Importing Affymetrix CEL Files with TAC for BRB-ArrayTools

Version 1.0

By

Dr. Qian Xie

BRB-ArrayTools Development Team

arraytools@emmes.com

Jan. 2021

Contents

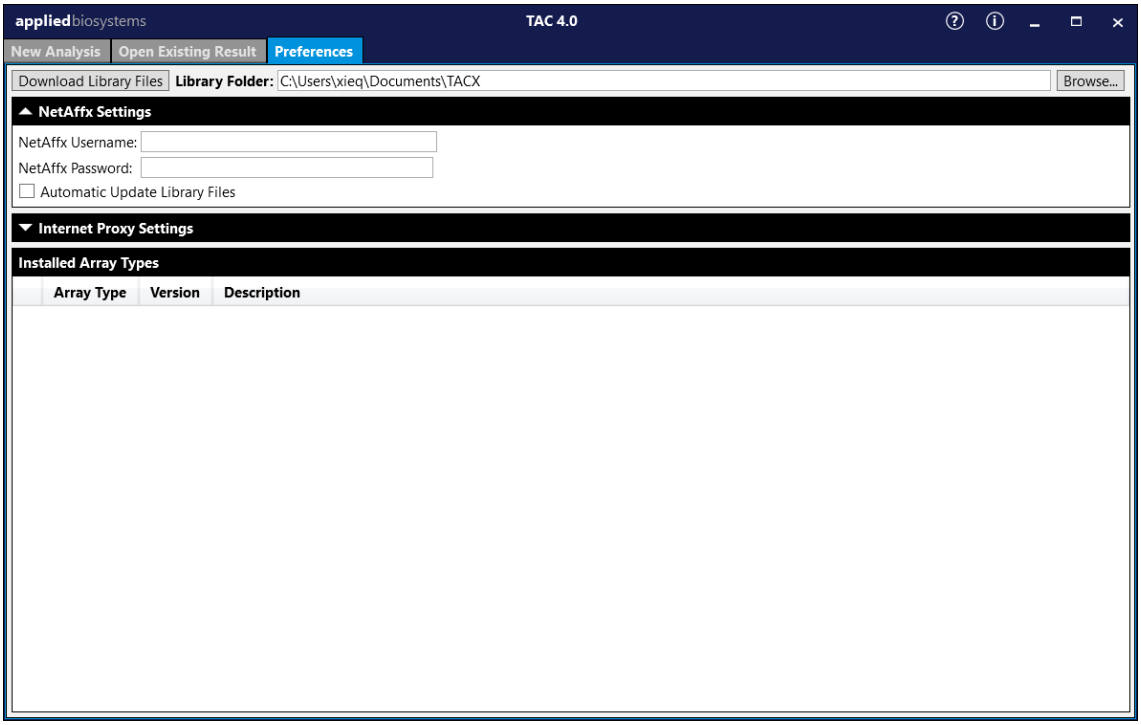
1. Introduction

Affymetrix Expression Console Software is now part of the Transcriptome Analysis Console (TAC) Software. In this instruction, you use the TAC software to convert Affymetrix CEL files to a probe summarization TXT file and then import the TXT file in BRB-ArrayTools to collate a BRB-ArrayTools project.

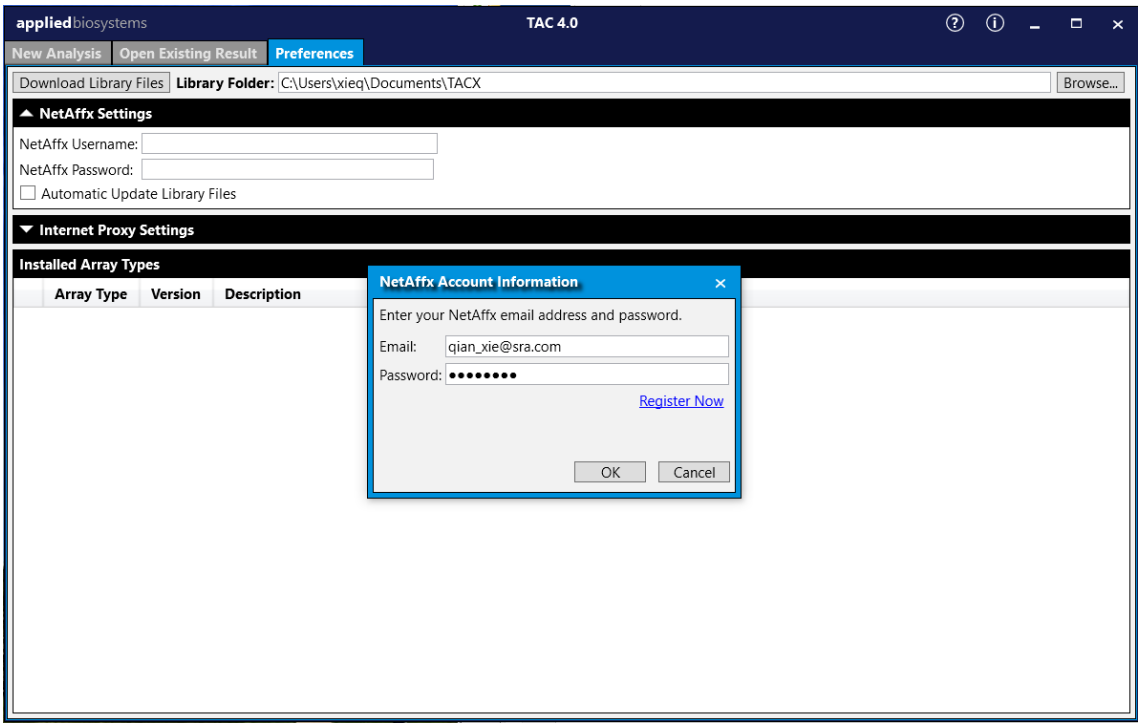
2. Download and install TAC Software

You can download the TAC software from the website https://www.thermofisher.com/us/en/home/life-science/microarray-analysis/microarray-analysis-instruments-software-services/microarray-analysis-software/affymetrix-transcriptome-analysis-console-software.html. Current version is TAC 4.0.2.15 (Nov. 2020). After installing the TAC software, you need to set up the software by clicking the tab "Preferences":

