1.-(10 points) Define the following file formats relevant in bioinformatics: Fasta format, Genbank format, fastq format.

FASTA format:

FASTA Format is a text based format for representing a peptide sequence or nucleotide sequence in which nucleotides or aminoacids are represented using single letter codes.

Fasta format starts with the description followed by the lines of sequence data .This description is only of one line. The description line is separated from the sequence data by a (">"). The word following the ">" symbol is the identifier of the sequence, and the other part of the line is the description . There should be no space between the ">" and the first letter of the identifier. The sequence ends if another line starting with a ">" appears; this indicates the start of another sequence.

A simple example of one sequence in FASTA format:

```
>gi|5524211|gb|AAD44166.1| cytochrome b [Elephas maximus maximus]
LCLYTHIGRNIYYGSYLYSETWNTGIMLLLITMATAFMGYVLPWGQMSFWGATVITNLFSAIPYIGTNLV
EWIWGGFSVDKATLNRFFAFHFILPFTMVALAGVHLTFLHETGSNNPLGLTSDSDKIPFHPYYTIKDFLG
LLILILLLLLALLSPDMLGDPDNHMPADPLNTPLHIKPEWYFLFAYAILRSVPNKLGGVLALFLSIVIL
GLMPFLHTSKHRSMMLRPLSQALFWTLTMDLLTLTWIGSQPVEYPYTIIGQMASILYFSIILAFLPIAGX
IENY
```

Genbank format: