Genetics v Genomics

C. Ryan Campbell

Duke University

c.ryan.campbell@duke.edu

05 Sept 2017

Overview

Today's Goals

- What is/are Genomics?
- How have techniques changed?
- What impact has that had on biological questions?

What is/are Genomics?

- What is Genomics?
- How is it different than Genetics?
- What allows us to do genomics instead of genetics?

What is/are Genomics?

1. Pair up

What is/are Genomics?

1. Pair up
2. Settle on a definition of genomics (or an answer to the other questions)

What is/are Genomics?

1. Pair up
2. Settle on a definition of genomics (or an answer to the other questions)
3. Post it to DotStorming

What is/are Genomics?

1. Pair up
2. Settle on a definition of genomics (or an answer to the other questions)
3. Post it to DotStorming
4. Vote on other answers

What is/are Genomics?

What is/are Genomics?

- Genomics is the study of genomes, the complete set of genetic material within an organism

What is/are Genomics?

- Genomics is the study of genomes, the complete set of genetic material within an organism
- Genomics involves the sequencing and analysis of genomes

What is/are Genomics?

- Genomics is the study of genomes, the complete set of genetic material within an organism
- Genomics involves the sequencing and analysis of genomes
- Genomics is also concerned with the structure, function, comparison, and evolution of genomes

What is/are Genomics?

- Genomics is the study of genomes, the complete set of genetic material within an organism
- Genomics involves the sequencing and analysis of genomes
- Genomics is also concerned with the structure, function, comparison, and evolution of genomes
- The field also includes studies of intragenomic (within the genome) phenomena such as heterosis (hybrid vigour), epistasis (effect of one gene on another), pleiotropy (one gene affecting more than one trait) and other interactions between loci and alleles within the genome.

What is/are Genomics?

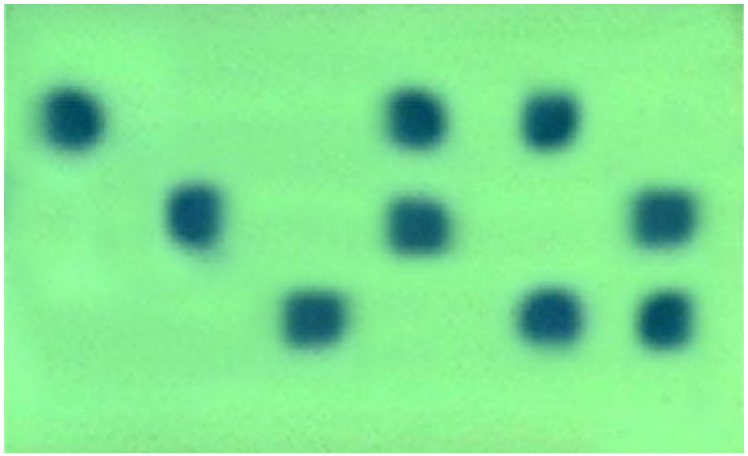
- Genomics is the study of genomes, the complete set of genetic material within an organism
- Genomics involves the sequencing and analysis of genomes
- Genomics is also concerned with the structure, function, comparison, and evolution of genomes
- The field also includes studies of intragenomic (within the genome) phenomena such as heterosis (hybrid vigour), epistasis (effect of one gene on another), pleiotropy (one gene affecting more than one trait) and other interactions between loci and alleles within the genome.
- In contrast to genetics, which refers to the study of individual genes and their roles in inheritance, genomics uses high throughput DNA sequencing and bioinformatics to assemble, and analyze the function and structure of entire genomes.

Brief History of Sequencing

- Allozymes
- Sanger Sequencing
- Next Generation Sequencing - NGS

Allozymes

- 1960's
- Electrophoresis separates different proteins by amino acid makeup
- First (limited) look at DNA composition

