INTERNSHIP PROGRAM FOR INTERNATIONAL STUDENTS

INTERNSHIP SUBJECT FORM

Name of the Host Laboratory	LIX (and BIOC)
Website of the Host Laboratory	https://www.lix.polytechnique.fr/
Research Group	AMIBio
Internship Supervisor	Mireille Régnier (with Alexis Gautreau)
Internship Subject	**Analysis of driver mutations in breast cancer**
Student's level	*☐ Advanced Undergraduate Students (3rd or 4th year) *☐ Master's students (1st or 2nd year) ☐ PhD students
Proposed Duration	☐ 3 months ☐ 4 months : from 4 to 6 months ☐ 5 months ☐ 6 months
Prerequisites	Script language (python) ; basis in probability/statistics
Internship description (max. 15 lines)	The human genome is composed of about 20 000 genes. Mutations of genes drive the progression of cancer, with several stages from a benign abnormal cell proliferation to a malignant and invasive tumor. However, mutations arise by chance and the few driver mutations that confer a growth advantage to the cell are found together with a majority of neutral mutations. Next generation sequencing of many tumors of all cancer types has identified the genes that are most frequently mutated. These genes are the most likely ones to drive cancer progression. Many authors have provided their list of cancer genes, which contains about 1% of total genes. In breast cancer, each tumor accumulates on average 33 mutations, from which only 2 to 7 drive cancer progression. The goal of this project is to develop a software that identifies the genes driving the progression of the tumor from all the mutated genes found in a single breast tumor. The output should a be the list of mutated genes ranked according to their probability of being driver. The tumoral genome should be compared to publicly available databases of tumor sequences. Mutation frequency determined from these databases would be a first (or main or initial) criteria for the algorithm. Mutation frequency, however, will be one criterium among several ad hoc criteria to be determined in close collaboration with the biology team. This bioinformatics project will unfold in parallel to an attempt to reconstitute the tumoral process by introducing mutations in the genome of a normal human breast cell in the biology lab.