II. GENE MAPPING

The ultimate goal of mapping is:

1) to identify the gene(s) responsible for a given phenotype or the mutation responsible for a specific variant;

2) to know which chromosome contains the gene(s) and precisely where the gene(s) lies on that chromosome and

3) to understand the genes function.

Gene mapping, describes any method used to identify new genes (determining the sequence), the locus of a gene and provides clues about the relative distances between genes on a chromosome either in linkage units or physical units. Early gene maps used linkage analysis (Genetic mapping = linkage mapping).The closer two genes are to each other on the chromosome, the more likely it is that they will be inherited together, so by following inheritance patterns, the relative positions of genes can be determined. More recently, scientists have used recombinant DNA (rDNA) techniques to establish the actual physical locations of genes on the chromosomes.

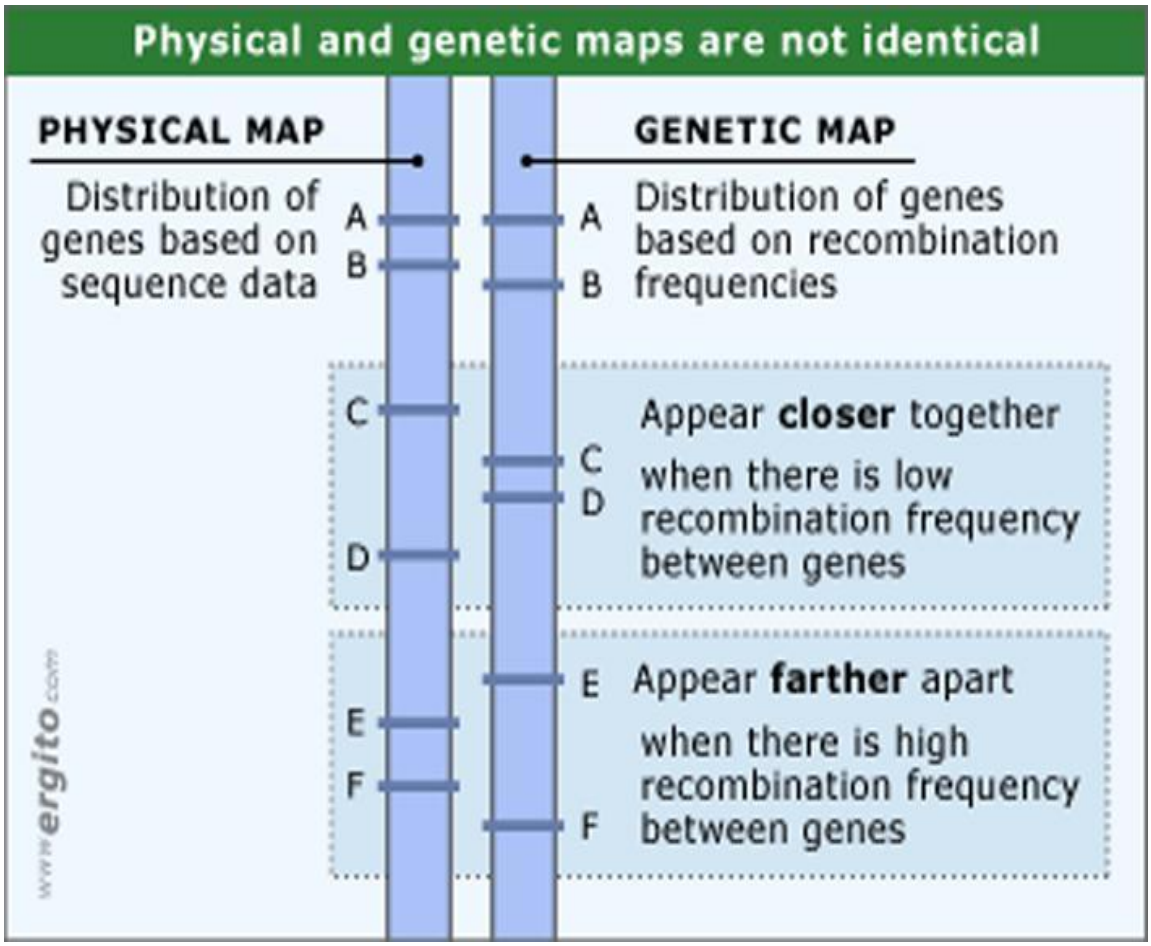
There are two distinctive types of "Maps" used in the field of genome mapping: genetic maps and physical maps.

