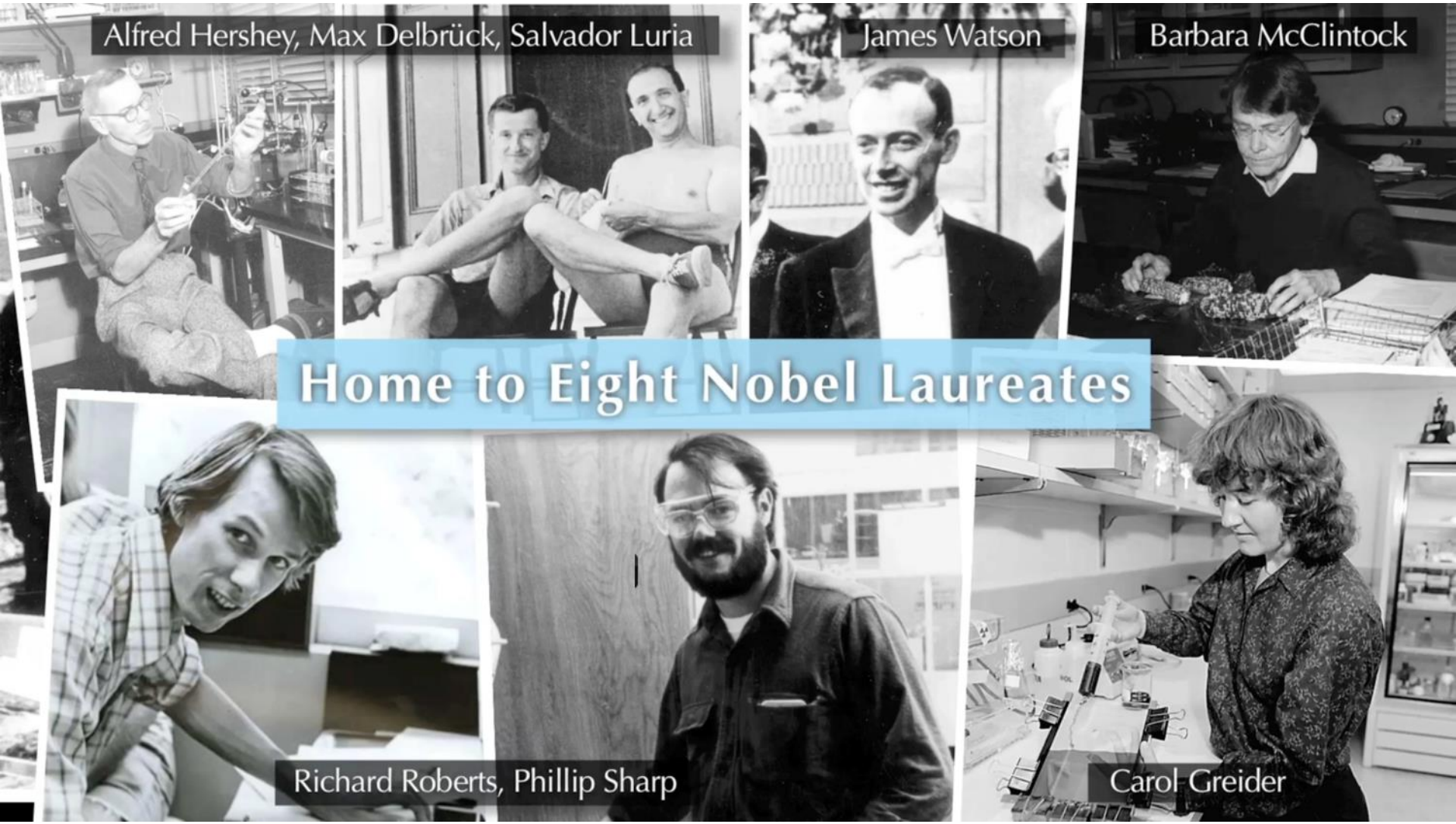
DNA Barcoding and DNA Subway: Infrastructure for Citizen Science

Dave Micklos
DNA Learning Center,
Cold Spring Harbor Laboratory
Banbury Nanopore Meeting, 9/23/24

Cold Spring Harbor Laboratory



The Egalitarian Gene

Agarose Gel Electrophoresis, 1973

1958
Matt Meselson &
Ultracentrifuge, $500,000

1973
Sharp, Sambrook, Sugden
Gel Electrophoresis
Chamber, $250

Detection of Two Restriction Endonuclease Activities in *Haemophilus parainfluenzae* Using Analytical Agarose–Ethidium Bromide Electrophoresis†

