# MBIO 4700 Computational Molecular Microbiology Fall 2021

## (Previously offered as: MBIO 4030 Applied Bioinformatics Resources for Microbial Based Research)

The introduction to bioinformatics concepts achieved by applying computational tools to analyse sequence/molecular data relevant to Microbiology. The course will introduce existing bioinformatics tools that can be applied to biological data such as: sequence alignments, RNA folding, gene and protein structure prediction, molecular phylogenetics, and genomics.

**Prerequisite: <u>Grade of B in MBIO 3410</u>**

**Students enrolled in this course must ensure they satisfy the following minimum technological requirements:**

1. A computing device where one can create and edit documents,
2. An internet connection capable of streaming videos and downloading software, and
3. Access to a web-cam and microphone.

**Course material:** Lectures and all course material online (UMlearn). Course might use several platforms such as Zoom or Webex - a core set of lectures (75 min.) introducing the basic concepts (theory) of bioinformatics [~ 19 (75 min.) lectures] will be recorded, the remaining class times are dedicated to discussion/trouble shooting/problem solving (via Zoom). Lecture materials and relevant scientific publications will be uploaded on UMlearn. Guidelines on how to prepare presentations and format for the Final Report (on the Project) will be posted on UMlearn.

**The students will work on a project in <u>teams of three</u>; the project is of the teams choosing but needs approval by the instructor and has to be based on a hypothesis and contain DNA (e.g., Phylogenetic analysis), RNA (i.e., RNA folding) and protein (i.e., protein folding and comparative analysis) components. The more elements you can add the better. Essentially the course material can be applied to solve a <u>hypothesis</u> driven problem. The "Project" and the "Tool assignments" lets you explore topics of your interest that may not have been covered in detail in the lectures. We are in an era of "big data/meta data etc." it is important to have some experience in the basic tools available to start "chipping away" at the available data and how to extract information. This course is an introduction but hopefully a starting point to get you interested in this discipline and start exploring this part of science on your own.**

**Gaol: Learn practical approaches for the analysis of DNA, RNA and Protein sequences**

**Instructor and co-ordinator: Georg Hausner; TA support TBA**

**Evaluation (<u>for the online version; all items will be graded online and feedback will be provided via UMlearn</u>)**

1. **Team Interim "report"** (teams prepare 10 ppt. slides on their project (Proposal, rational, what are the questions to be addressed?) slides to be uploaded on UMlearn assignment folder #1)
**due October 14th/2021** (5 %)

2. Online **midterm** (during scheduled class time) based on material presented in the lecture material)
**October 21st/2021** (15 %)

3. **Individual “Tool Presentation”** 10 ppt slides plus short 3 page (double spaced) write up - on a bioinformatics tool not covered in the course. (to be uploaded on UMlearn - assignment folder #3)
**due November 18th/2021** (10 %)

4. **Linux assignment** (to be submitted to UMlearn - assignment folder #2)
**due December 3rd/2021** (10 %)

5. **Final report** (written in the format of a scientific paper).
(to be uploaded on UMlearn – assignment folder #4) **due December 10th/2021** (35 %)

6. Online **Final Exam** (2 hours) – cumulative – covering all lectures presented in the course.
(date and times to be set by Registrars Office) (25 %)

**NOTE:** **November 22nd/2021** Voluntary Withdrawal (**VW**) deadline Fall term classes

**Topics (and Activities):**

1. **DNA** (5 Lectures)

Sequence alignments (NCBI resources (BLAST etc.), Generating contigs, Alignments, Ribosomal Database Project, MAFFT, PRALINE GENEDOC, AliView)

Phylogeny (PHYLIP, MEGA, MrBAYES, BEAST)

Comparative analysis (Sequence Logos or WEBlogo, UGENE, concept of databases such as KEGG and others)

2. Concepts of Bayesian **statistics**, Markov and Hidden Markov Models, and Monte Carlo simulations as relevant to Bioinformatics. (1 Lecture)

3. **RNA** (folding; mfold and related programs; VIENNA RNA package) (2 Lecture)

4. **Proteins** (4 Lectures)

Structure/function: EMBOSS, Swissmodel (EXPASY), Phyre2, MISTIC, PyMOL, CHIMERA and other relevant programs).

