Next Generation Sequencing

Collin Nguyen and Abby Laumer

What is DNA?

base pair

hydrogen bonds

Carries the genetic recipe that makes up all of life

Adenine (A)

Thymine (T)

Cytosine (C)

Guanine (G)
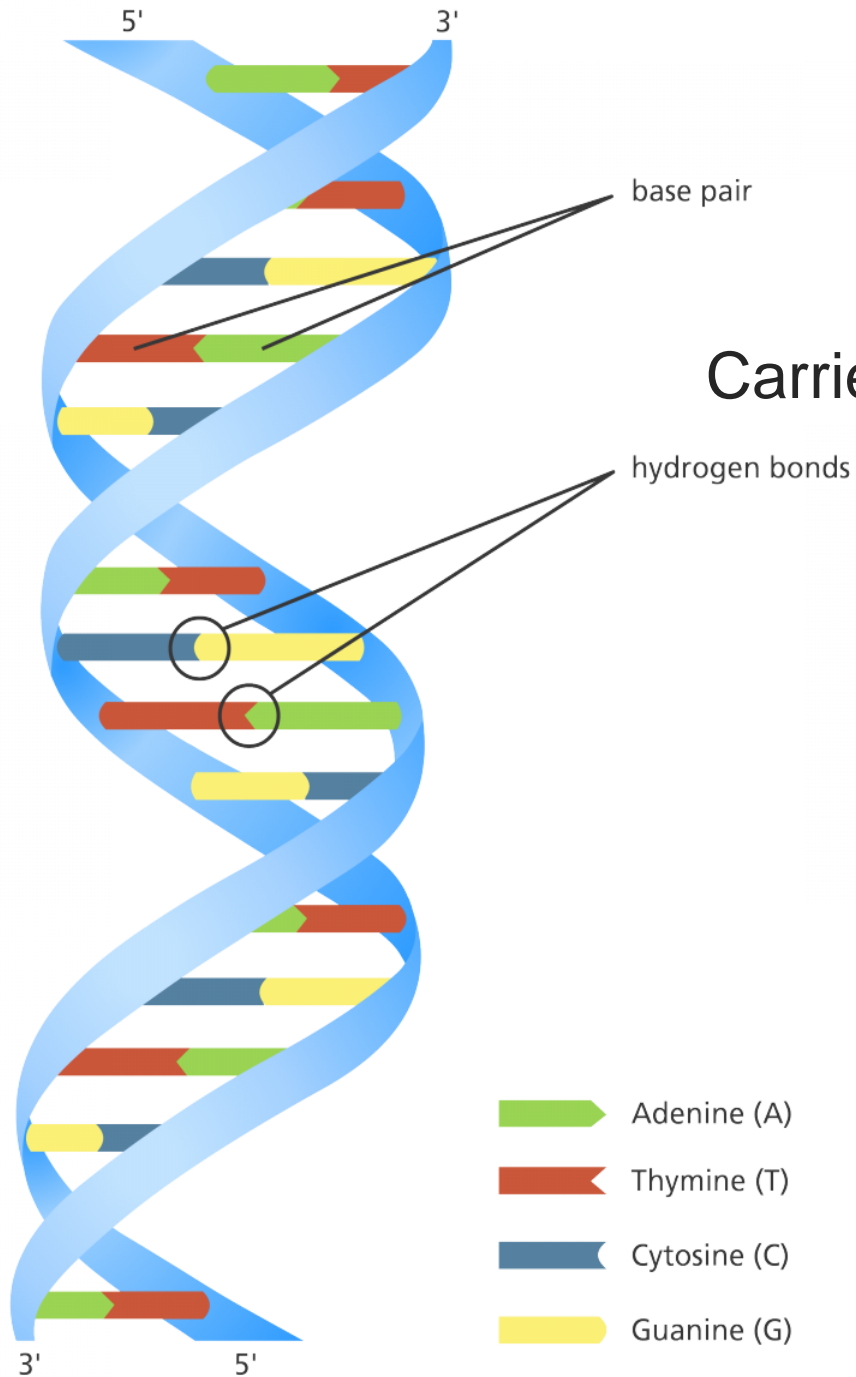
What is the genome?

