

Manual

DeMix : decovolution for mixed cancer genomes using raw measured data.

#Updated date : December 14, 2013
#Version : v1.0.1
#Jaeil Ahn, Wenyi Wang : email to (wwang7@mdanderson.org)

1 How to use :

To use demix, you need to move all files after unzipping Demix.zip to the designated folder where the dataset locates. Then, you need to have 'R' on your machine at version 3.0.0 or above. When you open R, follow the instruction below. Please note that there are two versions of DeMix: one for Mac and other for linux (for PC users). When you download the DeMix software, you must download appropriate version depending on what kind of operation system you have.

- a. First load a function by using 'source("TumorHv1.R");'
- b. The function name is 'Rdemix'. To use this function, you need to input parameters listed below
 - i Inputmat1 : Matrix of gene expressions after the RMA normalization. The $G \times S$ matrix where G is the number of genes and S is the number of samples.
 - ii Groupid: Vector of indicators. Element of 0 for normal and 1 for mixed samples. For 6 sample setup, we might have c(0,1,0,1,0,1).
 - iii Cit : Cit=64 tells the machine is built based on 64 bit. It can be 32 bit for old machines.
 - iv Matched: For matched samples, we need to specify matched=1, otherwise matched=0. For the matched case, the Groupid needs to be specified in the form of (0,1,0,1,...,0,1), i.e., the tumor samples need to be aligned next to the corresponding normal sample.
 - v Nref : If you have reference genes, set Nref=1 otherwise Nref=0
 - vi RefMat: If Nref=1, we need to define a matrix with reference gene expressions with variance (See example).
 - vii Nstart : Number of initial values for tumor proportion π . The running time is proportional to 'nstart'.
 - viii Nhuman : If the data are collected from human, specify Nhuman=1, otherwise put 0.

2 Return values

DeMix will return two types of outcomes.

- a. `phis` : Mixture proportions of the cell type that is denoted as 1.
- b. `Decov` : Deconvolved expressions after log-transformation corresponding to each individual sample.

3 Example when there is no reference

Please see the following example.

```
rm(list=ls(all=TRUE))
# Load maqc3 data
rnadata <- as.matrix(read.table("maqcdata3.txt", header=FALSE))
# Load R function
source("TumorHv1.R")
# Define group: 0 for normal 1 for mixed
cnvgroup <- c(rep(0,5), rep(1,10))
# Define other parameter
cbit=64
matched=0
nstart=2
ref=0
refmatrix=matrix(0, nrow(rnadata), 2)
nhuman=1
testr <- Rdemix(rnadata, cnvgroup, cbit, matched, ref, refmatrix,nstart, nhuman)
# Estimated proportions
> testr$phi
[1] 0.3577272 0.2827235 0.3477742 0.2712235 0.2666289 0.7731194 0.7769084
[8] 0.7373019 0.7488339 0.8050418
# Estimated deconvolved expressions (log2-transformed)
> testr$decov[1:5,1:5]
      [,1]      [,2]      [,3]      [,4]      [,5]
[1,] 9.852360 9.729444 9.798046 9.734874 9.789861
[2,] 9.204185 9.144845 9.155575 9.136025 9.156796
[3,] 9.966397 10.024052 10.007584 9.993843 10.050947
[4,] 9.061439 9.092339 9.041230 9.002670 9.039755
[5,] 9.123904 9.104556 9.151646 9.120139 9.113896
```

4 Example when there is reference

4.1 Example 2

```
rm(list=ls(all=TRUE))
```

```

# Example maqc1 data
rnadata<-as.matrix(read.table("Rmaqcddata1.txt", header=FALSE))
rnadata2<-as.matrix(read.table("Rmaqcddata1ori.txt", header=FALSE))

nabsv<-(abs(apply(rnadata[,1:5], 1, mean)-apply(rnadata2[,1:5],1,mean))>3 & abs(2^apply(rnadata[,1:5], 1,
rnadata=rnadata[nabsv,]
rnadataT<-rnadata2[nabsv,]
refdata2<-cbind((apply(rnadataT, 1, mean)), (apply(rnadataT, 1,
sd))^2);
source("TumorHv1.R");

# define group 0 for normal 1 for mixed
cnvgroup<-c(rep(0, 5), rep(1, 10)); # Maqc
cbit=64;
matched=0;
nstart=2
inputmat1<-rnadata
testr<-Rdemix(inputmat1, cnvgroup, cbit, matched, 1, refdata2, nstart, 1)

# Estimated proportions
> testr$phi
[1] 0.11095734 0.08985242 0.13702647 0.12959593 0.13757911 0.58713162
[7] 0.59679857 0.58952894 0.58964421 0.61047011

# Estimated deconvolved expressions
> testr$decov[1:5,1:5]
[,1] [,2] [,3] [,4] [,5]
[1,] 9.283528 9.326418 9.329793 9.352642 9.283690
[2,] 10.398591 10.362051 10.408206 10.385461 10.286755
[3,] 9.142444 9.256425 9.251844 9.245567 9.155768
[4,] 9.329873 9.291780 9.356418 9.335080 9.336665
[5,] 9.253079 9.316123 9.288251 9.227155 9.267332

```