Phillip A. Sharp,* Bill Sugden, and Joe Sambrook

ABSTRACT: A rapid assay for restriction enzymes has been developed using electrophoresis of DNA through 1.4% agarose gels in the presence of 0.5 μg/ml of ethidium bromide. The method eliminates lengthy staining and destaining procedures and resolves species of DNA which are less than 7 × 10^6 daltons. As little as 0.05 μg of DNA can easily be detected by direct examination of the gels in ultraviolet light. Using

activities have different chromatogra
phocellulose and Bio-Gel A-0.5m,
DNA at different sites. One activit
DNA at a single position situated
length from the insertion point of
adenovirus SV40 hybrid $Ad2^{++}ND_1$.
cleaves SV40 DNA at three sites w
3 of the 11 cleavage points atta
isolated from *H. influenzae* strain

lar to the axis of the DNA duplex; in other
ndromic sequences of the type

A B C C′ B′ A′
A′ B′ C′ C B A

use different enzymes attack different palindromes
hem generates a characteristic set of cleavage pr
n reacted with DNA. For any particular enzym
ber of fragments obtained is a measure of the num
ndromic sites in the DNA specific to the enzym
size of the fragments reflects the distribution of th
g the DNA.
he two principal methods which have been used to a
fragments of DNA produced by restriction enzym
city sedimentation and electrophoresis through

CHEMISTRY, VOL. 12, NO. 16, 1973

The Egalitarian Genome

Third Generation Sequencing 2014

2001:
ABI 3730 Sequencer

Human Genome:
$2.7 Billion, 13 Years,
1000s of people

2014
Oxford Nanopore
MinION

Human Genome:
$900, 6 Hours, one person

DNA Barcoding Begins:
CSHL Banbury Center Meetings 2003

Mark Stoeckle

Paul Hebert

Jesse Ausubel

The Inspiration

Mark and Kate Stoeckle

CNN.com /technology

DNA testing uncovers suspect sushi

STORY HIGHLIGHT
- Two teenage girls
- They used a genet
- Samples were col
- The results showe

Next Article in Techn

From Christina Chinnici
CNN

TEXT SIZE

NEW YORK (CNN) -- Two teenage girls used DNA bar coding to determine that some sushi on New York dinner plates was mislabeled with cheaper fish being passed off as a more expensive species.

Kate Stoeckle and Louisa Strauss were not science majors or even college students when they decided to take 60 samples of seafood and use a genetic fingerprinting technique to see whether the fish were labeled correctly.

The graduates of Manhattan's Trinity School in New York were inspired by Kate Stoeckle's father, Mark, a scientist and proponent of the use of DNA bar coding, a technique that greatly simplifies the process of identifying a species.

"Growing up, bar coding was dinner conversation, so I was familiar with it," Stoeckle said. "And then