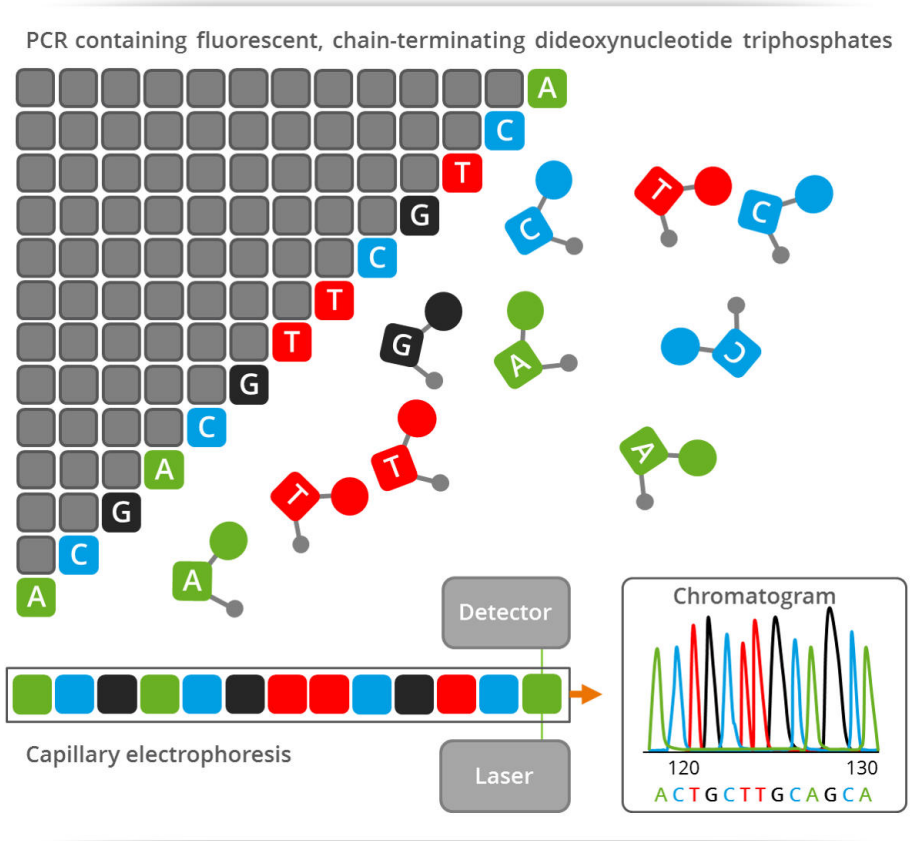


Sanger Sequencing

- 1977
- Determines the sequences a single piece of DNA up to 500bp
- Highly accurate but slow

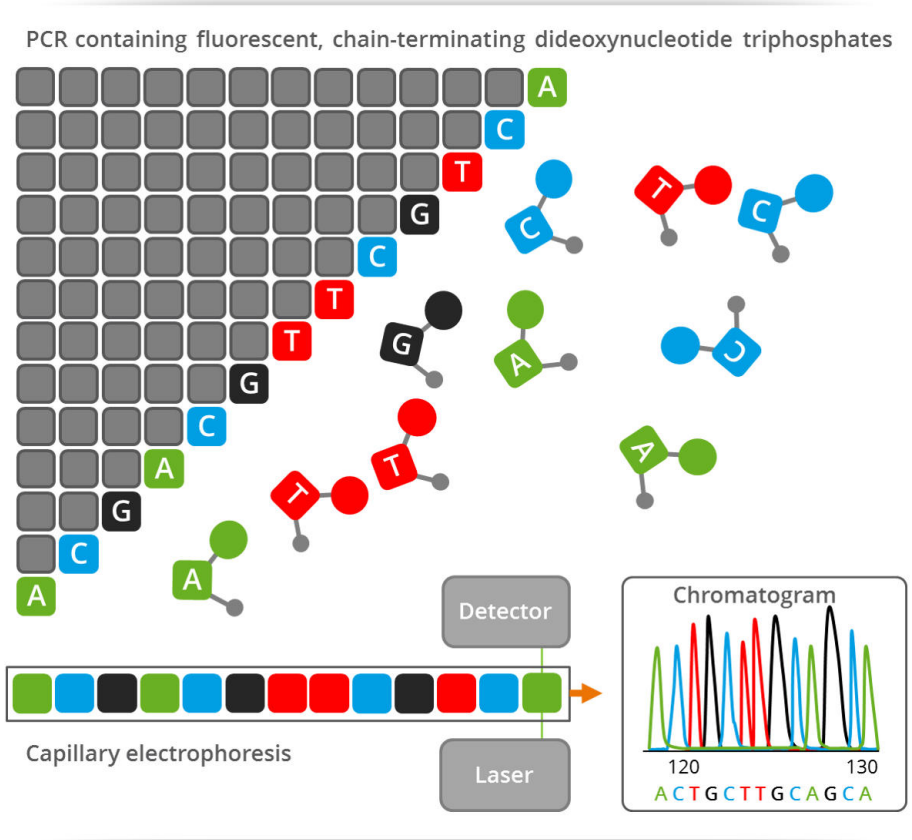
Sanger Sequencing

- Design a primer
- Run a PCR
- Chain-terminating dideoxynucleotide triphosphates



