DNA Methylation Practical 3 - MWAS

Allan McRae
a.mcrae@uq.edu.au

Sonia Shah
sonia.shah@uq.edu.au

Overview

This practical will cover

- Performing a Methylome-Wide Association Study (MWAS)

- Visualising results

- Advanced correction for batch effects

Read in Data

Load libraries, read in data, and remind yourself of its structure

```
library(meffil)
library(qqman)

load("norm.beta.Robj")

ls()
dim(norm.beta)
norm.beta[1:10,1:10]
```

Read in Data

Read sample sheet and cell counts. Remember some samples were removed – these need excluded from the sample sheet

```
load("samplesheet.Robj")
m = match(colnames(norm.beta), samplesheet$Sample_Name)
samplesheet = samplesheet[m,]

dim(samplesheet)
dim(norm.beta)


cc = read.table("cellcounts.txt", head=T)
cc = as.matrix(cc[,-1])
```

Basic MWAS

Perform a linear regression testing the effect of sex (M/F) on DNA methylation. Print a status update every 100 probes.

```
n = nrow(norm.beta)
sto = matrix(NA, nrow = n, ncol = 4)


for(i in 1:n) {
  if(i%%100 == 0) print(i)
  l = lm(norm.beta[i,] ~ samplesheet$Sex +
                         samplesheet$Slide +
                         samplesheet$sentrix_col +
                         cc)
  sto[i,] = summary(l)$coef[2,]
}
```

Visualise the Results – QQ

Create a QQ plot and calculate the genomic inflation factor

```
png("ewas-qq.png")
qq(sto[,4])
dev.off()


lambda = qchisq(1-median(sto[,4]),1)/qchisq(0.5,1)
lambda
```

Visualise the Results – Manhattan

Generate a Manhattan Plot

```
anno = meffil.get.features("450k")
anno = anno[match(row.names(norm.beta), anno$name),]

P = sto[,4]
CHR = substr(anno$chromosome,4,5)
CHR[which(CHR == "X")] = 23
CHR[which(CHR == "Y")] = 24
CHR = as.numeric(CHR)
BP = anno$position
x = data.frame(P,CHR,BP)

png("manhattan.png", width=960)
manhattan(x)
dev.off()
```

MWAS with SVA correction

Prepare phenotype and covariate data.frames for Meffil

```
phen = data.frame(samplesheet$Sex)
row.names(phen) = samplesheet$Sample_Name



covs = data.frame(samplesheet$Slide,
                  samplesheet$sentrix_col,
                  cc)
row.names(covs) = samplesheet$Sample_Name
```

MWAS with SVA correction

Run the MWAS, using most variable 20,000 probes to construct SVA

```
ewas.ret <- meffil.ewas(norm.beta,
                        variable=phen[,1],
                        covariates=covs,
                        winsorize.pct = NA,
                        most.variable = 20000)
```

Summarise Results

Prepare a summary of the MWAS results

```
ewas.parameters =
      meffil.ewas.parameters(
                sig.threshold=1e-7,
                max.plots=100,
                qq.inflation.method="regression",
                model="sva")


ewas.summary = meffil.ewas.summary(
                        ewas.ret,
                        norm.beta,
                        parameters=ewas.parameters)
```

Summarise Results

Write a summary file and save results for later analysis

```
meffil.ewas.report(ewas.summary,
                      output.file="sex.ewas.report.html")

save(ewas.ret, file="ewas.ret.Robj")
```