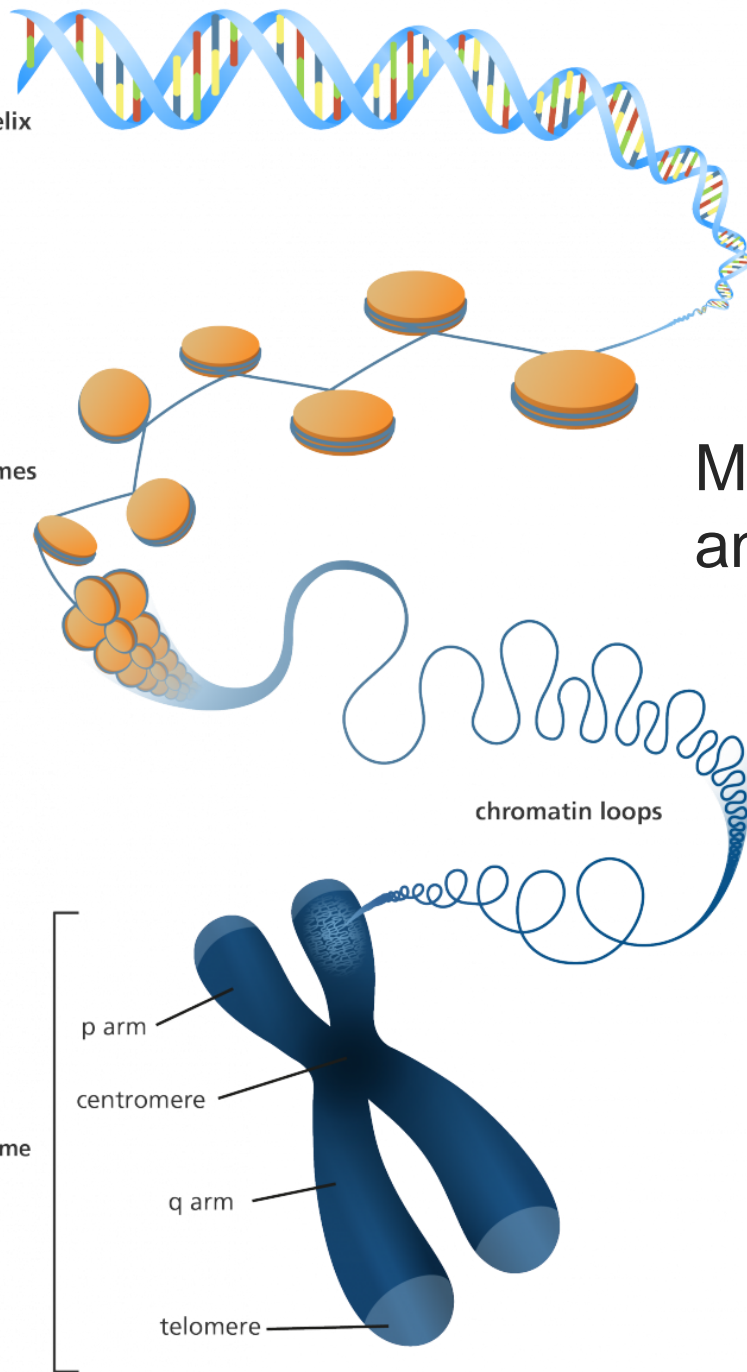DNA double helix

nucleosomes

chromatin loops

Chromosome
- p arm
- centromere
- q arm
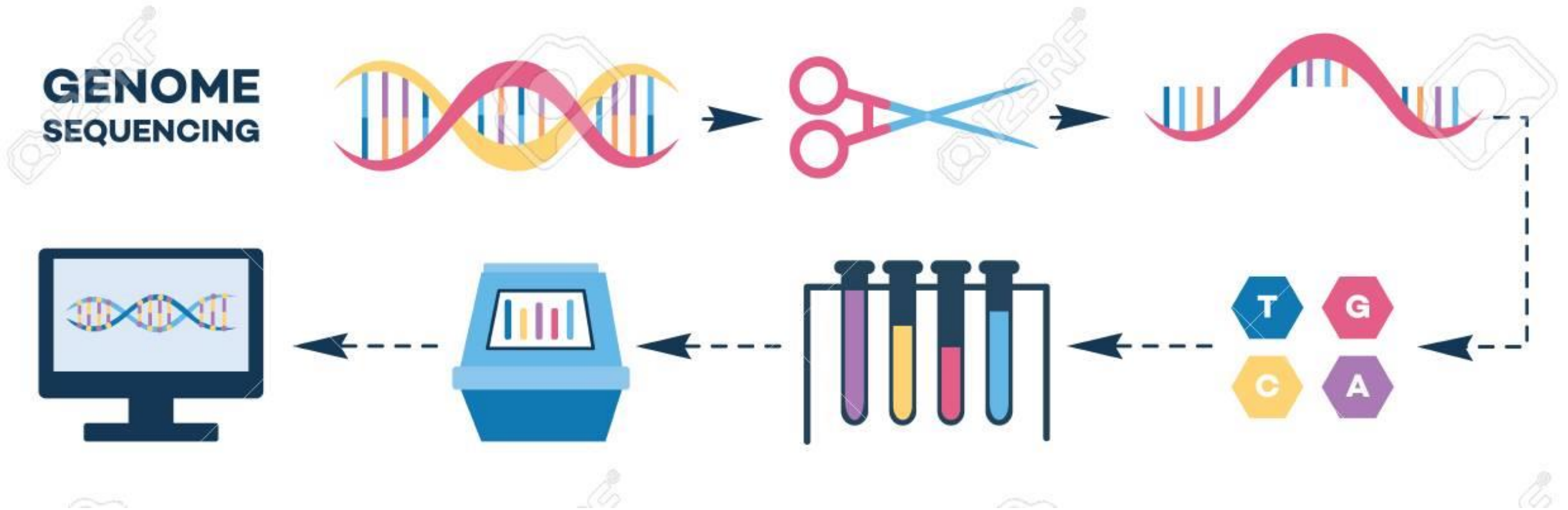- telomere

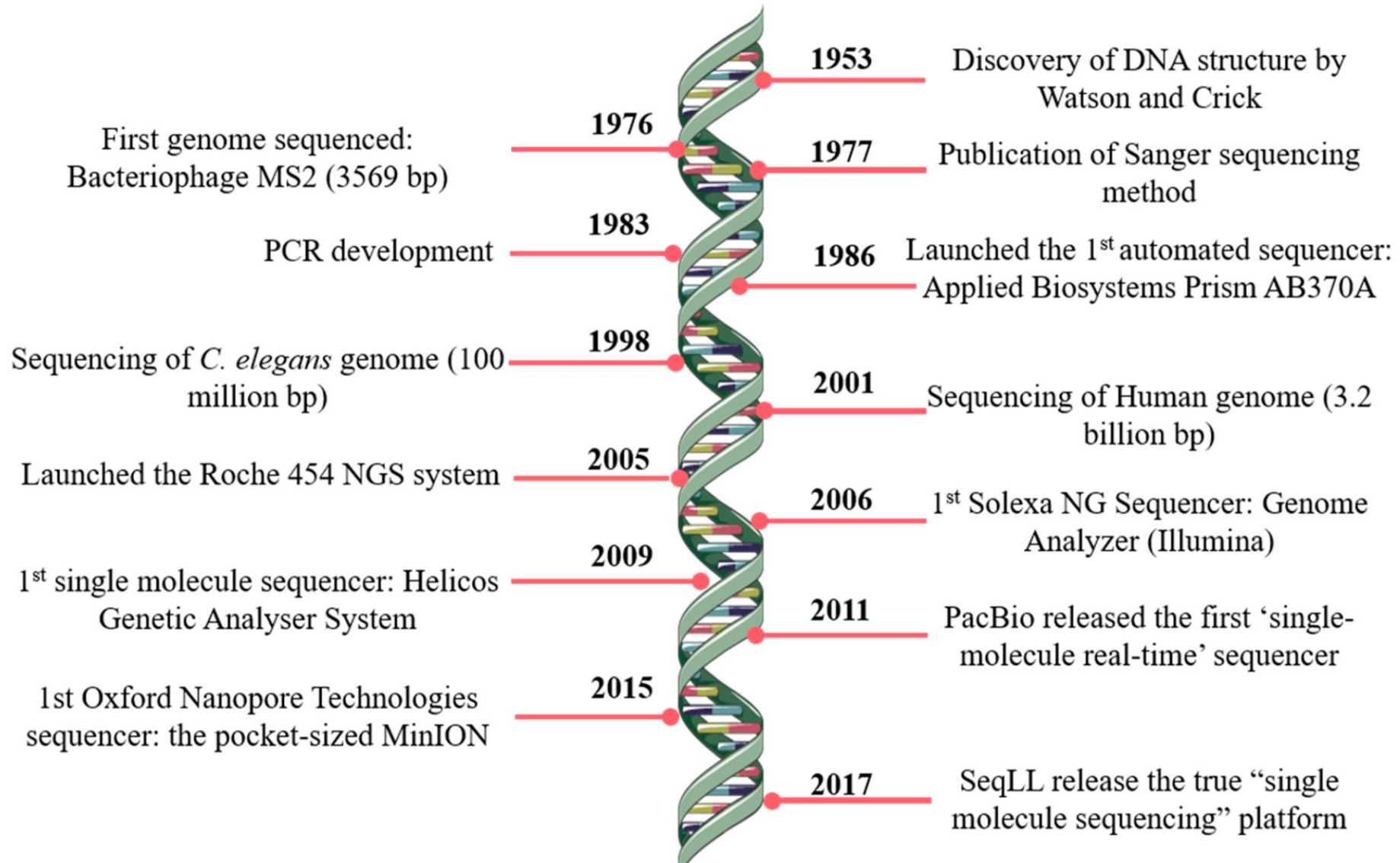Made up of chromosomes, complete set of genes in an organism

What is genetic sequencing?