Sanger Sequencing

- Run results on a gel
- Read with a laser, determines which base ended the PCR
- Color order is sequence order

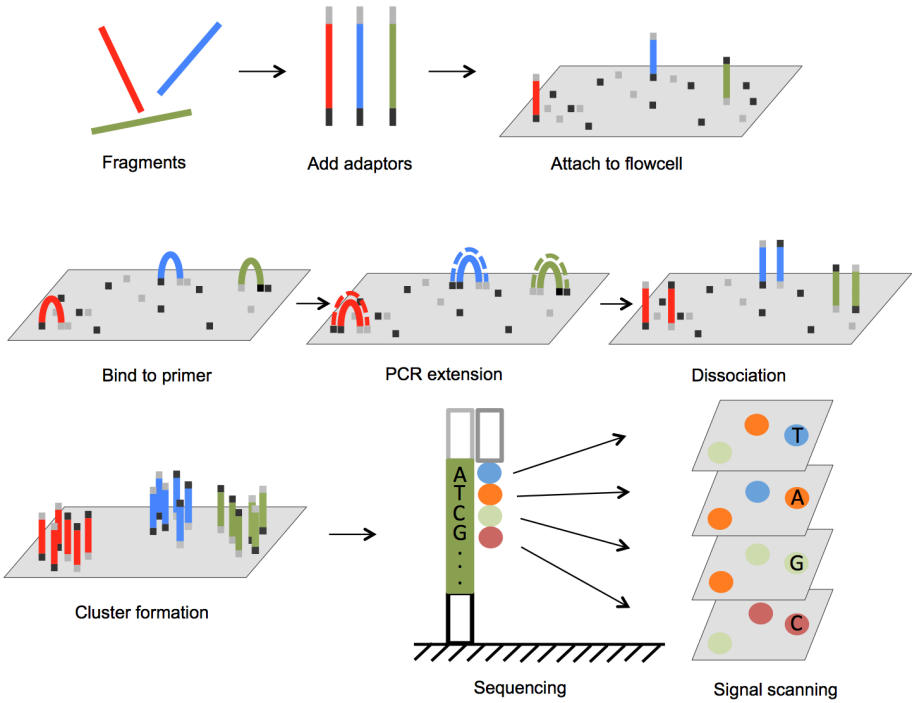


NGS Sequencing

- mid-2000's
- Many different companies and methods
- All generate far more data than Sanger

