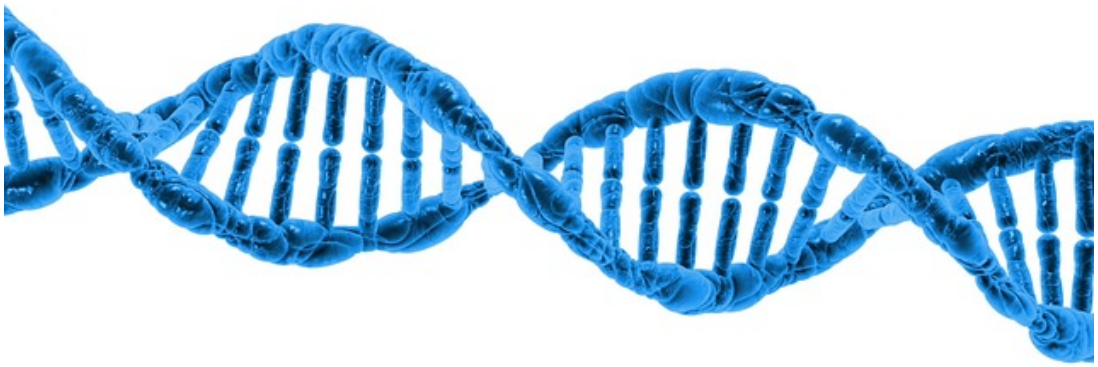
Genomics: Mapping and Sequencing Genomes

See online here

Genomics is the study of what the genome does. In this article, we include the topics of distinguishing between genetic and physical maps, describing the process of DNA sequencing and characterizing two different methods for sequencing genomes— the clone-by-clone vs. shotgun sequencing.



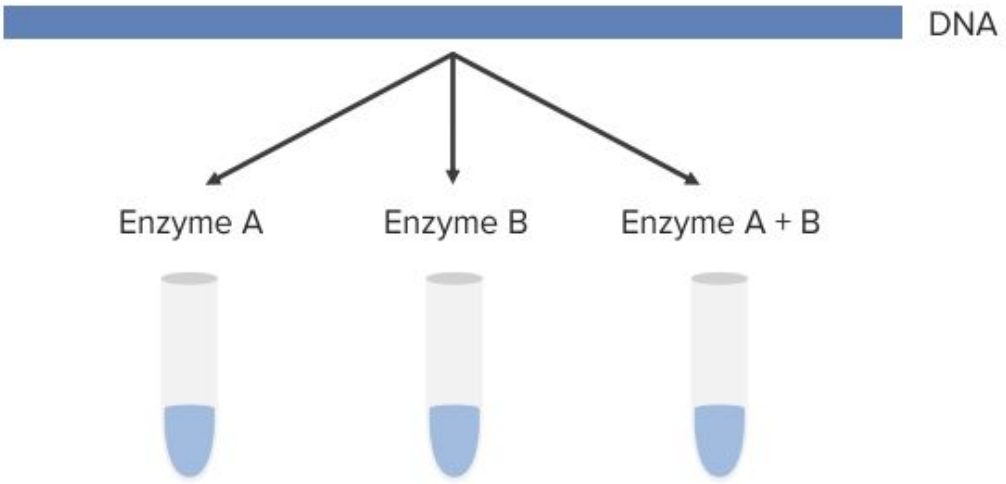
Restriction Mapping

Physical mapping allows us to find the **actual physical location of each gene** on the chromosome—all the way down to such a granular level we know precisely at what letter a gene begins.

- Physical mapping uses **landmarks** within DNA.
- Genetic maps provide the relative location of the genes, based on **recombination frequency**.
- Physical maps provide the **actual physical locations** of each gene.

