==========文件==========

1. *.newGene.longest_transcript.fa 新预测到的基因的最长转录本序列

2. *.newGene_final.filtered.gff 新预测到的基因的 gff 文件

注：gff 是纯文本文件，由 tab 键隔开的 9 列组成，以下是各列的说明：

Column 1:	“seqid”	序列的编号，此处用染色体编号表示
Column 2:	“source”	注释信息的来源，比如”Genescan”、”Genbank”等，可以为空，为空用”.”点号代替，由于新基因是用 cufflinks 预测的，此处用 Cufflinks 表示
Column 3:	“type”	注释信息的类型，比如 Gene、cDNA、mRNA 等，或者是 SO 对应的编号
Columns 4 & 5:	“start” and “end”	开始与结束的位置，注意计数是从 1 开始的。结束位置不能大于序列的长度
Column 6:	“score”	得分，数字，是注释信息可能性的说明，可以是序列相似性比对时的 E-values 值或者基因预测是的 P-values 值。”.”表示为空。
Column 7:	“strand”	序列的方向， +表示正义链, -反义链 , ? 表示未知.
Column 8:	“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。”.”表示为空。
Column 9:	“attributes”	以多个键值对组成的注释信息描述，键与值之间用”=“，不同的键值用”;“隔开，一个键可以有多个值，不同值用”,“分割。注意如果描述中包括 tab 键以及”,=;”，要用 URL 转义规则进行转义，如 tab 键用 %09 代替。键是区分大小写的，以大写字母开头的键是预先定义好的，在后面可能被其他注释信息所调用。

3. New_gene.fa 新预测基因的序列

==========文件夹==========

1. BMK_1_NewGene_Anno 新预测基因的功能注释