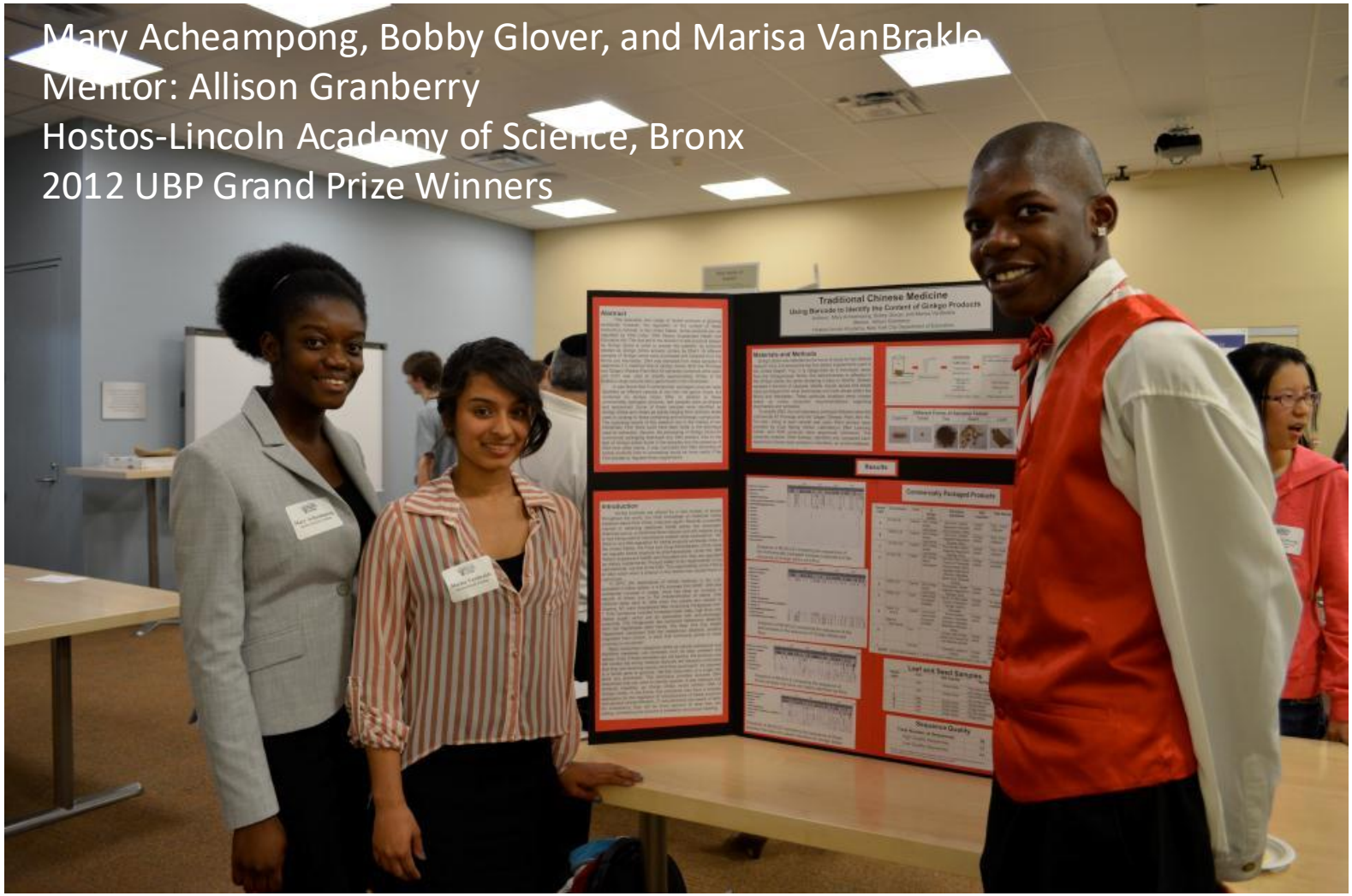
TIME Quotes of the Day

"If you're paying for white tuna and you're eating tilapia, I think you'd want to know that."

KATE STOECKLE, 19, after she and Louisa Strauss, 18, freelanced a science project in which they checked 60 samples of seafood using genetic fingerprinting and found that one-fourth of the fish samples with identifiable DNA were mislabeled

Friday, August 22, 2008

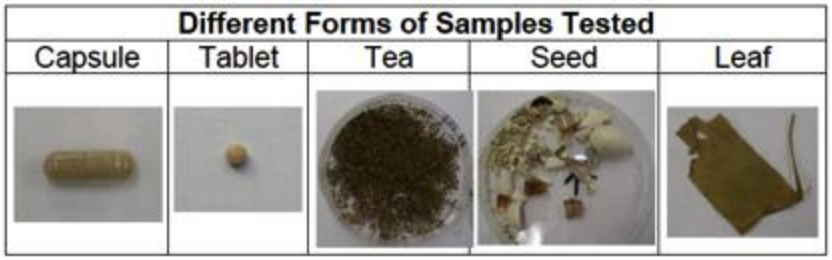
Ginkgo Product Fraud (2012)

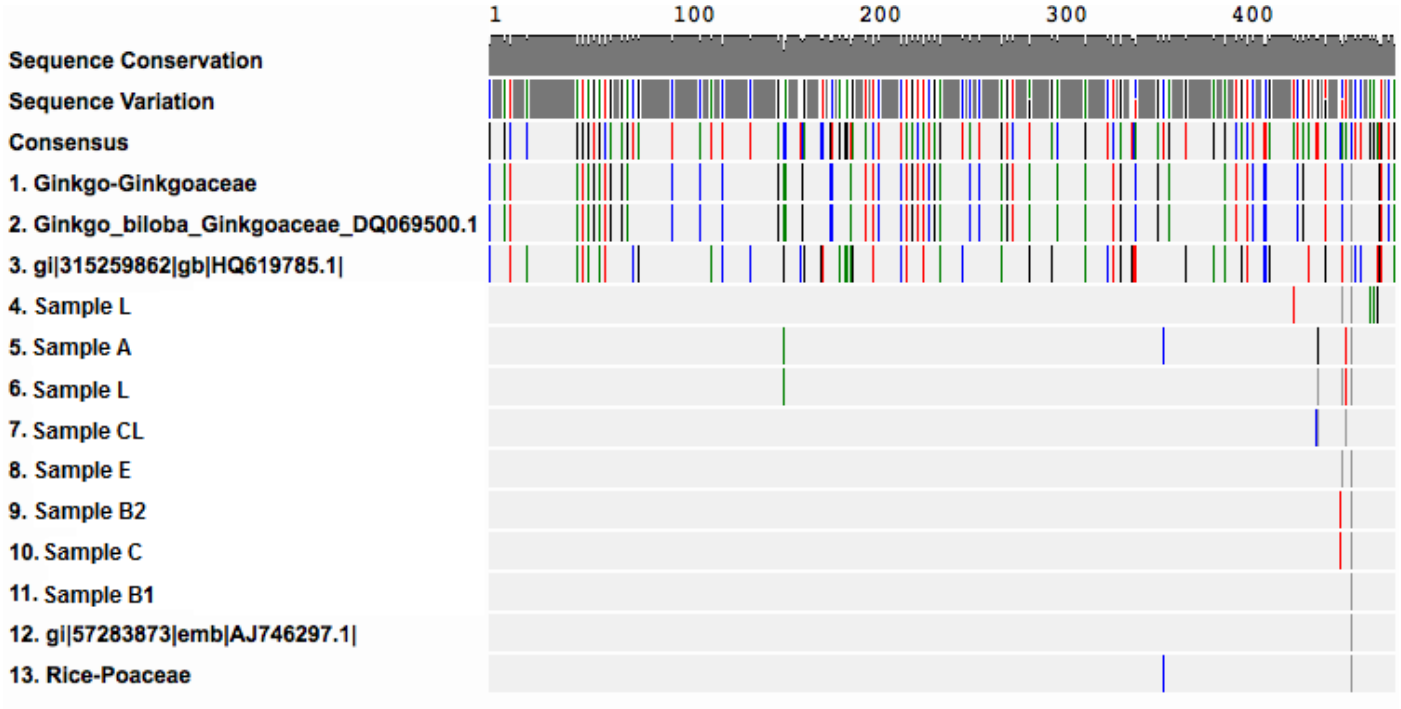


Mary Acheampong, Bobby Glover, and Marisa VanBrakle
Mentor: Allison Granberry
Hostos-Lincoln Academy of Science, Bronx
2012 UBP Grand Prize Winners

