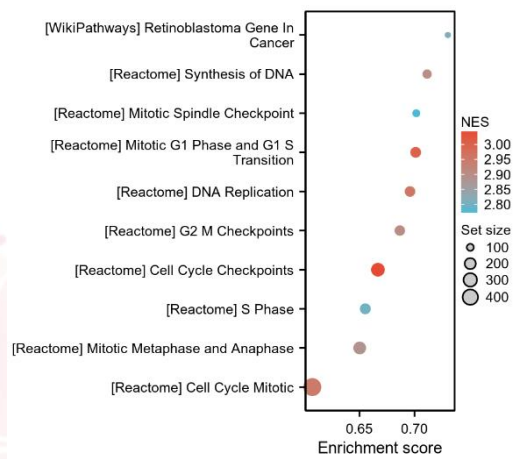


功能聚类 - GSEA 气泡图



网址: <https://www.xiantao.love>



更新时间: 2023.02.07

目录

基本概念	3
应用场景	3
主要结果	4
云端数据	5
参数说明	6
ID 列表	6
样式	7
点	8
标题	9
图注(Legend)	9
坐标轴	10
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

- 基因集富集分析 (Gene Set Enrichment Analysis, GSEA) : 用一个预先定义的基因集中的基因来评估在与表型相关度排序的基因表中的分布趋势, 从而判断其对表型的贡献。这个与表型相关度排序可以是 $\log FC$ 值。
- 数据集来自 (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) MSigDB 数据库, 如果想要了解数据集的选择以及细节, 可以到 MSigDB 数据库进一步了解。

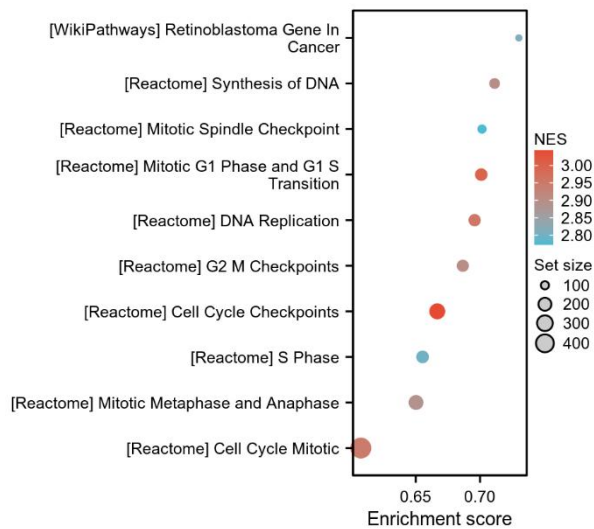
应用场景

想要知道进行了差异分析的两组别有什么功能和通路的差别, 并且手上已经有大部分的功能分子以及对应的值, 这个值可以是 $\log FC$ 。可以用这个 $\log FC$ 作为分子的排序, 从而来评估在预先定义的基因集中是否显著富集。

预先定义的基因集来自 MSigDB 数据库

(<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>), 这些预先定义的基因集中的分子基本为功能基因为主, 如果手上只有非功能基因(比如 miRNA、lncRNA、circRNA), 那么将由于缺少基因集而无法进行 GSEA 分析。

主要结果



通过气泡图展示各基因集的 GSEA 富集分析结果。

- 纵坐标为基因集名称，横坐标为所选择的 y 轴映射 内容。（如上图，对应基因集的富集评分(ES)，注意绘图的默认风格为 xy 轴颠倒）
- 图中点的颜色为所选择的 颜色映射 内容。（如上图，对应基因集的校正后归一化的富集评分(NES)）
- 图中点的大小为所选择的 大小映射 内容。（如上图，对应基因集包含的 ID 数量）

一般只要满足阈值（ $p.adjust < 0.05$ & $qvalue < 0.25$ ），就关注**基因集的名字**（最前面是对应的数据库或者分类）即可。可以挑选在满足阈值下的 NES top 的分子，或者一些感兴趣的分子。

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"/>		GSEA分析 @1.0	2023-02-02 22:26:07	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 GSEA 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。

