

# *QAmmerge* 1.0e Readme

It is widely accepted that any single microarray output is subjected to substantial variability. By pooling data from replicates, we can provide a more reliable classification of gene expression. Designing experiments with replications will greatly reduce misclassification rates.

The purpose of program *QAmmerge* reflects our efforts toward to trying a reasonable way to merge replicates data sets obtained from image analysis software QuantArray, so that a user can use the overall information provided by these replicated experiments.

## Features

1. Automatical file format detecting:

This program works on original QA output files in tab delimited ASCII text format, but QuantArray's default output file is in tab delimited unicode text format, so it must be converted to a tab delimited ASCII text file. *QAmmerge* will detect the input file format automatically, if the input file of a user is not in required format, the program will remind the user converting the file to ASCII text format.

2. QA data reduction:

*QAmmerge* has a incorporated background subtraction & normalization function. Individual channel intensities are background subtracted and normalized to total channel intensities. This works in the same way as Excel QA Data Reduction Macro, so the user don't need to conduct the background subtraction and normalization with the Excel Macro any more, *QAmmerge* will do all this automatically. The background subtracted & normalized channel intensities and related individual ratio for each data set are listed in output file *data.txt*, so that a user can use them for further studies.

3. Data sets merging:  
*QAMerge* calculates the mean of ratio and medium of ratio with all useful replicated spots obtained from multiple data sets. We listed the mean of ratio, the standard deviation of mean ratio, the median of ratio, the mean of Log-Ratio(base 2) and the standard deviation of mean Log-Ratio(base 2) of multiple replicated data sets in the output file *merge.txt*. In *merge.txt* you can also find the Gene Number, Gene Name, the mean of ratio for each data set, the number of spots actually used to calculate the overall merged mean and medium of ratio, and the number of outliers.
4. Outliers searching  
A very simple outlier searching algorithm was incorporated in *QAMerge*, those spots which leads the large difference between the mean of ratio and the medium of ratio are defined as outliers and eliminated from the calculation. The number of outlier spots for each gene is listed in *merge.txt*. Moreover, for some spots the channel intensities after background subtraction may be less than 0, we eliminate these spots also from overall ratio calculation. The number of spots actually used in ratio calculation for each gene is also listed in *merge.txt*.
5. Alternative Ratio choosing  
With *QAMerge* a user can choose which ratio should be considered(Channel 1/Channel 2 or Channel 2/Channel 1). The related channel(fluor) information is listed in the header of *merge.txt*.
6. Typing-error Checking  
*QAMerge* has a systematical typing-error checking function, if user input a file name which doesn't exist or has been already used, the program will remind the user checking the error and prompt the user to modify the input. If a user enter a unlike parameter, the program will also prompt the user to enter that parameter again.

## Usage

1. QuantArray's output file is in unicode tab delimited format, but *QAMerge* require input files as ASCII tab delimited format. First of all a user

must convert the unicode text file to ASCII text file. There are two ways to do this.

- (a) Open QA output files with **Windows Notepad**, dissect **unicode box**, and save it again as text file, they are converted then to ASCII format.
- (b) Sometime if a file has very large size, you may fail to open it with **Windows Notepad**. In this case open QA output files with **MS Excel**, delete the two weird symbols manually(important! don't delete any valid characters), and then save them as tab delimited text files(they are now in ASCII format).

If you are using **Unix** operating system, try **Emacs** or other Unix programs to convert a unicode file to a ASCII file.

2. Put *QAmERGE.exe* and all files to be merged in one directory.
  - IF you are using windows 95/98/2000/NT, just click *QAmERGE.exe*, then the program will be activated. If you prefer to run the program under DOS command line, just open a DOS window from **start menu** of your computer, go to the folder where your files and program are, enter *QAmERGE* and **Return** key, the program will be activated.
  - If you are using a Unix operating system, just go to the folder where your files and program are, enter *QAmERGE* and **Return** key, the program will be activated.

Follow the program and enter the parameters, in general the total computing takes only several seconds.

3. *QAmERGE* has two output files: *merge.txt* and *data.txt*. Both files are in tab delimited text format, you can open them with **MS Excel** for further studies.

## Parameters

The *QAmERGE* is very easy to use, it requires original QA output files in ASCII tab delimited format, so the only thing that a user should do is to

convert the output files from unicode format to ASCII format(see above), besides this a user should do nothing to QA output files.

Besides file names of data sets to be merged, the program requires only three parameters, which should be entered by user. They are: 1. The number of data sets to be merged; 2. The number of replicate spots in each data set. 3. Which kind of ratio should be considered (channel 2 :channel 1 or channel 1 : channel 2).

## Release and Version

The current version is *QAmmerge* 1.0e, which was released on Jan 16, 2000. *QAmmerge* was designed for the convenience of YMD(Yale Microarray Database) user to merge replicated data sets. No warranty is expressed or implied.

The *QAmmerge* 1.0e was released as a package in zip format. The package includes *QAmmerge* win32 executive program, a Readme in PDF format, two replicated data sets *test1.txt* and *test2.txt*. User can use these two files to test the program, in each data set there are 12 replicated spots.

## Contact

For bugs report and requirement of source code please contact *jinming.li@yale.edu*. For comments, suggestions and critics please contact *jinming.li@yale.edu*, *hongyu.zhao@yale.edu* and *kenneth.williams@yale.edu*.

## Last Update

Jan. 16, 2001.