

目录

基本概念	3
应用场景	4
主要结果	5
云端数据	6
参数说明	7
ID 列表	7
样式	8
方块	8
条带	9
标注	9
标题	10
图注(Legend)	10
坐标轴	11
风格	11
图片	12
结果说明	13
主要结果	13
补充结果	14
方法学	15
如何引用	16
常见问题	17

基本概念

- 富集分析：简单而言，就是取一部分有功能注释的分子与所有有功能注释的分子去比较（超几何分布检验），确定这一部分分子中都涉及了哪些功能作用。注意：单独几个分子做富集分析意义并不大。
- GO (Gene Ontology, 基因本体) 数据库：把基因的功能分成了三类：生物过程 (biological process, BP)、细胞组分 (cellular component, CC)、分子功能 (molecular function, MF)。利用 GO 数据库，可以得到目标基因在 CC, MF 和 BP 三个层面上有什么关联。
- KEGG (Kyoto Encyclopedia of Genes and Genomes) 数据库：一种通路数据库，收集了很多通路相关的数据库。通路数据库还包括 wikipathway, reactome 等。
- 超几何分布检验：超几何分布 (hypergeometric) 是统计学上一种离散概率分布。它描述了在 N 个物件中指定 M 个种类的物件，不放回的抽取 n 个，成功抽中指定类型物件的个数 (k) 的事件。
- 富集分析联合 logFC：就是在富集分析的基础上，利用提供的分子的 logFC，计算每个条目对应的 zscore，初步判断对应的条目是正调节 (zscore 为正) 还是负调节 (zscore 为负)。zscore 计算方法见下：

$$zscore = \frac{(Up - Down)}{\sqrt{Counts}}$$

- 其中，这里的 Up Down 代表对应条目分子的 logFC 为正以及为负分别对应数量，Counts 代表条目对应的分子总数（这里不是指 Z-score 标准化，是 GOplot 包所使用的概念和提供的方法）

（注意：相对于 GOKEGG 富集分析模块，这个模块只是在同样的富集方法的基础上，另外再计算了每个条目对应的 zscore 值）

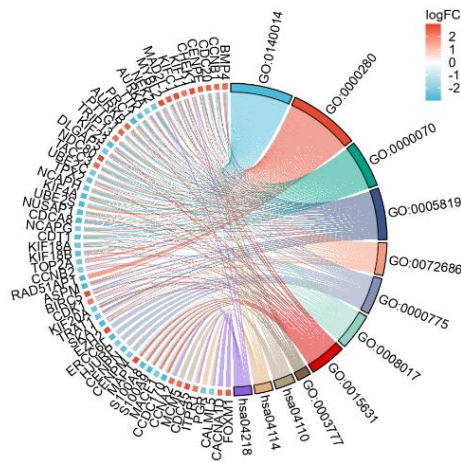
应用场景

本模块为【GO|KEGG 联合 FC】富集分析后结果的可视化展示。

注意：模块需要先进行【GO|KEGG 联合 FC】富集分析并保存结果后，此处的云端数据才会有结果记录，然后才能进行可视化的操作。主要基本参数中有 ID 列表 输入框，可以将对应云端数据记录的富集分析表格中的感兴趣的 ID 列复制到此处，进行可视化。



主要结果



通过弦图展示 **GOKEGG 联合 FC** 富集分析结果。

弦图分成左右两个部分，除了展示条目的情况外，还同时展示条目含有的分子以及分子对应的 logFC 情况：

- **左半边的部分**为基因色块，色块的不同颜色代表对应的 logFC 值，每个分子对应的色块的大小是一样的。
- **右半边的部分**为条目色块，色块的不同颜色代表对应的条目，色块的大小代表对应的 Counts（也就是本次富集分析中该条目中含有的分子数）
- 左半边和右半边色块之间的连线（弦线）代表条目中含有的分子情况，有连线就代表这个条目含有对应的分子。

默认展示各类别的 top 几个结果（默认 满足校正后 p 值<0.05）。可以挑选在满足阈值下的 top 的类目，或者一些感兴趣的类目。

云端数据

云端数据

	记录名称	来源模块	时间	补充说明
<input checked="" type="checkbox"/>	GOKEGG联合FC	GOKEGG联合FC @1.0	2023-02-13 11:22:56	数据记录可以在历史记录中找到

这里的云端数据与历史记录汇总【GOKEGG 联合 FC】富集分析模块的数据记录是保持一致的，可以在历史记录中找到相应的数据记录。

根据需要可视化的项目 选择好对应的云端数据记录。默认使用最近生成的分析记录。



参数说明

(说明：标注了颜色的为常用参数。)

ID 列表



- 标注 ID: 输入想要可视化的功能或者通路的条目 ID，默认为对应云端数据结果中每个类目的前几个条目，可以根据需要进行输入修改。注意：输入的 ID 来自所选云端数据记录的结果，需要先在历史记录中找到对应的记录，下载 excel 结果，复制想要展示的 ID 到这个输入框中，一行代表一个。最多支持 1 张图同时绘制 15 个类目。