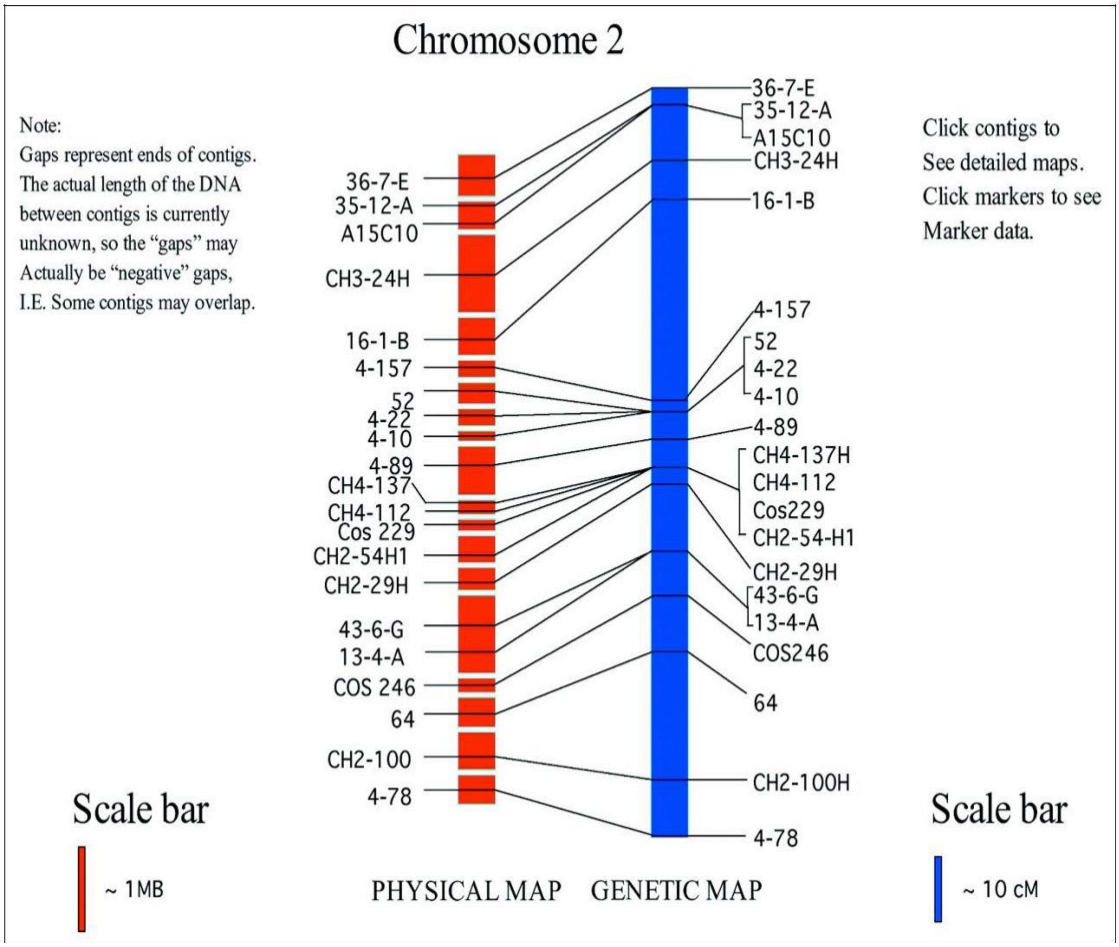
Genetic maps' distances are based on the genetic linkage information measured in centimorgans (CM) i.e. depends on the relative positions of loci based on the degree of recombination. This approach studies the inheritance/assortment of traits by genetic analysis.

Physical maps use actual physical distances between loci usually measured in number of nucleotides or base pairs. While the physical map could be a "more accurate" representation of the genome, genetic maps often offer insights into the nature of different regions of the chromosome. This approach applies techniques of molecular biology.