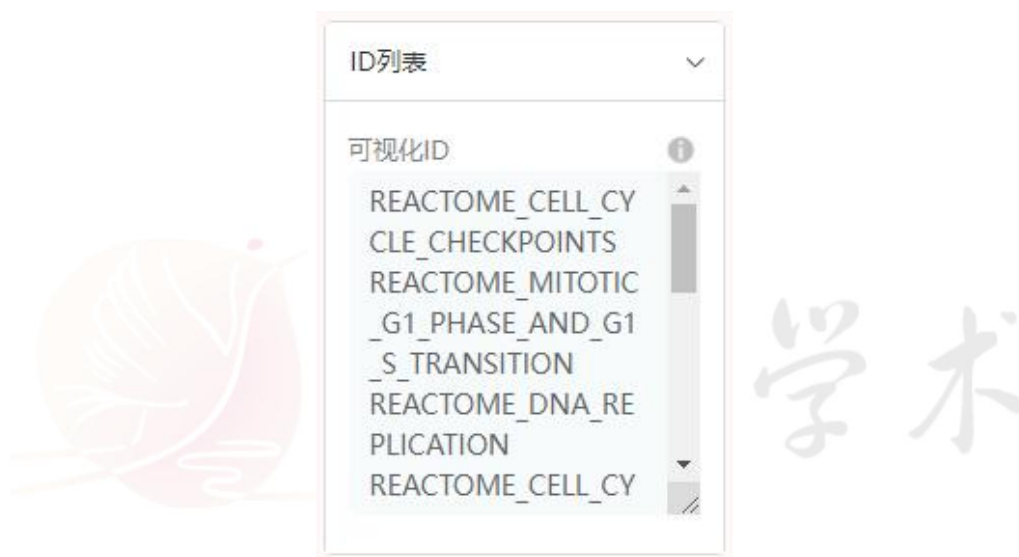
根据需要可视化的项目 选择好对应的云端数据记录。默认使用最近生成的分析记录。



参数说明

(说明：标注了颜色的为常用参数。)

ID 列表



- 可视化 ID：输入想要可视化的基因集 ID，默认为对应云端数据结果中每个类目前 10 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 20 个基因集。

样式

样式

ID换行 全名(自动换行) ▼

ID前缀是否去除 ☐

y轴映射 富集分数 ▼

颜色映射 NES ▼

大小映射 基因集ID数量 ▼

- ID 换行：ID 名称过长时，可以根据需要选择换行模式。可选择 全名(自动换行)、一行 20 长度、一行 30 长度、一行 40 长度、一行 50 长度、一行 60 长度、一行 70 长度、一行 80 长度、不换行。
- ID 前缀是否去除：默认不去除。
- y 轴映射：主要影响 y 轴（默认横坐标）的取值，具体数值可以通过**历史记录**中找到对应的记录，下载 excel 结果查看。可选择 富集分数 (enrichmentScore)、NES、p 值 (pvalue)、校正后 p 值 (p.adjust) 和 FDR (q 值, qvalue)。
- 颜色映射：主要影响点的颜色范围，可选择 富集分数 (enrichmentScore)、NES、p 值 (pvalue)、校正后 p 值 (p.adjust)、FDR (q 值, qvalue) 和 不映射。
- 大小映射：主要影响点的大小，可选择 基因集 ID 数量和不映射。

点



- **填充色**：点的填充色颜色选项，取决于 **颜色映射** 参数所选择的内容，展示数值型内容时，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。
- **描边色**：点的描边色颜色选项，取决于 **颜色映射** 参数所选择的内容，展示数值型内容时，修改第一和第二色卡作为数值从小到大的渐变色。默认黑色，**不受配色方案全局性修改**。
- **样式**：点的样式类型，可选择 圆形、正方形、菱形、三角形、倒三角。**单选，选择后全局修改**。
- **大小比例**：点的相对大小，取决于 **大小映射** 参数所选择的内容。
- **不透明度**：点的透明度。0 为完全透明，1 为完全不透明。

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（影响颜色映射的图例）

- 图注位置：可选择 默认、右、上。

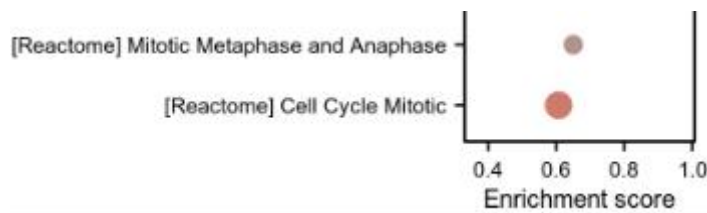
坐标轴

坐标轴

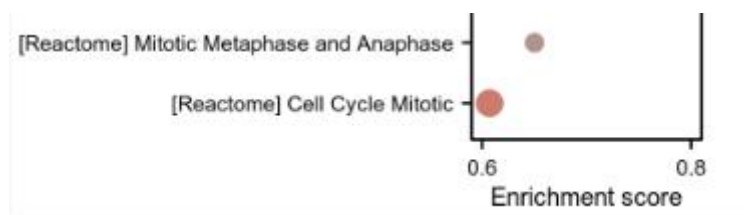
x轴标注旋
转
0

主要图对应的y
轴范围+刻度
逗号隔开

- x 轴标注旋转：支持对 x 轴文字进行旋转。适合于 x 轴文字过长的时候。（注意无论是否进行 xy 颠倒，均修改图形横坐标）
- 主要图对应的 y 轴范围+刻度：（默认对应横坐标，注意：范围的修改如果超过原本值范围的 20%会失效）
 - 如果只是想要修改范围，可以只输入两个范围值，比如 0,1