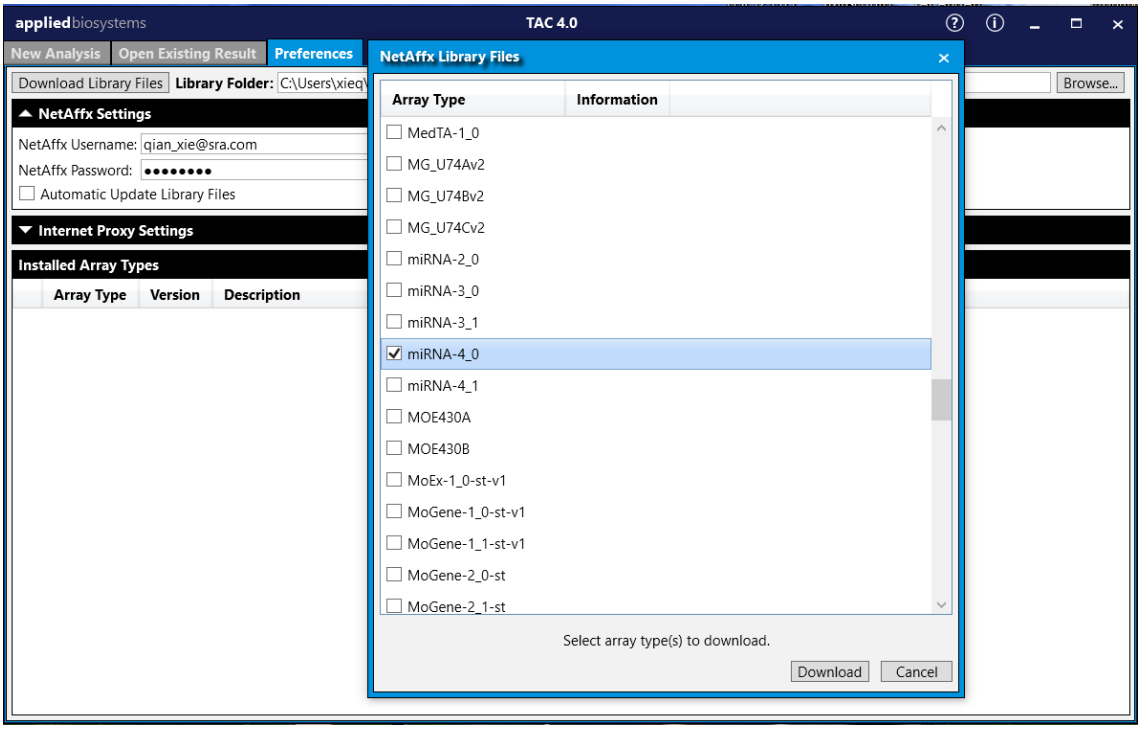
- Set Library Folders: C:/Users/%username%/Documents/TACX



- Download the Library files for your array type:
  - Click the "Download Library Files" button, a NetAffx Account Information window pops up. If you have an account, you may enter your account information here and click "OK". If you do not have an account, you need to click "Register Now" to create an account in your default browser and then enter here.



- The NetAffx Account Information will automatically fill in the NetAffx Settings field and a window pops up for selecting array types. Select miRNA-4_0 and click the “Download” button. The library files for the array type will be downloaded and the relevant information will keep under the “Installed Array Types” field.

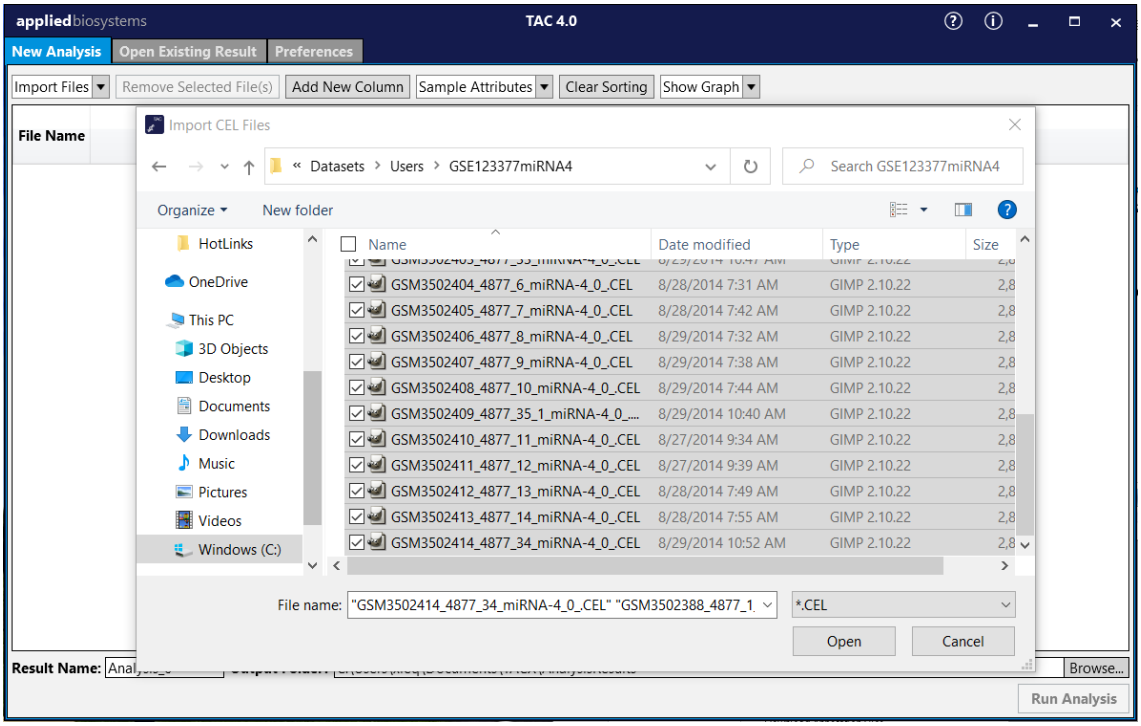


3. Convert CEL files to TXT file with TAC

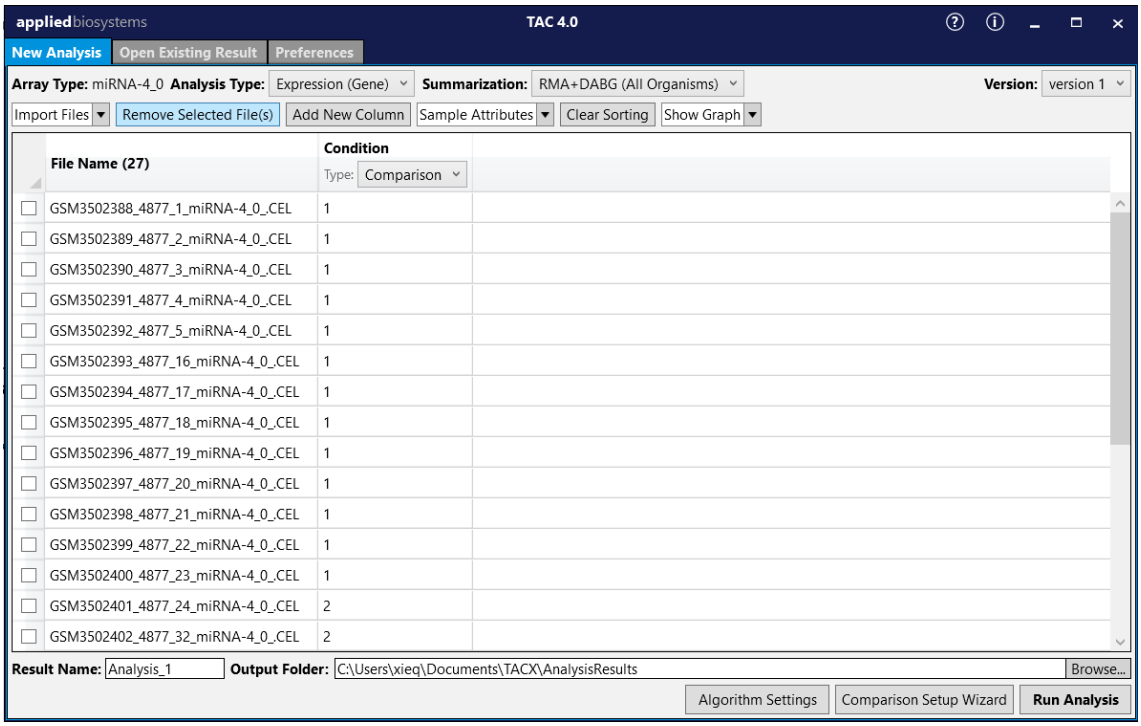
You keep all CEL files in a folder. Here uses the NCBI GEO GSE123377 dataset as an example. You may download the dataset https://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE123377&format=file.

