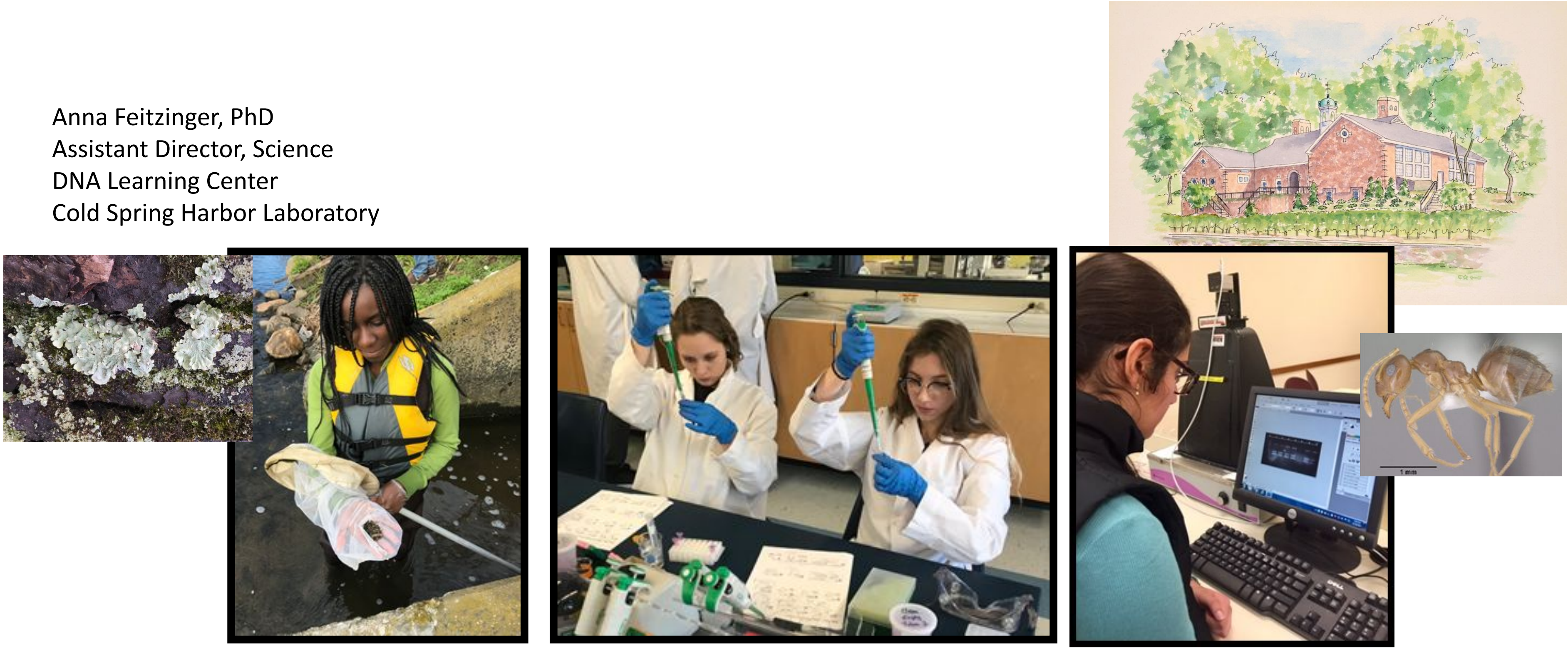
DNA Barcoding

Anna Feitzinger, PhD
Assistant Director, Science
DNA Learning Center
Cold Spring Harbor Laboratory

