deconvSeq: Deconvolution of cell mixture distribution in sequencing data

Rose Du, Vince Carey, Scott T. Weiss

2018-05-22

Introduction

`deconvSeq` is an R package for performing cell type deconvolution of whole tissue using RNA sequencing or bisulfite sequencing data. Required input includes the sequencing data for the individual cell types and the whole tissue of interest.

For this vignette, we will use the data included with the package for RNA sequencing (data_celltypes_rnaseq and data_tissue_rnaseq) and reduced representation bisulfite sequencing (data_celltypes_rrbs and data_tissue_rrbs).

```
library(deconvSeq)
```

Example using RNA sequencing data

First load the example data for individual cell types (T cells, B cells, monocytes, granulocytes) and create `DGEList` object using `EdgeR`.

```
data("data_celltypes_rnaseq")
```

Then compute projection matrix, `b0`.

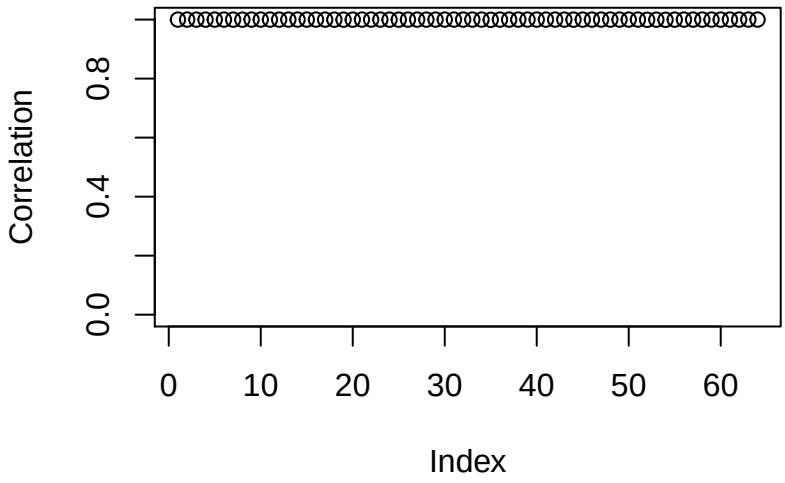
```
set.seed(1234)
b0 = getb0.rnaseq(dge.celltypes, design.rnaseq, ncpm.min=1, nsamp.min=4)
```

We will use the top 50 genes for the signature gene set. First, we obtain the predicted cell type mixture, `resultx1`, using the deconvolution method.

```
nb0 = 50
resultx1 = getx1.rnaseq(nb0,b0,dge.celltypes)
```

Then we compare our predicted results with the actual cell types, `x2`.

```
x2 = as.data.frame(design.rnaseq,row.names=sample.id.rnaseq)
cr = getcorr(resultx1$x1,x2)
plot(cr, ylim=c(0,1), ylab="Correlation")
```



We can now load the data for the whole tissue samples.

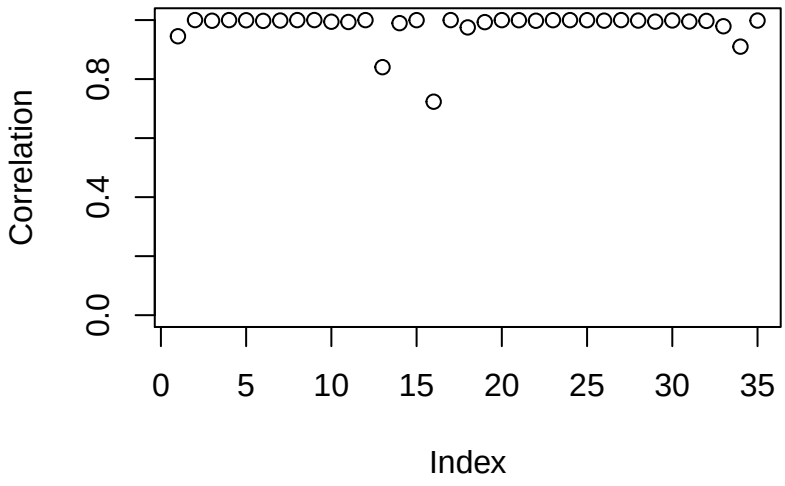
```
data("data_tissue_rnaseq")
dge.tissue = getdge(cnts.tissue,design=NULL, ncpm.min=1, nsamp.min=4)
```

Using the 50 signature genes from the projection matrix, `b0`, we compute the predicted cell type mixture for the whole tissue sample, `resultx1.tissue`.

```
nb0=50
resultx1.tissue = getx1.rnaseq(nb0,b0, dge.tissue)
```

We can then compare the predicted cell type mixture against the known cell type mixture of the tissue sample. In this case, the known cell types are obtained from a blood cell count differential where lymphocytes are a combination of T cells and B cells. The correlation per sample is obtained with `getcorr`.

```
x1 = cbind(lymph=resultx1.tissue$x1$tissuefBcell+resultx1.tissue$x1$tissuefTcell,
 mono = resultx1.tissue$x1$tissuefMonocytes, gran = resultx1.tissue$x1$tissuefGranulocytes)
x2 = as.matrix(cbc.rnaseq/100)
cr = getcorr(x1,x2)
plot(cr, ylim=c(0,1), ylab="Correlation")
```



Example using bisulfite sequencing data

First load the example data for individual cell types (T cells, B cells, monocytes, granulocytes).

```
data("data_celltypes_rrbs")
```

Then compute projection matrix, `b0`.

