What does analyze.iTRAQ( )?

Oct. 22, 2012
Lisa Chung

R function, do.iTRAQ( ) is written to take one run of 4-plex or 8-plex iTRAQ experiment. It performs cyclic-loess normalization [ref] and pair-wise comparison for differential protein-level expression. For each protein, peptide-level intensities are observed. Paired t-test is used to test for the significance of mean intensities (log2-scaled) between two experimental groups. (Each peptide is considered as a 'pair'.)

In the following, step-by-step procedure is described to use the function in R [ref].

Step 0. Pre-step to download the source code and install required package(s).

(1) Download following R script file from ***: iTRAQ_Normalization_FoldChange.R

(2) Start R. R package limma [ref] is required. To install this package, type the following on R Console

```
> install.packages("limma")
```

It will show a popup for CRAN mirror. Choose a CRAN mirror to download the package. The package only needs to be installed once. Alternatively, one can install the package from Bioconductor [ref]

```
> source("http://bioconductor.org/biocLite.R")
> biocLite("limma")
```

Step 1. Execute the source code.

In Windows R GUI, click File (on the top left menu) -> Source R code. On a pop-up, find and select the R script file (iTRAQ_Normalization_FoldChange.R). It runs the R script file and execute the function, do.iTRAQ( )

Alternatively, type the following on R Console (make sure to specify right path for the file):

```
> source("iTRAQ_Normalization_FoldChange.R")
```

If you want to use the function again (after sourcing the code once), following command will execute the function:

```
> analyze.iTRAQ()
```

Step 2. Use the do.iTRAQ().

(1) The first message will be printed to ask normalization process.

```
Do you want to Normalize data (Y if Yes):
```

Entering **y** or **Y** will start cyclic loess normalization [ref].

(2) As the following message is printed, a pop-up is opened. Choose the data file.

```
Choose the data file:
```

Let's assume that the data file name is **itraq.txt**. The function automatically creates a directory **itraq** where the data file is located. All result tables and graphics will be saved under this directory.

The file should be tab-delimited, *** output. This file should contain Accessions, Names, Annotation, Cleavages, and intensities (with column names Area.114, Area.115, etc). See the following example data:

Accessions	Names	Annotation	Cleavages	Area.114	Area.115	Area.116	Area.117
Meta990273986| Meta_contig_5537	hemocyanin subunit 1	auto		124.12	233.93	121.82	102.19
Meta990273986| Meta_contig_5537	hemocyanin subunit 1	auto		133.47	200.37	169.12	104.13

```
Only pick Annotation = auto and Cleavages = empty..
```

Annotation and Cleavages are used to filter out some low-quality (?) data. In the following, peptides with Annotation = auto and Cleavage = NA (empty) are considered.

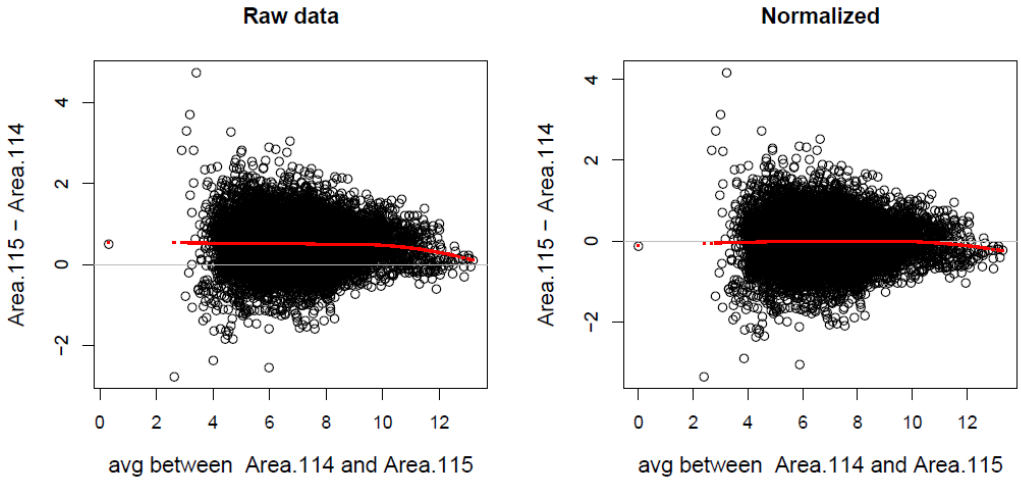
Step 3. Log-2 transformation and data normalization (if selected in Step 2 (1))

(1) Data is transformed in log2 scale. After the transformation, cycle-loess normalization is applied to remove intensity-dependent, systematic pattern. The result data is saved as:

AfterNormalization_**itraq**.txt if normalization is selected, otherwise, NoNormalization_**itraq**.txt. Normalized data will be shown in the very last 4 or 8 columns for 4-plex or 8-plex result file, respectively.