Ginkgo Product Fraud (2012)

Sample Letter	Form	DNA Expected	DNA Results
A	Capsule	*Ginkgo biloba*	Rice: *Oryza rufipogon*
B	Capsule	*Ginkgo biloba*	Rice: *Oryza rufipogon*
C	Capsule	*Ginkgo biloba*	Rice: *Oryza rufipogon*
D	Tablet	*Ginkgo biloba*	No sequence available.
E	Capsule	*Ginkgo biloba*	Rice: *Oryza rufipogon*
F	Liquid	*Ginkgo biloba*	No sequence available
G	Capsule	*Ginkgo biloba*	No sequence available
H	Tea	*Ginkgo biloba*	Other rbcL DNA present but not *Mentha piperita*
L	Capsule	*Ginkgo*	Rice: *Oryza*





Ginkgo Product Fraud Revisited

NY Attorney General Cease and Desist Order (2015)

STATE OF NEW YORK
OFFICE OF THE ATTORNEY GENERAL

ERIC T. SCHNEIDERMAN
ATTORNEY GENERAL

DIVISION OF REGIONAL AFFAIRS

February 2, 2015

Doug McMillon, President/CEO
Wal-Mart Stores, Inc.
702 SW 8th Street
Bentonville, Arkansas 72716

Certified—Return Receipt Requested

Re: ***CEASE & DESIST NOTIFICATION***
Spring Valley—Walmart Distributed Herbal Dietary Supplements

Dear Mr. McMillon:

This letter constitutes a demand to cease and desist engaging in the sale of adulterated and/or mislabeled herbal dietary supplements, and in particular to immediately stop the sale of six "Spring Valley" dietary supplements as identified by lot number in the exhibit annexed hereto.

Walmart "Spring Valley" Gingko Biloba

Gingko Biloba. Negative. No gingko biloba DNA was identified. The only DNA identified was "oryza" (commonly known as rice) in 6 of the fifteen tests, with other tests identifying dracaena (a tropical houseplant), mustard, wheat, and radish. Four of the tests revealed no plant DNA whatsoever.

DNALC's Three High School DNA Barcoding Projects

			
Students	1834		2042
Teams	671		702
Mentors	134		241
Schools Academic Institutions	91 32		68
DNA Samples	>10,600		>10,800
DNA Sequences	>16,000 111 million NextGen sequencing reads		>8,600 ~22 billion bp and 75 mil reads from microbiome projects
Sequences in GenBank	67 novel 15 with polymorphisms 115 GPS data		55 novel 275 with polymorphisms 966 GPS data
Years Active	*UBP* 2011- *UBRP* 2014-		2014-

Barcode Long Island Symposium

Cold Spring Harbor Laboratory

1st Genbank Record (2012)

NCBI Resources How To Sign in to NCBI

Nucleotide Nucleotide Search

Limits Advanced Help

Display Settings: GenBank Send:

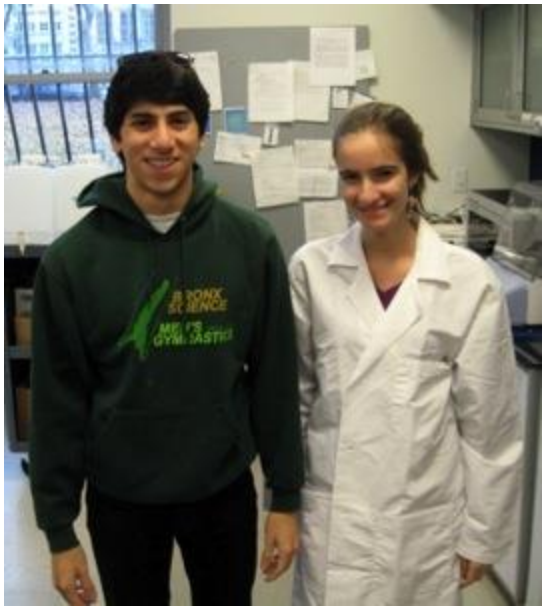
Leucoraja ocellata cytochrome c oxidase subunit I (COI) gene, partial cds; mitochondrial

