

EPI Programming



1. Introduction of R



Installing R-integration Package

- Download from Agilent web site.

<http://www.chem.agilent.com/scripts/generic.asp?lpage=34733&indcol=Y&prodcol=Y>

- It contains R2.1.0 with libraries, and sample programs.

The latest version of R is 2.1.1. (2005 Sep.)

- We provide it rather as examples of R implementation than standard function of GeneSpring GX.



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 3

2. Load data from GS



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 4

A Simple Program to Get Familiar with R

- Create “OpenRGUI.R” and save in the GeneSpring GX data directory.

```
#Prepare GeneSpring environment  
  
library(utils)  
library(GeneSpring)  
  
#Read the Experiment  
  
gs.int <- GSload.int()  
gs.exp <- GSload.exp()
```

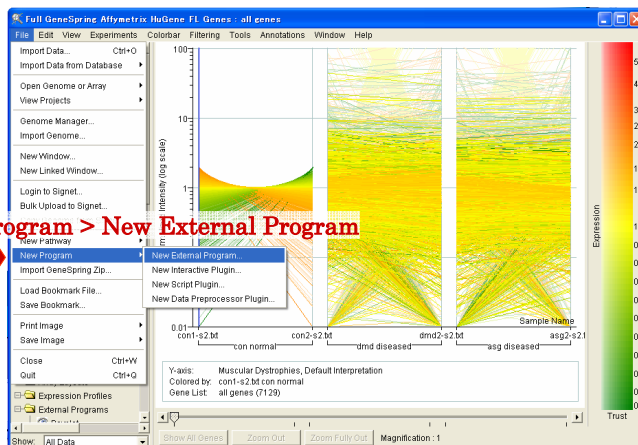


Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 5

Add New EPI Program to GeneSpring GX

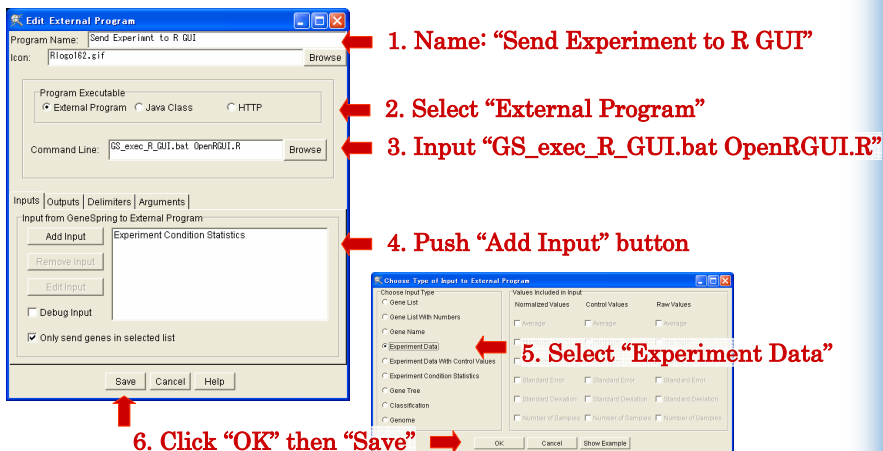
File > New Program > New External Program



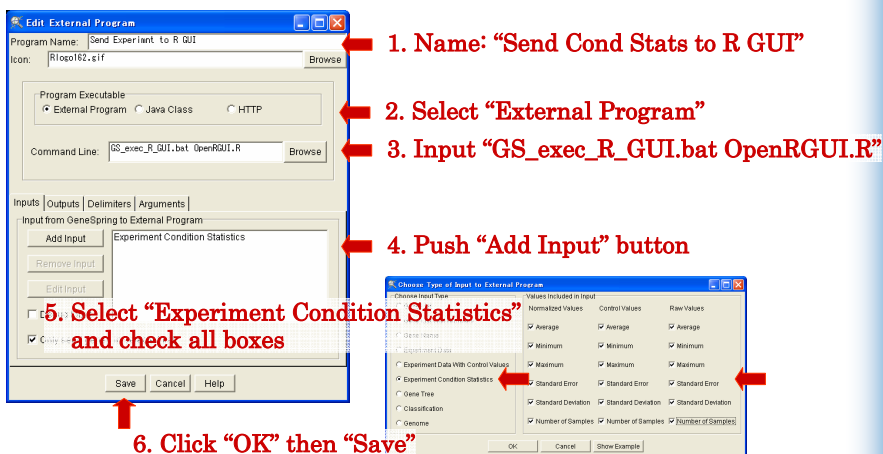
Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 6