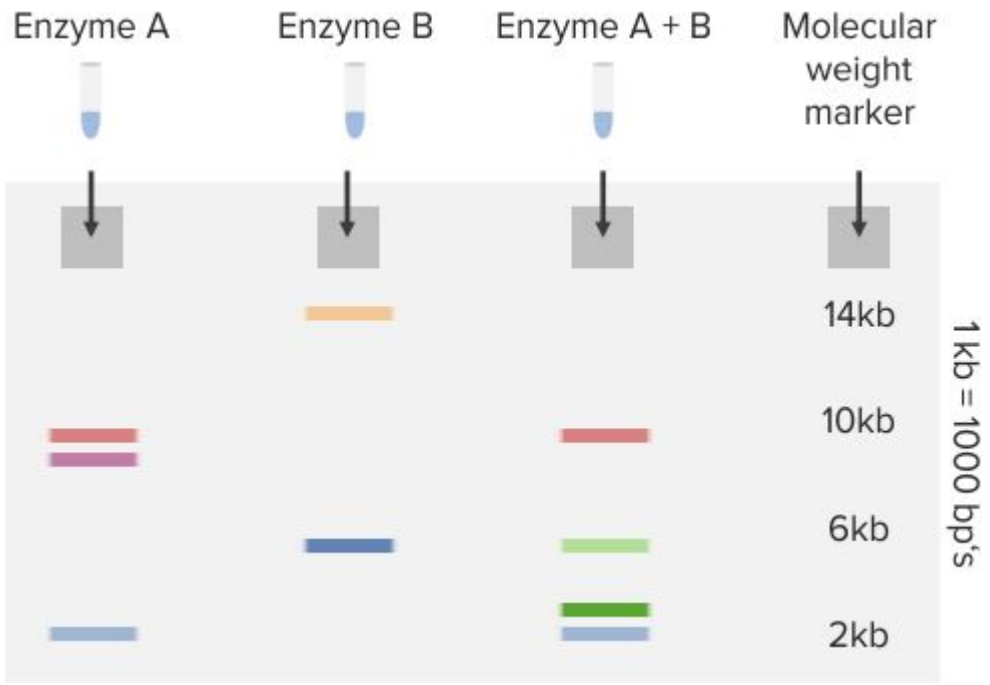
Restriction mapping provides physical maps of DNA fragments. The process is performed in the following steps:

1. **Multiple copies** of a segment of DNA are cut with **restriction enzymes**. A variety of different restriction enzymes are available to use for this purpose.



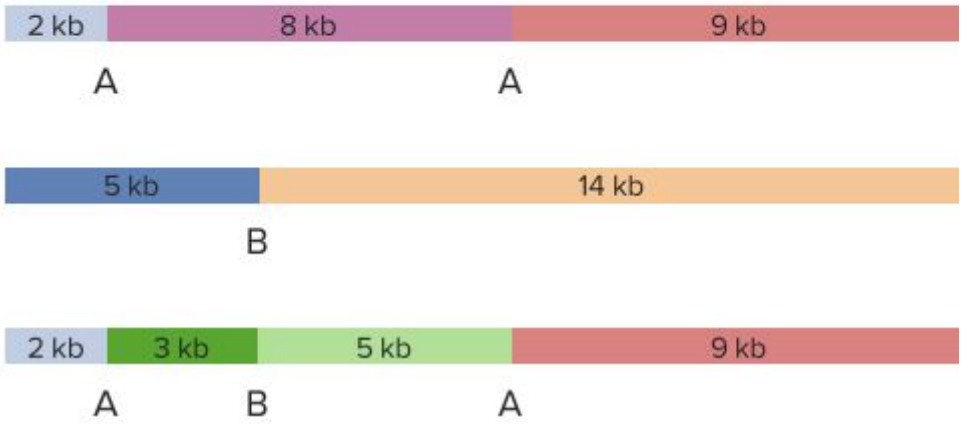
"Restriction mapping provides physical maps of DNA fragments" Image created by Lecturio

2. The fragments produced by enzyme A only, enzyme B only and by enzyme A and B together, are run side-by-side on a gel. The negatively charged DNA runs toward the positive pole and, since the larger fragments will move less distance through the gel, this procedural step **separates the fragments by size**.



