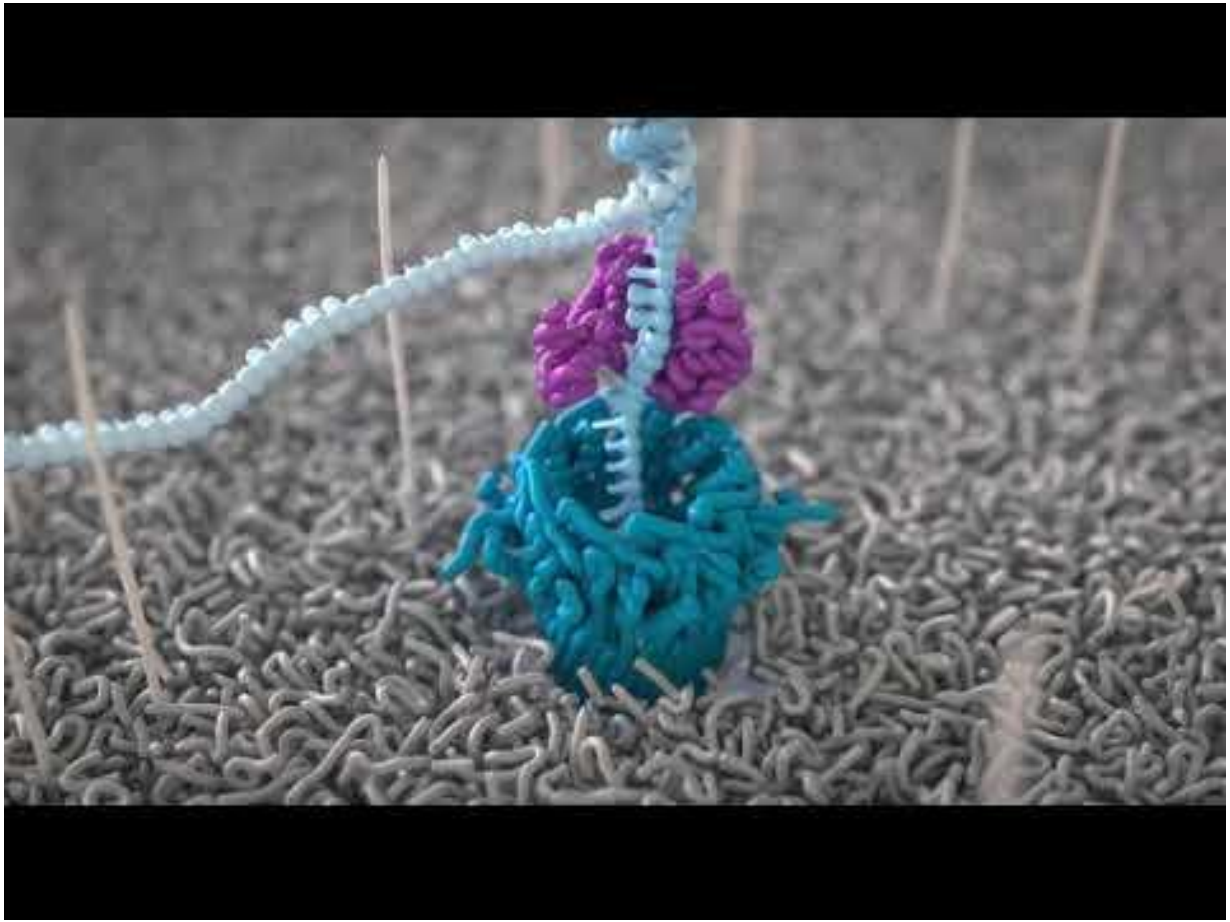
SequenceR

How nanopore sequencing works



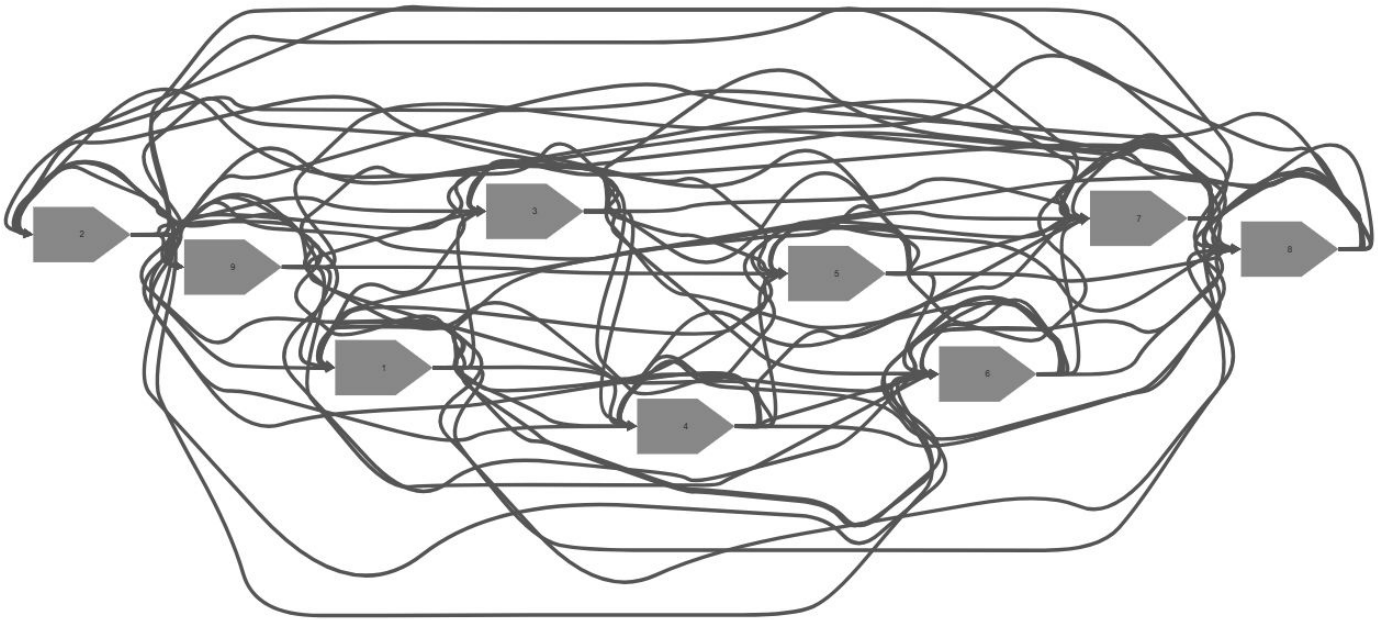
Interactive demo: STAMPS sequencer (10-15 minutes)

