

BIOCHEM 3BP3

Practical Bioinformatics in the Genomics Era

Instructor



Dr. Andrew McArthur

Teaching Assistants



Karyn Mukiri



Maddie McCarthy

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

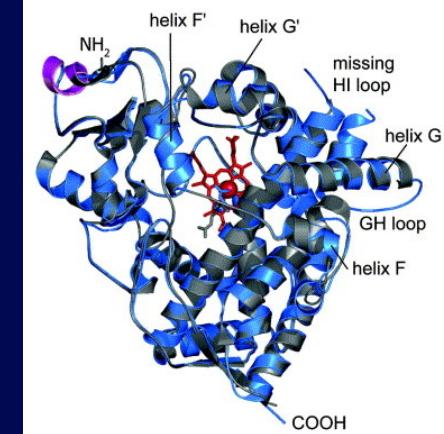
Biological Data | Computer Science | Math | Engineering | Statistics

Genomics

DNA | RNA | Sequencing Data | Genomes

For Example...

Two weeks from now...



Annotation of the toxicological defensome of a Shark

For Example...

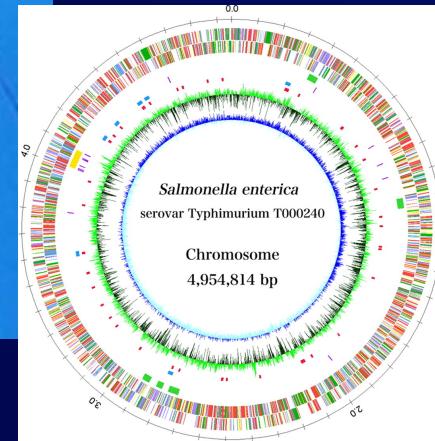
After Reading Break...

```
Terminal: ~ gb2genbank NP_058781 | head -n 10
NP_058781          503 aa      linear   ROD 10-AUG-2014
DEFINITION aromatase [Rattus norvegicus].
ACCESSION NP_058781
VERSION NP_058781.2 GI: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
>NP_058781|ref|NP_058781|1| aromatase [Rattus norvegicus]
MFLEMLNPYMHNVТИNVPETVPVSMPLLIMGLLILLIRNCESSSIFPGPGYCLGIGPLISHGRFLWMGI
GSACNYNNKMYGEFMRWVISGEETLISKSSESSHHVMMKHSNISSRGSKRGLOCIGHENGIFNNNNPSL
WRTVRPFMKAALTGFLIRHVEVCYESIKOHDLRLGOVTDSGYVVVTLMRHIMLDSTNLFLGIPDDE
SSIVKKI1QYNAWQALLIKPNIFFKLISWLRYKVERSVKRQKVSSAEKLECDMFAT
DLIAERGRGDLTKENVNOCILEMLTAAPDTMSVTLYVMMLLLIAEYEVETAILKEIHTVVGDIRI1G0V
QLNKVVENFINESLTVRQFVVDLVHRALEDDVIDGYPVVKKGTHILNIGRMHRLEYFPFKNEFTLENFEK
NVPVRYFQFFGPRSCAGKYIAMVMMKVVLTLKRFRVKTLQKRRCIEMMFKNNDLSLHLDDESPIVEI
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 and at 6 months. Genotyping was done by Taqman SNP allelic discrimination assay. RESULTS: Women with the AA genotype for the rs700518 (G/A at Val180) 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 < 0.05. 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...

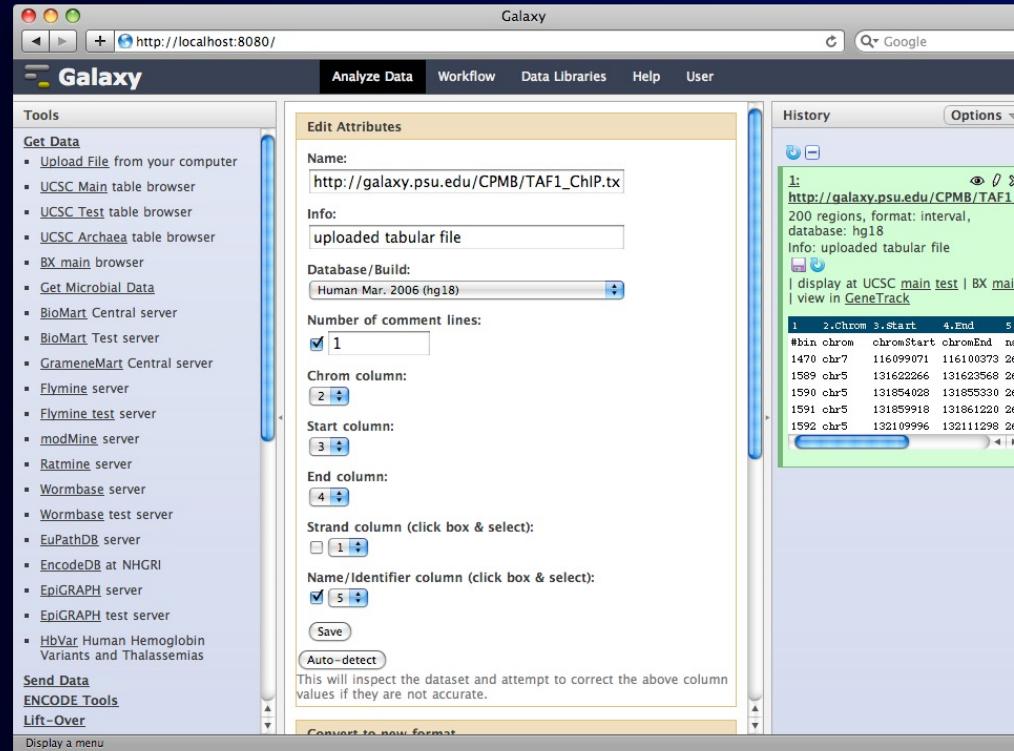
Early November...



Use genomics to analyze a pathogen outbreak

For Example...

By the end of November...



The screenshot shows the Galaxy web interface version 0.6.0. The main window displays a configuration dialog for a dataset named "http://galaxy.psu.edu/CPMB/TAF1_ChIP.txt". The dialog includes fields for Name, Info, Database/Build, Number of comment lines, Chrom column, Start column, End column, Strand column, and Name/Identifier column. A "Save" button is present at the bottom. To the right, the "History" panel shows a single dataset entry with the same name, detailing its content as a tabular file with 200 regions in hg18 format. Below the history is a preview of the dataset's contents, which include columns for bin, chrom, start, end, and name, along with several rows of genomic data.

Use cloud computing to analyze gene expression data

Hybrid Learning...

Product Solutions Open Source Pricing

agmcarthur / Biochem-3BP3 Public

Code Issues Pull requests Actions Projects Security Insights

master · 1 branch · 0 tags Go to file Code About

agmcarthur new Lecture 11 310ef1f 5 minutes ago 280 commits

Lab_1_Genome_Databases format for Flash Update lists last week

Lab_2_Gene_Finding format for Flash Update lists last week

Lab_3_Phlogenetics format for Flash Update lists last week

Lab_4_Ontologies format for Flash Update lists last week

Lab_5_Linux format for Flash Update lists last week

Lab_6_Genome_Assembly format for Flash Update lists last week

Lab_7_Epidemiology format for Flash Update lists last week

Lab_8_Microarrays format for Flash Update lists last week

Lab_9_RNASeq format for Flash Update lists last week

Lectures new Lecture 11 5 minutes ago

.gitignore gitignore 2 weeks ago

