

A large, abstract network graph composed of numerous small, semi-transparent blue circles connected by thin, dark blue lines. The graph is set against a light blue background and occupies the left half of the slide.

# **Biochem 3BP3**

**Practical Bioinformatics in the  
Genomics Era**

**Fall 2021**

---

Instructor:  
Dr. Jennifer Stearns



- BSc in Biology, University of Waterloo
- PhD in Biology – Plant-microbe interactions (UW)
- Post-Doc at UW working on the human microbiome and new methods for Illumina sequencing of 16S
- Post-Doc at McMaster University working on the human airway microbiome
- Assistant Professor in Medicine at McMaster University
- Farncombe Chair in Microbial Ecology and Bioinformatics
- Study microbial community succession in the human gut and bacterial interactions in the infant gut microbiota

## Teaching Assistants

### Jalees Nasir



- PhD student in the McArthur lab
- BSc in Chemical Biology (McMaster)
- Genomic surveillance of respiratory viruses

### Shahrokh Shekarriz



- PhD student in the Surettelab
- BSc in Molecular Biology (University of Tehran)
- MSc in x with Brian Golding (McMaster University)
- Computational microbiology with main focus on gut microbiota



# **Practical**

# **Bioinformatics**

# **Genomics**



# **Practical**

## **Bioinformatics**

**Biological Data | Computer Science | Math | Engineering | Statistics**

## **Genomics**

# **Practical**

## **Bioinformatics**

**Biological Data | Computer Science | Math | Engineering | Statistics**

## **Genomics**

**DNA | RNA | Sequencing Data | Genomes**

---

# **Practical**

**Problem-Based Learning | Real Experimental Data | Hands-On**

## **Bioinformatics**

**Biological Data | Computer Science | Math | Engineering | Statistics**

## **Genomics**

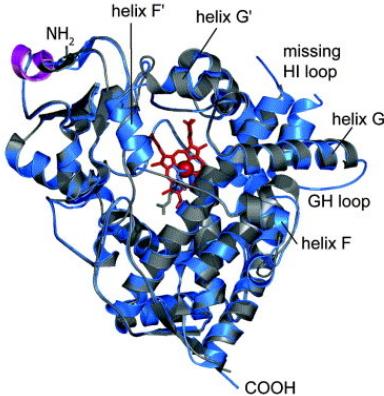
**DNA | RNA | Sequencing Data | Genomes**

