SmartArray

User Manual

Bioada Team[1]

October 8, 2021

[1] www.bioada.com

Contents

1

Introduction

There are many factors that make data preparation challenging, from understanding where to find the data to extracting it, then properly formatting it and finally loading it to a database management system (DBMS). SmartArray reduces data preparation time, so that you the focus is on using a high quality and enriched data for visualization, statistical analysis and predictive modeling. With Bioada, users have the full control on the data.

SmartArray provides a unique interactive data exploration platform for ad-hoc queries, data visualizations and statistical analysis. Users can explore data interactively and apply statistical models and analytical techniques to find potential biomarkers in the genomic data.Users can focus and emphasize on interactivity and effective integration of techniques from data science.

Genomic datasets are high-dimensional molecular profiles pose challenges to data interpretation and hypothesis generation. SmartArray has a method that discovers significantly enriched pathways and gene ontology terms across datasets. SmartArray enrichment is a versatile method that improves systems-level understanding of cellular entities in health and disease through integration of genomic datasets and pathway and gene ontology annotations.

2

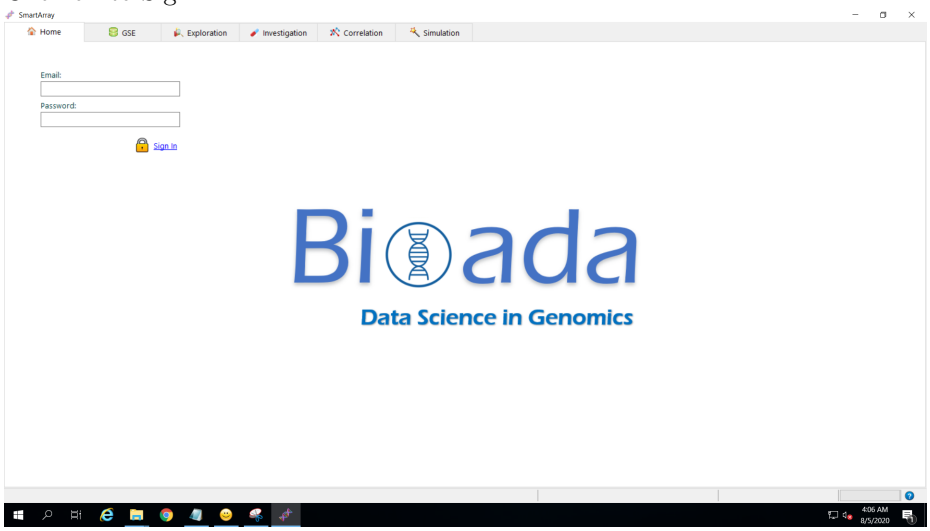
Home

The Home tab is the first screen of SmartArray. Every user is given a set of credentials to login to the application according to their roles.

2.1 Login

To login into the SmartArray application please enter the following details:

1. Enter Email and Password in the related fields.
2. Click on to Sign-In.

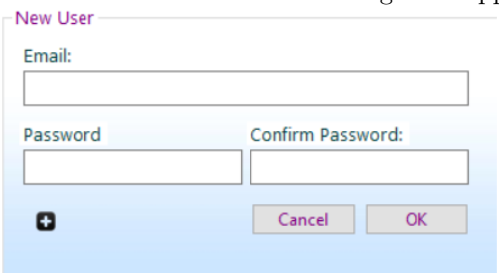


2.2 Add user

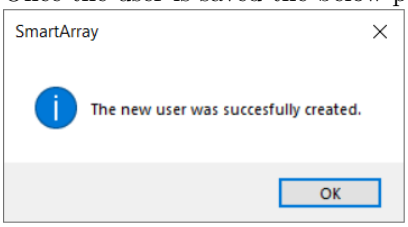
Admins have access to add/edit/delete users.

1. Sign-in into the application.

2. Click on "User" icon.A message box appears as below.



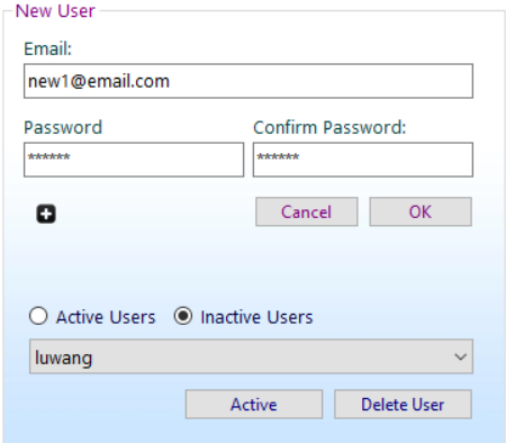
3. Enter the new email, password, and confirm password in the related fields. Once the user is saved the below popup appears



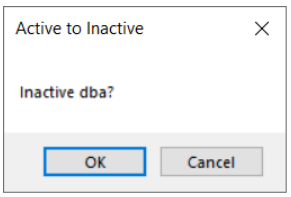
2.3 Edit user

This feature activates, inactivates, or deletes a user.

- **Activate User:**
  1. Sign-in to the application.
  2. Click on "User" icon.
  3. Click on the button.All the Inactive users will be displayed under "Inactive Users" list.



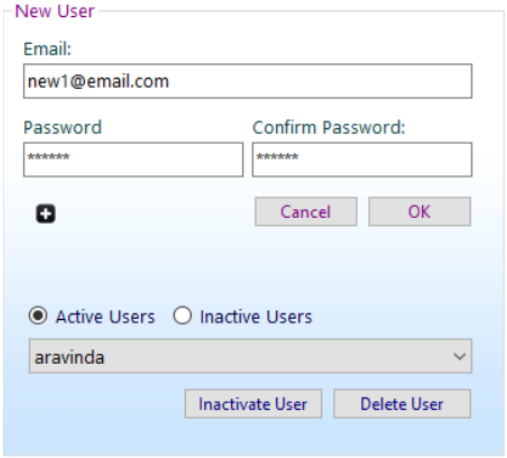
  4. Select an user to Activate and click on the Active button.A confirmation message appears as shown below.



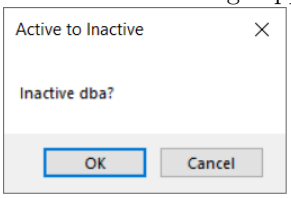
5. Once the user is activated, it will be displayed under active users list.

- **Inactivate User:**

1. Sign-in to the application.
2. Click on "User" icon.
3. Click on the button.All the active users will be displayed under "Active Users" list.



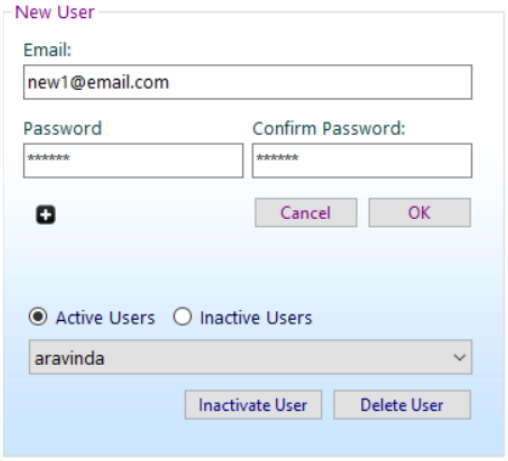
4. Select an user to inactivate and click on the Inactivate User button.A confirmation message appears as shown below.



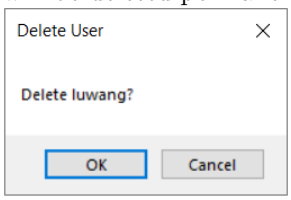
5. Once the user is inactivated, it will be displayed under inactive users list.

- **Delete User:**

1. Sign-in to the application.
2. Click on "User" icon.
3. Click on the button.All the active users will be displayed under "Active Users" and inactive users on "Inactive Users" related lists.



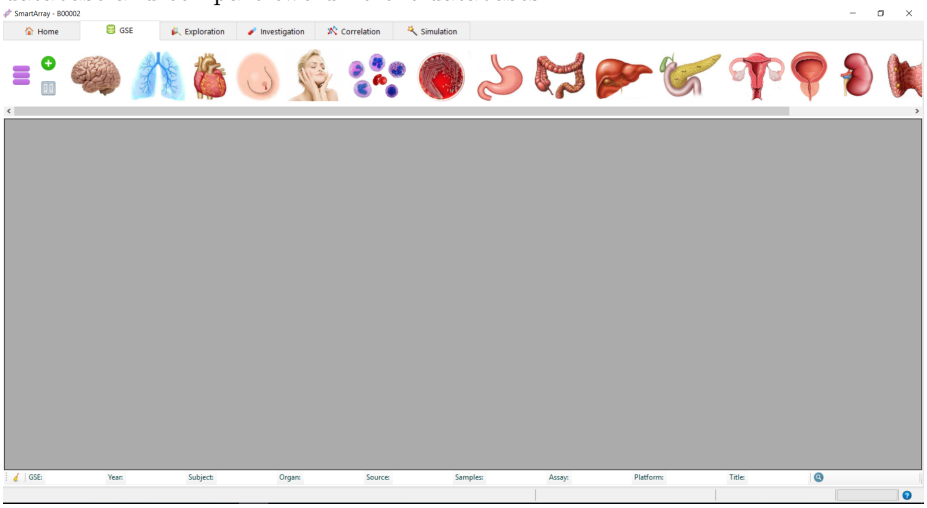
4. Select an user to delete and click on the Delete User button.A confirmation message appears as shown below.On click of "OK" the user will be deleted permanently.



3

Data

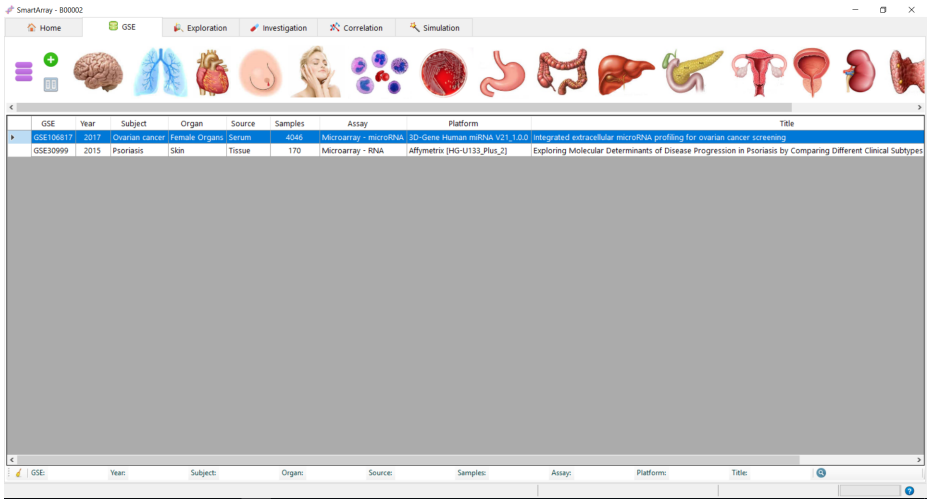
The Data(GSE) tab allows you to load data from database,add new data to database and compare two different databases.



3.1 Load data

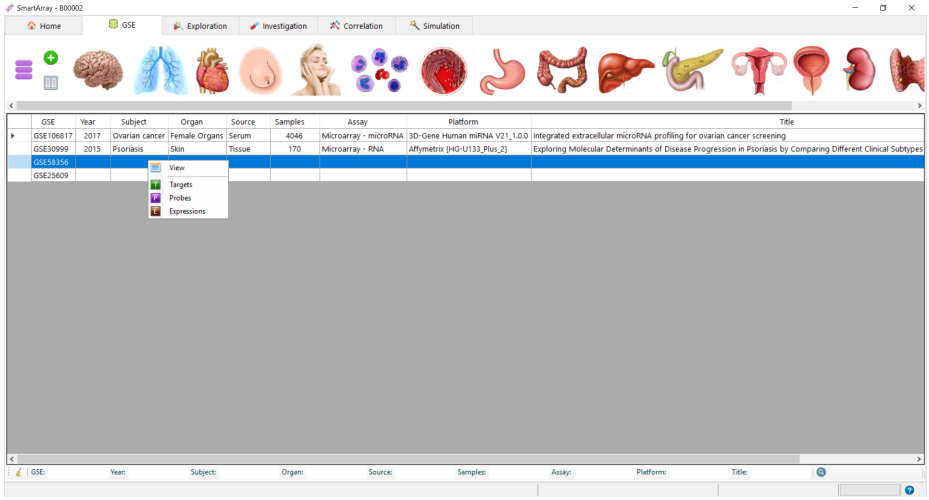
Loads all the existing databases.

1. Click on button.It loads the existing databases in the local database server.

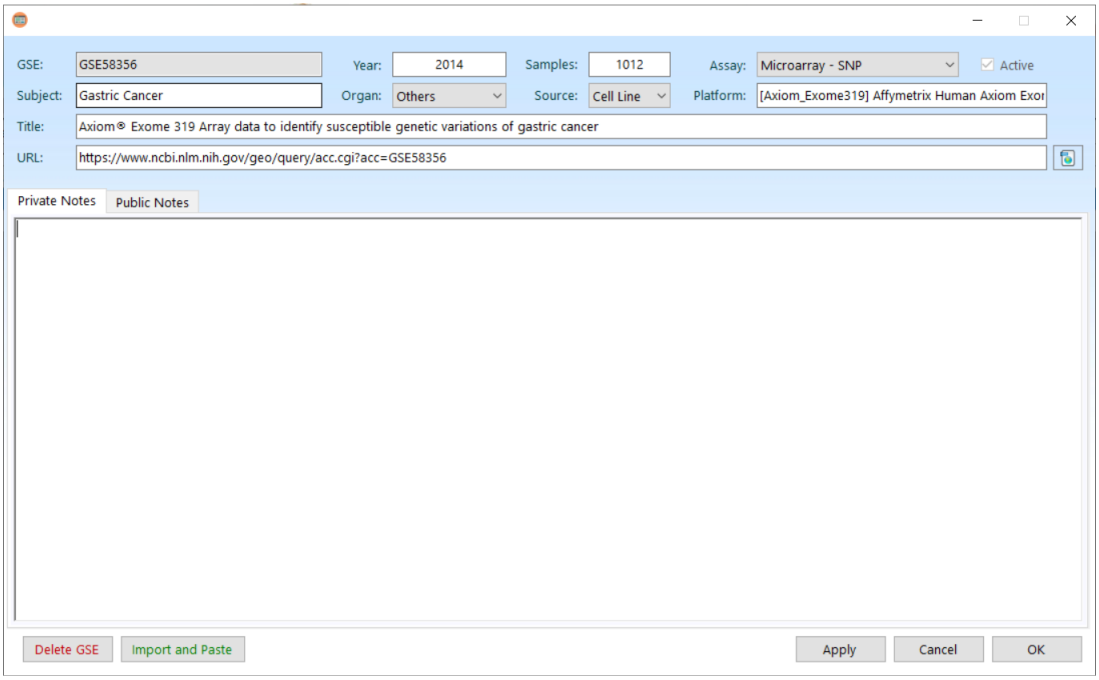


2. The multiple filters(GSE, Year, Subject, Organ, Source, Samples, Assay, Platform, Title) at the bottom of the screen can be used to filter the databases.