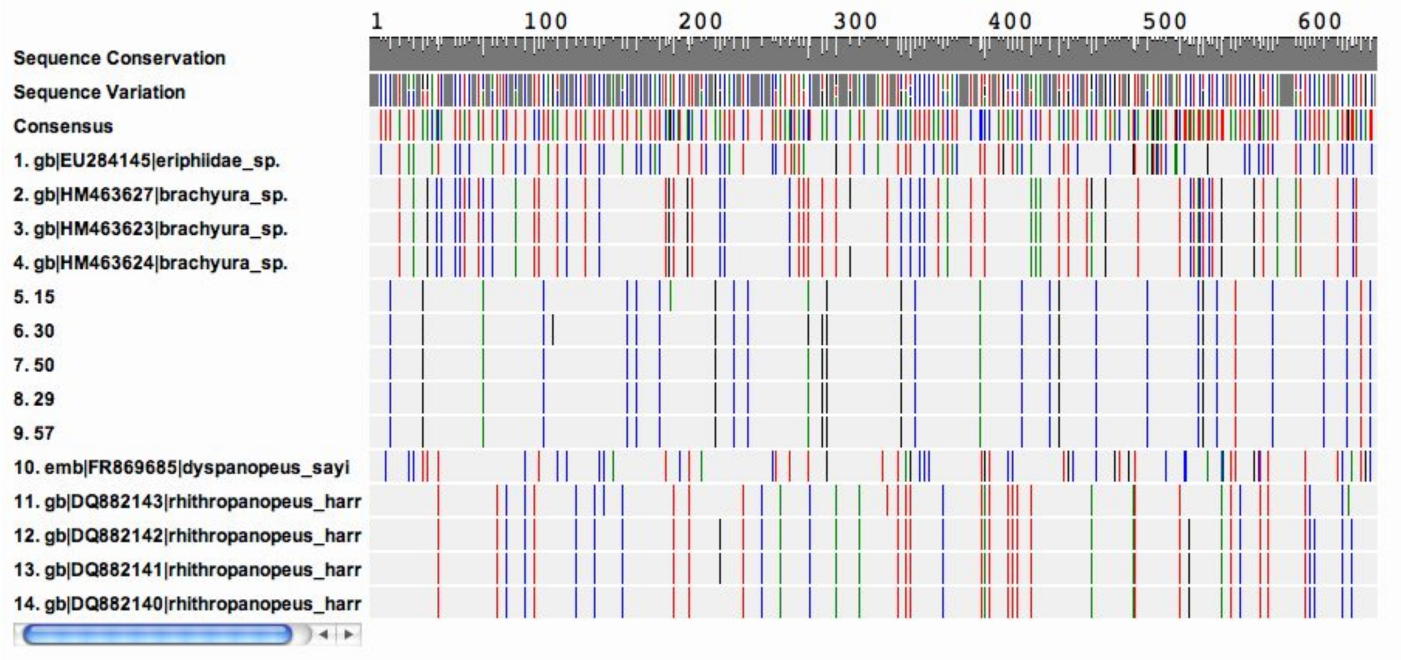


DNA Subway 2.0

CSH Cold Spring Harbor Laboratory
DNA Learning Center

What is DNA Barcoding?

Just as the unique pattern of bars in a universal product code (UPC) identifies each consumer product, a short “DNA barcode” (about 600 nucleotides in length) is a unique pattern of DNA sequence that can potentially identify each species



What is DNA Barcoding?

Organism is sampled

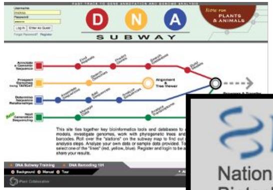
DNA is extracted

“Barcode” region amplified

ACGAGTCGGTAGCTGCCCTCTGACTGCATCGAA
TTGCTCCCCTACTACGTGCTATATGCGCTTACGA
TCGTACGAAGATTTATAGAATGCTGCTAGCTGC
TCCCTTATTCGATAACTAGCTCGATTATAGCTA

=

Sequenced DNA creates a unique “barcode” for each species



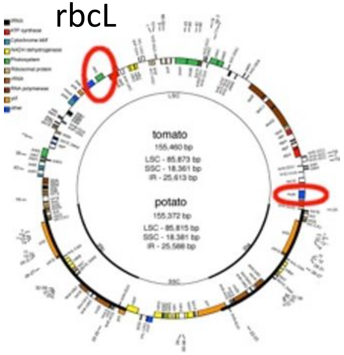
NCBI
National Center for Biotechnology Information

