

*Jongchul Seon*

*DNA Methylation example (Bisulfite Sequencing, Illumina 450k array)*

*Load*

*We are going to use public data, from <http://www.ncbi.nlm.nih.gov/gds/?term=GSE68777>. You can check The journal and main author here, <http://www.ncbi.nlm.nih.gov/pubmed/26147665> and <http://kasperdanielhansen.github.io/genbioconductor/html/minfi.html>*

*Microarray-based methods to determine pattern of methylation. The Illumina Methylation Assay is one such assay that applies the bisulfite sequencing technology on a microarray level to generate genome-wide methylation data.*

*We are going to compare DNA Methylation levels between two groups, Acute Mania(20) <-> Control(20)*

```
library(devtools)

## WARNING: Rtools is required to build R packages, but is not
## currently installed.
##
## Please download and install Rtools 3.3 from http://cran.r-
## project.org/bin/windows/Rtools/ and then run find_rtools().

library(GEOquery)

## Loading required package: Biobase
## Loading required package: BiocGenerics
## Loading required package: parallel
##
## Attaching package: 'BiocGenerics'
##
## The following objects are masked from 'package:parallel':
##
##   clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##   clusterExport, clusterMap, parApply, parCapply, parLapply,
##   parLapplyLB, parRapply, parSapply, parSapplyLB
##
## The following objects are masked from 'package:stats':
##
##   IQR, mad, xtabs
##
## The following objects are masked from 'package:base':
##
##   anyDuplicated, append, as.data.frame, as.vector, cbind,
##   colnames, do.call, duplicated, eval, evalq, Filter, Find, get,
```

```

##      grep, grepl, intersect, is.unsorted, lapply, lengths, Map,
##      mapply, match, mget, order, paste, pmax, pmax.int, pmin,
##      pmin.int, Position, rank, rbind, Reduce, rownames, sapply,
##      setdiff, sort, table, tapply, union, unique, unlist, unsplit
##
## Welcome to Bioconductor
##
##      Vignettes contain introductory material; view with
##      'browseVignettes()'. To cite Bioconductor, see
##      'citation("Biobase")', and for packages 'citation("pkgname")'.
##
## Setting options('download.file.method.GEOquery'='auto')
## Setting options('GEOquery.inmemory.gpl'=FALSE)

library(minfi)

## Loading required package: lattice
## Loading required package: GenomicRanges
## Loading required package: S4Vectors
## Loading required package: stats4
## Loading required package: IRanges
## Loading required package: GenomeInfoDb
## Loading required package: SummarizedExperiment
## Loading required package: Biostrings
## Loading required package: XVector
## Loading required package: bumpHunter
## Loading required package: foreach
## Loading required package: iterators
## Loading required package: locfit
## locfit 1.5-9.1    2013-03-22

library(doParallel) ##CRAN

## source("http://www.bioconductor.org/biocLite.R")
## biocLite(c("minfi"))

## getwd()

## data download

#getGEOSuppFiles("GSE68777")
#untar("GSE68777/GSE68777_RAW.tar", exdir = "GSE68777/idad")
#head(list.files("GSE68777/idad", pattern = "idad"))

#idadFiles <- list.files("GSE68777/idad", pattern = "idad.gz$", full =
TRUE)

#sapply(idatFiles, gunzip, overwrite = TRUE)

## data read

```

```
##rgSet <- read.450k.exp("GSE68777/idad")

rgSet <- read.450k.exp("D:/R/GSE68777/idad")
rgSet

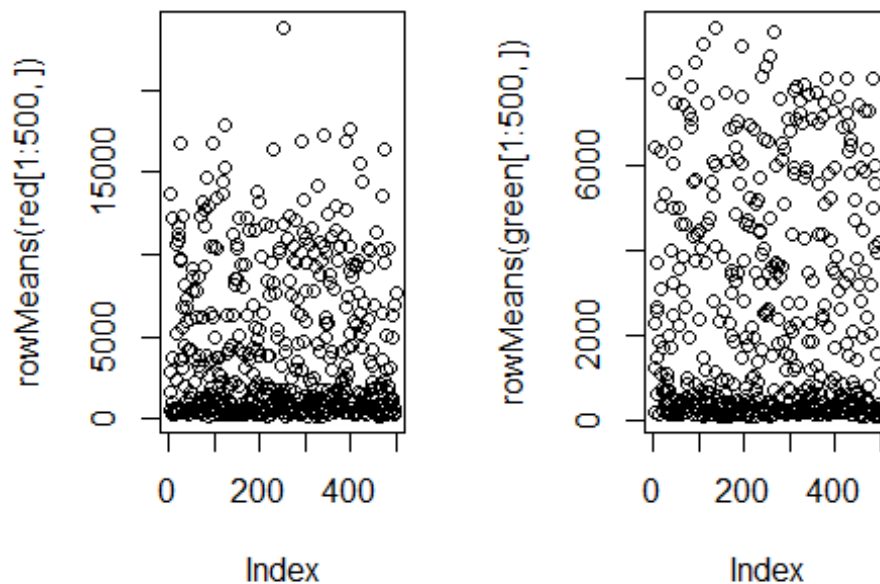
## RGChannelSet (storageMode: lockedEnvironment)
## assayData: 622399 features, 40 samples
##   element names: Green, Red
## An object of class 'AnnotatedDataFrame': none
## Annotation
##   array: IlluminaHumanMethylation450k
##   annotation: ilmn12.hg19
```