3. Right click on any GSE to view the details.

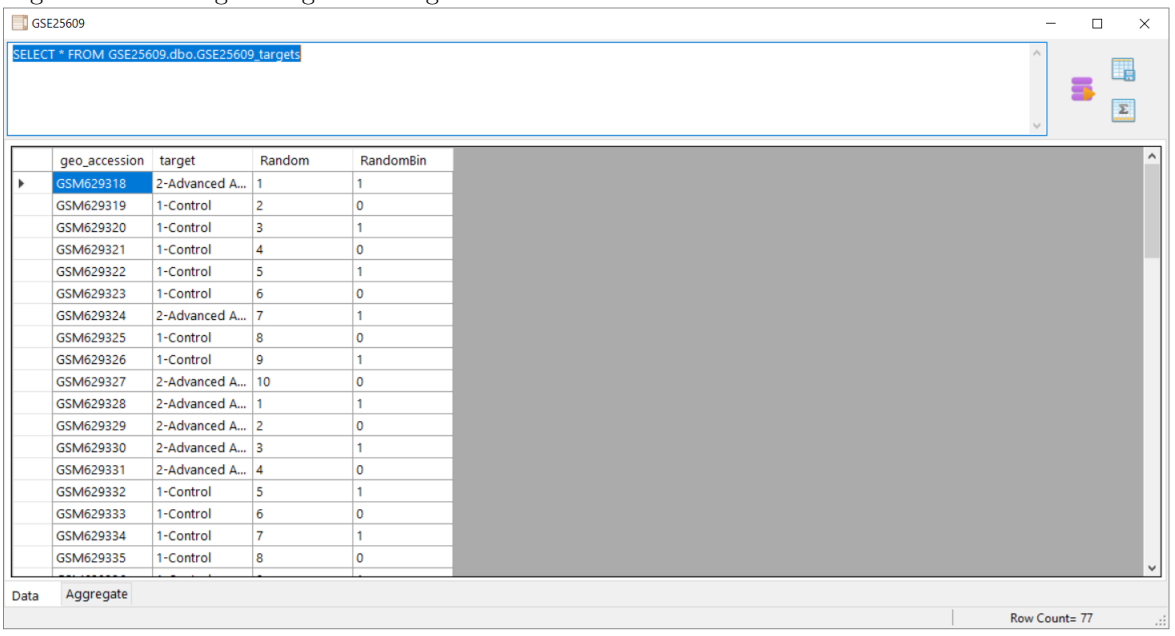


4. Enter the details of the data-set and click on Apply/OK.



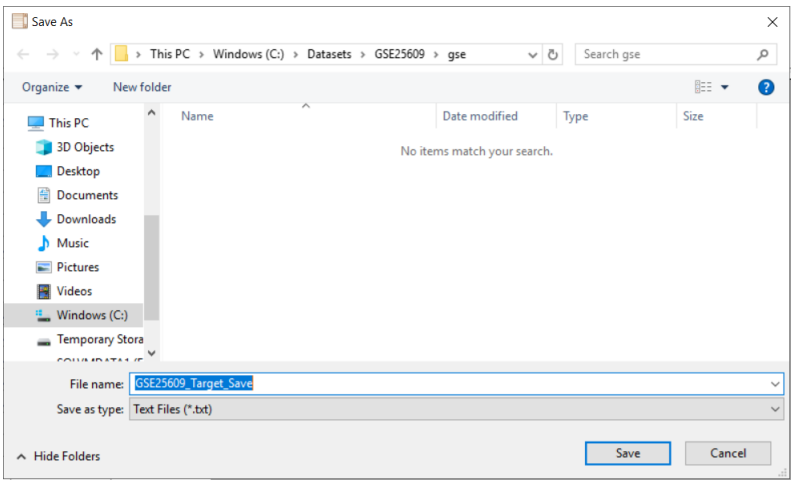
5. Delete GSE button can be used to delete the GSE from the database server.

6. Right click on Targets to get the target file details.



7. Click on  to run the query and the results are as shown above.

8. Click on  to save the query results.



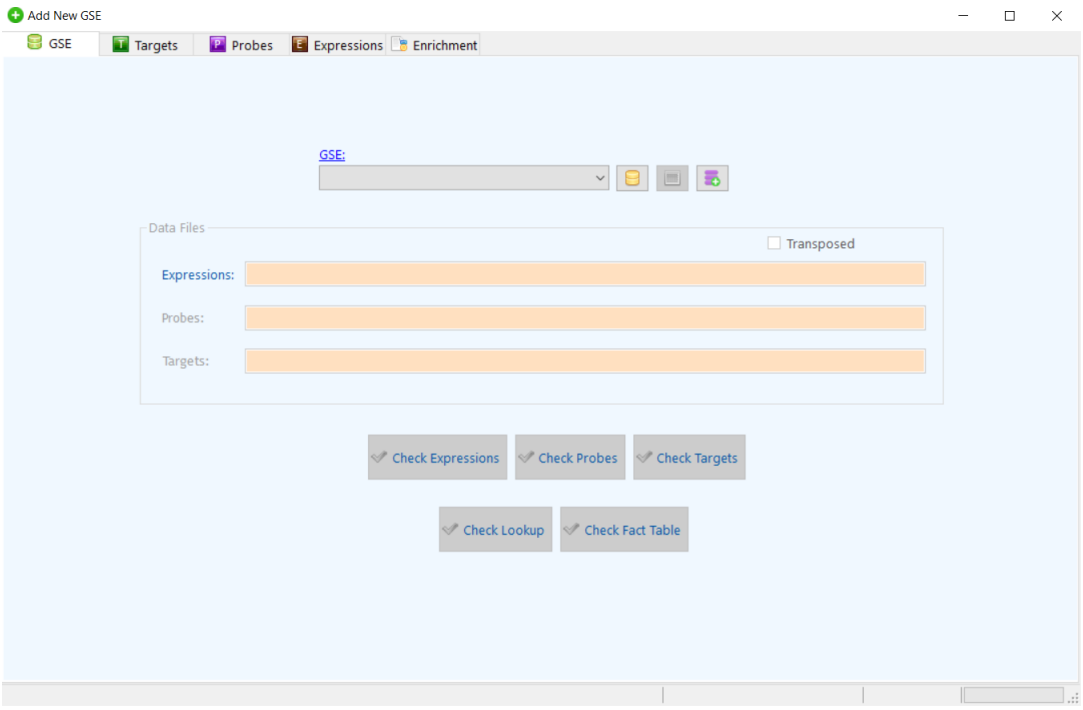
9. Click on  to get the aggregate of the selected column.

10. Similar results are obtained for Probes and Expressions.

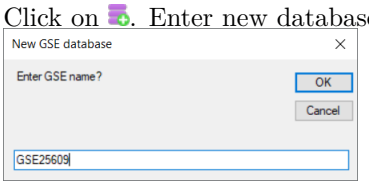
3.2 New Data

Adds a new database to the local database server.