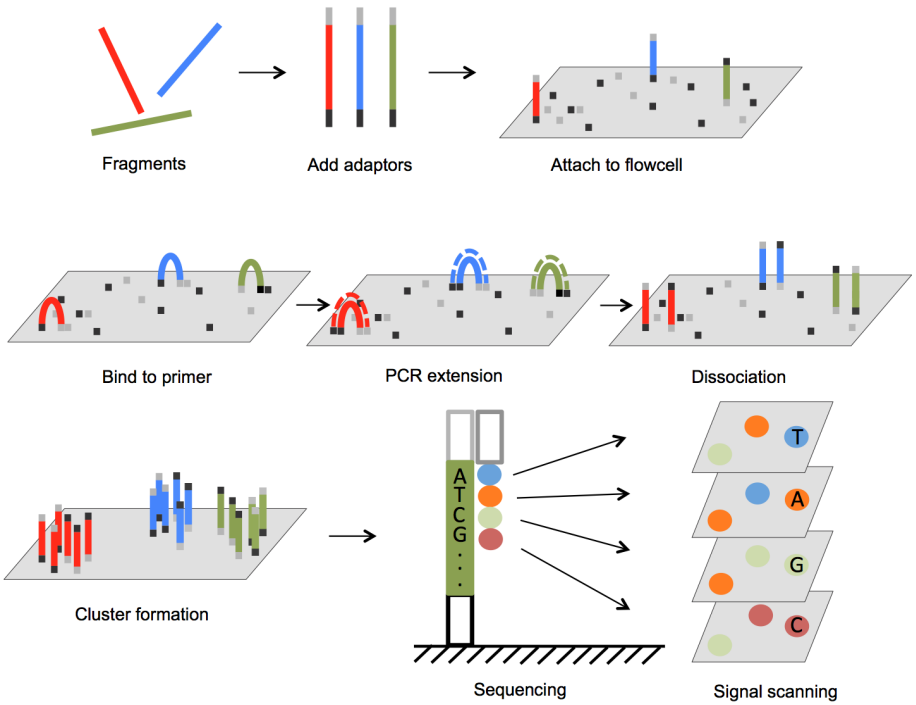
NGS Sequencing - Illumina

- Library Preparation
  - Fragment a sample of whole genomic DNA
  - Add adapters for the specific machine
- Amplify with PCR
- Read on machine (next slide)



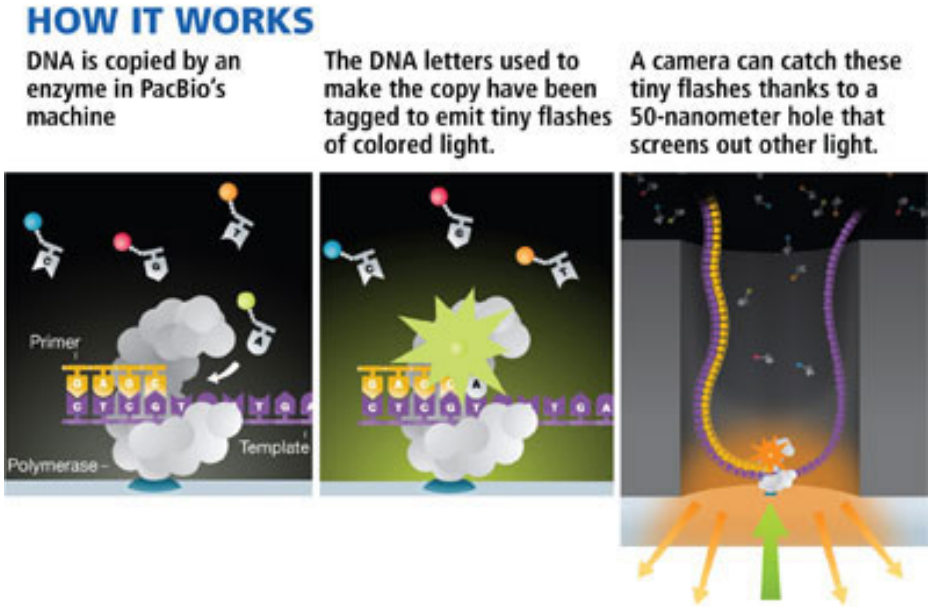
NGS Sequencing - Illumina

- Machine attaches adapter and DNA to a fixed surface
- Amplifies single strand
- Adds a new base each cycle and images for ID



