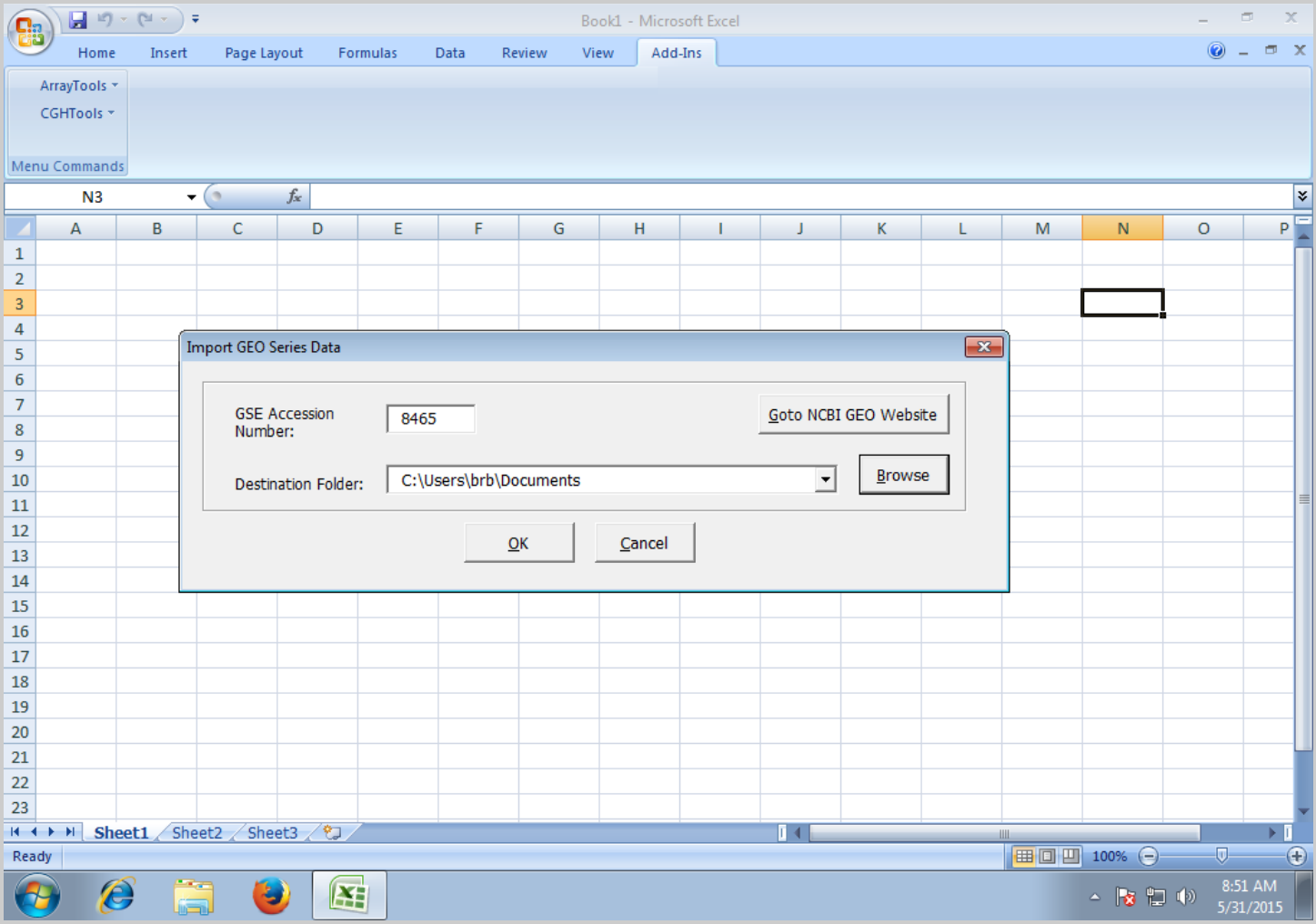
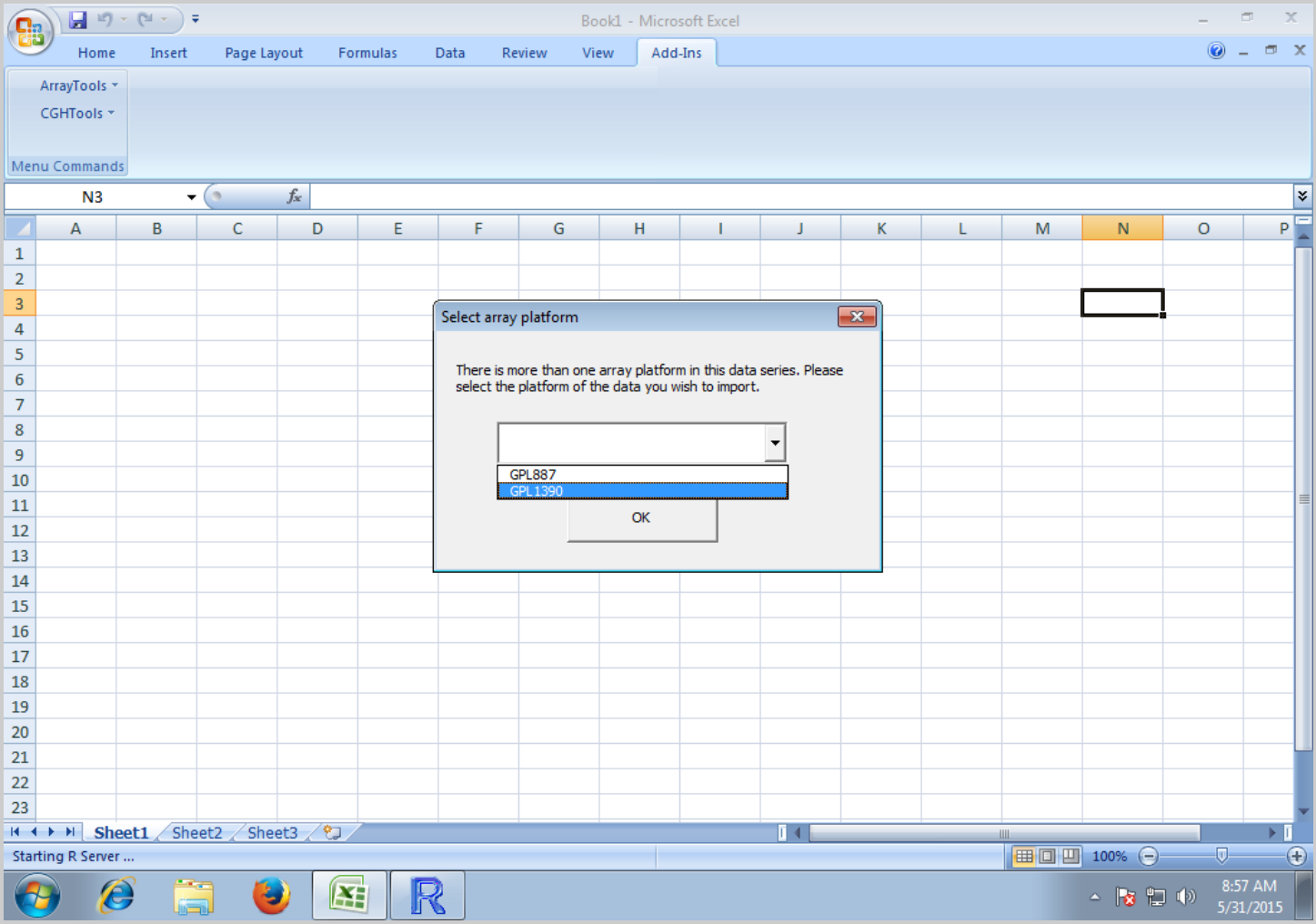
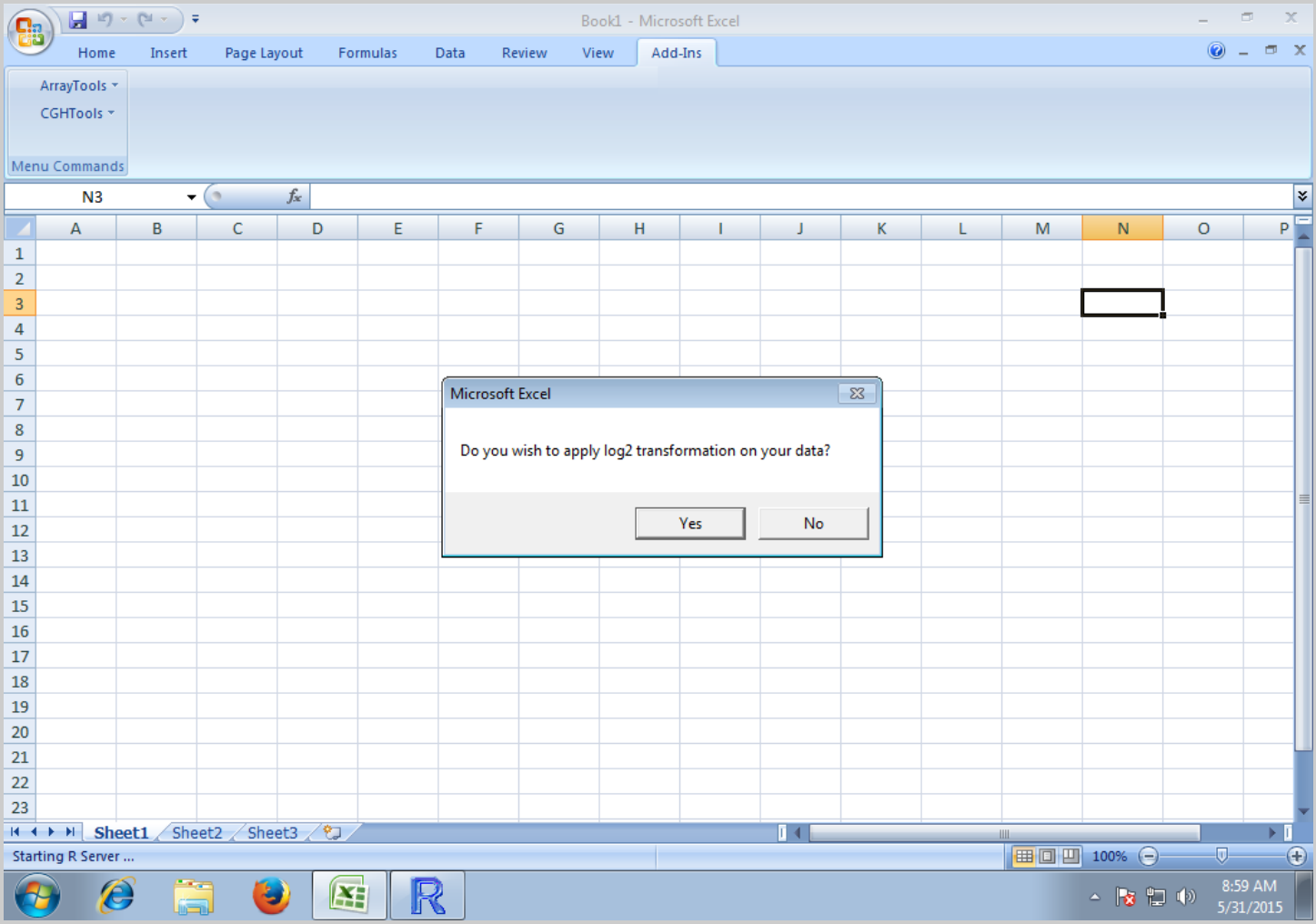
Importing GSE series is a new feature in BRB-ArrayTools v4.5.0