- 如果同时想要修改范围+刻度，可以输入比如：0.6,0.6,0.8,0.8。注意，此时最大和最小值会被当做范围值，不会作为刻度，如果需要刻度，需要类似于 0.8 那样同时写两次。



风格



- 外框：是否添加外框
- 网格：是否添加网格
- xy 颠倒：可以颠倒 xy 轴，[注意默认 颠倒](#)。
- 文字大小：针对图中所有文字整体的大小控制

图片

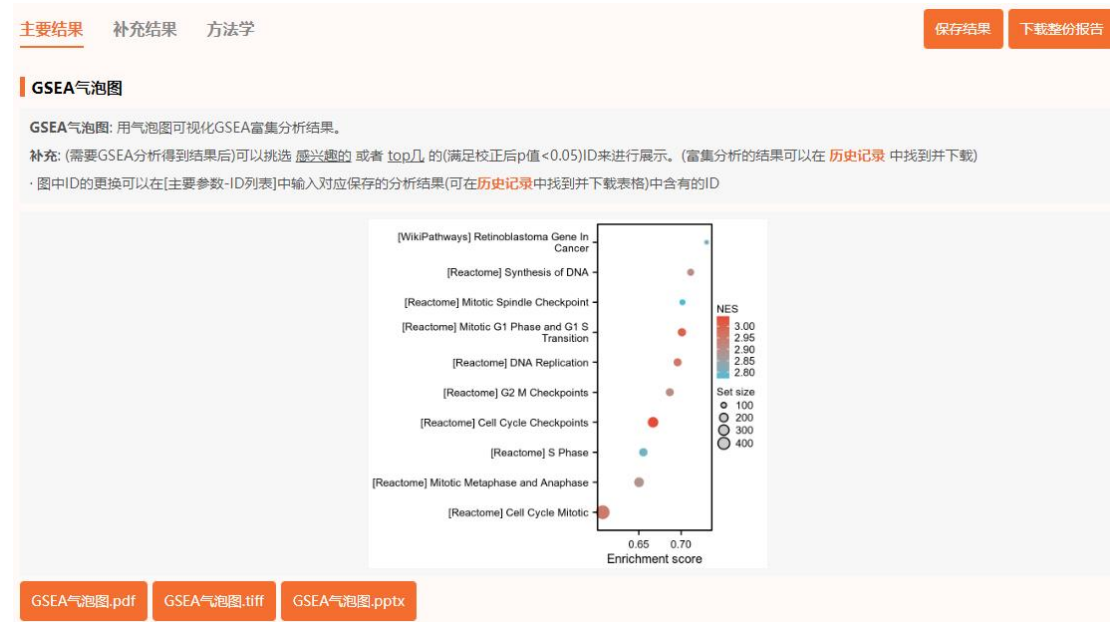
图片	▼
宽度 (cm)	8
高度 (cm)	7
字体	Arial ▼

- 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

ID	setSize	enrichmentScore	NES	pvalue	p.adjust	qvalue	rank	leading_edge
REACTOME_CELL_CYCLE_CH...	237	0.667	3.044	1e-10	7.58e-09	6.02e-09	1898	tags=43%, list=15%, signal=...
REACTOME_MITOTIC_G1_PH...	142	0.701	3.002	1e-10	7.58e-09	6.02e-09	1151	tags=41%, list=9%, signal=3...
REACTOME_DNA_REPLICATI...	137	0.696	2.956	1e-10	7.58e-09	6.02e-09	1763	tags=49%, list=14%, signal=...
REACTOME_CELL_CYCLE_MI...	458	0.607	2.947	1e-10	7.58e-09	6.02e-09	1763	tags=36%, list=14%, signal=...
REACTOME_G2_M_CHECKP...	134	0.687	2.901	1e-10	7.58e-09	6.02e-09	1898	tags=49%, list=15%, signal=...
REACTOME_SYNTHESIS_OF_...	110	0.711	2.901	1e-10	7.58e-09	6.02e-09	1898	tags=54%, list=15%, signal=...
REACTOME_MITOTIC_METAP...	201	0.650	2.887	1e-10	7.58e-09	6.02e-09	1850	tags=40%, list=15%, signal=...
WP_RETINOBLASTOMA_GE...	84	0.730	2.814	1e-10	7.58e-09	6.02e-09	1329	tags=51%, list=11%, signal=...
REACTOME_S_PHASE	145	0.655	2.799	1e-10	7.58e-09	6.02e-09	1763	tags=46%, list=14%, signal=...
REACTOME_MITOTIC_SPIND...	92	0.702	2.772	1e-10	7.58e-09	6.02e-09	449	tags=27%, list=4%, signal=2...