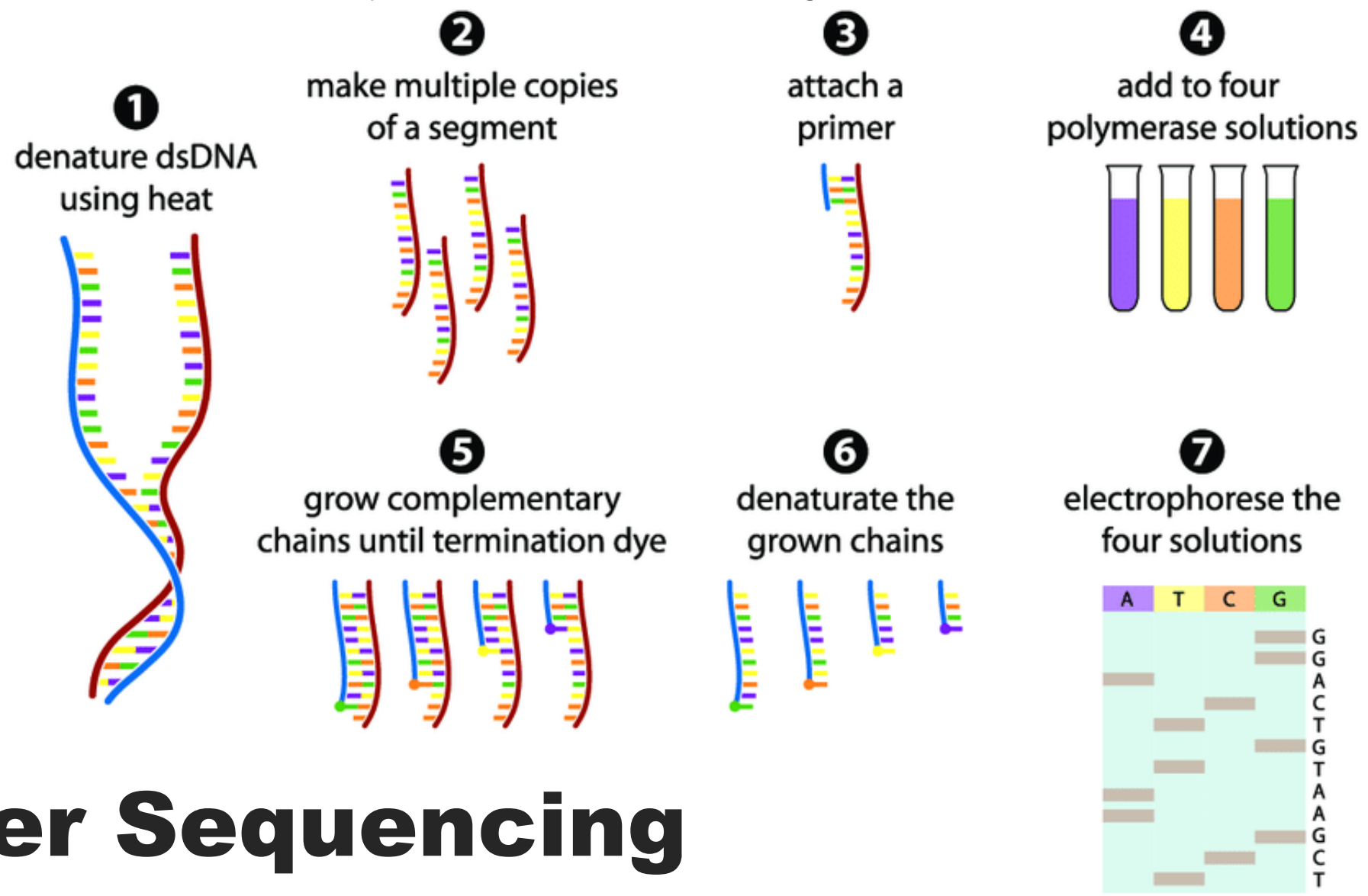Used to determine the order of nucleotides that make up DNA

Why is DNA sequencing important?

1953 — Discovery of DNA structure by Watson and Crick

1976 — First genome sequenced: Bacteriophage MS2 (3569 bp)

1977 — Publication of Sanger sequencing method

1983 — PCR development

1986 — Launched the 1st automated sequencer: Applied Biosystems Prism AB370A

1998 — Sequencing of *C. elegans* genome (100 million bp)

2001 — Sequencing of Human genome (3.2 billion bp)

2005 — Launched the Roche 454 NGS system

2006 — 1st Solexa NG Sequencer: Genome Analyzer (Illumina)

2009 — 1st single molecule sequencer: Helicos Genetic Analyser System

2011 — PacBio released the first 'single-molecule real-time' sequencer

2015 — 1st Oxford Nanopore Technologies sequencer: the pocket-sized MinION

2017 — SeqLL release the true "single molecule sequencing" platform

Can be used to compare genomes of individuals with disease and facilitate treatment
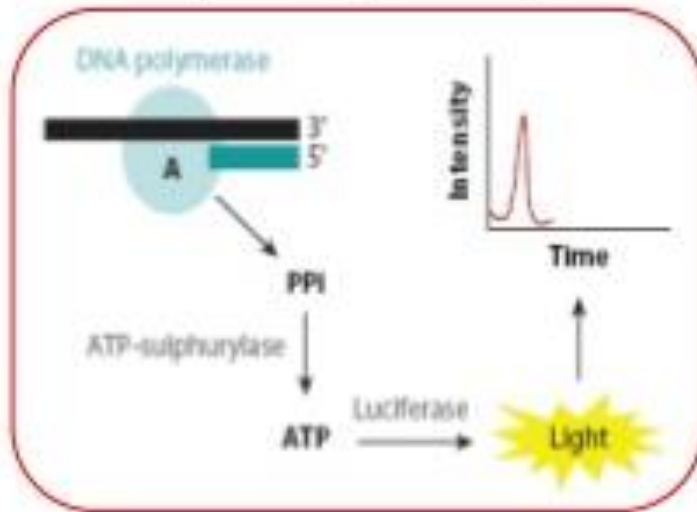
What was one of the most earliest sequencing methods?
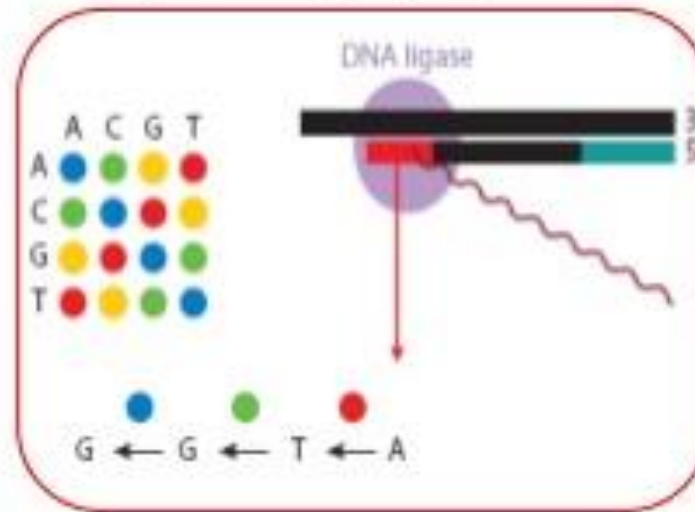
13 years to sequence human genome!

❶ denature dsDNA using heat

❷ make multiple copies of a segment

❸ attach a primer

❹ add to four polymerase solutions

❺ grow complementary chains until termination dye

❻ denaturate the grown chains

❼ electrophorese the four solutions

A	T	C	G

G
G
A
C
T
G
T
A
A
G
C
T

Sanger Sequencing

What is Next Gen Sequencing?