LICENSE README and License 4 years ago

README.md update tours 31 minutes ago

README.md

BIOCHEM 3BP3 Practical Bioinformatics in the Genomics Era

Department of Biochemistry & Biomedical Sciences, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada

This is a living document, content will be updated frequently.

Introduction to bioinformatics theory, tools, and practice with an emphasis on high-throughput DNA sequencing technologies. Areas of emphasis include gene sequence analysis, functional prediction, genome assembly and

About

Readme

GPL-3.0 license

Activity

7 stars

3 watching

3 forks

Report repository

Releases

No releases published

Packages

No packages published

Contributors 2

agmcarthur Andrew G. McArthur
jaleezyy Jalees A. Nasir

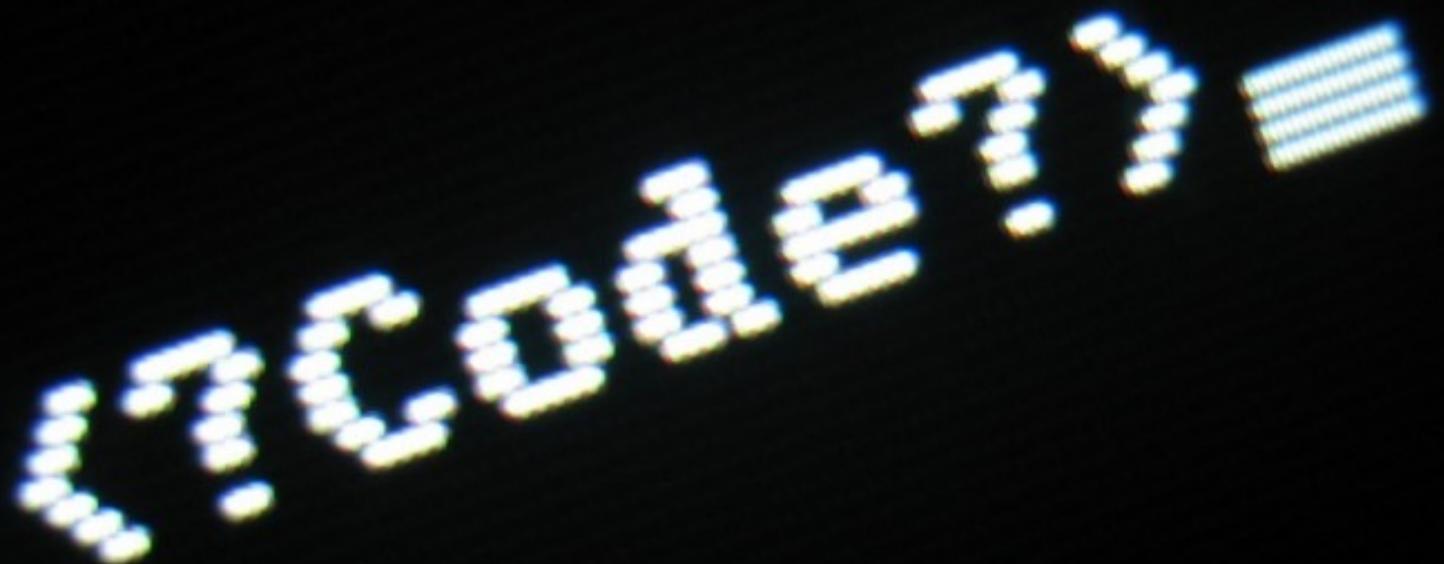


Supplementary Videos

Not official course content for Fall 2023, recorded during Fall 2020, requires McMaster authorized access. Please note, Stream (Classic) is being decommissioned and all videos will be migrated to Stream (on OneDrive and SharePoint) in August, 2023.

- Dr. Joanna Wilson, Department of Biology, McMaster University discusses their research program in aquatic toxicology and the role of genomics and bioinformatic in their research, [video ~10 minutes](#)
- Dr. Christine Mader, Farncombe Metagenomics DNA Sequencing Core, McMaster University provides an overview of McMaster high-throughput DNA sequencing facility, [video ~72 minutes](#)
- Mark Hahn, SHARCNET/Digital Alliance provides an overview of the SHARCNET high-performance computing platform, [video ~50 minutes](#)
- Dr. Robyn Lee, Dalla Lana School of Public Health discusses critical infectious disease analyses in the Canadian north, [video ~7 minutes](#)
- Dr. Fiona Whelan, University of Nottingham discusses genomics and bioinformatics of the human microbiome, [video ~6 minutes](#)
- Dr. Guillaume Paré, Population Health Research Institute, McMaster University discusses exome sequencing and the genetics of cardiovascular disease, [video ~17 minutes](#)

No Computer Science Experience Required



Registered students are from iSci, Biochemistry, BDC, Neuroscience, Biology & Mathematics, Biology, Mol. Biol. Genetics, Chemical Biology, BDC, Life Sciences, BioPharm

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

- We will use Avenue to Learn for submitting assessments and grading
 - Guest lecture material, course documents, supplementary tutorial files
- GitHub: <https://github.com/agmcarthur/Biochem-3BP3>
 - Schedule
 - Tutorials
 - Flash Update details
 - Lecture PowerPoint files
 - Supplementary Videos
- MS Teams for asynchronous communication

- All visual lecture and lab material (e.g., Powerpoint, PDF) will be provided in advance on A2L and/or GitHub.
- All lectures will be recorded and provided as videos in a timely manner to the entire class.
- There are no in-class evaluations, all assignments are take-home and not due for at least 7 days, including lecture quizzes.
- The exception are the 10-minute Flash Update presentations, which are scheduled for each student on a specific date.
- There are no tests or exams in this course.

Course Structure

Weekly Modules - Tuesdays

- Traditional lecture to introduce the topic and outline the key objectives
- Overview and (optional) readings available in GitLab:
<https://github.com/agmcarthur/Biochem-3BP3>
- PowerPoint files available on GitHub