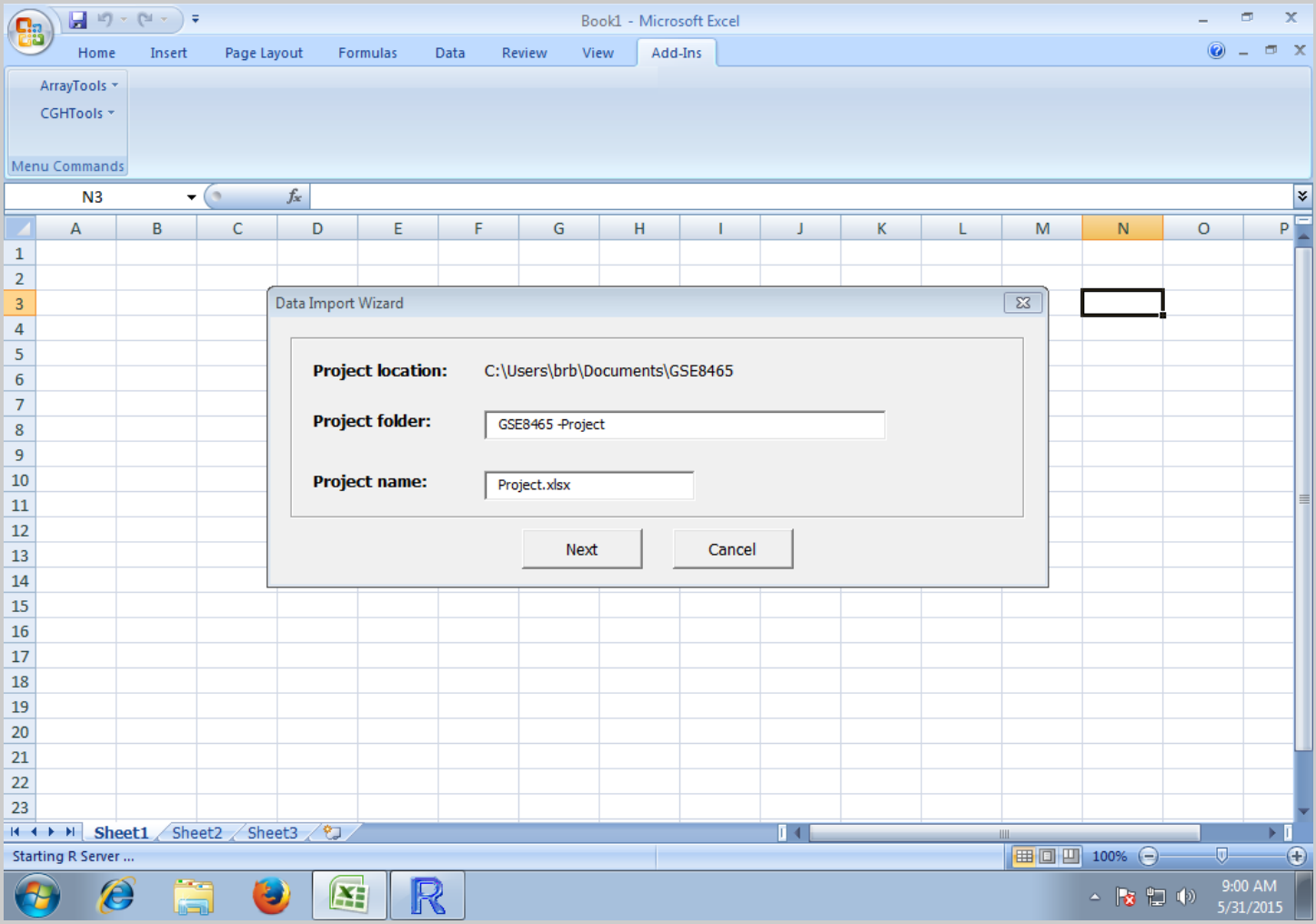
In this example, I pick up one GSE series contains more than one GPL.