*idad is raw data file format.*

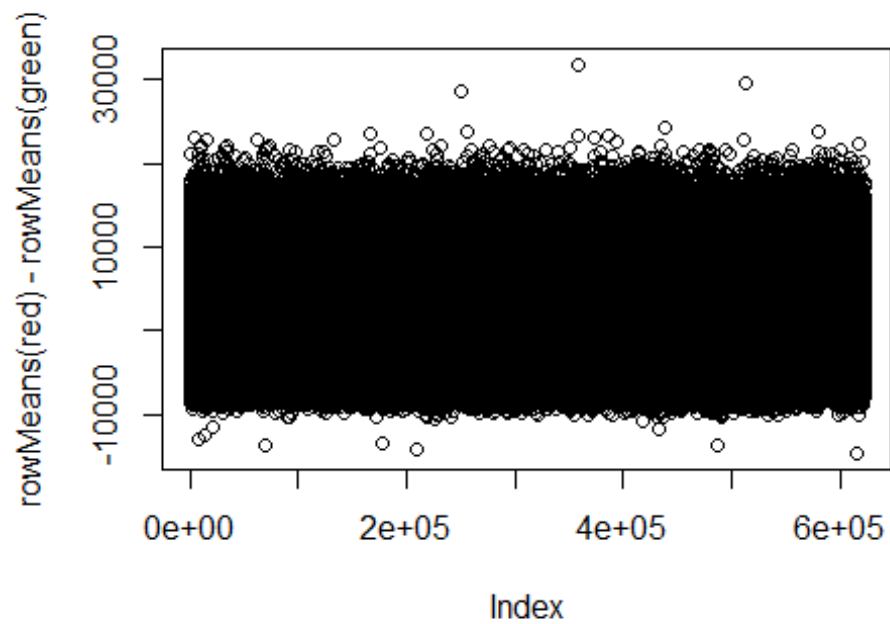
```
green <- getGreen(rgSet)
red <- getRed(rgSet)

par(mfrow=c(1,2))

plot(rowMeans(red[1:500,]))
plot(rowMeans(green[1:500,]))
```

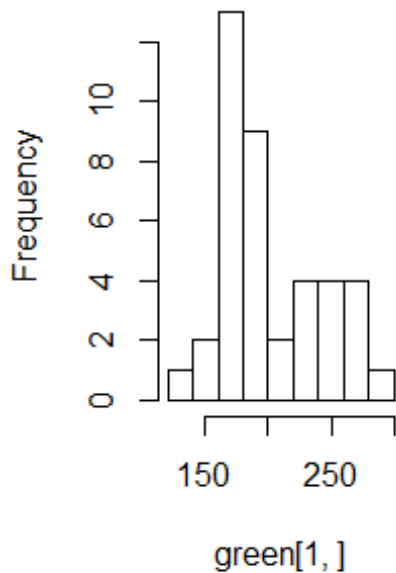


```
par(mfrow=c(1,1))  
plot(rowMeans(red)-rowMeans(green))
```

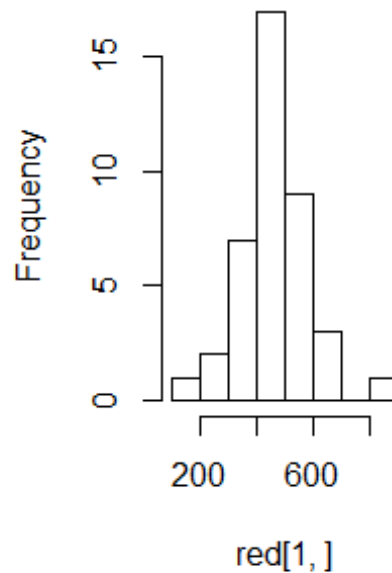


```
par(mfrow=c(1,2))  
hist(green[1,])  
hist(red[1,])
```

**Histogram of green[1,]**



**Histogram of red[1,]**



```
## probe locations can be divided into 4 categories of CpG island.
(Island, OpenSea, Shelf, Shore)
table(getIslandStatus(rgSet))

## Loading required package:
IlluminaHumanMethylation450kanno.ilmn12.hg19

##
##  Island OpenSea  Shelf  Shore
##  150254  176047  47144  112067

pData(rgSet)

## data frame with 0 columns and 40 rows

##class(rgSet)
##??RGChannelSet
##??minfi

## library(illuminaio)
## library(IlluminaDataTestFiles)

## head(list.files("GSE68777/idad", pattern = "idad"))

## green1 <-
readIDAT("D:/R/GSE68777/idad/GSM1681154_5958091019_R03C02_Grn.idat")

## red1 <-
```

```

readIDAT("D:/R/GSE68777/idad/GSM1681154_5958091019_R03C02_Red.idat")

## names(green1)
## head(green1$Quants)

##summary(green1$Quants[,3])
##summary(red1$Quants[,1])
##hist(log2(green1$Quants[,1]))
##hist(log2(red1$Quants[,1]))

##tempdir()

```

### *The phenotype data*

```

###getGEO
###geoMat <- getGEO("GSE68777")
### destdir = "D:\\R" => window OS directory

geoMat <- getGEO("GSE68777",destdir = "D:\\R")

## ftp://ftp.ncbi.nlm.nih.gov/geo/series/GSE68nnn/GSE68777/matrix/
## Found 1 file(s)
## GSE68777_series_matrix.txt.gz
## Using locally cached version: D:\R\GSE68777_series_matrix.txt.gz
## Using locally cached version of GPL13534 found here:
## D:\R\GPL13534.soft

pD.all <- pData(geoMat[[1]])

class(pD.all)

## [1] "data.frame"

head(pD.all)