1. Click on  button.A new window is opened as shown below.

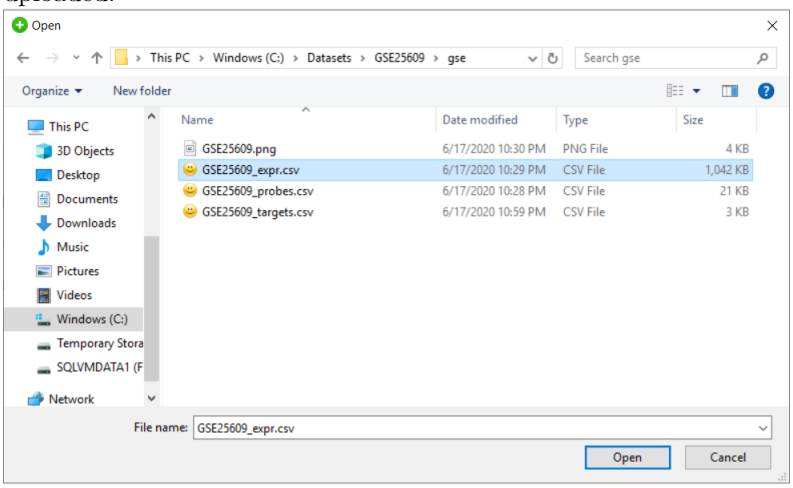


2. Click on . Enter new database name and click OK.

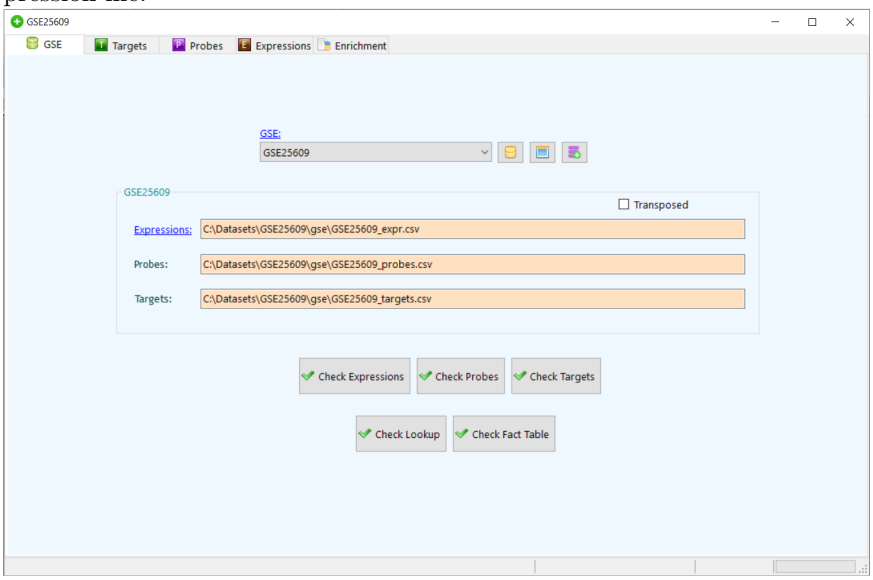


3. The newly added database appears in the GSE list. Select the database and click on to connect.

4. Click on the Expressions hyperlink and select the expression file to be uploaded.

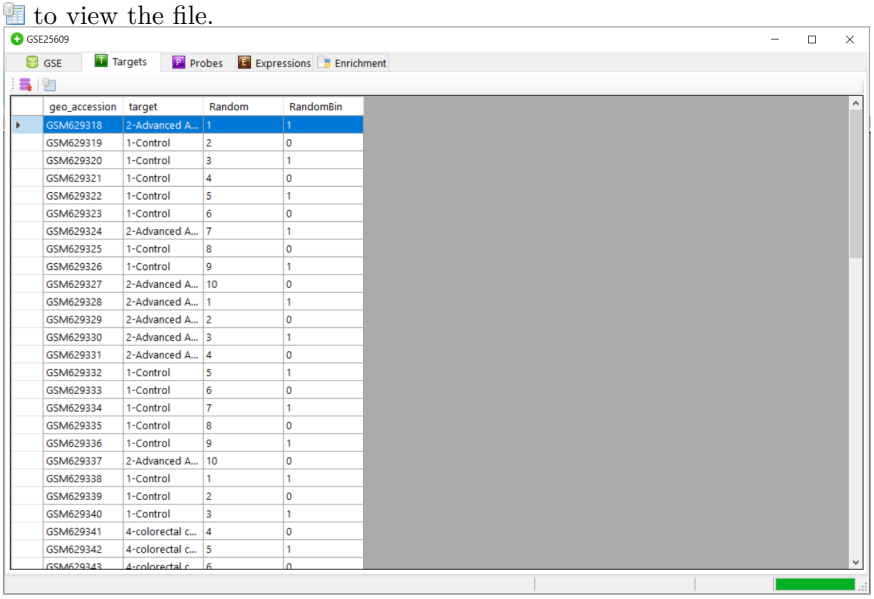


5. The probes and target files are automatically uploaded on selection of expression file.

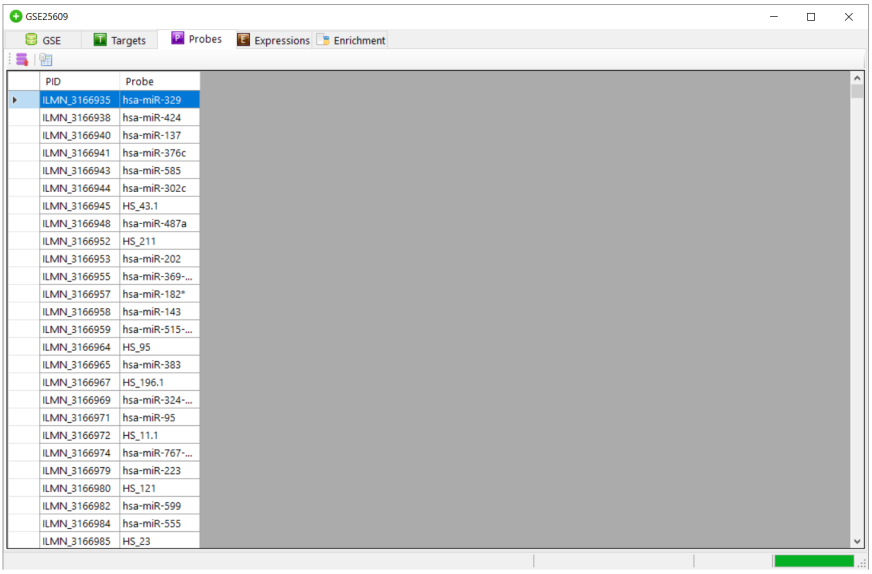


6. Click on various Check buttons to ensure all the files are intact.

7. Go to Targets tab and click on to upload the data to database.Click on to view the file.

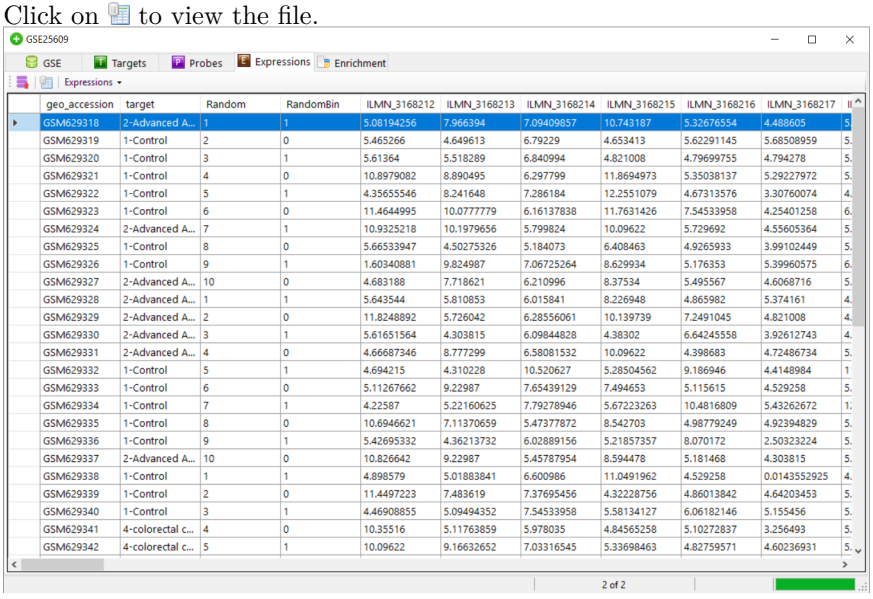


8. Go to Probes tab and click on to upload the data to database.Click on to view the file.



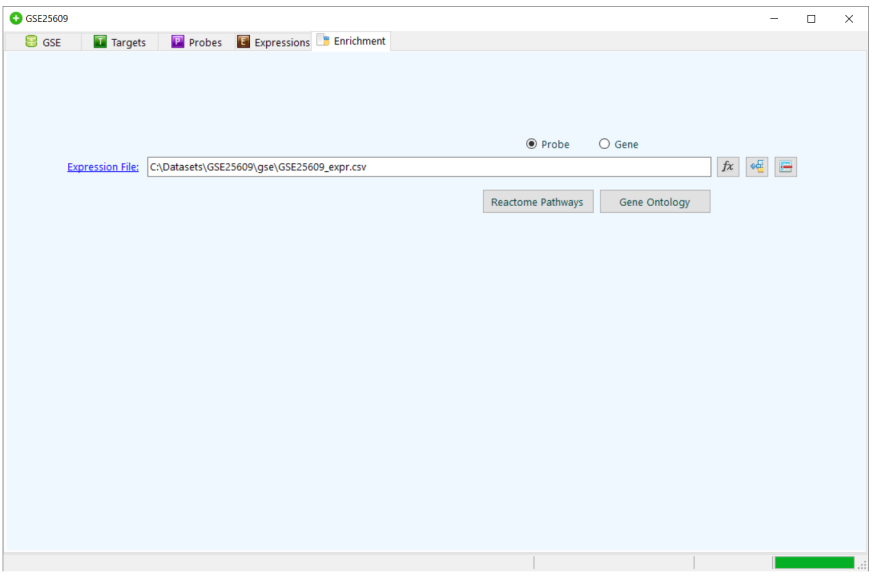
PID	Probe
ILMN_3166935	hsa-miR-329
ILMN_3166938	hsa-miR-424
ILMN_3166940	hsa-miR-137
ILMN_3166941	hsa-miR-376c
ILMN_3166943	hsa-miR-585
ILMN_3166944	hsa-miR-302c
ILMN_3166945	HS_43.1
ILMN_3166948	hsa-miR-487a
ILMN_3166952	HS_211
ILMN_3166953	hsa-miR-202
ILMN_3166955	hsa-miR-369-...
ILMN_3166957	hsa-miR-182*
ILMN_3166958	hsa-miR-143
ILMN_3166959	hsa-miR-515-...
ILMN_3166964	HS_95
ILMN_3166965	hsa-miR-383
ILMN_3166967	HS_196.1
ILMN_3166969	hsa-miR-324-...
ILMN_3166971	hsa-miR-95
ILMN_3166972	HS_11.1
ILMN_3166974	hsa-miR-767-...
ILMN_3166979	hsa-miR-223
ILMN_3166980	HS_121
ILMN_3166982	hsa-miR-599
ILMN_3166984	hsa-miR-555
ILMN_3166985	HS_23

9. Go to Expressions tab and click on to upload the data to database. Click on to view the file.



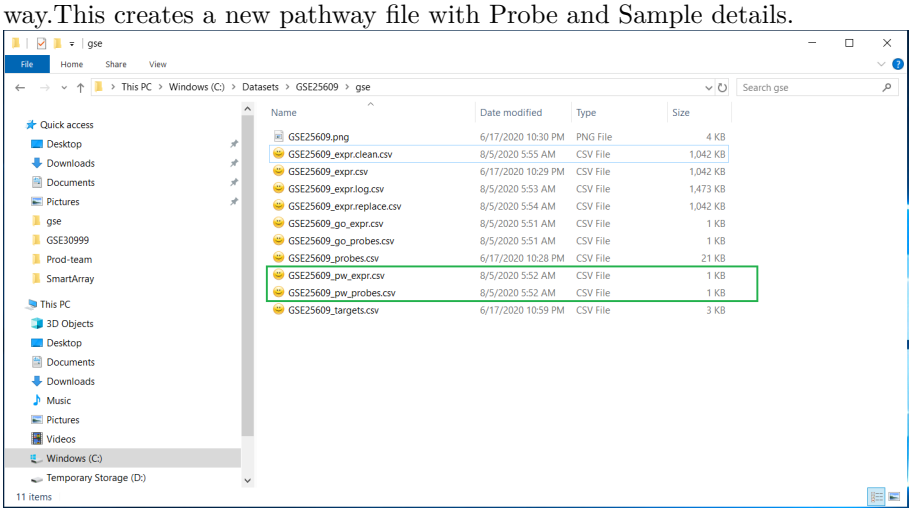
geo_accession	target	Random	RandomBin	ILMN_3168212	ILMN_3168213	ILMN_3168214	ILMN_3168215	ILMN_3168216	ILMN_3168217
GSM629318	2-Advanced A...	1	1	5.08194256	7.966394	7.09409857	10.743187	5.32676554	4.488605
GSM629319	1-Control	2	0	5.465266	4.649613	6.79229	4.653413	5.62291145	5.68508959
GSM629320	1-Control	3	1	5.61364	5.518289	6.840994	4.821008	4.79699755	4.794278
GSM629321	1-Control	4	0	10.8979082	8.890495	6.297799	11.8694973	5.35038137	5.29227972
GSM629322	1-Control	5	1	4.35655546	8.241648	7.286184	12.2551079	4.67313576	3.30760074
GSM629323	1-Control	6	0	11.4644995	10.0777779	6.16137838	11.7631426	7.54533958	4.25401258
GSM629324	2-Advanced A...	7	1	10.9325218	10.1979656	5.799824	10.09622	5.729692	4.55605364
GSM629325	1-Control	8	0	5.66533947	4.50275326	5.184073	6.408463	4.9265933	3.99102449
GSM629326	1-Control	9	1	1.60340881	9.824987	7.06725264	8.629934	5.176353	5.39960575
GSM629327	2-Advanced A...	10	0	4.683188	7.718621	6.210996	8.37534	5.495567	4.6068716
GSM629328	2-Advanced A...	1	1	5.643544	5.810853	6.015841	8.226948	4.865982	5.374161
GSM629329	2-Advanced A...	2	0	11.8248892	5.726042	6.28556061	10.139739	7.2491045	4.821008
GSM629330	2-Advanced A...	3	1	5.61651564	4.303815	6.09844828	4.38302	6.64245558	3.92612743
GSM629331	2-Advanced A...	4	0	4.66687346	8.777299	6.58081532	10.09622	4.398683	4.72486734
GSM629332	1-Control	5	1	4.694215	4.310228	10.520627	5.28504562	9.186946	4.4148984
GSM629333	1-Control	6	0	5.11267662	9.22987	7.65439129	7.494653	5.115615	4.529258
GSM629334	1-Control	7	1	4.22587	5.22160625	7.79278946	5.67223263	10.4816809	5.43262672
GSM629335	1-Control	8	0	10.6946621	7.11370659	5.47377872	8.542703	4.98779249	4.92394829
GSM629336	1-Control	9	1	5.42695332	4.36213732	6.02889156	5.21857357	8.070172	2.50323224
GSM629337	2-Advanced A...	10	0	10.826642	9.22987	5.45787954	8.594478	5.181468	4.303815
GSM629338	1-Control	1	1	4.898579	5.01883841	6.600986	11.0491962	4.529258	0.0143552925
GSM629339	1-Control	2	0	11.4497223	7.483619	7.37695456	4.32228756	4.86013842	4.64203453
GSM629340	1-Control	3	1	4.46908855	5.09494352	7.54533958	5.58134127	6.06182146	5.155456
GSM629341	4-colorectal c...	4	0	10.35516	5.11763859	5.978035	4.84565258	5.10272837	3.256493
GSM629342	4-colorectal c...	5	1	10.09622	9.16632652	7.03316545	5.33698463	4.82759571	4.60236931

10. Go to Enrichment tab and click on Expression File hyperlink to select the expression file to transform the data.

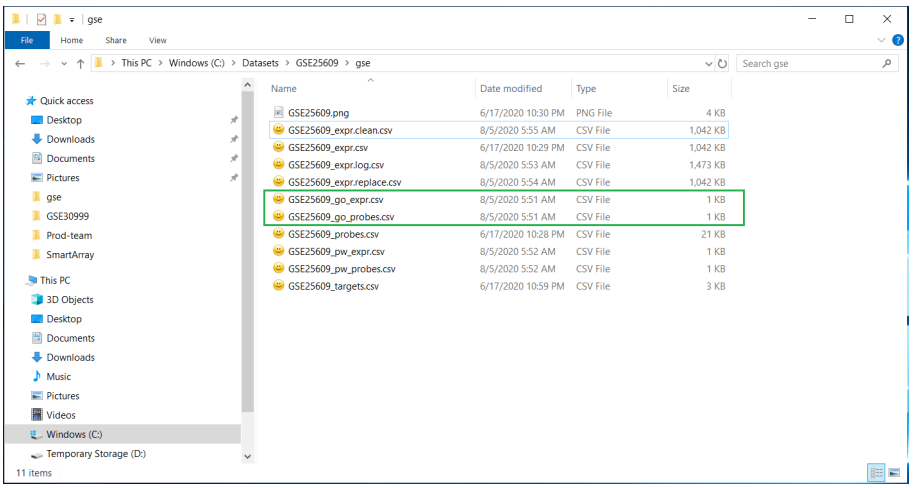


11. Select on the Probe/Gene depending on the data.

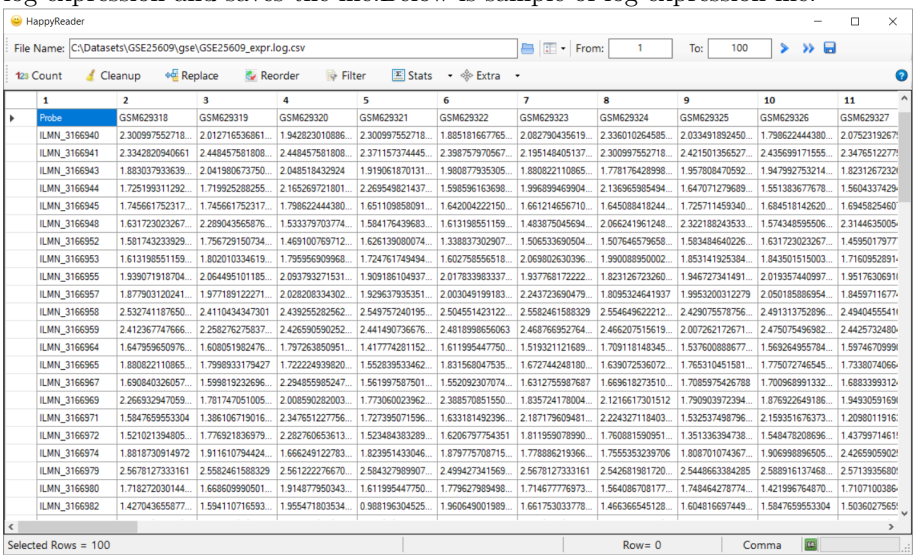
12. Click on Reactome Pathway button to compress the data to higher pathway.This creates a new pathway file with Probe and Sample details.



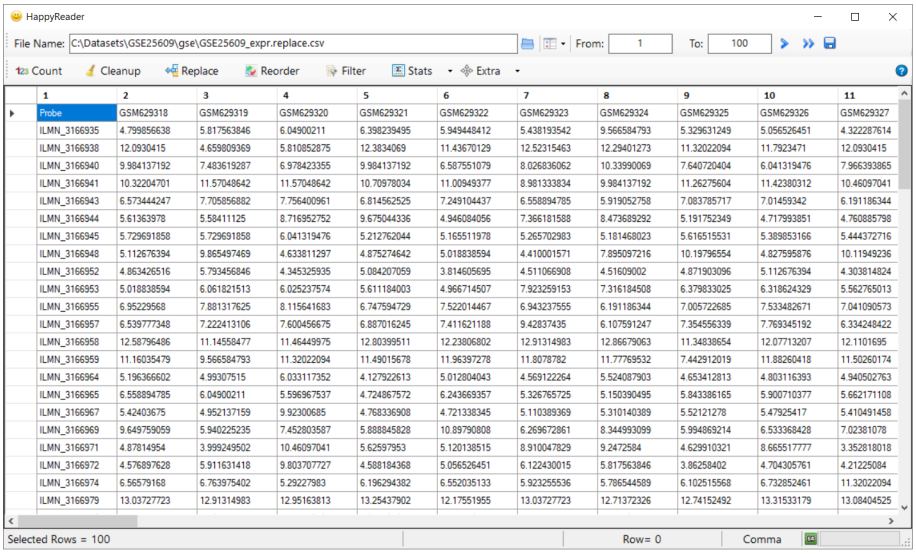
13. Click on Gene Ontology button.