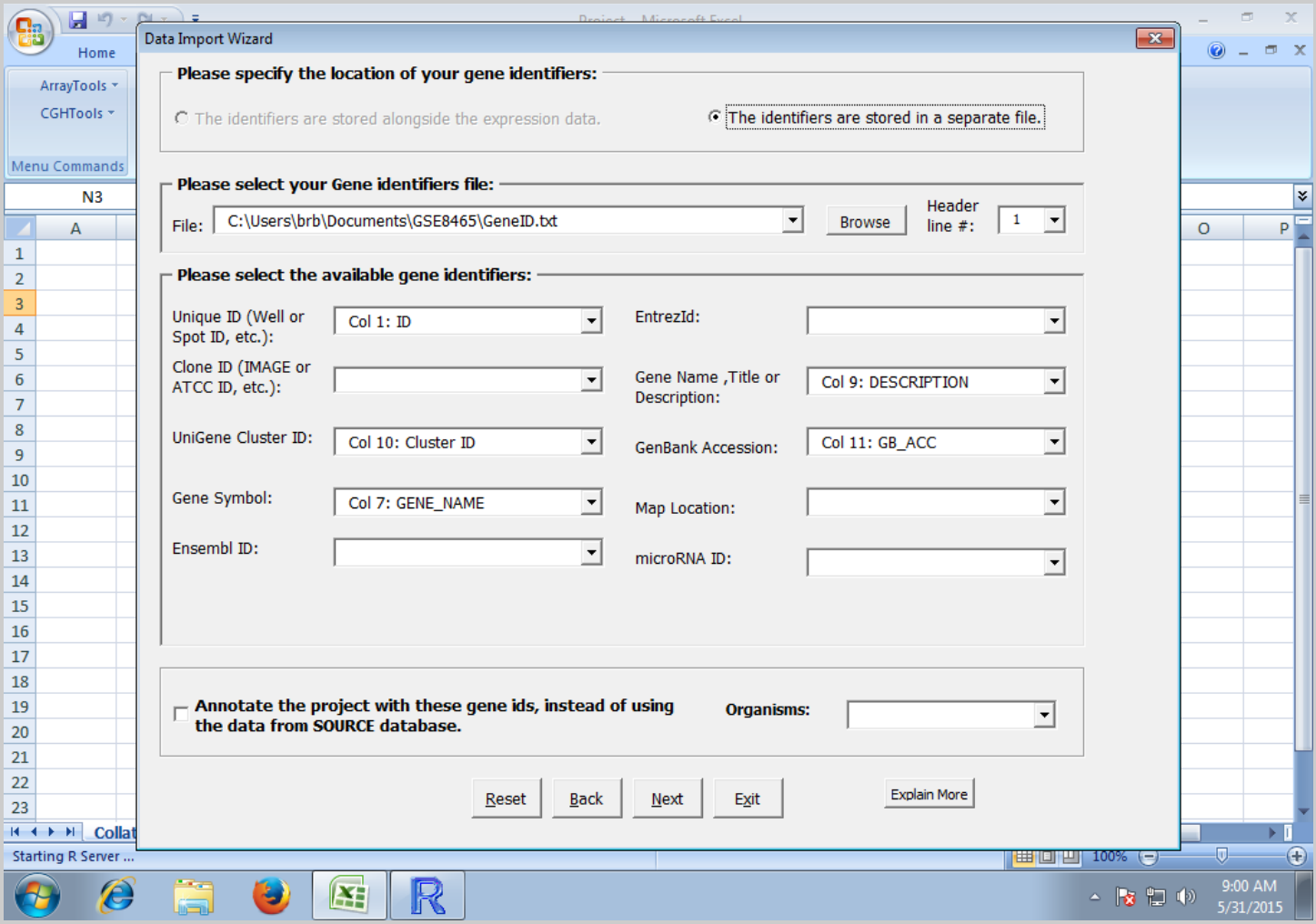
If multiple GPLs was found, it will ask for the GPL number for the series belonging to that one.



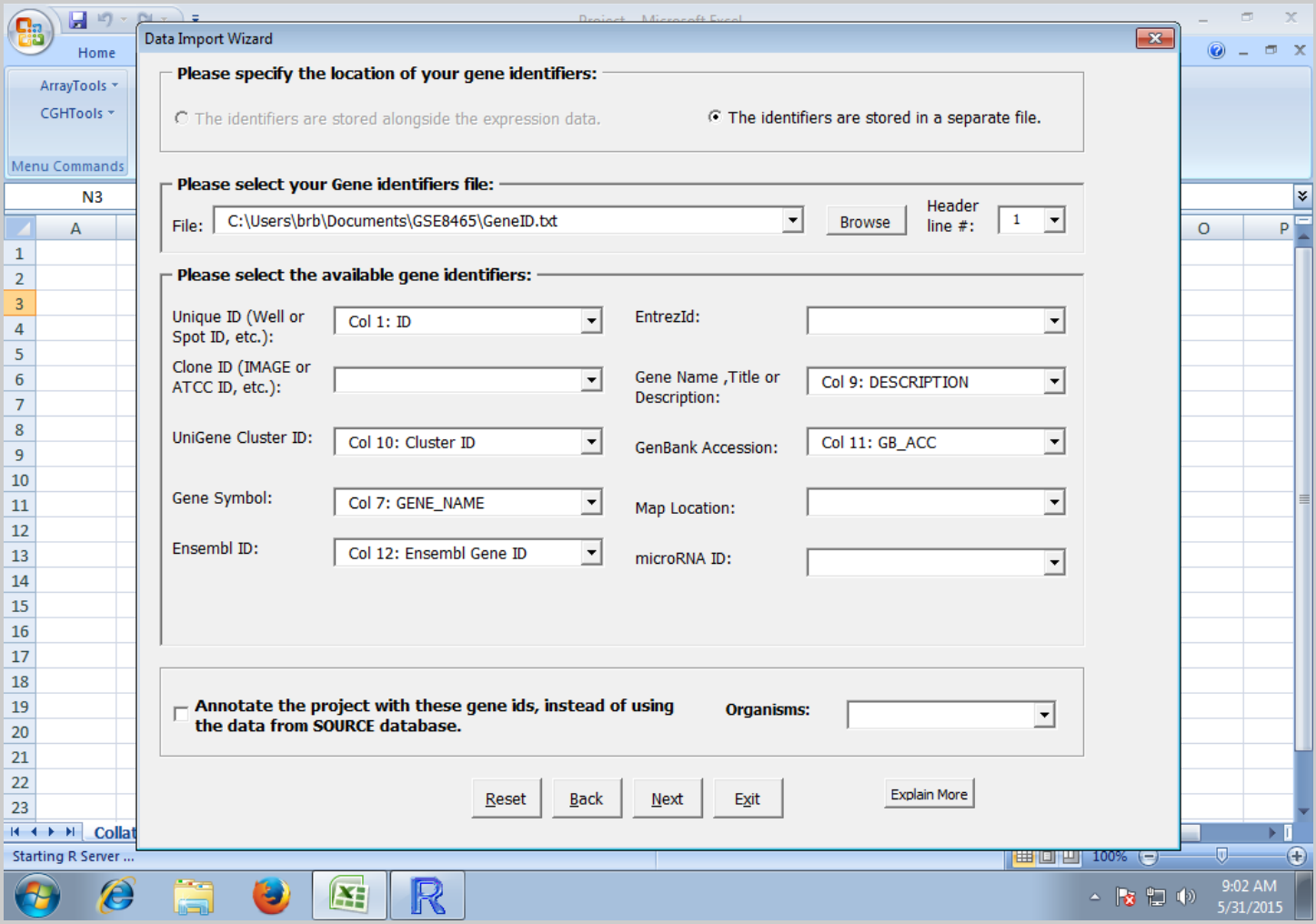

Book1 - Microsoft Excel
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-Ins
ArrayTools
CGHTools
Menu Commands
N3
Microsoft Excel
Do you wish to apply log2 transformation on your data?
Yes
No
Sheet1
Sheet2
Sheet3
Starting R Server ...
100%
8:59 AM
5/31/2015