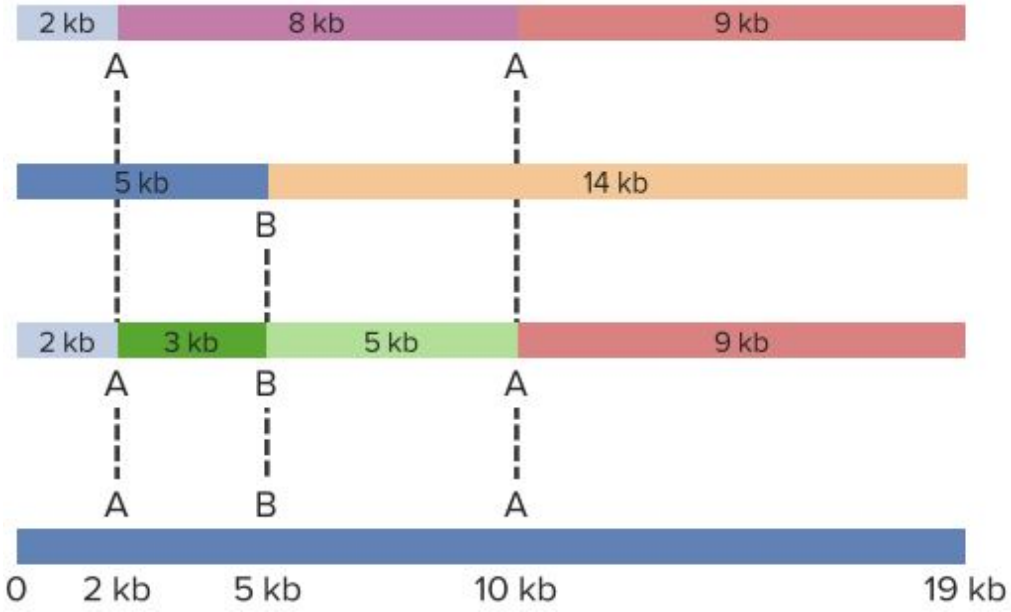
"Restriction Mapping" Image created by Lecturio

3. Thus, the fragments are arranged so that the smaller ones (produced by the simultaneous cut of enzymes A and B) can be grouped to generate the larger ones (produced by the individual enzymes). These pieces are taken and compared according to their **sizes and lengths**.



“Restriction-Mapping-Step-3” Image created by Lecturio

4. This is how a physical map is constructed.



“Restriction-Mapping-Step-4” Image created by Lecturio

Cytological Maps

In order to create genetic maps, it is necessary to use **labeling and tags**, so the result might resemble these **cytological maps**. The chromosome has been broken down into **sections,** which show the physical locations of the genes. Cytological maps use **staining** in order to mark places on the genome, thus allowing for a whole view of each chromosome and therefore, the entire genome.