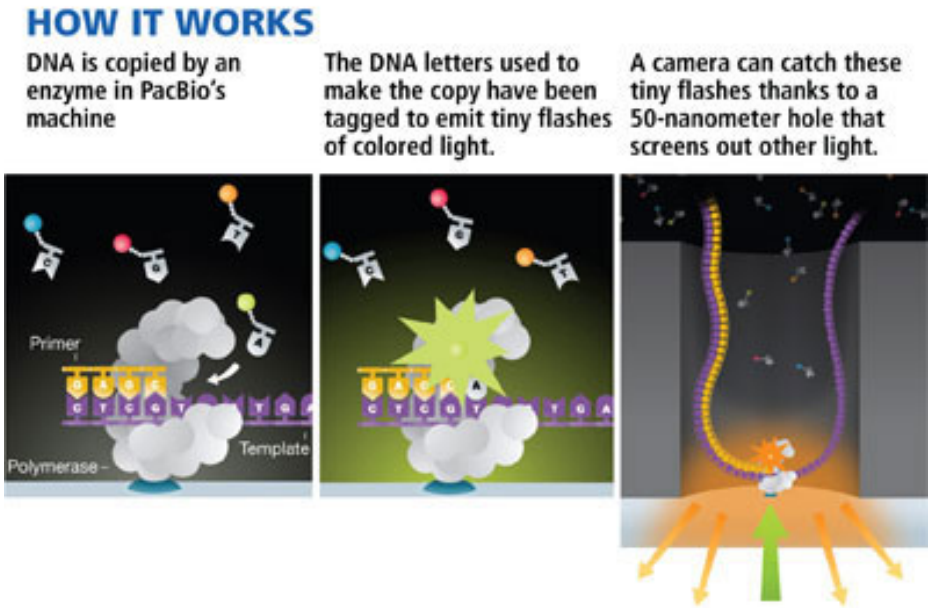
NGS Sequencing - PacBio

- Library Preparation
  - Fragment a sample of whole genomic DNA
  - Add adapters for the specific machine
- Amplify with PCR
- Read on machine



NGS Sequencing - PacBio

- A copy is made on the machine by an enzyme
- The bases used for the copy are flourescent
- As a new base is incorporated the color shows the identity



Next Generation Sequencing vs. Sanger

- Output for Quality Tradeoff
  - NGS = VERY High Output / Good Quality
  - Sanger = Low Output / Best Quality

Next Generation Sequencing vs. Sanger

- Output for Quality Tradeoff
  - NGS = VERY High Output / Good Quality
  - Sanger = Low Output / Best Quality
- NGS Methods and Machines

Next Generation Sequencing vs. Sanger

- Output for Quality Tradeoff
  - NGS = VERY High Output / Good Quality
  - Sanger = Low Output / Best Quality
- NGS Methods and Machines
  - PyroSeq - 454

Next Generation Sequencing vs. Sanger

- Output for Quality Tradeoff
  - NGS = VERY High Output / Good Quality
  - Sanger = Low Output / Best Quality
- NGS Methods and Machines
  - PyroSeq - 454
  - Illumina

Next Generation Sequencing vs. Sanger

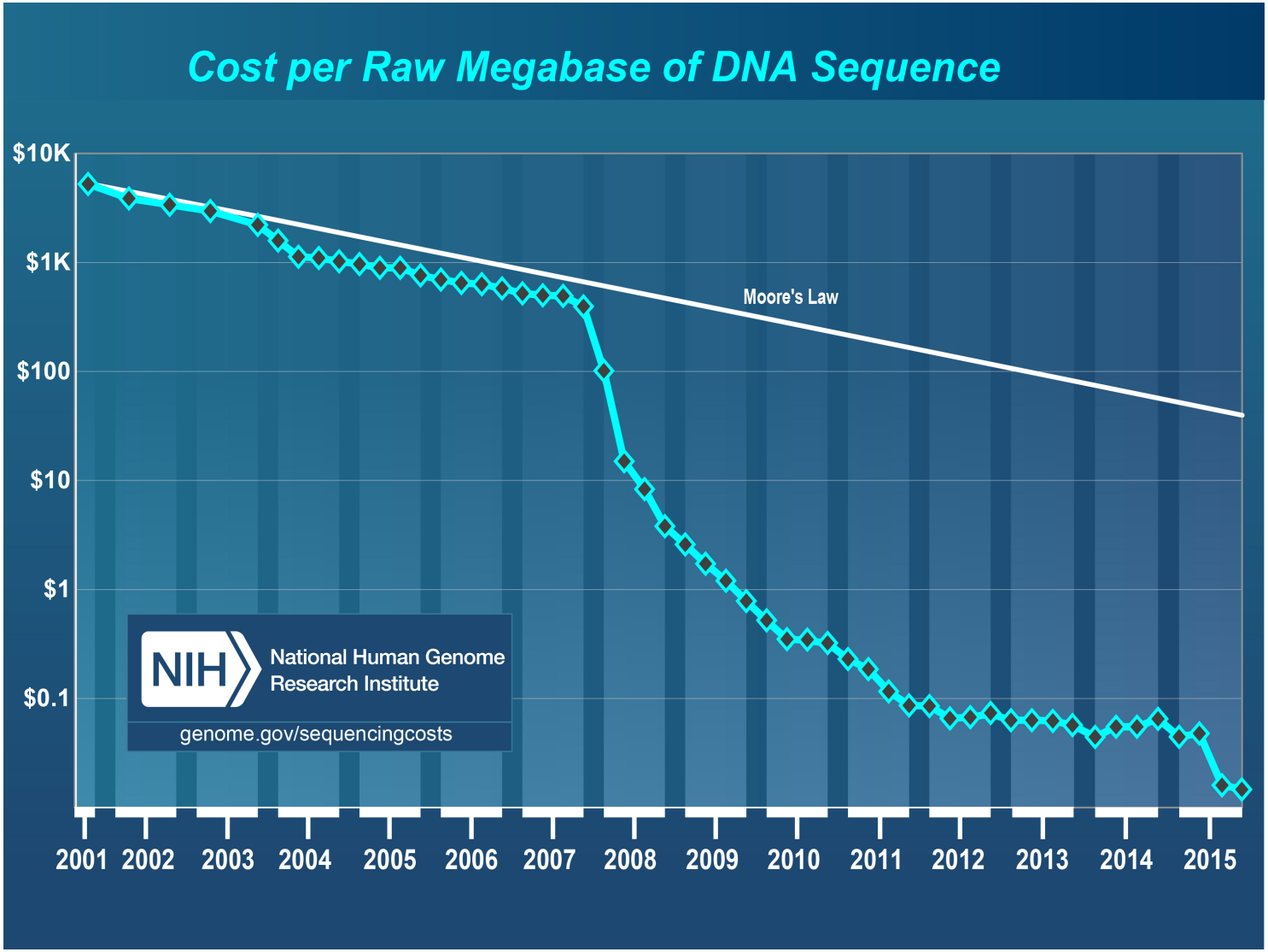
- Output for Quality Tradeoff
  - NGS = VERY High Output / Good Quality
  - Sanger = Low Output / Best Quality
- NGS Methods and Machines
  - PyroSeq - 454
  - Illumina
  - PacBio

