How To Use SubpathwayLNCE

Xinrui Shi, Chunquan Li, Xia Li

January 19, 2016

Contents

1 Overview 1

2 get candidate lncRNA-mRNA interaction 1

3 Reconstruct KEGG signal pathways 2

  3.1 Get the co-express lncRNA-mRNA interactions 2

  3.2 Embed competitively regulated lncRNAs to undirect KEGG signal pathway graphs 2

4 Locate KEGG metabolic subpathways 3

5 Identify the significantly enriched subpathways 3

6 Session Info 7

1 Overview

This vignette demonstrates how to easily use the SubpathwayLNCE package. This package can implement the identification of Kyoto Encyclopedia of Genes and Genomes (KEGG) signal subpathways competitively regulated by long non-coding RNAs (lncRNAs), by topologically locating lncRNAs and genes within reconstructed KEGG signal pathway graphs, which embedded by lncRNAs based on ceRNA theory. (1) This package provides the `getExampleData` to return example data and environment variables. (see the section 2). (2) This package provides the `getInterUMGraph` function to reconstruct KEGG signal pathways by embedding lncRNAs into undirect KEGG signal pathway graphs.(see the section 3). (3) This package provides the `getLocSubGraph` function to locate lncRNAs competitively regulated signal subpathways by topologically analyzing the “lenient distance” of lncRNAs and genes, based on reconstructed pathways.(see the section 4). (4) This package provides the `identifyGraphW` function to identify the significantly enriched signal subpathways, based on located subpathways.(see the section 5). (5) This package provides the `GetK2riData` function to get variable data in current environment.(see the section ??). (6) This package provides the `updateOrganEnvir` function to update the organism-specific environment variables.(see the section ??).

2 get candidate lncRNA-mRNA interaction

We can use function `getExampleData` to return example data and environment variables, such as the data of candidate lncRNA-mRNA interaction, the data of undirect KEGG metabolic pathway graphs with genes as nodes.
> # obtain the data for candidate lncRNA-mRNA interaction.
> interaction<-GetExampleData(exampleData="pp")
> # view first six rows of data
> interaction[1:6,]

<table>
<thead>
<tr>
<th>LncEns GeneEns</th>
<th>GeneEns</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSG00000189149 ENSG00000142192</td>
<td></td>
</tr>
<tr>
<td>ENSG00000214870 ENSG00000142192</td>
<td></td>
</tr>
<tr>
<td>ENSG00000237697 ENSG00000142192</td>
<td></td>
</tr>
<tr>
<td>ENSG00000212694 ENSG00000142192</td>
<td></td>
</tr>
<tr>
<td>ENSG00000244625 ENSG00000142192</td>
<td></td>
</tr>
<tr>
<td>ENSG00000247796 ENSG00000142192</td>
<td></td>
</tr>
</tbody>
</table>

> # obtain the data for undirect KEGG metabolic pathway graphs with genes as nodes
> g2<-GetExampleData(exampleData="g2")
> # obtain example data of matched mRNA-lncRNA expression profiles
> #GeneExp<-GetExampleData(exampleData="GeneExp")
> #LncExp<-GetExampleData(exampleData="LncExp")

3 Reconstruct KEGG signal pathways

We can use function `getInteUMGraph` to return the integrated KEGG signal pathway graph list. We first convert KEGG metabolic pathways to direct/undirect graphs with genes as nodes, then reconstructed pathways by linking lncRNAs to competitively regulated targets within it.

3.1 Get the co-express lncRNA-mRNA interactions

The function `getLncGenePairs` can calculated co-expression coefficient for any pair of relations in the candidate LncRNA-mRNA interaction based on matched LncRNA and mRNA expression profiles, those relations had reached a significant positive threshold were retained.

> # obtain example data of matched mRNA-lncRNA expression profiles
> GeneExp<-GetExampleData(exampleData="GeneExp")
> LncExp<-GetExampleData(exampleData="LncExp")
> # calculated co-expression coefficient, the significant positive threshold is 0.025
> LncGenePairs<-getLncGenePairs(GeneExp,LncExp,a=0.025)
> # obtain the data for undirect KEGG metabolic pathway graphs with genes as nodes
> g2<-GetExampleData(exampleData="g2")
> # get reconstructed undirect pathway graph list
> # interUMGraph<-getInteGraphList(g2,LncGenePairs)

3.2 Embed competitively regulated lncRNAs to undirect KEGG signal pathway graphs

The function `getInteGraphList` can competitively regulated lncRNAs into undirect KEGG signal pathway graphs with genes as nodes. With integrated graph list, we can offer the additional interested genes sets to identify the condition-specific pathways competitively regulated by lncRNAs.