---

# For Example...

# Two weeks from now...

**Annotation of the toxicological defensome of a Shark**

# For Example...

# After reading break...

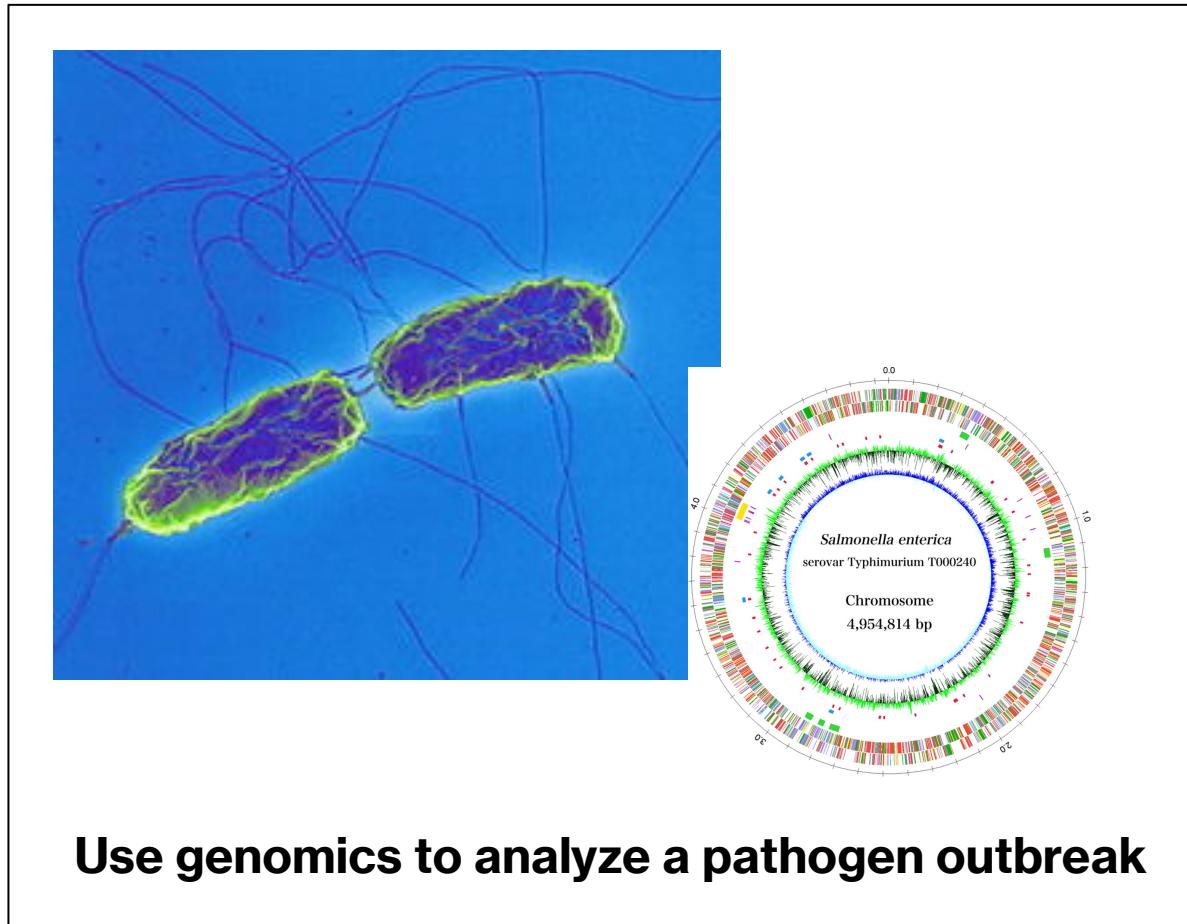
```
Terminal
. gb2genbank NP_058781 | head -n 10
LOCUS NP_058781          503 aa      linear ROD 10-AUG-2014
DEFINITION aromatase [Rattus norvegicus].
ACCESSION NP_058781
VERSION NP_058781.2 GT-281182626
DBSOURCE RefSeq; accession NM_017085.2
KEYWORDS RefSeq;
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
. gb2fasta NP_058781
>gi|281182626|Ref|NP_058781.2| aromatase [Rattus norvegicus]
MFLEMMLNPMHNYNTIMVFETVPSAMPILLIMGLLIRNCESSSITPGPGYCLGIGPLISHGRFLWMGI
GSACNYNKMKGFFMRWISGEETLIIKSSSMMVHKHSNYISRGFSKRGQLCIGHENGIIFFNNNPSL
WRTIVRPFFMKALTGPGLIRMVEYCYESIKOHLRLGQDVTNSGYVQDVVTLMRHIMLDTSNTLFLGTPLDE
SSVIKKIOGYFNAWQALLIKPNNFKISWLYRKYERSVKDLDEIEILVKEKROKVYSAEKLDCMDPFT
DLTFAERRGOLTKENVNQCTLEMLIAAPDTMSVTLYVMLLIAEYFEVETAILKEIHTVGDORDIRIGDV
QNLKVVENFINESLRYQFVQDLYMRALEDDVIDGYVKKGTNIILNIGRMHRLEYPKPNFETLENFEK
NVPYRYQFQFGFGFRSCAGKYIAMMMMKVLYTLLKRHVKTLLQKRCIENMPKNNNDLSLHLDDEDSPIVEI
IFSPRNSEKYLKQ

. pubmed 23643682
PMID 23643682
JOURNAL Bone
YEAR 2013
VOLUME 55
ISSUE 2
PAGES 309-314
TITLE Genetic polymorphism at Val80 (rs700518) of the CYP19A1 gene is associated with aromatase inhibitor associated bone loss in women with ER+ breast cancer
MINIREF Napoli N, et al. 2013. Bone 55(2): 309-314.
UNIQUEREF Napoli N, et al. 2013. Genetic polymorphism at Val80 (rs700518) of the CYP19A1 gene is associated with aromatase inhibitor associated bone loss in women with ER+ breast cancer. Bone 55(2): 309-314.
ABSTRACT PURPOSE: Polymorphisms in the CYP19A1 (aromatase) gene have been reported to influence disease-free survival and the incidence of musculoskeletal complaints in patients taking aromatase inhibitors (AIs) for estrogen receptor positive (ER+) breast cancer. Bone loss and fractures are well-recognized complications from AI therapy. The objective of this study is to determine the influence of polymorphisms in the CYP19A1 gene on bone loss among patients taking aromatase inhibitors for ER+ breast cancer. PATIENTS AND METHODS: The subjects consisted of 97 postmenopausal women with ER+ breast cancer who were initiated on third-generation AIs. Bone mineral density (BMD) was measured by dual energy X-ray absorptiometry at baseline and at 6 and 12 months. Twenty-four hour urine N-telopeptide (NTX) was measured by Elisa and serum estradiol was measured by ultrasensitive radioimmunoassay at baseline, a nd at 6 months. Genotyping was done by Tagman SNP allelic discrimination assay. RESULTS: Women with the AA genotype for the rs700518 (G/A at Val(80)) developed significant bone loss at the lumbar spine and the total hip at 12 months relative to patients carrying the G allele (GA/GG); both p = 0.03. There was a borderline greater increase in urinary NTX in those with the AA genotype compared to patients with the G allele, p = 0.05; but no significant difference in changes in estradiol levels among the genotypes. CONCLUSION: Patients with the AA genotype for the rs700518 polymorphism in the CYP19A1 gene are at risk for AI-associated bone loss and deserve close follow-up during long-term AI therapy
```

Process sequencing data using the command line

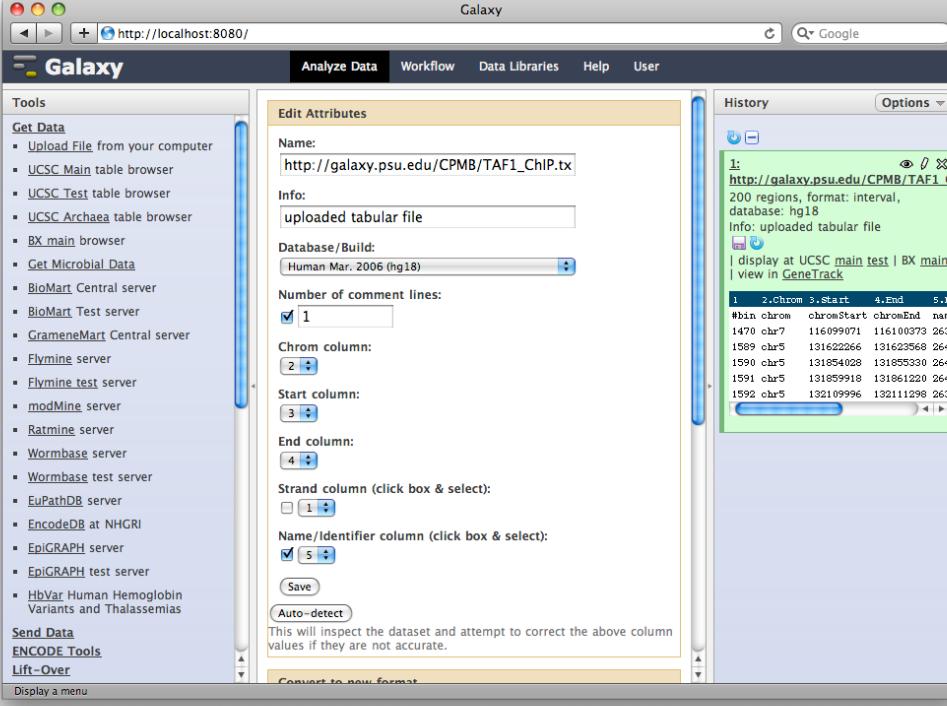
# For Example...