BOLDSYSTEMS

ID from databases

“Universal” DNA Barcodes

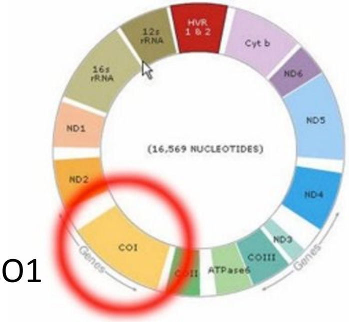
Plants: Chloroplast



rbcL

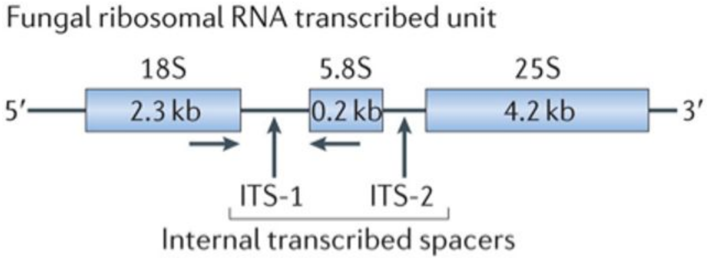
matK

Animals: Mitochondrion

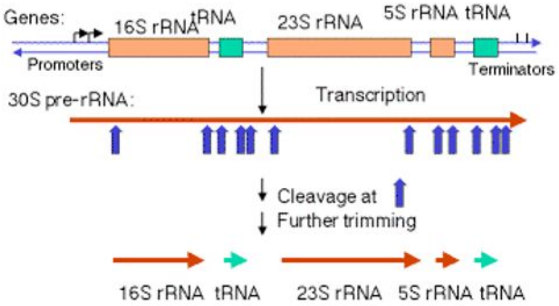


CO1

Fungi: Nucleus

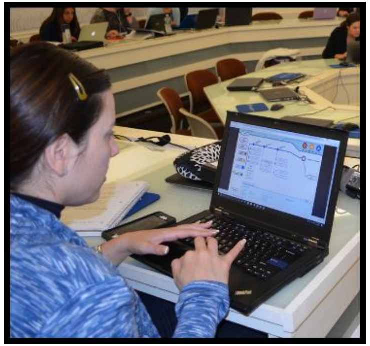


Bacteria



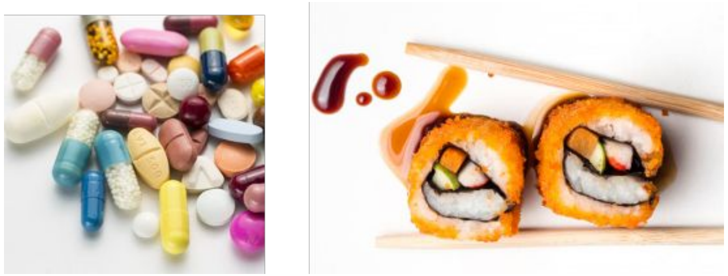
Advantages of DNA Barcoding: Identification

Non-experts can identify specimens



Can ID incomplete, damaged, or processed samples

Advantages of DNA Barcoding: Single infrastructure supports a range of distributed projects