*De novo* folding: Robetta/Rossetta (http://robetta.bakerlab.org/)

5. **Special Topics (Genomics)**

Genomes and annotation: Mobyle SNP workbench, **Galaxy**, CyVerse and Artemis, tools for the identification of mobile elements, phages etc. in bacterial genomes, annotation of organellar genomes (MFannot, RNAweasel). (~4 Lectures – at least 2 Lectures on Galaxy)

Introduction to Metagenomics (1 Lecture)

Introduction to **LINUX** (3 Lectures) (revisiting Galaxy)

**Online Activities:** Presentations and discussion plus problem solving (trouble shooting projects).

### Using Copyright material

**Please respect copyright.** We will use copyrighted content in this course. University guidelines state that copyrighted works, **including those created by instructors of the course are made available for private study and research and must not be distributed in any format without permission**. For more information, see the University's Copyright Office website at http://umanitoba.ca/copyright/ or contact um_copyright@umanitoba.ca .

## To get Started: "Build your basic tool box" - You need access to the internet and download several programs (all freeware for academic use).

| | | |
|---|---|---|
| 1. | Clustal-x | http://www.clustal.org/download/current/ (Win, Linux, or Mac OS version) |
| 2. | Genedoc | https://genedoc.software.informer.com/2.7/ (or AliView) |
| 3. | AliView | http://www.ormbunkar.se/aliview/<br>or http://www.ormbunkar.se/aliview/#DOWNLOAD |
| 4. | FigTree | http://tree.bio.ed.ac.uk/software/figtree/ (older version TreeView) |
| 5. | MrBayes | http://nbisweden.github.io/MrBayes/download.html |
| 6. | PyMol | https://pymol.org/2/ |
| 7. | MEGAX | http://megasoftware.net/ (Phylogenetics) |

**PLUS bookmark the following (more to come but for a start):**

**8. MISTIC (http://mistic.leloir.org.ar/index.php) Co-evolution (protein structure)**
**9. PHYRE2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index) Protein fold prediction**
**10. PRALINE (protein alignment tool) http://www.ibi.vu.nl/programs/pralinewww/ Alignment**
**11. MAFFT (https://mafft.cbrc.jp/alignment/server/) Alignment**
**12. GALAXY (https://usegalaxy.org/ or mirror sites in Europe or Australia) Genomics**
**https://usegalaxy.org.au/**
**https://usegalaxy.eu/ ***(most reliable site)**
**https://galaxy.pasteur.fr/**
**[Register for an account]**
**13. UNAfold web server (mFOLD) (http://www.unafold.org/mfold/applications/rna-folding-form.php) – RNA folding**
**14. SWISS Model (https://swissmodel.expasy.org/) Protein structures/folding**
**15. Robetta https://robetta.bakerlab.org/ Protein structures/folding**

## *Comments on individual "tool" Presentation Assignment:

Find a "tool" (computational methodology) – explain it in detail so that you can teach it to the class and note its relevance and importance (prepare ~ **10 power point slides and a short write up ~ 3 pages (double spaced)** to explain your "tool"). **AVOID BASIC tools**: such as NCBI or components within NCBI like BLAST etc. Write up should have in text citations and references in the format of the journal Genome (see further below NRC web page#).

Grade based on the quality of the power point slides/presentation (5 marks) and the short report (5 marks); complexity of the topic is also a consideration. Submit your ppt slides and report to UMlearn.