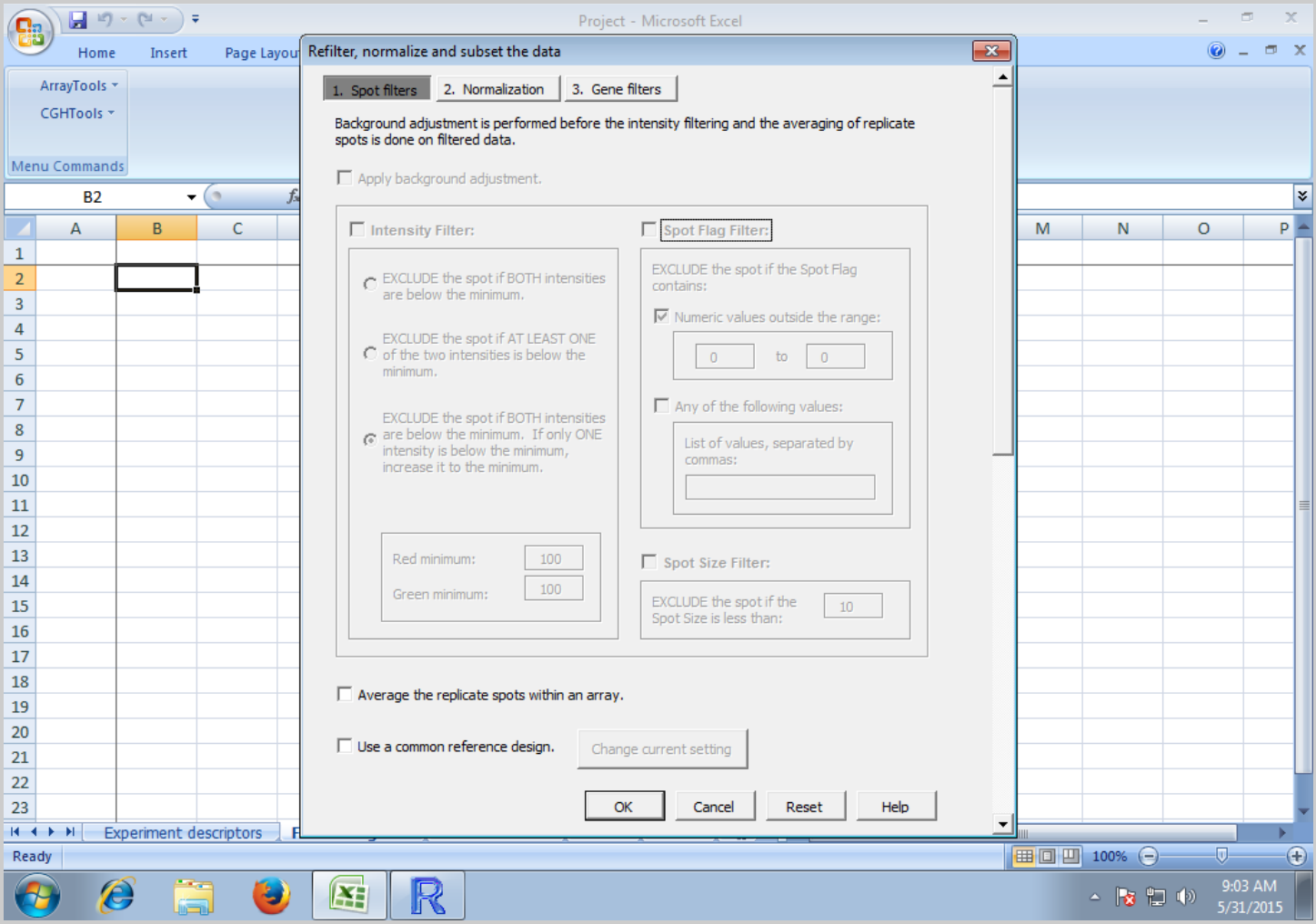

Book1 - Microsoft Excel
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-Ins
ArrayTools
CGHTools
Menu Commands
N3
Data Import Wizard
Project location:
C:\Users\brb\Documents\GSE8465
Project folder:
GSE8465 -Project
Project name:
Project.xlsx
Next
Cancel
Sheet1
Sheet2
Sheet3
Starting R Server ...
100%
9:00 AM
5/31/2015

The available gene identifiers depend on <GeneID.txt> file. In this case, five IDs were used.