Identify food fraud


Monitor disease vectors


Morphological dopplegangers


Establish biodiversity inventories

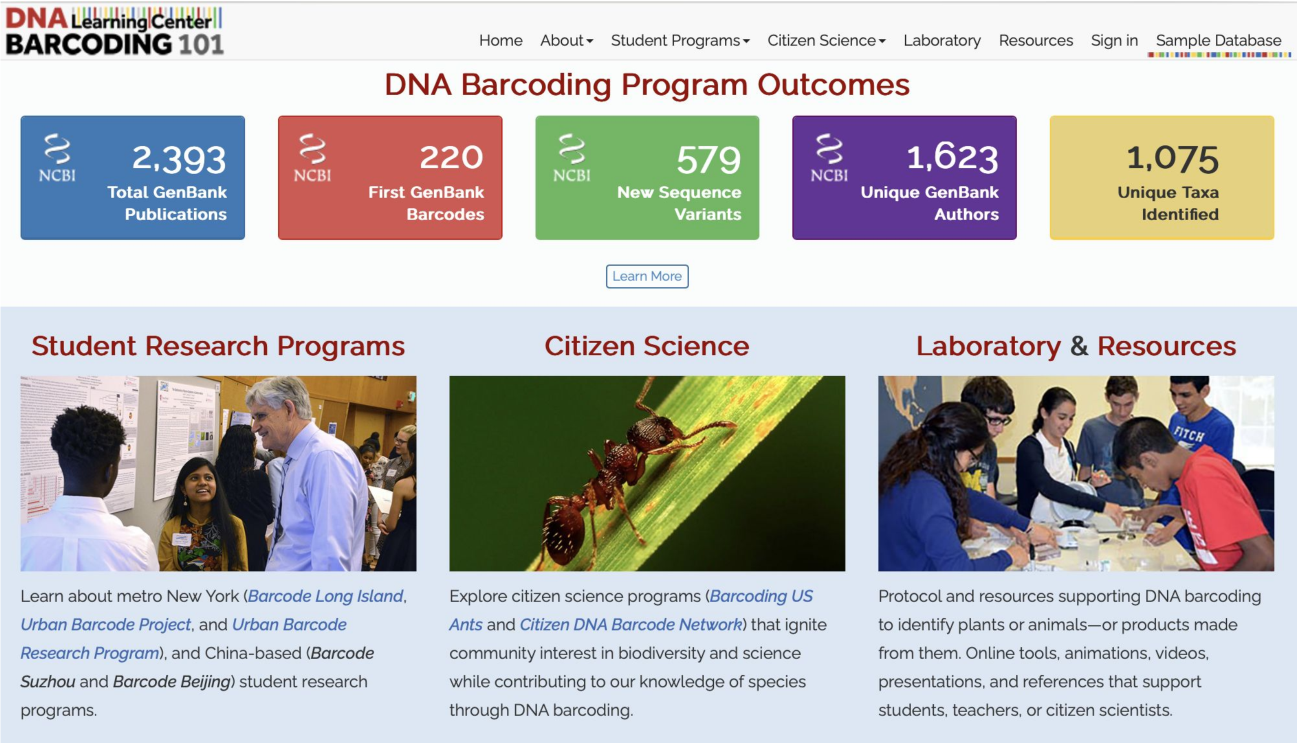

Under-described life stages


Determine unknown diets

...and much more!


Combat poaching

DNA Barcoding 101



DNA Barcoding Program Outcomes

2,393 Total GenBank Publications

220 First GenBank Barcodes

579 New Sequence Variants

1,623 Unique GenBank Authors

1,075 Unique Taxa Identified

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.

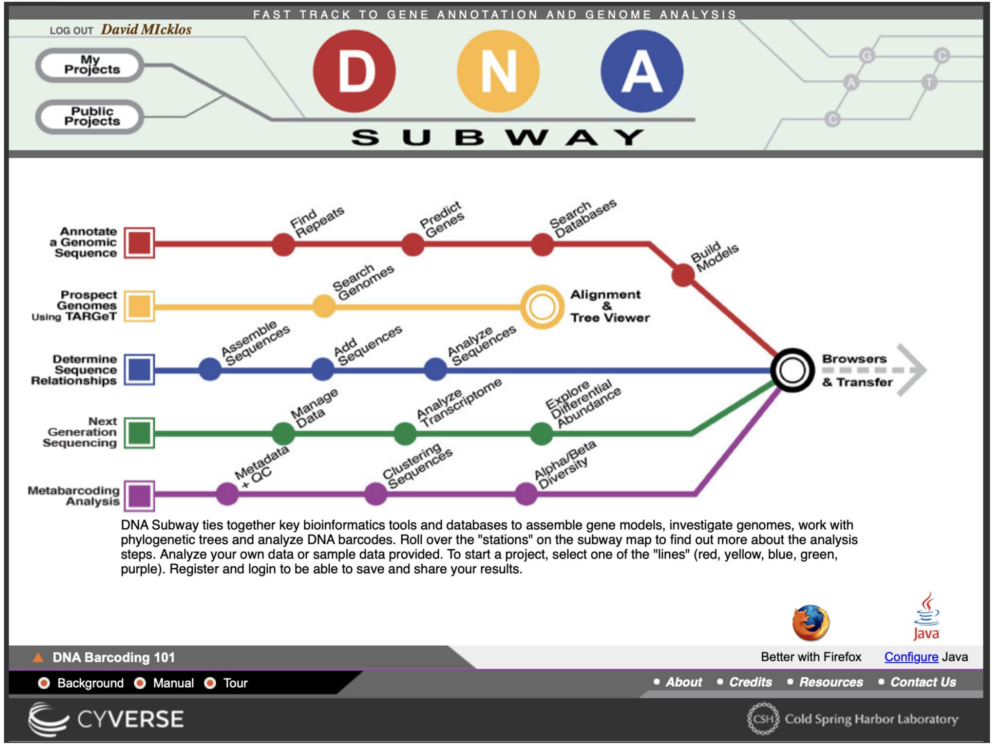
Citizen Science

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.

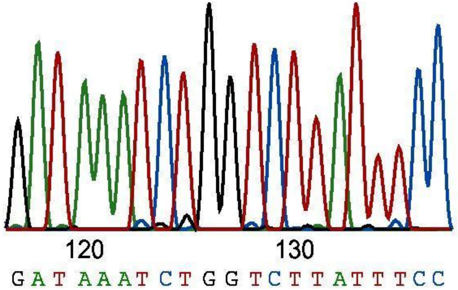
DNA Subway 1.0

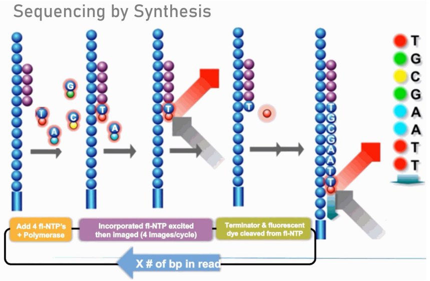


DNA Subway ties together key bioinformatics tools and databases to assemble gene models, investigate genomes, work with phylogenetic trees and analyze DNA barcodes. Roll over the "stations" on the subway map to find out more about the analysis steps. Analyze your own data or sample data provided. To start a project, select one of the the "lines" (red, yellow, blue, green, purple). Register and login to be able to save and share your results.