You click the tab “New Analysis” to start a new study.

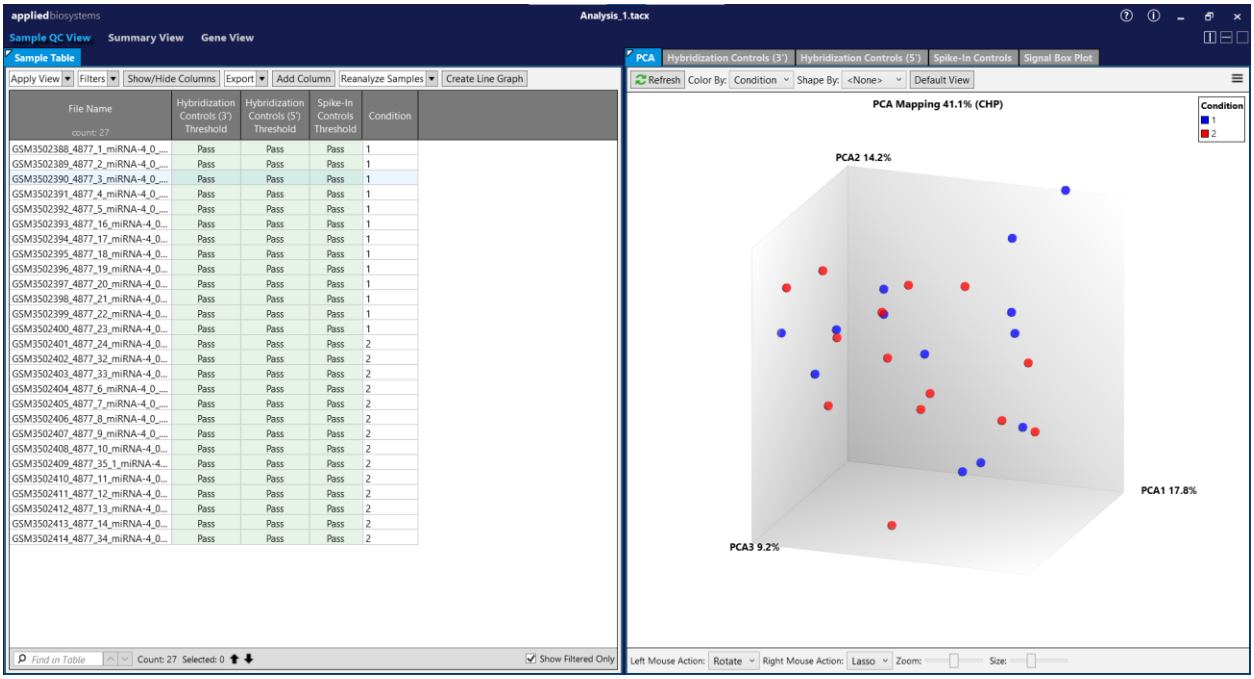
- Import Files > Import CEL Files: a file browser window pops up > Select CEL files (NCBI GEO GSE123377) > Open.



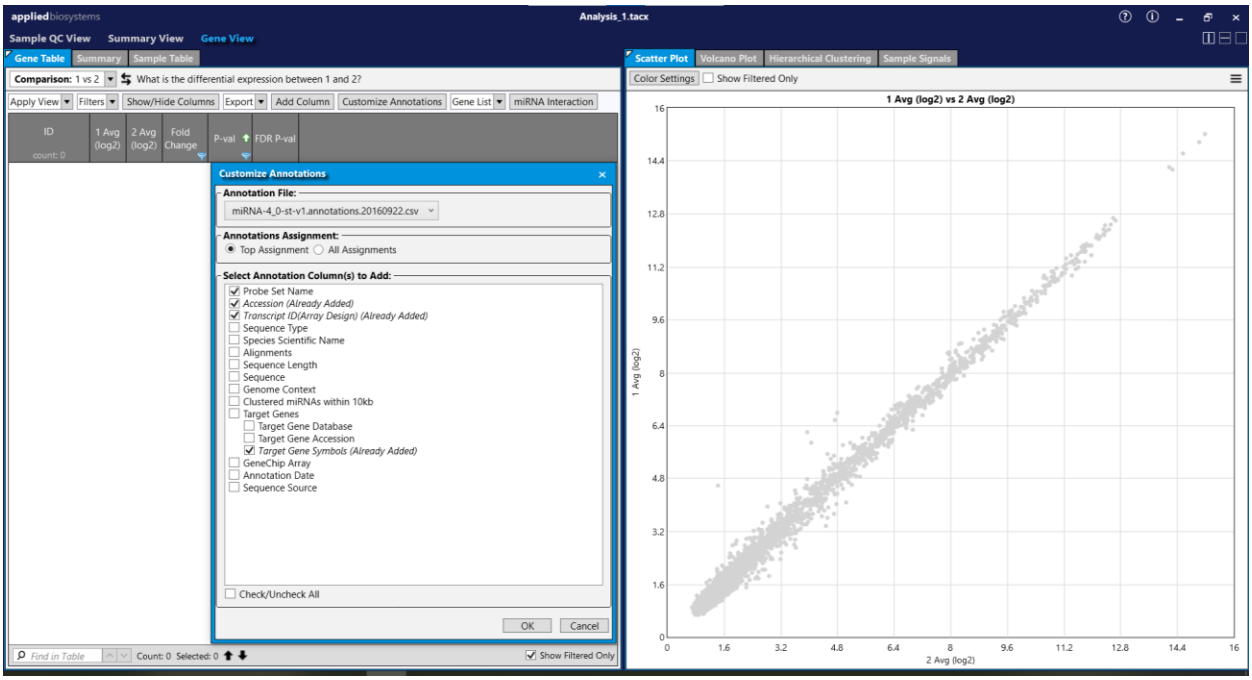
- In the column “Condition”, change Type: Comparison and assign the column values. Half of the files are “1” and another half of the files are “2”. These assignments are random. It is for the program to process. Click the lower-left button “Run Analysis”.