## Early November...



# For Example...

# By the end of November...



The screenshot shows the Galaxy web interface version 0.6.0. The main window has a title bar "Galaxy" and a URL "http://localhost:8080/". The top menu includes "Analyze Data", "Workflow", "Data Libraries", "Help", and "User". The left sidebar is titled "Tools" and lists various UCSC table browsers, BX main browser, Get Microbial Data, BioMart Central server, and several Flymine, Ratmine, Wormbase, and EuPathDB servers. The central panel is titled "Edit Attributes" for a dataset named "http://galaxy.psu.edu/CPMB/TAF1\_ChIP.tx". It contains fields for "Info" (uploaded tabular file), "Database/Build" (Human Mar. 2006 (hg18)), and column definitions for "Chrom", "Start", "End", "Strand", and "Name/Identifier". A "Save" button is at the bottom. The right panel is titled "History" and shows a list of datasets. The first dataset is "1: http://galaxy.psu.edu/CPMB/TAF1\_ChIP.tx", which is a tabular file with 200 regions, format: interval, database: hg18. Below it is a "UCSC main test | BX main" entry. The bottom section of the history panel displays the first few lines of the tabular data:

1	2.Chrom	3.Start	4.End	5.Name
1470	chr7	116099071	116100373	263
1589	chr5	131622266	131623568	264
1590	chr5	131854028	131855330	264
1591	chr5	131859918	131861220	264
1592	chr5	132109996	132111298	263

**Use cloud computing to analyze gene expression data**

# Online Learning...

## Research highlights



## 9 hands-on labs:

- Genome databases
- Phylogeny
- Analysis of predicted proteins
- Sequencing
- Genome assembly
- Gene expression analysis

Human Genetics Week  
Pandemic Week

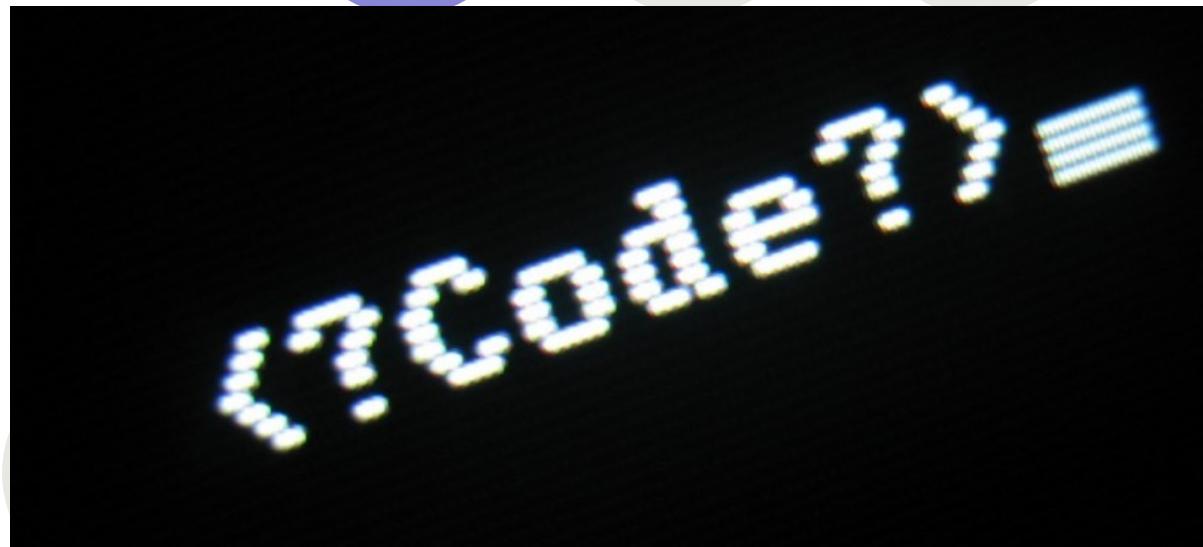
Analytics Guest Speakers

# No Computer Science Experience Required

Registered students are from Biomedical Discovery & Commercialization, Biochemistry, iSci, Molecular Biology & Genetics, Medical & Biological Physics

The course goal is to teach you the theory and practice of analysis of real world research data in the biological / biomedical sciences.

The course will provide an introduction to Bioinformatics plus expand your research toolkit.





# Course Structure

## Electronic Resources

- Avenue to Learn:
  - Assignments, quizzes and grading
  - Flash updates will be uploaded here
- Tutorials and links to lecture .ppt: <https://github.com/jstearns/Biochem-3BP3>
- MS Teams: for asynchronous communication and videoconferencing
  - Course Schedule & Resources – links to all recorded content
  - Flash Update Presentation Schedule

---

# Course Structure

## Weekly Modules

- Traditional lecture to introduce the topic and outline the key objectives
- Lectures will be recorded in advance of each module
- Teams Word document will provide links to lectures and supplementary presentations.
- Please attend Wednesday 12:30 pm sessions for live updates
- All Live content will be recorded for later access

# Course Structure – MS Teams



BIOCHEM 3BP3 C01 FAL  
2021 Practical...