New EPI Program 1



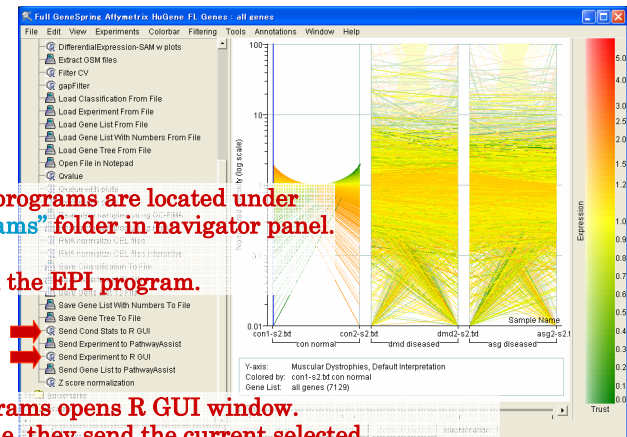
New EPI Program 2



Run The New EPI Programs

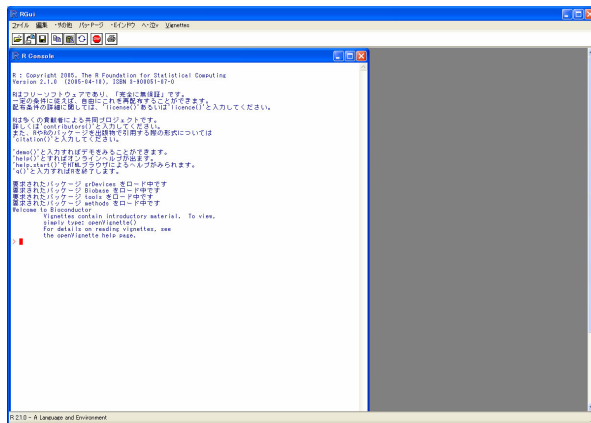
The saved EPI programs are located under
"External Programs" folder in navigator panel.

Click one to run the EPI program.



These EPI programs opens R GUI window.
At the same time, they send the current selected
Gene List (or all genes) and Interpretation to create R object.

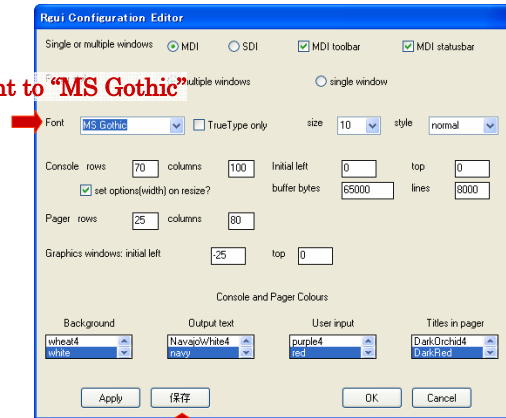
Automatically Opens R GUI Window



R GUI Japanese environment

1. Open “Rgui Configuration Editor” by selecting “編集 > GUIプリファレンス”

2. Change Font to “MS Gothic”



3. Save to [R directory]\etc\Rconsole



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 11

Basics of GeneSpring objects in R

- Type “`help(GSint)`” and return to see help of GeneSpring interpretation object.

You'll see what kind data is available from a **GSint** object.

- Try these helps

`help(GSint),`

`help(GSload.exp), help(GSload.int)`

`help(GSsave.exp), help(GSsave.genelist)`

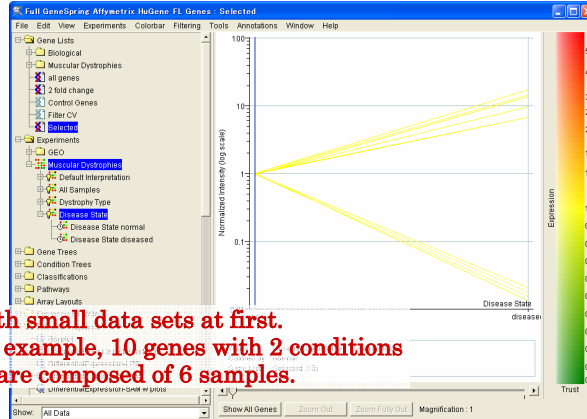
`help(GSint2BC), help(BC2GSint)`



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 12

Run "Send Experiment to R GUI"



