

miRNA 预测

=====文件夹 1=====

|-- BMK_1_miR_Seq 预测 miRNA 序列

 |-- All_miRNA.expressed.fa 预测得到的所有 miRNA 的 mature 序列（重要文件）

 |-- All_miRNA_Pre.expressed.fa 预测得到的所有 miRNA 的前体序列（重要文件）

 |-- Known_mature.expressed.fa 已知 miRNA 的 mature 序列

 |-- Known_Pre.expressed.fa 已知 miRNA 的前体序列

 |-- Novel_mature.expressed.fa 新预测的 miRNA 的 mature 序列

 |-- Novel_Pre.expressed.fa 新预测的 miRNA 的前体序列

=====文件夹 2=====

|-- BMK_2_Len_Distribution miRNA 长度分布结果

 |-- All_miRNA.length.pdf 所有 miRNA 长度分布图（PDF 格式）

 |-- All_miRNA.length.png 所有 miRNA 长度分布图（PNG 格式）

 |-- All_miRNA.stat.xls 所有 miRNA 长度分布统计文件

 |-- Length 长度

 |-- Number 此长度的 miRNA 的数目

 |-- Known_miRNA.length.pdf 已知 miRNA 长度分布图（PDF 格式）

 |-- Known_miRNA.length.png 已知 miRNA 长度分布图（PDF 格式）

 |-- Known_miRNA.stat.xls 已知 miRNA 长度分布统计文件

 |-- Novel_miRNA.length.pdf 新预测 miRNA 长度分布图（PDF 格式）

 |-- Novel_miRNA.length.png 新预测 miRNA 长度分布图（PDF 格式）

 |-- Novel_miRNA.stat.xls 新预测 miRNA 长度分布统计文件

 |-- *.known_miRNA_len.stat.xls 单样本已知 miRNA 长度分布统计文件

 |-- *.predicted_miRNA_len.stat.xls 单样本新预测 miRNA 长度分布统计文件

=====文件夹 3=====

|-- BMK_3_Base_Distribution miRNA 碱基偏好

|-- All.base_bias.pdf

|-- All.base_bias.png 所有 miRNA 各位点碱基分布图（PNG 格式）

图中横坐标表示序列的各位点；纵坐标表示 miRNA 各位点碱基所占百分比。

|-- All.base_bias.stat.xls 所有 miRNA 各位点碱基分布统计文件

文件第一列为 pos，之后每一列是该 pos 是分别是 A/U/G/C 的 miRNA 的数目。

|-- All.first_base.pdf

|-- All.first_base.png 所有 miRNA 首位碱基的分布图（PNG 格式）

|-- All.first_base.stat.xls 所有 miRNA 首位碱基的统计文件

|-- Known.base_bias.pdf

|-- Known.base_bias.png 已知 miRNA 各位点碱基分布图（PNG 格式）

|-- Known.base_bias.stat.xls 已知 miRNA 各位点碱基分布统计文件

|-- Known.first_base.pdf

|-- Known.first_base.png 已知 miRNA 首位碱基的分布图（PNG 格式）

|-- Known.first_base.stat.xls 已知 miRNA 首位碱基的统计文件

|-- Novel.base_bias.pdf

|-- Novel.base_bias.png 新预测的 miRNA 各位点碱基分布图（PNG 格式）

|-- Novel.base_bias.stat.xls 新预测的 miRNA 各位点碱基分布统计文件

|-- Novel.first_base.pdf

|-- Novel.first_base.png 新预测的 miRNA 首位碱基的分布图（PNG 格式）

|-- Novel.first_base.stat.xls 新预测的 miRNA 首位碱基的统计文件

|-- BMK_4_Summary_Stat

|-- Summary_Known_miRNA.xls 已知 miRNA 的总结文件

鉴定 miRNA 时，我们将比对上参考基因组的 reads 序列与已知 miRNA 数据库 miRBase (v21) 中的成熟 miRNA 序列进行比对。序列与已知 miRNA 完全相同的 reads 被认为是本项目中鉴定到的已知 miRNA，因此此文件中的部分列为"--"；

```
|-- miRNA : miRNA ID
|-- score_total : "--"
|-- mature_sequence : 从 miRbase 数据库中提取的已知 miRNA 的成熟序列
|-- star_sequence : "--"
|-- GenomeID : "--"
|-- strand : "--"
|-- start : "--"
|-- end : "--"
|-- pre_seq : "--"
|-- hairpin_stru : "--"
|-- hairpin_energy : "--"
|-- S*   miRNA 在各个样品中的表达量信息
```

|-- Summary_Novel_miRNA.xls 新预测 miRNA 的总结文件

新 miRNA 用 miRDeep2 软件包进行预测。

```
|-- miRNA          miRNA ID
|-- score_total    软件打分
|-- mature_sequence 预测得到的 miRNA 的 mature 序列
|-- star_sequence  Star 序列
|-- GenomeID       序列所在染色体
|-- strand         序列所在链的方向
|-- start          前体序列的开始位置
|-- end            前体序列的结束位置
|-- pre_seq        前体序列
|-- hairpin_stru    发夹结构
|-- hairpin_energy  发夹结构能量值
|-- S*   miRNA 在各个样品中的表达量信息
```

|-- Total_miRNA.stat.xls 所有 miRNA 数据统计文件

统计了每个样品中预测得到的已知和 Novel miRNA 的数量及 Total miRNA 的数量。

=====文件夹 4=====

|-- BMK_5_PDF

 |-- *.pdf 二级结构预测