Sequencing Technologies

Short-read





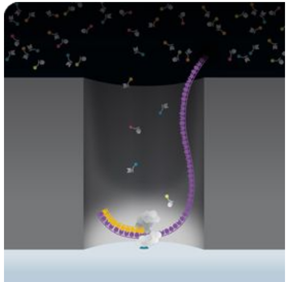
Sanger

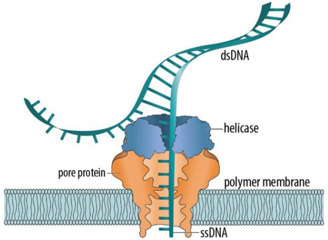
- 100-1000 bp
- Low-throughput
- Low error

illumina

- 100-300 bp
- High-throughput
- Low error
- Substantial startup ($10K+)

Long-read





PacBio

- 1-50 kbp
- High-throughput
- Low error
- Substantial startup ($10K+)

Nanopore

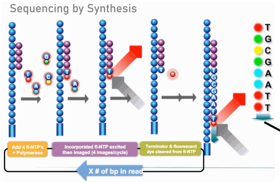
- 0.5 kbp - millions
- High-throughput
- Moderate error
- Low startup ($2K+)

*Only sequencing technology that can be done in the classroom

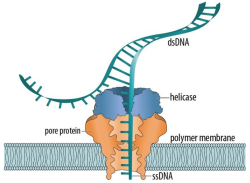
Nanopore Compatible DNA Subway

DNA is extracted

“Barcode” region amplified



Sanger Sequencing



Nanopore Sequencing (in the classroom)

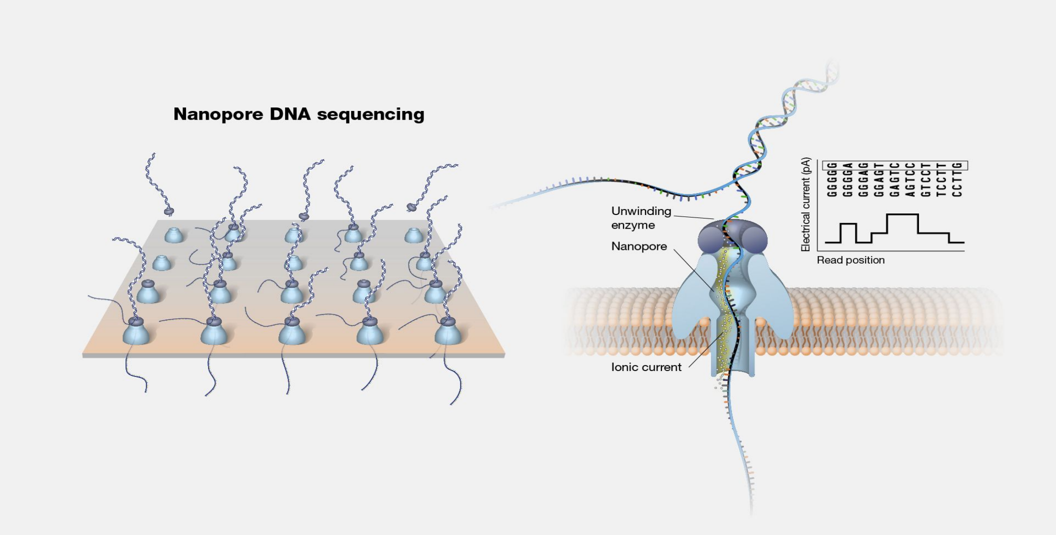
DNA Subway

Bioinformatics Analysis
Blue Line

ID from databases

What is Nanopore Sequencing?

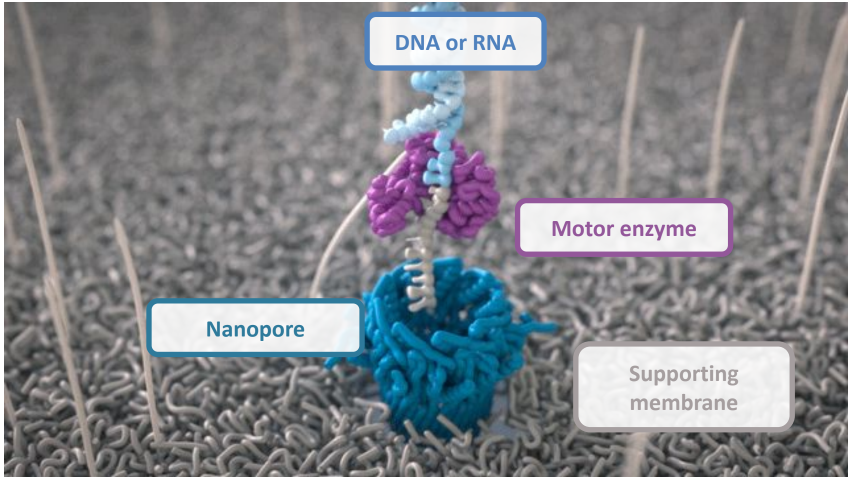
Nanopore sequencing is a technology that enables direct, real-time analysis of any length of DNA or RNA fragments. It works by monitoring changes to an electric current as nucleic acids are passed through a protein nanopore.