1. Try with small data sets at first.
In this example, 10 genes with 2 conditions
which are composed of 6 samples.

2. Click on "Send Experiment to R GUI" under ExternalPrograms folder.



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 13

Inside of gs.exp 1

```

> gs.exp
GeneSpring Expression Interpretation set (GSInt)
  Muscular Dystrophies (slot: expName)
  with 10 genes (slot: ngenes),
  6 Conditions/samples (slot: numConditions)
  and 3 experimental parameters (slot: expParam).
Available expression data slots:
  a.nor (Normalized expression values).
  a.expBa.nor
  1 2 3 4 5 6
1 1.384835e+01 1.284564e+01 1.0083529 0.9396472 2.065375e+01 1.204419e+01
2 1.502888e+01 1.653902e+01 0.3708378 1.0293621 1.873575e+01 1.764516e+01
3 4.556256e+00 5.770994e+00 0.8446832 1.1553167 3.323806e+00 8.386924e+00
4 8.89141e+00 1.005974e+01 0.7567988 1.2432032 2.231335e+01 1.732586e+01
5 1.194405e+01 1.141790e+01 0.8132818 1.1867182 8.252713e+00 7.025959e+00
6 7.692926e+00 7.374499e+00 1.0046107 0.9958893 6.759110e+00 5.336187e+00
7 3.204820e-05 3.461205e-05 1.1257318 0.8742681 6.687343e-02 3.101438e-05
8 3.165413e-05 5.469834e-02 0.7487916 1.2532083 3.683886e-05 3.063303e-05
9 2.226753e-05 1.827719e-01 0.7850927 1.2148073 2.591480e-05 2.154922e-05
10 5.953708e-03 1.266001e-05 0.8739647 1.0261553 1.395777e-05 2.880926e-02
  
```

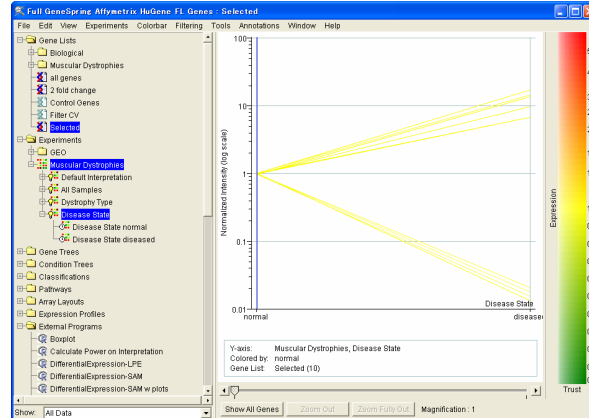
"gs.exp" has normalized values.
To access these values, type [gs.exp@a.nor](#)
But "gs.int" is empty.



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 14

Run "Send Cond Stats to R GUI"



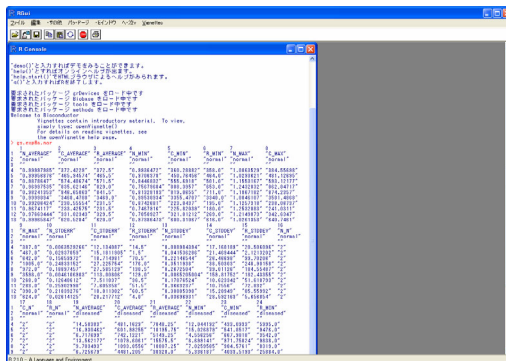
Click on "Send Cond Stats to R GUI" under ExternalPrograms folder.



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 15

Inside of gs.exp 2



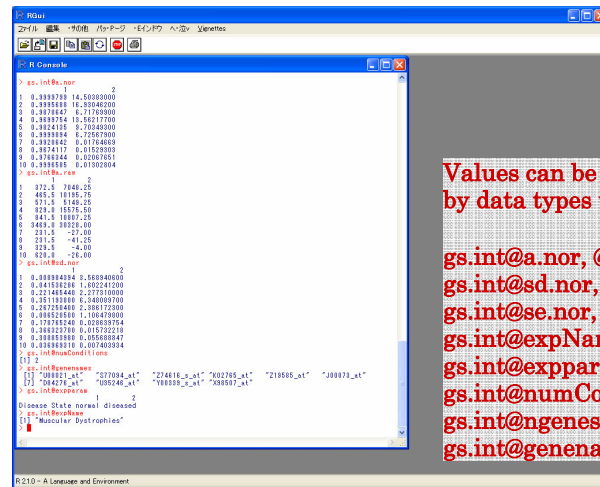
This time, not only average of normalized value but also all other type of values are available.