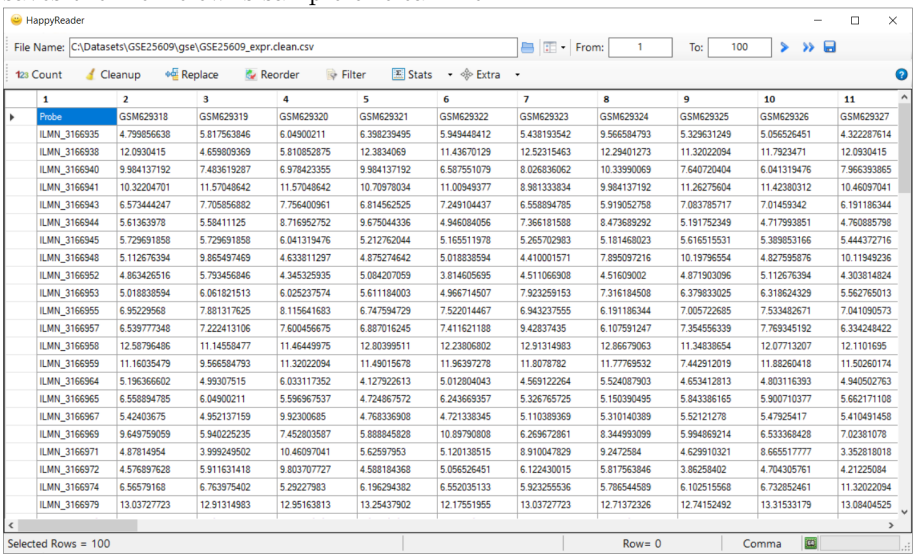
14. Click on fx for Log transformation. This transforms the expression file to log expression and saves the file.Below is sample of log expression file.



15. Click on to replace any missing value with the average value of corresponding row. This modifies the expression file and saves the file.Below is sample of replaced file.



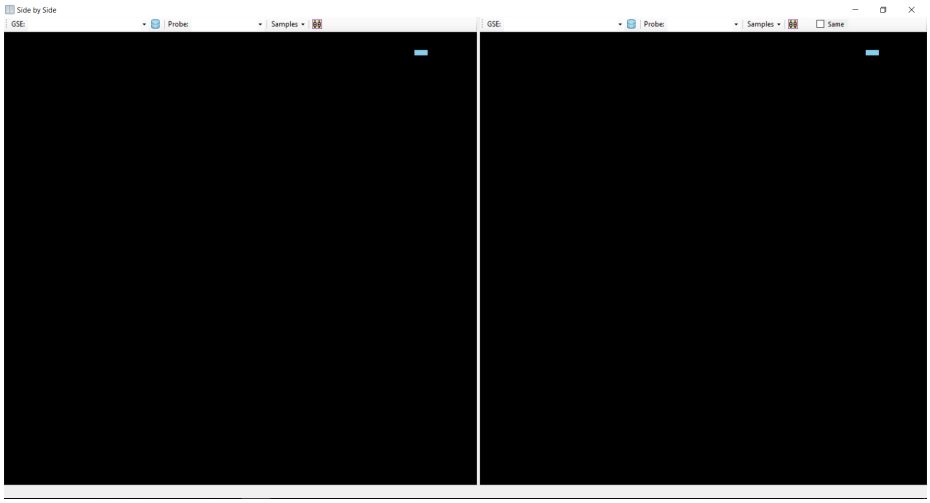
16. Click on to delete any empty row . This changes the expression file and saves the file.Below is sample of clean file.



3.3 Compare GSE

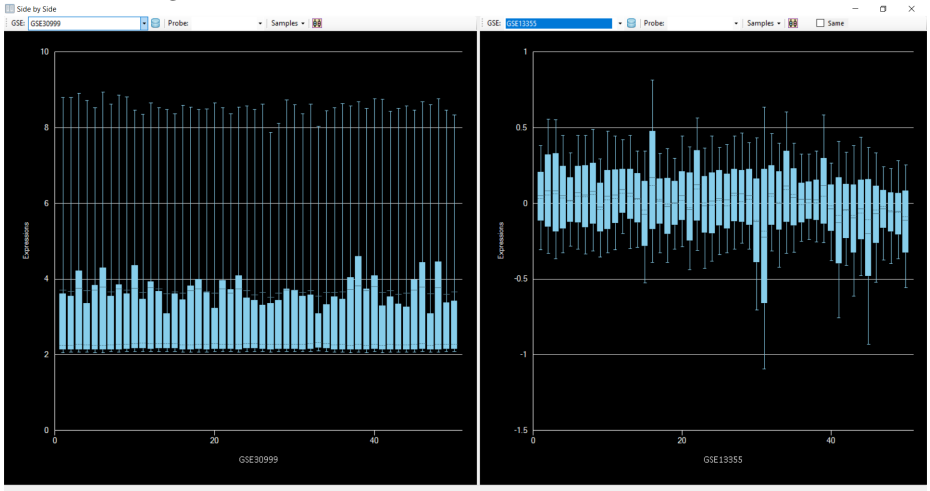
Compare two databases of same category.

1. Click on .A new compare window opens up.



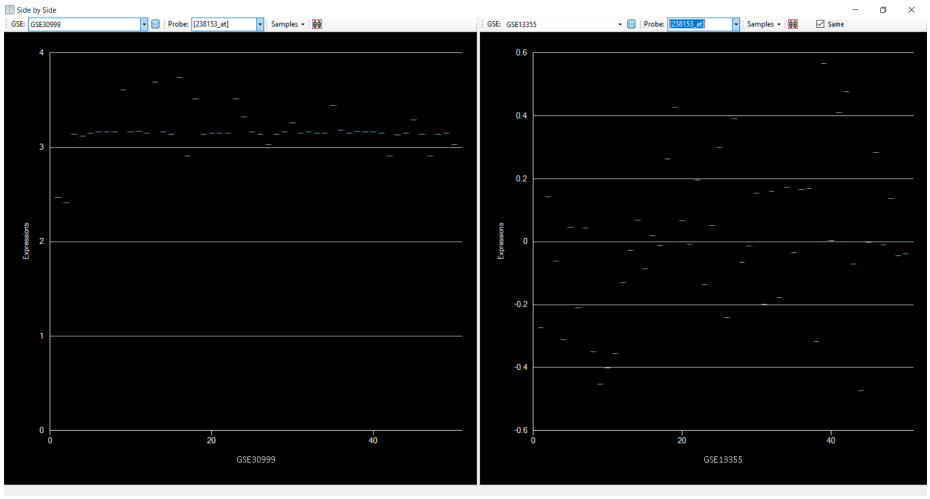
2. Select two databases that similar to same category for comparison.

3. Click on to get the distribution of data from both data-sets.

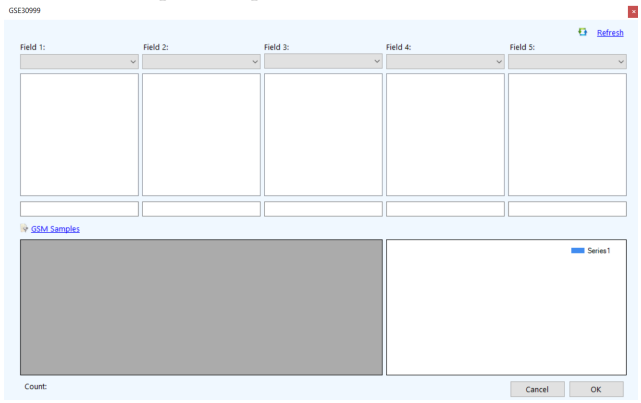


4. Click on Same checkbox to compare the same samples on either sides if available.

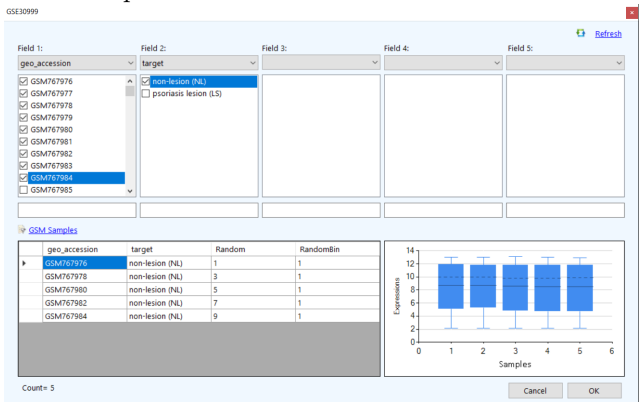
5. Select probes on each side to compare individual probes in the data-sets.



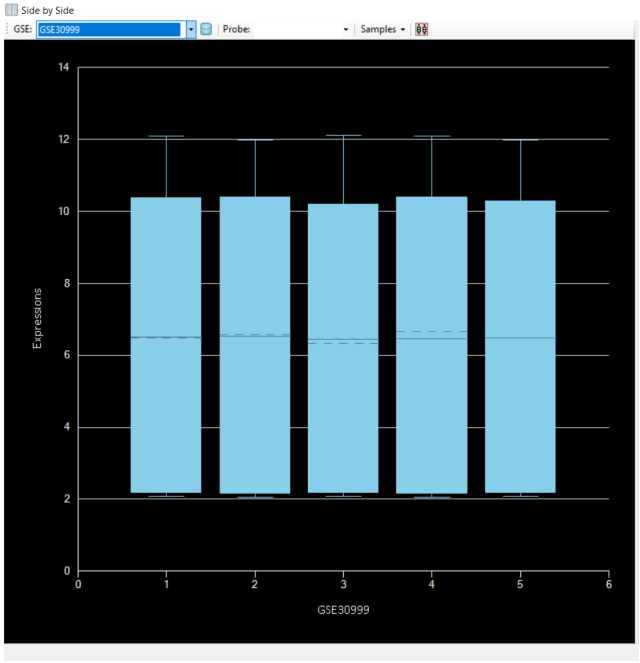
6. Click on Samples drop-down for filters.Click on Add Filter.



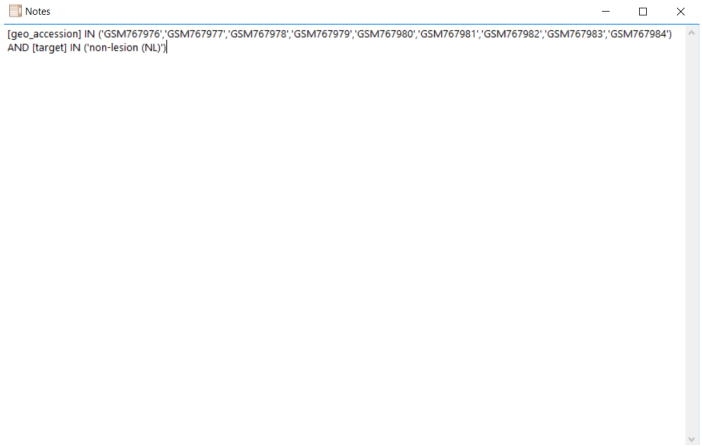
7. Select the filters to apply and the filtered data will be displayed under GSM Samples on click.



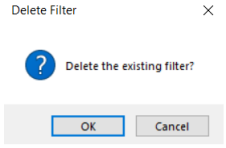
8. Click on Refresh link to clear all the filters or click on "OK" to apply the filter.



9. Click on View Filter to check the filters applied.



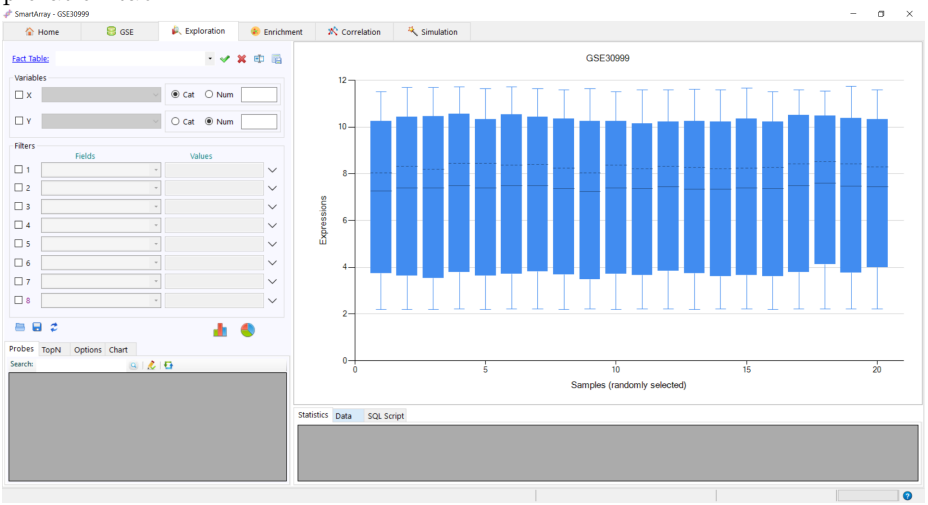
10. Click on Delete Filter to delete the existing filter.



4

Exploration

The exploration tab performs the statistical analysis on the database.Upon double clicking on the database in GSE tab, it redirects to the default page of Exploration tab.



