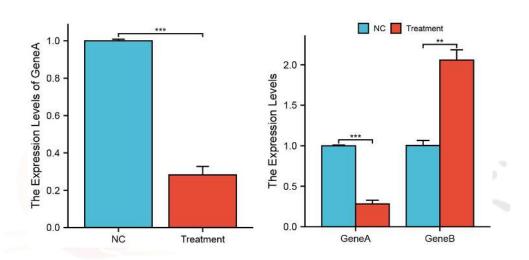


# 实验分析 - qPCR 分析



网址: <a href="https://www.xiantao.love">https://www.xiantao.love</a>



更新时间: 2023.02.02



### 目录

基本概念 3
应用场景 7
主要结果/解读 8
数据格式/数据要求 9
参数说明 12
统计分析 12
间距设置 13
点
箱/柱15
小提琴 16
误差线 17
标题文本 18
图 <mark>注(Legend)</mark>
<mark>坐标轴</mark> 19
风格
图片22
结果说明 23
主要结果 23
说明文本 24
方法学 26
如何引用 27
结果验证 27
常见问题



### 基本概念

- ▶ qPCR: 荧光定量 PCR, 在 DNA 扩增反应中,通过加入荧光化学物质,监测每次聚合酶链式反应循环后的总量,通过内参法对目标分子进行表达定量的方法。
  - 一般在一次实验中,会同时设置对照组和实验组(或者多个实验组),每个组内分别检测内参分子和目标分子(可多个目标分子),每个组会设置多个样本(生物学重复),每个样本会设置多个复孔。上机完成后,会得到每个复孔对应的 Ct 值。
  - 定量靶标分子的表达过程:
  - 1. 计算每个样本每个基因对应的 Ct 均值(每个组 每个样本 每个基因的复 孔 Ct 均值)
  - 2. (每个组 每个样本 靶标分子的 Ct 值) 减去 (同一个样本 内参分子的 Ct 值),得到 每个组 每个样本 靶标分子的 ΔCt 值
  - 3. 单独计算(上一步) 对照组 所有样本中靶标分子的ΔCt 算术均值
    - ➤ 备注: 一般是计算算术均值(求和然后求平均值)(工具计算的方法是用的这个),但是也有文章推荐用几何均数(值相乘然后开平方) https://toptipbio.com/qpcr-multiple-reference-genes
  - 4. (每个组 每个样本 靶标分子的 ΔCt 值) (第二步) 减去 (对照组 所有样本中靶标分子的ΔCt 均值) (第三步),得到 每个组 每个样本 靶标分子的 ΔΔCt 值

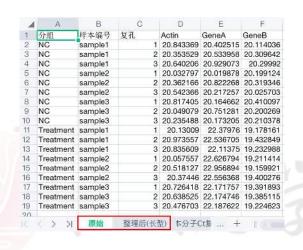


5. 由第 4 步得到的 每个组 每个样本 靶标分子的 ΔΔCt 值, 经过 2<sup>^</sup>-ΔΔCt 计算, 最终得到 靶标分子在每个组每个样本对应的相对表达量

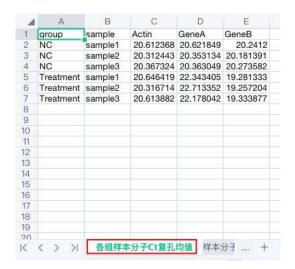
上述这些步骤,均可以在分析后的 下载区找到 qPCR 整理.xlsx 文件中得到:



#### 原始上传数据:



每个组 每个样本 各个分子对应的 Ct(复孔)均值(<mark>第一步</mark>)





每个组 每个样本 各个靶标分子的 ΔCt 值 (<mark>第二步</mark>)

(每个组 每个样本 靶标分子的 Ct 值) 减去 (同一个样本 内参分子的 Ct 值)



对照组 所有样本中各个靶标分子的ΔCt 算术均值(<mark>第三步</mark>





每个组 每个样本 各个靶标分子的 ΔΔCt 值(<mark>第四步</mark>):

(每个组 每个样本 靶标分子的 ΔCt 值) (第二步) 减去 (对照组 所有样本中靶标分子的ΔCt 均值) (第三步) 注意: 因为这一步是计算相对值(要减去对照组的均值情况),对照组的均值会校正成均值为 1。对照组默认是是上传数据第 1 列中第 1 个出现的分组。