- GitHub may provide links to supplementary videos
- All lectures will be recorded and provided as videos in a timely manner to the entire class

Course Structure

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](#)

Flash Updates

- *GenBank*
- *Ensemble*
- *Growth of Sequencing Data*

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](#)

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

Weekly Modules – Thursday

- A core take-home assignment due on Avenue to Learn in 7 days
- 3 students give a 10 minute “Flash Update” presentation each week
- TAs overview the Tutorials: <https://github.com/agmcarthur/Biochem-3BP3>
- AL2 Quizzes used for submission of Tutorial answers + multiple choice questions

Event	Day	Time	Location	Participants
Lecture	Tuesdays	3:30-4:20 pm	ABB 136	All students
Tutorial Group 1	Thursdays	2:30-4:20 pm	ABB 164	Group 1 students & Maddie
Tutorial Group 2	Thursdays	4:30-6:20 pm	ABB 164	Group 2 students & Karyn

Flash Updates

- See the Flash Update Presentation Schedule
- 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 Thursdays!

Week	GROUP 1		GROUP 2	<u>Flash Update</u>
	<u>Thursday 2:30-4:20 (Madeline McCarthy)</u>		<u>Thursday 4:30-6:20 (Karyn Mukiri)</u>	
2	Yuen, Nathan		Ruan, Celina	Ensemble
2	Ahmed, Ameer		Neku, Ayushma	Growth of Sequencing Data
2	Khan, Zahrah		Peterson, Keaton	GenBank
3	Hossain, Abrar		Haris, Abdullah	BLAST
3	Wang, Spring		Graham, Andrew	Pfam
3	Murphy, Avery		Li, Danielle	PROSITE
4	Williams, Abby		Bleecker, Will	Terminology
4	Ta, Tiffany		Siraj, Kabir	Sequence Alignment
4	Subramaniam, Asha		Babbar, Samiksha	Phylogenetic Trees
5	Mohammad, Ali		Yang, Jingyi	Gene Ontology
5	Hiscott, Mackenzie		Marlatt, Matt	KEGG

A2L 

GitHub 

WEEK 2 - GenBank, Ensembl, Growth of Sequencing Data

- **NCBI & GenBank.** Provide a review of the GenBank resource, with an emphasis on the variety of tools and data it offers. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D29-D38. [PMID 36370100](#), *Nucleic Acids Res.* 2022 Jan 7;50(D1):D161-D164. [PMID 34850943](#).
- **Ensembl.** Provide a review of the Ensembl resource, with an emphasis on the variety of tools and data it offers. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D933-D941. [PMID 36318249](#).
- **Growth of Sequencing Data.** Provide an overview of the growth of DNA sequencing data as well as predicted growth. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D141-D144. [PMID 36350640](#), *GenBank and WGS Statistics*, *The Cost of Sequencing a Human Genome*, and *In The Year 2030—Looking at How Genomic Data Might Evolve*.

Essay Assignments

Critical Review – Due October 24, 2023

This is a critical review exercise, worth 20% 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, 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.

Post your preprint on MS Teams



Reflective Exercise – Due December 6, 2023

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.

Essay Assignments

Grading

This course does not have tests, a mid-term, or a final exam.

Item Graded	% of Final Grade	Due Date
Lab Assignments (7)	35%	weekly
Flash Update Presentation	15%	varies
Lecture Quizzes (3)	15%	varies
Critical Review	20%	October 24, 2023
Reflective Essay	15%	December 6, 2023

- No textbook – primary literature only
- TAs will post their office hours– use MS Teams too!
- All work submitted via Avenue to Learn
- Late penalties will be assessed at 10% per day, including weekends.