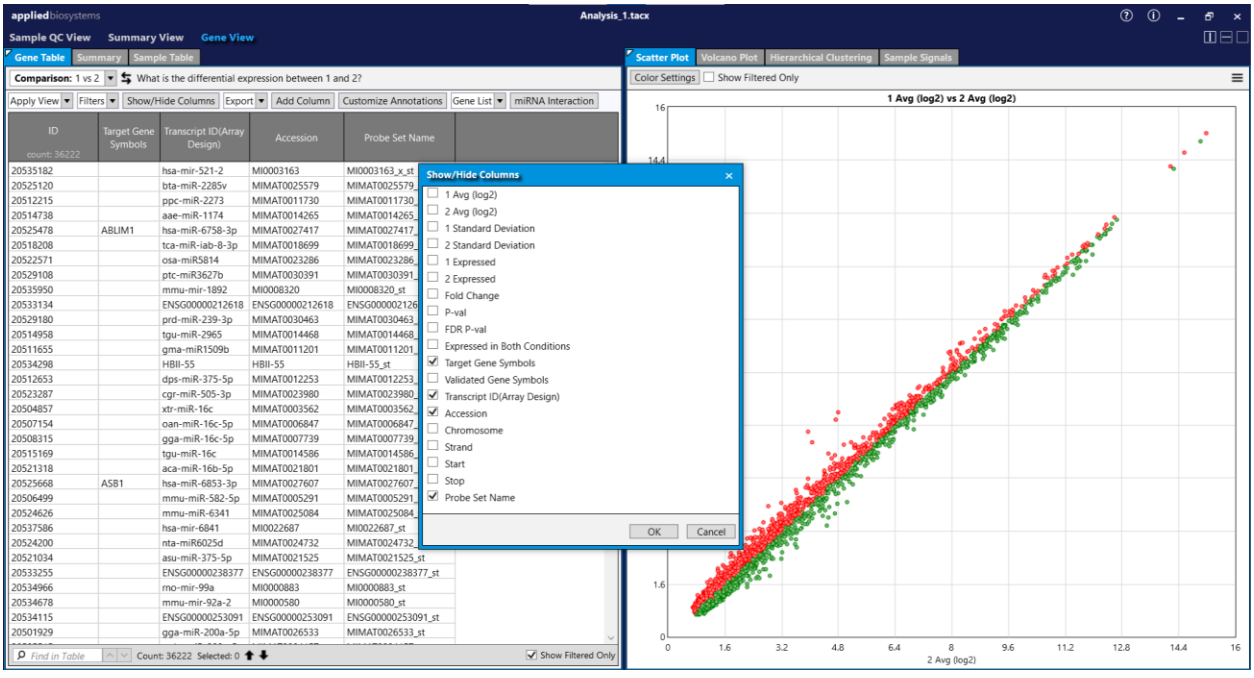
- When the analysis finishes, a new window pops up with the results.



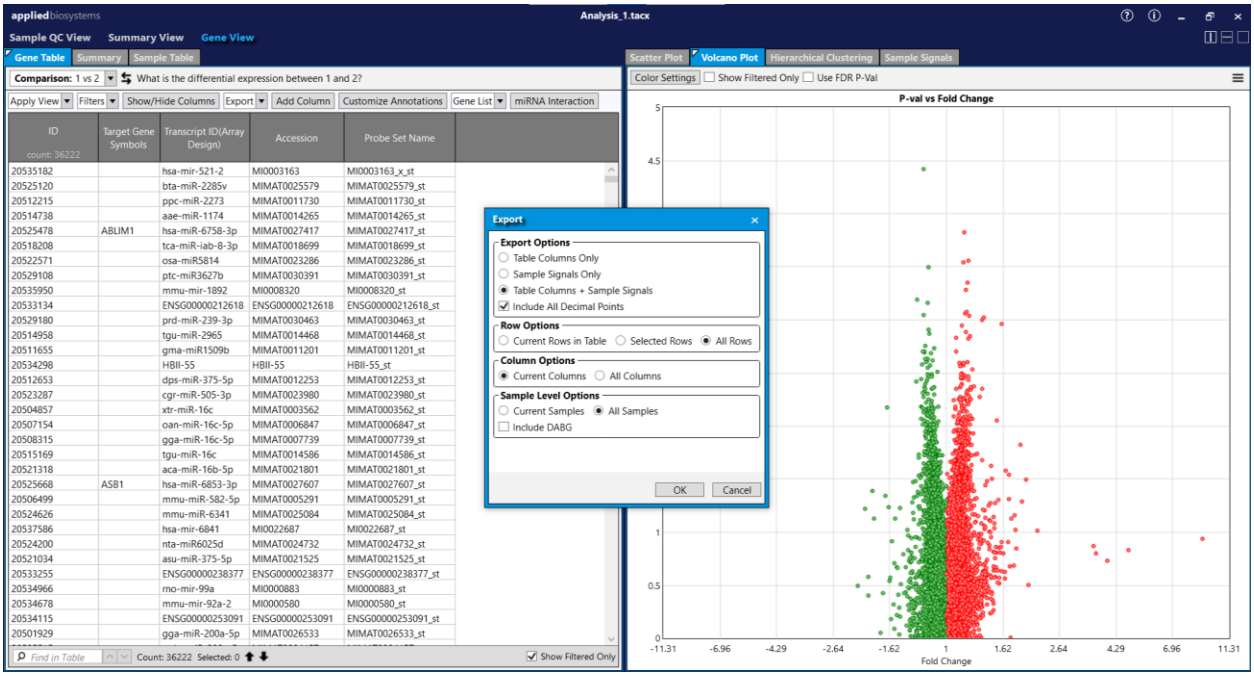
- Click the “Gene View” in the new window. And click “Customize Annotations” to select information columns for output. In the popped up window, check “Top Assignment” to pick the first subfield in each annotation column and select gene information columns you need. Click “OK”.



- Click “Show/Hide Columns”, a window will pop up. You can select columns for output. Click “OK”.



- Click “Export” > Export…, an Export window pops up. You can select the output contents. Click “OK”.