4	A	В	C	D	
1	group	sample	GeneA	GeneB	
2	NC	sample1	-0.005818	-0.172514	
3	NC	sample2	0.0253919	0.067602	
4	NC	sample3	-0.019574	0.1049118	
5	Treatment	sample1	1.6816872	-1.166432	
6	Treatment	sample2	2.3813387	-0.860857	
7	Treatment	sample3	1.5488604	-1.081351	
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
<	< > >	各组样2	<b>本分子ΔΔCt</b>	样本分子2个	1-2

各个靶标分子 在每个组每个样本对应的相对表达量( $2^{-}\Delta \Delta Ct$ )(<mark>第五步</mark>):





# 应用场景

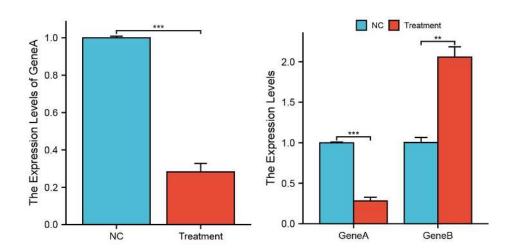
➤ qPCR 上机后得到 Ct 值,对分子进行定量分析

(如果是已经计算好每个组每个样本靶标分子的相对表达量,可直接到基础绘图的分组比较图模块 或者 豆荚图模块 进行可视化)





## 主要结果/解读

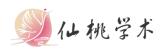


- > 每个柱子代表一个分组
- ▶ 左图为单个分子在对照组和实验组中的相对表达均值情况 (2<sup>-</sup>-ΔΔCt)
  - 可以看到 GeneA 分子在实验组中相对低表达, 差异具有统计学意义 (p < 0.001)
- ▶ 右图为两个分子分别在对照组和实验组中的相对表达均值情况(2<sup>-</sup>-ΔΔCt)
  - 可以看到 GeneA 分子在实验组中相对低表达,差异具有统计学意义(p < 0.001); GeneB 分子在实验组中相对高表达,差异具有统计学意义 (p<0.01)

还有多种可视化类型 (同分组比较图)



下载区中的 qPCR 整理表 的结果解读,可以见"基本概念"部分的内容。



## 数据格式/数据要求

4	А	В	С	D	E	F
1	分组	样本编号	复孔	Actin	GeneA	GeneB
2	NC	sample1	1	20.84336858	20.4025154	20.1140361
3	NC	sample1	2	20.35352879	20.533958	20.3096425
4	NC	sample1	3	20.64020556	20.929073	20.2999204
5	NC	sample2	1	20.03279669	20.0198775	20.1991238
6	NC	sample2	2	20.36216626	20.8222681	20.3193459
7	NC	sample2	3	20.54236647	20.217257	20.025703
8	NC	sample3	1	20.81740464	20.1646616	20.4100972
9	NC	sample3	2	20.04907874	20.7512815	20.2002694
10	NC	sample3	3	20.23548838	20.1732046	20.210378
11	Treatment	sample1	1	20.13009021	22.3797603	19.1781608
12	Treatment	sample1	2	20.97355723	22.5367053	19.4328488
13	Treatment	sample1	3	20.83560876	22.1137497	19.2329882

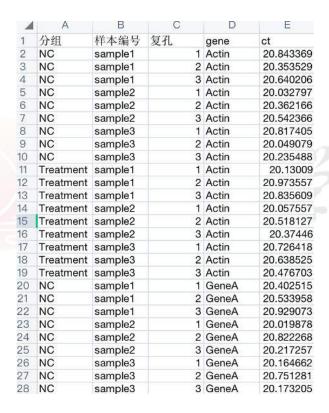
### 宽型数据类型 (推荐)

#### 一共至少需要提供5列数据:

- ▶ 第1列(命名为"分组"或者"group"): 这一列定义分组信息,每个分组有多少个样本就会对应有多少行(如果每个样本还有复孔重复,则为每组样本数×复孔重复数),至少需要提供两组(对照组和实验组),其中对照组需要放前面(第2行),实验组放后面(如果顺序错误,则有可能对照组会识别错误)。
- ▶ 第2列(命名为"样本编号"或者"sample"):这一列定义样本编号,样本编号
  不需要在对照组和实验组中都一样(对照组和实验组的样本编号一样说明是
  配对样本,该模块不会用配对的属性)
  - 如果每组样本(样本编号不满足3个以上)不满足3个以上的重复,则 结果中不会进行统计推断的内容,没有统计学检验部分的内容。只有满 足3个以上的样本,才会有统计检验的内容。



- ➤ 第3列(命名为"复孔"或者"well"): 这一列定义复孔编号,代表同一个样本对分子进行了多个复孔的检测,可提供每个复孔对应的 Ct 值,也可以手动计算复孔均值后,只提供复孔的均值,此处可以填入任意的编号
- ▶ 第4列以及后面的列(命名为对应的分子名):这几列依次填入内参分子、 靶标分子的 Ct 值,注意, 内参分子要放在第4列,后面再放靶标分子,否 则有可能会识别内参分子错误!