[GSEA可视化ID.xlsx](#)
[GSEA可视化ID.docx](#)

此表格提供当前可视化的 GSEA 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)

基因集数据库: MSigDB Collections

(<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>)

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。

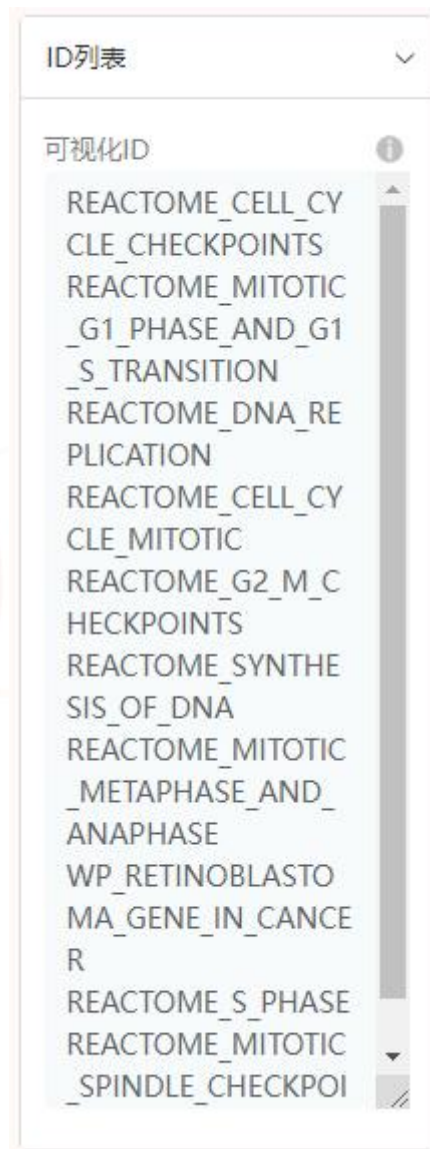


常见问题

1. 可视化结果能否更换别的 ID?

答:

在“ID 列表”选项卡中，有基因集 ID 的输入框：

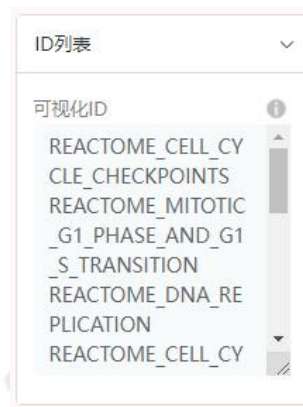


选项框内默认选择对应云端记录结果中前 10 个 ID, 可以在此处选择想要可视化的 ID。



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 20 个。

	A	B	C
1	ID	Description	setSize
2	REACTOME_CELL_CYCLE_CHECKPOINTS	REACTOME_C	237
3	REACTOME_MITOTIC_G1_PHASE_AND_G1_S_TRANSITION	REACTOME_M	142
4	REACTOME_DNA_REPLICATION	REACTOME_C	137
5	REACTOME_CELL_CYCLE_MITOTIC	REACTOME_C	458
6	REACTOME_G2_M_CHECKPOINTS	REACTOME_C	134
7	REACTOME_SYNTHESIS_OF_DNA	REACTOME_S	110
8	REACTOME_MITOTIC_METAPHASE_AND_ANAPHASE	REACTOME_M	201
9	WP_RETINOBLASTOMA_GENE_IN_CANCER	WP_RETINOB	84
10	REACTOME_S_PHASE	REACTOME_S	145
11	REACTOME_MITOTIC_SPINDLE_CHECKPOINT	REACTOME_M	92
12	REACTOME_DNA_REPLICATION_PRE_INITIATION	REACTOME_C	111