```

##		title	geo_accession	status
##	GSM1681154	5958091019_R03C02	GSM1681154	Public on Jun 22 2015
##	GSM1681155	5935446005_R05C01	GSM1681155	Public on Jun 22 2015
##	GSM1681156	5958091020_R01C01	GSM1681156	Public on Jun 22 2015
##	GSM1681157	5958091020_R03C02	GSM1681157	Public on Jun 22 2015
##	GSM1681158	5935403032_R05C01	GSM1681158	Public on Jun 22 2015
##	GSM1681159	5958091019_R04C02	GSM1681159	Public on Jun 22 2015
##		submission_date	last_update_date	type channel_count
##	GSM1681154	May 12 2015	Jun 23 2015	genomic 1
##	GSM1681155	May 12 2015	Jun 23 2015	genomic 1
##	GSM1681156	May 12 2015	Jun 23 2015	genomic 1
##	GSM1681157	May 12 2015	Jun 23 2015	genomic 1
##	GSM1681158	May 12 2015	Jun 23 2015	genomic 1
##	GSM1681159	May 12 2015	Jun 23 2015	genomic 1
##		source_name_ch1	organism_ch1	characteristics_ch1

```

## GSM1681154 human serum DNA Homo sapiens sample type: serum DNA
## GSM1681155 human serum DNA Homo sapiens sample type: serum DNA
## GSM1681156 human serum DNA Homo sapiens sample type: serum DNA
## GSM1681157 human serum DNA Homo sapiens sample type: serum DNA
## GSM1681158 human serum DNA Homo sapiens sample type: serum DNA
## GSM1681159 human serum DNA Homo sapiens sample type: serum DNA
##
## characteristics_ch1.1 characteristics_ch1.2 molecule_ch1
## GSM1681154 diagnosis: Mania Sex: Female genomic DNA
## GSM1681155 diagnosis: Mania Sex: Female genomic DNA
## GSM1681156 diagnosis: Ctr Sex: Male genomic DNA
## GSM1681157 diagnosis: Ctr Sex: Female genomic DNA
## GSM1681158 diagnosis: Mania Sex: Female genomic DNA
## GSM1681159 diagnosis: Mania Sex: Male genomic DNA
##
extract_protocol_ch1
## GSM1681154 genomic DNA was extracted and purified from serum samples
using EpiCenter MasterPure kit according to standard instructions
## GSM1681155 genomic DNA was extracted and purified from serum samples
using EpiCenter MasterPure kit according to standard instructions
## GSM1681156 genomic DNA was extracted and purified from serum samples
using EpiCenter MasterPure kit according to standard instructions
## GSM1681157 genomic DNA was extracted and purified from serum samples
using EpiCenter MasterPure kit according to standard instructions
## GSM1681158 genomic DNA was extracted and purified from serum samples
using EpiCenter MasterPure kit according to standard instructions
## GSM1681159 genomic DNA was extracted and purified from serum samples
using EpiCenter MasterPure kit according to standard instructions
##
## label_ch1 label_protocol_ch1 taxid_ch1
## GSM1681154 Cy5 and Cy3 Standard Illumina Protocol 9606
## GSM1681155 Cy5 and Cy3 Standard Illumina Protocol 9606
## GSM1681156 Cy5 and Cy3 Standard Illumina Protocol 9606
## GSM1681157 Cy5 and Cy3 Standard Illumina Protocol 9606
## GSM1681158 Cy5 and Cy3 Standard Illumina Protocol 9606
## GSM1681159 Cy5 and Cy3 Standard Illumina Protocol 9606
##
hyb_protocol
## GSM1681154 bisulphite converted DNA was amplified, fragmented and
hybridised to Illumina Infinium Human Methylation450K Beadchip using
standard Illumina protocol
## GSM1681155 bisulphite converted DNA was amplified, fragmented and
hybridised to Illumina Infinium Human Methylation450K Beadchip using
standard Illumina protocol
## GSM1681156 bisulphite converted DNA was amplified, fragmented and
hybridised to Illumina Infinium Human Methylation450K Beadchip using
standard Illumina protocol
## GSM1681157 bisulphite converted DNA was amplified, fragmented and
hybridised to Illumina Infinium Human Methylation450K Beadchip using
standard Illumina protocol
## GSM1681158 bisulphite converted DNA was amplified, fragmented and
hybridised to Illumina Infinium Human Methylation450K Beadchip using

```

standard Illumina protocol

## GSM1681159 bisulphite converted DNA was amplified, fragmented and hybridised to Illumina Infinium Human Methylation450K Beadchip using standard Illumina protocol