GenBank: KC109830.1

FASTA Graphics

Go to:

```
LOCUS       KC109830               708 bp    DNA     linear   VRT 12-NOV-2012
DEFINITION  Leucoraja ocellata cytochrome c oxidase subunit I (COI) gene,
            partial cds; mitochondrial.
ACCESSION   KC109830
VERSION     KC109830.1  GI:410991995
DBLINK      BioProject: PRJNA159485
KEYWORDS    .
SOURCE      mitochondrion Leucoraja ocellata (winter skate)
  ORGANISM  Leucoraja ocellata
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
            Elasmobranchii; Batoidea; Rajiformes; Rajidae; Leucoraja.
REFERENCE   1  (bases 1 to 708)
  AUTHORS   Burg,N., Jorquera,A., Atlas,E. and Pineda-Catalan,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-OCT-2012) DNALC, Cold Spring Harbor Laboratory, 1
            Bungtown Rd., Cold Spring Harbor, NY 11724, USA
FEATURES             Location/Qualifiers
     source          1..708
                     /organism="Leucoraja ocellata"
                     /organelle="mitochondrion"
                     /mol_type="genomic DNA"
                     /isolate="DNAS-15-25477"
                     /isolation_source="skate traded in market as food"
                     /db_xref="taxon:173042"
                     /dev_stage="adult"
                     /country="USA: New York, Queens, Flushing Meadows China
                     Town"
                     /lat_lon="40.44 N 73.49 W"
                     /collection_date="09-Apr-2012"
                     /collected_by="Noah Burg, Andrew Jorquera, Emily Atlas,
                     Oscar Pineda-Catalan"
                     /identified_by="Noah Burg and Oscar Pineda-Catalan"
                     /PCR_primers="fwd_name: FishF2_t1, fwd_seq:
                     tgtaaaacgacggccagtcgactaatcataaagatatcggcac, rev_name:
                     FR1d_t1, rev_seq:
                     caggaaacagctatgaccacctcagggtgtccgaaraaycaraa"
     gene            <1..>708
                     /gene="COI"
     CDS             <1..>708
                     /gene="COI"
                     /codon_start=3
                     /transl_table=2
                     /product="cytochrome c oxidase subunit I"
                     /protein_id="AFV95092.1"
                     /db_xref="GI:410991996"
                     /translation="STFNWKDIGTLYLIFGAWAGMVGTGLSLLIRAELSQPGSLLGDD
                     QIYNVLVTAHAFLMIFFMVMPIMIGGFGNWLVPLMIGSPDMAFPRMNNMSFWLLPPSF
                     LLLLASAGVEAGAGTGWTVYPPLAGNLAHAGASVDLTIFSLHLAGISSILASINFITT
                     IINMKPPAISQYQTPLFVWSILVTTVLLLMALPVLAAGITMLLTDRNLNTTFFDPAGG
                     GDPILYQHLFWFFGHPE"
ORIGIN
        1 gcaggtagac tcatcataaa gatatcggca acctatacct aatattttgt gcatgagcag
       61 gtatagtcgg aactggacta agtatttttc ttcgagcagc gttgagccag ccaggcatcc
      121 ttataggtga tgatcagatt tataatgtcc ttgttacagc ccatgcattt gtaataatct
      181 tttttatggt tatacccatt ataattgggg ggttttggca ttgactagtc actttaatga
      241 ttggctatcc agcataggcc tttccccagc a tgaataatat gagcattttga attttcaccc
      301 atcaatttct tcttctacta gcatccgctg gagttgaagc tggggcgggg acagggttga
      361 ctgtctacca cctctagca ggcaatctag cccacgcagg ggcatccgta gacttcaccc
      421 tttctctat ccatttggca ggtatctcct ccatcatggc atccattaac tttatcacca
      481 caattattaa catgaaaccc cctgcaatct ctcaatcacc aacacctctt ttagtatgat
      541 caatacttgt tcacaactgt ctacttattt tagccatccc agttatagca gctggcatta
      601 ctatgcttct cacagatcgt aacctcaata ccactttatt tgacccagca ggaggaggag
      661 accccatttt ttcccagcac ctgttctgat tcttcggcca cccgaag
```



Andrew Jorquera, Bronx HSS
Emily Atlas, Stuyvesant HSS
Noah Burg, AMNH
Oscar Pineda-Catalan, DNALC

PLOS ONE

RESEARCH ARTICLE

DNA barcoding Brooklyn (New York): A first assessment of biodiversity in Marine Park by citizen scientists

Christine Marizzi[1☯*], Antonia Florio[2☯], Melissa Lee[1], Mohammed Khalfan[3], Cornel Ghiban[1], Bruce Nash[1], Jenna Dorey[1,4], Sean McKenzie[5], Christine Mazza[6], Fabiana Cellini[6], Carlo Baria[7], Ron Bepat[8], Lena Cosentino[7], Alexander Dvorak[9], Amina Gacevic[10], Cristina Guzman-Moumtzis[11], Francesca Heller[12], Nicholas Alexander Holt[8†], Jeffrey Horenstein[13], Vincent Joralemon[11], Manveer Kaur[10], Tanveer Kaur[10], Armani Khan[8], Jessica Kuppan[8], Scott Laverty[7], Camila Lock[14], Marianne Pena[10], Ilona Petrychyn[14], Indu Puthenkalam[14], Daval Ram[8], Arlene Ramos[10], Noelle Scoca[15], Rachel Sin[12], Izabel Gonzalez[10], Akansha Thakur[14], Husan Usmanov[12], Karen Han[8], Andy Wu[12], Tiger Zhu[13], David Andrew Micklos[1]