 - After 4 days, the assignment will not be accepted and a grade of 0 will be assigned.
- MSAF (Self Report) for absences not longer than 3 days & work worth not greater than 25% of the final grade
- MSAF (Administrative) for absences longer than 3 days or work worth more than 25% of the final grade

Andrew

- Undergraduate in Ecology & Evolution (UWO)
- PhD in Deep-Sea Biology (UVic)
- Post-Doc at the Smithsonian Institution (Washington, DC)
- Faculty at the Marine Biological Laboratory (Woods Hole, MA)
- Left academia in 2006 – 10 years in the private sector with my own Bioinformatics company
- Joined McMaster in late 2014 as the Cisco Research Chair in Bioinformatics
- Sequenced the *Giardia lamblia* genome, leads the Comprehensive Antibiotic Resistance Database

Karyn

- PhD student in the McArthur lab
- UG degree in Biotechnology (McMaster)
- Software engineering & analytics of antibiotic resistome prediction

Maddie

- PhD student in the McArthur lab
- UG degree in Microbiology (UGuelph), MSc Micro & Immunology (USask)
- Advanced sequencing methods for genomic surveillance of bacterial pathogens

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 materials
- Ask questions and use your Teaching Assistant!

A few things...



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:
[https://uts.mcmaster.ca/services/computers-
printers-and-software/virtual-private-
networking/](https://uts.mcmaster.ca/services/computers-printers-and-software/virtual-private-networking/)

Using PubMed

WEEK 2 - GenBank, Ensembl, Growth of Sequencing Data

- **NCBI & GenBank.** Provide a review of the GenBank resource, with an emphasis on the variety of tools and data it offers. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D29-D38. [PMID 36370100](#), *Nucleic Acids Res.* 2022 Jan 7;50(D1):D161-D164. [PMID 34850943](#).
- **Ensembl.** Provide a review of the Ensembl resource, with an emphasis on the variety of tools and data it offers. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D933-D941. [PMID 36318249](#).
- **Growth of Sequencing Data.** Provide an overview of the growth of DNA sequencing data as well as predicted growth. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D141-D144. [PMID 36350640](#), *GenBank and WGS Statistics*, *The Cost of Sequencing a Human Genome*, and *In The Year 2030—Looking at How Genomic Data Might Evolve*.

Using PubMed

National Library of Medicine
National Center for Biotechnology Information

Log in

PubMed®

Advanced

Search

User Guide

Save Email Send to Display options

> Nucleic Acids Res. 2023 Jan 6;51(D1):D29–D38. doi: 10.1093/nar/gkac1032.

Database resources of the National Center for Biotechnology Information in 2023

Eric W Sayers ¹, Evan E Bolton ¹, J Rodney Brister ¹, Kathi Canese ¹, Jessica Chan ¹, Donald C Comeau ¹, Catherine M Farrell ¹, Michael Feldgarden ¹, Anna M Fine ¹, Kathryn Funk ¹, Eneida Hatcher ¹, Sivakumar Kannan ¹, Christopher Kelly ¹, Sunghwan Kim ¹, William Klimke ¹, Melissa J Landrum ¹, Stacy Lathrop ¹, Zhiyong Lu ¹, Thomas L Madden ¹, Adriana Malheiro ¹, Aron Marchler-Bauer ¹, Terence D Murphy ¹, Lon Phan ¹, Shashikant Pujar ¹, Sanjida H Rangwala ¹, Valerie A Schneider ¹, Tony Tse ¹, Jiyao Wang ¹, Jian Ye ¹, Barton W Trawick ¹, Kim D Pruitt ¹, Stephen T Sherry ¹

Affiliations + expand

PMID: 36370100 PMCID: PMC9825438 DOI: 10.1093/nar/gkac1032

Free PMC article

Abstract