长型数据类型 (不推荐,有可能样本编号对应不上导致计算错误)

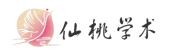
### 一共至少需要提供5列数据:

▶ 第1列(命名为"分组"或者"group"): 这一列定义分组信息,每个分组有多少个样本就会对应有多少行(如果每个样本还有复孔重复,则为每组样本数×复孔重复数),至少需要提供两组(对照组和实验组),其中对照组需要



放前面(第2行),实验组放后面(<u>如果顺序错误,则有可能对照组会识别</u> 错误)。

- ➤ 第2列(命名为"样本编号"或者"sample"):这一列定义样本编号,样本编号 不需要在对照组和实验组中都一样(对照组和实验组的样本编号一样说明是 配对样本,该模块不会用配对的属性)
- ➤ 第3列(命名为"复孔"或者"well"): 这一列定义复孔编号,代表同一个样本对分子进行了多个复孔的检测,可提供每个复孔对应的 Ct 值,也可以手动计算复孔均值后,只提供复孔的均值,此处可以填入任意的编号
- ▶ 第4列(分子名):这几列依次填入内参分子、靶标分子名字,注意,内参 分子要放在前面,后面再放靶标分子,否则有可能会识别内参分子错误!
- 》 第 5 列 (Ct 值): 每个组 每个样本 每个复孔 每个分子 对应的 Ct 值



## 参数说明

(说明:标注了颜色的为常用参数。)

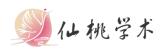
### 统计分析





- 统计方法: 统计方法默认为 auto (自动选择),当上传数据验证成功并点击确认后,会自动替换成适合于上传数据的统计方法,之后可以自行选择和修改别的统计方法。统计方法的选择依据可以参考"基本概念"中统计方法的说明。
- ▶ 分组比较: 统计学差异标注的分组,默认为 all(全部都标注)。当上传数据验证成功并点击确认后,会自动替换成对应上传数据的分组。之后可以自行选择想要保留和去掉的比较。(如果分组不满足>3 个观测以及标准差>0 的情况,则可能不会出现在此处。)允许都去掉不标注分组比较的内容。
- ▶ <mark>显著性显示的类型</mark>:可选择星号或者 p 值以及其他,影响分组比较中显著性标注,默认为星号。可以根据需要进行修改。

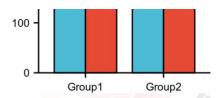
▶ 显著性大小:可以修改显著性标注的大小

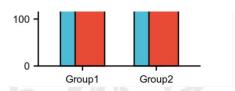


# 间距设置



▶ 组内总宽度: 只有在二维分组比较的时候才会有作用,用于控制整个组的大小,影响大组和大组的间距







点



- ▶ 显示: 可选择是否展示。
- ▶ 填充色:点的填充色颜色选项,有多少个分组会提取多少个颜色,最多支持修改8个颜色。受配色方案全局性修改
- 描边色:点的描边色颜色选项,有多少个分组会提取多少个颜色,最多支持修改8个颜色。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。可以多选,多选后不同的分组中点的类型也会有不同。
- ▶ 大小:点的大小。
- ▶ 透明度:点的透明度。0为完全透明,1为完全不透明。
- ▶ 分布宽度:图中的点会在一个水平线上随机分布,此处影响点能随机水平移动的范围。



## 箱/柱



▶ 显示:可选是否展示。

类型:可选择柱状图或者箱式图

▶ 填充色: 箱/柱的填充色颜色选项,有多少个分组会提取多少个颜色,最多支持修改8个颜色。受配色方案全局性修改

描边色: 箱/柱的的描边色颜色选项,有多少个分组会提取多少个颜色,最多支持修改8个颜色。受配色方案全局性修改。

▶ 描边粗细:箱/柱描边的粗细,默认为 0.75pt

▶ 透明度: 箱/柱的透明度。0 为完全透明, 1 为完全不透明。

▶ 宽度:箱/柱的宽度。



### 小提琴



- ▶ 显示:可选是否展示。影响图中可视化的类型。
- 填充色:小提琴的填充色颜色选项,有多少个分组会提取多少个颜色,最多支持修改8个颜色。受配色方案全局性修改
- ▶ 描边色: 小提琴的的描边色颜色选项,有多少个分组会提取多少个颜色,最多支持修改8个颜色。受配色方案全局性修改。
- ▶ 描边粗细: 小提琴描边的粗细, 默认为 0.75pt
- ▶ 透明度:小提琴的透明度。0为完全透明,1为完全不透明。
- ▶ 宽度:小提琴的宽度。
- ▶ 宽度校正:用于提高小提琴中较窄位置的宽度和整体宽度



