Meth27QC: QC Illumina Infinium HumanMethylation27 BeadChip Data Based on Control Probes

Ling Teng, Chao Chen and Chunyu Liu
January 15, 2011

Contents

1 Introduction

This document describes an R package for analyzing Illumina Infinium HumanMethylation27 BeadChip Data and generating QC reports. This package allows users quickly assess data quality of the Assay. Users can evaluate the data quality in the way that Illumina GenomeStudio/BeadStudio recommended based on the control probes. The package reads files exported from the GenomeStudio/BeadStudio software, generating intensity and standard deviation plots grouped by the types of the control probes. Meth27 carries 40 control probes for staining, hybridization, target removal, extension, bisulfite conversion, specificity, negative and non-polymorphic controls. Details of those control probes can be found in the Infinium Assay for Methylation Protocol Guide from Illumina.
We also used the other non-control probes to plot intensity of detected genes, signal average for green and red. Outliers can be identified.

2. Usage

After starting R, the package should be loaded by the following command.
> library(Meth27QC)

This will load Meth27QC as well as the gplots, and tcltk packages and their dependencies.

Then, define the path and filenames.
Dir: the location of the directory in which the input files are stored.
controlfile: control probe profile file name, exported from BeadStudio/GenomeStudio
samplefile: sample table file name, exported from BeadStudio/GenomeStudio
>Dir =”actual path”; controlfile=”control filename”; samplefile =”sample filename”;

Last, run the package.
> Meth27QC (Dir, controlfile, samplefile)

3 Inputs and Outputs files

Meth27QC takes in input the two files from GenomeStudio/BeadStudio plus a file in the inst/datafile directory, provided by Meth27QC package.

- Sample table
- Control probe profile table
- BeadtypeIDs (in inst/datafile directory of the package)

Inputs

Sample table - Required columns from BeadStudio:

- Index
- Sample ID
- Sample Group
- Sentrix Barcode
- Sample Section
- Detected Genes (0.01)
- Detected Genes (0.05)
- Signal Average GRN
- Signal Average RED
- Signal P05 GRN
- Signal P05 RED
- Signal P25 GRN
- Signal P25 RED
- Signal P50 GRN
- Signal P50 RED
- Signal P75 GRN
- Signal P75 RED
- Signal P95 GRN
- Signal P95 RED
- Sample Well
- Sample Plate

Control probe profile table - Required columns from BeadStudio (<Sn> = Sample Name):

- Index
- TargetID
- ProbeID
- <Sn>.Signal Grn
- <Sn>.Signal Red
- <Sn>.Detection Pval
- ...

BeadtypeIDs – Obtained from Illumina protocol (rows), provided by the package:

- BISULFITE CONVERSION (4 rows)
- EXTENSION (4 rows)
- HYBRIDIZATION (3 rows)

- NEGATIVE (16 rows)
- NON-POLYMORPHIC (4 rows)
- SPECIFICITY (4 rows)
- STAINING (4 rows)
- TARGET REMOVAL

Outputs

Meth27QC creates as output different plots (saved in a subdirectory named AnalyzedResults) to assess the quality of the samples:

1. ***Intensity plots for each internal control***. (Bisulfite_Control, Extension_Control, Hybridization_Control, Negative_Control, Non-Polymorphic_Control, Specificity_Control, Stain_Control, Target_Removal_Background)

2. ***Sample Intensity Plot.*** (intensity plot and box plot of Dectected Genes(0.01), Detected Genes(0.05), Signal Average GRN, Siagnal_Average RED, outliers are identified)

3. ***Pvalue_misMatchtable.*** (mismatch table between Detected_Pvalue and Expected_Intensity defined in BeadtypeIDs.txt file by Illumina protocol. Setting P value <0.05 as “detected”)

4 Acknowledgments

This package was inspired by the HumMeth27QCReport package created by Francesco Mancuso and the Illumina BeadStudio QC Reports.