What is the relationship between the genetic map and the physical map of a chromosome? (figure)

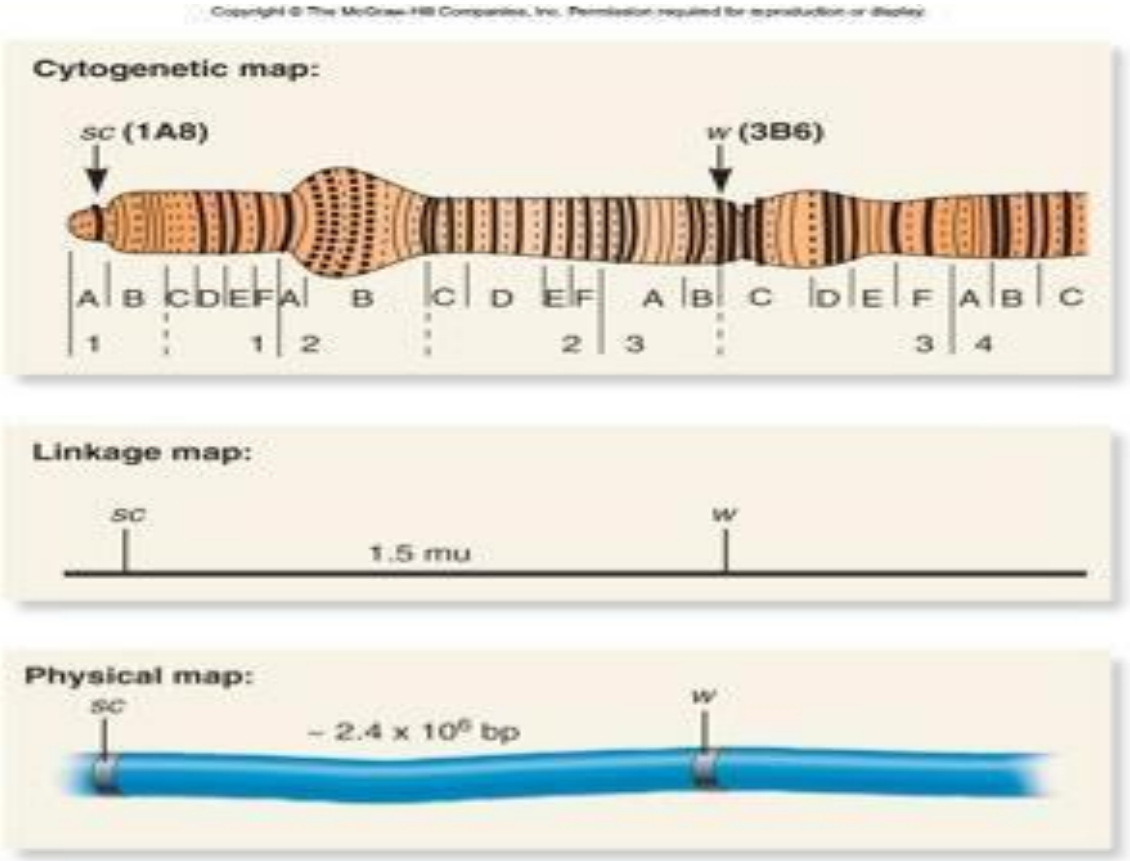
As a very rough rule; 1 cM on a chromosome encompasses 1 megabase (1 Mb = 10^6 bp) of DNA. But for the reasons mentioned above, this relationship is only approximate. Although the genetic maps of human females average 90% longer than the same maps in males, their chromosomes contain the same number of base pairs. So their **physical maps** are identical.





How do geneticists indicate the location of a gene?

Geneticists use **3 maps** to describe the location of a particular gene on a chromosome. One type of map uses the **cytogenetic location** to describe a gene's position. The cytogenetic location is based on a distinctive pattern of bands created when chromosomes are stained with certain chemicals. Another type of map (**linkage or genetic mapping**) is based on recombination frequencies by using the order and relative distances between genes. Unit of distance in linkage map is a map unit = 1cM = 1 map unit is equal to 1% recombination. The third type uses the molecular location (**physical mapping**), a precise description of a gene's position on a chromosome. The molecular location is based on the sequence of DNA building blocks (base pairs) that make up the chromosome.



Cytogenetic location (Cytogenetic mapping)

Geneticists use a standardized way of describing a gene's cytogenetic location. In most cases, the location is based on the distinctive banding patterns of stained chromosomes. The combination of numbers and letters provide a gene's "address" on a chromosome.

This address is made up of several parts:

1. The chromosome on which the gene can be found: from Chromosomes 1 to 22 (the autosomes) are designated by their chromosome number. The sex chromosomes are designated by X or Y.
2. The arm of the chromosome: Each chromosome is divided into two sections (arms) based on the location of the centromere (mentioned in previous lecture). By convention, the shorter arm is called p, and the

longer arm is called q. The chromosome arm is the second part of the gene's address.

Examples:

5q is the long arm of chromosome 5,

and Xp is the short arm of the X chromosome.