Pyrosequencing — 454 sequencing
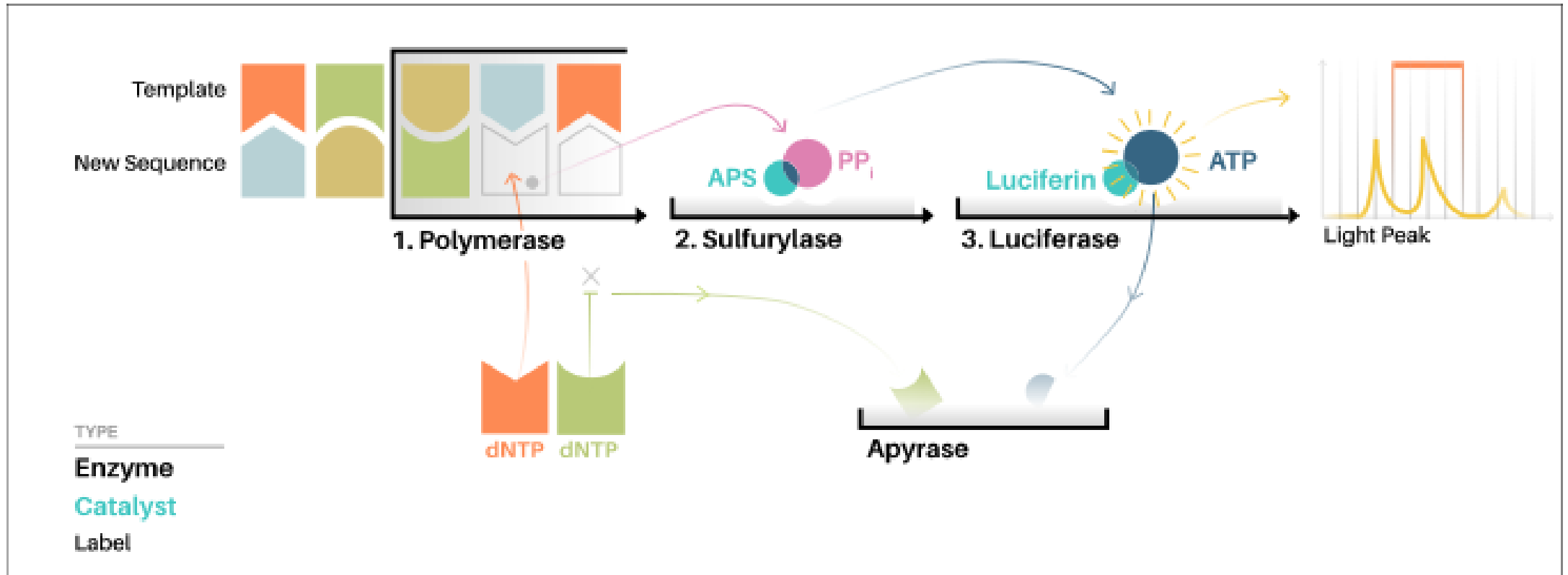
Sequencing-by-ligation — SOLiD platform
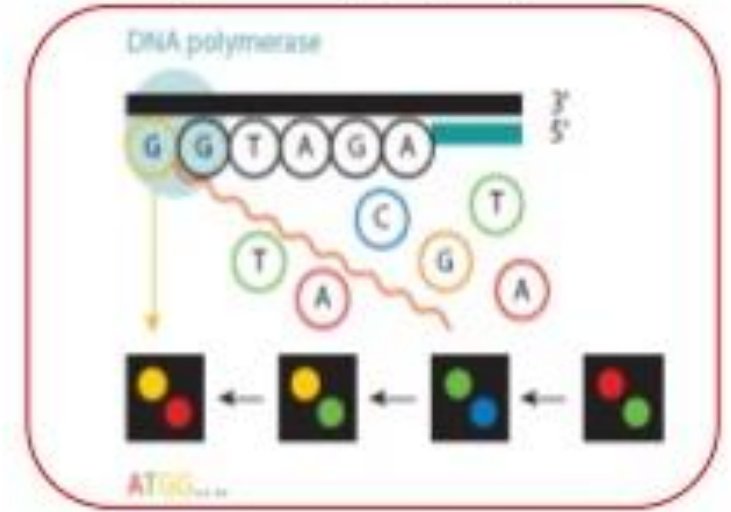
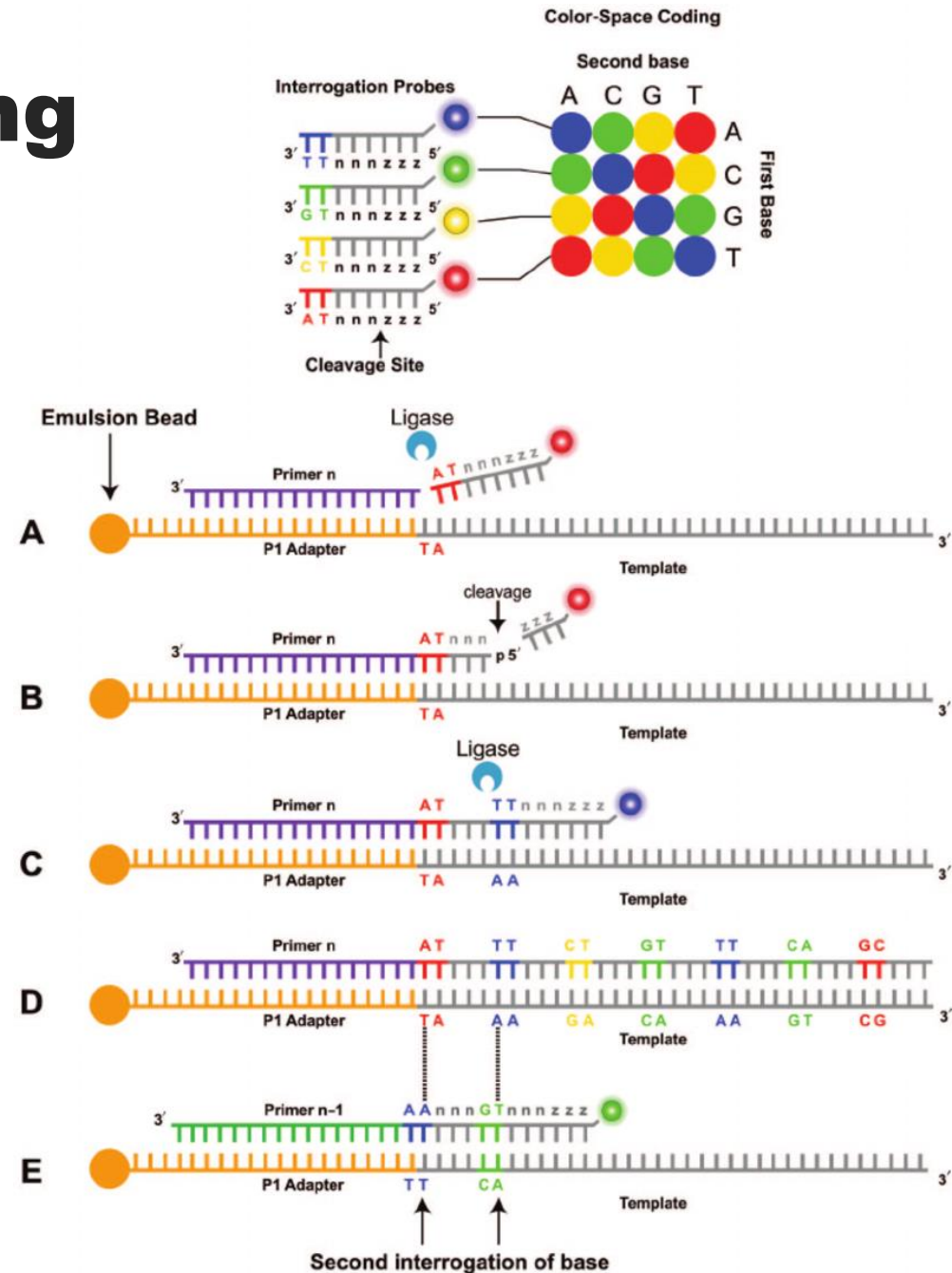Sequencing-by-synthesis — Solexa technology

Second generation sequencing, faster, cheaper

How does Pyrosequencing (454) work?

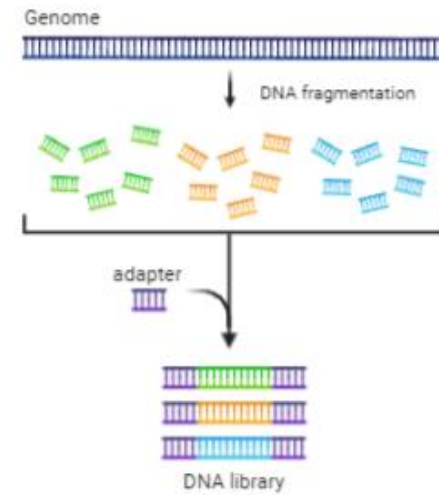Release of pyrophosphate is used in Luciferase reaction to form light
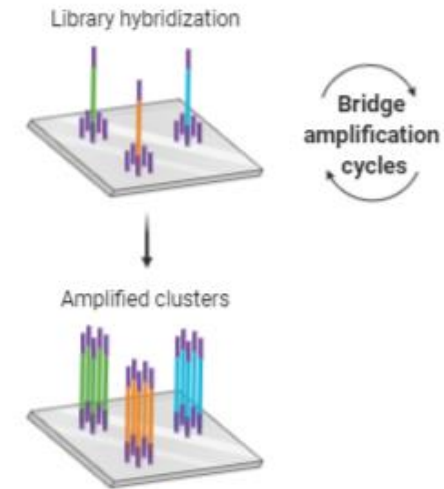
How does sequencing by ligation (SOLiD) work?