样式



- ID 展示: ID 名称过长时, 可以根据需要选择换行模式。可选择 ID 号、全名(自动换行)、全名(一行 20 长度)、全名(一行 30 长度)、全名(一行 40 长度)、全名(一行 50 长度)、全名(一行 60 长度)、全名(一行 70 长度)、全名(一行 80 长度)、全名(不换行)。

方块



- **右侧颜色**: 右侧色块 (代表条目) 的填充色颜色选项, 有多少个功能类别会提取多少个颜色, 最多支持修改 10 个颜色。受配色方案全局性修改。
- **左侧色阶**: 左侧色块 (代表分子对应 logFC 值) 的填充色颜色选项, 对应数值型内容, 修改第一和第二色卡作为数值从小到大的**渐变色**。受配色方案全局性修改。

- 右侧描边：右侧色块描边的粗细，默认为 0.50pt。
- 左侧描边：左侧色块描边的粗细，默认为 0.00pt。

条带

条带

描边粗细 0.00pt

不透明度 0.5

- 描边粗细：弦线描边的粗细，默认为 0.00pt。
- 不透明度：弦线的透明度。0 为完全透明，1 为完全不透明。

标注

标注

右侧文字 6pt

左侧文字 6pt

- 右侧文字：右侧（条目名称）标注的字体大小
- 左侧文字：左侧（分子名称）标注的字体大小

标题

标题 ▼

大标题

大标题内容

x轴标题

x轴标题内容

y轴标题

y轴标题内容

- 大标题：大标题文本
- x 轴标题：x 轴标题文本
- y 轴标题：y 轴标题文本
- 补充：在要换行的中间插入\n。如果需要上标，可以用两个英文输入法下的大括号括住，比如 {{2}}；如果需要下标，可以用两个英文输入法下的中括号括住，比如 [[2]]。

图注(Legend)

图注 ▼

是否展示

☒

图注标题

图注标题内容

图注位置

默认 ▼

- 是否展示：是否展示图注

- 图注标题：可以添加图注标题
- 图注位置：可选择 默认、右、上、右上、右下。

坐标轴



- x 轴范围：可以控制 x 轴范围和刻度，提供 2 个值来控制范围。形如 -1, 1 (最小值和最大值不能超过 3，调整过大可能会无作用)。
- y 轴范围：可以控制 y 轴范围和刻度，提供 2 个值来控制范围。形如 -1, 1 (最小值和最大值不能超过 3，调整过大可能会无作用)。

风格



- 文字大小：针对图中图例文字的大小控制

图片

图片

▼

宽度 (cm)

8

高度 (cm)