## 误差线



误差线只有在没有箱式图时才会显示(箱式图本身自带类似误差线)。

▶ 展示: 可选是否展示。

▶ 样式:可选上、上下。

》 类型: 可选均值±标准差、均值±标准误、中位数~上下四分位,建议选择均值±标准差。

▶ 颜色:误差线颜色,默认为纯黑,不受配色方案全局性影响。

▶ 粗细:误差线粗细,默认为 0.75pt

▶ 宽度:误差线的宽度。



## 标题文本



> 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)

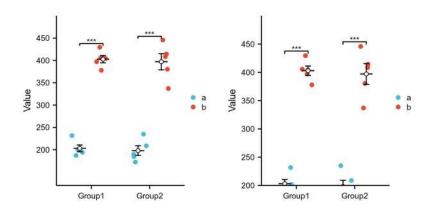
图注		~
是否展示		
图注标题	图注标题内容	
图注位置	默认	~

▶ 展示: 是否展示图注

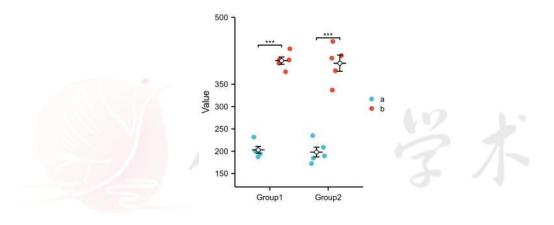
▶ 图注位置:可选右、上,默认为右。

▶ 图注标题:可以添加图注标题

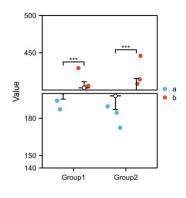
- ➤ X 轴分组名: 支持直接修改 x 轴各个分组的名字,每个名字之间需要用英文输入法的逗号隔开,比如 group1, group2。这里支持换行,需要换行的位置可以插入\n
- ➤ X 轴标注旋转: 支持对 x 轴文字进行旋转。适合于 x 轴文字过长的时候
- ➤ Y轴范围+刻度:(注意:范围的修改如果超过原本值范围的 30%会失效)
  - 如果只是想要修改范围,可以只输入两个范围值,比如 120,500



■ 如果同时想要修改范围+刻度,可以输入比如: 120, 150, 200, 250, 300, 350, 500, 500。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 500 那样同时写两次



■ 如果需要实现截断的功能,需要用小括号分别括住对应的两个坐标(最多支持同时分割成2个),具体写法类似:(140,140,150,180,200)(400,450,500,450,500)





# 风格

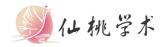


▶ 外框:是否添加外框

▶ 网格:是否添加网格

▶ 是否颠倒 XY 轴: 可以颠倒 xy 轴

文字大小: 针对图中所有文字整体的大小控制



# 图片



▶ 宽度: 图片横向长度,单位为 cm

▶ 高度:图片纵向长度,单位为 cm

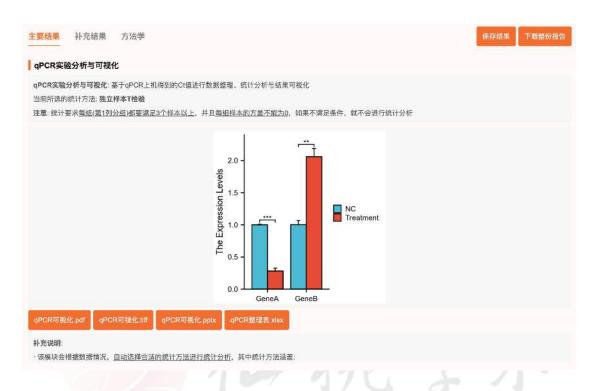
▶ 字体:可以选择图片中文字的字体





### 结果说明

### 主要结果



- 如果数据可以进行统计分析,将会进行统计分析。统计分析默认是根据数据情况选择合适的统计方法,涵盖:
  - 两组: T test(满足正态+方差齐) | Welch t' test(满足正态+不满足方差齐性) | Wilcoxon rank sum test(不满足正态, 非参数检验)
  - 多组整体: One-way ANOVA(满足正态+方差齐) | Welch one-way ANOVA(满足正态+不满足方差齐性) | Kruskal-Wallis test(不满足正态, 非参数检验)
- 同如果数据可以进行统计分析,将会进行统计分析。统计分析默认是根据数据情况选择合适的统计方法,涵盖:
  - 两组: T test(满足正态+方差齐) | Welch t' test(满足正态+不满足方差齐性) | Wilcoxon rank sum test(不满足正态,非参数检验)



