MAQC-II – Rat titration Experiment Agilent 1-Color

Data processing explanation and recommendations

Anne Bergstrom Lucas
Agilent Technologies
August 19th, 2008

Raw data are provided in the form of one tab-delimited text (.txt) file per microarray – a total of 96 files. These files are output from Agilent's Feature Extraction (v9.5.3.1) software.

Each .txt file contains three tables of input parameters and output results. These tables are FEPARAMS, STATS, and FEATURES. The FEPARAMS table contains input parameters and options used to run Feature Extraction. The STATS table gives results derived from statistical calculations that apply to all features on the microarray. The last table, FEATURES contains the feature-level data. These tables and their columns are described in detail in the Agilent G2567AA Feature Extraction Software (v9.5) - Reference Guide.

Data Flagging

Data points should be flagged in the following order:

1. Features that meet the following criteria should be marked as ABSENT (**A**):

ControlType <> 0
OR
gProcessedSignal = gSurrogateUsed

(ControlType: FEATURES table col 11
gProcessedSignal: FEATURES table col 20
gSurrogateUsed: FEATURES table col 18)

2. Features that meet the following criterion should be marked as MARGINAL (**M**):

gIsFeatNonUnifOL <> 0

(gIsFeatNonUnifOL: FEATURES table col 37)

3. All other features should be marked as PRESENT (**P**).

Raw Data

The values from the gProcessedSignal column should be used as the raw (background-subtracted, spatially detrended) data value for each feature.

(gProcessedSignal: FEATURES table col 20)

Data Normalization

A 75th Percentile Normalized gProcessedSignal value should be calculated for data points marked PRESENT (**P**). All measurements on every microarray should be divided by the 75th Percentile gProcessedSignal value from that microarray. The 75th Percentile value for each microarray can be found in the STATS table in the column called gPercentileIntensityProcessedSignal.

(gProcessedSignal: FEATURES table col 20
gPercentileIntensityProcessedSignal: STATS table col 95)

Optional Data Filtering Flags

Features that have signals that are statistically significantly higher than the background signal can be filtered with the “Positive And Significant” boolean and the features that are 2.6 times the background standard deviation can be filtered with the “Well Above Background” boolean:

gIsPosAndSignif<>0
OR
gIsWellAboveBG<>0

(gIsPosAndSignif: FEATURES table col 44
gIsWellAboveBG: FEATURES table col 47)