DNA is amplified via PCR and attached to emulsion beads, beads then attached to glass slides

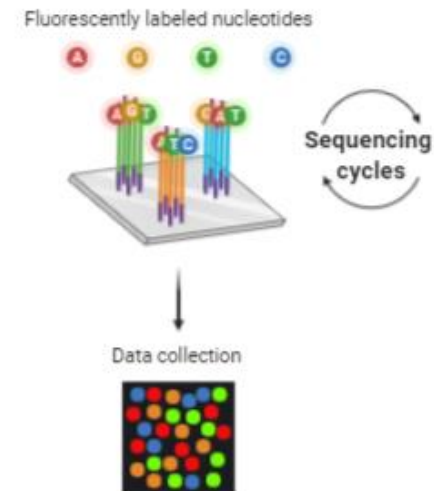# How does sequencing by synthesis (Solexa/Illumina) work?

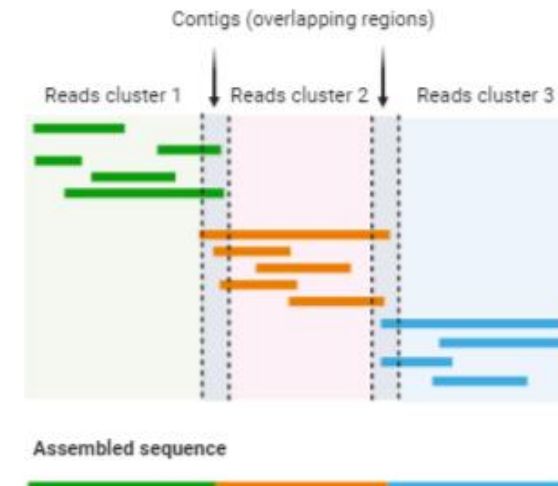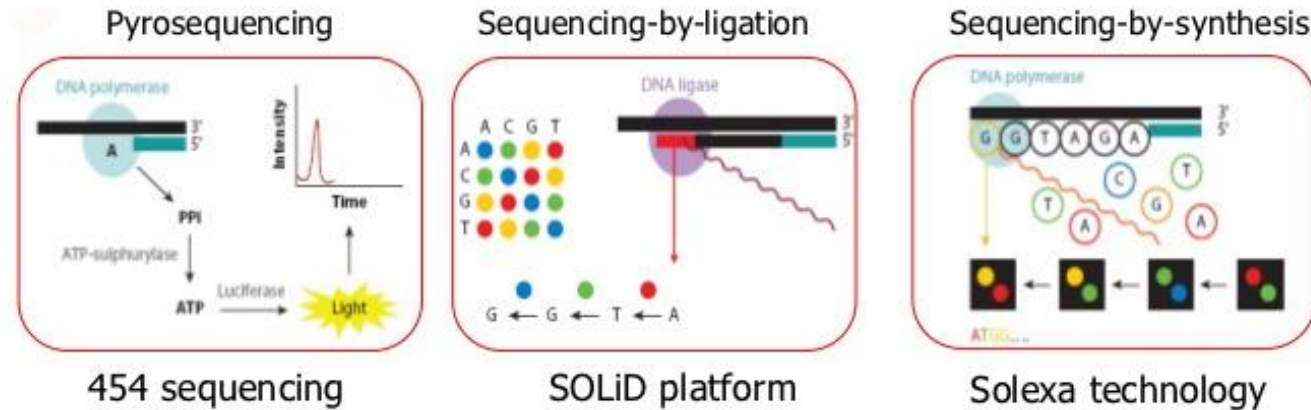DNA clusters are amplified on chip and free floating dNTPs bind and give off light

Advantages and Disadvantages of NGS

Pyrosequencing	Sequencing-by-ligation	Sequencing-by-synthesis
454 sequencing	SOLiD platform	Solexa technology

454

1 million reads/run
$10,000 per run

24 Hours

Long read sizes, fast

Homopolymer errors

SOLiD

1.2-1.4 billion reads/run
$60 - $130

1-2 weeks

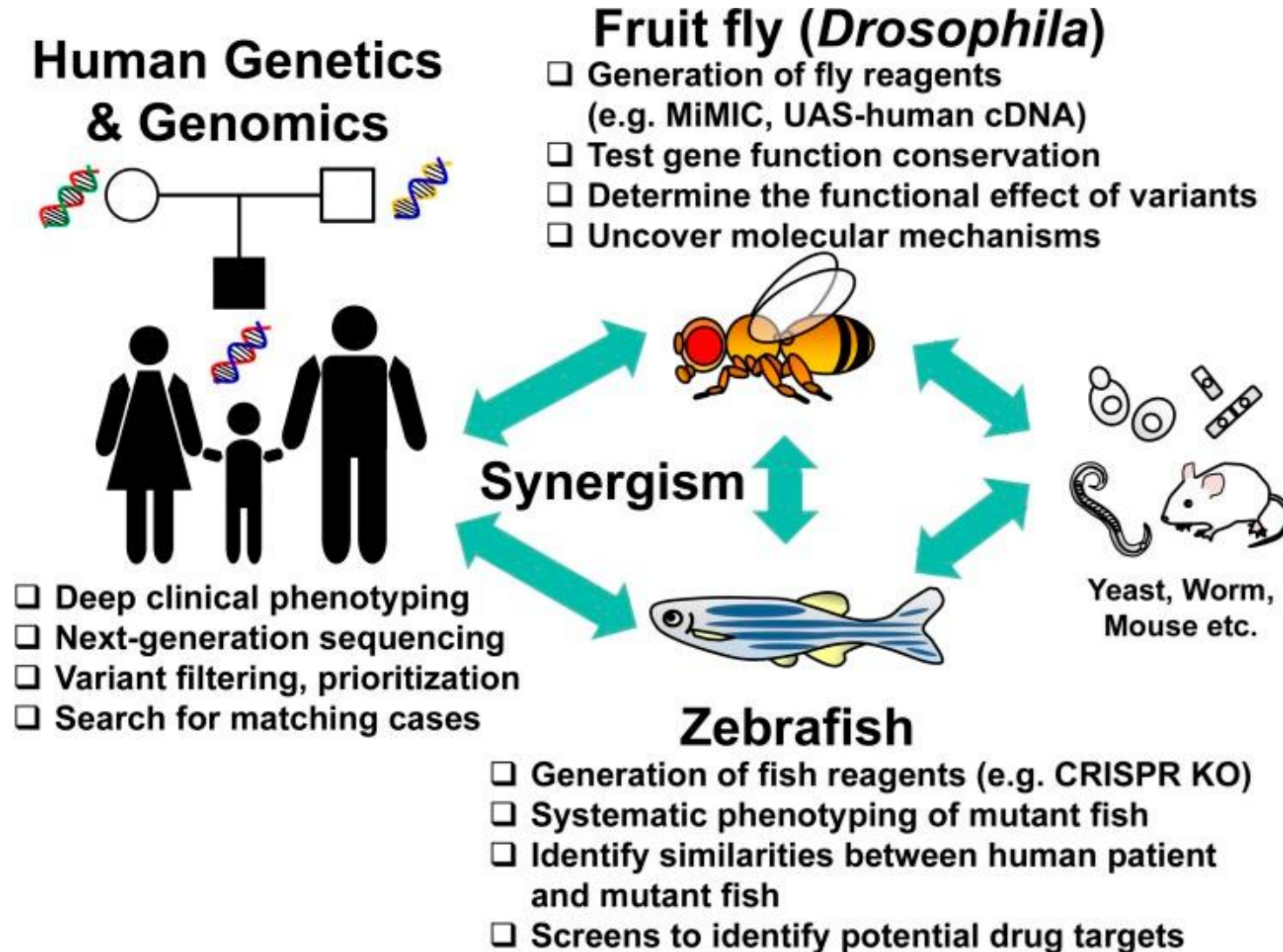Low cost per base

Slower than other methods

Solexa/Illumina

1 million – 3 billion reads/run
$5-$150

1-11 days

High sequence yield

High equipment cost

How can we use model organisms and NGS to facilitate disease diagnosis?

