```
pvalue_scz <- SCZ[SCZ$snpid %in% snp_result,]$pval
BIP_jiaoji <- BIP[BIP$snpid %in% snp_result,]
SCZ_jiaoji <- SCZ[SCZ$snpid %in% snp_result,]
BIP_SCZ <- inner_join(BIP[,c('snpid','pval')],SCZ[,c('snpid','pval')],by="snpid")
```

```
# The parameter estimates are obtained using the EM algorithm
theta_est <- EM1(pvalue_bip, pvalue_scz,pi00_ini = 0.8, pi01_ini=0.01,pi10_ini=0.01,pi11_ini=0.1,alpha1
theta_est
```

```
## $pi00
## [1] 0.8173783
##
## $pi01
## [1] 0.006304867
##
## $pi10
## [1] 0.001133071
##
## $pi11
## [1] 0.1751837
##
## $alpha1
## [1] 0.5985608
##
## $alpha2
## [1] 0.5039367
```

```
gamma_post <- theta_est$pi00+theta_est$pi01*theta_est$alpha2*pvalue_scz^(theta_est$alpha2-1)+theta_est$a
```

```
z11_post <- theta_est$pi11*theta_est$alpha1*theta_est$alpha2*pvalue_bip^(theta_est$alpha1-1)*pvalue_scz
z01_post <- theta_est$pi01*theta_est$alpha2*pvalue_scz^(theta_est$alpha2-1)/gamma_post
z10_post <- theta_est$alpha1*pvalue_bip^(theta_est$alpha1-1)*theta_est$pi10/gamma_post
z00_post <- theta_est$pi00/gamma_post
z1_post <- z11_post+z10_post
z2_post <- z11_post+z01_post
```

```
# SNPS associated with both diseases
Z_est_11 <- assoc(z11_post, 0.1)
sum(Z_est_11)
```

```
## [1] 6956
```

```
# SNPS associated with BIP
Z_est_10 <- assoc(z1_post, 0.1)
sum(Z_est_10)
```

```
## [1] 7018
```

```
# SNPS associated with SCZ
Z_est_01 <- assoc(z2_post, 0.1)
sum(Z_est_01)
```

```
## [1] 8210
```

The coincidence of the ids of the three classes of SNPS was judged

```
snpid_bip <- BIP[BIP$snpid %in% snp_result,]$snpid
snp_bip = data.frame(snpid_bip = snpid_bip,bip_11 = Z_est_11,bip_10 = Z_est_10)
snpid_bip_11 <- snp_bip[which(snp_bip$bip_11 == 1),'snpid_bip']
snpid_bip_10 <- snp_bip[which(snp_bip$bip_10 == 1),'snpid_bip']
```

```
length(intersect(snpid_bip_11,snpid_bip_10))
```

```
## [1] 6956
```

```
# The high degree of overlap between those associated with both diseases and those associated with BIP
```

```
setdiff(snpid_bip_10,snpid_bip_11)
```

```
##  [1] "rs4908147"  "rs387176"   "rs263908"   "rs10921107" "rs2270543"
##  [6] "rs17017120" "rs3845684"  "rs906867"   "rs13386455" "rs1526645"
## [11] "rs7583278"  "rs4430889"  "rs1435845"  "rs1161474"  "rs6729882"
## [16] "rs7641289"  "rs17718783" "rs9288851"  "rs16825602" "rs810471"
## [21] "rs6829771"  "rs2646269"  "rs370831"   "rs391529"   "rs13128519"
## [26] "rs7733288"  "rs256014"   "rs9393672"  "rs9275523"  "rs7767277"
## [31] "rs2842625"  "rs2138707"  "rs1934124"  "rs4357169"  "rs9387090"
## [36] "rs13227417" "rs11761050" "rs6958911"  "rs2196"     "rs12342040"
## [41] "rs10760846" "rs7895653"  "rs3858136"  "rs10791838" "rs11174379"
## [46] "rs7300029"  "rs11174386" "rs11109376" "rs1262775"  "rs17107699"
## [51] "rs10518779" "rs16957422" "rs7190307"  "rs12943984" "rs7209032"
## [56] "rs156430"   "rs4277599"  "rs2378249"  "rs6030341"  "rs11906231"
## [61] "rs2836779"  "rs926761"
```

```
# The only SNP associated with BIP but not with SCZ
```

```
snpid_scz <- SCZ[SCZ$snpid %in% snp_result,]$snpid
snp_scz = data.frame(snpid_scz = snpid_scz,scz_11 = Z_est_11,scz_10 = Z_est_01)
snpid_scz_11 <- snp_scz[which(snp_scz$scz_11 == 1),'snpid_scz']
snpid_scz_01 <- snp_scz[which(snp_scz$scz_10 == 1),'snpid_scz']
```

```
length(intersect(snpid_scz_11,snpid_scz_01))
```

```
## [1] 6956
```

