

LncRNA 的预测

=====文件=====

|-- LncRNA_circos.png LncRNA 在染色体上分布环状图

描述不同种类 LncRNA 在染色体上的分布；图片的最外层是基因组的染色体环，由外到内依次是正义链 LncRNA 环(绿色),基因间区 LncRNA 环(红色),反义 LncRNA 环(灰色),内含子 LncRNA 环(蓝色)。

|-- LncRNA_classification.png LncRNA 的分类柱状图（重要图）

横坐标为 4 种不同类型的 LncRNA（其中 lincRNA 表示基因间区长链非编码 RNA，Antisense-lncRNA 表示反义长链非编码 RNA，Intronic-lncRNA 表示内含子长链非编码 RNA，sense_lncRNA 表示正义长链非编码 RNA）；纵坐标为对应的 LncRNA 数量。

|-- LncRNA_classification.xls LncRNA 的分类表

|-- 第一列：LncRNA ID；
|-- 第二列：LncRNA 的分类注释信息。

=:Complete match of intron chain（已知 lncRNA）

c:Contained(已知 lncRNA)

j:Potentially novel isoform (fragment): at least one splice junction is shared with a reference transcript(对应图中 sense_lncRNA:正义长链非编码 RNA)

e:Single exon transfrag overlapping a reference exon and at least 10 bp of a reference intron, indicating a possible pre/ mRNA fragment.

i:A transfrag falling entirely within a reference intron（对应图中 Intronic/ lncRNA:内含子长链非编码 RNA）

o:Generic exonic overlap with a reference transcript(对应图中 sense_lncRNA:正义长链非编码 RNA)

u:Unknown, intergenic transcript（对应图中 lincRNA:基因间区长链非编码 RNA）

x:Exonic overlap with reference on the opposite strand(对应图中 Antisense/ lncRNA：反义长链非编码 RNA)

s:An intron of the transfrag overlaps a reference intron on the opposite strand (likely due to read mapping errors)

|-- LncRNA_classification_name.xls 每个 LncRNA 的对应四种类型名称

|-- 第一列：LncRNA ID；
|-- 第二列：对应四种类型名称。

lincRNA :基因间区长链非编码 RNA ;Antisense-lncRNA :反义长链非编码 RNA ;Intronic-lncRNA :
内含子长链非编码 RNA ; sense_lncRNA : 正义长链非编码 RNA

|-- LncRNA.fa lncRNA 的预测结果，fasta 文件（重要文件）

FASTA 格式每一个序列单元以">"开头，直到出现下一个">"之前为止。">"开头的行为序列 ID 行，后面紧接着基因 ID；下面一行或多行为该基因的碱基序列。

|-- LncRNA.gff lncRNA 的预测结果，gff 文件（重要文件）

|-- 第一列：#Seq_ID—染色体号；
|-- 第二列：Source—注释信息的来源；
|-- 第三列：Type—注释特征(Feature)类型；
|-- 第四,五列：Start/End—特征序列的起止位置；
|-- 第六列：Score—得分，数字，注释信息可能性的说明，“,”表示缺失值；
|-- 第七列：Strand—特征序列所在的正负链；
|-- 第八列：Phase—仅对注释类型为 CDS 有效，表示起始编码的位置,有效值为 0、1、2，“,”表示缺失值；
|-- 第九列：Attributes—以多个键值对组成的注释信息描述。

|-- LncRNA.gtf lncRNA 的预测结果，gtf 文件（重要文件）

|-- 第一列：Chr—染色体序号；
|-- 第二列：Source—来源；
|-- 第三列：Type—转录本或者其外显子；
|-- 第四列：Start Site—起始坐标；
|-- 第五列：End Site—终止坐标；
|-- 第六列：Score—对应位置得分；
|-- 第七列：Strand—链的信息；
|-- 第八列：frame—frame 的信息；
|-- 第九列：Description—描述信息。

=====文件夹=====

|-- BMK_1_Software_Result 四种软件编码能力预测