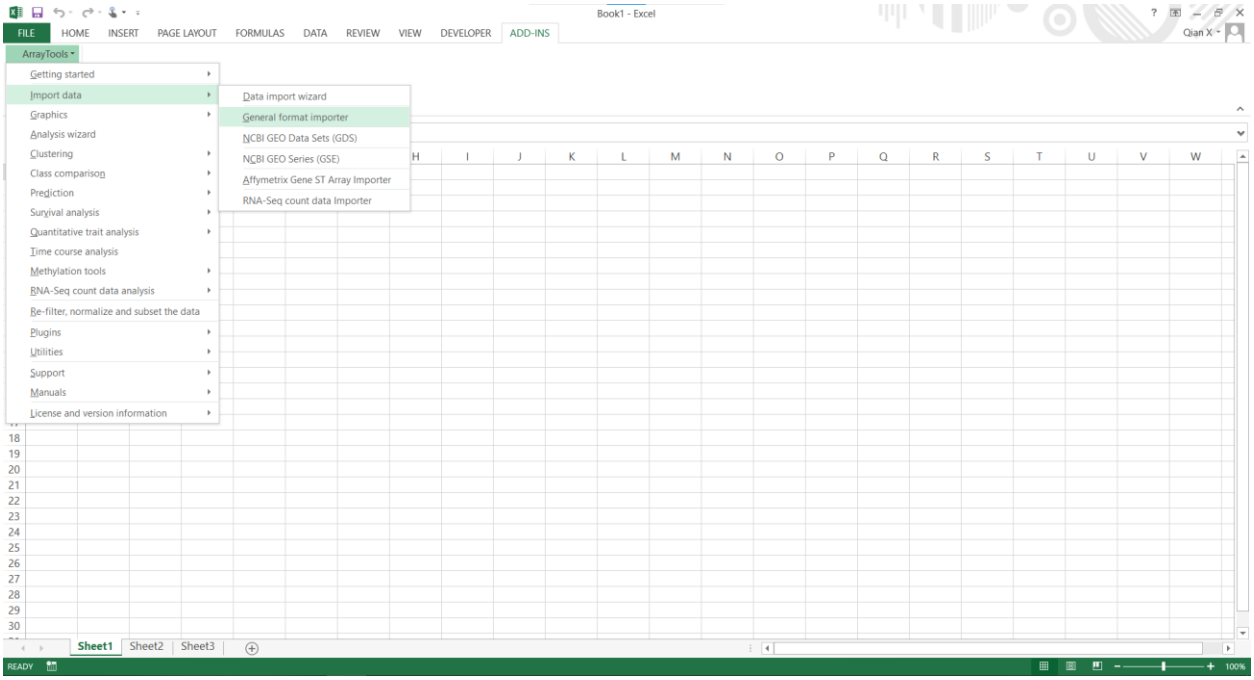


- A file browser window pops up. Give an Export File name: GSE123377.txt. Click “Save”. The probe summarization TXT file will be saved onto your local drive.

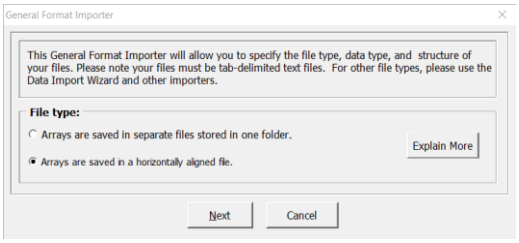
4. Import the TXT file in BRB-ArrayTools

You may modify the title names of the array signals in the TXT file with a text editor before importing it in BRB-ArrayTools. For example, you may shorten the title name “GSM3502388_4877_1_miRNA-4_0_.rma-dabg.chp Signal” to be “GSM3502388”.

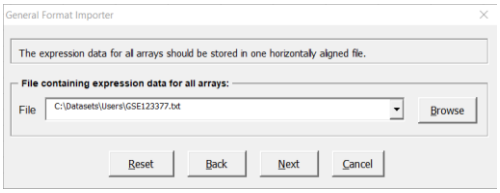
Start Excel and import the TXT file with the BRB-ArrayTools “General format importer”.



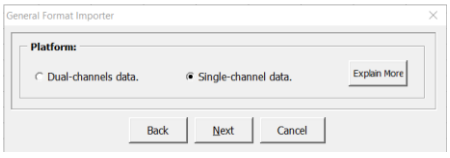
Select “Arrays are saved in a horizontally aligned file.” And click “Next”.



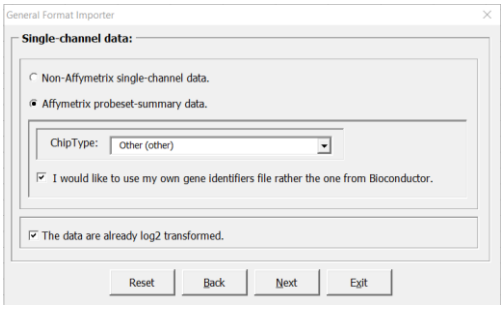
Browse for the TXT file and click “Next”.



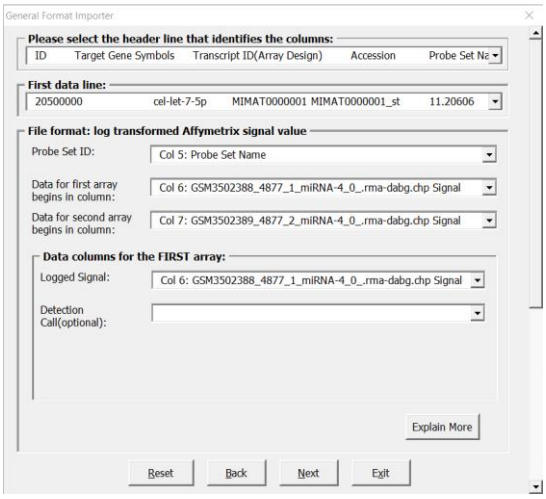
Select “Single-channel data.” And click “Next”.



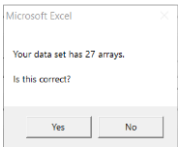
Select “Affymetrix probeset-summary data.” And click “Next”.



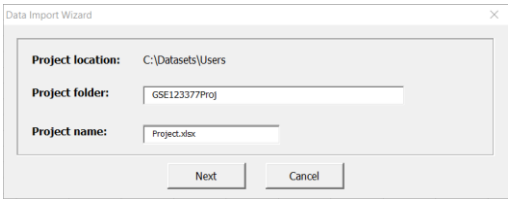
Select correct lines and columns for the fields. And click “Next”.



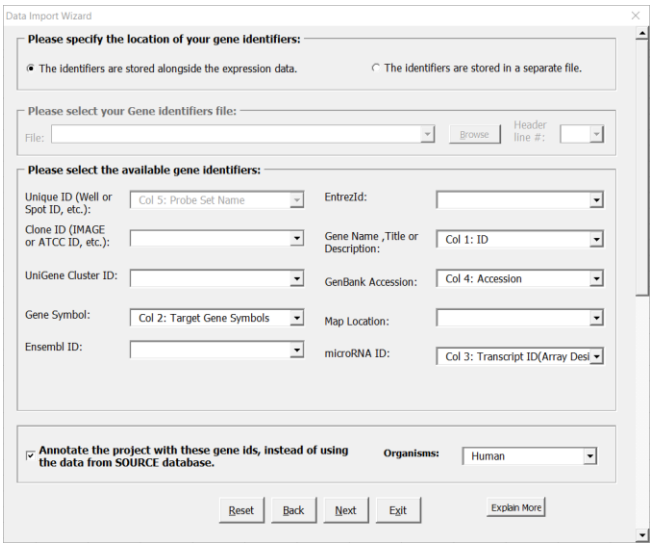
Confirm by clicking “Yes”.



