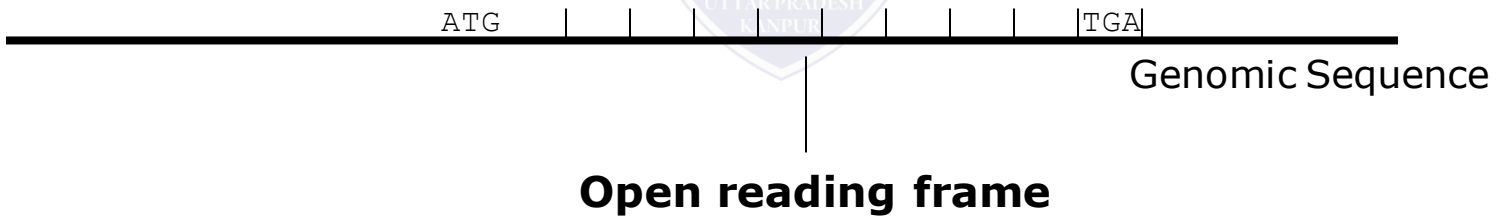
FACULTY OF ENGINEERING AND TECHNOLOGY

Department Biotechnology

ORFs

- In molecular genetics, an **open reading frame** (ORF) is the part of a reading frame that contains no stop codons or region of amino acids coding triplet codons
- **An ORF starts with a start codon ATG (Met) in most species and ends with a stop codon (TAA, TAG, TGA)** Transcription
- termination pause site is located after ORF



- One common use of open reading frames is as one piece of evidence to assist in gene prediction
- Potential coding regions of a gene are detected by looking at **ORFs** in a DNA sequence

Simple First Step In Gene Finding

- A genome of length *n* is comprised of (*n*/3) codons
- Stop codons break genome into segments between consecutive Stop codons
- Segments of these that start from Start codon (**ATG**) are ORFs
  - ORFs in different frames may overlap

- Genomic sequence is translated in codon frames
- Stop codons are identified in each frame
- Regions without stop codons are called ORFs
- All of the likely ORFsin a sequence are located and tagged
- Longest ORFfrom a Met codon is a good prediction of a protein encoding sequence

Reading sequence of DNA in six reading frames

- Every region of DNA has six possible **reading frames;**
  - Three in the forward direction
  - Three in the reverse direction

- Frame **1** starts with the "**a**", Frame **2** with the "**t**" and Frame **3** with the "**g**"
- 5' atgcccaagctgaatagcgtagaggggttttcatcatttgaggacgatgtataa 3'

1 atg ccc aag ctg aat agc gta gag ggg ttt tca tca ttt gag gac gat gta taa M P K
L N S V E G F S S F E D D V *

2 tgc cca agc tga ata gcg tag agg ggt ttt cat cat ttg agg acg atg tat C P
S * I A * R G F H H L R T M Y

3 gcc caa gct gaa tag cgt aga ggg gtt ttc atc att tga gga cga tgt ata AQ A E
* R R G V F I I * G R C I

- Stop codons are indicated by an "*" in the protein sequence

The longest ORF is in Frame **1**

SOFTWARES TO FIND ORFs

- **ORF Finder:** It is a graphical analysis tool which finds all open reading frames of a selectable minimum size in a user's sequence or in a sequence already in the database
- **ORF Investigator:** It is a program which not only gives information about the coding and non-coding sequences but also can perform pairwise global alignment of different gene/DNA regions sequences
- **ORFPredictor:** It is a web server designed for identifying protein-coding regions in expressed sequence tag (EST)-derived sequences

XYZ

abc

XYZ

abc

XYZ

abc

XYZ

abc

XYZ

abc

XYZ

abc

XYZ

abc

XYZ

abc

XYZ

abc

MCQs

1. A
2. A
3. A
4. A
5. A
6. A
7. A
8. A
9. A
10. A