The National Center for Biotechnology Information (NCBI) provides online information resources for biology, including the GenBank® nucleic acid sequence database and the PubMed® database of citations and abstracts published in life science journals. NCBI provides search and retrieval operations for most of these data from 35 distinct databases. The E-utilities serve as the programming interface for most of these databases. New resources include the Comparative Genome Resource (CGR) and the BLAST ClusteredNR database. Resources receiving significant updates in the past year include PubMed, PMC, Bookshelf, IgBLAST, GDV, RefSeq, NCBI Virus, GenBank type assemblies, iCn3D, ClinVar, GTR, dbGaP, ALFA, ClinicalTrials.gov, Pathogen

FULL TEXT LINKS

OXFORD ACADEMIC

FREE Full text PMC

ACTIONS

Cite Collections

SHARE

Twitter Facebook LinkedIn

PAGE NAVIGATION

>Title & authors

Abstract

Figures

Similar articles

Using PubMed

An official website of the United States government. [Here's how you know.](#)

 National Library of Medicine
National Center for Biotechnology Information

[Log in](#)

 [Search PMC Full-Text Archive](#) [Search in PMC](#) [Advanced Search](#) | [User Guide](#)

[Journal List](#) > [Nucleic Acids Res](#) > [v.51\(D1\); 2023 Jan 6](#) > [PMC9825438](#)

As a library, NLM provides access to scientific literature. Inclusion in an NLM database does not imply endorsement of, or agreement with, the contents by NLM or the National Institutes of Health.
Learn more: [PMC Disclaimer](#) | [PMC Copyright Notice](#)

Nucleic Acids Research

Nucleic Acids Res. 2023 Jan 6; 51(D1): D29–D38.
Published online 2022 Nov 12. doi: [10.1093/nar/gkac1032](https://doi.org/10.1093/nar/gkac1032)

PMCID: [PMC9825438](#) PMID: [36370100](#)

Database resources of the National Center for Biotechnology Information in 2023

Eric W Sayers,  Evan E Bolton, J Rodney Brister, Kathi Canese, Jessica Chan, Donald C Comeau, Catherine M Farrell, Michael Feldgarden, Anna M Fine, Kathryn Funk, Eneida Hatcher, Sivakumar Kannan, Christopher Kelly, Sunghwan Kim, William Klimke, Melissa J Landrum, Stacy Lathrop, Zhiyong Lu, Thomas L Madden, Adriana Malheiro, Aron Marchler-Bauer, Terence D Murphy, Lon Phan, Shashikant Pujar, Sanjida H Rangwala, Valerie A Schneider, Tony Tse, Jiyao Wang, Jian Ye, Barton W Trawick, Kim D Pruitt, and Stephen T Sherry

► Author information ► Article notes ► Copyright and License information [PMC Disclaimer](#)

Associated Data

► [Data Availability Statement](#)

ABSTRACT

Go to: ►

OTHER FORMATS
[PubReader](#) | [PDF \(1003K\)](#)

ACTIONS
“ Cite
Collections

SHARE
  

RESOURCES
[Similar articles](#) +
[Cited by other articles](#) +
[Links to NCBI Databases](#) +

Using PubMed

Libraries
Off-Campus Access: Login

Login to access Library e-Resources

[Log in with your MacID@mcmaster.ca →](#)

For current McMaster Students, Faculty, and Staff.

Or

[Log in for Retired Faculty, Courtesy and Alumni](#)

McMaster University  + 

Install [Lean Library](#) for easier off campus access!

McMaster University  Libraries

[Off-Campus Access Information](#)

If you have questions or need assistance, please contact us:
[University Library](#) | [Health Sciences Library](#).

McMaster University  BRIGHTER WORLD

[CONTACT](#) [TERMS & CONDITIONS](#) [PRIVACY POLICY](#)

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

Using PubMed

