INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

DPT BIOSCIENCES
INSA Campus LyonTech - 11 Avenue Jean Capelle
Batiment Louis Pasteur - 69621 VILLEURBANNE
Tel. 0472436448
Courriel : bs-secretariat@insa-lyon.fr

Bioinformatique

Génomique médicale

IDENTIFICATION

CODE : BS-5-S1-EC-COGEMED
ECTS : 2.0

HORAIRES

Cours :	8.0 h
TD :	16.0 h
TP :	0.0 h
Projet :	0.0 h
Face à face pédagogique :	24.0 h
Travail personnel :	26.0 h
Total :	50.0 h

ÉVALUATION

1 projet

SUPPORTS PÉDAGOGIQUES

Les supports pédagogiques seront disponibles sur la page Moodle dédié à ce cours.

LANGUE D'ENSEIGNEMENT

Français

CONTACT

M. PARISOT Nicolas
nicolas.parisot@insa-lyon.fr

OBJECTIFS RECHERCHÉS PAR CET ENSEIGNEMENT

OBJECTIFS :
A l'issue de ce module l'étudiant devra être capable de s'insérer dans un programme de recherche ou de développement en génomique médicale avec une complète autonomie de travail.
Les objectifs pédagogiques de ce module sont :
- d'apporter les concepts biologiques et informatiques fondamentaux en bioinformatique pour la génomique médicale.

PROGRAMME

Theoretical part:
1) Genomics
- concepts: genotype-phenotype association, Mendelian and complex diseases, rare diseases, germline and somatic alterations, cell-free DNA
- resources: human reference genome, annotations, and databases (TCGA) for medical genomics; cloud computing resources
- techniques: arrays and whole-exome/genome sequencing of cohorts, trio sequencing
- methods: calling germline and somatic variants (SNVs, indels, structural variants, mutational signatures, cf DNA), genotype imputation, Genome-Wide Association Studies
2) Transcriptomics, multi-omics and beyond
- concepts: inter- and intra-tissue heterogeneity, cancer and microenvironment; complementarity of different 'omic' layers; clinical data and digital pathology
- resources: large reference databases (GTEx, TCGA, scRNA atlases)
- techniques: bulk sequencing, single-cell sequencing
- methods: calling somatic variants (SNVs, indels, structural variants from RNA-seq), deconvolution and quantifying the tumor microenvironment; multi-omic integration and classification (associations between layers, embeddings), deep learning and integration with image analysis
3) Epigenomics (course in English)
- concepts: chromatin and histone modification, methylation, cell-type signatures, effect of environmental factors
- resources: annotations and databases (tissue-specific profiles)
- techniques: ATAC-seq, bisulfite sequencing, methylation arrays
- methods: mapping, methylation quantification, peak calling, differentially methylated positions and regions, deconvolution and identification of cell types, inference of environmental risk factors
4) Metabolomics (course in English)
- concepts: experimental design, metabolites and disease, biomarkers
- resources: databases (MetaboLights repository), workflow repositories and cloud resources (PHENOMenal, workflows4metabolomics)
- techniques: NMR, mass-spectrometry, Liquid Chromatography Mass Spectrometry
- methods: peak detection, metabolite identification; clustering and regression, metabolic pathway and network analysis; identifying biomarkers

Practical part:
1) Developing and deploying an open-source medical genomic bioinformatic workflow
- using the Nextflow DSL to run parallel, scalable analyses on High Performance Computing and cloud computing facilities
- efficient use of github for open-source development and Continuous Integration automated tests
- reliance on conda and docker/singularity containers for reproducibility
2) Performing a multi-omic analysis of cancer data with R
- accessing public cancer resources from the R environment
- performing uni-omic and multi-omic molecular classifications
- interpreting the results and finding clinical implications

Projects:
Several projects will be proposed to process (bioinformatic workflow development) and analyze (R scripting) cancer data, related to the interests of researchers of the International Agency for Research on Cancer - WHO (rare cancers, pan-cancer analyses, intra-tumor heterogeneity and evolution). Students will work in small groups (~4 people). Weekly meetings (in person or remotely) will take place with the supervisor. A final project restitution and debriefing will be held at the end of the module.

BIBLIOGRAPHIE

-

PRÉ-REQUIS

- Connaissances de base en bioinformatique (analyse de séquences NGS)
- Connaissances de base en biologie moléculaire

INSA LYON
Campus LyonTech La Doua
20, avenue Albert Einstein - 69621 Villeurbanne cedex - France
Tel. +33 (0)4 72 43 83 83 - Fax +33 (0)4 72 43 85 00
www.insa-lyon.fr