It may be hard to extract values that users exactly want.



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 16



Values can be extracted more easily by data types with “gs.int”.

gs.int@a.nor, @a.ctr, @a.raw
gs.int@sd.nor, @sd.ctr, @sd.raw
gs.int@se.nor, @se.ctr, @se.raw
gs.int@expName
gs.int@expparam
gs.int@numConditions
gs.int@ngenues
gs.int@genenames

Summary of loading data from GS

■ Sending “Experiment” from GS

Normalized data of selected genes can be extracted.

```
gs.exp <- GSload.exp()
gs.exp@a.nor
```

■ Sending “Condition Statistics” form GS

Various kind of data of selected genes can be extracted.

```
gs.int <- GSload.int()
gs.int@a.nor
gs.int@a.raw
gs.int@sd.nor ...
```

3. Return data to GS



Save Gene List

■ GSSave.genelist(genelist)

Argument: genelist would be

- simple list or vector of gene names or **data.frames**.
- To save a gene list with **associated numbers**, **only data.frame**.

■ Example (CV filter)

```
genelist <- data.frame (genes = gs.int@genenames, pass = as.numeric(CV))  
GSSave.genelist(genelist)
```



Save Experiment

■ “GSsave.exp(interpretation)”

Argument: interpretation would be

Gene Spring interpretation object (**GSint class**)

■ Example (Z score normalization)

```
newGSint <- GSint(a.nor=a.nor.new, expName="Z score experiment",  
expparam=gs.int@expparam, genenames=gs.int@genenames)  
GSsave.exp(newGSint)
```



Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 21

4. Create New Programs Using Existing Programs

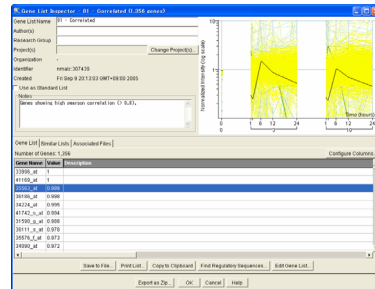


Agilent Technologies

GeneSpring GX EPI Programming
© Agilent 2005
Page 22

Find Synchronized Genes

- Extracts genes which has similar expression pattern in two groups of conditions using R correlation function. (e.g. similar time course pattern in different dose.)



Use cv_filter.R as template

- Make a copy “cv_filter.R” in the data directory and name “Find_Synchronized_Genes_by_Correlation.R.”
- Then modify notes and code between “gs.int<-GSload.int()” and “GSsave.genelist(genelist)”

Install these sample programs

- Drag and Drop
“FindSynchronizedGenes.zip” onto
GeneSpring GX window.
- Open
“Find_Synchronized_Genes_by_Correlati
on.R” and
“Find_Synchronized_Genes_by_Fold.R”
to see the modifications.



5. Bioconductor



What is Bioconductor?

- Bioconductor (<http://www.bioconductor.org/>) is an open source and open development software project for the analysis and comprehension of genomic data.
- Bioconductor is primarily based on the [R](#) programming language. There are two releases of Bioconductor every year. In addition there are a large number of [meta-data packages](#) available. They are mainly, but not solely oriented towards different types of microarrays.
- * Quoted from the Bioconductor's web site.
(<http://www.bioconductor.org/whatisit>)



Search Current Installed Package

- Get a list of current installed packages.
`library()`
- Get a list of current loaded packages.
`search()`
- Load a package.
`library(NameOfLibrary)`
- GeneSpring R-integration package includes some of Bioconductor's.



Install Bioconductor Packages

- Open R GUI and select “select download site” under “Package” menu.
- Select “Bioconductor.”
- Select “Install Package.”
- Select a package you’d want to install and click “OK.”



Load a Library

- After installing a library, file is saved on local computer. BUT the library is not loaded automatically.
- Load the library by command like,
`library(aCGH)`
- To see “help” of a library, command like,
`library(help=aCGH)`



Reference

- **Bioinformatics And Computational Biology Solutions Using R And Bioconductor. (Statistics for Biology & Health)**
Springer-Verlag ; ISBN: 0387251464 ; (2005/09)
- **Bioconductor Mailing list**
<http://www.bioconductor.org/docs/mailList.html>