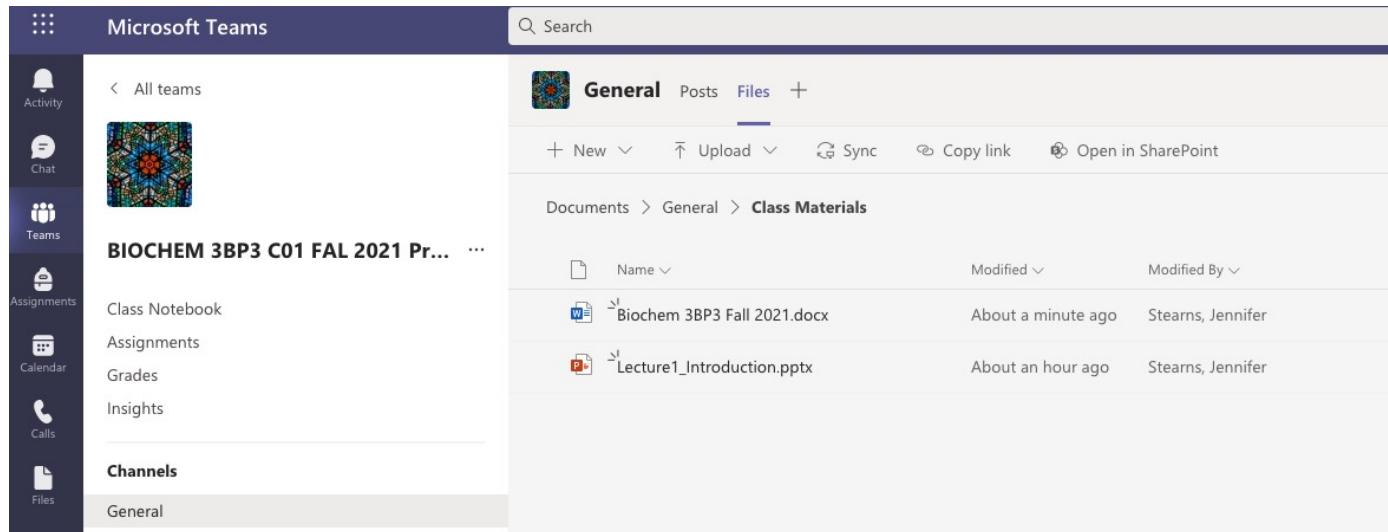


BIOCHEM 3BP3 T01 FAL  
2021 Practical...



BIOCHEM 3BP3 T02 FAL  
2021 Practical...

# Course Structure – MS Teams



Microsoft Teams

Activity Chat Teams Assignments Calendar Calls Files

All teams

BIOCHEM 3BP3 C01 FAL 2021 Pr... Class Notebook Assignments Grades Insights

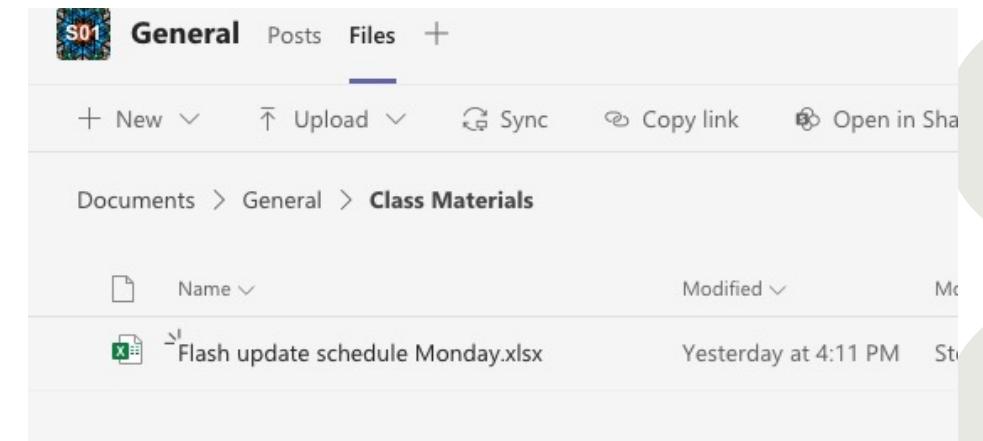
General

Posts Files +

+ New Upload Sync Copy link Open in SharePoint

Documents > General > Class Materials

Name	Modified	Modified By
Biochem 3BP3 Fall 2021.docx	About a minute ago	Stearns, Jennifer
Lecture1_Introduction.pptx	About an hour ago	Stearns, Jennifer



s01 General Posts Files +

+ New Upload Sync Copy link Open in SharePoint

Documents > General > Class Materials

Name	Modified	Modified By
Flash update schedule Monday.xlsx	Yesterday at 4:11 PM	Stearns, Jennifer

# Course Structure – Files in Teams

## Biochemistry 3BP3: Practical Bioinformatics in the Genomics Era

### Course Schedule & Resources Fall 2021

*This is a living document, content and links to recorded material will be updated frequently.*

#### COURSE INFORMATION

##### Instructor

Jennifer C. Stearns, [stearns@mcmaster.ca](mailto:stearns@mcmaster.ca)

Virtual Office Hours: Tues and Thurs 9 - 10 am, [https://teams.microsoft.com/l/meetup-join/19%3a3glwN0H\\_RQW2WYKZm3TrGdbAFT-qKcRFP-49ELPzzM1%40thread.tacv2/1630608185668?context=%7b%22Tid%22%3a%2244376307-b429-42ad-8c25-28cd496f4772%22%2c%22Oid%22%3a%229e61ee6-ed06-47ab-8084-bbfd60626551%22%7d](https://teams.microsoft.com/l/meetup-join/19%3a3glwN0H_RQW2WYKZm3TrGdbAFT-qKcRFP-49ELPzzM1%40thread.tacv2/1630608185668?context=%7b%22Tid%22%3a%2244376307-b429-42ad-8c25-28cd496f4772%22%2c%22Oid%22%3a%229e61ee6-ed06-47ab-8084-bbfd60626551%22%7d)

##### Teaching Assistants

Jalees Nasir, [nasirja@mcmaster.ca](mailto:nasirja@mcmaster.ca)

Shahrokh ~~Shekarriz~~, [shekas3@mcmaster.ca](mailto:shekas3@mcmaster.ca)

*Virtual Office Hours by appointment only.*

##### Electronic Resources

- Avenue to Learn for submitting assessments and grading, <http://a2l.mcmaster.ca>
- Tutorial content can be found at <https://github.com/jstearns/Biochem-3BP3>
- MS Teams for asynchronous communication and hosting videoconferencing

# Course Structure – Files in Teams

## WEEK 1 (September 8) - INTRODUCTORY SESSIONS

LIVE Lecture during Wednesday lecture session: Introduction to the Course

Recorded Content:

- Research Focus - Dr. Joanna Wilson, Department of Biology, McMaster University,  
<https://web.microsoftstream.com/video/d8efccec-57d6-4439-ba13-2d5bff4f365b>