Brief History of Sequencing

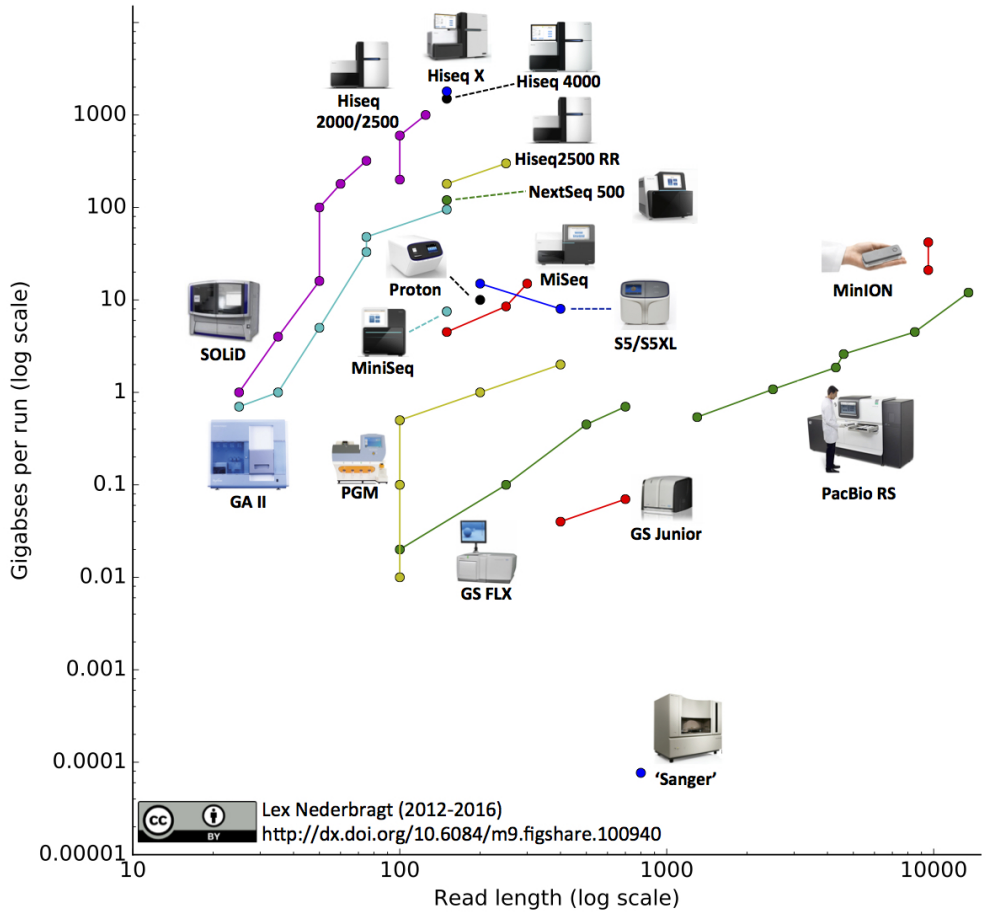
- Allozymes
  - 1960's
- Sanger Sequencing
  - 1977
- NGS - Next Generation Sequencing
  - 2000