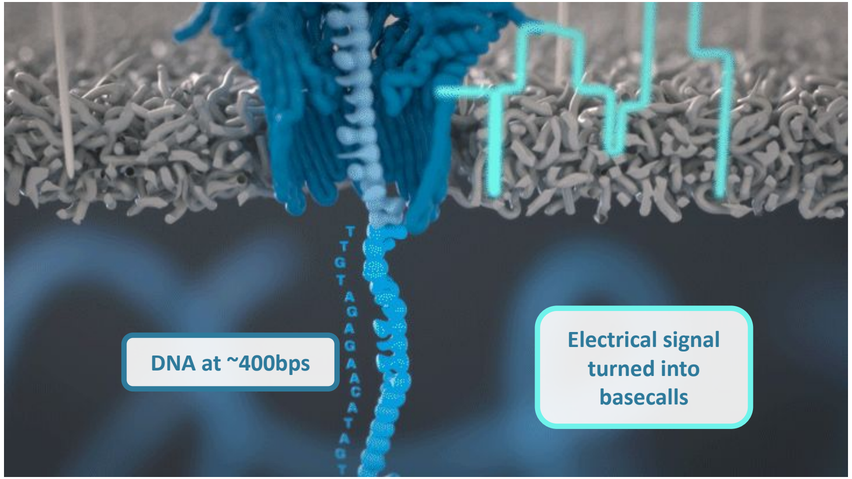


How Does Nanopore Sequencing Work?