FISH – Fluorescence In-Situ Hybridization

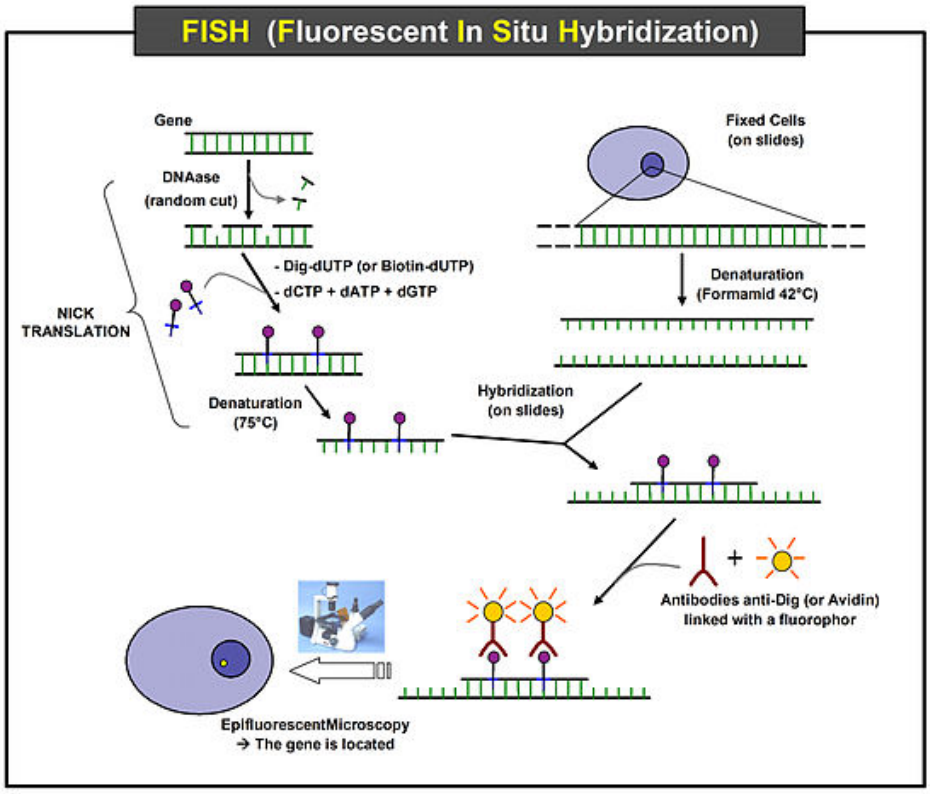


Image: “Scheme of the principle of the FISH (Fluorescent in-situ hybridization) experiment to localize a gene in the nucleus.” by Mr. Matze. License: CC BY-SA 3.0

FISH is one of the staining techniques used to mark chromosomes with **fluorescent dyes** and see if they contain particular genes. The primary purpose of the cytological maps is to characterize chromosomal abnormalities.

Sequence Tagged Sites (STSs)

STSs can provide a sort of **scaffolding that shows how the pieces in the genome go together.** This helps **to investigate locations** of known DNA sequences on a chromosome. STSs comprise of **200 to 500 base pair** sequences that have a **single occurrence**. DNA fragments from DNA libraries are cut with **restriction enzymes** and run on the **gel,** with **electrophoresis** separating the resulting pieces by **size**. Each clone, provides different pieces of DNA, which **can be aligned because of the STSs.**

In order to **identify these STSs, polymerase chain reaction (PCR)** is used with probes; the probes will attach when the DNA separates, then they can be located by using visualization techniques.

The Ultimate Physical Map: The Sequencing of the Entire Genome

The ultimate physical map represents the **exact DNA sequence on a chromosome**. We can say exactly where a gene is located on a chromosome by using DNA sequencing. **Vectors** containing cloned DNA from libraries can be used to sequence a

genome.

