



## Home assignment 1

You have been given the designation of one specific type of human cancer. You will explore this disease and all its related data and information until the end of the semester.

For the first assignment the purpose is solely to explore the available data about the disease within the most relevant bioinformatics repositories

- A. Identify the disease in OMIM database about the disease and identify known biological sequences implied
- B. Identify the **most relevant proteins related to the disease** over the known repositories and extract the most important information
  - a. Protein and Gene names and IDs
  - b. Protein GO Annotation (molecular function, biological process and cellular component. Suggestion: Summarize the information found with term counts
  - c. Known ligands (where appropriate)
  - d. Variants and known mutations relevant to the disease
  - e. Structure (secondary and tertiary if available)
  - f. Any other relevant information deemed relevant
  - g. Relevant publications – Check PubMed and search for articles that are relevant within the last 2 years. Do they confirm or disprove what you have so far found? Are there contradictory information?
- C. Write a report (4 pages maximum) relative to the information found. The report should have the following structure:
  - a. Introduction with 200-400 words referring the most important information relative to the proteins and genes in study.
  - b. Methods – describing the repositories and the work performed and the most relevant difficulties
  - c. Discussion - describe and illustrate the findings. Use of figures and tables is encouraged to summarize information

**Deadline: March 12**

**The submission will be made online within the course's moodle**