioinformatics suite of tools

Ramón Díaz-Uriarte$^{1,2}$, Andrés Cañada$^1$, Edward R. Morrissey$^1$,
Oscar Rueda$^1$, Andreu Alibés$^1$, David Casado$^1$

$^1$Bioinformatics Unit, Spanish National Cancer Center (CNIO), Melchor Fernández Almagro 3,
Madrid, 28029, Spain

Abstract

Asterias (http://www.asterias.info) is an integrated collection of freely-accessible web tools for the analysis of microarray gene expression and aCGH data. Six of the applications use R (and many R packages) for the computations and graphics, and four of those rely heavily on Rmpi and/or snow for parallelization (Asterias runs on a computing cluster with 60 CPUs). R has shown, once again, that it is an ideal system to “turn ideas into software, quickly and faithfully” (Chambers, 1998, “Programming with data”), but setting up and maintaining the system up and running has presented several challenges. In this talk we will discuss some of the features of our set-up, including load-balancing and high-availability, the combination of R and Python for the web-based applications, checking the status of MPI and launching MPI-dependent applications in a web-serving context, and automated testing of web-based applications.

$^2$Corresponding author. Email: rdiaz@ligarto.org