No Tutorial

## WEEK 2 (SEPTEMBER 13 and 15) - COMPUTATION & GENOME DATABASES

LIVE Lecture during Wednesday lecture session

- Mark Hahn of SHARCNET - Overview of High-Performance Computation,

Recorded Content

- Overview & Demo of Laboratory #1 - Genome Databases,  
<https://web.microsoftstream.com/video/00da3224-9a5e-40a9-91d9-ad78dfb838f7>

Tutorial

- LIVE session with Teaching Assistants
- Tutorial content can be found at GitHub, answers due on A2L

No Flash Updates

# Course Structure - GitHub

## Lab # 1 - Introduction to the Laboratory & Genome Databases

### Table of Contents

1. [Introduction](#)
2. [Moore's and Kryder's Laws](#)
3. [National Center for Biotechnology Information](#)
4. [PubMed](#)
5. [GenBank / Entrez](#)
6. [BLAST](#)
7. [Ensembl](#)

### Introduction

The goal of this lab is to introduce two key bioinformatics databases – GenBank & Ensembl – plus a few additional online resources.

#### Lectures - [Introduction to Bioinformatics & the Course](#)

#### Background Reading (optional)

- J. Chang. 2015. Core services: Reward bioinformaticians. [Nature 520:151-2](#)
- Goodman et al. 2014. Ten simple rules for the care and feeding of scientific data. [PLoS Comput Biol. 10\(4\):e1003542](#)
- Burge et al. 2012. Biocurators and biocuration: surveying the 21st century challenges. [Database Mar 20:bar059](#)
- The GenBank resource and the variety of tools and data it offers [Nucleic Acids Res. 2019 Jan 8;47\(D1\):D23-D28](#) & [Nucleic Acids Res. 2019 Jan 8;47\(D1\):D94-D99](#).
- Ensemble resource and the variety of tools and data it offers [Nucleic Acids Res. 2019 Jan 8;47\(D1\):D745-D751](#).
- Growth of Sequencing Data and predicted growth [Nucleic Acids Res. 2019 Jan 8;47\(D1\):D94-D99](#),  
<http://www.ncbi.nlm.nih.gov/genbank/statistics>, <http://www.genome.gov/sequencingcosts/>, and [Science 2011 331:728-9](#).

#### Links

- NCBI & GenBank, <http://www.ncbi.nlm.nih.gov>
- Ensembl, <http://www.ensembl.org>

#### Computer Resources

- You can complete this entire lab by using your web browser

# Course Structure – A2L

## Avenue to Learn

- Documents needed to complete each tutorial
- Assignments and marking rubrics
- Presentation marking rubric
- Quizzes

## Biochemistry 3BP3: Practical Bioinformatics in the Genomics Era

### Lab # 1 - Introduction to the Laboratory & Genome Databases

Lab Available on GitHub: <https://github.com/agmcarthur/biochem3BP3>

Student Name:

**Question #1.** Take a look at [www.submarinecablemap.com](http://www.submarinecablemap.com). How many submarine cable landing points does Japan have (excluding Okinawa)?

*<insert answer>*

**Question #2.** How many cable links are between Russia and Japan?

*<insert answer>*

**Question #3.** If you were accessing the deCODE database in Iceland ([www.decode.com](http://www.decode.com)), name three countries that the data could traverse to get to you here in Canada.

*<insert answer>*

**Question #4.** Undersea cables can be broken by dragging ships anchors or ill-positioned fishing trawlers. If you were doing your PhD on the genomics of the endemic ranid frog *Platymantis* in French Polynesia, how many undersea cables

# Course Structure

## Weekly Modules – Tutorial sessions Mondays & Wednesdays

- Tutorials will start **LIVE** at the tutorial session start time
  - 3 students give a 10 minute “Flash Update” presentation each week beginning on **Sept 20th**
- Live sessions will be recorded for later access
- Tutorials: <https://github.com/jstearns/Biochem-3BP3>
- A core take-home assignment due on Avenue to Learn that will include:
  - Answers to tutorial problems
  - multiple choice questions on the lecture and the flash presentations

# Flash Updates

- See the Flash Update Presentation Schedule in Team files for each tutorial session
  - A Flash Update is a 10 minute PowerPoint presentation summarizing the key points of the assigned topic
  - Must include 3 <https://kahoot.com> questions!
  - PowerPoint file must be uploaded to Avenue to Learn by start of lab on Mondays & Wednesdays!
-

# Essay Assignments

## Critical Review – Due October 27, 2021

*This is a critical review exercise, worth 25% of the total course grade. Please follow the guidelines provided in the grading rubric and use the template WORD file provided.*

***Excluding references, the Critical Review cannot exceed 2 pages in length.***

You are being asked to review a pre-publication manuscript submitted to [www.biorxiv.org](http://www.biorxiv.org), an open access preprint repository for the biological sciences. Papers in the bioRxiv have generally not undergone peer review and thus are not considered formal publications. A pre-print at bioRxiv may latter appear as a publication in a scientific journal after peer-review.

## Reflective Exercise – Due December 8, 2021

*This is a reflective exercise, worth 15% of the total course grade. Please follow the guidelines provided in the grading rubric and use the template WORD file provided.*

***Excluding references, the Reflective Essay cannot exceed 1 page in length.***

You are being asked to write in the style of a CIHR Bioinformatics Graduate School Fellowship Application. You are competing with hundreds for this prestigious funding that will allow you the pick of graduate schools across Canada or internationally.

Questions to answer in your essay:

- If you were to build a bioinformatics career, what research questions would

# Grading

Course Outline & Marking Rubrics are on Avenue to Learn:

Item graded	% of final grade	Due date
Lab assignments (7)	35%	weekly
10 min presentation	15%	varies
Lecture quizzes (2)	10%	varies
Critical review	25%	October 27, 2021
Essay assignment	15%	December 8, 2021

- No textbook – primary literature only
- TAs are available by appointment – see them first!
- All work submitted via Avenue to Learn
- Late penalty of 10% per day (grade of zero after 4 days)
- Use McMaster Student Absence Form for absence less than 3 days
- No exams!



