Molecular Biology and Genetics Undergraduate Program
Department of Biological Sciences
Faculty of Arts and Sciences
Eastern Mediterranean University

This lecture is part of Spring 2020 – Covid-19 Pandemic – Online Education

Following rules are to be strictly followed:

- **Unauthorized photography, video or audio recordings of these sessions are NOT allowed.**
- **Lecture materials are not to be in-part or in-full distributed IN ANY FORM, ON ANY PLATFORM whatsoever.**

DETECTION OF SEQUENCE POLYMORPHISM AND SNP DATABASE

Mutations can be;

- Point mutations
- Small deletions
- Insertions
- Large-scale changes

SNP (Single Nucleotide Polymorphism)

- A point mutation that occurs in at leats 1% of the population is called SNP
- Important class of mutations
- Occur at a frequency of at least 0.1% (1/1000 bases)
- May occur more frequently in certain regions
- In human genome, >65% of all SNPs involve

C T transition mutations

→

- A set of linked SNPs that tend to inherit together as a unit is referred to as SNP haplotype.
- May occur in both coding and non-coding regions
- SNPs in coding regions may alter the characteristics of the protein
- SNPs in the regulatory regions may alter the expression profile of genes

International HapMap Project

- Is a multi-country (USA, UK, Canada, Japan, China, Nigeria) effort to identify and catalog genetic similarities and differences in humans.
- Identify and catalog SNPs and SNP haplotypes that confer susceptibility/resistance to disease or therapy.

- Sequence polymorphisms can be detected through pairwise alignment of two DNA sequences from two individuals
- Deep resequencing of specific regions of the genome can also identify sequence polymorphism

dbSNP

- NCBI SNP database is the largest public database of short genetic variations (SNVs).
- Includes;
  - Single-base nucleotide substitution (SNPs)
  - Small scale mutti-base deletions or insertions (deletion-insertion polymorphism or DIPs)
  - Retroposable element insertions and microsatellite repeat variations (short tandem repeats or STRs)

Real Life Applications: Detecting and Understanding the Effects of DNA Mutations at DNA and Protein Level

- You are a scientist working on identifying mutations related to breast cancer. You collect DNA from patients and test them hoping to identify novel mutations.
- You performed DNA sequencing for the GBE1 gene in your samples. You know that DNA sequencing may not read the start and the end part of the sequence tested well. You hope to find a SNP in your sample. You are using NM_000158.4 as a reference sequence.
- This means, you compare your DNA sequence and this reference sequence to identify any SNPs. You are very curious to see if any SNPs you find change the protein sequence, structure and thus function.