Human Genetics & Genomics

- ❑ Deep clinical phenotyping
- ❑ Next-generation sequencing
- ❑ Variant filtering, prioritization
- ❑ Search for matching cases

Fruit fly (*Drosophila*)

- ❑ Generation of fly reagents (e.g. MiMIC, UAS-human cDNA)
- ❑ Test gene function conservation
- ❑ Determine the functional effect of variants
- ❑ Uncover molecular mechanisms

Synergism

Yeast, Worm, Mouse etc.

Zebrafish

- ❑ Generation of fish reagents (e.g. CRISPR KO)
- ❑ Systematic phenotyping of mutant fish
- ❑ Identify similarities between human patient and mutant fish
- ❑ Screens to identify potential drug targets

Major concern in medicine is finding the gene that causes human disease

What is the best model organism for your disease?

Price

Easy Manipulation

Short life span
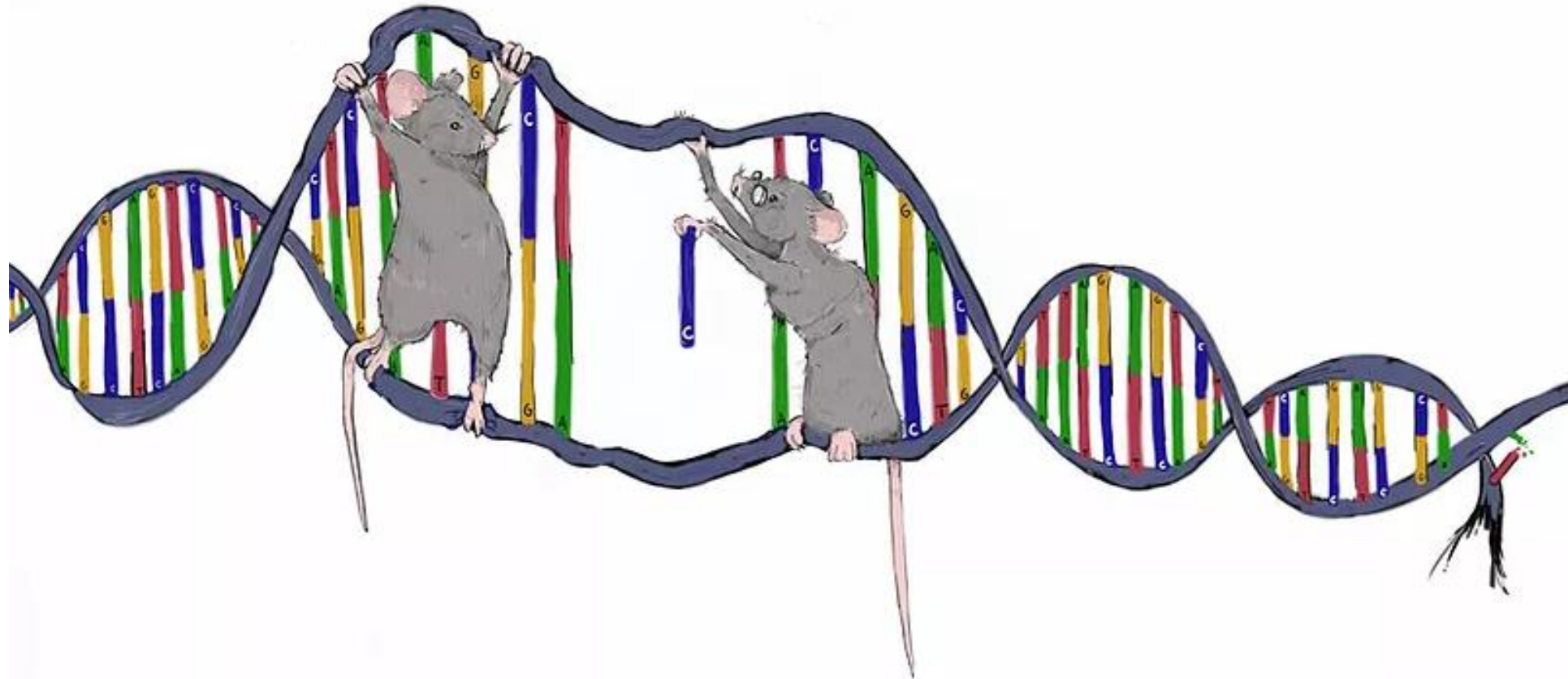
Large # of offspring

Well understood genomes

Illustration by Lindsay M. Davison. Lab Member

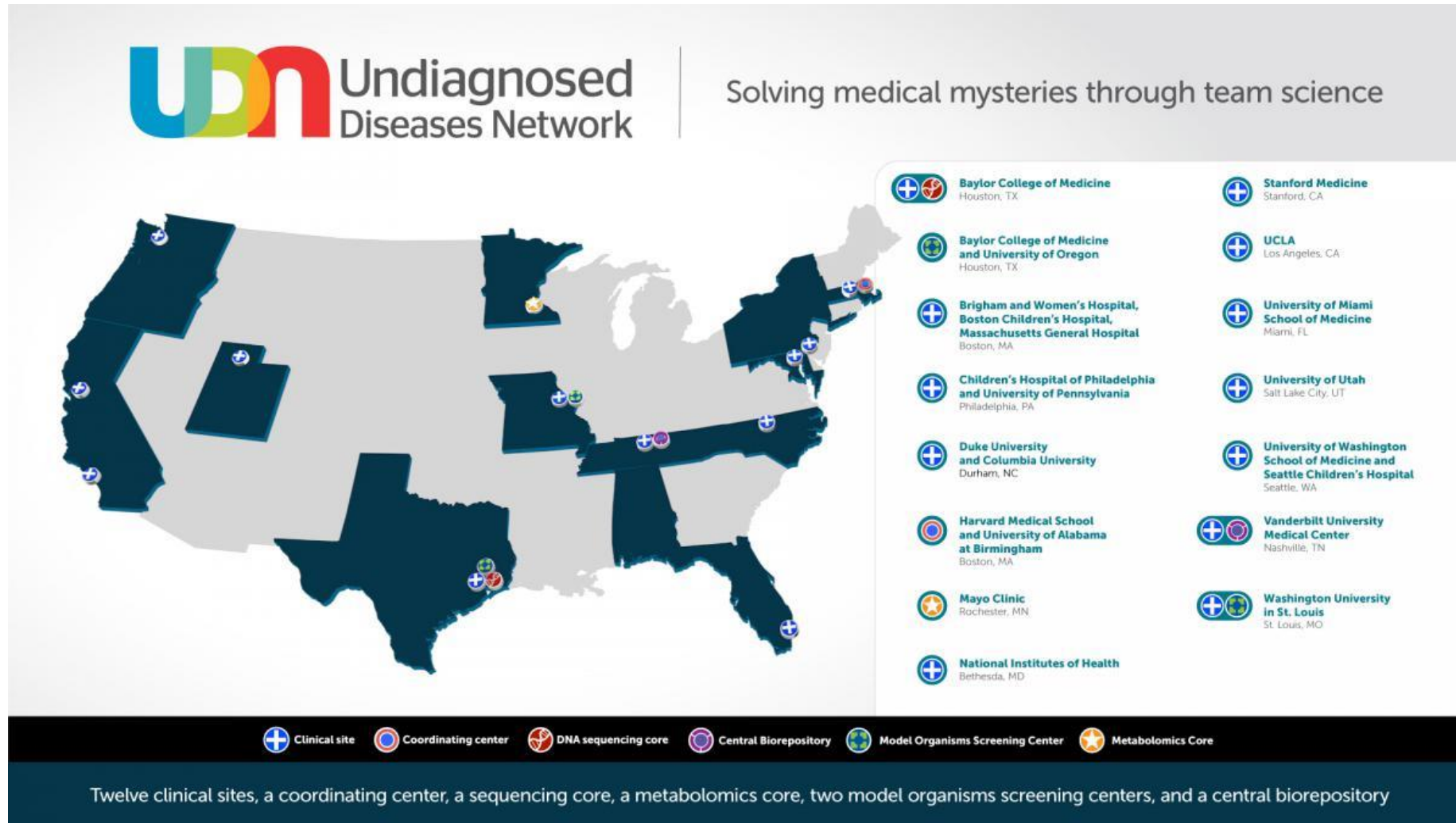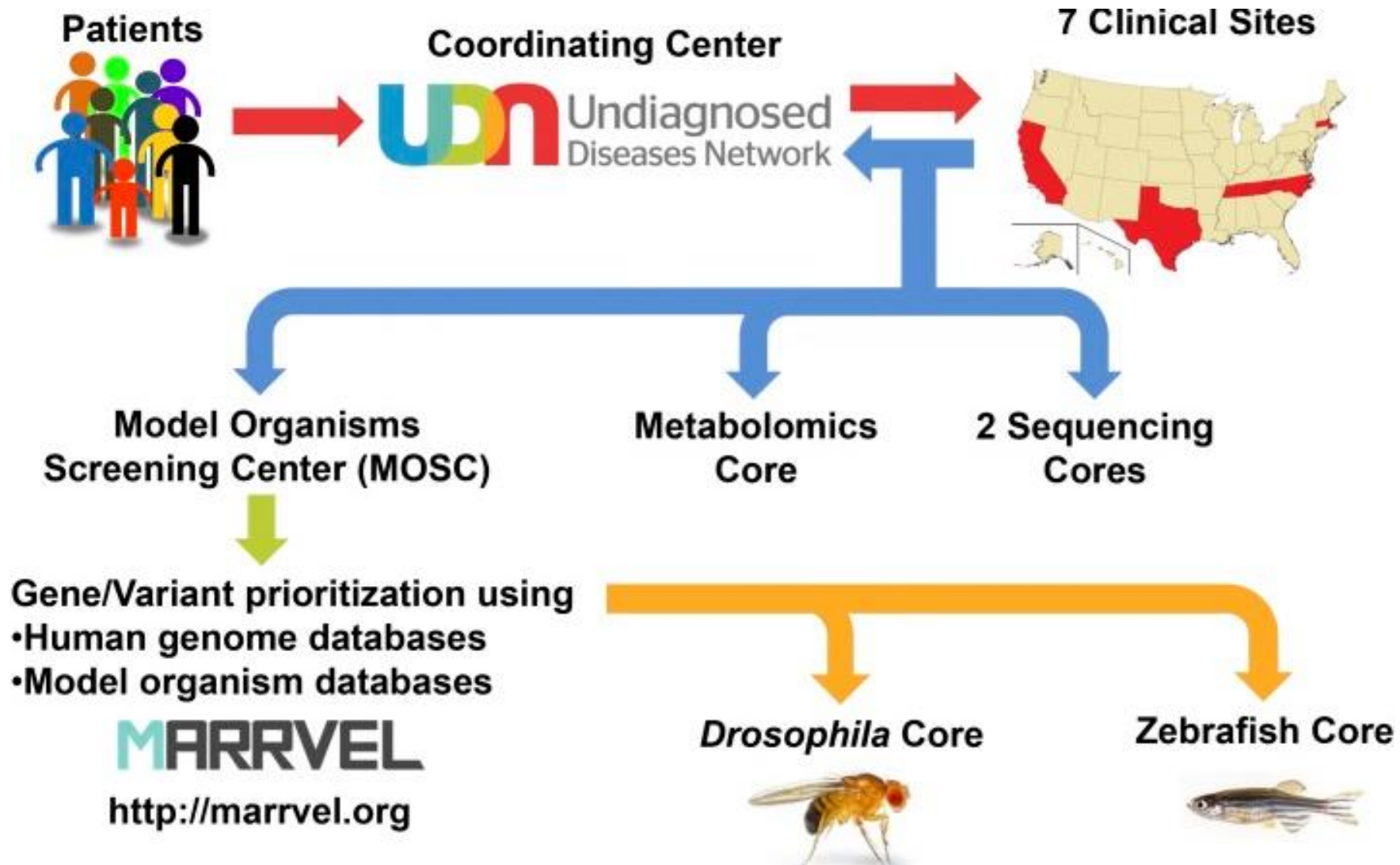# What is an undiagnosed disease?

Millers Syndrome

Kabuki Syndrome

What is UDN?

"Is a research study backed by the National Institutes of Health Common Fund that seeks to provide answers for patients and families affected by these mysterious conditions"

How does UDN work?

Summary

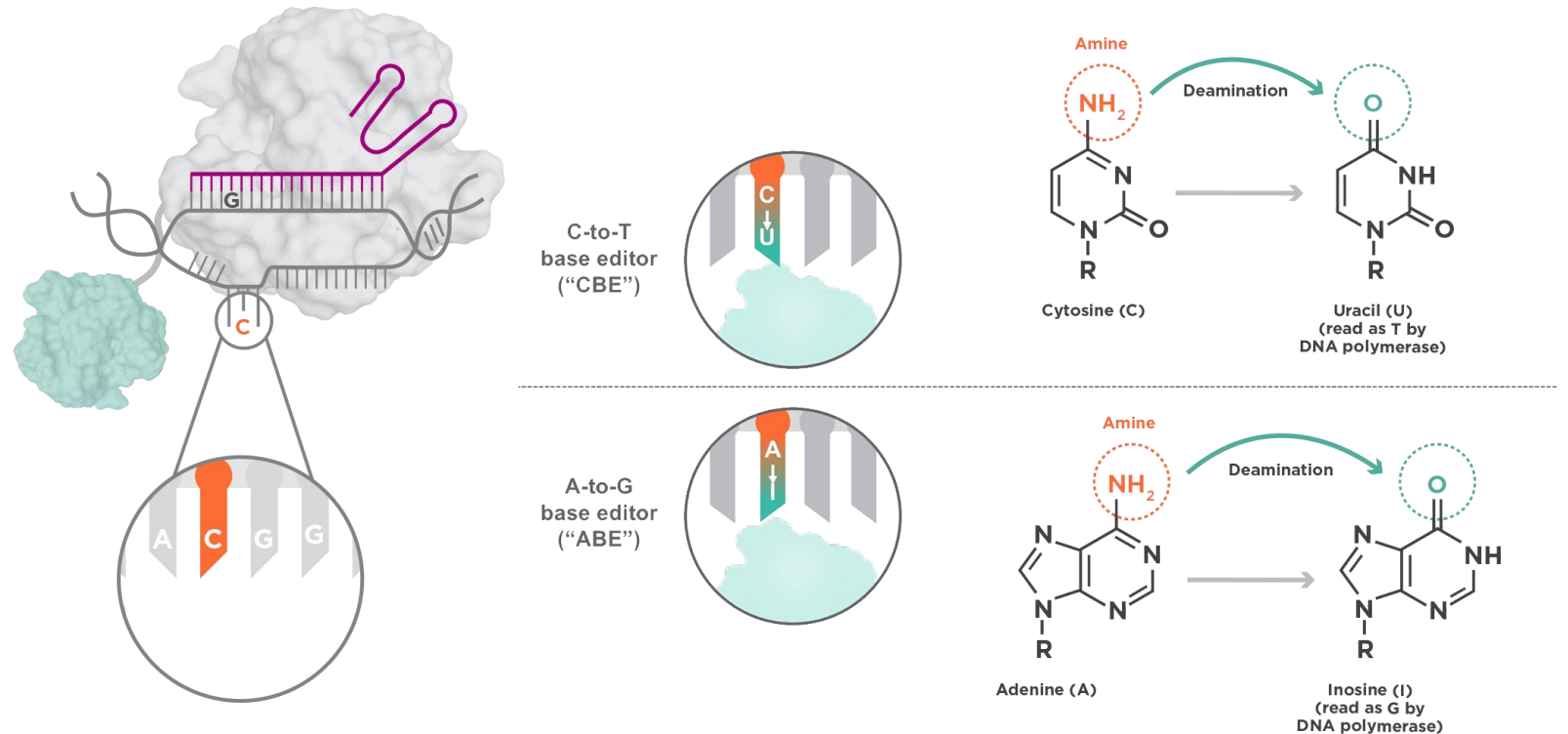Next Gen Sequencing can be used to identify whole genomes

Genomes can be studied for mutations and cross checked with model organisms' genomes

Undiagnosed diseases are being treated with the help of model organisms

Who are the scientists behind this work?