- A DNA / RNA strand is passed through a nanopore

An electrical signal is interpreted into sequence data





The advantages of nanopore sequencing

Real-time Analysis

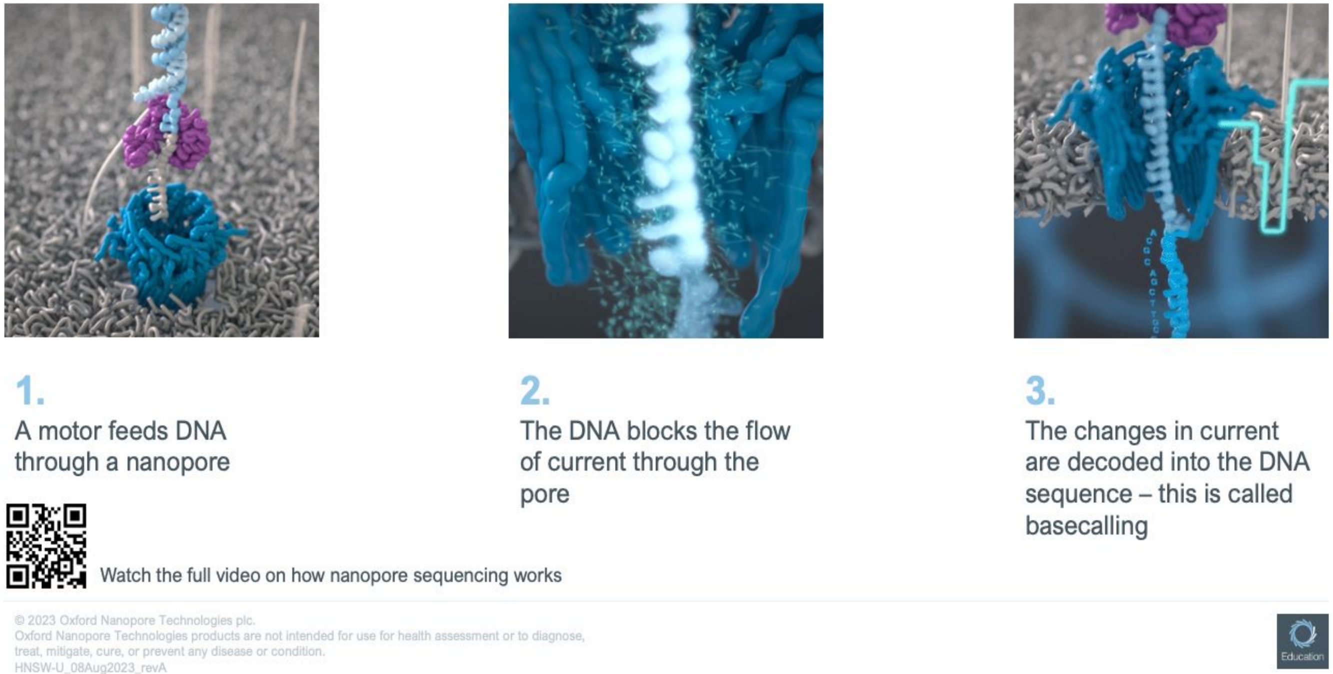
PCR free, no amplification bias

Modified base detection

Read length-agnostic

Direct sequencing of DNA / RNA

Sequencing Technologies



1.
A motor feeds DNA through a nanopore

2.
The DNA blocks the flow of current through the pore

3.
The changes in current are decoded into the DNA sequence – this is called basecalling

Watch the full video on how nanopore sequencing works

HNSW-U_08Aug2023_revA

Sequencing Technologies

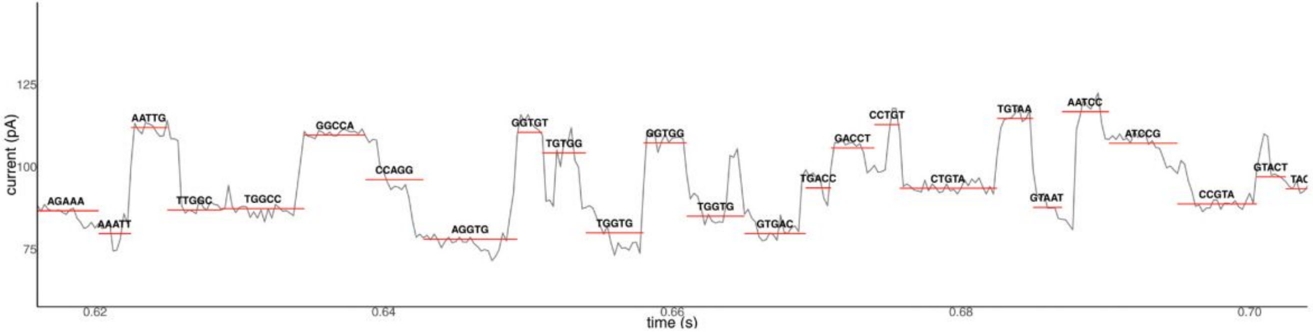
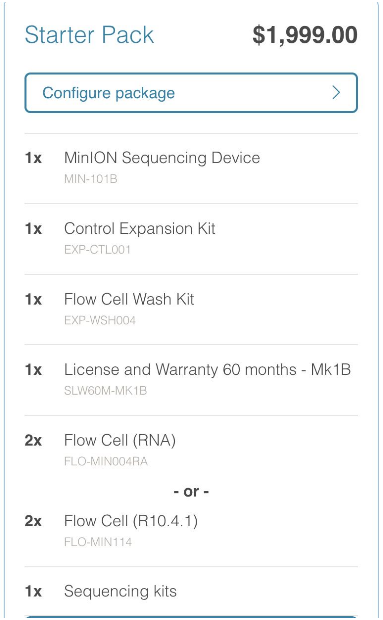


Figure 3. Example of a MinION DNA read as raw data (grey line) and the event data (red lines) extracted from it, corresponding to discrete sets of bases. For the sake of illustration it is assumed that five bases influence the current at a given time, although in reality this assumption may not always hold. Data used in this figure was obtained from the Nanopore WGS consortium (third release)[3].

Current Change Across the Pore → DNA Sequences (4 base pairs)