> # obtain the data for undirect KEGG metabolic pathway graphs with genes as nodes
> g2<-GetExampleData(exampleData="g2")
> # obtain example data of matched mRNA-lncRNA expression profiles
> #GeneExp<-GetExampleData(exampleData="GeneExp")
> #LncExp<-GetExampleData(exampleData="LncExp")
> #calculated co-expression coefficient, the significant positive threshold is 0.025
> #LncGenePairs<-getLncGenePairs(GeneExp,LncExp,a=0.025)
> # get reconstructed undirect pathway graph list
> # To improve efficiency, a fraction of signal pathway as case
> LncGenePairs<-GetExampleData(exampleData="LncGenePairs")
> interUMGraph<-getInterGraphList(g2[42:45],LncGenePairs)
> ### Integrate lncRNAs of competitive regulation into KEGG pathway graphs ###
> ##LncGenePairs<-GetExampleData(exampleData="LncGenePairs")
> ##interUMGraph<-getInteUMGraph(LncGenePairs)
> # To improve efficiency, a fraction of signal pathway as case
> LncGenePairs<-GetExampleData(exampleData="LncGenePairs")
> interUMGraph<-getInterGraphList(g2[42:45],LncGenePairs)
> ### Integrate lncRNAs of competitive regulation into KEGG pathway graphs ###
> #LncGenePairs<-GetExampleData(exampleData="LncGenePairs")
> #inteUMGraph<-getInteUMGraph(LncGenePairs)

The following commands can show the reconstructed pathway graph with genes and lncRNAs as nodes.

> # visualize the reconstructed undirect pathway
> #LncGenePairs<-GetExampleData(exampleData="LncGenePairs")
> #inteUMGraph<-getInteUMGraph(LncGenePairs)
> plotGraphL(interUMGraph[[1]],vertex.label=getNodeLabel)

Figure 1 shows the reconstructed undirect p53 signaling pathway.

4 Locate KEGG metabolic subpathways

We can use function getLocSubGraph to locate signal subpathways by topologically analyzing the "lenient distance" of lncRNAs and/or genes based on reconstructed pathways.

> ### Integrate lncRNAs of competitive regulation into KEGG pathway graphs ###
> #LncGenePairs<-GetExampleData(exampleData="LncGenePairs")
> #inteUMGraph<-getInteUMGraph(LncGenePairs)
> ### get user-interested lncRNAs and genes sets.
> ##geneLnc<-c(getBackground(type="gene")[1:3000],unique(LncGenePairs[1,]))
> geneLnc<-GetExampleData(exampleData="geneLnc")
> # get locate subpathways.
> sub<-getLocSubGraphLnc(geneLnc,interUMGraph,type="gene_lncRNA",n=1,s=8)

5 Identify the significantly enriched subpathways

We can use function identifyGraphW to identify the significantly enriched subpathways based on located direct/undirect signal subpathways.

> ### Integrate lncRNAs of competitive regulation into KEGG pathway graphs ###
> #LncGenePairs<-GetExampleData(exampleData="LncGenePairs")
> #inteUMGraph<-getInteUMGraph(LncGenePairs)
> ### get user-interested lncRNAs and genes sets.
> ##geneLnc<-c(getBackground(type="gene")[1:3000],unique(LncGenePairs[1,]))
> geneLnc<-GetExampleData(exampleData="geneLnc")
Figure 1: The visualization of reconstructed undirect p53 signaling pathway.
The following commands can show the reconstructed pathway graph with genes and miRNAs as nodes.

```r
> plotAnnGraph("path:04916_1", sub, SubcodeLncResult, gotoKEGG=FALSE, vertex.label=getNodeLabel)
```

**Figure 2** shows the reconstructed undirect Calcium signaling pathway.
Figure 2: The visualization of reconstructed undirect Calcium signaling pathway.
6 Session Info

The script runs within the following session:

R version 3.1.2 (2014-10-31)
Platform: i386-w64-mingw32/i386 (32-bit)

locale:
[1] LC_COLLATE=C
[4] LC_NUMERIC=C

attached base packages:
[1] stats graphics grDevices utils datasets methods base

other attached packages:
[1] SubpathwayLNCE_1.0 BiasedUrn_1.0 BiasedUrn_1.07 RBGL_1.46.0 graph_1.44.1
[5] igraph_0.7.1

loaded via a namespace (and not attached):
[1] BiocGenerics_0.12.1 parallel_3.1.2 stats4_3.1.2
tools_3.1.2

References