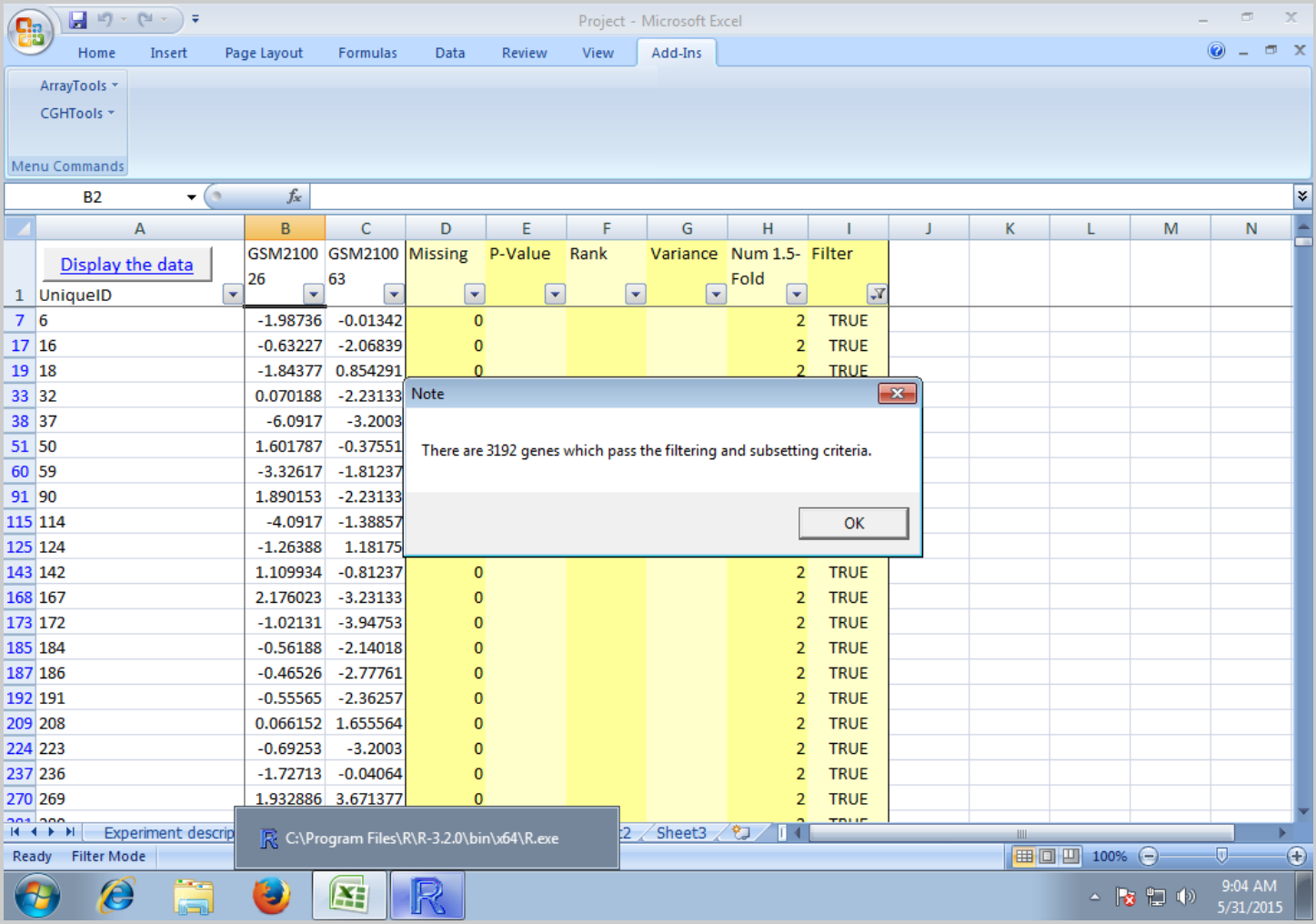


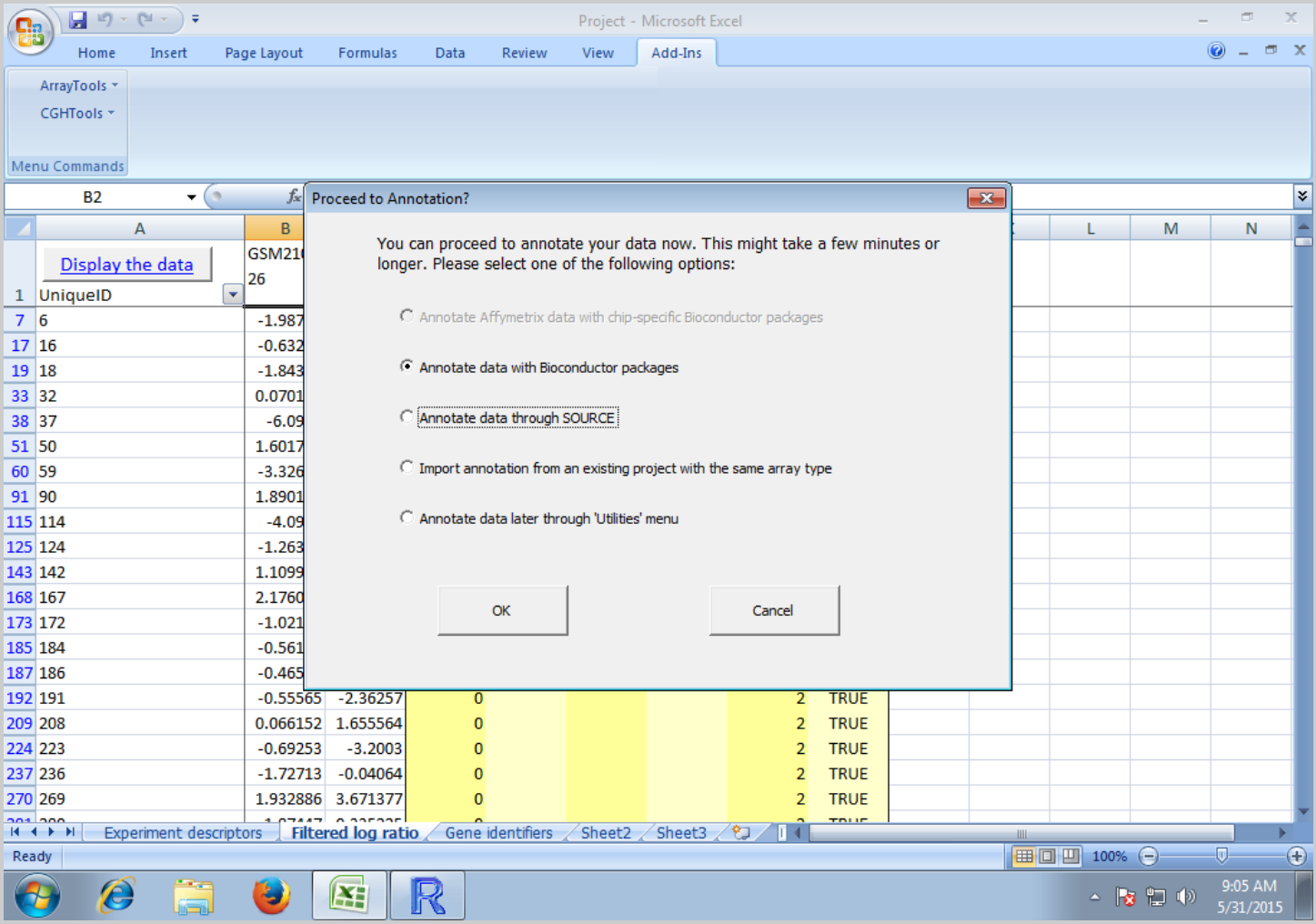
Since the <GeneID.txt> contains Ensembl column, we manually add it to the 'Ensemebl ID' entry.

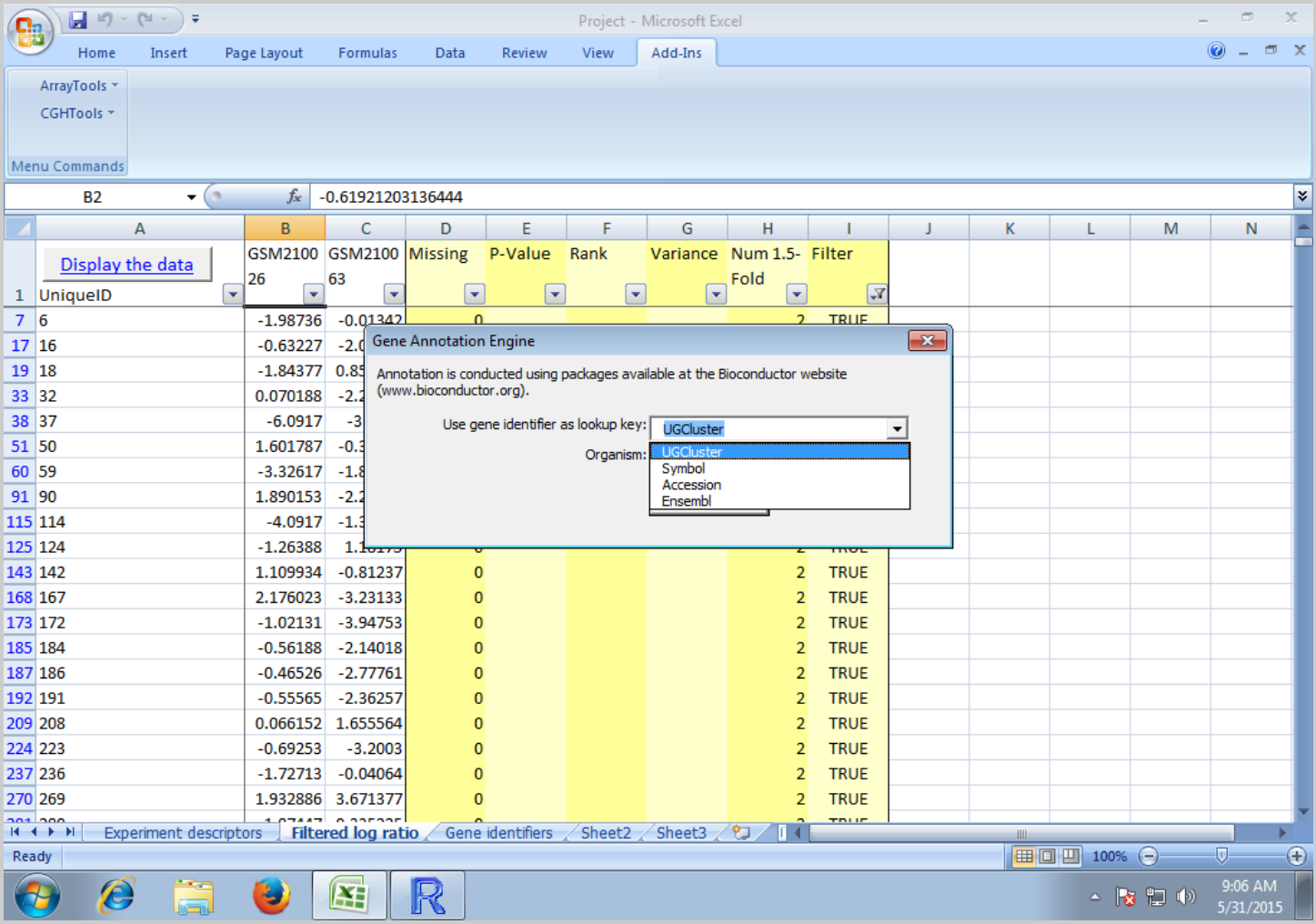



Project - Microsoft Excel
Home
Insert
Page Layou
ArrayTools
CGHTools
Menu Commands
B2
Refilter, normalize and subset the data
1. Spot filters
2. Normalization
3. Gene filters
Background adjustment is performed before the intensity filtering and the averaging of replicate spots is done on filtered data.
Apply background adjustment.
Intensity Filter:
EXCLUDE the spot if BOTH intensities are below the minimum.
EXCLUDE the spot if AT LEAST ONE of the two intensities is below the minimum.
EXCLUDE the spot if BOTH intensities are below the minimum. If only ONE intensity is below the minimum, increase it to the minimum.
Red minimum: 100
Green minimum: 100
Spot Flag Filter:
EXCLUDE the spot if the Spot Flag contains:
Numeric values outside the range:
0 to 0
Any of the following values:
List of values, separated by commas:
Spot Size Filter:
EXCLUDE the spot if the Spot Size is less than: 10
Average the replicate spots within an array.
Use a common reference design.
Change current setting
OK
Cancel
Reset
Help
Experiment descriptors
Ready
100%
9:03 AM
5/31/2015