# A ‘Living’ Course

- Each term – evolving lectures, data, tools
- Problem-Based Learning, Inverted Teaching Style
- Real Experimental Data
- Not everything is going to go perfectly
- Experimental data is messy – there may not be an “answer”
- The schedule & tutorial details are tentative

# Being Successful

- Attend all live sessions, watch all recorded sessions
- Ask questions and use your Teaching Assistant
- Come to office hours or schedule a time to talk to me if things aren't going well
- “Take chances, make mistakes, get messy” – Miss Frizzle



# A few things...

**avenue** to learn

Avenue to Learn is central to the course – handing in assignments, schedules, Flash Update powerpoints, Lecture powerpoints, lab assignment downloads, etc.

**PubMed**

All readings can be found on PubMed via a McMaster IP address or via the McMaster LibAccess portal if off campus: <https://libraryssl.lib.mcmaster.ca> (E-Journals)



Some aspects of the course will require use of McMaster's Virtual Private Network:  
<http://www.mcmaster.ca/uts/network/vpn/>

# Using PubMed

The screenshot shows the PubMed search results for PMID 25855439. The search bar at the top contains the identifier "25855439[uid]". The main content area displays the abstract of the article published in Nature, April 2015. The abstract text is as follows:

Nature. 2015 Apr 9;520(7546):151-2. doi: 10.1038/520151a.  
**Core services: Reward bioinformaticians.**  
Chang J<sup>1</sup>.  
Author information  
PMID: 25855439 [PubMed - indexed for MEDLINE]  
[Facebook](#) [Twitter](#) [Google+](#)

Below the abstract are sections for MeSH Terms and LinkOut - more resources. At the bottom left is a link to PubMed Commons, which has 0 comments. On the right side, there is a "Send to:" section with a "Full text links" button, which is circled in red. This button leads to a "nature" page. There are also "Save items" and "Add to Favorites" options. The "Similar articles" section lists several related publications, including "Perspectives on an education in computational biology and medicine." by Chang J (Yale J Biol Med. 2012), "Computing: A vision for data science." by Nature (Nature. 2013), "Recruiters and industry. Microsoft's European perspective." by Nature (Nature. 2005), and two "Review" articles: "Effectiveness, quality and transfer: further development" (Dtsch Med Wochenschr. 2011) and "An interdisciplinary and interactive online tool to..." (Evid Fortbild Qual Gesundhw...).

# Using PubMed

nature.com : Sitemap      Login : Register

ADVERTISEMENT      Or register with an existing account      ADVERTISEMENT

ORCID   Natureasia.com   OpenID   G+   f   in   tw

**nature INDEX**

Search   Go   Advanced search

**nature** International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

Archive > Volume 520 > Issue 7546 > Comment > Article

NATURE | COMMENT

Core services: Reward bioinformaticians

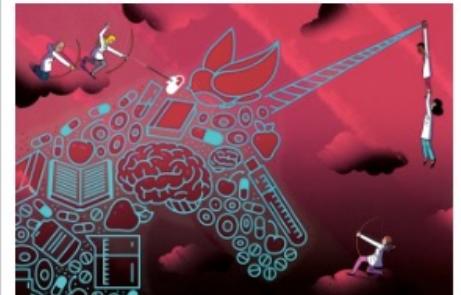
Jeffrey Chang

08 April 2015

Biological data will continue to pile up unless those who analyse it are recognized as creative collaborators in need of career paths, says Jeffrey Chang.

 PDF    Rights & Permissions

Subject terms: Computational biology and bioinformatics · Research management · Careers

  
Mythical beasts  
The science myths that will not die  
False beliefs and wishful thinking about the human experience are common. They are hurting people — and holding back science.

# Using PubMed



OFF-CAMPUS ACCESS: LOGIN TO E-RESOURCES - MCMASTER LIBRARIES  
Library

The image shows a login interface for McMaster University's library resources. At the top, it says "OFF-CAMPUS ACCESS: LOGIN TO E-RESOURCES - MCMASTER LIBRARIES" and "Library". Below this, the main title is "Login to access Library e-Resources". A note " McMaster Users Only" is present. There are two input fields: "Username:" and "Password:", each with a corresponding text input box. Below the password field is a "Log in" button.

<https://libraryssl.lib.mcmaster.ca>

# Using PubMed

The screenshot shows the University Library website's "Off-campus Access" page. At the top, there is a navigation bar with links for Home, Service, Collections, Spaces, Research Help, About, and a search bar labeled "Search Resources". Below the navigation bar, there is a yellow banner with the text "COVID-19: The University Library services continue online. Read the [Library's Response to COVID-19](#)". On the left side, there is a sidebar with links for Archives & Rare Books, Catalogue, Databases, Data / Statistics, Digital Archive, and Manuscripts. The main content area has a heading "Off-campus Access" and a sub-section titled "Access". It contains text explaining how to find resources and log in with a MacID. A red circle highlights the "SEE ALL COLLECTIONS →" button, which is located below the sidebar. To the right of the main content, there is a button labeled "Report Off-campus access problem →". Below the main content, there is a section titled "Off-campus Access Overview" with a "Expand/Collapse All" button and three items: "Directly from the Library website", "Install a browser bookmarklet", and "Configure Google Scholar", each preceded by a plus sign icon.

COVID-19: The University Library services continue online. Read the [Library's Response to COVID-19](#).

Archives & Rare Books

Catalogue

Databases

Data / Statistics

Digital Archive

Manuscripts

**SEE ALL COLLECTIONS →**

Off-campus Access Overview

Expand/Collapse All

+ Directly from the Library website

+ Install a browser bookmarklet

+ Configure Google Scholar

# Using PubMed

The screenshot shows the McMaster University Library website. At the top left is the McMaster University logo. To its right is the word "Library" in a large, bold, maroon font. The top navigation bar is dark maroon with white text, featuring links for Home, Services, Collections, Spaces, Research Help, About, a search bar labeled "Search Resources", and a "Mobile" icon. A yellow banner below the navigation bar contains the text: "COVID-19: The University Library buildings are closed but library services continue online. Read the [Library's Response to COVID-19](#)." The main content area has a blue-toned background image of people in a library. It features four white rectangular buttons: "Search", "Databases", "Online Journals" (which is highlighted in maroon), and "Research Guides". Below these buttons is a search bar with the placeholder text "Find online full-text journals, magazines and news" and a magnifying glass icon. Underneath the search bar is a link to "Library Catalogue". At the bottom of the page, there is a horizontal navigation bar with "Home" and "Collections" on the left, and "Are we missing something? Request a book purchase." on the right.

