**Packages:**

library(XLConnect)

library(limma)

library(Biobase)

library(meffil)

library(FDb.InfiniumMethylation.hg19)

hm450.hg19 <- getPlatform(platform='HM450', genome='hg19')

**Normalised methylation and phenotype data**:

load("methdata.RData")

pheno<-read.csv('pheno.csv')

cellcounts<-data.frame(counts)

cellcounts$id<-rownames(cellcounts)

pheno<-pheno[match(cellcounts$id, pheno$id),]

pheno<-pheno[match(colnames(norm.beta), pheno$id),]

counts<-cellcounts[match(pheno$id, cellcounts$id),]

meffil.list.featuresets()

annotate<-meffil.get.features("450k")

**Removing snps within CpG probes**

load("featsAll\_SNPs\_Xr\_dens.RData")

featsAll<-subset(featsAll, snp10=='TRUE'|Xreactive=='TRUE')

norm.beta<-subset(norm.beta, !(featsAll$IlmnID %in% rownames(norm.beta)))

**Analysis:**

variable<-pheno[,c(33)] # diagnosis IBD vs non-IBD

cov<-pheno[,c(8,9)] # covariates age and gender

covariates<-cbind(cov, counts[,c(1:6)]) # combining cell counts to covariates

**Ulcerative colitis**

UC<-subset(pheno, !(dx=='CD'|dx=='IBD\_U'))

UCcov<-covariates[match(UC$id, rownames(covariates)),]

betauc<-norm.beta[,match(UC$id, colnames(norm.beta))]

variableUC<-UC[,c(33)]

ewasUC<- meffil.ewas(betauc, variable=variableUC, covariates=UCcov)

tophitsUC<-data.frame(ewasUC$analyses$all$table)

probe\_id <- rownames(tophitsUC)

probes <- hm450.hg19[probe\_id]

gs <- getNearestGene(probes)

gs <- gs[,3:4]

chr <- probes@seqnames

coord <- probes@elementMetadata@listData$probeTarget

UCtop<- cbind(probe\_id,chr,coord,gs, tophitsUC)

**Sensitivity analysis UC:**

UC1<-UC[!(UC$smoking=='Unknown'),]

rownames(UC1)<-UC1$id

counts1<-counts[match(UC1$id, counts$X),]

UC.beta<-norm.beta[,match(UC1$id, colnames(norm.beta))]

variable<-UC1[,c(33)]

cov<-UC1[,c(8,9)] # age and gender

covariates<-cbind(cov, counts1[,c(1:6)])

ewas.UC <- meffil.ewas(UC.beta, variable=variable, covariates=covariates)

tophitsUC<-data.frame(ewas.UC$analyses$all$table)

toptableUC<-subset(tophitsUC, p.value<0.05)

x<-annotate[match(rownames(toptableUC), annotate$name),]

toptableUC<-cbind(x,toptableUC)

probe\_id <- rownames(toptableUC)

probes <- hm450.hg19[probe\_id]

gs <- getNearestGene(probes)

gs <- gs[,3:4]

chr <- probes@seqnames

coord <- probes@elementMetadata@listData$probeTarget

UChits<- cbind(probe\_id,chr,coord,gs, toptableUC)

**Crohn’s disease**

CD<-subset(pheno, !(dx=='UC'|dx=='IBD\_U'))

CDcov<-covariates[match(CD$id, rownames(covariates)),]

betacd<-norm.beta[,match(CD$id, colnames(norm.beta))]

variableCD<-CD[,c(33)]

ewasCD<- meffil.ewas(betacd, variable=variableCD, covariates=CDcov)

tophitsCD<-data.frame(ewasCD$analyses$all$table)

probe\_id <- rownames(tophitsCD)

probes <- hm450.hg19[probe\_id]

gs <- getNearestGene(probes)

gs <- gs[,3:4]

chr <- probes@seqnames

coord <- probes@elementMetadata@listData$probeTarget

TopTableCD<- cbind(probe\_id,chr,coord,gs, tophitsCD)

**Sensitivity analysis CD:**

CD1<-CD[!(CD$smoking=='Unknown'),]

counts1<-counts[match(CD1$id, cellcounts$X),]

CD.beta<-norm.beta[,match(CD1$id, colnames(norm.beta))]

variable<-CD1[,c(33)]

cov<-CD1[,c(8,9)]

covariates<-cbind(cov, counts1[,c(1:6)])

ewasCD <- meffil.ewas(CD.beta, variable=variable, covariates=covariates)

tophitsCD<-data.frame(ewasCD$analyses$all$table)

toptableCD<-subset(tophitsCD, p.value<0.05)

x<-annotate[match(rownames(toptableCD), annotate$name),]

toptableCD<-cbind(x,toptableCD)

probe\_id <- rownames(toptableCD)

probes <- hm450.hg19[probe\_id]

gs <- getNearestGene(probes)

gs <- gs[,3:4]

chr <- probes@seqnames

coord <- probes@elementMetadata@listData$probeTarget

CDhits<- cbind(probe\_id,chr,coord,gs, toptableCD)