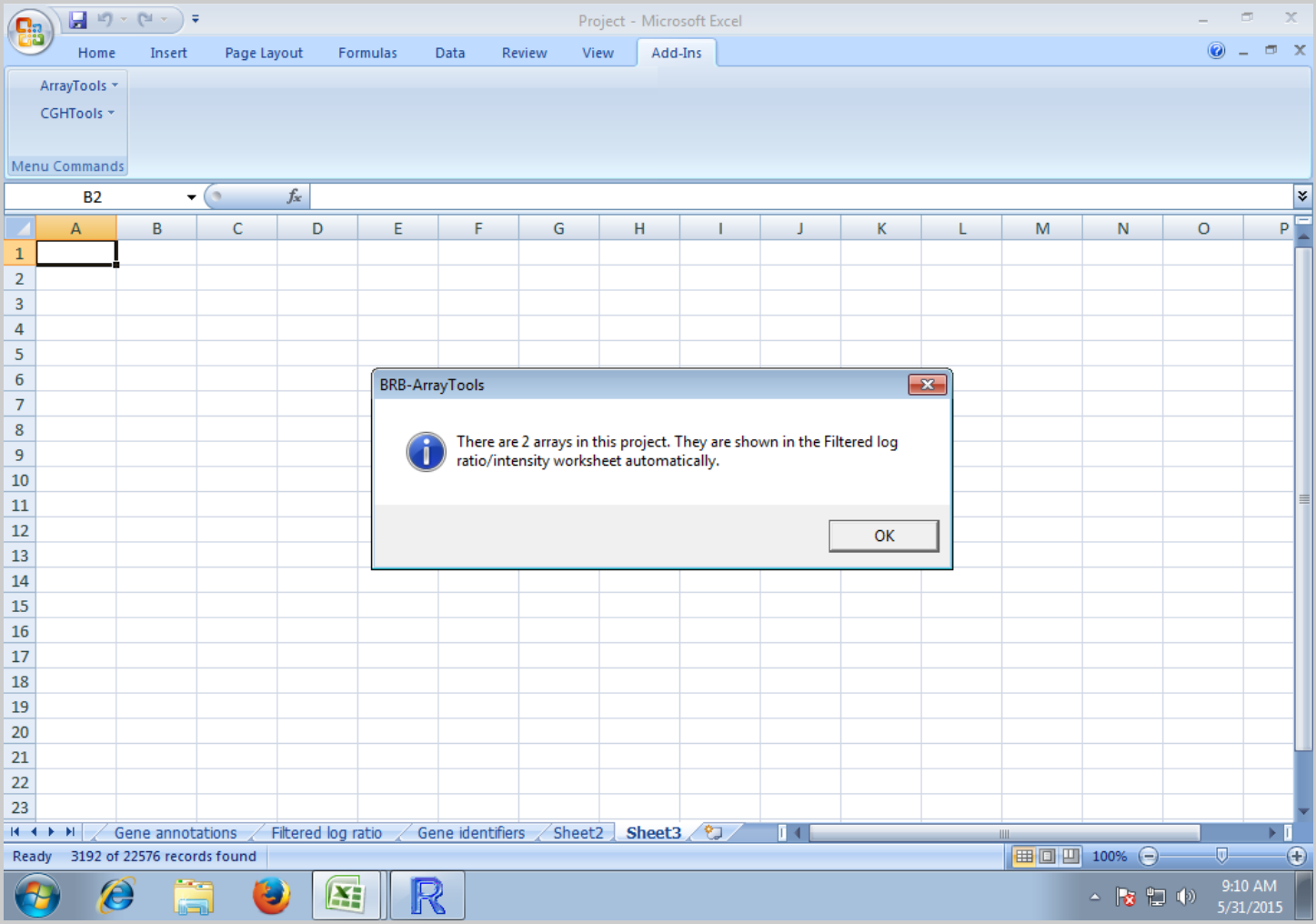
Hit the OK button to continue.



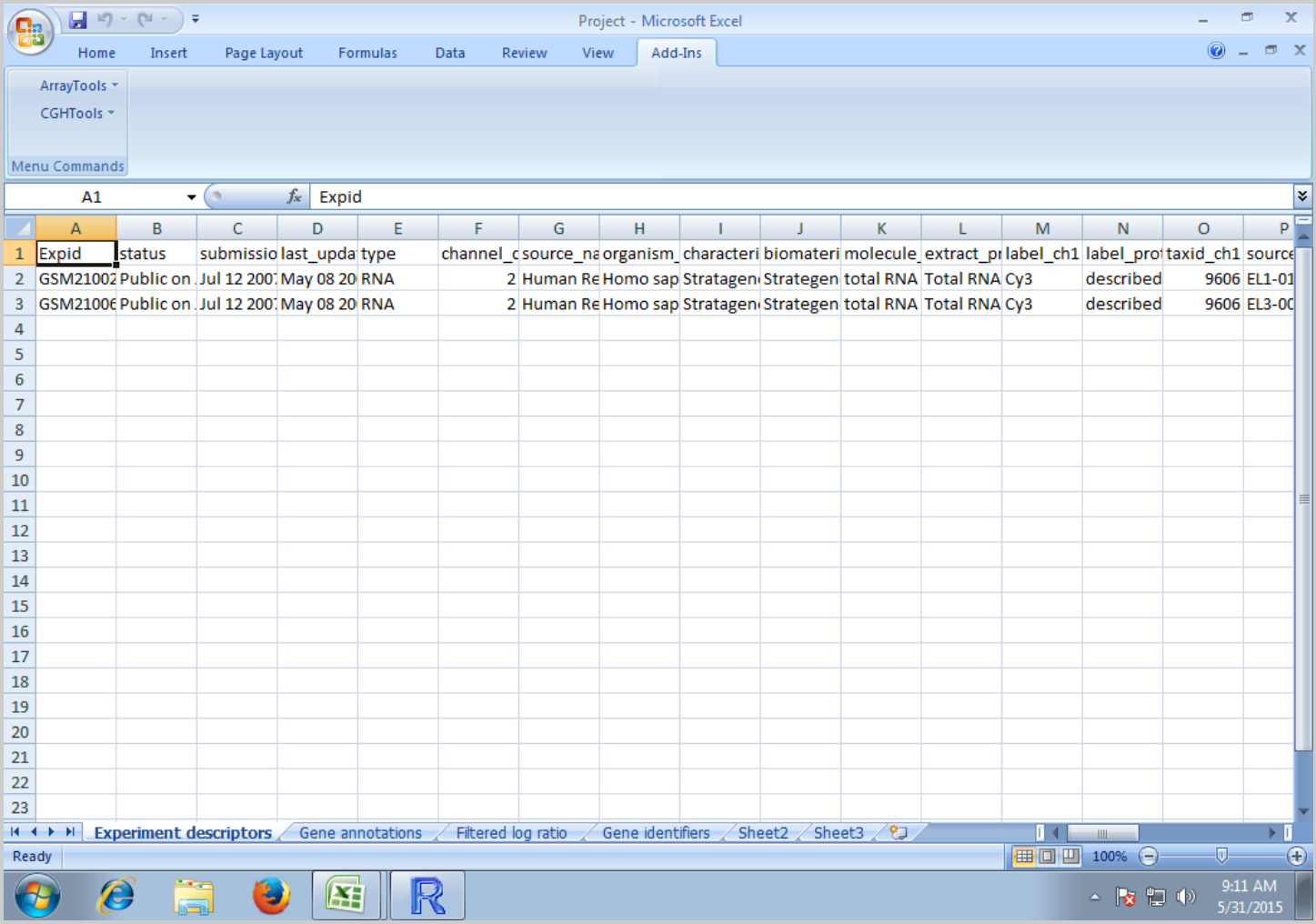

Project - Microsoft Excel
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-Ins
ArrayTools
CGHTools
Menu Commands
Proceed to Annotation?
You can proceed to annotate your data now. This might take a few minutes or longer. Please select one of the following options:
Annotate Affymetrix data with chip-specific Bioconductor packages
Annotate data with Bioconductor packages
Annotate data through SOURCE
Import annotation from an existing project with the same array type
Annotate data later through 'Utilities' menu
OK
Cancel
Display the data
UniqueID
Experiment descriptors
Filtered log ratio
Gene identifiers
Sheet2
Sheet3
Ready
100%
9:05 AM
5/31/2015