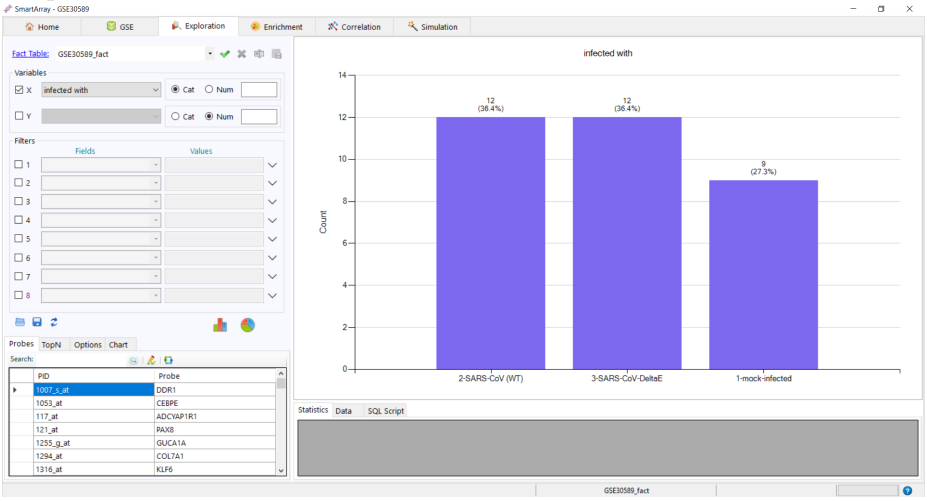
1. Click on the Fact Table hyperlink to load the fact tables from the database.Select the corresponding fact table and click on Click on ✔ to load the table. The original fact table cannot be deleted,renamed or exported.
2. Click on ✖ to delete the fact table.
3. Click on to rename the fact table.
4. Click on to export the fact table.

4.1 Univariate Analysis

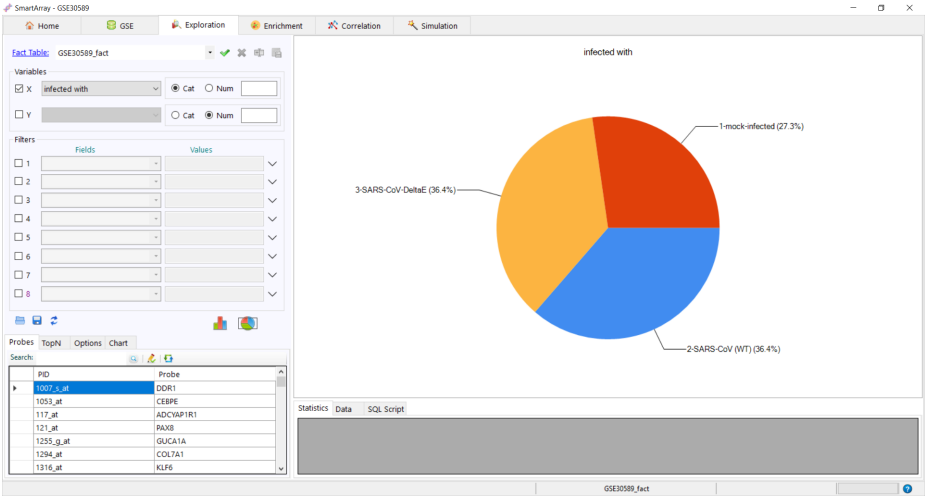
Univariate analysis on a numerical or categorical variant can be performed.

1. Select any variant as X variable to perform univariate analysis. Depending on the variant selected, the type is changed to Cat(Categorical) or Num(Numerical).

2. Click on  to get the distribution of the variant. The data and the SQL Script used can be seen at the bottom of the screen.



3. Click on  to get the pie chart distribution.

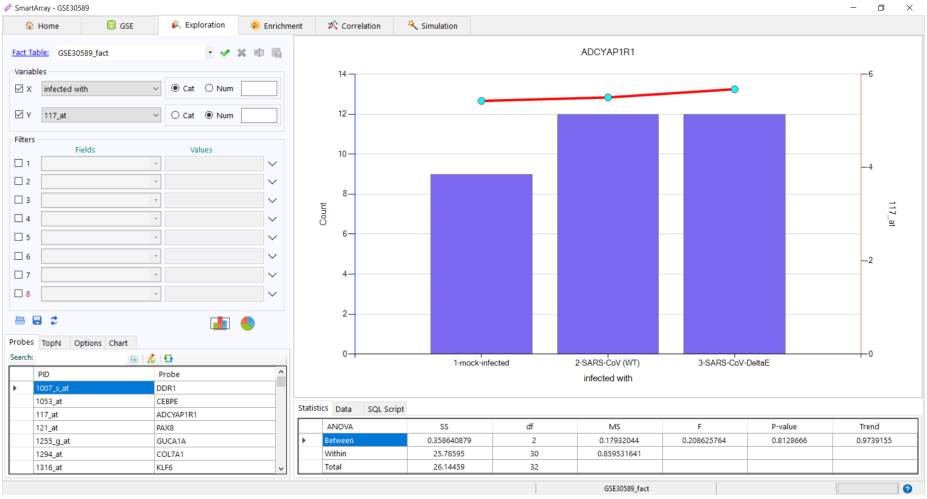


4.2 Bivariate Analysis

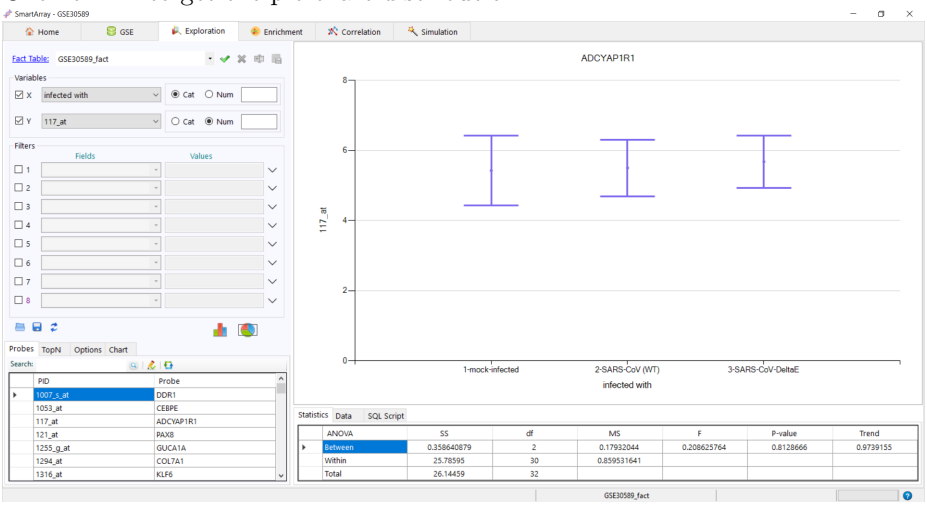
Bivariate analysis on a numerical or categorical variant can be performed.

1. Select any variant as Y variable to perform bivariate analysis.Depending on the variant selected, the type is changed to Cat(Categorical) or Num(Numerical).

2. Click on  to get the distribution of the variant. The statistics, data and the SQL Script used can be seen at the bottom of the screen.



3. Click on  to get the pie chart distribution.

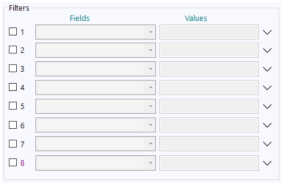


4. Keep changing the values in the Y variable to get the correlation of each value with X variant.

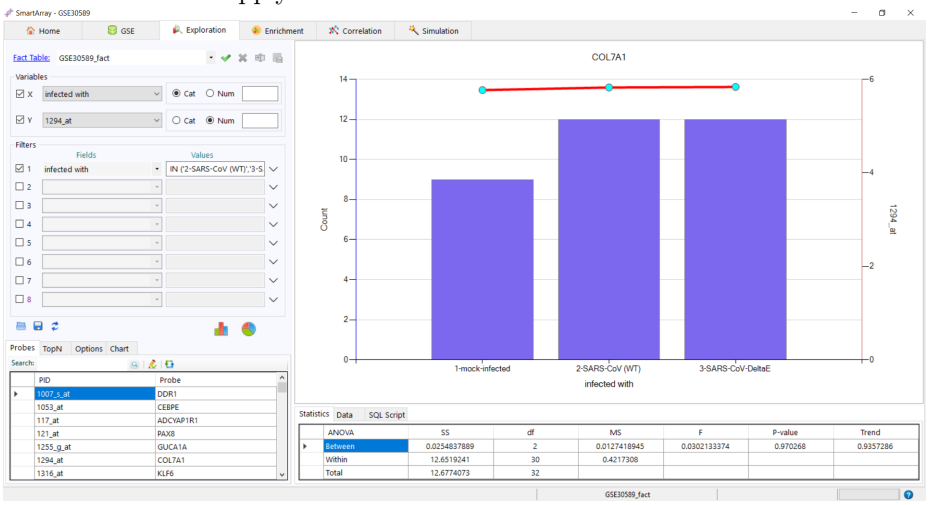
4.3 Visualization

The top 5,100 or 800 subsets can be sliced to form another dataset.

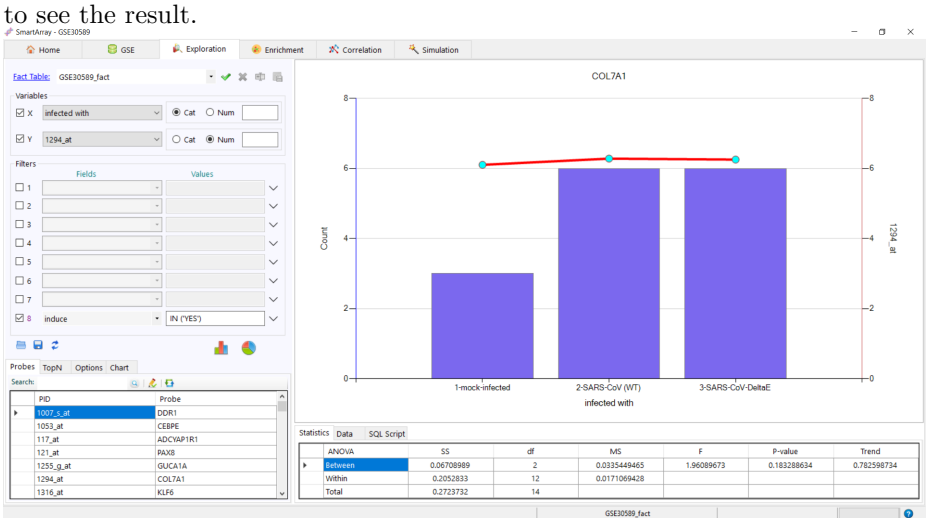
1. Filters: Apply filters for more clear view of the data.7 Filters (Filter 1- 8) can be applied for filter functionality.



- Select a variable to apply the filter.Click on  or  to see the result.

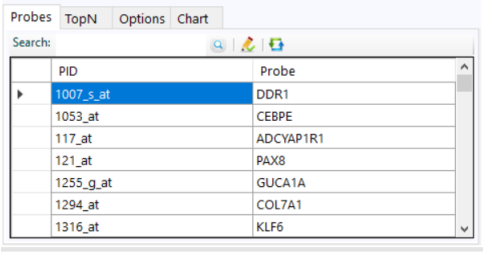


- Filter No. 8 adds another dimension to the data.Click on  or  to see the result.

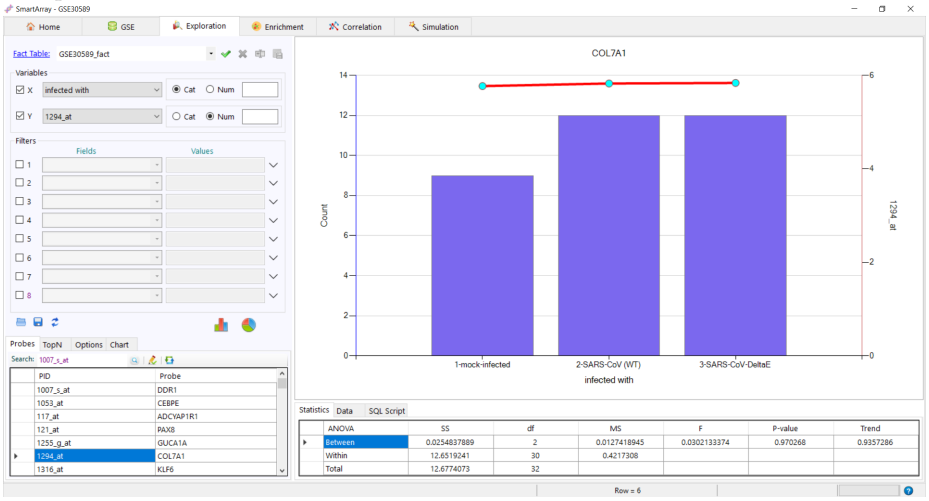


- Click on  to refresh the existing filters.
- Click on to save the data.
- Click on  to load the saved filters.

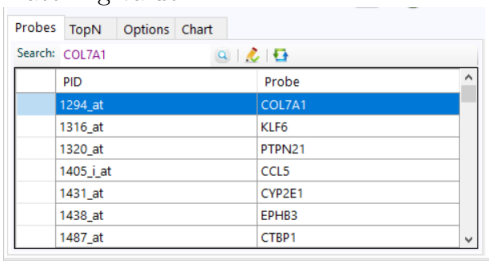
2. Probes: The correlation between each value of Y variable with any X variant can be seen in the Probes Tab.



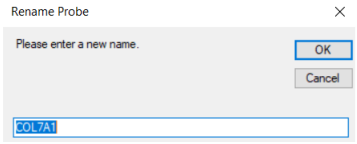
- Keep clicking on different values to see the distribution changing in box plot.



- Enter the probe ID/name to search and click on to search the next matching value.

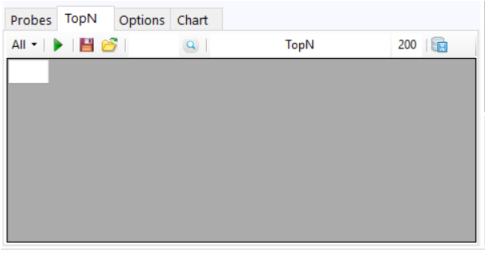


- Select the probe to be renamed and click on to rename the probe.



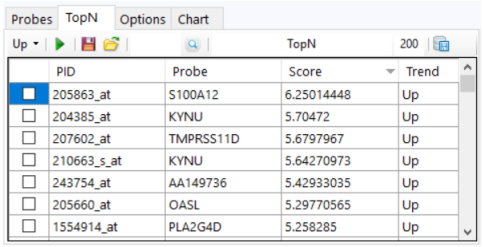
- Click on to refresh the data.