3. The position of the gene on the p or q arm: The position of a gene is based on a distinctive pattern of light and dark bands that appear when the chromosome is stained in a certain way. The position is usually designated by two digits (representing a region and a band), which are sometimes followed by a decimal point and one or more additional digits (representing sub-bands within a light or dark area).

   Examples:

   14q21 represents position 21 of the gene on the long arm of chromosome 14.

   17q12-q21 represents a range of bands from 12-21 (if less is known about the exact location of the gene) on the long arm of chromosome 17.

   CFTR **gene** position was **7q31.2** (self-doing) ……?

   N.B. The number indicating the gene position increases with distance from the centromere: 14q21 is closer to the centromere than 14q22.

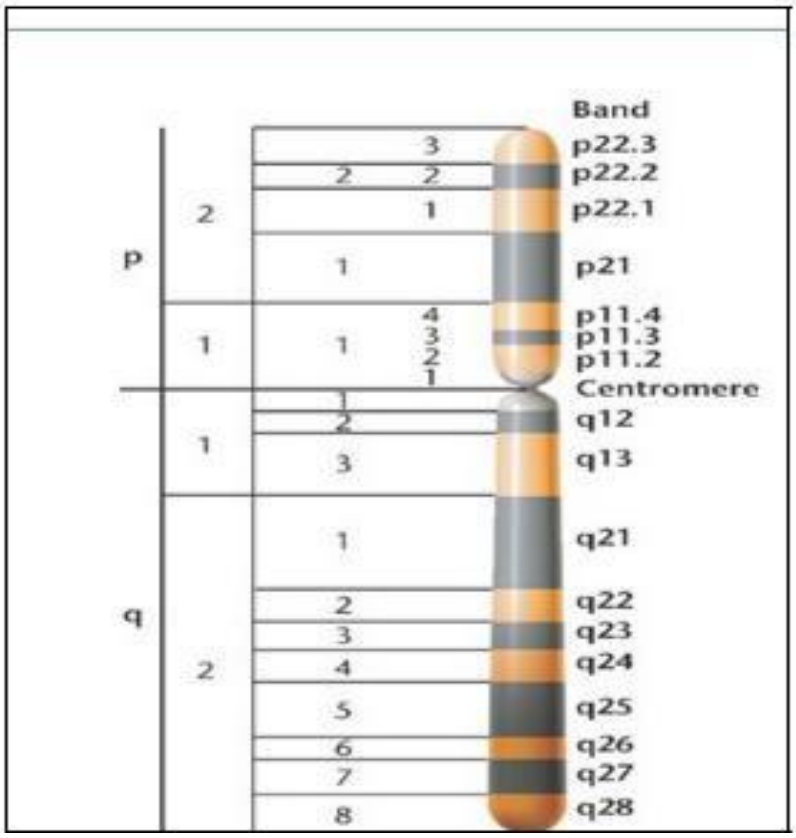
4. The abbreviations "cen" or "ter" or "tel" are sometimes used to describe a gene's cytogenetic location: "Cen" indicates that the gene is very close to the centromere. "Ter" stands for terminus, which indicates that the gene is very close to the end of the p or q arm.

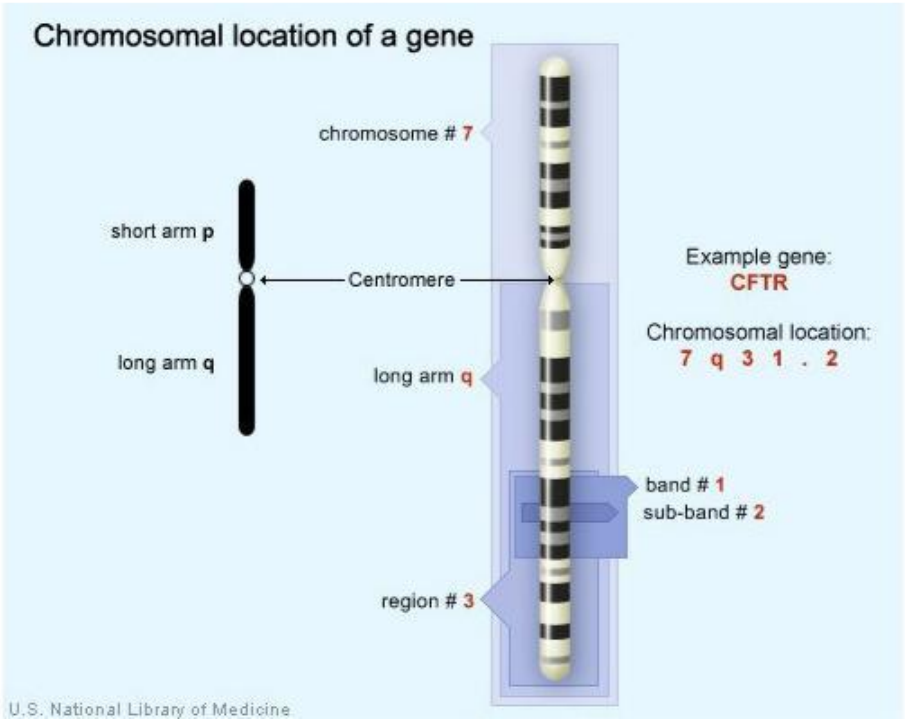
   "Tel" stands for telomeres, which are at the ends of each chromosome.

