Microarray Analysis and Normalization LOWESS v_2.3

DMG/JWM/AAP/RAK/JB
Updated 091803

Step 1: Scan slides using QuantArray software and assign USER DEFINED treatments to Cy3 or Cy5*

*example channels: High gainers/Low gainers.

*example channels: High exercisers/Low exercisers, High fat diet/Low fat diet

Step 2: "copy" any intensities > 64500 to output file, calculate raw intensity values, get USER DEFINED normalization , retrieve normalization choice, calculate normalization channel medians or means per sub-array. (pooled oligo (default)

a.) copy any channel intensities X_{ch1} , X_{ch2} > 64500 to output error file.
b.) calculate all raw intensity channel values X_{ch1} , X_{ch2} per sub-array:
X_{ch1} = (intensity-background) , X_{ch2} = (intensity-background)
c.) get user defined normalization choice $NX_{(pooled\ oligo)}$. (pooled oligo=default)
d.) retrieve $NX_{ch1\ (pooled\ oligo)}$, $NX_{ch2\ (pooled\ oligo)}$ per subarray
e.) sort $NX_{ch1\ (pooled\ oligo)}$, $NX_{ch2\ (pooled\ oligo)}$ per sub-array.
f.) get medians/mean:
$m_{ch1\ (pooled\ oligo)}$, $m_{ch2\ (pooled\ oligo)}$ of $NX_{ch1\ (pooled\ oligo)}$, $NX_{ch2\ (pooled\ oligo)}$
g.) remove any subarrays with zero $m_{ch1\ (pooled\ oligo)}$, $m_{ch2\ (pooled\ oligo)}$

Step 3: Normalize all gene intensities per channel per subarray [48].

a.) Normalize all gene intensities per channel C_{ch1} , C_{ch2} per sub-array.

$$N_{ch1} = \frac{C_{ch1}}{m_{ch1(pooled\ oligo)}} \qquad N_{ch2} = \frac{C_{ch2}}{m_{ch2\ (pooled\ oligo)}}$$

Step 4: Remove unwanted genes, remove intensities < default = 0,

a.) Remove USER DEFINED genes , ie: salmon sperm, pos/neg controls, etc. (default = salmon Sperm DNA (Cy5), 'blanks', neg/pos).
b.) remove any genes with raw intensity values X_{ch1} (intensity-background) , X_{ch2} (intensity-background) equal to zero (or USER DEFINED intensity threshold).

NOTE: luciferase may be used in the future for analyses

Step 5: normalize each subarray using LOWESS

a.) remove any N_{ch1} $N_{ch2} \leq 0$

a.) calculate $LogM_{ch1}$ $Log\ A_{ch2}$ from N_{ch1} N_{ch2} per gene, per subarray.

b.) sort by $LogA_{ch2}$ per subarray

c.) lowess each subarray (resulting in a M_{fit} per gene, per subarray)

$M_{lowess} = M_{ch1} - M_{fit}$

All slides Processed? — N / Y

Step 7: Analyze gene normalized ratios.

a.) get list of M_{new} per gene per subarray per number per file (slide).
b.) get **Mean** & **S.D.**
c.) (option: get USER DEFINED μ), default $\mu = 0$
d.) perform 2-sided Students **T-test** (test of mu, small n)
e.) get **p-value** for informational reference.
f.) get USER DEFINED α
g.) determine Confidence Intervals for μ:

$$t = \frac{(\bar{X}_R - \mu)}{(s/\sqrt{n})}$$

$$\bar{X} - t_{\alpha/2,\,(n-1)\,d.f.}(s/\sqrt{n}) < \mu < \bar{X} + t_{\alpha/2,\,(n-1)\,d.f.}(s/\sqrt{n})$$

h.) return p-value, t-score, and Confidence Intervals to the user for analysis.

Math equations:

$$\text{Sample variance} = S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{(n-1)}$$

$$\text{Standard Deviation} = S = \sqrt{S^2}$$

T-Test : (2-sided) $H_a: \mu \neq 0$ $H_o: \mu = 0$ $(\log_2(1) = 0)$