3. Top: Allows users to see the correlation between all the values in Y variable with respect to X variable.



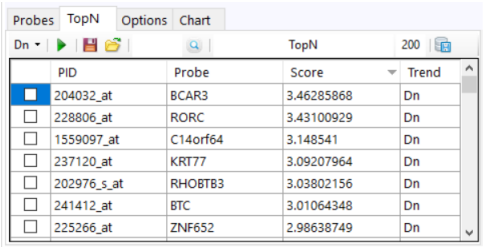
- Click on All dropdown to filter the data to be displayed.
  (a) All - Displays the complete data.

(b) Upregulated - Displays the data that has trend Up.



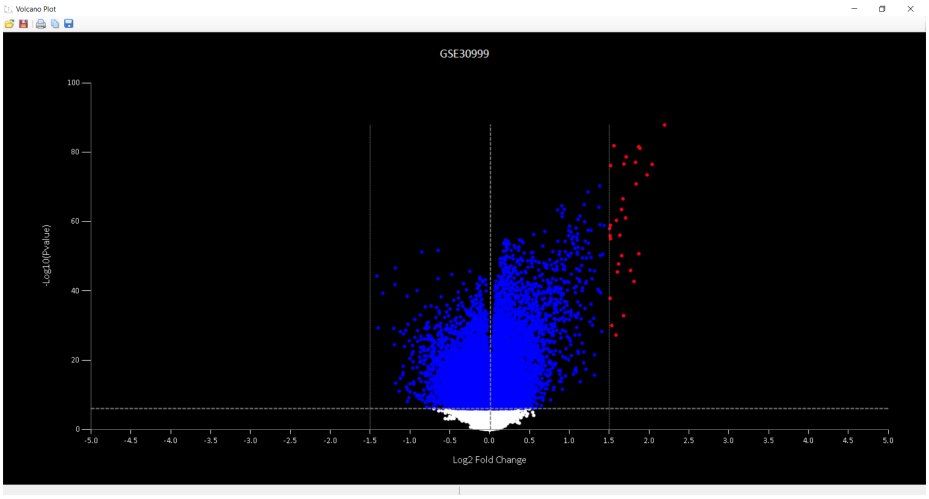
	PID	Probe	Score	Trend
☐	205863_at	S100A12	6.25014448	Up
☐	204385_at	KYNU	5.70472	Up
☐	207602_at	TMPRSS11D	5.6797967	Up
☐	210663_s_at	KYNU	5.64270973	Up
☐	243754_at	AA149736	5.42933035	Up
☐	205660_at	OASL	5.29770565	Up
☐	1554914_at	PLA2G4D	5.258285	Up

(c) Downregulated - Displays the data that has trend Down(Dn).

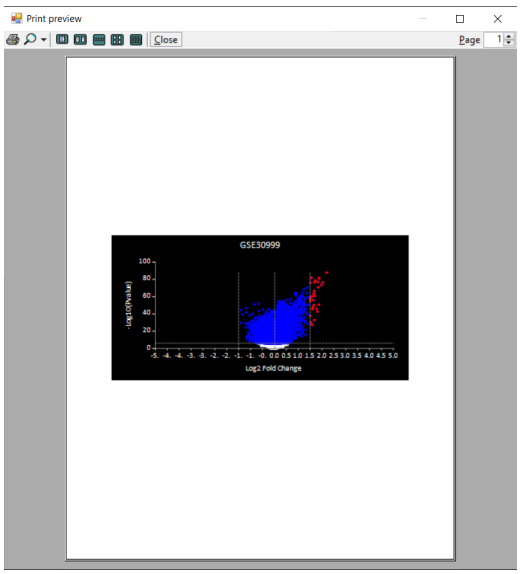


	PID	Probe	Score	Trend
☐	204032_at	BCAR3	3.46285868	Dn
☐	228806_at	RORC	3.43100929	Dn
☐	1559097_at	C14orf64	3.148541	Dn
☐	237120_at	KRT77	3.09207964	Dn
☐	202976_s_at	RHOBTB3	3.03802156	Dn
☐	241412_at	BTC	3.01064348	Dn
☐	225266_at	ZNF652	2.98638749	Dn

- Click on ▶ to get the Volcano/Manhattan plot.



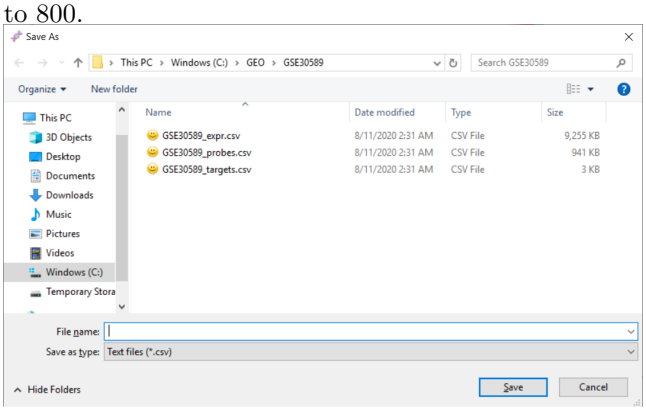
(a) Click on 📂 to load the data used for obtaining volcano plot.
(b) Click on 💾 to save the data used for obtaining volcano plot.
(c) Click on 🖨 to print the volcano plot.



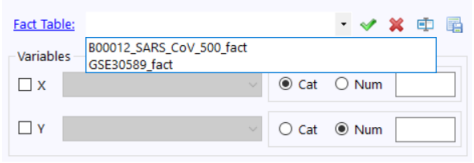
(d) Click on to copy the volcano plot.

(e) Click on to save the volcano plot.

- Click on to search for Probe ID or Probe name.
- Click on to save the top N data where N can be in the range of 5 to 800.



- Click on to load the existing subset dataset.
- Click on to save the Top N data to database.The textbox can be used to name the table and the number of rows(N) to be saved. The saved dataset appears on the FactTable list.

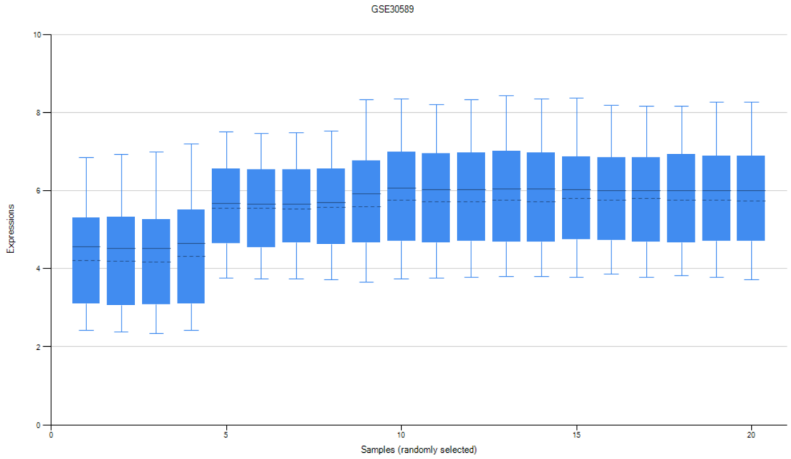


4. Options tab provides more features for SNPs.

- Exclude NC (No Call): Excludes NC genotype for the statistical analysis

- Encode SNP: Encodes the SNPs
- In-memory Aggregation: This feature allows users to use the in-memory of their local computers to run SmartArray application

5. Chart tab provides more features on visualization.
   - Font Size: Increasing or decreasing of font size changes the font displayed on the charts.
   - Click on Density Plot to get the density plot for the data-set.

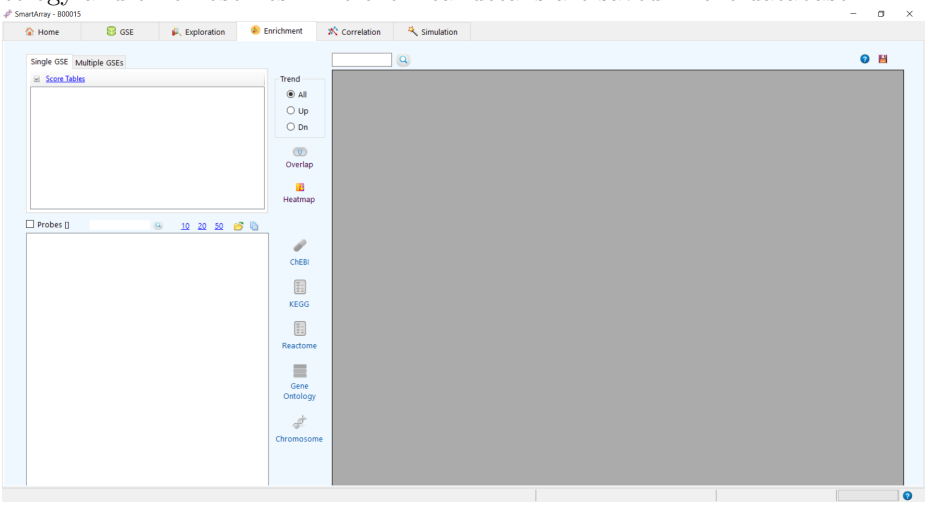


   - To copy the density plot, use and edit it in any image editor.
   - Click on to print the density plot.
   - Click on to save the density plot.
   - X-axis Show Range: Shows the range of values in X-axis.
   - Y_2 % Format: Shows the plot with y-axis values in terms of percentage.
   - Volcano Plot: Check on this item to show the volcano plot after running. Uncheck to not show the plot. Click on to show the saved plot
   - X-axis Sort by Count: Sorts the data according to the x-axis value count.
   - Minimum Y-Axis: Use minimum value of Y-axis instead of 0 as minimum value.
   - Log_2FC: It takes the value of $(Log_2(B) - Log_2(A))$ if checked and B - A if unchecked in X-axis of the volcano plot.
   - X-axis Reversed: Reverses the X-axis and displays the plot.
   - $-Log_{10}$(Pvalue): On check considers $-Log_{10}$(Pvalue) as Y-axis value else takes the score as Y-value in the volcano plot.

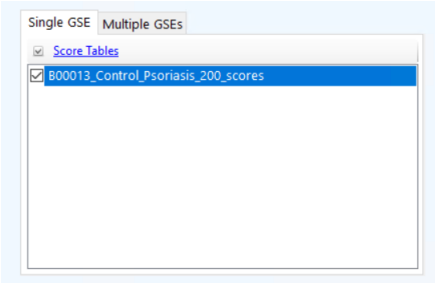
5

Enrichment

Enrichment displays the relation of probes with chemicals, pathways, gene ontology and chromosomes. The chemical details are saved in the database.



Click on Score Tables to load the subset data-sets.



1. Click on Overlap to see the probes.
2. Click on Heat-map to see the heat-map of all probes.

Probe	B00013_Control
CDH3	322.3567
VNN3	322.2463
DCUN1D5	321.5381
GSPT1	320.966
EPSTI1	319.9707
ELAC2	319.795
A2ML1	319.6004
ARMET	316.6617
GTF3C6	316.3443
RAB10	315.2776
RGS20	315.2238
RAI14	-316.2119
GLTSCR2	-321.4535
PTPN14	-323.9662
PLLP	-327.4801
CCL27	-337.463
INPP5A	-345.9203
PARD3	-349.1754
COBL	-355.4759

3. Trend Options allows to see the trend of probes.

   - All - Displays all the probes of subset.

Probe	Trend	B00013_Control
AI819863	Up	352.07009887...
N4BP1	Up	350.67520141...
GZMB	Up	350.61669921...
ARSF	Up	350.57559204...
UBE2C	Up	350.31600952...
SLC7A11	Up	349.76889038...
PARD3	Dn	349.17541503...
NFKBIZ	Up	348.44610595...
ATP12A	Up	346.82611083...
INPP5A	Dn	345.92028808...
FOXM1	Up	344.40719604...
PGM2	Up	344.28781127...
TRIM14	Up	343.84820556...
ABCC1	Up	343.31509399...
PYCARD	Up	343.07269287...
UBE2N	Up	342.91378784...
FBXO9	Up	342.68978881...
SPTSSA	Up	342.10281372...

   - Up - Displays probes that are up regulated.

Probe	Trend	B00013_Control
PI3	Up	1924.9759521...
S100A7A	Up	1596.9289550...
DEFB4A	Up	1554.3420410...
AKR1B10	Up	1270.1479492...
S100A12	Up	1151.4560546...
TCN1	Up	1060.0539550...
KYNU	Up	1053.6300048...
IL36G	Up	1046.6469726...
SPRR2C	Up	862.453125
SERPINB4	Up	834.02960205...
C10orf99	Up	805.28082275...
SERPINB3	Up	801.14093017...
SPRR2D	Up	780.85211181...
ADAMDEC1	Up	750.51721191...
HPSE	Up	732.58959960...
ZC3H12A	Up	674.53540039...
PDZK1IP1	Up	673.92468261...
S100A9	Up	628.39172363...

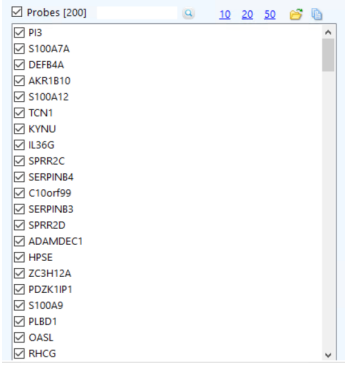
   - Down - Displays probes that are down regulated.

Probe	Trend	B00013_Control
BTC	Dn	489.20410156...
CLDN1	Dn	450.40460205...
SAPCD1	Dn	442.81350708...
C5orf46	Dn	436.66619873...
KRT77	Dn	414.00021362...
BG109230	Dn	406.63888549...
ID4	Dn	401.94720458...
IL37	Dn	398.33929443...
RPL15	Dn	392.80709838...
GATA3	Dn	387.85140991...
PTPN21	Dn	370.34640502...
CHP2	Dn	356.32058715...
COBL	Dn	355.47589111...
PARD3	Dn	349.17541503...
INPP5A	Dn	345.92028808...
CCL27	Dn	337.46301269...
PLLP	Dn	327.48010253...
PTPN14	Dn	323.96618652...
GLTSCR2	Dn	321.45349121...
RAI14	Dn	316.21191406...