##

scan\_protocol

## GSM1681154 Arrays were imaged using BeadArray Reader using standard recommended Illumina scanner setting at

## GSM1681155 Arrays were imaged using BeadArray Reader using standard recommended Illumina scanner setting at

## GSM1681156 Arrays were imaged using BeadArray Reader using standard recommended Illumina scanner setting at

## GSM1681157 Arrays were imaged using BeadArray Reader using standard recommended Illumina scanner setting at

## GSM1681158 Arrays were imaged using BeadArray Reader using standard recommended Illumina scanner setting at

## GSM1681159 Arrays were imaged using BeadArray Reader using standard recommended Illumina scanner setting at

## data\_processing platform\_id contact\_name

## GSM1681154 Bioconductor minfi package GPL13534 Sarven,,Sabunciyan

## GSM1681155 Bioconductor minfi package GPL13534 Sarven,,Sabunciyan

## GSM1681156 Bioconductor minfi package GPL13534 Sarven,,Sabunciyan

## GSM1681157 Bioconductor minfi package GPL13534 Sarven,,Sabunciyan

## GSM1681158 Bioconductor minfi package GPL13534 Sarven,,Sabunciyan

## GSM1681159 Bioconductor minfi package GPL13534 Sarven,,Sabunciyan

## contact\_email contact\_department

contact\_institute

## GSM1681154 ssabunc1@jhmi.edu Pediatrics Johns Hopkins University

## GSM1681155 ssabunc1@jhmi.edu Pediatrics Johns Hopkins University

## GSM1681156 ssabunc1@jhmi.edu Pediatrics Johns Hopkins University

## GSM1681157 ssabunc1@jhmi.edu Pediatrics Johns Hopkins University

## GSM1681158 ssabunc1@jhmi.edu Pediatrics Johns Hopkins University

## GSM1681159 ssabunc1@jhmi.edu Pediatrics Johns Hopkins University

## contact\_address contact\_city

contact\_state

## GSM1681154 600 N. Wolfe Street, Blalock 1146 Baltimore MD

## GSM1681155 600 N. Wolfe Street, Blalock 1146 Baltimore MD

## GSM1681156 600 N. Wolfe Street, Blalock 1146 Baltimore MD

## GSM1681157 600 N. Wolfe Street, Blalock 1146 Baltimore MD

## GSM1681158 600 N. Wolfe Street, Blalock 1146 Baltimore



MD

## GSM1681159 600 N. Wolfe Street, Blalock 1146 Baltimore

MD

## contact\_zip/postal\_code contact\_country

## GSM1681154 21287 USA

## GSM1681155 21287 USA

## GSM1681156 21287 USA

## GSM1681157 21287 USA

## GSM1681158 21287 USA

## GSM1681159 21287 USA

##

supplementary\_file

## GSM1681154

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681154/GSM1681154\_5958091019\_R03C02\_Grn.idat.gz

## GSM1681155

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681155/GSM1681155\_5935446005\_R05C01\_Grn.idat.gz

## GSM1681156

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681156/GSM1681156\_5958091020\_R01C01\_Grn.idat.gz

## GSM1681157

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681157/GSM1681157\_5958091020\_R03C02\_Grn.idat.gz

## GSM1681158

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681158/GSM1681158\_5935403032\_R05C01\_Grn.idat.gz

## GSM1681159

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681159/GSM1681159\_5958091019\_R04C02\_Grn.idat.gz

##

supplementary\_file.1

## GSM1681154

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681154/GSM1681154\_5958091019\_R03C02\_Red.idat.gz

## GSM1681155

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681155/GSM1681155\_5935446005\_R05C01\_Red.idat.gz

## GSM1681156

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681156/GSM1681156\_5958091020\_R01C01\_Red.idat.gz

## GSM1681157

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681157/GSM1681157\_5958091020\_R03C02\_Red.idat.gz

## GSM1681158

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681158/GSM1681158\_5935403032\_R05C01\_Red.idat.gz

## GSM1681159

ftp://ftp.ncbi.nlm.nih.gov/pub/geo/DATA/supplementary/samples/GSM1681nn  
n/GSM1681159/GSM1681159\_5958091019\_R04C02\_Red.idat.gz

```
##          data_row_count
## GSM1681154          485512
## GSM1681155          485512
## GSM1681156          485512
## GSM1681157          485512
## GSM1681158          485512
## GSM1681159          485512
```

```
names(pD.all)
```

```
## [1] "title"          "geo_accession"
## [3] "status"         "submission_date"
## [5] "last_update_date" "type"
## [7] "channel_count"  "source_name_ch1"
## [9] "organism_ch1"   "characteristics_ch1"
## [11] "characteristics_ch1.1" "characteristics_ch1.2"
## [13] "molecule_ch1"  "extract_protocol_ch1"
## [15] "label_ch1"      "label_protocol_ch1"
## [17] "taxid_ch1"      "hyb_protocol"
## [19] "scan_protocol"  "data_processing"
## [21] "platform_id"    "contact_name"
## [23] "contact_email"  "contact_department"
## [25] "contact_institute" "contact_address"
## [27] "contact_city"   "contact_state"
## [29] "contact_zip/postal_code" "contact_country"
## [31] "supplementary_file" "supplementary_file.1"
## [33] "data_row_count"
```

```
rgSet
```

```
## RGChannelSet (storageMode: lockedEnvironment)
## assayData: 622399 features, 40 samples
##   element names: Green, Red
## An object of class 'AnnotatedDataFrame': none
## Annotation
##   array: IlluminaHumanMethylation450k
##   annotation: ilmn12.hg19
```