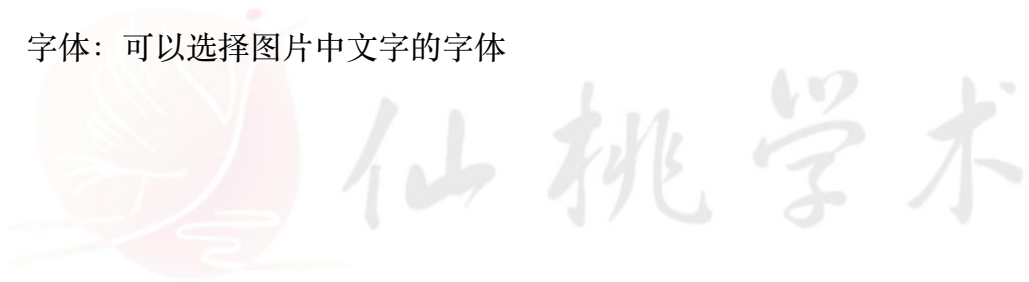
8

字体

Arial

▼

- 宽度：图片横向长度，单位为 cm
- 高度：图片纵向长度，单位为 cm
- 字体：可以选择图片中文字的字体



结果说明

主要结果



主要结果格式为图片格式，提供 PDF、TIFF 和 PPTX 格式下载，结果报告可以下载包括 pdf 以及说明文本的内容。

补充结果

可视化ID

当前模块可视化所选ID

ONTOLOGY	ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue	geneID
BP	GO:0140014	mitotic nuclear division	31/197	293/18800	2.61e-22	7.82e-19	6.71e-19	BMP4/CCNB1/CDC20/...
BP	GO:0000280	nuclear division	35/197	446/18800	8.47e-21	1.27e-17	1.09e-17	BMP4/CCNB1/CDC20/...
BP	GO:0000070	mitotic sister chromatid se...	24/197	171/18800	2.26e-20	1.83e-17	1.57e-17	CCNB1/CDC20/CENPE...
CC	GO:0005819	spindle	26/203	402/19594	9.72e-14	2.88e-11	2.53e-11	BIRC5/CCNB1/CDK1/...
CC	GO:0072686	mitotic spindle	17/203	160/19594	8.73e-13	1.29e-10	1.14e-10	CDK1/CENPE/KIF11/...
CC	GO:0000775	chromosome, centromeric r...	19/203	227/19594	2.81e-12	2.77e-10	2.44e-10	BIRC5/CCNB1/CENPE...
MF	GO:0008017	microtubule binding	19/192	272/18410	7.14e-11	3.28e-08	3e-08	BIRC5/CENPE/KIF11...
MF	GO:0015631	tubulin binding	19/192	376/18410	1.57e-08	3.6e-06	3.3e-06	BIRC5/CENPE/KIF11...
MF	GO:0003777	microtubule motor activity	8/192	67/18410	4.66e-07	6.74e-05	6.18e-05	CENPE/KIF11/KIFC1...
KEGG	hsa04110	Cell cycle	11/95	126/8164	1.99e-07	4.18e-05	3.98e-05	CCNA2/CCNB1/CDK1/...
KEGG	hsa04114	Oocyte meiosis	10/95	131/8164	2.55e-06	0.0003	0.0003	CCNB1/CDK1/CDC20/...
KEGG	hsa04218	Cellular senescence	10/95	156/8164	1.22e-05	0.0009	0.0008	CACNA1D/CCNA2/CCN...

GOKEGG可视化ID.xlsx GOKEGG可视化ID.docx

此表格提供当前可视化的 GOKEGG 联合 FC 富集分析结果, 提供 Excel、Docx 格式下载。



方法学

所有分析和可视化均在 R 4.2.1 中进行

涉及的 R 包：ggplot2 包（用于可视化）

处理过程：使用 ggplot2 包对富集分析结果进行可视化。



如何引用

生信工具分析和可视化用的是 R 语言，可以直接写自己用 R 来进行分析和可视化即可，可以无需引用仙桃，如果想要引用仙桃，可以在致谢部分 (Acknowledge) 致谢仙桃学术 (www.xiantao love)。

方法学部分可以参考对应说明文本中的内容以及一些文献中的描述。



常见问题

1. 图中标注的 ID 能否更换别的 ID? 可视化结果的 ID 从哪里获得?

答:

在“ID 列表”选项卡中，有标注 ID 的输入框:



选项框内默认选择对应云端记录结果中前几个条目，可以在此处选择想要可视化的 ID。



注意：输入的 ID 来自所选云端数据记录的结果，需要先在历史记录中找到对应的记录，下载 excel 结果，复制想要展示的 ID 到这个输入框中，一行代表一个。最多同时支持 15 个。GOKEGG 联合 FC 气泡图展示的是所有的富集结果，需要标注的 ID 通过输入框修改。

	A	B	C	D
1	ONTOLOGY	ID	Description	GeneRatio
2	BP	GO:0140014	mitotic nucle	31/197
3	BP	GO:0000280	nuclear divis	35/197
4	BP	GO:0000070	mitotic siste	24/197
5	BP	GO:0048285	organelle fiss	36/197
6	BP	GO:0000819	sister chrom	25/197
7	BP	GO:0007059	chromosome	28/197
8	BP	GO:1902850	microtubule	20/197
9	BP	GO:0098813	nuclear chro	25/197
10	BP	GO:0007052	mitotic spinc	17/197
11	BP	GO:0007051	spindle organ	19/197

ID列表
标注ID
GO:0140014
GO:0000280
GO:0000070
GO:0048285
GO:0000819
GO:0007059
GO:1902850
GO:0098813

2. 为什么出来的图中少了 KEGG (或者 BP 或者 CC 或者 MF)，明明已经选了 GO+KEGG? (为什么出来的图里面某个分类只有 1 个或者没有?)

答:

GOKEGG 可视化模块仅仅只是对已经完成 GOKEGG 富集分析的数据进行可视化，如果对应保存的数据中就不存在某些类（没有富集出来某些类），可视化是不可能会有这些类的。而在 GOKEGG 富集分析模块中，最终的结果表格只保留了满足较宽的阈值 ($p < 0.1$ 以及 $qvalue < 0.2$) 的结果，而不满足这一较宽阈值下的条目都会被过滤，如果整个类 (BP、CC、MF、KEGG) 都不满足这个阈值，那么最终的表格中就会缺少这个类。

可以先检查 GOKEGG 富集分析的结果，在历史记录中找到保存的记录：

工具首页 分析工具 历史记录 拼图工具

① 历史记录中超过30天的记录会自动清理!

批量删除 刷新

ID	名称	模块	状态	类型	时间	操作
1	GOKEGG联合FC	GOKEGG联合FC	完成	表格		GOKEGG联合logFC.xlsx GOKEGG联合logFC.docx ID转换情况.xlsx 下载整份报告

下载 | 查看

3. 能否上传自己的富集数据进行可视化?

答:

这个模块没有办法上传自己的数据，也没有对应的上传数据版本（[因为要求的数据比较特殊](#)）。

