**Excercise1:**

**1-4：**

long-orfs:

-n no header

-t 1.15 cutoff 1.15, considering only genes with an entropy score less than 1.15)

glimmer3:

-o50 maximum overlap bases = 50

-g100 minimum gene length = 100

-t30 Threshold score = 30

**5.**

Other two sources: known genes in the genome, genes in homologous species/strains

**6.**

No. In 24.fa and 49.fa, Glimmer predicts very few ORFs.

Glimmer can only be used in prokaryotic genomes. The prokaryotic genome and the eukaryotic genome is very different. Because of the high-gene-density property, it is easier to predict genes in prokaryotic genomes and the prediction of Glimmer is based primarily on the location of the six possible reading frames. It can't handle the introns and other complex situations in eukaryotic genome.

**8.**

No. ORFs in all genomes follow right-skewed distribution. However, in 24.fa and 49.fa, very few ORFs are found and there is no specific distribution pattern for them.

**9.**

Homo sapiens TRAP1 transcript variant 1, mRNA

Homo sapiens TRAP1 transcript variant X1, mRNA