1 DNA Learning Center, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, United States of America, **2** Department of Herpetology, American Museum of Natural History, New York, New York, United States of America, **3** New York University, New York, New York, United States of America, **4** The New York Botanical Garden, Bronx, New York, United States of America, **5** The Rockefeller University, New York, New York, United States of America, **6** Genovesi Environmental Study Center, New York City Department of Education, Brooklyn, New York, United States of America, **7** CSI for International Studies, New York City Department of Education, Staten Island, New York, United States of America, **8** High School for Construction Trades, Engineering and Architecture, New York City Department of Education, Queens, New York, United States of America, **9** International High School at Union Square, New York City Department of Education New York, New York, United States of America, **10** High School for Health Professions and Human Services, New York City Department of Education, New York, New York, United States of America, **11** Frank McCourt High School, New York City Department of Education, New York, New York, United States of America, **12** Franklin D. Roosevelt High School, New York City Department of Education, Brooklyn, New York, United States of America, **13** Stuyvesant High School, New York City Department of Education, New York, New York, United States of America, **14** Forest Hills High School, New York City Department of Education, Queens, New York, United States of America, **15** Brooklyn International High School, New York City Department of Education, Brooklyn, New York, United States of America

☯ These authors contributed equally to this work.
† Deceased.
* cmarizzi@cshl.edu

Citation: Marizzi C, Florio A, Lee M, Khalfan M, Ghiban C, Nash B, et al. (2018) DNA barcoding Brooklyn (New York): A first assessment of biodiversity in Marine Park by citizen scientists. PLoS ONE 13(7): e0199015. https://doi.org/10.1371/journal.pone.0199015

Editor: Matjaž Kuntner, Scientific Research Centre of the Slovenian Academy of Sciences and Art, SLOVENIA

Received: September 21, 2017

Accepted: May 30, 2018

Published: July 18, 2018

DNA Learning Center
BARCODING 101

Home | About | Student Programs | Citizen Science | Laboratory | Resources | Sign in | Sample Database

DNA Barcoding Program Outcomes

NCBI **2,393** Total GenBank Publications

NCBI **220** First GenBank Barcodes

NCBI **579** New Sequence Variants

NCBI **1,623** Unique GenBank Authors

1,075 Unique Taxa Identified

Learn More

Student Research Programs

Learn about metro New York (*Barcode Long Island*, *Urban Barcode Project*, and *Urban Barcode Research Program*), and China-based (*Barcode Suzhou* and *Barcode Beijing*) student research programs.

Citizen Science

Explore citizen science programs (*Barcoding US Ants* and *Citizen DNA Barcode Network*) that ignite community interest in biodiversity and science while contributing to our knowledge of species through DNA barcoding.