The screenshot shows the McMaster University Library homepage. At the top, the McMaster University logo is on the left, and the word "Library" is on the right. A navigation bar below has links for Home, Services, Collections, Spaces, Research Help, About, and a My account link. The main content area features the Omni search interface with a search bar containing "nature" and a search button. Below the search bar are links for Advanced Search, New Items, and Omni Help. To the right of the search interface is a sidebar with links for Databases, Journal Search, and Research Guides, with "Journal Search" highlighted by a large orange oval. The page also includes a "Collections" section with links for Archives & Rare Books, Local Digital Collections, and Maps & GIS, each with a corresponding icon. A "Recommend a book, film, or journal" button is located on the right side.

McMaster University

Library

Home Services Collections Spaces Research Help About My

nature

Advanced Search | New Items | Omni Help

Databases

Journal Search

Research Guides

Home · Collections

Collections

McMaster's four libraries provide books and journals in the humanities and social sciences (Mills), science and engineering (Thode), business (Innis), and medicine (Health Sciences Library). Online resources - and print - are available from our Catalogue and Databases. Need something more specialized? Check below!

Are we missing something? Request a book purchase.

Recommend a book, film, or journal →

Archives & Rare Books

Local Digital Collections

Maps & GIS

Using PubMed

The screenshot shows the McMaster University Library's Journal Search interface. The top navigation bar includes links for 'New Search', 'Journal Search', 'Browse', 'Databases', 'Research Guides', 'Omni Help', and '...'. A search bar at the top right contains the text 'Nature'. Below the search bar, a large white box displays the search results for 'Nature'. The results include the journal title 'Nature' and several other entries related to it. On the left side of the page, there is a sidebar listing various academic disciplines and general categories for browsing journals.

Journal Search | Nature

Search for journals

Use the following options to find journals, newspapers and magazines:

- Enter a title or ISSN (International Standard Serial Number) in the search box.
- Use the Journals by category option to browse titles by subject area.

Omni | McMaster University

Collections | McMaster University Library

New Search | Journal Search | Browse | Databases | Research Guides | Omni Help | ...

Journal Search | Nature

Search for journals

Use the following options to find journals, newspapers and magazines:

- Enter a title or ISSN (International Standard Serial Number) in the search box.
- Use the Journals by category option to browse titles by subject area.

Journals by category

- Arts, Architecture & Applied Arts
- Business & Economics
- Earth & Environmental Sciences
- Engineering & Applied Sciences
- General
- Health & Biological Sciences
- History & Archaeology
- Journalism & Communications
- Languages & Literatures
- Law, Politics & Government
- Music, Dance, Drama & Film
- Philosophy & Religion
- Physical Sciences & Mathematics
- Social Sciences

Using PubMed

Collections | McMaster University Library

Journal Search - Nature

mni McMaster University

New Search Journal Search Browse Databases Research Guides Omni Help ...

Journal Search Nature x 🔍

Sign in to get complete results and to request items Sign in DISMISS