## *Comments on "Project" Assignment:

**The group has to prepare a manuscript style paper in the format as outlined for the Journal Genome. Please refer to the #NRC web page for the journal Genome:**

https://cdnsciencepub.com/journal/gen/authors#guidelines **(accessed July 26th/2021)**

**See example:** https://cdnsciencepub.com/doi/full/10.1139/gen-2020-0157

General comments: All manuscripts should contain a title page (Your name(s) and title), an abstract, followed by **Introduction**, **Materials and methods**, **Results**, **Discussion** plus **References, tables, figure captions, and appendices (for alignments or large data files)**, in that order. (See descriptions of each part of the manuscript, below.) See Genome (NRC) guidelines for preparing Tables and Figures. Figure captions can be on separate pages. **Submit your paper to UMlearn (GH will set up a folder - UMlearn).**

This assignment is worth 35 marks. Twenty-five (25) marks are assigned to the quality of the manuscript, following the guidelines provided (see link above), providing a sufficient literature review as part of the introduction, clearly stating the rational for the work and the outcome. Ten marks (10) are assigned for creativity, level of difficulty and incorporating tools not covered in the course.

**Additional Comments (the fine print):**

Note: No make-up midterms - missed work (assignments) will be assigned 0 marks, unless documentable reasons can be provided, for missed midterm exams the final exam mark will be adjusted to 40 %. Late assignments will NOT be accepted.

-> The Mid-term examination will be held (online) during the regular scheduled class period.

-> The Final examination will be comprehensive (i.e., covers all lectures), and will be scheduled by Student Records (Registrar's Office) during the December examination period. Permission to write a deferred final exam is granted by your faculty - the instructor is not involved in this process. If it is necessary for you to write your final exam at an alternate date, you must visit your faculty office with appropriate documentation to request permission for a deferred exam. This is a strict university policy, and there are no exceptions. If a deferral is granted it is your responsibility to contact the instructor immediately for the date of the deferred exam, missing the deferred exam will result in a grade of F.

All written answers will be graded based on quality of understanding, originality of thought, and clearness of presentation. Good writing skills certainly help!

Students requiring accommodations are directed to Student Accessibility Services to facilitate the implementation of accommodations. Course instructors are willing to meet with Students to discuss the accommodations recommended by Student Accessibility Services.

All work is to be completed independently unless otherwise specified. Please remember that group projects are subject to the rules of academic dishonesty and every group member must ensure that a group project adheres to the principles of academic integrity.

**Academic dishonesty** guidelines are stated in your calendar regarding University of Manitoba policies with respect to academic dishonesty (particularly plagiarism and cheating) and behaviour and absence from final exams. All work is to be completed independently unless otherwise specified. Please remember that group projects are subject to the rules of academic dishonesty and every group member must ensure that a group project adheres to the principles of academic integrity.

The Faculty of Science web page has detailed information: https://www.sci.umanitoba.ca/undergraduate-students/academic-resources/academic-integrity-2/. Please read and follow these guidelines, and ask if you have any questions.

**For information:**

**If you experience depression, anxiety, or other health or stress related issues – you are not alone - please consider the following resources:**

**Student Counselling Centre**
474 University Centre
University of Manitoba, Winnipeg, MB R3T 2N2 Canada
Phone: 204 474-8592 Fax: 204 474-7558
http://umanitoba.ca/student/counselling/

**CMHA Manitoba and Winnipeg**
930 Portage Avenue, Winnipeg MB R3G 0P
E-mail: office@cmhawpg.mb.ca
https://mbwpg.cmha.ca/mental-health-resources-for-winnipeg/

**Klinic Community Health**
870 Portage Avenue
Winnipeg, MB, R3G 0P1
Phone: (204) 784-4090
Admin Fax: (204) 772-7998
Medical Fax: (204) 784-4013
http://klinic.mb.ca

**First Nations and Inuit Hope for Wellness Help Line**
1-855-242-3310
Counselling available in English and French - upon request, in Cree, Ojibway, and Inuktut
Crisis Response Centre
817 Bannatyne, Winnipeg; attend in person

**Jack.org**
General information about student mental health, useful for sharing with friends and start the conversation!
https://jack.org/Home
Phone: 416-425-2494

## Urgent Help

**University Security Services (24 hs)** #555 (from MTS or Roger wireless)
**On Campus Suicide Crisis Klinic (24 hs)** 4-(204) 986-6222
**Adult Mobile Crisis Service** 204-940-1781
**Crisis Stabilization Unit** 204-940-3633
**Crisis services Canada** http://www.crisisservicescanada.ca/en/