*The phenoData(pD) and the featureData(rgSet) are modified for further analysis.*

```
pD <- pD.all[, c("title", "geo_accession", "characteristics_ch1.1",
"characteristics_ch1.2")]
head(pD)
```

```
##          title geo_accession characteristics_ch1.1
## GSM1681154 5958091019_R03C02   GSM1681154      diagnosis: Mania
## GSM1681155 5935446005_R05C01   GSM1681155      diagnosis: Mania
## GSM1681156 5958091020_R01C01   GSM1681156      diagnosis: Ctr
## GSM1681157 5958091020_R03C02   GSM1681157      diagnosis: Ctr
## GSM1681158 5935403032_R05C01   GSM1681158      diagnosis: Mania
## GSM1681159 5958091019_R04C02   GSM1681159      diagnosis: Mania
##          characteristics_ch1.2
```

```

## GSM1681154          Sex: Female
## GSM1681155          Sex: Female
## GSM1681156          Sex: Male
## GSM1681157          Sex: Female
## GSM1681158          Sex: Female
## GSM1681159          Sex: Male

names(pD)[c(3,4)] <- c("group", "sex")
head(pD)

##              title geo_accession          group
sex
## GSM1681154 5958091019_R03C02    GSM1681154 diagnosis: Mania Sex:
Female
## GSM1681155 5935446005_R05C01    GSM1681155 diagnosis: Mania Sex:
Female
## GSM1681156 5958091020_R01C01    GSM1681156  diagnosis: Ctr   Sex:
Male
## GSM1681157 5958091020_R03C02    GSM1681157  diagnosis: Ctr Sex:
Female
## GSM1681158 5935403032_R05C01    GSM1681158 diagnosis: Mania Sex:
Female
## GSM1681159 5958091019_R04C02    GSM1681159 diagnosis: Mania   Sex:
Male

##pD$group
pD$group <- sub("^diagnosis: ", "", pD$group)
##pD$group

##pD$sex
pD$sex <- sub("^Sex: ", "", pD$sex)
##pD$sex

##sampleNames(rgSet)
sampleNames(rgSet) <- sub(".*_5", "5", sampleNames(rgSet))
##sampleNames(rgSet)

rownames(pD) <- pD$title
pD <- pD[sampleNames(rgSet),]

head(pD)

##              title geo_accession group    sex
## 5958091019_R03C02 5958091019_R03C02    GSM1681154 Mania Female
## 5935446005_R05C01 5935446005_R05C01    GSM1681155 Mania Female
## 5958091020_R01C01 5958091020_R01C01    GSM1681156   Ctr   Male
## 5958091020_R03C02 5958091020_R03C02    GSM1681157   Ctr Female
## 5935403032_R05C01 5935403032_R05C01    GSM1681158 Mania Female
## 5958091019_R04C02 5958091019_R04C02    GSM1681159 Mania   Male

```

```
pData(rgSet) <- pD
head(pData(rgSet))
```

```
##               title geo_accession group sex
## 5958091019_R03C02 5958091019_R03C02   GSM1681154 Mania Female
## 5935446005_R05C01 5935446005_R05C01   GSM1681155 Mania Female
## 5958091020_R01C01 5958091020_R01C01   GSM1681156   Ctr   Male
## 5958091020_R03C02 5958091020_R03C02   GSM1681157   Ctr Female
## 5935403032_R05C01 5935403032_R05C01   GSM1681158 Mania Female
## 5958091019_R04C02 5958091019_R04C02   GSM1681159 Mania   Male
```

```
rgSet
```

```
## RGChannelSet (storageMode: lockedEnvironment)
## assayData: 622399 features, 40 samples
##   element names: Green, Red
## An object of class 'AnnotatedDataFrame'
##   sampleNames: 5958091019_R03C02 5935446005_R05C01 ...
##     5935403032_R04C01 (40 total)
##   varLabels: title geo_accession group sex
##   varMetadata: labelDescription
## Annotation
##   array: IlluminaHumanMethylation450k
##   annotation: ilmn12.hg19
```

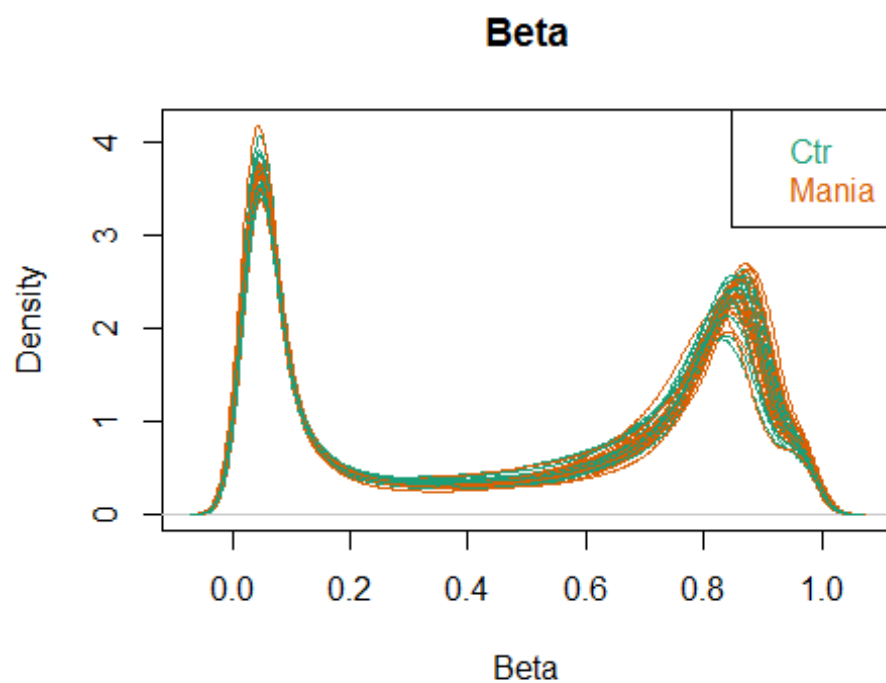
*Beta = Meth / (Meth + Unmeth + offset)*