PAGE 1 1-10 of 252 Results ▾

1 JOURNAL [Nature](#)
Lockyer, Norman, Sir, 1836-1920, editor.
1869
Available online and Bertrand Russell Archives Bertrand Russell Archives (88 Forsyth Ave. N) - Russell Library (RUSS LIB 0056) and other locations >
Available Online >

2 JOURNAL [Nature \(Makassar, Indonesia\)](#)
Ikatan Arsitek Indonesia, issuing body.
2014-
PEER-REVIEWED JOURNAL OPEN ACCESS
Available Online >

3 JOURNAL [Nature photonics \(Online\)](#)
c2007-
Available Online >

4 JOURNAL [Democracy & nature \(Online\)](#)
1995 - 2003
PEER-REVIEWED JOURNAL
Available Online >

5 JOURNAL [Nature materials \(Online\)](#)
Nature Publishing Group
2002-

Modify your results

Sort by Relevance ▾

Availability ▾

- Available online
- Peer-Reviewed
- Open Access
- Available in McMaster Libraries

Journals by category

- Arts, Architecture & Applied Arts
- Business & Economics
- Earth & Environmental Sciences
- Engineering & Applied Sciences
- General
- Health & Biological Sciences
- History & Archaeology
- Journalism & Communications
- Languages & Literatures
- Law, Politics & Government
- Music, Dance, Drama & Film

Using PubMed

The screenshot shows a library search results page for the journal "Nature". The page includes a sidebar with navigation links like "Search by category" and "View Online". The main content area displays the journal details, sharing options, and full-text availability information. A yellow oval highlights the "Nature Online Journals" entry.

JOURNAL
Nature.
Lockyer, Norman, Sir, 1836-1920, editor.
1869

Available at Bertrand Russell Archives Bertrand Russell Archives (88 Forsyth Ave. N) - Russell Library (RUSS LIB 0056) and other locations >
Available Online >

Top
Send to
Search inside
View Online
Get It
Details
Virtual Browse
Links

Citation Email Permalink Export BibTeX Export to Excel Export RIS EndNote Mendeley Print

Article title or keyword

View Online

Full text availability

Factiva
Available from 05/20/2022.
 [Terms of Use](#)

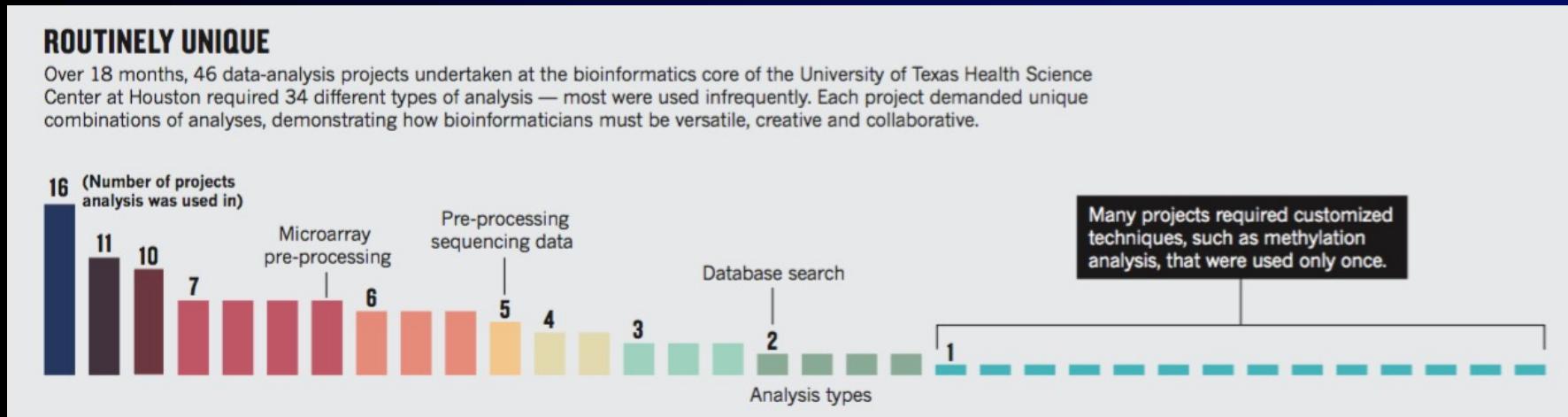
Gale Academic OneFile
Available from 2000.
Most recent 1 year(s) not available.
 [Terms of Use](#)

Gale OneFile: CPI.Q
Available from 01/06/2000 until 12/23/2021.
 [Terms of Use](#)

Nature Online Journals
Available from 1869 volume: 1 issue: 1.
 [Terms of Use](#)
 [Terms of Use](#)

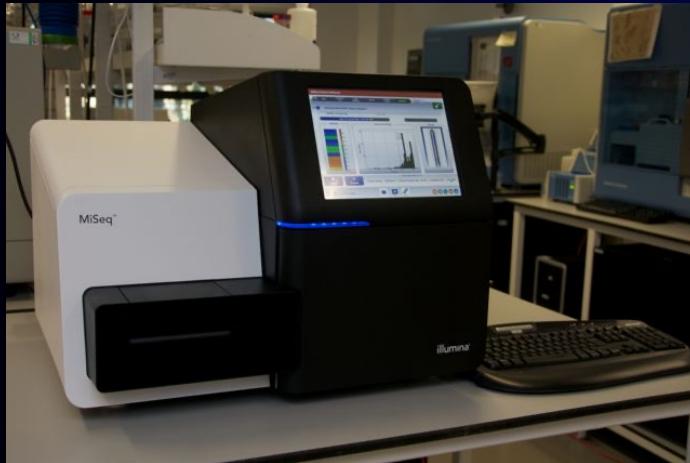
Bioinformatics

- 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



Bioinformatics

- Even 10 years ago collecting low- to mid-volume data (e.g. DNA sequencing) was slow and expensive
- Now – large scale data collection fast and inexpensive



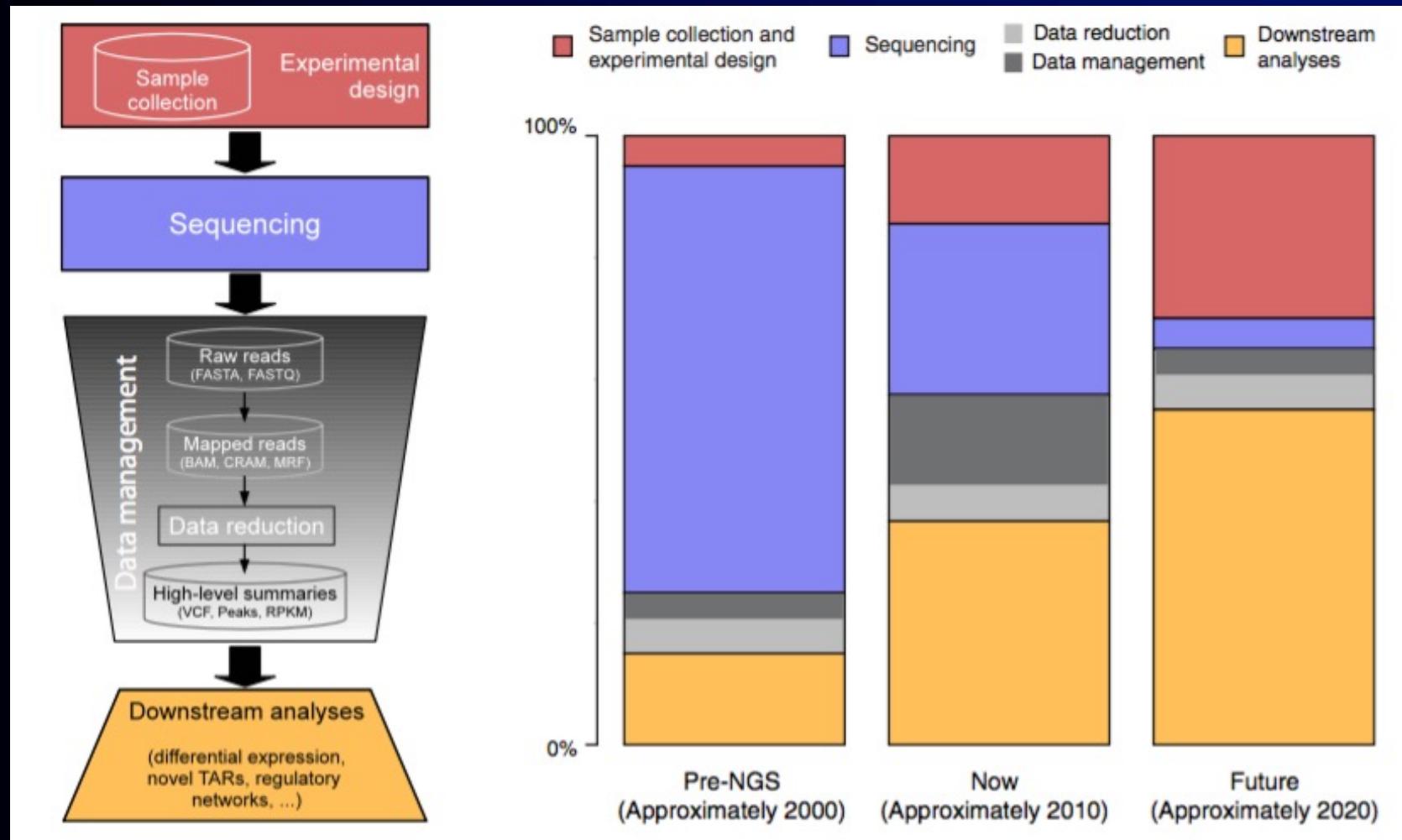
Illumina DNA Sequencer



Nanopore DNA Sequencer

- DNA sequencing is quickly becoming an “assay” – performed by many researchers whose primary goal is not bioinformatics
- We now need a balance between bioinformatics labs / core facilities and translating the skill set to researchers in all areas of bioscience

Bioinformatics – Relative Costs



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

Next Generation Sequencing (NGS) – Low cost, high volume advanced sequencing methods, e.g. Solexa/Illumina massively parallel sequencing
Pre-NGS – High cost, low volume DNA sequencing based on the Sanger sequencing method

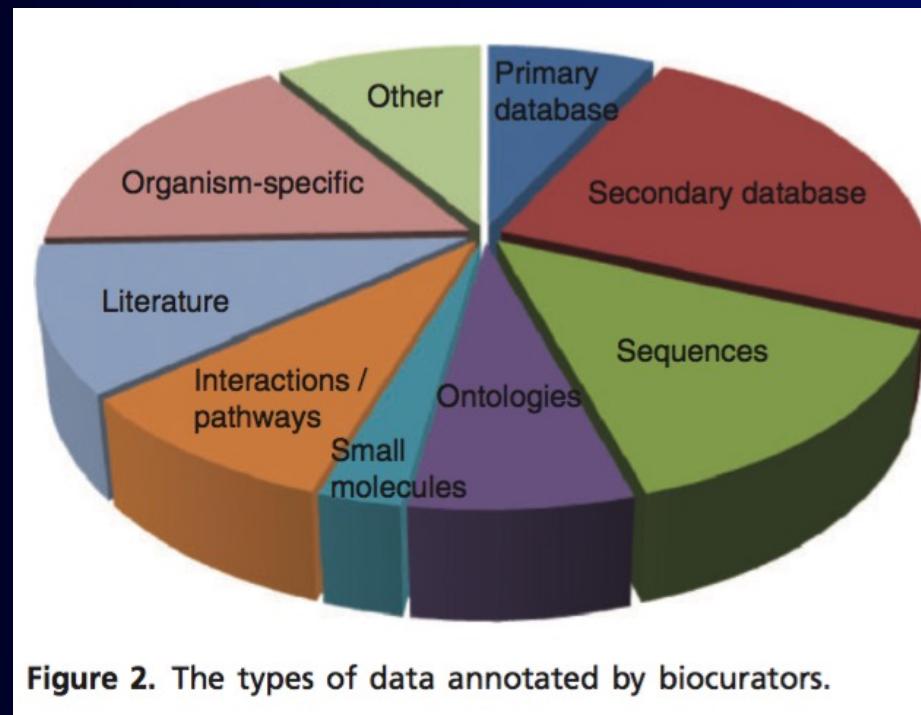
Bioinformatics

- 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