Here I pick up the first one 'UGCluster' for the lookup key. BRB-ArrayTools will start to download required R packages now.

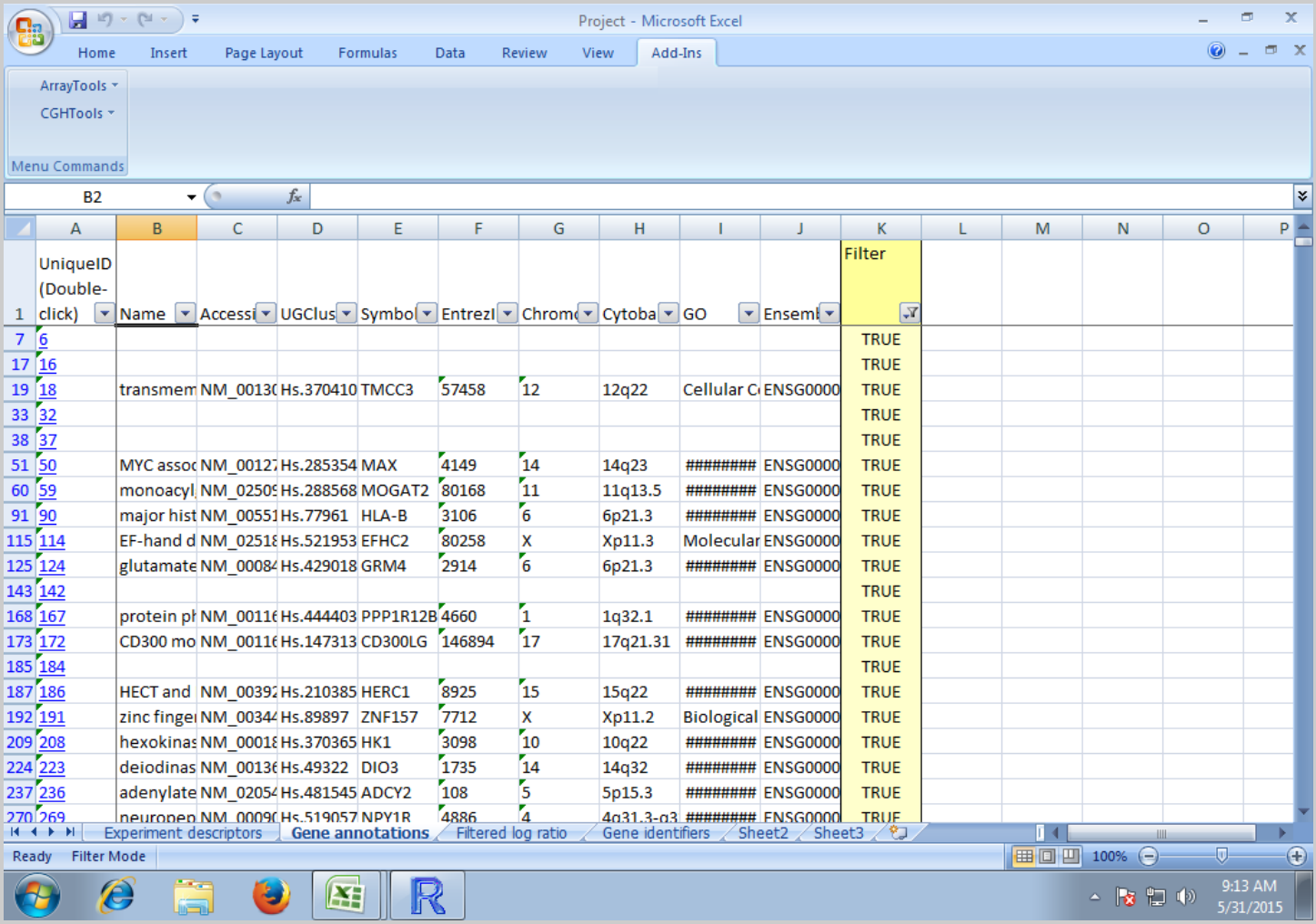

Project - Microsoft Excel
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-Ins
ArrayTools
CGHTools
Menu Commands
B2
BRB-ArrayTools
There are 2 arrays in this project. They are shown in the Filtered log ratio/intensity worksheet automatically.
OK
Gene annotations
Filtered log ratio
Gene identifiers
Sheet2
Sheet3
Ready
3192 of 22576 records found
100%
9:10 AM
5/31/2015

At the end, the importing will be finished and displays the 'Experiment descriptors' worksheet.