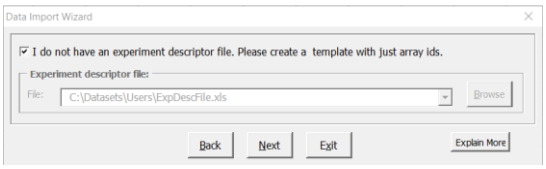
Give a BRB-ArrayTools project folder name or use the default one and click “Next”.



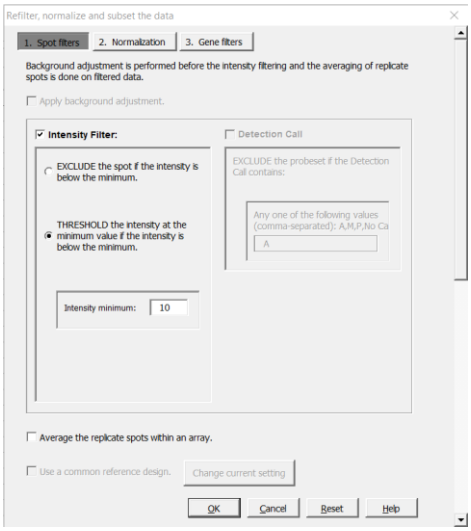
Select correct columns for different gene identifiers and click “Next”.



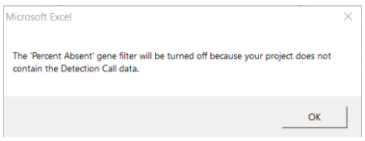
Check “I do not have an experiment descriptor file. Please create a template with just array Ids.” And click “Next”.



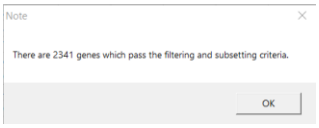
Confirm by clicking “OK”.



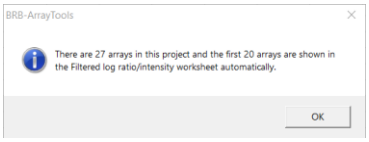
Confirm by clicking “OK”.



Confirm by clicking “OK”.

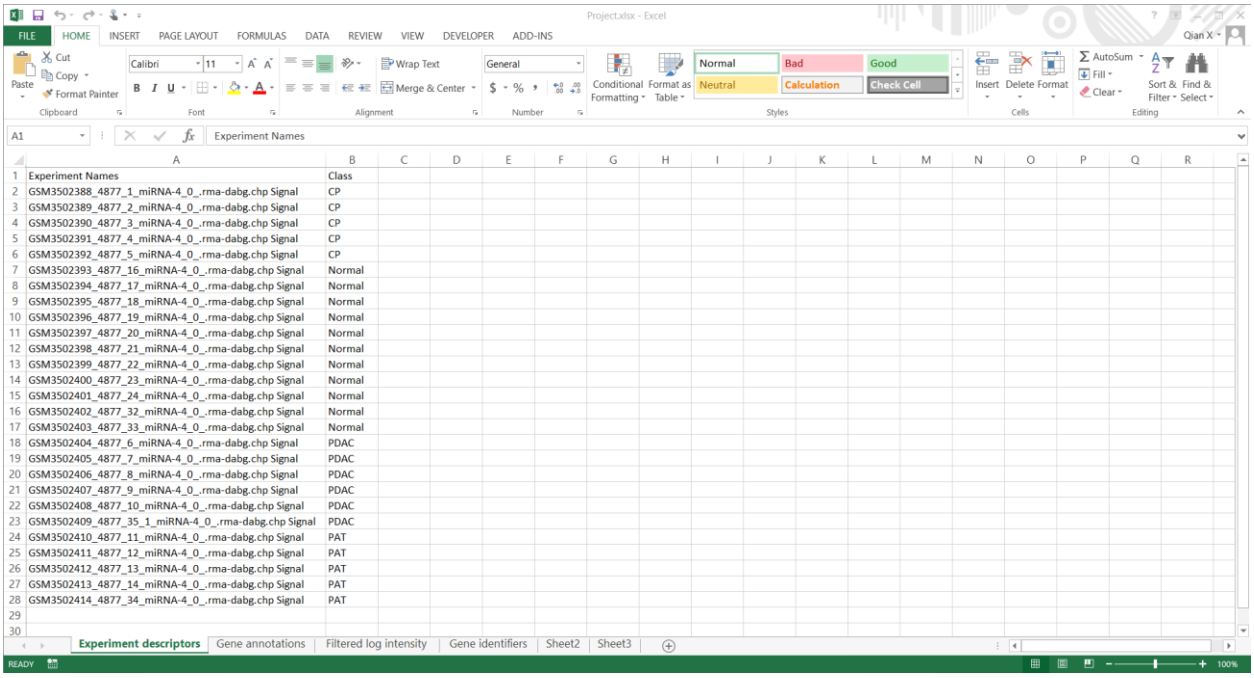


Confirm by clicking “OK”.



Now the BRB-ArrayTools project is collated successfully.

You may edit the worksheet “Experiment descriptors” to add sample information. You can run all the analysis tools in BRB-ArrayTools.



	A	B
1	Experiment Names	Class
2	GSM3502388_4877_1_miRNA-4_0_.rma-dabg.chp Signal	CP
3	GSM3502389_4877_2_miRNA-4_0_.rma-dabg.chp Signal	CP
4	GSM3502390_4877_3_miRNA-4_0_.rma-dabg.chp Signal	CP
5	GSM3502391_4877_4_miRNA-4_0_.rma-dabg.chp Signal	CP
6	GSM3502392_4877_5_miRNA-4_0_.rma-dabg.chp Signal	CP
7	GSM3502393_4877_16_miRNA-4_0_.rma-dabg.chp Signal	Normal
8	GSM3502394_4877_17_miRNA-4_0_.rma-dabg.chp Signal	Normal
9	GSM3502395_4877_18_miRNA-4_0_.rma-dabg.chp Signal	Normal
10	GSM3502396_4877_19_miRNA-4_0_.rma-dabg.chp Signal	Normal
11	GSM3502397_4877_20_miRNA-4_0_.rma-dabg.chp Signal	Normal
12	GSM3502398_4877_21_miRNA-4_0_.rma-dabg.chp Signal	Normal
13	GSM3502399_4877_22_miRNA-4_0_.rma-dabg.chp Signal	Normal
14	GSM3502400_4877_23_miRNA-4_0_.rma-dabg.chp Signal	Normal
15	GSM3502401_4877_24_miRNA-4_0_.rma-dabg.chp Signal	Normal
16	GSM3502402_4877_32_miRNA-4_0_.rma-dabg.chp Signal	Normal
17	GSM3502403_4877_33_miRNA-4_0_.rma-dabg.chp Signal	Normal
18	GSM3502404_4877_6_miRNA-4_0_.rma-dabg.chp Signal	PDAC
19	GSM3502405_4877_7_miRNA-4_0_.rma-dabg.chp Signal	PDAC
20	GSM3502406_4877_8_miRNA-4_0_.rma-dabg.chp Signal	PDAC
21	GSM3502407_4877_9_miRNA-4_0_.rma-dabg.chp Signal	PDAC
22	GSM3502408_4877_10_miRNA-4_0_.rma-dabg.chp Signal	PDAC
23	GSM3502409_4877_35_1_miRNA-4_0_.rma-dabg.chp Signal	PDAC
24	GSM3502410_4877_11_miRNA-4_0_.rma-dabg.chp Signal	PAT
25	GSM3502411_4877_12_miRNA-4_0_.rma-dabg.chp Signal	PAT
26	GSM3502412_4877_13_miRNA-4_0_.rma-dabg.chp Signal	PAT
27	GSM3502413_4877_14_miRNA-4_0_.rma-dabg.chp Signal	PAT
28	GSM3502414_4877_34_miRNA-4_0_.rma-dabg.chp Signal	PAT
29		
30		