Requirements



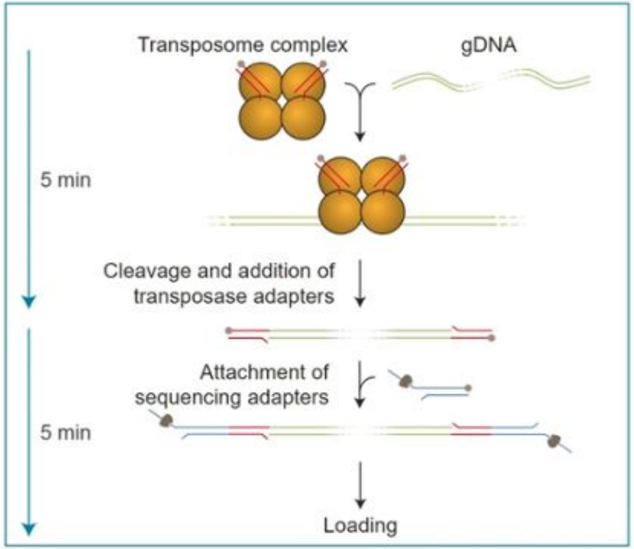
Starter Pack **$1,999.00**

1x	MinION Sequencing Device MIN-101B
1x	Control Expansion Kit EXP-CTL001
1x	Flow Cell Wash Kit EXP-WSH004
1x	License and Warranty 60 months - Mk1B SLW60M-MK1B
2x	Flow Cell (RNA) FLO-MIN004RA
	- or -
2x	Flow Cell (R10.4.1) FLO-MIN114
1x	Sequencing kits

Workflow

PCR

Library Preparation
Rapid Barcoding Kit

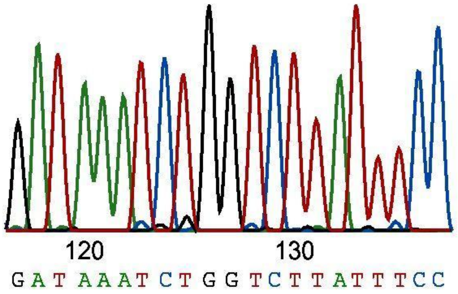


Load On Flongle

Sequence

- Rapid barcoding kit quick enough to do during class - add rapid barcode directly to PCR product
- Flongle - No need for very high through put
- Even with fragmentation can get consensus of amplicons ~600bps

Data output

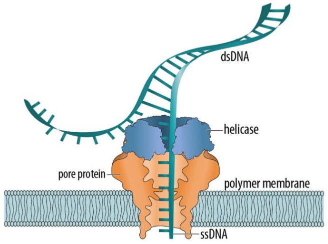


Sanger

.abi files (electropherograms)

↓

.fasta (one forward sequence and one reverse sequence per sample)



Nanopore

.fast5 (current change)

↓

.fastq (many sequences with associated quality scores)

Forensics Crime Lab

- Forensic Entomology – using the identity of insects found on a cadaver to estimate the post-mortem interval (PMI)
- Specimens field collected from students of the Forensic Anthropology Center (FAC) at the University of TN, Knoxville.

- Use DNA Barcoding to identify samples from the Forensics Anthropology Center to help determine post-mortem interval.

Jeffry Petracca

Sanger Sequences didn't come back in time!

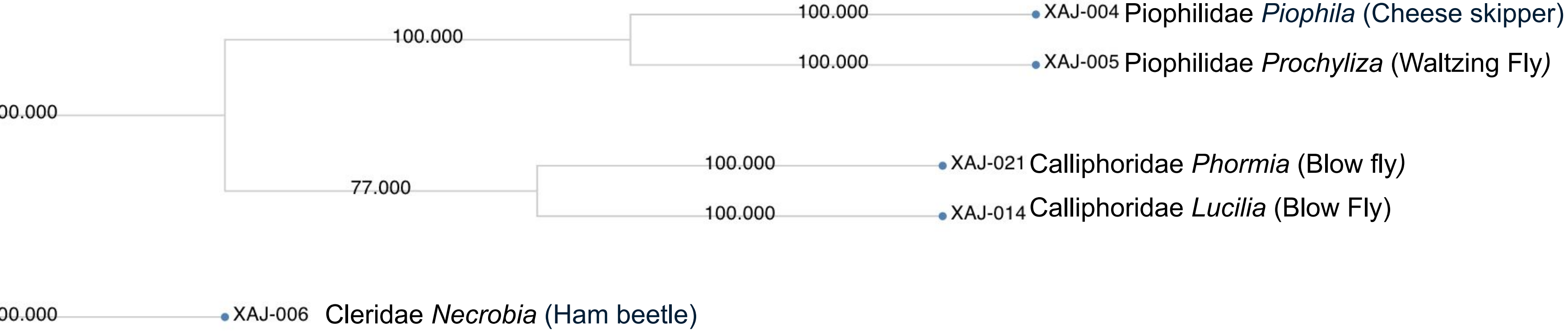
Demo Link

https://dnasubwayv2.dnalc.org/

Username: barcoding

Password: atemeeting

Forensics Crime Lab



Piophilidae *Piophila* (Cheese skipper)

Piophilidae *Prochyliza* (Waltzing Fly)

Calliphoridae *Phormia* (Blow fly)

Calliphoridae *Lucilia* (Blow Fly)

Cleridae *Necrobia* (Ham beetle)