*Beta value of Each sample should be left and right skewed.*

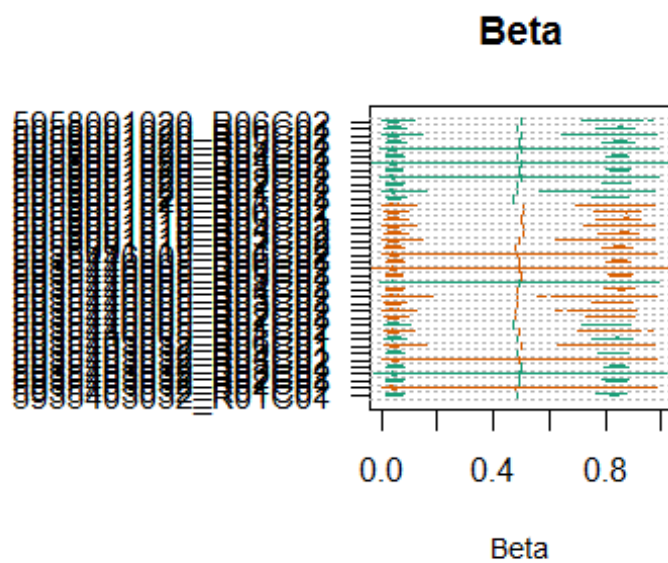
```
par(mfrow=c(1,1))
```

```
densityPlot(rgSet, sampGroups = pD$group, main = "Beta", xlab = "Beta")
```

```
## Loading required package: IlluminaHumanMethylation450kmanifest
```



```
par(oma=c(2,10,1,1))
densityBeanPlot(rgSet, sampGroups = pD$group,sampNames = pD$title)
```



## Normalization

$Beta = M/(M + U + 100) \Rightarrow$  Illumina formula

*M-values are perhaps an unfortunate terminology, but it seems to be standard in the methylation array world. These are computed as  $\logit(Beta)$  and are obtained by `getM`.*

```
MSet.raw <- preprocessRaw(rgSet)
```

```
getMeth(MSet.raw)[1:4,1:3]
```

```
##           5958091019_R03C02 5935446005_R05C01 5958091020_R01C01
## cg00050873           223           401           9722
## cg00212031           266           132           187
## cg00213748           151           242           1617
## cg00214611           237           237           200
```

```
getUnmeth(MSet.raw)[1:4,1:3]
```

```
##           5958091019_R03C02 5935446005_R05C01 5958091020_R01C01
## cg00050873           166           222           2162
## cg00212031           223           219           5281
## cg00213748           138           89           246
## cg00214611           65           303           5075
```

```
getBeta(MSet.raw, type = "Illumina")[1:4,1:3]
```

```
##           5958091019_R03C02 5935446005_R05C01 5958091020_R01C01
## cg00050873           0.4560327           0.5546335           0.81124833
## cg00212031           0.4516129           0.2926829           0.03358477
## cg00213748           0.3881748           0.5614849           0.82373917
## cg00214611           0.5895522           0.3703125           0.03720930
```

```
getM(MSet.raw)[1:4,1:3]
```

```
##           5958091019_R03C02 5935446005_R05C01 5958091020_R01C01
## cg00050873           0.4258605           0.8530426           2.168887
## cg00212031           0.2543825          -0.7303929          -4.819701
## cg00213748           0.1298803           1.4431298           2.716589
## cg00214611           1.8663754          -0.3544307          -4.665336
```

```
MSet.norm <- preprocessIllumina(rgSet, bg.correct = TRUE, normalize =
"controls", reference = 2)
```

## Finding differentially methylated positions (DMPs)

```
M <- getM(MSet.norm[1:20000,], type = "beta", betaThreshold = 0.001)
```

```
dmp <- dmpFinder(M, pheno=pD$group, type="categorical")
```

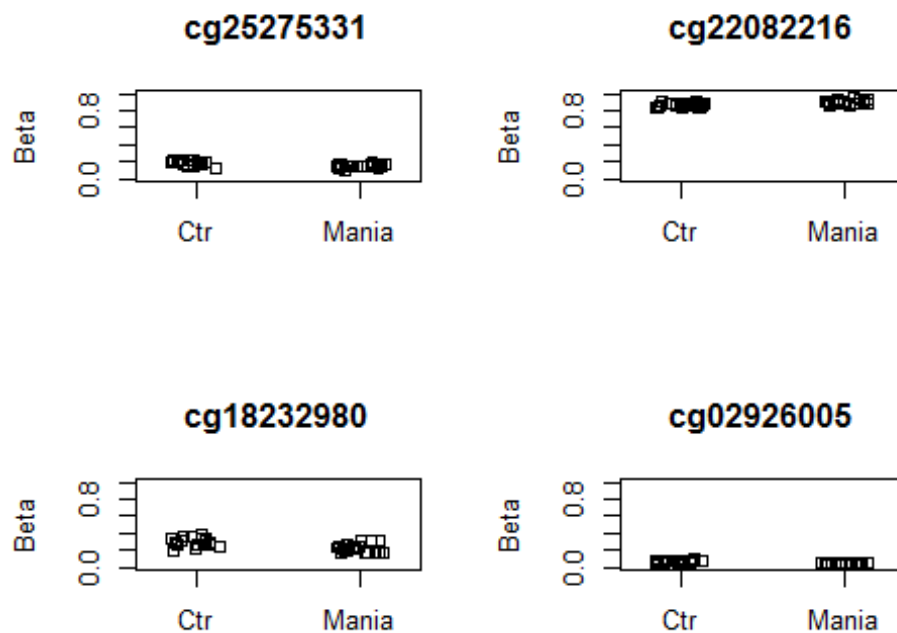
```
head(dmp)
```

```
##          intercept      f      pval      qval
## cg25275331 -2.163578 26.38418 8.659808e-06 0.04168854
## cg22082216  2.621089 24.56916 1.519647e-05 0.04168854
## cg18232980 -1.248161 23.65499 2.030851e-05 0.04168854
## cg02926005 -3.940681 23.52492 2.117194e-05 0.04168854
## cg17820022  4.182657 23.51310 2.125229e-05 0.04168854
## cg15466587 -8.250225 23.13320 2.401446e-05 0.04168854
```

```
cpgs <- rownames(dmp)[1:4]
```

```
par(mfrow=c(2,2))
```

```
plotCpg(MSet.norm[1:20000,], cpg=cpgs, pheno=pD$group)
```