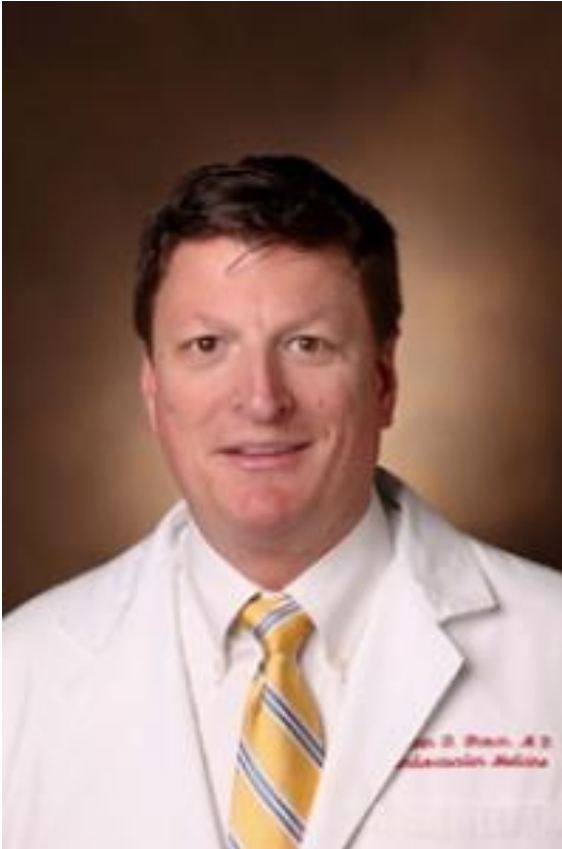**David R. Liu, Ph.D.**
Harvard University

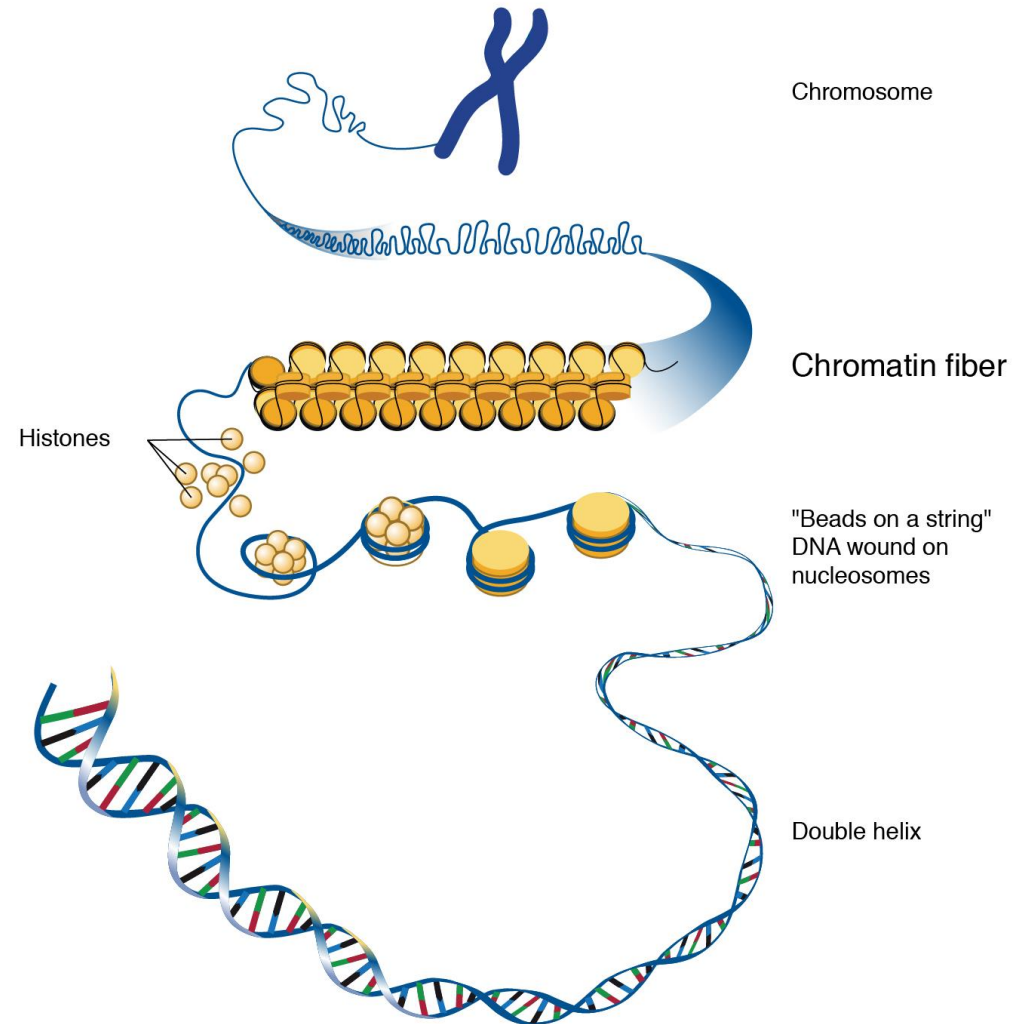ABEs and CBEs were invented in this lab, also developed prime editing

Who are the scientists behind this work?

Liu Lab Gallery

Who are the scientists behind this work?

Chromosome

Chromatin fiber

Histones

"Beads on a string" DNA wound on nucleosomes

Double helix

Jonathan D. Brown, MD
Vanderbilt University Medical Center

How chromatin-dependent signaling drives transcriptional programs that change cell state, due to stress or inflammation

References

- Rabbani, Bahareh, et al. "Next-Generation Sequencing: Impact of Exome Sequencing in Characterizing Mendelian Disorders." *Journal of Human Genetics*, vol. 57, no. 10, Oct. 2012, pp. 621–32. *DOI.org (Crossref)*, doi:10.1038/jhg.2012.91.

- Wangler, Michael F., et al. "Model Organisms Facilitate Rare Disease Diagnosis and Therapeutic Research." *Genetics*, vol. 207, no. 1, Sept. 2017, pp. 9–27. *DOI.org (Crossref)*, doi:10.1534/genetics.117.203067.

- Ng, Sarah & Buckingham, Kati & Lee, Choli & Bigham, Abigail & Tabor, Holly & Dent, Karin & Huff, Chad & Shannon, Paul & Jabs, Ethylin & Nickerson, Deborah & Shendure, Jay & Bamshad, Michael. (2009). Exome sequencing identifies the cause of a mendelian disorder. Nature genetics. 42. 30-5. 10.1038/ng.499.

- Phenotypic spectrum and management issues in Kabuki syndrome Kawame, Hiroshi et al. The Journal of Pediatrics, Volume 134, Issue 4, 480 - 485