第6次上机作业

1.从 FASTA文件中过滤空序列

现在有一个序列文件out22.fas，其中有一些序列只有文件头而没有序列，使用Biopython中的函数完成对空序列的过滤，将过滤后的结果写入一个新的文件中。

2.创建随机序列的FASTA文件

有一些统计检验需要使用随机序列。随机序列在当你没有足量的真实数据时，也可以被用来检验程序。  
我们假定需要产生500条序列，每一条的长度在4000到15000之间，试使用Biopython生成fasta序列文件。

3.提取选定的蛋白的上游序列

给定一个Genbank文件MTtabacum.gbk，保存了烟草的线粒体信息。使用Biopython中gbk文件处理的函数，完成编程。

在基因起始端的上游区域大多数是调控元素所在区域。如聚腺苷酸信号，TATA盒，增强子等。这个程序我们有一个Genbank文件和一列基因（cox2,atp6,atp9,cob）,我们需要提取它们的上游序列，直到1000个碱基。