

=====图片=====

1. *.insertSize.r.png 样品*插入片段长度模拟分布图(png)

注:通过插入片段两端的 Reads 在参考基因组上的比对应起止点之间的距离计算插入片段长度。

图中横坐标为双端 Reads 在参考基因组上的比对应起止点之间的距离,范围为 0 到 800bp;纵坐标为比对应起止点之间不同距离的双端 Reads 的数目。

插入片段长度检验插入片段长度的离散程度能直接反映出文库制备过程中磁珠纯化的效果。

2. *.rankcheck.png 样品*测序 Reads 在转录本上的位置分布图

注:横坐标为标准化后的 mRNA 位置,纵坐标为对应位置区间内 Reads 在总 Mapped Reads 中所占百分比。

由于参考的 mRNA 长度不同,作图时把每个 mRNA 按照长度划分成 100 个区间,进而统计每一区间内的 Mapped Reads 数目及所占的比例,图中反映的是所有 mRNA 各个区间内的 Mapped Reads 比例的汇总。

3. *.Total.rankchecks.png 所有样品测序 Reads 在转录本上的位置分布图

注:图中不同颜色的曲线代表不同的样品,横坐标为标准化后的 mRNA 位置,纵坐标为对应位置区间内 Reads 在总 Mapped Reads 中所占百分比。

由于参考的 mRNA 长度不同,作图时把每个 mRNA 按照长度划分成 100 个区间,进而统计每一区间内的 Mapped Reads 数目及所占的比例,图中反映的是所有 mRNA 各个区间内的 Mapped Reads 比例的汇总。

4. *_Saturation.png 样品*测序数据的饱和度分布图

注:使用各样品的 Mapped Data 对检测到的不同表达情况的基因数目饱和情况进行模拟,绘制曲线图,可查看随着测序数据量的增加,检测到的不同表达量的基因数目是否趋于饱和。

图中是随机抽取 10%、20%、30%.....90%的总体测序数据单独进行基因定量分析的结果;横坐标代表抽取数据定位到基因组上的 Reads 数占总定位的 reads 数的百分比,纵坐标代表所有抽样结果中表达量差距小于 15%的 Gene 在各个 FPKM 范围的百分比。不同颜色代表不同的基因表达量范围。

=====文件=====

5.Lib_assess.txt 数据链特异性建库评估信息表

注:第一列 sample_id:样品名称;

第二列 read1_vs_gene:该列数值代表同源基因的链信息一致的 read1 数量占总的 read1 的比例;

第三列 read 2 _vs_gene:该列数值代表同源基因的链信息一致的 read 2 数量占总的 read2 的比例;

第四列 undefined:在基因组中可能 含有双向的转录本存在,该区域的 read 无法判断是否同源基因链信息是否一致,所以命名为 undefined。