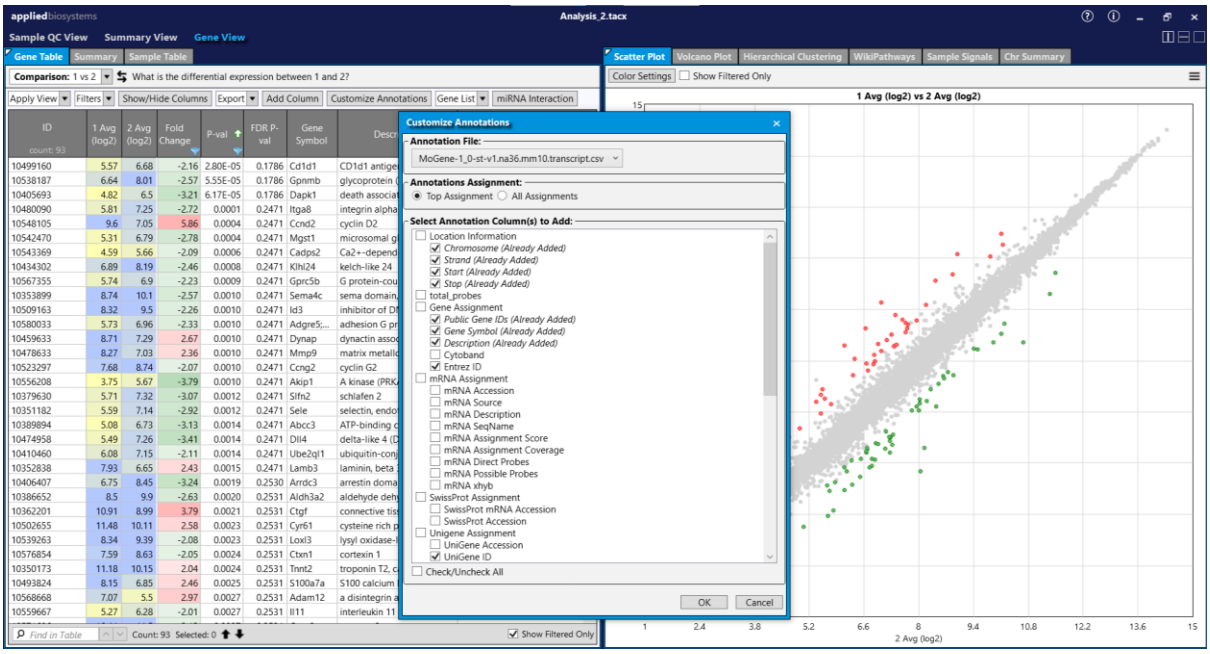
5. Examples of other array types

Other types of arrays data can be imported into BRB-ArrayTools by following the same steps given above. Different array types may have different annotation items for output. In the following examples, popular gene information columns are selected. You may select different columns for your needs.

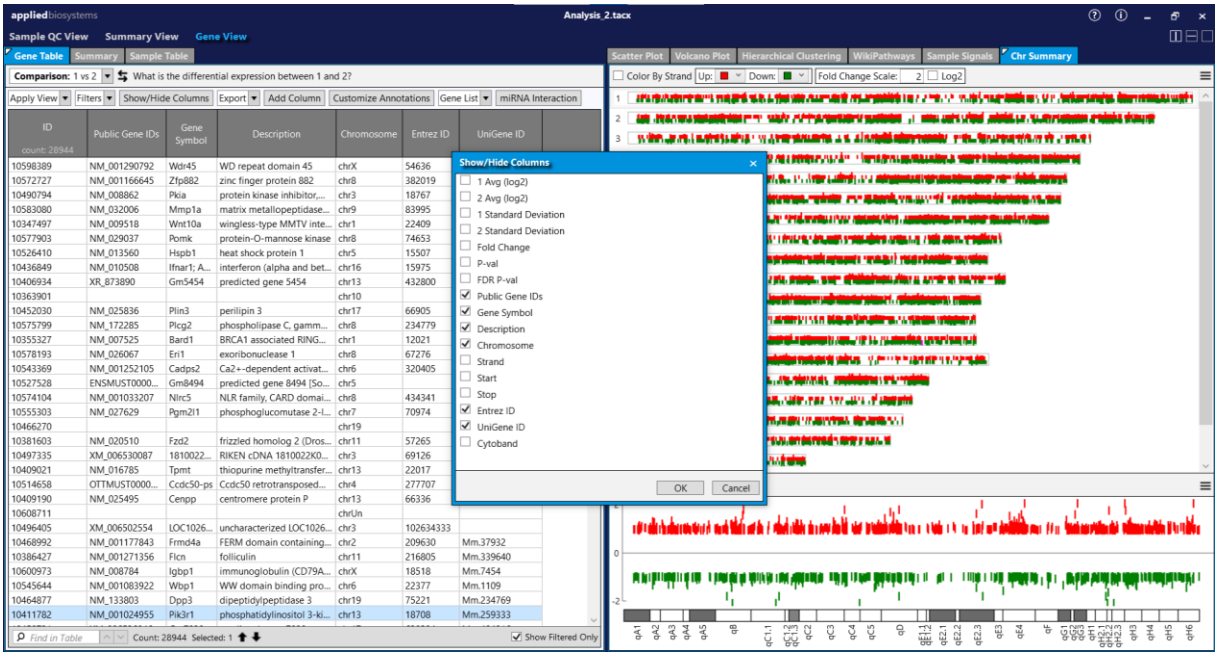
5.1. Affymetrix Mouse Gene 1.0 ST Array

NCBI GEO GSE54773: Affymetrix Mouse Gene 1.0 ST Array. You can follow the above steps to convert CEL files to a probe summarization TXT file and import it in BRB-ArrayTools. You may pay attention to the choices in the following steps:

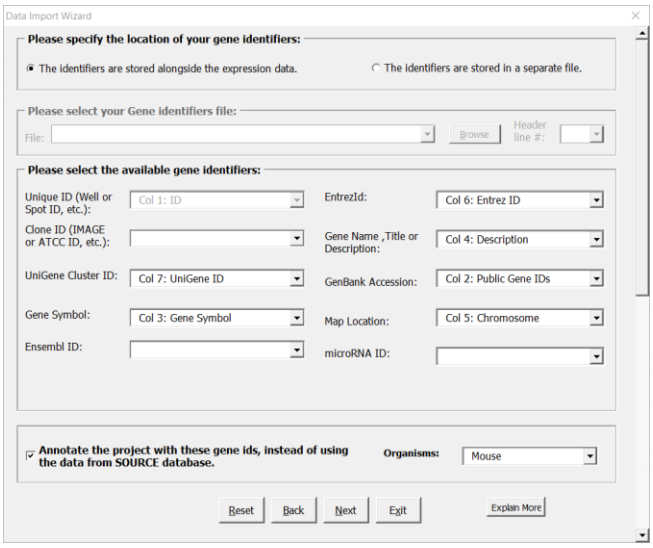
Click “Customize Annotations” to select information columns for output in TAC.