- Each of you represents a single nanopore (we should have 35? total)
- TAs will hand out the microbial cell
- Once I say go:
  - Lyse DNA (open egg)
  - Sequence DNA with your STAMPS sequencer (pass yarn through plastic needle)
  - Record how long it takes for you to lyse, extract, and pass your DNA through your nanopore:
    - https://docs.google.com/spreadsheets/d/1biKZKh17Ro8MyPL_xYXdEEIx1SuAsa_hLu92OzEsppY/edit?usp=sharing
- **Red** sticky note if you are unable to lyse the cell, **Green** sticky note once finished.

AsseMBLr

Example 1: AAAAAAAAAAAA (no unique 3-mer)

- 4bp reads:
  - (1) AAAA, (2) AAAA, (3) AAAA, (4) AAAA, (5) AAAA, (6) AAAA, (7) AAAA, (8) AAAA, (9) AAAA
- 3-mers:
  - AAA
- Overlaps:
  - (1)->(2),(1)->(3),(1)->(4),(1)->(5),(1)->(6),(1)->(7),(1)->8,(1)->(9)
  - (2)->(1),(2)->(3),(2)->(4),(2)->(5),(2)->(6),(2)->(7),(2)->8,(2)->(9)
  - ….
  - (9)->(1),(9)->(2),(9)->(3),(9)->(4),(9)->(5),(9)->(6),(9)->(7),(9)->(8)


2
9
3
1
4
5
6
7
8

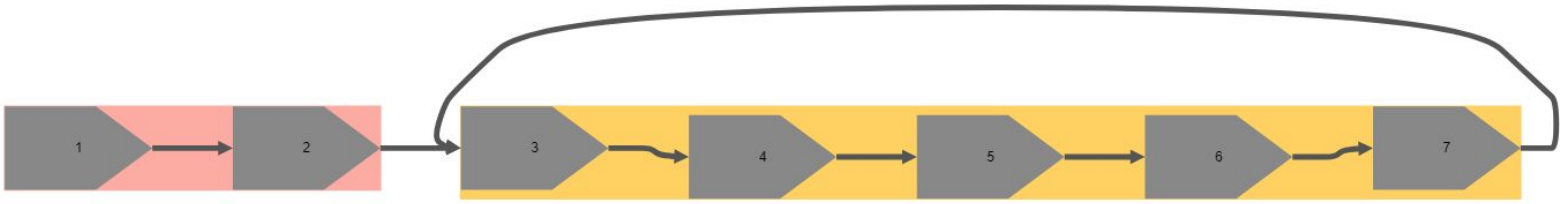
Example 2: AATCCGTTCGGA (no 3-mer repeats)

- 4bp reads:
  - (1) AATC, (2) ATCC, (3) TCCG, (4) CCGT, (5) CGTT, (6) GTTC, (7) TTCG, (8) TCGG, (9) CGGA
- 3-mers (10)
  - (i) AAT, (ii) ATC, (iii) TCC, (iv) CCG, (v) CGT, (vi) GTT, (vii) TTC, (viii) TCG, (ix) CGG, (x) GGA
- Overlaps
  - (1)->(2)
  - (2)->(3)
  - (3)->(4)
  - (4)->(5)
  - (5)->(6)
  - (6)->(7)
  - (7)->(8)
  - (8)->(9)


1
2
3
4
5
6
7
8
9

Example 3: AATCCGTTCGGA (sequencing error)

- 4bp reads:
  - (1) AATC, (2) ATCC, (3) TCCG, (4) CCGT, (5) CGTT, (6) GTTC, (7) TTC**C**, (8) TCGG, (9) CGGA
- 3-mers (10)
  - (i) AAT, (ii) ATC, (iii) TCC (X2), (iv) CCG, (v) CGT, (vi) GTT, (vii) TTC, (viii) TCG, (ix) CGG, (x) GGA
- Overlaps
  - (1)->(2)
  - (2)->(3)
  - (3)->(4)
  - (4)->(5)
  - (5)->(6)
  - (6)->(7)
  - ~~(7)->(8)~~
  - (7)->(3)
  - (8)->(9)


1
2
3
4
5
6
7

Mystery “metagenomic” sample has been sequenced by the STAMPS sequencing machine

- The STAMPS sequencer is only capable of generating **30 bp** reads
- You’re in luck, we know that each read has a 12bp match with the subsequent read (dovetail overlap)

- (discuss: why is this helpful for metagenomic assembly?)

The STAMPS assembler is YOU; there are 38 cores that work in parallel on this task