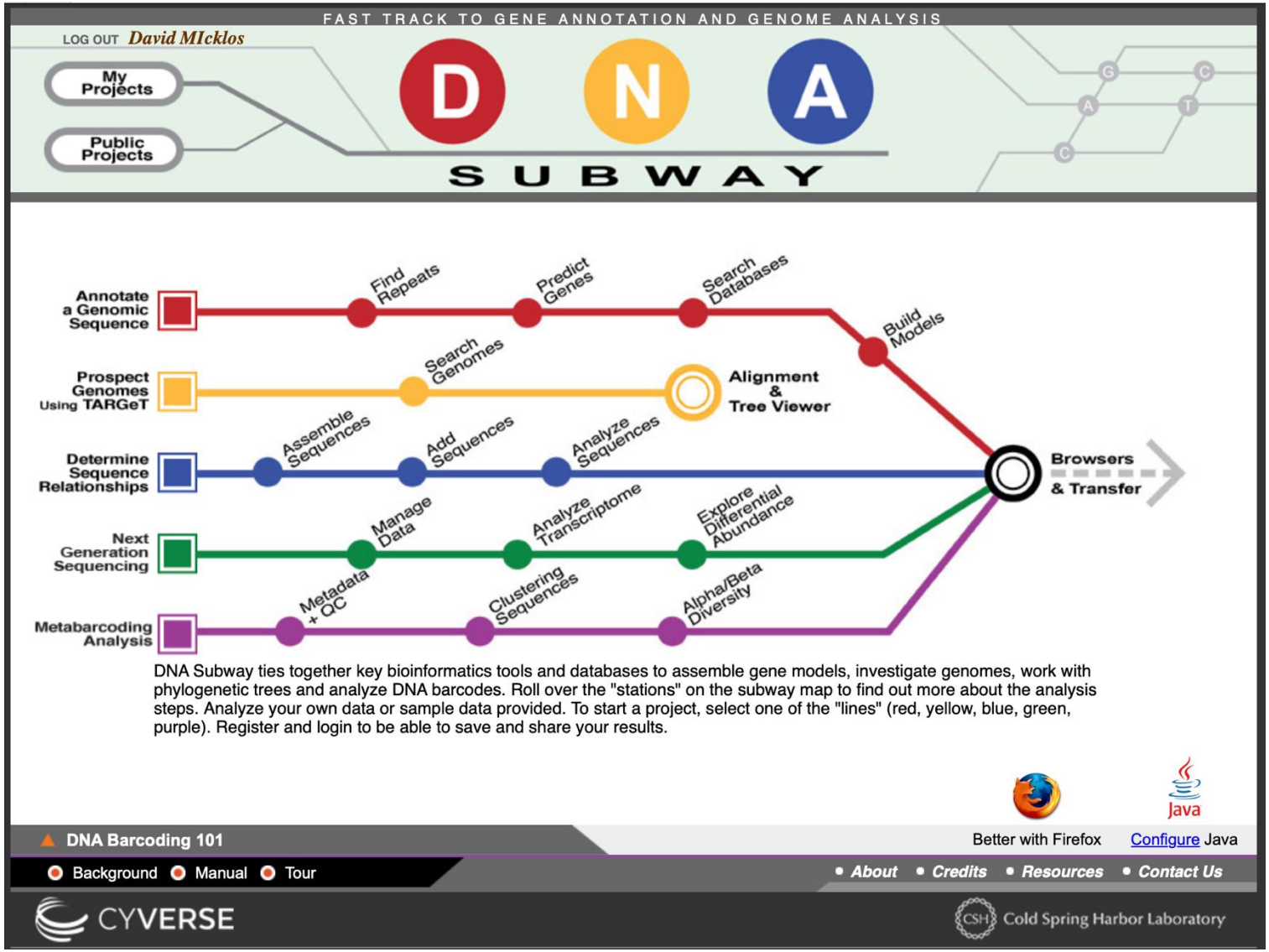
Laboratory & Resources

Protocol and resources supporting DNA barcoding to identify plants or animals—or products made from them. Online tools, animations, videos, presentations, and references that support students, teachers, or citizen scientists.

CSH Cold Spring Harbor Laboratory
DNA Learning Center

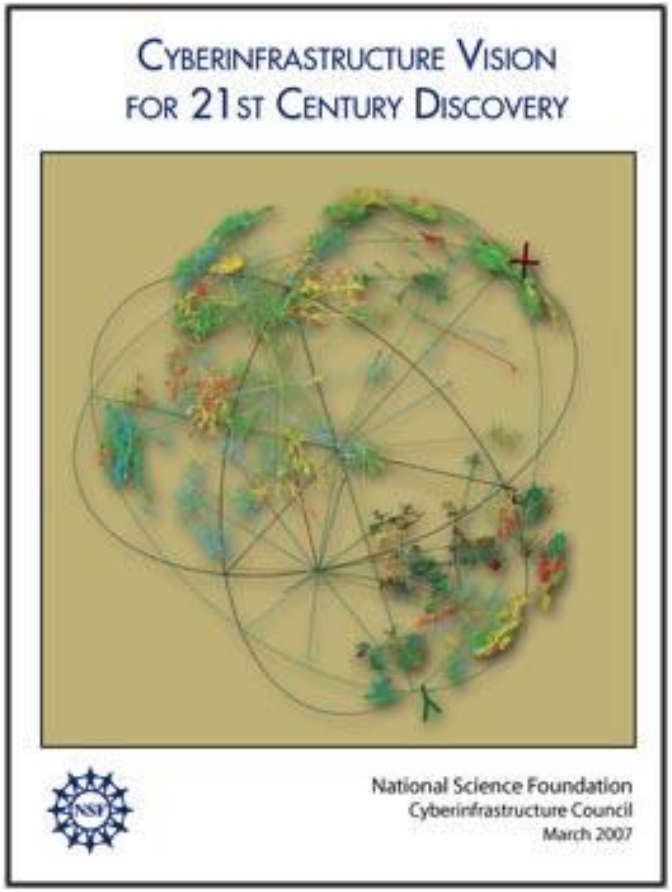
Walk?

…ride!

