**XPEB**

**R package**

**Manual version 1.0**

November 21st 2014

Authors: Marc Coram, Sophie Candille, and Hua Tang

Contacts: [huatang@stanford.edu](mailto:huatang@stanford.edu) and [scand@stanford.edu](mailto:scand@stanford.edu)

**Description**

This R package implements the cross-population empirical Bayes method, XPEB, described in Coram et al. 2015. XPEB takes as input P-value summary statistics from two GWAS, a target-GWAS from for example an ethnic minority population of primary interest and an auxiliary base-GWAS such as a larger GWAS in Europeans, and reprioritizes SNPs in the target population to compute local false discovery rates. XPEB also estimates the degree of overlap in the genetic architecture underlying the trait in the two populations.

**Software requisites**

It is assumed that R 2.15 or later has been installed on your machine. R is a free software, and can be downloaded from the R website (<http://www.r-project.org/>) for Linux, Mac OS X, and Windows machines.

**Download**

Download the tar.gz source file from

http://med.stanford.edu/tanglab/software/

Downloading the source file is not necessary when installing from the CRAN repository.

**Installation**

There are three ways to install the XPEB package.

1. At the command prompt, using the local tar.gz

close R

cd to the directory that contains the tar.gz source file

type at the command prompt:

> R CMD INSTALL XPEB\_0.1.tar.gz

2. From the R console, using the local tar.gz

set the working directory to that of the tar.gz source file

type at the prompt in the R console

>install.packages("XPEB\_0.1.tar.gz",repos=NULL,type="source")

3. From the R console, using CRAN: \*\*\*not available yet\*\*\*

type in the R console

>install.packages("XPEB")

**Input files requirements**

Two input text files are required, one for the target-GWAS and one for the base-GWAS, that contain the P-values for the target and base-GWAS. These files are white space delimited, and have a header.

The target-GWAS file includes four columns with names: SNP, CHR, BP, and P.

The base-GWAS file includes two columns named: SNP and P.

These files contain The GWAS results for each marker, one marker per line. Missing data are coded as NA.

Example: target-GWAS

SNP CHR BP P

rs1110052 1 863421 0.4288

rs3748595 1 877423 0.7718

...

Example: base-GWAS

SNP P

rs1110052 0.02246

rs3748595 0.6035

...

**Usage Example**

#In the R console, load the XPEB package:

>library(XPEB)

#Unzip the example files from the package and retrieve their path

#target-GWAS

>path.target <- system.file("extdata", "target.gwas.txt.zip", package="XPEB")

>unzip(path.target,exdir="TMPinput")

>path.target <- "TMPinput/target.gwas.txt"

#base-GWAS

>path.base <- system.file("extdata", "base.gwas.txt.zip", package="XPEB")

>unzip(path.base,exdir="TMPinput")

>path.base <- "TMPinput/base.gwas.txt"

#Run XPEB on the example files

>res <- run.xpeb(path.target=path.target,path.base=path.base,n.target=1e4,n.base=1e5)

#Save the locfdr calculation results to a text file

>write.table(res$locfdr, file="locfdrResults.txt", sep="\t", quote=F, row.names=F)

#For help on the run.xpeb() function:

>?run.xpeb

**Options for run.xpeb()**

path.target and path.base:

string indicating the path to the GWAS result files in the target and base populations.

gc.target and gc.base:

T or F to indicate whether a genomic control correction should be applied or not. Default is to apply the genomic control correction.

n.target and n.base:

median sample size in target and base.

n.iter:

number of iterations for the MCMC. Default is the recommended 1e6.

**Ouput**

The function run.xpeb() outputs a list with 2 elements. res$overlap is the estimated overlap and res$locfdr is a data frame with the XPEB locfdr calculation.

For example:

SNP CHR BP LOCFDR

rs13302982 1 851671 0.2628274

rs4040604 1 852987 0.2296354

**Citation**

Marc A Coram, Sophie I Candille, Qing Duan, Katie Chan, Yun Li, Charles Kooperberg, Alex P Reiner, Hua Tang. An Ethnicity-Aware Approach for Mapping Complex Traits Using Multi-ethnic Data. American Journal of Human Genetics (submitted)

The manuscript presents the description of the XPEB method and testing on simulated and real datasets. Please cite our manuscript if you use our software.