Miles Zhang
@ymilesz
Following
This meme is everywhere, so I thought I'd add a biology twist to it.
BIOLOGISTS
NGS
SANGER SEQUENCING
7:11 PM - 25 Aug 2017

Sequencing Cost



Sequencing Output



Generational Shift

- More and more data can be generated

Generational Shift

- More and more data can be generated
- Data length and quality are both improving

Generational Shift

- More and more data can be generated
- Data length and quality are both improving
- How does this change the scope of research?

In-Class Activity

1. Make 4 groups (rearrange desks)

In-Class Activity

1. Make 4 groups (rearrange desks)
2. Divide up the following Biological Questions

In-Class Activity

1. Make 4 groups (rearrange desks)
2. Divide up the following Biological Questions
   - How similar are two species?

In-Class Activity

1. Make 4 groups (rearrange desks)
2. Divide up the following Biological Questions
   - How similar are two species?
   - What genes underlie a specific function?

In-Class Activity

1. Make 4 groups (rearrange desks)
2. Divide up the following Biological Questions
   - How similar are two species?
   - What genes underlie a specific function?
   - What is the code of the human genome?

In-Class Activity

1. Make 4 groups (rearrange desks)
2. Divide up the following Biological Questions
   - How similar are two species?
   - What genes underlie a specific function?
   - What is the code of the human genome?
   - Group's Choice:
     - Transcription Factors
     - CpG islands
     - DNA Methylation

In-Class Activity

1. Make 4 groups (rearrange desks)
2. Divide up the following Biological Questions
   - How similar are two species?
   - What genes underlie a specific function?
   - What is the code of the human genome?
   - Group's Choice:
     - Transcription Factors
     - CpG islands
     - DNA Methylation
3. How would your group assess this question
   - in 1997?
   - in 2017?

- Group Choice Order:

- Group Choice Order:
- Close your laptops/phones

- Group Choice Order:
- Close your laptops/phones
- In your group take 60 seconds to answer the following:

- Group Choice Order:
- Close your laptops/phones
- In your group take 60 seconds to answer the following:
- How many basepairs are in the human genome?

- Group Choice Order:
- Close your laptops/phones
- In your group take 60 seconds to answer the following:
- How many basepairs are in the human genome?
- _____

- Group Choice Order:
- Close your laptops/phones
- In your group take 60 seconds to answer the following:
- How many basepairs are in the human genome?
- _____
- 3,234.83 Mb

Function

- 2017
- easily ask question bc you can completely sequence the genome
- RNAseq
- knockout screen - knock sections out and see changes
- 1997
- asl the question in a limited way - how does a single gene alter a single function

Transcription Factor

- Transcription Factor - protien that regulates the transcription of a gene
- feedback loops in gene pathways
- 1997
- knockouts and broad genetic methods
- ask about a known TF
- 2017
- CHIPseq - chromatin immunoprecipitation seq
- digest DNA, protect areas that are bound to protein
- assess which TFs are present
- correlative but guides future research

Genome Assembly

- 1997
- broke down the genome into 150k bp frags
- cloned fragments
- broke down again
- cloned and Sanger sequenced
- 2017
- easy to get a lot of data
- hard to assess repetitive regions
- combine illumina and pacbio data

PopGen

- 1997
- sanger sequence conserved genes
- 2017
- whole genome seq
- RADseq - focus on shared areas across species

The End