■ 多组整体: One-way ANOVA(满足正态+方差齐) | Welch one-way ANOVA(满足正态+不满足方差齐性) | Kruskal-Wallis test(不满足正态, 非参数检验)

### 说明文本

个组常见	「统计描述指	标山								
组别1	组别2	数目	最小值	最大值	中位数(Median)	四分位距(IQR)	下四分位	上四分位	均值(Mean)	标准差(SD)
GeneA	NC	3	0.98255	1.0137	1.004	0.015553	0.9933	1.0089	1.0001	0.015926
GeneA	Treatment	3	0.19193	0.34178	0.31172	0.074924	0.25182	0.32675	0.28181	0.079275
GeneB	NC	3	0.92986	1.127	0.95422	0.098579	0.94204	1.0406	1.0037	0.10749
GeneB	Treatment	3	1.8161	2.2446	2.116	0.21422	1.9661	2.1803	2.0589	0.21986

此表格提供统计描述的结果,可以自行提取所需要的。



此表格异常值情况表,可以判断数据是否存在异常值。

: Shapiro-Wilk normali	ty test			
组别1	组别2	自由度(df)	统计量	p值
GeneA	NC	3	0.95372	0.5859
GeneA	Treatment	3	0.89325	0.3643
GeneB	NC	3	0.84108	0.2169
GeneB	Treatment	3	0.94938	0.5666

此表格为正态性检验的结果



方法: Levene's test				
se on Mean				
组别	自由度1(df1)	自由度2(df2)	统计量	p值
GeneA	1	4	7.3628	0.0533
GeneB	1	4	1.8194	0.2487

#### 此表格为方差齐性检验的表格

I条件: 两组组	R立数据, 满足正	态性检验和方差齐性检	194				
组别	组别	组别J	自由度(df)	统计量t	差值(J-I)	置信区间(95%CI)	p值
GeneA	NC	Treatment	4	-15.386	-0.71828	-0.847890.58866	0.000
GeneB	NC	Treatment	4	7,4681	1.0552	0.6629 - 1.4475	0.0017

此表格为独立样本t检验的结果

(不同的分组数量、不同的分子数 对应的统计部分的说明文本都有可能不同, <mark>多组统计学比较还会有两两多重比较的结果</mark>。)

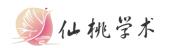


# 方法学

统计分析和可视化均在 R 4.2.1 中进行,具体可根据返回的方法学中的内容进行 查看。

涉及的 R 包: ggplot2 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。

### 结果验证

具体可以查看 分组比较图模块教程文档 中结果验证的内容



## 常见问题

#### 1. 工具计算的结果和自己手动计算的 2^-ΔΔCt 结果不同?

答:

请先下载"qPCR 整理表.xlsx",按照本文档"基本概念"部分提供的计算过程 逐步 核对 自己计算的过程 和 qPCR 整理表 中 各个步骤对应的值。

如果自己的计算过程不同于"基本概念"中的计算过程, (或者因为粗心纳入数据错误导致的计算有问题),则有可能会出现结果不同的情况。

另外还要<u>注意第1列分组的顺序(参考组需要放在前面),以及分子的顺序(内</u>参基因需要放前面)

#### 2. 如何修改 y 轴或者 x 轴标题?

答:

可以在右边找到"标题文本"参数进行输入,如果想要换行,可以中间加入\n



#### 3. 为什么没有统计检验的结果?

答:



当上传的数据各组的<mark>样本数不满足 3 个</mark>以上,或者存在各组样本对应的 Ct 值方 差为 0,则工具不会进行统计检验部分的内容。

4. 能否各组只提供 1 例样本(样本不足,不能一次性提供 3 个以上的样本)? 答:

每组可以只提供 1 列样本或者不足 3 例样本,但是如果提供的样本数不足 3 个,则<u>不会进行统计检验</u>部分的内容。

5. 部分样本没有进行 3 次复孔的检测,是否可以计算?

答:

复孔只会用在第一步计算复孔的均值的时候会用到,一般 qPCR 是会做多个复孔的,如果提供少于 3 次复孔的数据,工具也是会先计算复孔的均值,然后再往下计算的,最终也是可以拿到分析的结果。