Sanger Sequencing: The Enzymatic Method

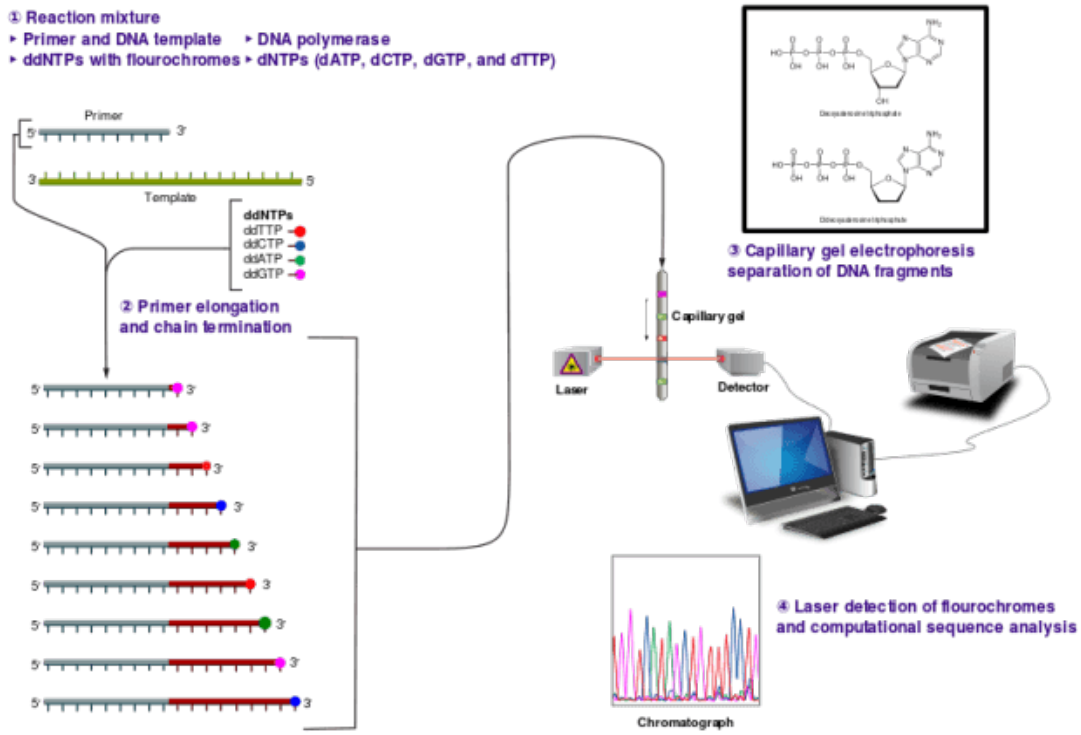


Image: “The Sanger (chain-termination) method for DNA sequencing. (1) A primer is annealed to a sequence, (2) Reagents are added to the primer and template, including DNA polymerase, dNTPs and a small amount of all four dideoxynucleotides (ddNTPs) labeled with fluorophores. During primer elongation, the random insertion of a ddNTP instead of a dNTP terminates synthesis of the chain because DNA polymerase cannot react with the missing hydroxyl. This produces all possible lengths of chains. (3) The products are separated on a single lane capillary gel, where the resulting bands are read by an imaging system. (4) This produces several hundred thousand nucleotides a day, data which requires storage and subsequent computational analysis” by Estevezj. License: CC BY-SA 3.0

Dideoxynucleotides are a critical component. **3′ OH** is needed for **DNA polymerase** to add new nucleotides.

- DNA replication in vitro.
- Termination of replication occurs every time a dideoxynucleotide shows up.

Genome Sequencing: The Development of Artificial Chromosomes

Methods that can be used in genome sequencing can be broken into two categories: the clone-by-clone sequencing method and the shotgun sequencing method.

Clone-by-clone sequencing: physical mapping

Also known as **BAC to BAC** (because we are putting it from bacterial artificial chromosome to bacterial artificial chromosome), or **hierarchical** sequencing. It requires a little **less lining up** of fragments.

1. As a first step, large DNA clones are isolated. These are arranged into **contiguous sequences** based on **overlapping tagged sites**.

2. Large clones are then **fragmented** into smaller clones for sequencing.
3. The entire sequence is assembled from the overlapping larger clones.

Shotgun sequencing: advanced computing

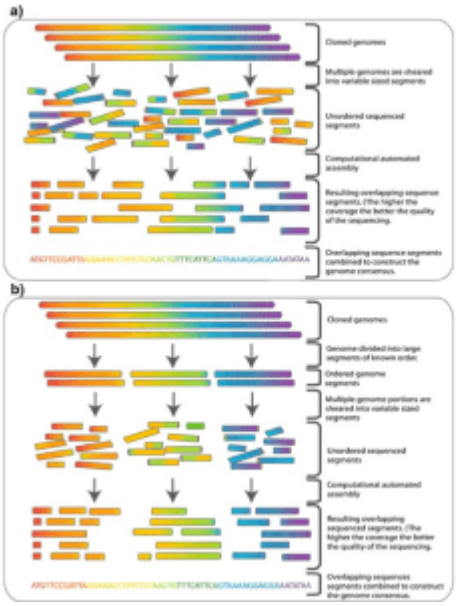


Image: “In whole genome shotgun sequencing (top), the entire genome is sheared randomly into small fragments (appropriately sized for sequencing) and then reassembled. In hierarchical shotgun sequencing (bottom), the genome is first broken into larger segments. After the order of these segments is deduced, they are further sheared into fragments, appropriately sized for sequencing.” by Commins, J., Toft, C., Fares, M. A. License: CC BY-SA 2.5

It’s also known as **whole genome shotgun sequencing,** and is only possible because of the **higher levels of computing** that are possible these days.

Rather than doing BAC to BAC, we just hack up the entire genome into millions of little pieces. Previously, this method consumed a lot of computer resources, but nowadays, with the advantage of powerful

References

Restriction mapping by nature education.

Cytological map. The free dictionary by farlex.

Notes