**ceRNA分析结果**

**=====================================文件=====================================**

**|-- \*ceRNA\_pair\_adjust\_p\_Sig.xls ceRNA关系对**

**注：若样品数目不少于5，会将共表达的结果考虑到ceRNA中，此文件为”共表达ceRNA结果文件”；否则不考虑共表达结果，此文件为”ceRNA结果文件”。**

|-- ceRNA1 : ceRNA1的ID

|-- ceRNA2 : ceRNA2的ID

|-- miRNA\_target\_num1 : ceRNA1所结合的miRNA数目

|-- miRNA\_target\_num2 : ceRNA2所结合的miRNA数目

|-- overlap\_miRNA\_num : ceRNA1和ceRNA2共同结合的miRNA数目

|-- miRNA\_target1\_uniq : ceRNA1所结合的特有miRNA

|-- miRNA\_target2\_uniq : ceRNA2 所结合的特有miRNA

|-- overlap\_miRNA : ceRNA1和ceRNA2共同结合的miRNA

|-- pvalue : 超几何检验p值

|-- FDR : 错误发现率FDR值

|-- coexp\_cor : ceRNA关系对之间的皮尔森相关系数（样品数目小于4不存在）

|-- coexp\_pvalue : ceRNA关系对之间的皮尔森相关性的显著性p值（样品数目小于4不存在）

这里，ceRNA1、ceRNA2可由lncRNA、circRNA、mRNA充当。

**====================================文件夹=====================================**

**|-- \*\_vs\_\* 差异ceRNA关系及关键基因通路整合分析结果**

注：从ceRNA关系对中提取每组差异RNA的一步紧邻网络，即得到差异ceRNA关系对（即sub\_\*\_vs\_\*\_network.xls文件）。

采用随机游走中的经典算法PageRank来获取网络中所有节点（即差异ceRNA）的打分（即重要性）。对打分进行排序，选择前（百分比）%作为关键节点作为重点研究对象。

对关键节点中的基因+lncRNA的靶基因+circRNA的Host gene(以下简称为”关键基因”，即\*\_vs\_\*\_Key\_RNA\_gene.xls)进行通路富集分析，选择最显著富集的前5个通路。提取这5个通路中基因与基因之间的关系整合成通路网络，并将关键基因map到通路。

在这里，节点可以看做是ceRNA(gene(基因)/lncRNA/circRNA)的ID。

**|-- BMK\_1\_KEGG 关键基因KEGG通路富集分析**

**|-- \*\_KEGG\_pathway\_enrich\_barplot.pdf 关键基因KEGG富集条形图（PDF格式）**

**|-- \*\_KEGG\_pathway\_enrich\_barplot.png 关键基因KEGG富集条形图（PNG格式）**

横坐标代表关键基因注释在通路中的基因数，纵坐标代表通路，柱的颜色代表校正后的p值。

**|-- \*\_KEGG\_pathway\_enrich\_dotplot.pdf 关键基因KEGG富集气泡图（PDF格式）**

**|-- \*\_KEGG\_pathway\_enrich\_dotplot.png 关键基因KEGG富集气泡图（PNG格式）**

**|-- \*\_KEGG\_pathway\_enrich.list.xls 关键基因KEGG富集结果文件**

|--ID ko\_id；

|--Description ko 名称；

|--GeneRatio 富集到此通路上的差异gene的数目/所有差异gene的数目；

|--BgRatio 富集到此通路上的所有基因的数目/所有基因的数目；

|--enrich\_factor 富集因子；

|--pvalue Pvalue；

|--qvalue Qvalue；

|--geneID 富集到此通路上的关键基因的ID。

**|-- sub\_\*\_vs\_\*\_network.xls 差异ceRNA关系列表**

**|-- \*\_vs\_\*\_Key\_RNA\_gene.xls 关键基因列表**

|-- Gene “关键基因”ID（基因ID+lncRNA靶基因+circRNA Host gene）

|-- type 关键RNA类型

|-- Key\_id 关键RNA ID

**|-- \*\_vs\_\*\_edge.color.xls**

**|-- \*\_vs\_\*\_network.pdf 网络图（PDF格式）**

**|-- \*\_vs\_\*\_network.png 网络图（PNG格式）**

每个圆点代表一个基因，每个长方形代表一个通路，线代表通路中基因与基因或其他通路之间的关系。不同线的颜色代表关系来自不同的通路。红色的点为关键 基因。本报告中所有网络图均采用R语言的iGraph包通过脚本进行绘制，同时也会提供绘制网络图所需文件，研究者可采用Cytoscape绘制自己所感兴趣的样式。

**|-- \*\_vs\_\*\_network.xls iGraph作图输入文件（最富集的前５个通路中基因与基因、基因与其他通路关系对）**

|-- path：最富集的前5个通路的ID

|-- entry1：通路中的基因ID（Symbol）

|-- entry2：基因/其他通路

|-- type

|-- subtype：根据subtype定义边的类型，粗细以及箭头方向,以及边上的字

|-- lty

|-- width

|-- arrow

|-- lab

**|-- \*\_vs\_\*\_node.xls iGraph作图输入文件（节点(edge)属性）**

|-- Gene：基因或者通路ID

|-- shape：节点形状属性（基因：circle；通路：rectangle）

|-- color：节点颜色（关键基因：red；通路：blue；非关键基因：grey）

#Subtype lty width arrow lab

compound 1 2 0

hidden|dissociation|missing 1 1 2

activation|expression 2 2 2

inhibition|repression 5 2 2

indirect effect 3 2 2

indirect effect,activation 4 2 2

state change 3 2 0

binding/association 2 2 0

phosphorylation 1 2 2 +p

dephosphorylation 1 2 2 -p

glycosylation 1 2 2 +g

数字代表的含义：

数字:线的类型|箭头类型

0：no edges|no arrows(-)

1：soild lines|backward arrows(<-)

2：dashed|forward arrows(->)

3：dotted|both(<->)

4：dotdash

5：longdash

6：twodash