

关键基因通路整合分析

采用随机游走中的经典算法 PageRank 来获取网络中所有节点的打分（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(以下简称为“**关键基因**”)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。

=====文件=====

|-- Net_score.xls 网络中所有节点的打分（即重要性）

第一列是 circRNA/lncRNA/gene 的 ID,第二列是对应的 Symbol,第三列是打分值（重要性）。

|-- Key_RNA_top（百分比）.xls 关键节点（前（百分比）%的节点）

第一列是 circRNA/lncRNA/gene 的 ID,第二列是对应的 Symbol,第三列是打分值（重要性）。

|-- Key_gene.xls **关键基因**

第一列是关键基因的 ID,第二列是对应的 Symbol,第三列是打分值（重要性）。

=====文件夹=====

|-- KEGG 对**关键基因**进行 KEGG 通路分析 (ClusterProfile)

|-- Key_KEGG_pathway_enrich_barplot.pdf 关键基因 KEGG 富集条形图 (PDF 格式)

|-- Key_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。

|-- Key_KEGG_pathway_enrich_dotplot.pdf 关键基因 KEGG 富集气泡图 (PDF 格式)

|-- Key_KEGG_pathway_enrich_dotplot.png 关键基因 KEGG 富集气泡图 (PNG 格式)

|-- Key_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。

|-- network.pdf KEGG 整合通路网络 (PDF 格式)

|-- network.png KEGG 整合通路网络 (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。

|-- 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 :

| -- 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 : solid lines | backward arrows(<-)

2 : dashed | forward arrows(->)

3 : dotted | both(<->)

4 : dotdash

5 : longdash

6 : twodash