Only the bipolar GWAS data was used for analysis

```
## Define the function that calculates the log-likelihood
# input: p, pi1, alpha
log_P_single <- function(P, pi1, alpha){
  return(sum(log(pi1*dbeta(P, alpha, 1) + (1 - pi1))))#
```

```r
}
## Define functions that implement the EM algorithm for the analysis of a single dataset
# input: P, pi1_ini, alpha_ini, max_iter, tol
# ouput: pi1_est, alpha_est
EM_single <- function(P, pi1_ini, alpha_ini, max_iter = 1e4, tol = 1e-6){
 L_ini <- log_P_single(P, pi1_ini, alpha_ini)
 for (iter in 1:max_iter){
   if (iter == 1){
     pi1_old <- pi1_ini
     alpha_old <- alpha_ini
     L_old <- L_ini
   }

   # E step
   comp_gamma <- pi1_old*dbeta(P, alpha_old, 1)
   gamma <- comp_gamma/(comp_gamma + 1 - pi1_old)

   # M step
   pi1_new <- mean(gamma)
   alpha_new <- -sum(gamma)/sum(gamma*log(P))
   L_new <- log_P_single(P, pi1_new, alpha_new)
   # whether the algorithm converges
   if (L_new < L_old){
     print("Error: log likelihoood is not increasing!")
     break
   }
   if ((L_new - L_old)/abs(L_new) < tol){
     break
   } else {
     pi1_old <- pi1_new
     alpha_old <- alpha_new
     L_old <- L_new
   }
 }

 return(list(pi1 = pi1_new, alpha = alpha_new))
}
```

```r
theta_est_BIP <- EM_single(pvalue_bip, pi1_ini = 0.1, alpha_ini = 0.2)
theta_est_BIP
```

```
## $pi1
## [1] 0.2194204
##
## $alpha
## [1] 0.6628928
```

```r
posterior_bip <- (theta_est_BIP$pi1*theta_est_BIP$alpha*pvalue_bip^(theta_est_BIP$alpha - 1))/(theta_est
Z_est_bip <- assoc(posterior_bip, 0.1)
sum(Z_est_bip)
```

```
## [1] 403
```

```r
snp_bip_single = data.frame(snpid_bip = snpid_bip,bip_1 = Z_est_bip)
snpid_bip_1 <- snp_bip_single[which(snp_bip_single$bip_1 == 1),'snpid_bip']
```

```r
length(intersect(snpid_bip_1,snpid_bip_10))
```

```
## [1] 387
```

Only the schizophrenia GWAS data was used for analysis

```r
theta_est_SCZ <- EM_single(pvalue_scz, pi1_ini = 0.1, alpha_ini = 0.2)
theta_est_SCZ
```

```
## $pi1
## [1] 0.2419776
##
## $alpha
## [1] 0.5668053
```

```r
posterior_scz <- (theta_est_SCZ$pi1*theta_est_SCZ$alpha*pvalue_scz^(theta_est_SCZ$alpha - 1))/(theta_es
Z_est_scz <- assoc(posterior_scz, 0.1)
sum(Z_est_scz)
```

```
## [1] 4012
```

```r
snp_scz_single = data.frame(snpid_scz = snpid_scz,scz_1 = Z_est_scz)
snpid_scz_1 <- snp_scz_single[which(snp_scz_single$scz_1 == 1),'snpid_scz']
```

```r
length(intersect(snpid_scz_1,snpid_scz_01))
```

```
## [1] 3692
```

Plot the Manhattan

```r
FDR <- function(posterior, alpha){
  M          <- length(posterior)
  fdr        <- 1 - posterior
  rank.fdr   <- rank(fdr)
  sort.fdr   <- sort(fdr)
  cumsum.fdr <- cumsum(sort.fdr)
  sort.FDR   <- cumsum.fdr/seq(1, M, 1)
  FDR        <- sort.FDR[rank.fdr]

  return(FDR)
}
```

BIP