

# Resources Cancer Detective

## What is Cancer Detective?

**Cancer Detective** is an analysis simulation based on Cancer Genome Interpreter, for an educational purpose.



## Cells

**The cells** are the "building blocks of life", from the Latin cellula 'small room', they are the basic structural, functional, and biological unit of all known organisms.

**The nucleus** is the "control center of the cell", a membrane-bound organelle found in eukaryotic cells, enclosed by the nuclear envelope, isolating its contents from the cytoplasm; it contains all of the cell's genome (except for the small amount of mitochondrial DNA), it maintains the integrity of the genes and controls all the activities of the cell.

**The cytoplasm** is all of the material of eukaryotic cells, enclosed by the cell membrane, except for the cell nucleus; mainly composed of cytosol (a gel-like substance) and the organelles (the cell's internal sub-structures).

**The ribosomes** macromolecular machines that perform biological protein synthesis (mRNA translation).

## The genome

**The genome** comprises all genetic information of an organism, consisting of nucleotide sequences of DNA (or RNA in RNA viruses).

**The DNA** is a molecule composed of two polynucleotide chains that coil around each other to form a double helix, carrying genetic instructions for the development, functioning, growth and reproduction of all known organisms.

**A gene** is a basic unit of heredity and a sequence of nucleotides in DNA or RNA that encodes the synthesis of a gene product, either RNA or protein.

**The RNA** is a molecule composed of one polynucleotide chain, essential in various biological roles in coding, decoding, regulation and expression of genes; cellular organisms use messenger RNA (mRNA) to convey genetic information that directs synthesis of specific proteins.

## The proteins

**The proteins** are macromolecules that comprise one or more long chains of amino acid residues, they perform a vast array of functions within organisms; they are assembled from amino acids using

information encoded in genes; the information contained in the mRNA is used as a template to synthesize proteins by ribosomes.

## **The mutations**

**Mutations** are errors or changes in the DNA sequence of cells. Mutations are naturally occurring in the cells of our bodies. Most mutations do not have an effect, although some can alter the function of proteins.

**The mutational processes** are a set of mechanisms that introduce mutations in the DNA. They can be classified as external (exogenous) or internal (endogenous) processes according to their origin.

**External mutational** processes are caused by agents that originate outside our bodies such as tobacco smoking or UV light. **Internal mutational** processes are caused by mechanisms inside the tissues and cells inside our bodies.

## **Cancer**

**Cancer** is the name of a group of diseases characterized by the uncontrolled growth and proliferation of cells within the body. Cancer cells can divide without control and invade organs and tissues, causing metastasis.

**Cancer driver mutations** are mutations that confer selective advantage to the cells bearing them, allowing them to survive and proliferate more.

**Cancer driver genes** are genes or proteins that are affected by cancer driver mutations.

## **Tumors**

**A tumor** is a mass of cells, that can be benign (lacking the ability to either invade neighboring tissue or spread throughout the body) or malignant (cancer).

## **Selective advantage**

Regarding cells, **the selective advantage** refers to the characteristic that enables them to survive and proliferate better than other cells in a population in a given environment; the basis for evolution by natural selection.

## **Precision cancer medicine**

**Precision cancer medicine** is a form of medical care that uses information about a patient's own genes or proteins (molecular profiling) to prevent, diagnose, or treat cancer.

**Targeted therapies** are the type of treatments that use drugs or other substances to identify and attack specific types of cancer cells with less harm to normal cells.

## **Biopsy**

**A biopsy** is a removal of cells or tissues for examination by a pathologist; the most common types include: (1) **incisional biopsy**, in which only a sample of tissue is removed; (2) **excisional biopsy**, in which an entire lump or suspicious area is removed; and (3) **needle biopsy**, in which a sample of tissue or fluid is removed with a needle.

## **Bioinformatics**

**Bioinformatics** is the field of science that uses computers, databases, math, and statistics to collect, store, organize, and analyze large amounts of biological, medical, and health information.

**Machine learning** is the use and development of computer systems that are able to learn and adapt without following explicit instructions, by using algorithms and statistical models to analyse and draw inferences from patterns in data.

**An algorithm** is a process or set of rules to be followed in calculations or other problem-solving operations, especially by a computer.

## **BBGLab bioinformatic tools**

**IntOGen** is a unified resource that merges information from different statistical methods and identifies cancer driver proteins.

**BoostDM** is a machine learning algorithm that can predict which mutations within cancer driver proteins are actually driver mutations.

**The Cancer Genome Interpreter** is a algorithm that uses the data generated by IntOGen and other resources to inform us whether a mutation in a protein is a cancer driver and if there is a specific therapy for cancers with this alteration.