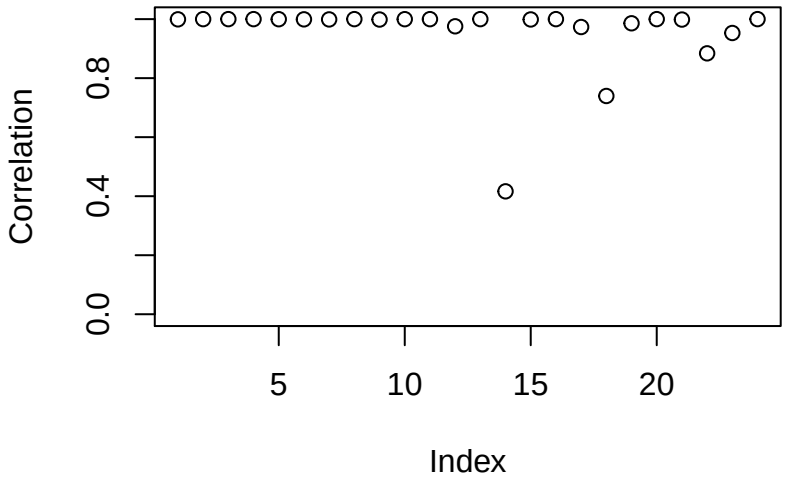
```
set.seed(1234)
b0 = getb0.biseq(methmat, design.rrbs)
```

We will use the top 250 CpG sites for the signature CpG set. First, we obtain the predicted cell type mixture, `resultx1`, using the deconvolution method.

```
nb0=250
resultx1 = getx1.biseq(nb0,b0,methmat,sample.id.rrbs,celltypes.rrbs)
```

Then we compare our predicted results with the actual cell types, `x2`.

```
x2 = as.data.frame(design.rrbs,row.names=sample.id.rrbs)
cr=getcorr(resultx1$x1,x2)
plot(cr, ylim=c(0,1), ylab="Correlation")
```



We can now load the data for the whole tissue samples.

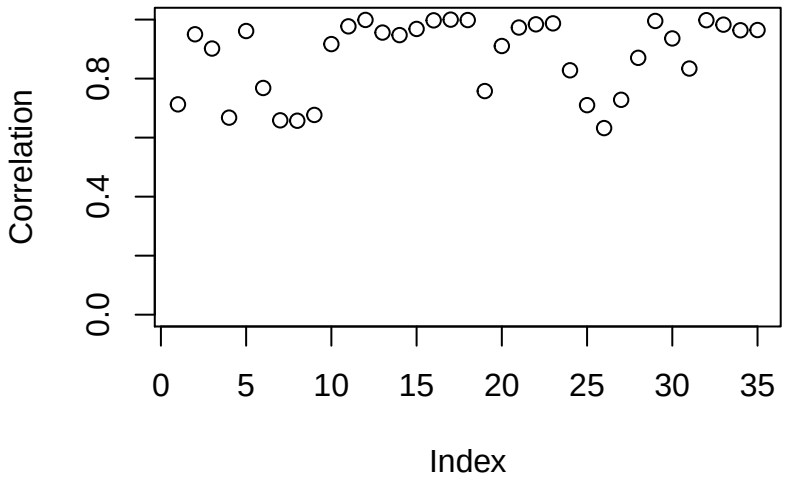
```
data("data_tissue_rrbs")
```

Using the 250 signature CpGs from the projection matrix, `b0`, we compute the predicted cell type mixture for the whole tissue sample, `resultx1.tissue`.

```
nb0=250
resultx1.tissue = getx1.biseq(nb0,b0,methmat.tissue,sample.id.tissue,celltypes.rrbs)
```

We can then compare the predicted cell type mixture against the known cell type mixture of the tissue sample. In this case, the known cell types are obtained from a blood cell count differential where lymphocytes are a combination of T cells and B cells. The correlation per sample is obtained with `getcorr`.

```
x1 = cbind(lymph=resultx1.tissue$x1[,1]+resultx1.tissue$x1[,2], mono = resultx1.tissue$x1[,3], gran = re
x2 = as.matrix(cbc.rrbs/100)
cr = getcorr(x1,x2)
plot(cr, ylim=c(0,1), ylab="Correlation")
```


Correlation
0.0
0.4
0.8
0
5
10
15
20
25
30
35
Index