## Collections

Are we missing something? Request a book purchase.

# Using PubMed

## Find e-Journal

Titles      Subject      Locate      CitationLinker

[0-9](#) [A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#) [Others](#)

Title:   Starts with  Contains  Exact

GO

Total number of e-Journals: 165984

[Switch to Table View](#)

Showing page 5518 of 8300 pages.

|< [<< Previous](#) [Next >](#) >|

[Nature](#) [0028-0836] 

Full Text Online [Gale Academic OneFile](#)

Available from 2000

Most recent 1 year(s) not available

Full Text Online [Gale OneFile: CPI.Q](#)

Available from 2000/01/06

Most recent 1 year(s) not available

Full Text Online [Nature Online Journals](#)

Available from 1869 volume: 1 issue: 1

Full Text Online [Nursing & Allied Health Database](#)

Available from 1990/01/04

Most recent 1 year(s) not available

Full Text Online [Publisher's Website \(Free\)](#)

Available from 1869 volume: 1 issue: 1 until (and including) 1875 volume: 12 issue: 313



Chemistry: General and Others

Earth Sciences: General and Others

Environmental Sciences: General and Others

Life Sciences: Biology

Physics: Astronomy

Physics: General and Others

Absorbed: Nature Physical Science [0300-8746]

Absorbed in part by: Nature New Biology [0090-0028]



# Using PubMed

Source: NATURE [0028-0836]

## Basic



[Full Text Online](#) Nature Online Journals

Available from 1869 volume: 1 issue: 1

[Terms of Use](#)



[Full Text Online](#) Nursing & Allied Health Database

Available from 1990/01/04

Most recent 1 year(s) not available



[Full Text Online](#) Publisher's Website (Free)

Available from 1869 volume: 1 issue: 1 until 1875 volume: 12 issue: 3



[Check for a print copy](#) in our library catalogue

● [Request document on Interlibrary Loan via RACER Scholars Portal](#)

Note: You have to have registered in RACER previously in order to u

● [Broken links? Report problems](#) to librarians

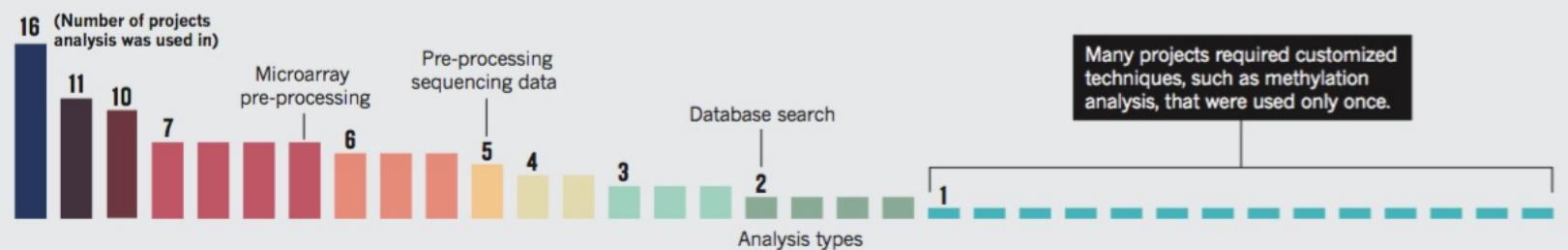


# Bioinformatics

- Application of mathematical, statistical, and computational approaches to biological data
- Modern emphasis (and emphasis of this course) is upon genomic data, dominated by DNA sequencing.
- Yet, bioinformatics is very broad & diverse, for example:
  - chemoinformatics
  - protein structure, prediction of ligand binding
  - whole cell simulation
  - evolutionary biology
  - assay development
  - analytics and machine learning

## ROUTINELY UNIQUE

Over 18 months, 46 data-analysis projects undertaken at the bioinformatics core of the University of Texas Health Science Center at Houston required 34 different types of analysis — most were used infrequently. Each project demanded unique combinations of analyses, demonstrating how bioinformaticians must be versatile, creative and collaborative.



# Bioinformatics

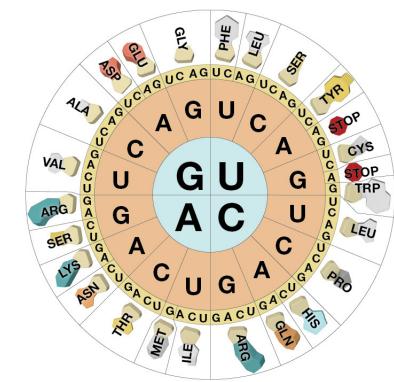
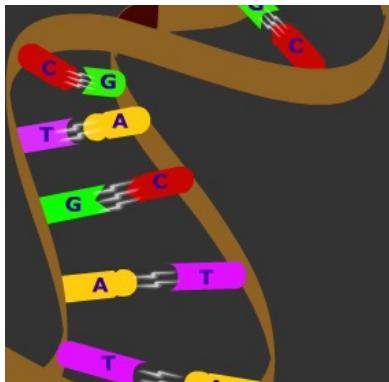
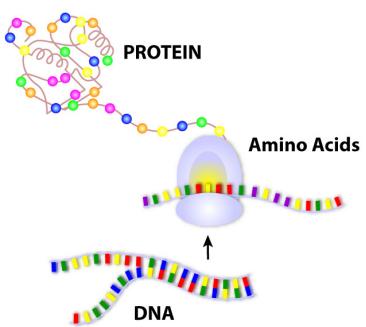
- Bioinformatics can be...
  - Hypothesis testing using experimentally generated data
  - Data curation and knowledge integration
  - Data mining as part of experimental design
- There are no ‘one size fits all’ analyses
- “Pure” bioinformatics is rare – new algorithm or data schema development
- Most bioinformatics is “applied” and highly collaborative
- Bioinformaticians need to understand the biology and lab work