Click “Show/Hide Columns” to select output columns for the TXT file in TAC.



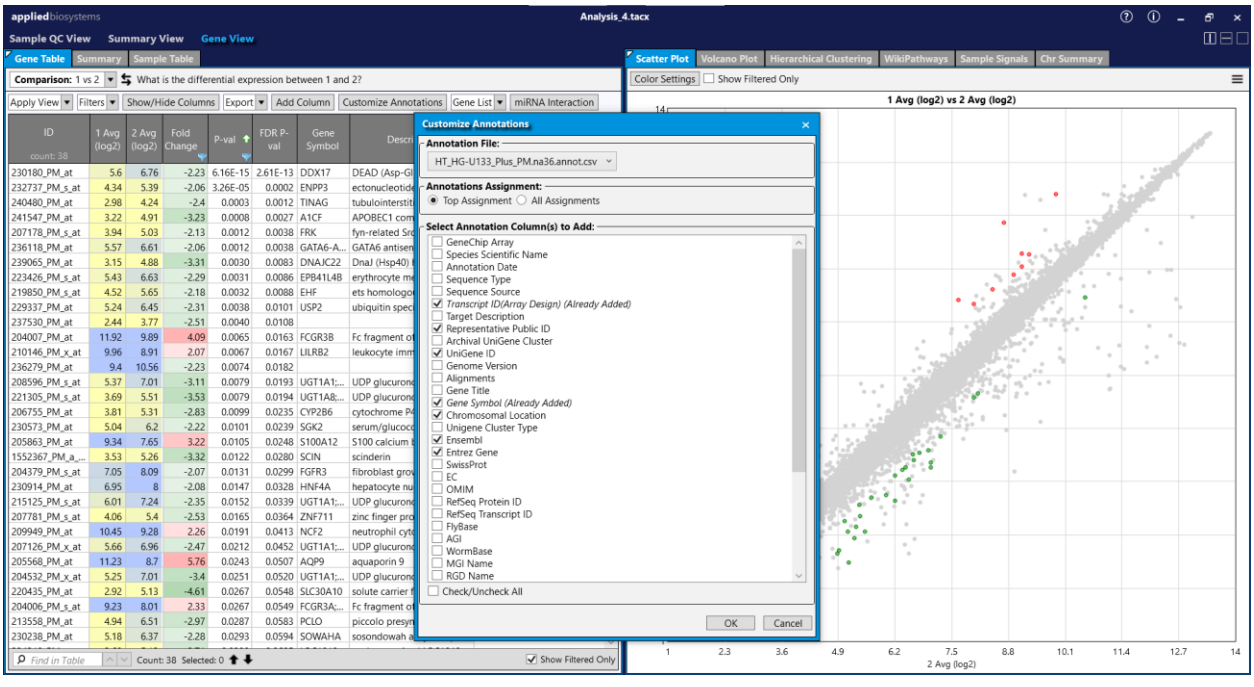
Assign different gene identifiers in BRB-ArrayTools.



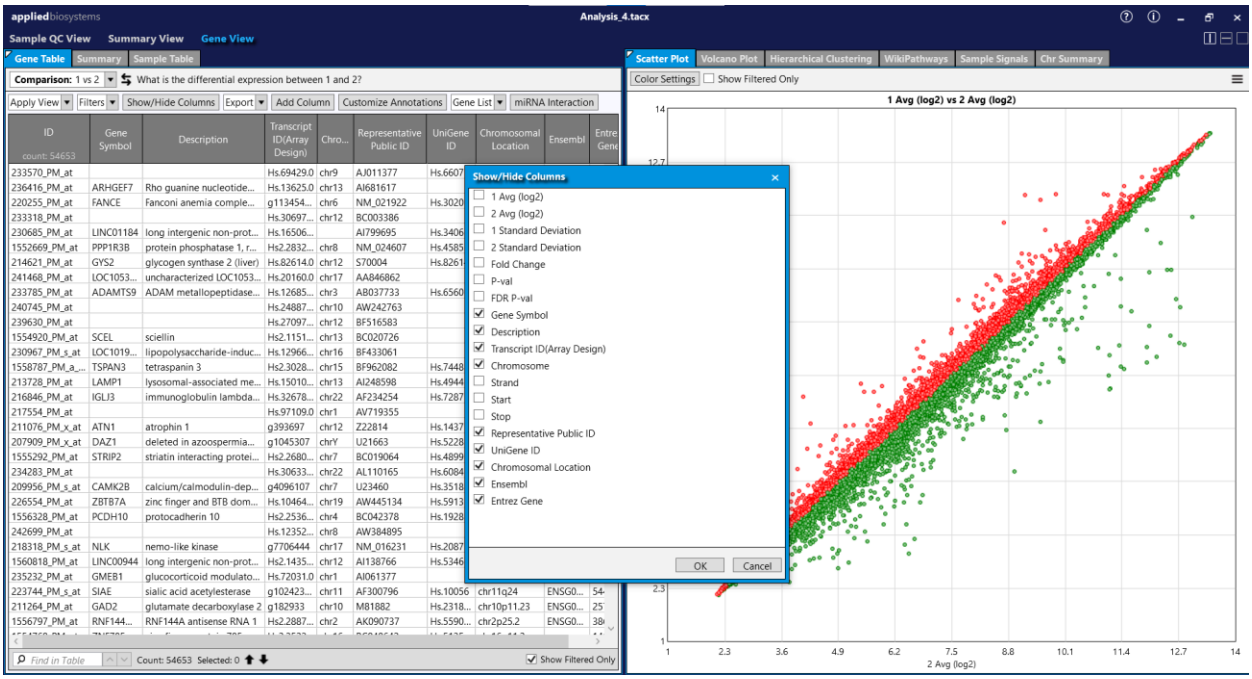
5.2. Affymetrix HT HG-U133+ PM Array

NCBI GEO GSE100833: Affymetrix HT HG-U133+ PM Array Plate. You can follow the above steps to convert CEL files to a probe summarization TXT file and import it in BRB-ArrayTools. You may pay attention to the choices in the following steps:

Click “Customize Annotations” to select information columns for output in TAC.



Click “Show/Hide Columns” to select output columns for the TXT file in TAC.



Assign different gene identifiers in BRB-ArrayTools.