(2) This function further provides pair-wise scatter plots before and after the normalization. Please see the following graph (on the next page) as an example.

```
Making scatter plot for before/after normalization
```



Graphics for all pair-wise comparisons are provided on pairwiseScatterPlotForNormalization.pdf.

Step 4. Pair-wise comparison

(1) First, choose a cutoff for fold-change (unlogged scale). It will be used to make a result graphics and tables.

```
What cutoff do you want?
1.2
```

(2) Users can specify a pair of area intensities to be compared. When the following is asked, one can type **y** or **Y** to specify the pair. Not choosing the pair will make all possible comparisons (6 pairs for 4-plex and 28 pairs for 8-plex):

```
Want to choose two groups to be compared? Not doing it will
make ALL pairwise comparison. (y/n)
y
```

Then, two groups should be specified:

```
write the label names in GROUP1 to be compared, if more than
one, separate by COMMAs:
114,115

write the label names in GROUP2 to be compared, if more than
one, separate by COMMAs:
116,117
```

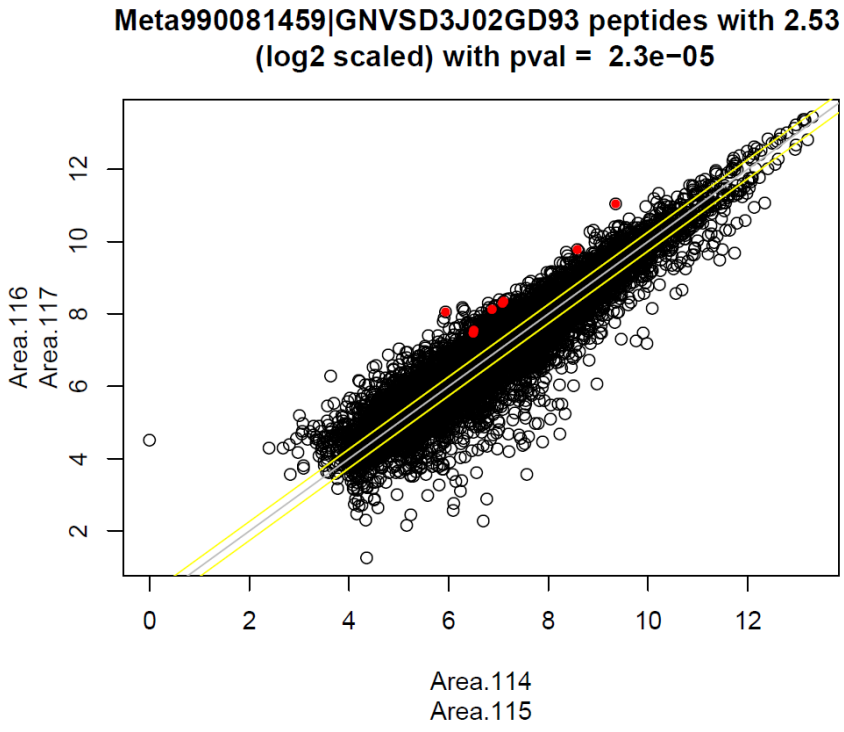
In this example, Area.114 and Area.115 are assigned to be Group 1, and Area.116 and Area.117 are Group 2. This example can be used if Area.114 and Area.115 are technical replicates and similarly for Area.116 and Area.117. Choosing more than one channel (? Isotope?) will use the average intensities for each peptide-level expression.

The comparison provides the following result table, summarized at protein level. Each of two tables gives analysis result from all proteins and significant proteins. Significant proteins are chosen at a pre-specified cutoff among proteins with at least two peptides.

- Area.114Area.115vsArea.116Area.117_1.2foldChange2012-10-25_All.txt
- Area.114Area.115vsArea.116Area.117_1.2foldChange2012-10-25_Signif.txt

Accessions	Names	unlog.fc	avg.log2.fc	median.log2.fc	sd.log2.fc	cv.log2.fc	p.val.from.ttest	num. peptides
gi|23598403; Meta990236143| Meta_contig_1753 gi|33469507	profilin [Euprymna scolopes]; beta-actin [Euprymna scolopes]	1.580	0.6606	0.570	0.443	0.670	4.34E-19	69
		0.654	-0.605	-0.605	0.065	-0.108	0.048	2

(3) For significant proteins at a pro-specified level of cutoff, do.iTRAQ( ) provides a summary graphic. In the following graph, each dot indicates each peptide in the iTRAQ experiment. Average of Area.114 and Area.115 (Group 1) is plotted on x-axis against its average of Area.116 and Area.117 (Group 2) on the y-axis. Red dots illustrate intensities of peptide from a given protein, Meta99008149|GNVSD3J02GD93. It has 6 peptides and the average of fold changes across these 6 peptides is 2.53 in log2 scale with a t-test p-value 2.3×10^{-5}. Grey line is a diagonal line and two yellow lines indicates the pre-chosen fold change cutoff of 1.2 (in log2 scale, +/- 0.263).



References: will be updated..