### *Finding differentially methylated positions (Bumphunting)*

```
dat = mapToGenome(MSet.norm, mergeManifest = TRUE)
```

```
dat = ratioConvert(dat,type="Illumina")
```

```
detectCores()
```

```
## [1] 4
```

```
registerDoParallel(cores = 4)
```

```

group =pData(dat)$group
X= model.matrix(~group)
index = which(seqnames(dat)=="chr15")

dat = dat[index,] ## for illustrative purposes
res=bumphunter(dat,X,cutoff=0.1,B=1000)

## [bumphunterEngine] Parallelizing using 4 workers/cores (backend:
doParallelSNOW, version: 1.0.10).
## [bumphunterEngine] Computing coefficients.
## [bumphunterEngine] Performing 1000 permutations.
## [bumphunterEngine] Computing marginal permutation p-values.
## [bumphunterEngine] cutoff: 0.1
## [bumphunterEngine] Finding regions.
## [bumphunterEngine] Found 16 bumps.
## [bumphunterEngine] Computing regions for each permutation.
## [bumphunterEngine] Estimating p-values and FWER.

res

## a 'bumps' object with 16 bumps

res$tab

```

##	chr	start	end	value	area	cluster	indexStart
## 8	chr15	90927939	90927939	0.1838362	0.1838362	6034	13132
## 11	chr15	101593831	101593831	0.1588504	0.1588504	6864	14973
## 9	chr15	94147555	94147555	0.1475215	0.1475215	6348	13815
## 2	chr15	25337657	25337657	0.1382332	0.1382332	280	651
## 16	chr15	102025217	102025217	-0.1357141	0.1357141	6958	15161
## 6	chr15	76016056	76016056	0.1304861	0.1304861	4671	10225
## 7	chr15	81453270	81453270	0.1234999	0.1234999	5260	11424
## 10	chr15	101509442	101509442	0.1207726	0.1207726	6844	14931
## 15	chr15	99709980	99709980	-0.1163038	0.1163038	6681	14539
## 3	chr15	35529473	35529473	0.1109124	0.1109124	1123	2476
## 4	chr15	62899159	62899159	0.1089428	0.1089428	3077	6843
## 13	chr15	74345103	74345103	-0.1086848	0.1086848	4403	9457
## 12	chr15	28200653	28200653	-0.1076619	0.1076619	552	1344
## 14	chr15	81412880	81412880	-0.1040853	0.1040853	5253	11407
## 5	chr15	74494900	74494900	0.1020061	0.1020061	4423	9548
## 1	chr15	24043142	24043142	0.1006007	0.1006007	190	410
##	indexEnd	L	clusterL	p.value	fwer	p.valueArea	fwerArea
## 8	13132	1	1	0.09875178	0.834	0.1129194	0.861
## 11	14973	1	4	0.20529836	0.986	0.2184653	0.986
## 9	13815	1	1	0.26207405	0.997	0.2742403	0.997
## 2	651	1	3	0.33206931	1.000	0.3431822	1.000
## 16	15161	1	1	0.35845579	1.000	0.3692000	1.000
## 6	10225	1	5	0.42544899	1.000	0.4354032	1.000
## 7	11424	1	2	0.52235740	1.000	0.5307842	1.000
## 10	14931	1	1	0.56254279	1.000	0.5701796	1.000
## 15	14539	1	1	0.63374941	1.000	0.6397008	1.000



```
## 3      2476 1      5 0.72681308 1.000 0.7300258 1.000
## 4      6843 1      1 0.76662980 1.000 0.7690525 1.000
## 13     9457 1      2 0.77221257 1.000 0.7745299 1.000
## 12     1344 1      2 0.79496498 1.000 0.7969663 1.000
## 14    11407 1      1 0.88186654 1.000 0.8824986 1.000
## 5      9548 1     12 0.93980092 1.000 0.9399589 1.000
## 1       410 1      2 0.98098699 1.000 0.9810397 1.000
```

```
table(getIslandStatus(dat))
```

```
##
##  Island OpenSea  Shelf  Shore
##    4510   6379   1315   3055
```