   Example:

16pcen refers to the short arm of chromosome 16 near the centromere.

14qter (self-doing) ……..?





How is a genetic mapping created?

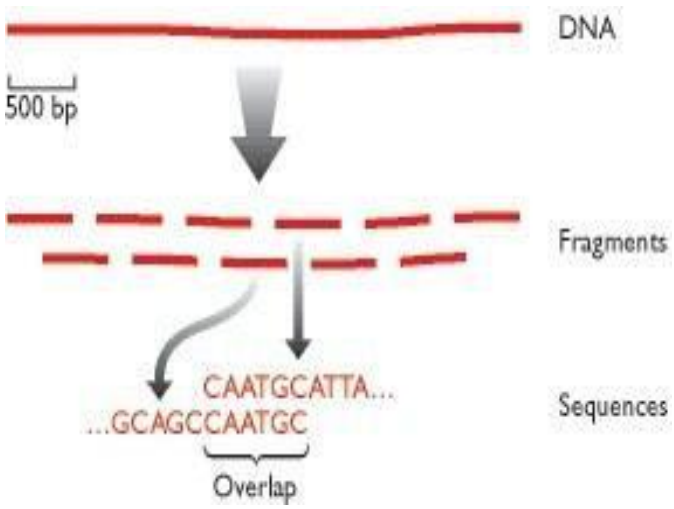
As we mentioned previously, it is done using the rule of Morgan: 1 cM = 1% recombination frequency.

How is a physical mapping created?

Physical mapping involves breaking the chromosomes into smaller DNA fragments that can be handled individually, and ordering these fragments to their respective locations on the chromosomes.

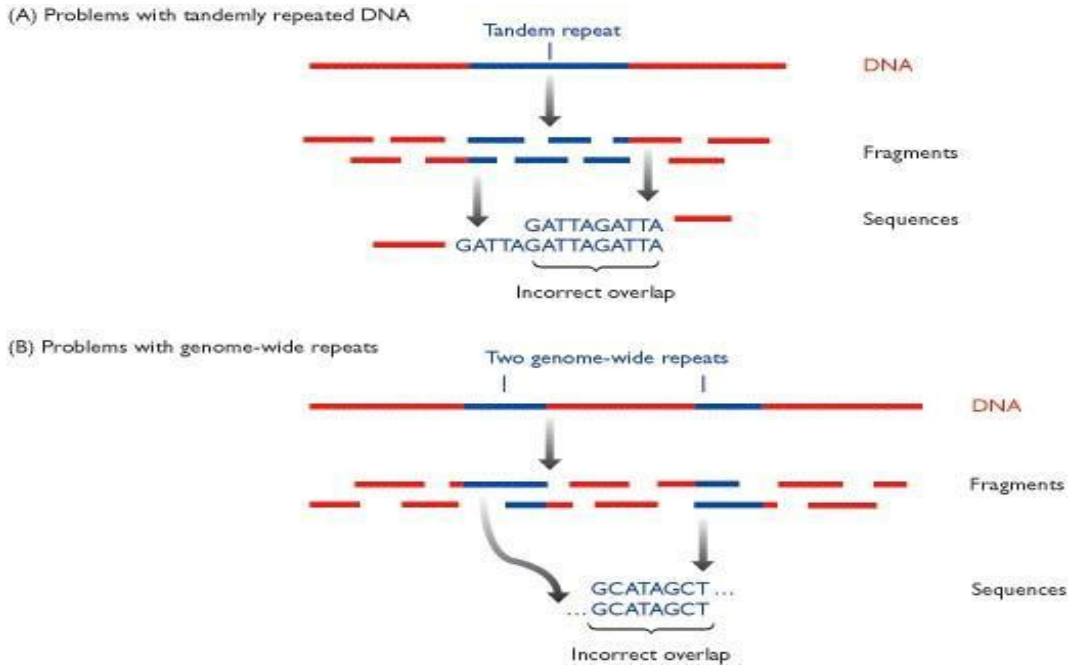
Since actual base-pair distances are generally hard or impossible to directly measure, physical maps are actually constructed by first shattering the genome into hierarchically smaller pieces. By characterizing each single piece and assembling back together, the overlapping path or "tiling path" of these small fragments would allow researchers to infer physical distances between genomic features.

