



Cold Spring Harbor Laboratory

CURES and Data

Good data training in the classroom

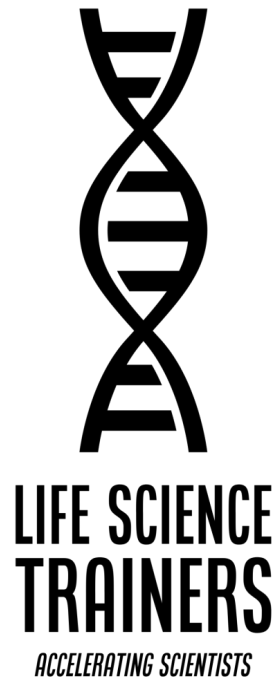
Jason Williams

Cold Spring Harbor Laboratory, DNA Learning Center

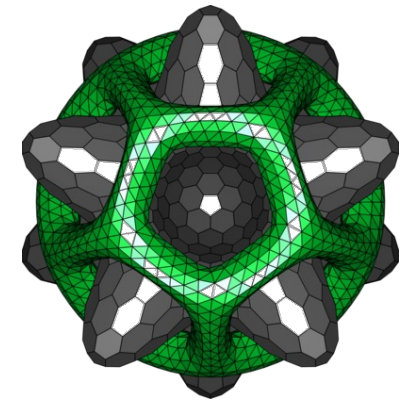


@JasonWilliamsNY

AgBioData – ASPB workshop Aug 2023

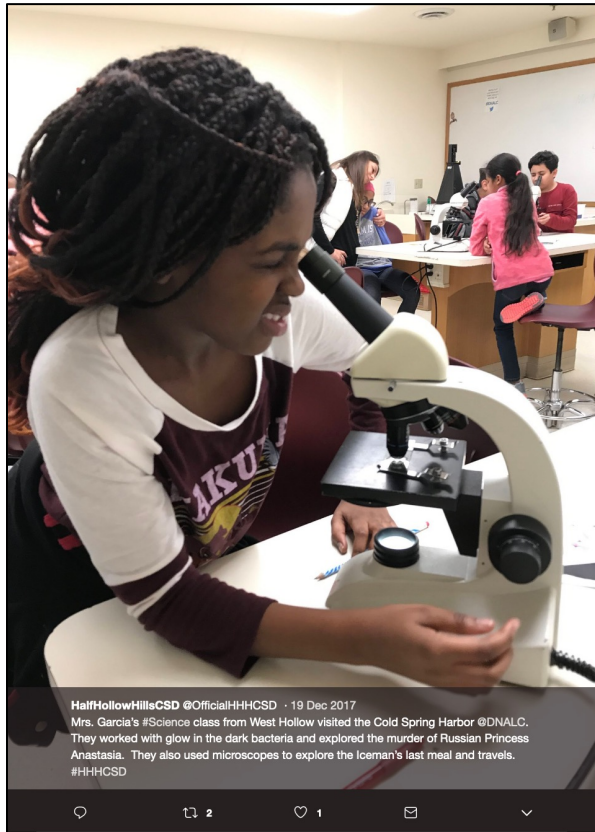


CyVerse



**Journal of
Open Source Education**

I went to school for...



I need to know this for my research...


```
import urllib2
eutils = 'http://www.ncbi.nlm.nih.gov/entrez/eutils/'
efetch = 'efetch.fcgi?'
s = eutils + efetch

targets = ['J04243', 'M60064']
idString = 'id=' + ','.join(targets)
s += idString + '&db=nucleotide&rettype=fasta'
fileObject = urllib2.urlopen(s)
data = fileObject.read().strip()

entries = data.split('\n\n')
title, sequence = entries[0].split('\n', 1)
print title.split(' ', 1)[0]

# prints:
# >gi|154102|gb|J04243.1|STYHEMAPRF
```

I don't know what the future looks like



Community Discussion – What will bioinformatics training and learning look like in the age of ChatGPT and AI?

April 7, 2023 /// No Comments

What will it mean to teach and learn about bioinformatics with the assistance of AI?

[Continue reading »](#)



April post/Meeting info

CSHL DNA Learning Center




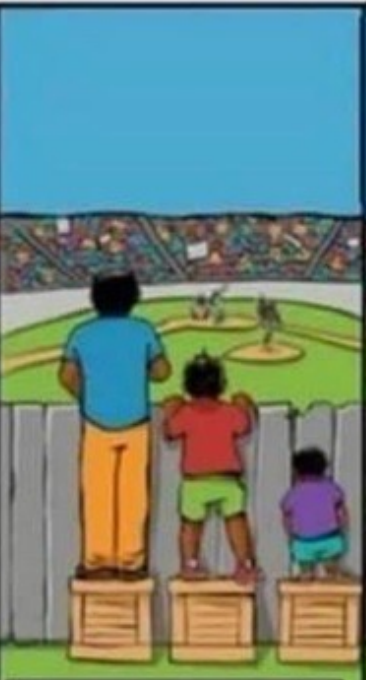



Hands on education in molecular biology/bioinformatics for secondary students (Grade 6-12); secondary and undergraduate faculty training; websites/multimedia



95%

of respondents indicate that bioinformatics should be integrated into the life science curriculum;
32% of faculty report achieving this

These gaps multiply and perpetuate

 <p>REALITY</p>	 <p>EQUALITY</p>	 <p>EQUITY</p>	 <p>JUSTICE</p>	<p>@ClinPsychDavid</p>  <p>INCLUSION</p>
<p>One gets more than is needed, while the other gets less than is needed. Thus, a huge disparity is created.</p>	<p>The assumption is that everyone benefits from the same supports. This is considered to be equal treatment.</p>	<p>Everyone gets the support they need, which produces equity.</p>	<p>All 3 can see the game without supports or accommodations because the cause(s) of the inequity was addressed. The systemic barrier has been removed.</p>	<p>Everyone is INCLUDED in the game. No one is left on the outside; we <u>didn't</u> only remove the barriers keeping people out, we made sure they were valued & involved.</p>


