Using R with GeneSpring

Thon de Boer

October 3, 2007

1 Introduction

This package contains a number of functions to facilitate the integration of R code into the Gene Expression analysis program GeneSpring, by Agilent Technologies. Available functions include:

- `GSload.int()` - Read GeneSpring experiment interpretation from file and return a GeneSpring gene expression object (GSint)

- `GSload.intBC()` - Read GeneSpring experiment interpretation from file and return a BioConductor gene expressioon object (exprSet)

- `GSload.exp()` - Read GeneSpring experiment from file and return a Gene-Spring gene expression object (GSint)

- `GSload.expBC()` - Read GeneSpring experiment from file and return a BioConductor gene expressioon object (exprSet)

- `GSint()` - Create a GeneSpring gene expression object (GSint)

- `GSint2BC()` - Convert a GeneSpring gene expression object (GSint) to a BioConductor gene expression object (exprSet)

- `BC2GSint()` - Convert a BioConductor gene expression object (exprSet) to a GeneSpring gene expression object (GSint)

- `GSload.genelist()` - Read a GeneSpring GeneSpring gene list from file

- `GSsave.genelist()` - Save a GeneSpring GeneSpring gene list to file

- `GSsave.exp()` - Save a GeneSpring gene expression object (GSint) to file

Most of the functions are conversion functions that convert files in the proprietary format of GeneSpring, into objects that can be used by generic R functions or the BioConductor packages. The BioConductor GeneSpring library should be used in conjunction with the GeneSpring-R integration package that will

complete the incorporation of R scripts into GeneSpring. This GeneSpring-R integration package will install some auxillary scripts and programs that are needed to fully and seamlessly integrate R scripts into GeneSpring. For more information on integrating R scripts into the GeneSpring environment see the support section on our website at www.silicongenetics.com.

2 GeneSpring Experiments and Interpretations

Some defintions of GeneSpring terms will help in the construction of useful scripts.

- `Sample` - Expression values for a single hybridization.

- `Experiment` - Set of samples that are related to each other. (For instance, all the samples for a time series or drug concentration study. Each replicate will be considered a separate sample)

- `Interpretation` - One or more representations of the same experiments. An experiment can be represented by combining all the Time points together, be defining the experimental earameter Time as the parameter that defines the condtions. Each experiment can have one or more interpretations.

- `Condition` - Group of samples that have a particular parameter in common. (For instance, all the replicates for the measurements (samples) for time 0 hours can be combined into one condition for Time 0)

Basically, the main difference between Experiments and Interpretations is that in Experiments, the samples are always represented as individual samples, while in Interpretations, one or more samples can be combined into one condition, where the expression values are represented by the mean or average of expression values of the samples in each of the conditions. Depending on the pupose of the analysis, either Experiments or Interpretations can be used.

See the GeneSpring-R Integration document for more information on how to select the type of expression data to be used in the R scripts.