

AGENDA

[Online Workshop] Data Analysis for Genomics

Available on website: <https://caotranbichngoc.wixsite.com/dagws/program>

Time: 19:00-21:00 from August 24th to August 30th, 2020

Venue: IT room 502 and 504, USTH, A21 Building, VAST Campus, 18 Hoang Quoc Viet, Hanoi

(Optional) Tran Bich Ngoc CAO and Cyril Matthey-Doret

Monday 24th August

Pack 1 : Biostatistics for Dummies - Part 1

Announcement of the Workshop, Set up Jupyter Notebook, version control Python/Anaconda, program description. Introduction to Biostatistics

Tuesday 25th August

Pack 1 : Biostatistics for Dummies - Part 2

Revision of key concepts of statistics

Wednesday 26th August

Pack 1 : Biostatistics for Dummies - Part 3

Small project with JP Notebook

Choose your Room :

Thursday 27th August

Room 001 : Invited Speaker (Cyril Matthey-Doret)

Pack 2 : Sequence algorithms Part 1 Introduction to sequence algorithms

** Algorithms basics * Time complexity * Data structures * Hashing*

Room 002 : Tran Bich Ngoc CAO

Pack 3 : Scikit-learn Python : Introduction of Machine Learning in Genomics

Friday 28th August

Room 001 : Invited Speaker (Cyril Matthey-Doret)

Pack 2 : Sequence algorithms Part 2 Sequence alignments

** Exact matching - naive - suffix arrays and trees*

** Inexact matching - Dynamic programming*

Room 002 : Tran Bich Ngoc CAO

Pack 3 : Scikit-learn Python : Unsupervised learning

Saturday 29th August

Room 001 : Invited Speaker (Cyril Matthey-Doret)

Pack 2 : Sequence algorithms Part 3 - Computational models in genomics

* Markov models and HMM

* Classification problem - Tree-based models - Intro to machine learning

Room 002 : Tran Bich Ngoc CAO

Pack 3 : Scikit-learn Python Part 3 : Supervised learning

Sunday 30th August

Room 001 : Invited Speaker (Cyril Matthey-Doret)

Pack 2 : Sequence algorithms Part 4 - Network analysis

* Graph representations * Graph traversal * Network clustering

Room 002 : Tran Bich Ngoc CAO

Pack 3 : Scikit-learn Python Part 4 : Cross-validation and model assessment

* Keynote speaker

Speakers of the workshop:

1st speaker: **Ms. Cao Tran Bich Ngoc**, Bachelor Gen 7 at USTH, Master student in Bioinformatics at the École Normale Supérieure Paris (ENS Paris); Author of a publication published in Nucleic Acids Research (IF: 11.5); research experience at Institut Pasteur, Paris, Centre de Recherche en Cancérologie de Toulouse (CRCT), Toulouse (France), National Yang-Ming University (Taiwan).

2nd speaker: **Cyril Matthey-Doret**, PhD Student of the PPU Program (International Doctoral Program) at the Institut Pasteur, Paris; Author of 4 publications, 1 software (Chromosight); Specialist in Computer vision and Pattern recognition; Master in Bioinformatics and worked as Bioinformatician at University of Lausanne, Switzerland.