- 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

Bioinformatics

- Bioinformatics is one of the data sciences
 - data – actual measurements and observations
 - metadata – information on methods, observer ID, date time, etc.
 - software & algorithms
 - databases and data formats
 - computing hardware & networks
- Publish and share your data or it might as well not exist!
- Permanent identifiers are essential
- Follow data and file format standards
- Document and share your methods, software & workflows – this is important metadata
- Love your data scientist – Biocurators are the underdogs of modern biological research

BioCuration & Databases

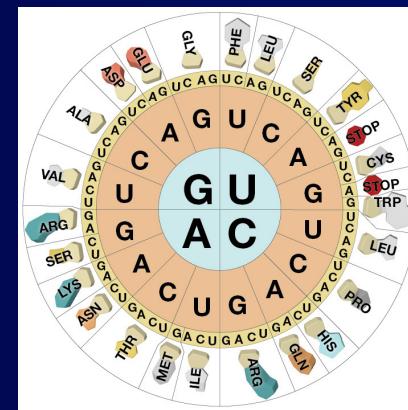
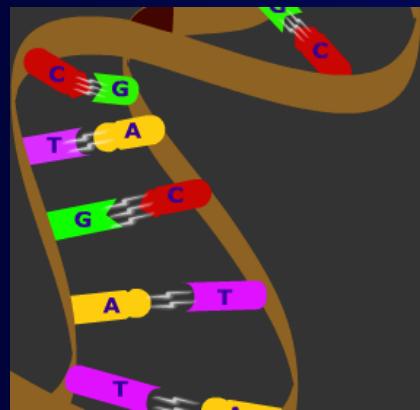
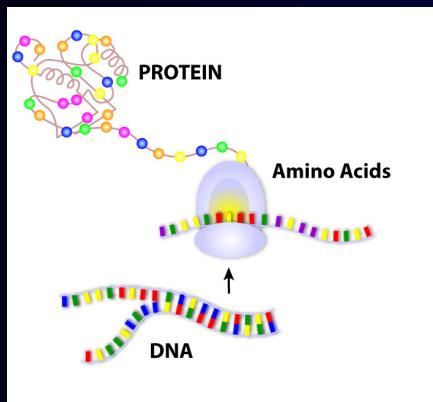


Burge et al. 2012. Biocurators and biocuration: surveying the 21st century challenges. Database Mar 20:bar059

- One of the most important tools in bioinformatics today is the database
 - Primary – raw data – DNA sequence, protein structure, chemical structure, gene expression measurements, etc
 - Secondary – integration of many sources of primary data, often organism-specific (e.g. *Drosophila* DB) or area specific (e.g. drug-drug interaction DB)
 - Literature curation (human or automated) and higher-level conceptual organization of knowledge (e.g. ontologies) increasingly important

Next Week's Lab focuses on Databases

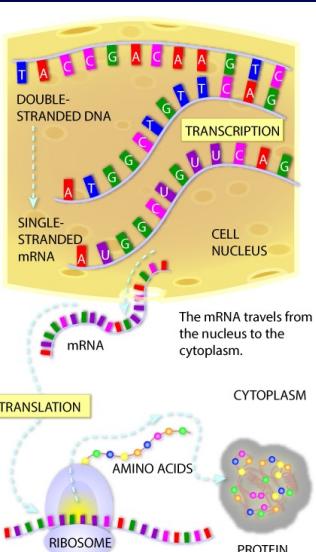
- Central Dogma



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

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).

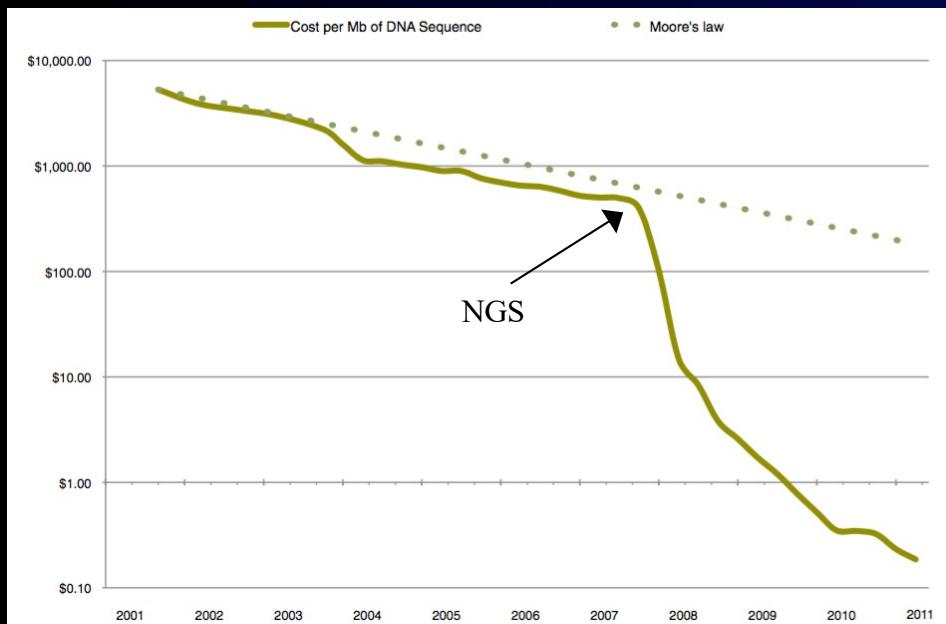
TRANSLATION: The protein-making machinery, called the ribosome, reads the mRNA sequence and translates it into the amino acid sequence of the protein. The ribosome starts at the sequence AUG, then reads three nucleotides at a time. Each three-nucleotide codon specifies a particular amino acid. The "stop"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



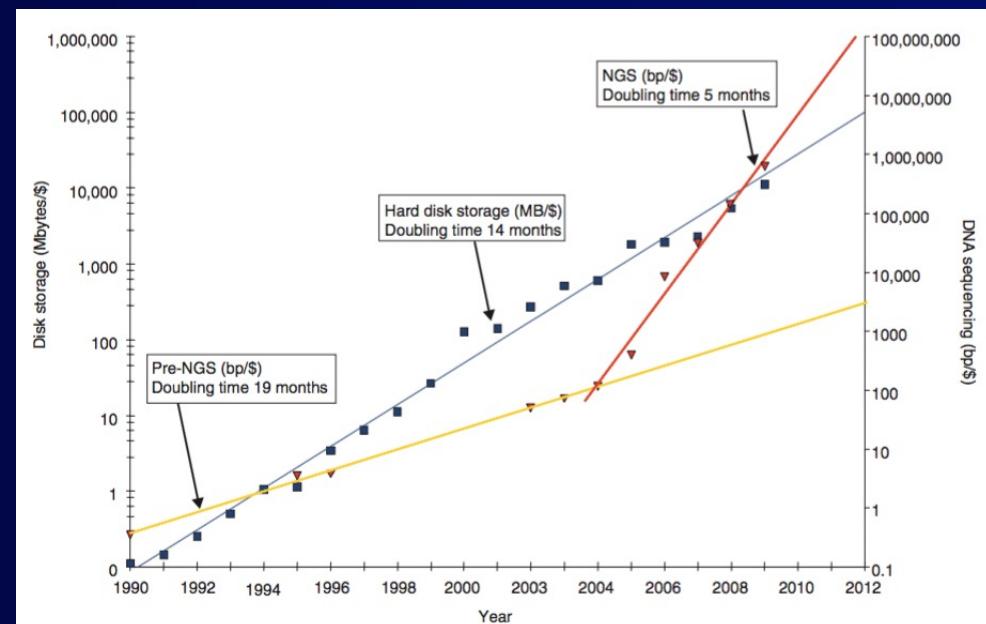
- adenine and guanine are purines
- cytosine, thymine, and uracil are pyrimidines

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?



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.

This week & next...

HSC 3N49

GROUP 1:
Yuen, Nathan
Ahmed, Ameer
Khan, Zahrah

GROUP 2:
Ruan, Celina
Neku, Ayushma
Peterson, Keaton

September 5 & 7	Lecture 1: Introduction to Bioinformatics & the Course	Tours of FHS SeqCore	
September 12 & 14	Tours of SHARCNET	Lab 1: Introduction to Lab & Genome Databases	GenBank, Ensembl, Growth of Sequencing Data

WEEK 2 - GenBank, Ensembl, Growth of Sequencing Data

- **NCBI & GenBank.** Provide a review of the GenBank resource, with an emphasis on the variety of tools and data it offers. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D29-D38. [PMID 36370100](#), *Nucleic Acids Res.* 2022 Jan 7;50(D1):D161-D164. [PMID 34850943](#).
- **Ensembl.** Provide a review of the Ensembl resource, with an emphasis on the variety of tools and data it offers. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D933-D941. [PMID 36318249](#).
- **Growth of Sequencing Data.** Provide an overview of the growth of DNA sequencing data as well as predicted growth. See *Nucleic Acids Res.* 2023 Jan 6;51(D1):D141-D144. [PMID 36350640](#), [GenBank and WGS Statistics](#), [The Cost of Sequencing a Human Genome](#), and [In The Year 2030—Looking at How Genomic Data Might Evolve](#).