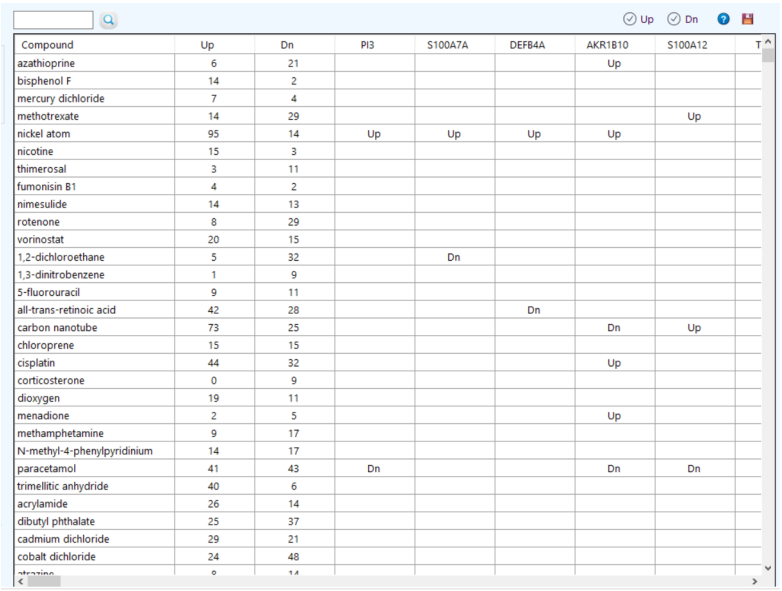
4. Text box can be used to search for probes.

5.1 Chemical Interaction

Multiple chemical interactions can be seen in Enrichment tab. Click on numbers or the checkbox to select the probes.



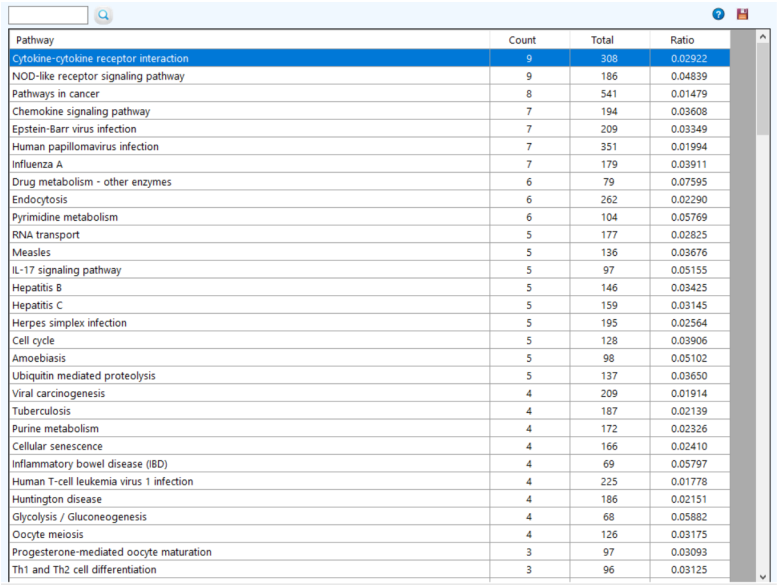
1. Click on ChEBI to see the Gene-Compound interactions in probes.



Compound	Up	Dn	PI3	S100A7A	DEFB4A	AKR1B10	S100A12
azathioprine	6	21				Up	
bisphenol F	14	2					
mercury dichloride	7	4					
methotrexate	14	29					Up
nickel atom	95	14	Up	Up	Up	Up	
nicotine	15	3					
thimerosal	3	11					
fumonisin B1	4	2					
nimesulide	14	13					
rotenone	8	29					
vorinostat	20	15					
1,2-dichloroethane	5	32		Dn			
1,3-dinitrobenzene	1	9					
5-fluorouracil	9	11					
all-trans-retinoic acid	42	28			Dn		
carbon nanotube	73	25				Dn	Up
chloroprene	15	15					
cisplatin	44	32				Up	
corticosterone	0	9					
dioxygen	19	11					
menadione	2	5				Up	
methamphetamine	9	17					
N-methyl-4-phenylpyridinium	14	17					
paracetamol	41	43	Dn			Dn	Dn
trimellitic anhydride	40	6					
acrylamide	26	14					
dibutyl phthalate	25	37					
cadmium dichloride	29	21					
cobalt dichloride	24	48					

- Click on Up to see the up regulated probes.
- Click on Dn to see the down regulated probes.
- Click on to see the ChEBI qualifiers.
- Click on to save the results.

2. Click on to see the KEGG pathway enrichment.



Pathway	Count	Total	Ratio
Cytokine-cytokine receptor interaction	9	308	0.02922
NOD-like receptor signaling pathway	9	186	0.04839
Pathways in cancer	8	541	0.01479
Chemokine signaling pathway	7	194	0.03608
Epstein-Barr virus infection	7	209	0.03349
Human papillomavirus infection	7	351	0.01994
Influenza A	7	179	0.03911
Drug metabolism - other enzymes	6	79	0.07595
Endocytosis	6	262	0.02290
Pyrimidine metabolism	6	104	0.05769
RNA transport	5	177	0.02825
Measles	5	136	0.03676
IL-17 signaling pathway	5	97	0.05155
Hepatitis B	5	146	0.03425
Hepatitis C	5	159	0.03145
Herpes simplex infection	5	195	0.02564
Cell cycle	5	128	0.03906
Amoebiasis	5	98	0.05102
Ubiquitin mediated proteolysis	5	137	0.03650
Viral carcinogenesis	4	209	0.01914
Tuberculosis	4	187	0.02139
Purine metabolism	4	172	0.02326
Cellular senescence	4	166	0.02410
Inflammatory bowel disease (IBD)	4	69	0.05797
Human T-cell leukemia virus 1 infection	4	225	0.01778
Huntington disease	4	186	0.02151
Glycolysis / Gluconeogenesis	4	68	0.05882
Oocyte meiosis	4	126	0.03175
Progesterone-mediated oocyte maturation	3	97	0.03093
Th1 and Th2 cell differentiation	3	96	0.03125

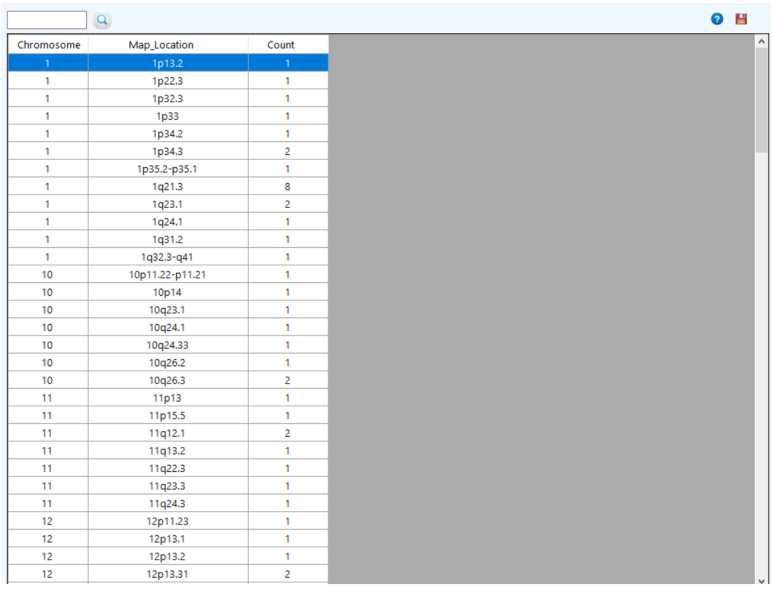
3. Click on to see the reactome pathway enrichment.

Pathway	Count	Total	Ratio
Immune System	62	2130	0.02911
Innate Immune System	36	1057	0.03406
Metabolism	34	2123	0.01602
Signal Transduction	33	2758	0.01197
Metabolism of proteins	29	2039	0.01422
Cytokine Signaling in Immune system	27	859	0.03143
Post-translational protein modification	21	1438	0.01460
Neutrophil degranulation	20	482	0.04149
Disease	20	1406	0.01422
Developmental Biology	19	1113	0.01707
Gene expression (Transcription)	19	1510	0.01258
Cell Cycle	17	682	0.02493
Signaling by Interleukins	17	462	0.03680
RNA Polymerase II Transcription	16	1372	0.01166
Cell Cycle, Mitotic	16	570	0.02807
Generic Transcription Pathway	15	1249	0.01201
Transport of small molecules	14	729	0.01920
Infectious disease	13	741	0.01754
Adaptive Immune System	13	757	0.01717
Signaling by GPCR	12	1150	0.01043
Interferon Signaling	11	199	0.05528
GPCR downstream signalling	11	1078	0.01020
Interleukin-1 family signaling	10	140	0.07143
Formation of the cornified envelope	10	130	0.07692
Cellular responses to external stimuli	10	608	0.01645
Cellular responses to stress	10	594	0.01684
Class I MHC mediated antigen processing & presentation	10	373	0.02681
Membrane Trafficking	10	634	0.01577
Metabolism of lipids	10	743	0.01346
Keratinization	10	214	0.04673

4. Click on  to see the gene ontology pathway enrichment.

Go_term	GO_type	Count	Total	R
protein binding	MF	117	8551	0
cytosol	CC	81	4874	0
cytoplasm	CC	59	4268	0
nucleus	CC	57	4556	0
nucleoplasm	CC	54	3605	0
plasma membrane	CC	49	4357	0
extracellular exosome	CC	47	2081	0
extracellular region	CC	36	1596	0
extracellular space	CC	28	1424	0
membrane	CC	28	2042	0
identical protein binding	MF	23	1296	0
integral component of membrane	CC	23	3082	0
cell	CC	21	1088	0
neutrophil degranulation	BP	20	475	0
inflammatory response	BP	19	377	0
ATP binding	MF	19	1445	0
innate immune response	BP	18	417	0
mitochondrion	CC	18	1217	0
protein-containing complex	CC	17	635	0
RNA binding	MF	16	1301	0
metal ion binding	MF	16	2225	0
signal transduction	BP	15	956	0
positive regulation of transcription by RNA polymerase II	BP	14	1124	0
integral component of plasma membrane	CC	14	1278	0
cell surface	CC	13	605	0
cytoskeleton	CC	13	384	0
apoptotic process	BP	12	543	0
immune response	BP	12	308	0
nucleolus	CC	12	815	0
positive regulation of cell proliferation	BP	12	400	0

5. Click on  to see the chromosomes frequency.



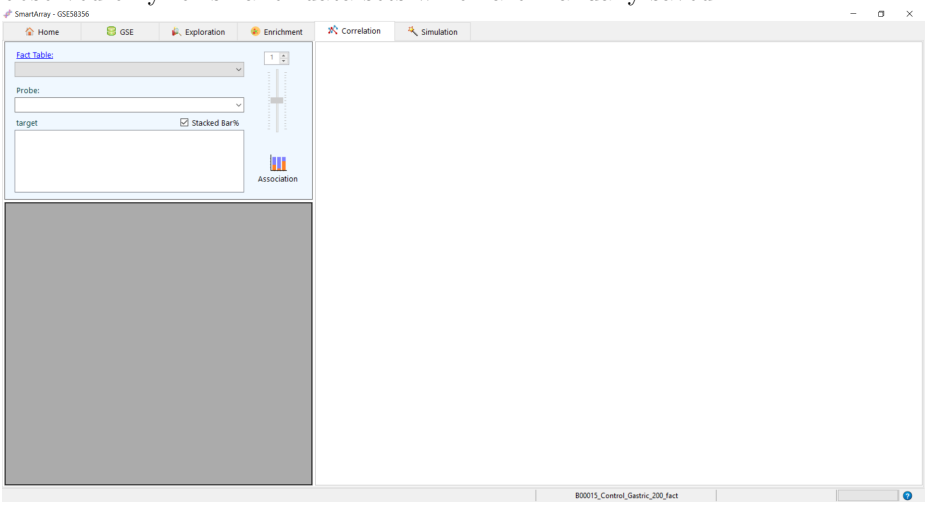
Chromosome	Map_Location	Count
1	1p13.2	1
1	1p22.3	1
1	1p32.3	1
1	1p33	1
1	1p34.2	1
1	1p34.3	2
1	1p35.2-p35.1	1
1	1q21.3	8
1	1q23.1	2
1	1q24.1	1
1	1q31.2	1
1	1q32.3-q41	1
10	10p11.22-p11.21	1
10	10p14	1
10	10q23.1	1
10	10q24.1	1
10	10q24.33	1
10	10q26.2	1
10	10q26.3	2
11	11p13	1
11	11p15.5	1
11	11q12.1	2
11	11q13.2	1
11	11q22.3	1
11	11q23.3	1
11	11q24.3	1
12	12p11.23	1
12	12p13.1	1
12	12p13.2	1
12	12p13.31	2

6. Click on Map hyperlink to see the chromosome map.

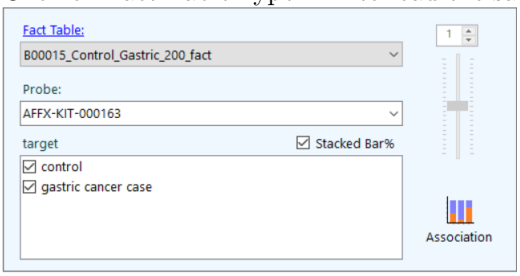
6

Correlation

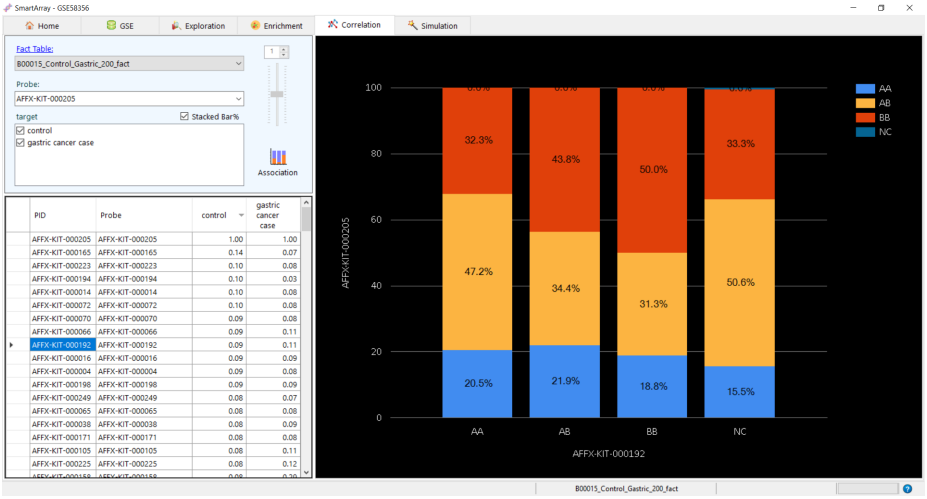
This tab gives the correlation between probes in a data-set.Correlation can be observed only for smaller data-sets which are manually saved.