2. 要选择哪些 ID 来进行可视化？每个 ID 是什么含义？

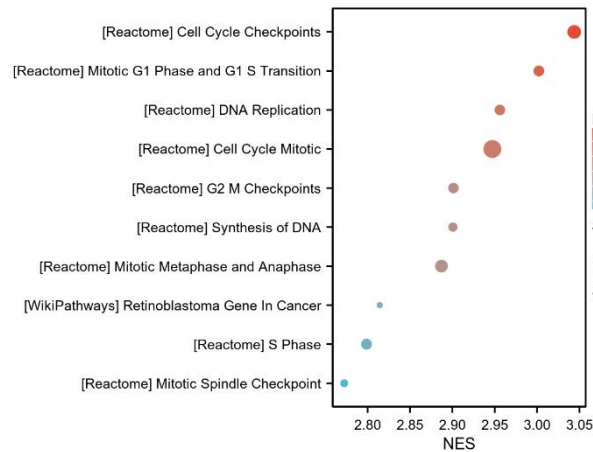
答：

在满足阈值 ($p_{adj} < 0.05$ & $qvalue < 0.25$) 下，可以是 TOP 几，也可以是自己感兴趣的想要展示的条目。具体数据集可以通过 **MSigDB 数据库** (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) 进行了解。

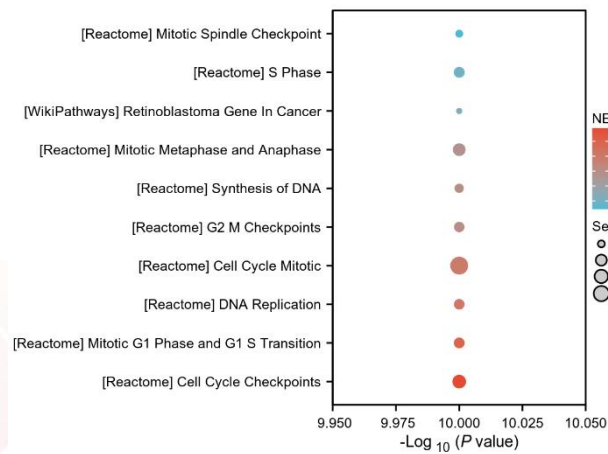
3. 如何修改展示的数据？

答：

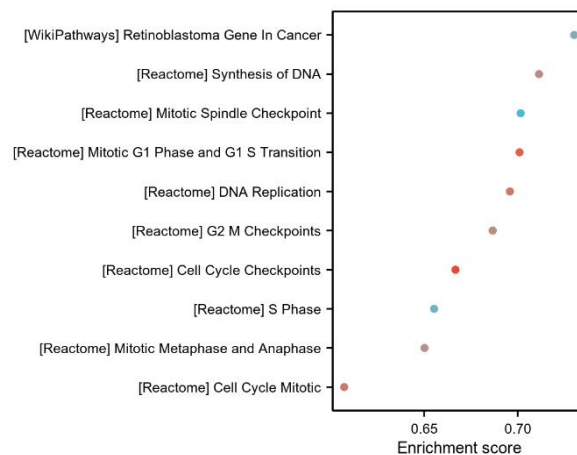
可以通过样式中的 **y 轴映射**、**颜色映射** 和 **大小映射** 参数，下拉框选择 GSEA 富集分析结果中不同的结果指标进行展示。



上图：y 轴映射 - NES，颜色映射 - NES



上图：y 轴映射 - p 值，颜色映射 - NES

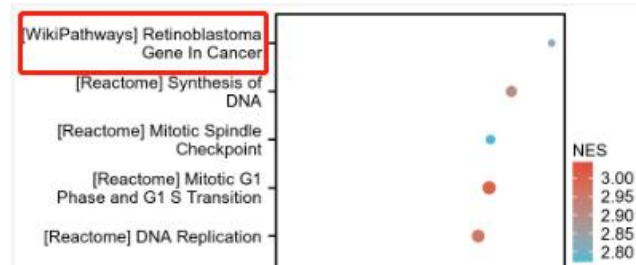


上图：大小映射 - 不映射

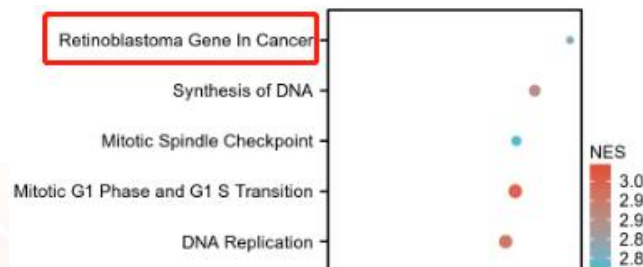
4. 基因集名称太长了，如何修改？

答：

当基因集名称太长时，可以在样式中的 ID 换行 或 ID 前缀是否去除 参数中进行换行和修改。



上图：ID 换行 - 一行 20 长度



上图：ID 前缀是否去除 - 去除