The **fragmentation** of the genome can be achieved by restriction enzyme cutting or through shattering it by sonication. Once cut, the DNA fragments (contigs) forming the genetic fingerprint are assembled by automated or manual means into overlapping DNA stretches (figure below). Now a good choice of clones can be made to efficiently sequence the clones to determine the DNA sequence of the organism under study.



Genome sequencing has one major limitation: even with the most sophisticated technology it is rarely possible to obtain a sequence of more than about 750 bp in a single experiment. This means that the sequence of a long DNA molecule has to be constructed from a series of shorter sequences. The **shotgun sequencing** method is the standard approach for sequencing small prokaryotic genomes, but is much more difficult with larger genomes because the required data analysis becomes disproportionately more complex as the number of fragments increases (for n fragments the number of possible overlaps is given by $2n^2 - 2n$).

A second problem with the shotgun method is that it can lead to errors when repetitive regions of a genome are analyzed. When a repetitive sequence is broken into fragments, many of the resulting pieces contain the same, or very similar, sequence motifs. It would be very easy to reassemble these sequences so that a portion of a repetitive region is
left out, or even to connect together two quite separate pieces of the same or different chromosomes (figure below).



These two limitations of genetic mapping mean that for most eukaryotes a genetic map must be checked and supplemented by alternative mapping

procedures before large-scale DNA sequencing begins. Plenty of physical mapping techniques have been developed to address this problem using molecular techniques.

BENEFITS OF PHYSICAL MAPPING

A genetic map is used to locate and identify the gene or group of genes that determines a particular inherited trait. The techniques developed for genetic mapping have had great impact on the life sciences, and particularly in medicine. But genetic mapping technologies also have useful applications in other fields (agriculture, energy and environment, forensic, …). A round-up of genetic mapping applications would include (but not be limited to) the areas below.

1. Medicine

Scientists have become more proficient in genetic sequencing, the detailed genetic maps that help find and identify the risk genes of genetic diseases for a host such as cystic fibrosis and Duchenne muscular dystrophy or those responsible for more common disorders such as asthma, heart disease, diabetes, cancer, and psychiatric conditions. The ability to investigate the root cause of diseases may one day allow medical researchers to develop strategies for gene therapy, to avoid the environmental conditions that serve as triggers to disease, formulate customized drugs, and techniques and can offer firm evidence that a disease transmitted from parent to child is linked to one or more genes.

The technique can also be used in organ transplants to achieve better matches between recipients and donors**,** thus minimizing the risks of complications and maximizing the use of donated healthy organs, a scarce resource.

2. Agricultural Applications

Knowledge of the genetic maps of plants and animals leads to the development of agricultural crops and animal breeds that are more nutritious, productive and can better resist diseases, insects and drought. Researchers can breed special plants that help clean up wastes that are difficult to break down. Commercialization of the fruits of genomics research promises immense opportunities for industry.

3. Energy and the Environment

Genetic maps of microbes enable researchers to harness the power of bacteria for producing energy from bio-fuels, reducing toxic waste, and developing environment-friendly products and industrial processes.

4. Forensics

You are already familiar with the use of genetic mapping in crime investigations, paternity tests, and identification.

For more delectable applications, genetic mapping can authenticate the origins of consumer goods like caviar, fruits, and wine or the pedigree of livestock and animal breeds.

For more readings:

1. Essentials of Genetics, 8th edition, Klug, Cummings, Spencer, Palladino (eds.), 2013 (online).
2. Biology, 8th edition, 2008, Losos, Mason and Singer (eds.) (in Botany Department Library).
3. Genomes, 2nd edition, 2002, Brown (ed.). (online)

Animations:

Shotgun sequencing

https://www.youtube.com/watch?v=vg7Y5EeZsjk