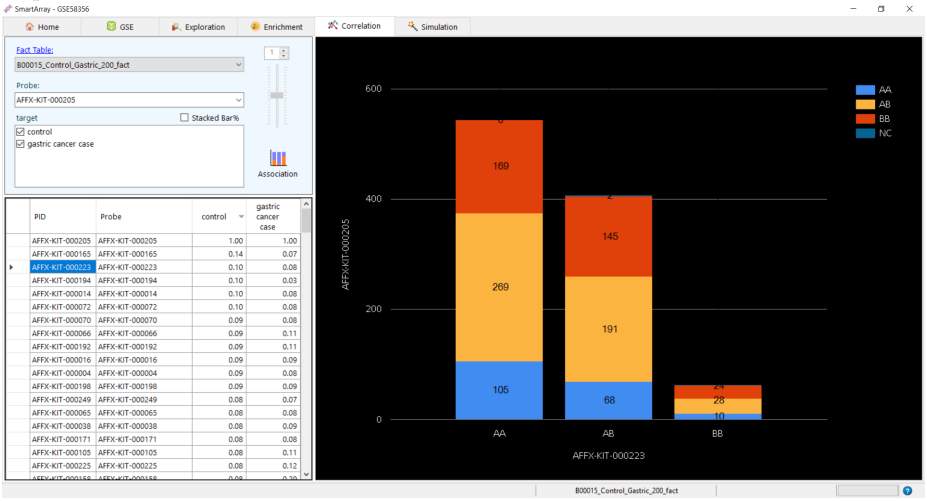
1. Click on Fact Table hyperlink to load the saved data-sets.



2. Select a probe/SNP and click on - Correlation in case of probes or Association in case of SNPs.



3. Unchecking of Stacked Bar% gives the plot with the count of SNPs instead of percent.

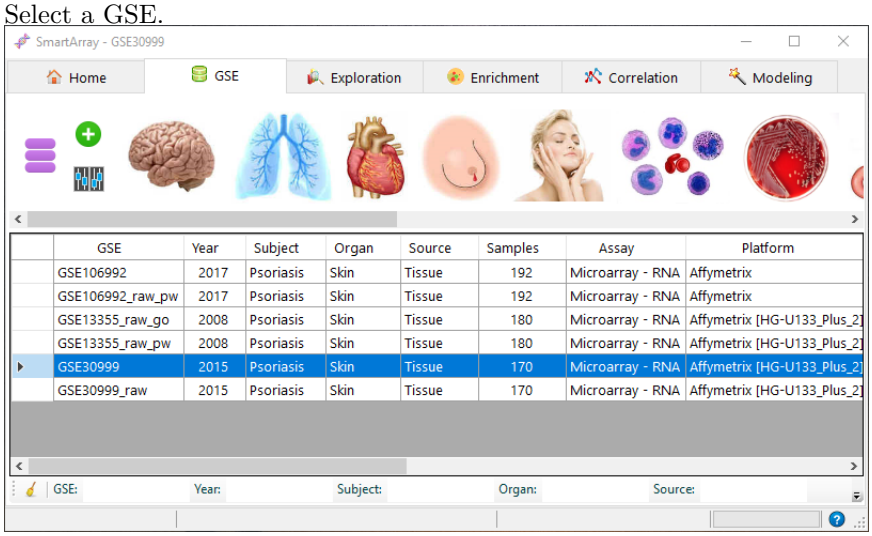


4. The zoom option is enabled for probes and can be used to change the offset of Y-axis.

7

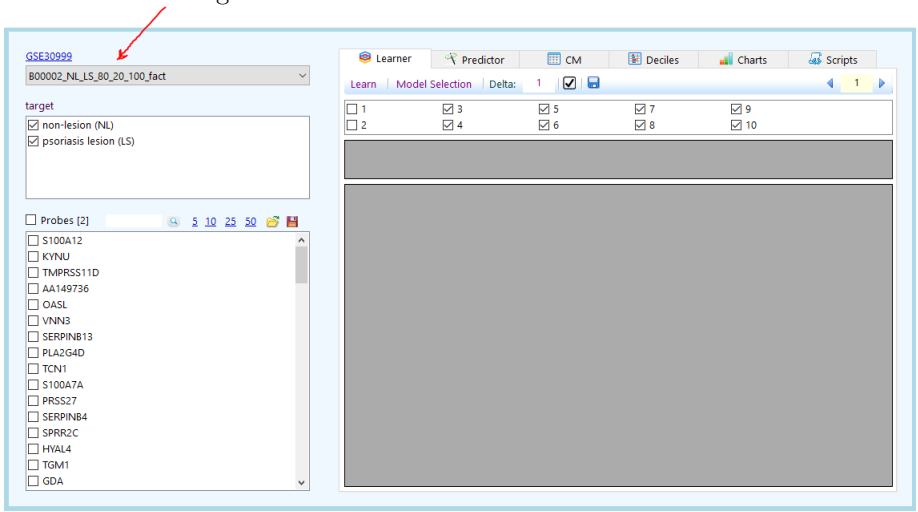
Modelling

1. Select a GSE.

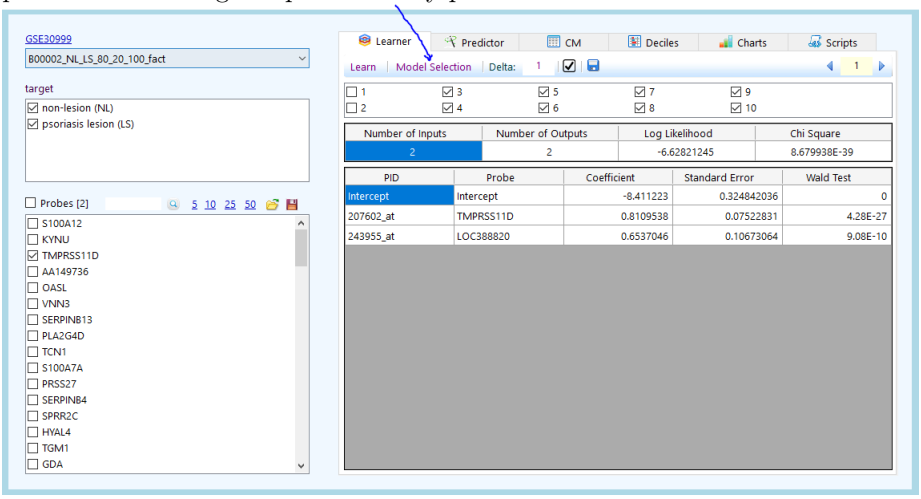


GSE	Year	Subject	Organ	Source	Samples	Assay	Platform
GSE106992	2017	Psoriasis	Skin	Tissue	192	Microarray - RNA	Affymetrix
GSE106992_raw_pw	2017	Psoriasis	Skin	Tissue	192	Microarray - RNA	Affymetrix
GSE13355_raw_go	2008	Psoriasis	Skin	Tissue	180	Microarray - RNA	Affymetrix [HG-U133_Plus_2]
GSE13355_raw_pw	2008	Psoriasis	Skin	Tissue	180	Microarray - RNA	Affymetrix [HG-U133_Plus_2]
GSE30999	2015	Psoriasis	Skin	Tissue	170	Microarray - RNA	Affymetrix [HG-U133_Plus_2]
GSE30999_raw	2015	Psoriasis	Skin	Tissue	170	Microarray - RNA	Affymetrix [HG-U133_Plus_2]

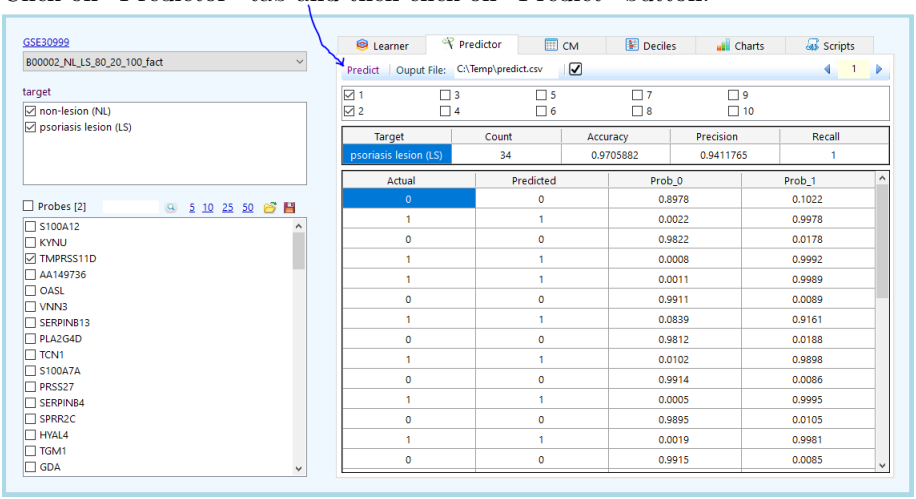
2. Click on “Modeling” tab and select a subset.



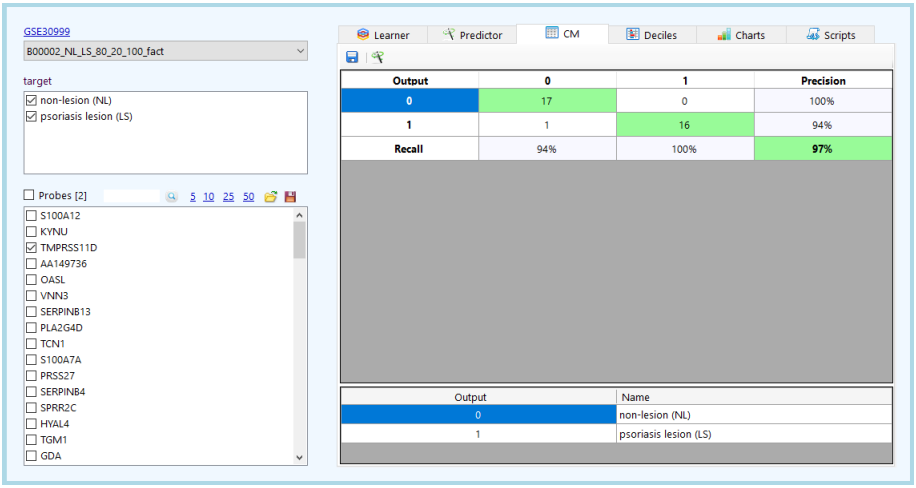
3. Click on “Model Selection” button. SmarrtArray automatically chooses probes with the highest predictability power.



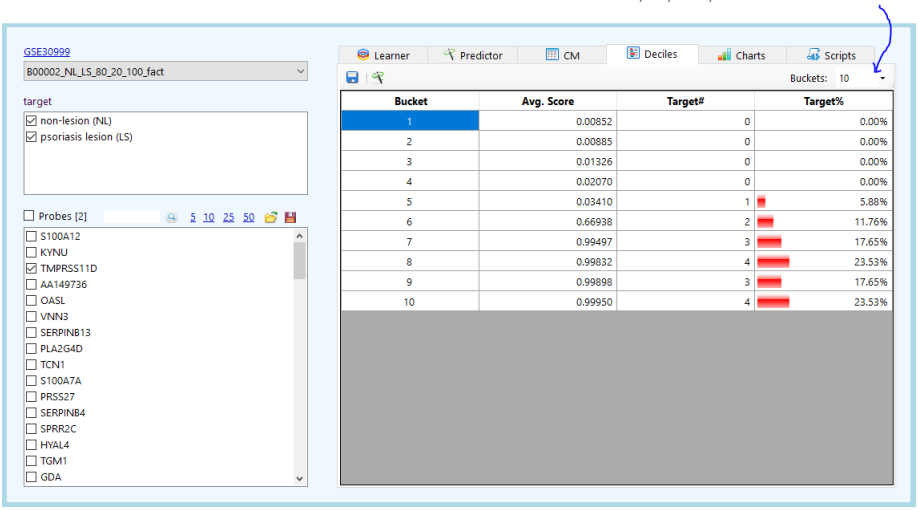
4. Click on “Predictor” tab and then click on “Predict” button.



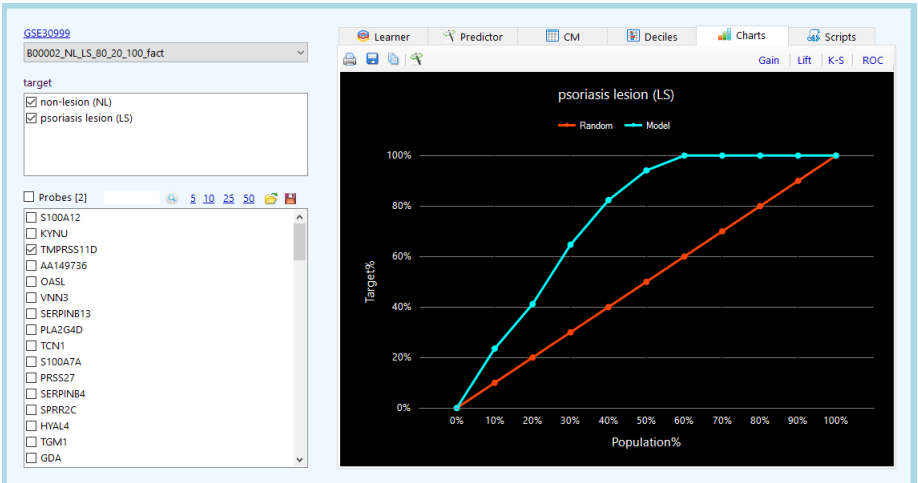
5. Click on “CM” tab (Confusion Matrix).



6. Click on "Deciles" tab. The number of buckets can be 2, 5, 10, 20 and 100.



7. Click on "Charts" tab to view Lift, Gain, K-S and ROC charts.



8. Click on “Scripts” tab to view the predictive model scripts in SQL, Python, Java, C# and VB.