

## 单个样本 circRNA 预测结果

=====文件=====

|-- \*.circ.intersect.xls CIRC 和 find\_circ 取交集后整理的结果（如果同时使用这两款软件进行 circRNA 预测；否则是经过整理后的某一款软件的结果）

|-- chr : circRNA 所在染色体；  
|-- start : circRNA 开始位置；  
|-- end : circRNA 结束位置；  
|-- Mean\_junction : 平均 junction reads 数目，即同一个 circRNA,求两款软件的平均 junction reads 数目；  
|-- circRNA\_type : circRNA 在基因组上的位置信息，有 exon、intergenic\_region 和 intron 三种；  
|-- gene\_id : circRNA 的来源基因 ID；  
|-- strand : circRNA 所在链。

=====文件夹 1 =====

## |-- BMK\_1\_CIRI CIRI 预测结果（如果使用了此款软件进行 circRNA 预测）

CIRI 预测环状原理：CIRI 软件使用 BWA 软件与参考基因序列比对，生成 SAM 文件，并对 SAM 文件中的 CIGAR 值进行分析的，从 SAM 文件中扫描 PCC 信号（paired chiastic clipping signals）。CIGAR 值在 junction read 的特征是 xS/HyM 或者 xMyS/H，其中 x,y 代表碱基数目，M 是 mapping 上的，S 是 soft clipping，H 是 hard clipping。对于双端 Reads，CIRI 算法会考虑一对 reads，其中一条可以 mapping 到 CircRNA 上，另一条也需要 mapping 到 CircRNA 的区间内。对于单外显子成环，或者“长外显子 1-短外显子-长外显子 2”形成的环形结构，CIGAR 值应该是 xS/HyMzS/H 以及 (x+y)S/HzM 或者 xM(y+z)S/H，CIRI 软件可以很好的将这两种情况分开。对于 splicing 信号(GT,AG) CIRI 也会考虑其他弱 splicing 信息例如（AT-AC），算法会从 GTF/GFF 文件中抽取外显子边界位置，并用已知的边界来过滤假阳性。

## |-- \*.CIRI\_Predict.xls CIRI 软件预测环状 RNA 结果

```
|-- circRNA_ID : circRNA 名称，格式:染色体：start|end
|-- chr : 所在染色体
|-- circRNA_start : 开始位置
|-- circRNA_end : 结束位置
|-- #junction_reads : junction reads 的数目
|-- SM_MS_SMS : CIGAR
|-- #non_junction_reads : 非 junction reads 的数目
|-- junction_reads_ratio : junction_reads/(non_junction_reads/2+junction_reads)
|-- circRNA_type : circRNA 类型
|-- gene_id : 来源基因 ID
|-- strand : 方向
|-- junction_reads_ID : junction reads 的 ID（测序 clean reads ID）
```

|-- BMK\_2\_find\_circ      find\_circ 软件预测结果（如果使用了此款软件进行 circRNA 预测）

|-- \*.find\_circ\_Predict.xls      find\_circ 软件预测环状 RNA 结果

|-- chrom：染色体名称  
|-- start：左剪切位点（以 0 开始）  
|-- end：右剪切位点（以 0 开始）  
|-- name：环状 RNA 名称  
|-- n\_reads：number of reads supporting the junction (BED'score')  
|-- strand：+/-；n\_uniq：支持 junction 的唯一 reads 数目  
|-- uniq\_bridges：锚点比对唯一的 reads 数目  
|-- best\_qual\_left：支持左剪切 junction 最好的锚点比对分数  
|-- best\_qual\_right：支持右剪切 junction 最好的锚点比对分数  
|-- tissues：样品名称，以逗号分割  
|-- tiss\_counts：每个样品对应的 reads 数，以逗号分割  
|-- edits：锚点延伸过程的错配数  
|-- anchor\_overlap：断点在一个锚点内的核苷酸数目  
|-- breakpoints：以侧翼 GT/AG 方式打断 reads 的数目  
|-- signal：侧翼核苷酸剪切信号(GT/AG)  
|-- strandmatch：'MATCH','MISMATCH' or 'NA'（非链特异性分析）  
|-- category：描述 junction 的关键字