



Department of Computer Science  
College of Engineering  
University of the Philippines  
Diliman, Quezon City



COURSE NO.	: BIO 397
COURSE TITLE	: Current Topics in Biology - Advanced Bioinformatics
COURSE CREDIT	: 3 units
SEMESTER	: Second Semester, AY 2018-2019
INSTRUCTOR	: Jan Michael C. Yap
EMAIL ADDRESS	: jcyap@dcs.upd.edu.ph

### COURSE DESCRIPTION

This course serves as an advanced course to bioinformatics methods. In particular, the course will be focusing on methods for sequence analysis and phylogenetics, genome assembly, SNP analysis, QTL mapping, and transcriptomic analyses.

### CLASS SCHEDULE

Wednesdays , 2:30PM - 5:30PM

### COURSE OUTLINE

1. Introduction
  - (a) Overview
  - (b) Programming Foundations
    - i. Introduction to Linux Command Line Interface and Scripting
    - ii. Quick Introduction to Programming Using R
2. Analyses and Methods
  - (a) Phylogenetic Analyses: Sequence Similarity and Phylogeny
    - i. Multiple Sequence Alignment
    - ii. Maximum Likelihood Tree Reconstruction
  - (b) Sequence Assembly
    - i. Reference-Based Assembly
    - ii. De Novo Assembly
  - (c) SNP Calling
  - (d) QTL Mapping
    - i. Single QTL Model
    - ii. Multiple QTL Model
  - (e) Transcriptomic Analysis
    - i. Microarray/Chip and RNA-seq/Next Gen Seq Transcriptomic Data Analysis
    - ii. Gene Coexpression Network Analysis

## GRADING SYSTEM

Requirements:

Problem Sets	40%
Project	60%

$92 \leq FG \leq 100$	1.00
$88 \leq FG < 92$	1.25
$84 \leq FG < 88$	1.50
$80 \leq FG < 84$	1.75
$76 \leq FG < 80$	2.00
$72 \leq FG < 76$	2.25
$68 \leq FG < 72$	2.50
$64 \leq FG < 68$	2.75
$59 < FG < 64$	3.00
$FG \leq 59$	5.00

## CLASS POLICIES

**CONSULTATIONS.** Schedule of the instructor's free times for the semester would be **Tuesdays, Thursdays, and Fridays, 8:00am-10:00am, 3:00-5:00pm**. Due to other commitments apart from teaching and to ensure proper queueing of multiple requests for consultations, students who wish to schedule a consultation must first send an e-mail to the e-mail address mentioned in the first part of the syllabus setting an appointment **at least 24 hours before** the requested date and time of consultation. Note however that a rescheduling of the consultation may be done subject to the availability of instructor.

**WEBSITE.** The link to the class website is <http://bit.ly/BIO397NiJan>. Announcements and certain materials will be posted there. Pertinent announcements will be posted **Tuesdays, 9:00pm at latest**. If no announcements are posted by that time, then classes are assumed to proceed as usual. In certain cases, announcement via e-mail (using the indicated e-mail address above).

**DELIVERABLES.** Submission of deliverables should be done on or before the designated deadlines (to be announced in future sessions). Late submissions would incur a 20% deduction per day late, and can incur a maximum of 100% deduction. Note though that even if a deliverable will no longer merit any points, it is considered a requirement, and non-submission of the deliverable on or before the deadline of submission of grades would mean a grade of INC assuming FG is passing.

**INTELLECTUAL DISHONESTY.** Intellectual dishonesty of any form will be subjected to proper disciplinary action as per Section V.2 of the Code of Student Conduct (download link posted in our website). If found guilty, results of the requirement/s in question will be nullified, final grade will be recomputed, and a corresponding change of grade will be filed if warranted.

**CLASS HOURS.** The official designated hours for the class is from 2:30pm until 5:30pm. However, to take into consideration possible activities of some students before and after class, the session will normally start at 2:50pm and will end at around 5:20pm. In the event of unforeseen cancellation of classes (e.g. natural disasters) and/or when needed, we will adjust the schedule of activities for a number of future sessions to maximize the allotted number of hours for the class to cover lost time and avoid holding additional make-up classes.

**DROPPING THE COURSE.** A student is considered dropped upon completion of the dropping procedure (see <https://crs.upd.edu.ph/downloads/Dropping%20Flowchart.pdf>). A student may officially drop the course **on or before 11 April 2019**.

## REFERENCES AND RESOURCES

- Sequence Alignment and Phylogenetics
  - Autenrieth F, Isralewitz B, Luthey-Schulten Z, Sethi A, Pogorelov T (2005). **Bioinformatics and Sequence Alignment**. <http://www.ks.uiuc.edu/Training/Tutorials/science/bioinformatics-tutorial/bioinformatics.pdf>
  - Mullan LJ (2002). **Multiple sequence alignment – the gateway to further analysis**. Briefings in Bioinformatics 3(3): 303-305.
  - Baldauf SL (2003). **Phylogeny for the faint of heart: a tutorial**. TRENDS in Genetics 19(6): 345-351.
  - Husmeier D (n.d.). **A Brief Tutorial in Phylogenetics** (slides). [http://www.bioss.ac.uk/~dirk/talks/tutorial\\_phylogenetics.pdf](http://www.bioss.ac.uk/~dirk/talks/tutorial_phylogenetics.pdf)
- Genome Assembly
  - Miller J, Koren S, Sutton G (2010). **Assembly algorithms for next-generation sequencing data**. Genomics, 95(6): 315-327.
  - Salberg SL, et al. (2012). **GAGE: A critical evaluation of genome assemblies and assembly algorithms**. Genome Res. 22(3): 557-567.
  - Schatz M (2012). **Whole Genome Assembly and Alignment** (slides). <http://schatzlab.cshl.edu/teaching/2012/CSHL.Sequencing/Whole%20Genome%20Assembly%20and%20Alignment.pdf>
  - Zimin A, et al. (2014). **Sequencing and Assembly of the 22-Gb Loblolly Pine Genome**. Genetics 196: 875-890.
- SNP Calling
  - Bioinformatics Team at The University of Texas. **Variant calling tutorial**. <https://wikis.utexas.edu/display/bioiteam/Variant+calling+tutorial>
  - Danecek P, et al. (2011). **The variant call format and VCFtools**. Bioinformatics 27(15): 2156-2158.
  - Hutchinson E (2016). **How To Analyze Your Single Nucleotide Polymorphism (SNP) Chip Data**. <http://bitesizebio.com/8223/how-to-analyze-your-single-nucleotide-polymorphism-snp-chip-data/>.

- Wondji CS, Hemingway J, Ranson H (2007). **Identification and analysis of Single Nucleotide Polymorphisms (SNPs) in the mosquito Anopheles funestus, malaria vector.** BMC Genomics 8(5). doi: 10.1186/1471-2164-8-5.
- QTL Mapping
  - Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005). **An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts.** Euphytica 142(1-2): 169-196.
- Transcriptomic Analysis
  - Babu MM (2006) **Chapter 11: An Introduction to Microarray Data Analysis.** In: Computational Genomics. Cambridge University Press, UK.
  - Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szczesniak MW, Gaffney DJ, Elo LL, Zhang X, Mortazavi A (2016). **A survey of best practices for RNA-seq data analysis.** Genome Biology 17:13. doi: 10.1186/s13059-016-0881-8.
  - Langfelder P, Horvath S (2008) **WGCNA: an R package for weighted correlation network analysis.** BMC Bioinformatics 2008, 9:559.