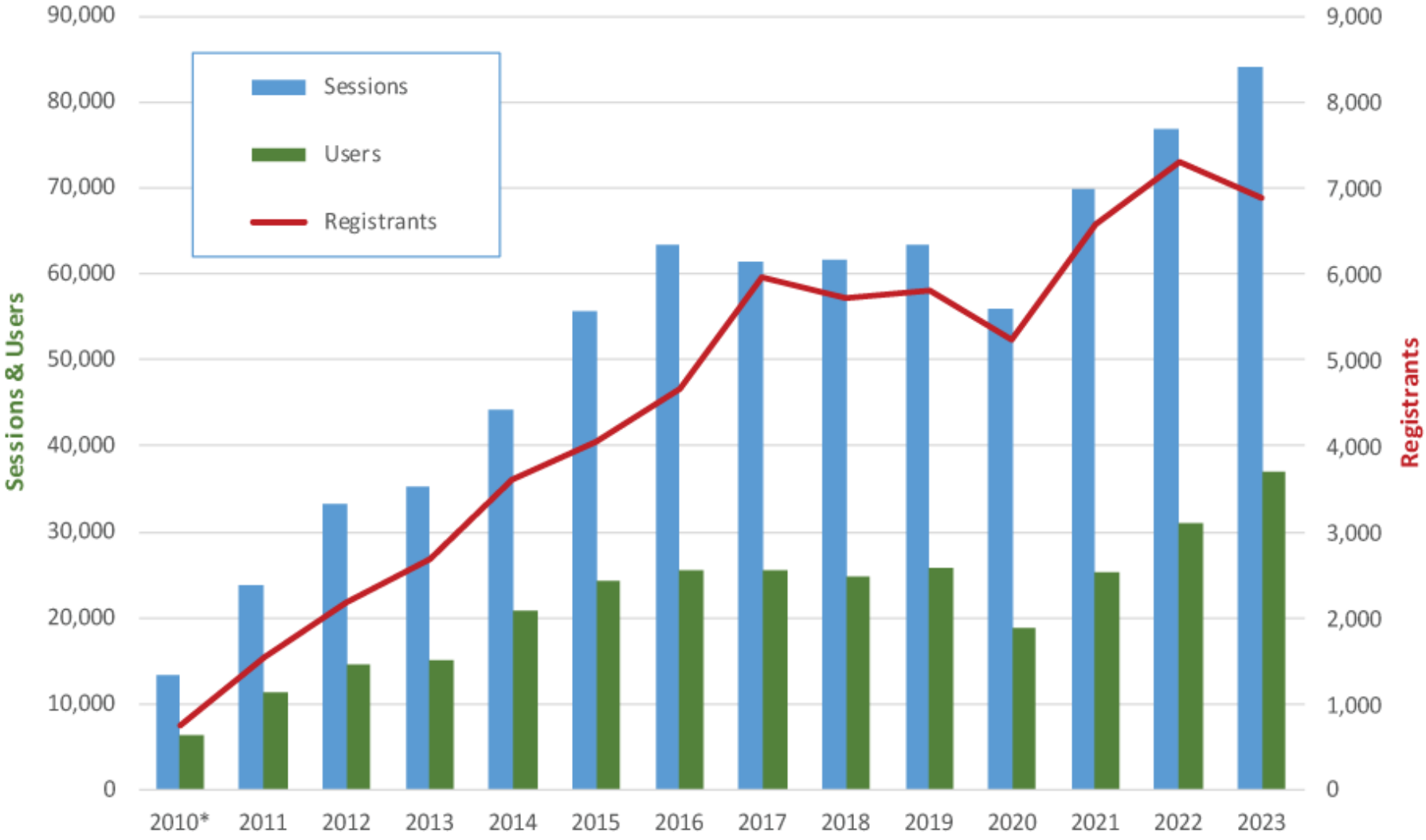

FAST TRACK TO GENE ANNOTATION AND GENOME ANALYSIS
LOG OUT David MIcklos
My Projects
Public Projects
D N A
SUBWAY
Annotate a Genomic Sequence
Find Repeats
Predict Genes
Search Databases
Build Models
Prospect Genomes Using TARGeT
Search Genomes
Alignment & Tree Viewer
Determine Sequence Relationships
Assemble Sequences
Add Sequences
Analyze Sequences
Browsers & Transfer
Next Generation Sequencing
Manage Data
Analyze Transcriptome
Explore Differential Abundance
Metabarcoding Analysis
Metadata + QC
Clustering Sequences
Alpha/Beta Diversity
DNA Subway ties together key bioinformatics tools and databases to assemble gene models, investigate genomes, work with phylogenetic trees and analyze DNA barcodes. Roll over the "stations" on the subway map to find out more about the analysis steps. Analyze your own data or sample data provided. To start a project, select one of the "lines" (red, yellow, blue, green, purple). Register and login to be able to save and share your results.
Java
DNA Barcoding 101
Better with Firefox
Configure Java
Background
Manual
Tour
About
Credits
Resources
Contact Us
CYVERSE
Cold Spring Harbor Laboratory

DNA Subway
an educational *Discovery Environment*

- Developed as educational outreach for iPlant (Cyverse), a 10 year NSF project to develop a computer infrastructure to apply computational thinking to solve big-data biological
- Students “ride” on any of five different lines to access simplified workflows for gene discovery, annotation, and comparison
- 25 collaborators at 11 institutions
- Launched in March 2010



DNA Subway Visitation to Date



DNA Subway Visitation to Date

- 66,652 registered users
- 329,653 total users (including guests)
- 799,058 user sessions, averaging 17 minutes (25.8 person years)
- Blue line analyses account for 71% of traffic
- 277,534 DNA sequences uploaded from GENEWIZ (mainly DNA barcodes)

Pop-up Survey Results (n=5,000)

- 73% said *DNA Subway* is easy to use
- 76% were likely to use in again
- 73% K-16 students, 27% teachers
- Students and faculty were: 75% four-year, 10% CC, 15% precollege
- 41% of non-K-16 students were researchers or graduate students