Undergraduate curriculum

PLOS ONE

 OPEN ACCESS  PEER-REVIEWED

RESEARCH ARTICLE

Bioinformatics core competencies for undergraduate life sciences education

Melissa A. Wilson Sayres, Charles Hauser, Michael Sierk, Srebrenka Robic, Anne G. Rosenwald, Todd M. Smith, Eric W. Triplett, Jason J. Williams, Elizabeth Dinsdale, William R. Morgan, James M. Burnette III, Samuel S. Donovan, Jennifer C. Drew, [...], Mark A. Pauley  [view all]

Published: June 5, 2018 • <https://doi.org/10.1371/journal.pone.0196878>

[See the preprint](#)

C1. Role of Bioinformatics

C2. Computational Concepts

C3. Statistical Concepts

C4. Bioinformatics Tools

C5. Data Retrieval

C6. Model

C7. Scripting

C8. Data Types

C9. Implications

POWER

(number of uses/applications)



 **Read the Docs**



SUBWAY

Green Line
RNA-Seq

reStructuredText



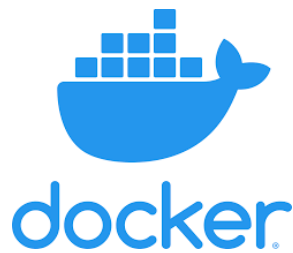
SUBWAY

Purple Line
Metabarcoding



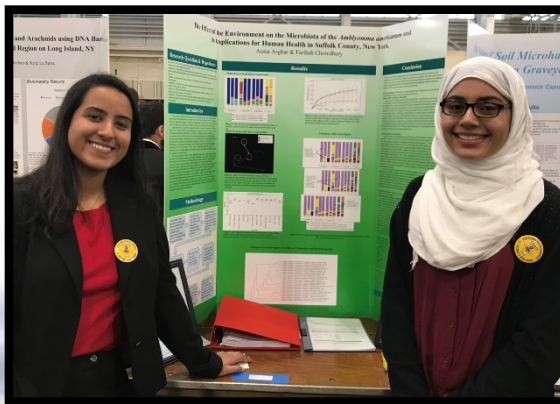
SUBWAY

Blue Line
DNA Barcoding

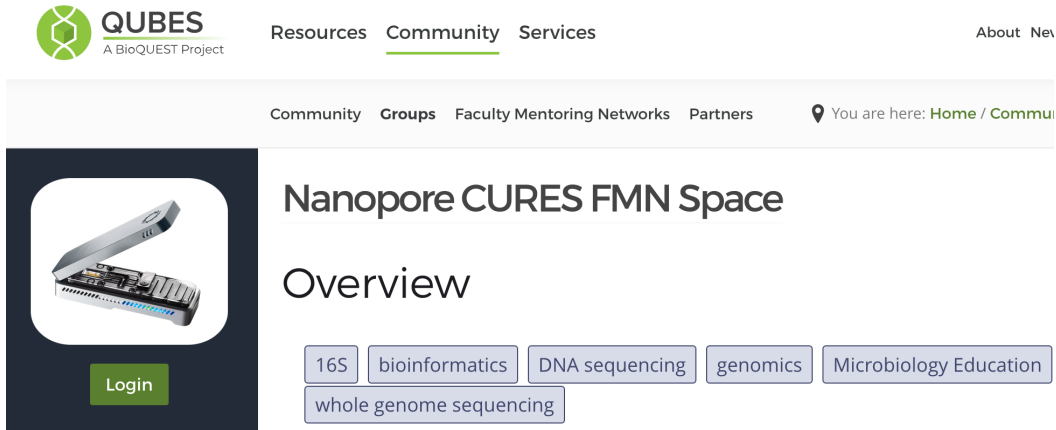


Ease of use
(Skills to get started)

DNA Barcoding Programs



Nanopore Network



The screenshot shows the QUBES website interface. At the top left is the QUBES logo with the text "A BioQUEST Project". Navigation links include "Resources", "Community" (highlighted), and "Services". A breadcrumb trail shows "You are here: Home / Commu". The main content area features a "Nanopore CURES FMN Space" overview with a "Login" button and a list of topics: "16S", "bioinformatics", "DNA sequencing", "genomics", "Microbiology Education", and "whole genome sequencing".



DUE# 2216349

Developing Foundations for Nanopore DNA Sequencing Course-based Undergraduate Research Experiences at Minority-Serving Institutions

- Pilot (2-years)
- Simplify lab and bioinformatics protocols
- Support faculty needs and understand barriers to use



Using DNA Barcodes to Identify and Classify Living Things

DNA barcodes allow non-experts to objectively identify species—even from small, damaged, or industrially processed material. A "DNA barcode" is a unique DNA sequence that identifies each living thing. Short DNA barcodes, about 700 nucleotides in length, can be quickly processed from thousands of specimens and unambiguously analyzed by computer programs. With DNA barcodes, we can discover and catalog biodiversity on our planet using tools developed at the DNA Learning Center.

[Learn More](#)



DNA Barcoding Program Outcomes



NCBI

1,640

Total GenBank
Publications



NCBI

146

First GenBank
Barcodes



NCBI

279

New Sequence
Variants



NCBI

1,331

Unique GenBank
Authors

544

Total Species
Identified

Banbury working group (May 2022)



*“The illiterate of the 21st
century will not be those
who cannot read and write,
but those who cannot
learn, unlearn, and relearn”*
– A. Toffler

Thanks!