There should be more necessary works related with Methylation Level Analysis.

```
devtools::session_info()
```

```
## Session info
```

```
-----
## setting  value
## version  R version 3.2.2 (2015-08-14)
## system   x86_64, mingw32
## ui       RTerm
## language (EN)
## collate  English_United States.1252
## tz       Asia/Seoul
## date     2015-12-20
```

```
## Packages
```

```
-----
## package      * version      date
## annotate     1.48.0        2015-10-14
## AnnotationDbi 1.32.2        2015-12-09
## base64        1.1           2011-12-03
## beanplot      1.2           2014-09-19
## Biobase        * 2.30.0        2015-10-14
## BiocGenerics  * 0.16.1        2015-11-06
## BiocParallel  1.4.3         2015-12-18
## biomaRt        2.26.1        2015-11-23
## Biostrings    * 2.38.2        2015-11-21
## bitops         1.0-6         2013-08-17
## bumpHunter    * 1.10.0        2015-10-14
## codetools      0.2-14        2015-07-15
## colorspace     1.2-6         2015-03-11
## corpcor        1.6.8         2015-07-08
## DBI            0.3.1         2014-09-24
## devtools      * 1.9.1         2015-09-11
```

## digest	0.6.8	2014-12-31
## doParallel	* 1.0.10	2015-10-14
## doRNG	1.6	2014-03-07
## ellipse	0.3-8	2013-04-13
## evaluate	0.8	2015-09-18
## foreach	* 1.4.3	2015-10-13
## formatR	1.2.1	2015-09-18
## futile.logger	1.4.1	2015-04-20
## futile.options	1.0.0	2010-04-06
## genefilter	1.52.0	2015-10-14
## GenomeInfoDb	* 1.6.1	2015-11-03
## GenomicAlignments	1.6.1	2015-10-22
## GenomicFeatures	1.22.7	2015-12-18
## GenomicRanges	* 1.22.2	2015-12-12
## GEOquery	* 2.36.0	2015-10-14
## ggplot2	2.0.0	2015-12-18
## gtable	0.1.2	2012-12-05
## htmltools	0.2.6	2014-09-08
## igraph	1.0.1	2015-06-26
## IlluminaHumanMethylation450kanno.ilmn12.hg19	* 0.2.1	2015-12-20
## IlluminaHumanMethylation450kmanifest	* 0.4.0	2015-10-20
## illuminaio	0.12.0	2015-10-14
## IRanges	* 2.4.6	2015-12-12
## iterators	* 1.0.8	2015-10-13
## knitr	1.11	2015-08-14
## lambda.r	1.1.7	2015-03-20
## lattice	* 0.20-33	2015-07-14
## limma	3.26.3	2015-11-16
## locfit	* 1.5-9.1	2013-04-20
## magrittr	1.5	2014-11-22
## MASS	7.3-45	2015-11-10
## matrixStats	0.50.1	2015-12-15
## mclust	5.1	2015-10-27
## memoise	0.2.1	2014-04-22
## minfi	* 1.16.0	2015-10-14
## mixOmics	5.2.0	2015-11-19
## multtest	2.26.0	2015-10-14
## munsell	0.4.2	2013-07-11
## nlme	3.1-122	2015-08-19
## nor1mix	1.2-1	2015-07-27
## pkgmaker	0.22	2014-05-14
## plyr	1.8.3	2015-06-12
## preprocessCore	1.32.0	2015-10-14
## quadprog	1.5-5	2013-04-17
## RColorBrewer	1.1-2	2014-12-07
## Rcpp	0.12.2	2015-11-15
## RCurl	1.95-4.7	2015-06-30
## registry	0.3	2015-07-08
## reshape	0.8.5	2014-04-23
## rgl	0.95.1429	2015-12-07

## rmarkdown	0.8.1	2015-10-10
## rngtools	1.2.4	2014-03-06
## Rsamtools	1.22.0	2015-10-14
## RSQLite	1.0.0	2014-10-25
## rtracklayer	1.30.1	2015-10-22
## S4Vectors	* 0.8.5	2015-12-11
## scales	0.3.0	2015-08-25
## siggenes	1.44.0	2015-10-14
## stringi	1.0-1	2015-10-22
## stringr	1.0.0	2015-04-30
## SummarizedExperiment	* 1.0.1	2015-11-06
## survival	2.38-3	2015-07-02
## XML	3.98-1.3	2015-06-30
## xtable	1.8-0	2015-11-02
## XVector	* 0.10.0	2015-10-14
## yaml	2.1.13	2014-06-12
## zlibbioc	1.16.0	2015-10-14
## source		
## Bioconductor		
## Bioconductor		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## Bioconductor		
## Bioconductor		
## Bioconductor		
## Bioconductor		
## Bioconductor		
## CRAN (R 3.2.2)		
## Bioconductor		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.0)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## CRAN (R 3.2.2)		
## Bioconductor		
## Bioconductor		
## Bioconductor		
## Bioconductor		
## Bioconductor		
## CRAN (R 3.2.3)		

```
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## local
## Bioconductor
## Bioconductor
## Bioconductor
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## Bioconductor
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.3)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## Bioconductor
## CRAN (R 3.2.2)
## Bioconductor
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## Bioconductor
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## Bioconductor
## CRAN (R 3.2.2)
## Bioconductor
## Bioconductor
## CRAN (R 3.2.2)
## Bioconductor
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## Bioconductor
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## CRAN (R 3.2.2)
## Bioconductor
```

```
## CRAN (R 3.2.2)
## Bioconductor
```