# Bioinformatics

- This course will survey a number of key applied bioinformatics questions, with an introduction to the underlying theory
  - The course emphasizes breadth instead of depth
  - Go deeper...
    - Bioinformatics 4th year thesis project
    - Co-Op in a Bioinformatics lab
    - Graduate school
-

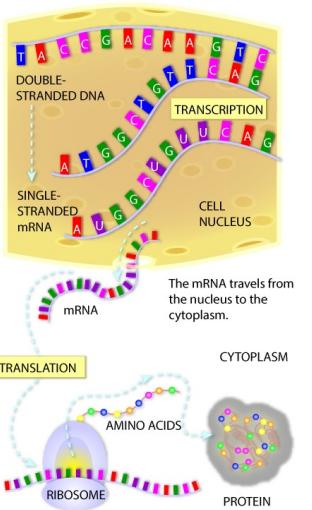
# Next Week's Lab focuses on Databases

# Central Dogma



**TRANSCRIPTION:** In the nucleus, the cell's machinery copies the gene sequence into messenger RNA (mRNA), a molecule that is similar to DNA. Like DNA, mRNA has four nucleotide bases - but in mRNA, the base uracil (U) replaces thymine (T).

**TRANSLATION:** The protein-making machinery, called the ribosome, reads the mRNA sequence and translates it into the amino acid sequence of the protein. The ribosome starts at the sequence AUG, then reads three nucleotides at a time. Each three-nucleotide codon specifies a particular amino acid. The "stop" codons (UAA, UAG and UGA) tell the ribosome that the protein is complete.

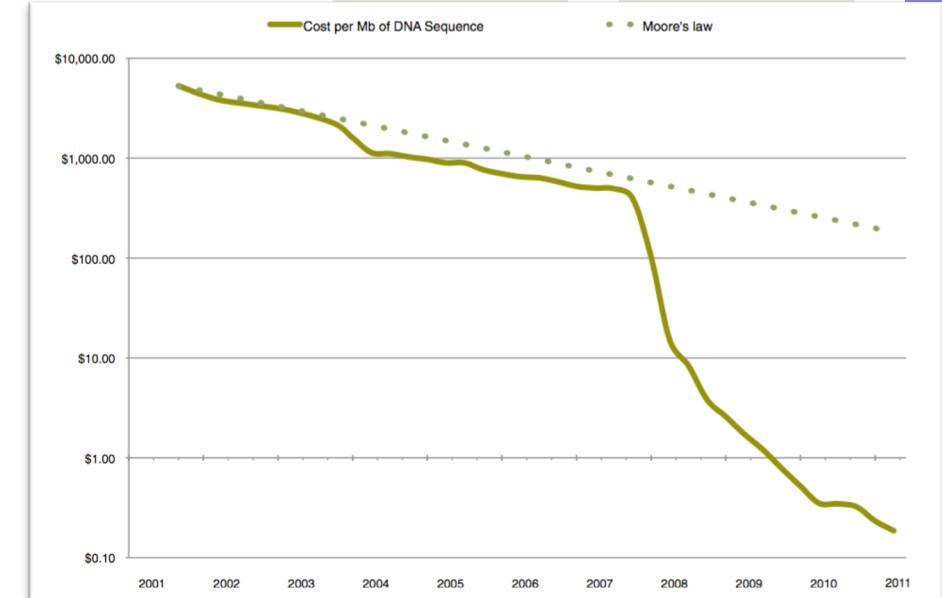


Images from "Genetic Science Learning Center, University of Utah, <http://learn.genetics.utah.edu>."

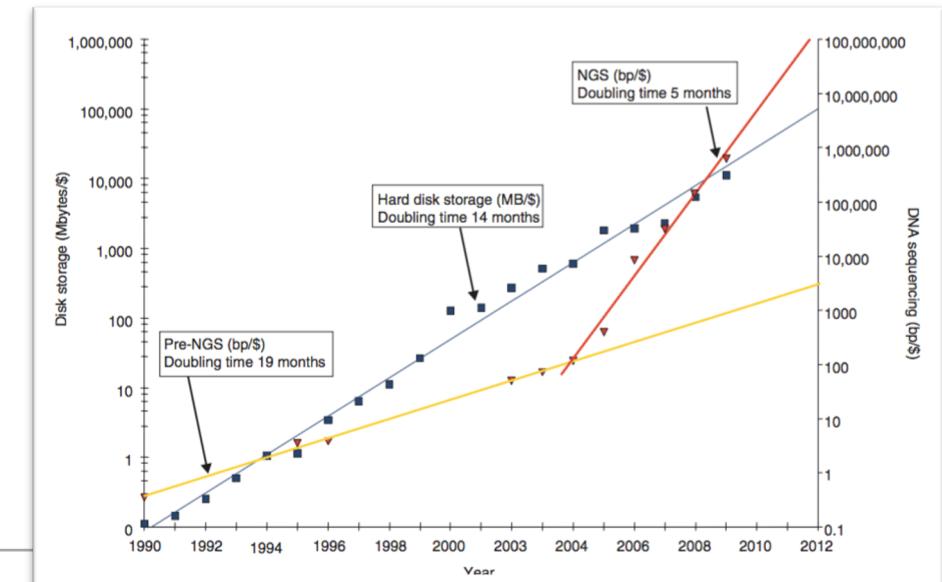
# Next Week's Lab focuses on Databases

- Moore's Law (computer processor speed doubles every 18 months) and Kryder's Law (disk storage capacity doubles every 14 months) are critical to the future of biological research, see <http://www.scientificamerican.com/article/kryders-law/>
- Will advances in DNA sequencing outstrip Moore's and Kryder's law?

NGS



Sboner et al. 2011. The real cost of sequencing: higher than you think! *Genome Biol.* 12(8):125.



Stein. 2010. The case for cloud computing in genome informatics. *Genome Biol.* 11(5):207.