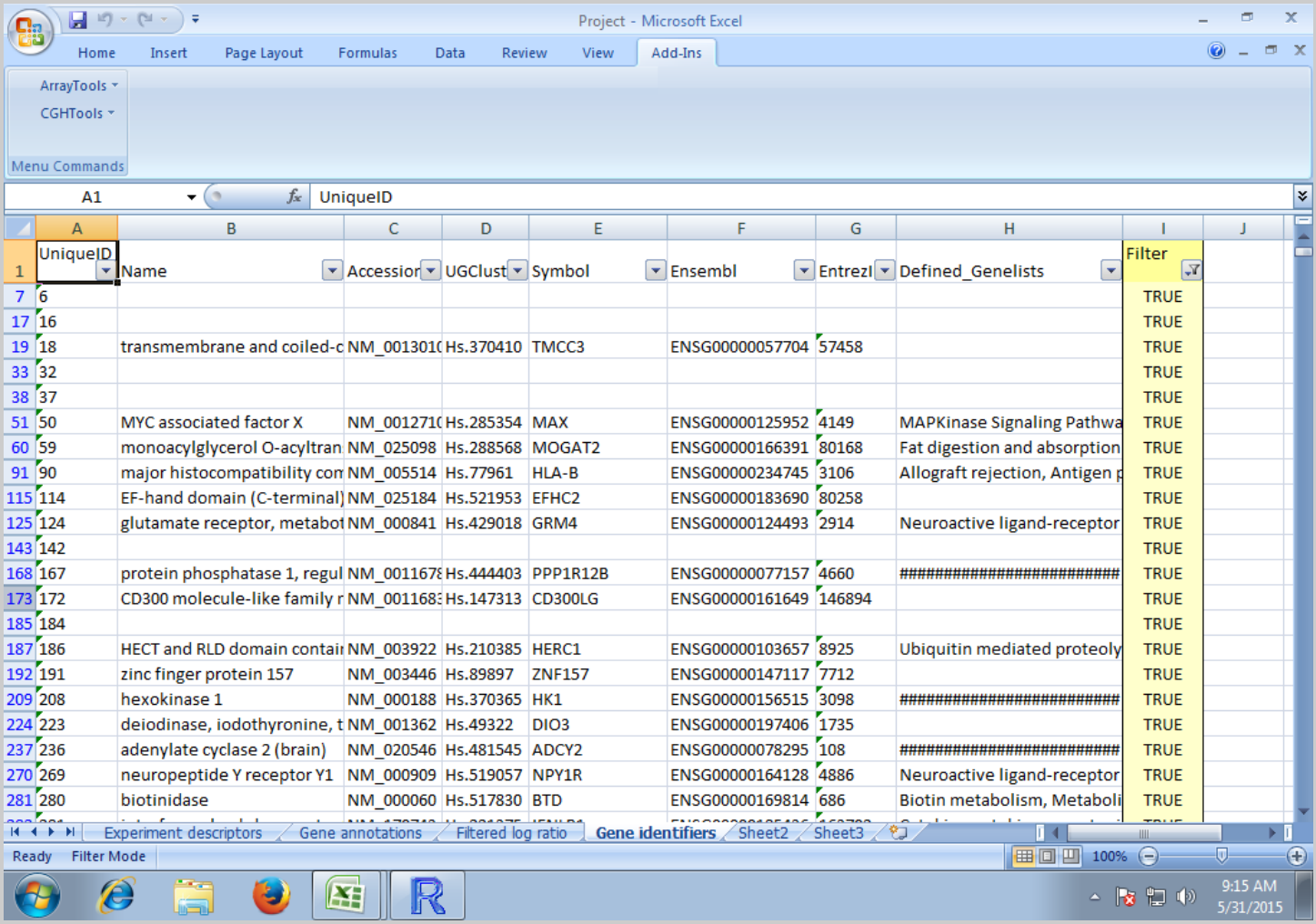


We can now go around different worksheet to see if everything looks right.

'Gene annotations' worksheet.

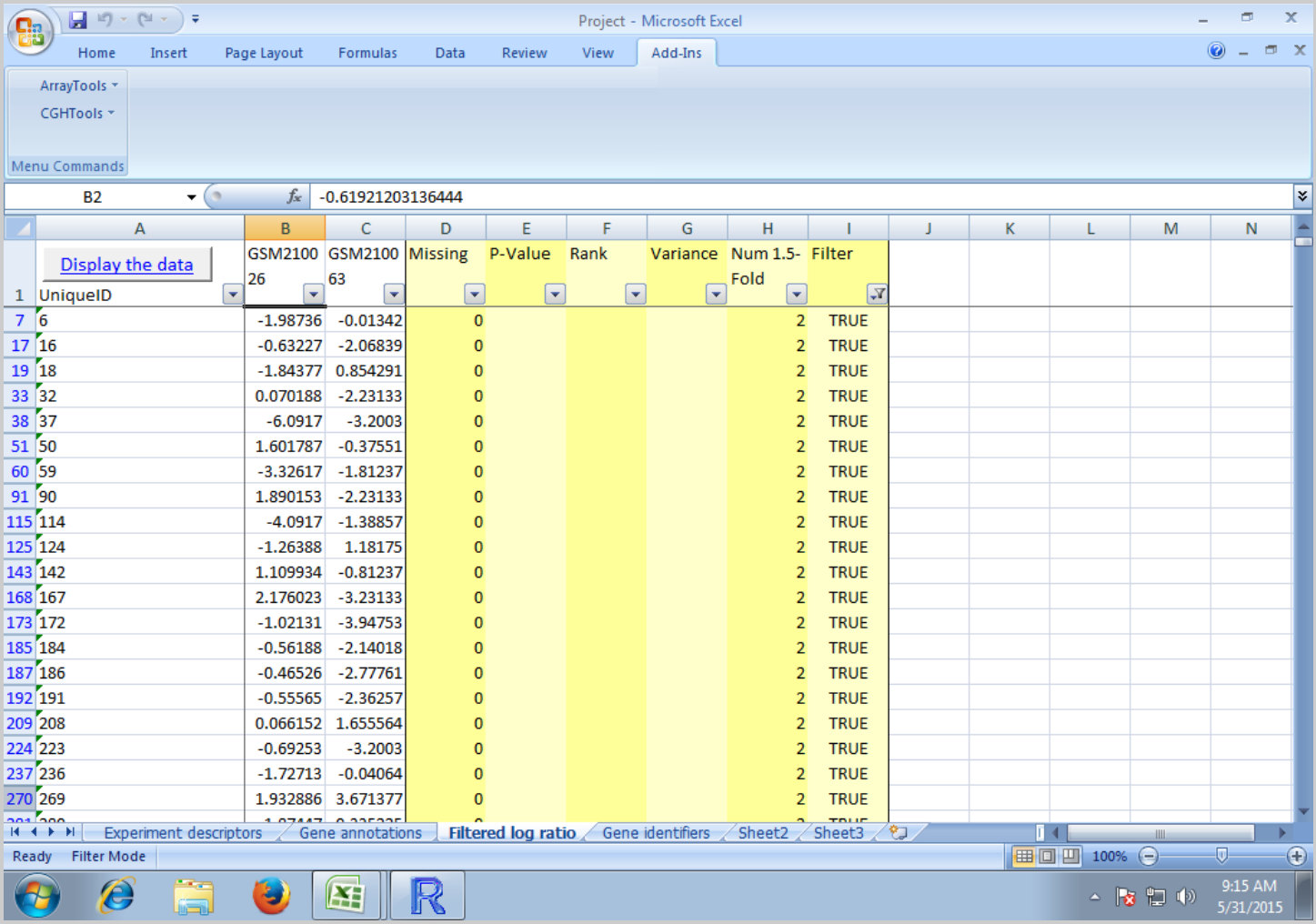


'Gene identifiers' worksheet.



	A	B	C	D	E	F	G	H	I	J
1	UniqueID	Name	Accessior	UGClust	Symbol	Ensembl	Entrezl	Defined_Genelists	Filter	
7	6								TRUE	
17	16								TRUE	
19	18	transmembrane and coiled-c	NM_001301(	Hs.370410	TMCC3	ENSG00000057704	57458		TRUE	
33	32								TRUE	
38	37								TRUE	
51	50	MYC associated factor X	NM_001271(	Hs.285354	MAX	ENSG00000125952	4149	MAPKinase Signaling Pathwa	TRUE	
60	59	monoacylglycerol O-acyltran	NM_025098	Hs.288568	MOGAT2	ENSG00000166391	80168	Fat digestion and absorption	TRUE	
91	90	major histocompatibility com	NM_005514	Hs.77961	HLA-B	ENSG00000234745	3106	Allograft rejection, Antigen p	TRUE	
115	114	EF-hand domain (C-terminal)	NM_025184	Hs.521953	EFHC2	ENSG00000183690	80258		TRUE	
125	124	glutamate receptor, metabot	NM_000841	Hs.429018	GRM4	ENSG00000124493	2914	Neuroactive ligand-receptor	TRUE	
143	142								TRUE	
168	167	protein phosphatase 1, regul	NM_001167{	Hs.444403	PPP1R12B	ENSG00000077157	4660	##########	TRUE	
173	172	CD300 molecule-like family r	NM_001168:	Hs.147313	CD300LG	ENSG00000161649	146894		TRUE	
185	184								TRUE	
187	186	HECT and RLD domain contair	NM_003922	Hs.210385	HERC1	ENSG00000103657	8925	Ubiquitin mediated proteoly	TRUE	
192	191	zinc finger protein 157	NM_003446	Hs.89897	ZNF157	ENSG00000147117	7712		TRUE	
209	208	hexokinase 1	NM_000188	Hs.370365	HK1	ENSG00000156515	3098	##########	TRUE	
224	223	deiodinase, iodothyronine, t	NM_001362	Hs.49322	DIO3	ENSG00000197406	1735		TRUE	
237	236	adenylate cyclase 2 (brain)	NM_020546	Hs.481545	ADCY2	ENSG00000078295	108	##########	TRUE	
270	269	neuropeptide Y receptor Y1	NM_000909	Hs.519057	NPY1R	ENSG00000164128	4886	Neuroactive ligand-receptor	TRUE	
281	280	biotinidase	NM_000060	Hs.517830	BTD	ENSG00000169814	686	Biotin metabolism, Metaboli	TRUE	

'Filtered log ratio' worksheet.



That's it!