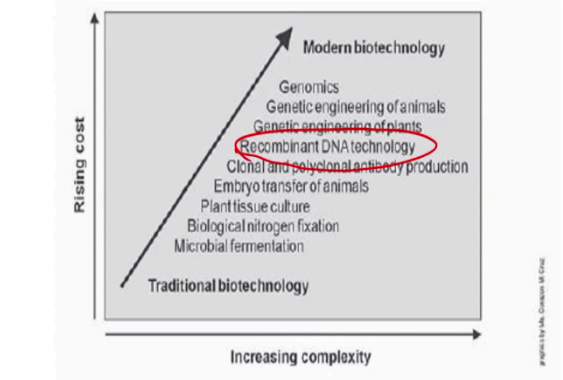
Genome Editing and Engineering
Prof. Utpal Bora
Department of Bioscience and Bioengineering
Indian Institute of Technology, Guwahati

Module - 01
Introduction to genetics and genetic engineering
Lecture - 02
History and Basics of Genetic Engineering

Welcome to my 2nd lecture under module 1 which is on the History and Basics of Genetic Engineering. Here we will discuss a little bit of the historical development of the various scientific events that played a critical role in the development of this discipline as well as some of the important tools that are critical for carrying out genetic engineering work.

(Refer Slide Time: 01:00)



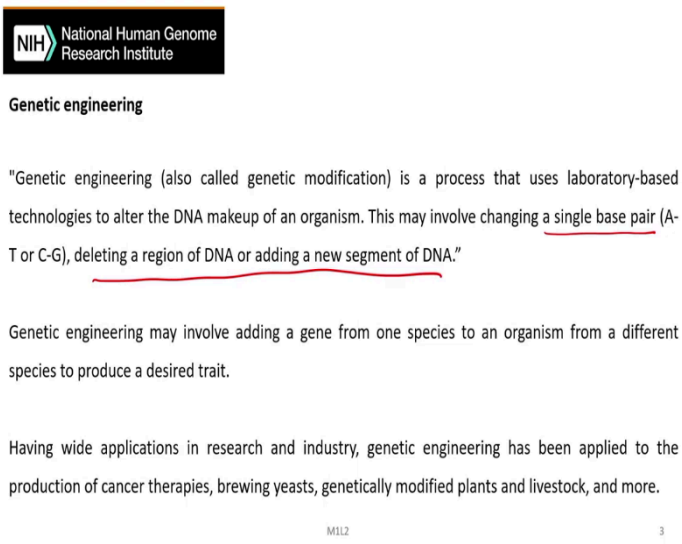
Source: Doyle, J.J. and G.J. Persley (eds.). 1996. *Enabling the Safe Use of Biotechnology: Principles and Practices.* Washington, D.C.: The World Bank.

https://commons.wikimedia.org/wiki/File:Eprimer-genes_figure_1.jpg

Let us begin with a discussion on a historical development starting with traditional biotechnology, which is several thousand years old, where humanity acquired the art of microbial fermentation. It was simple technique and the cost involve for this kind of biotechnology was quite low. But as you can see the movement through this graph upwards from the transition from traditional biotechnology towards modern biotechnology of the current era, you can see that the cost of the technology going up by several folds and in between you can see the positioning of Recombinant DNA technology which we are discussing in this particular module. Missing here is of course, the cost positioning of gene editing and genome engineering which we will discuss at a later point of time. Now, in this

entire developmental journey, we are going to focus on the history of recombinant DNA technology and little bit of the genetic engineering of plants and animals and what are the tools which are required to carry out the same.

(Refer Slide Time: 02:21)

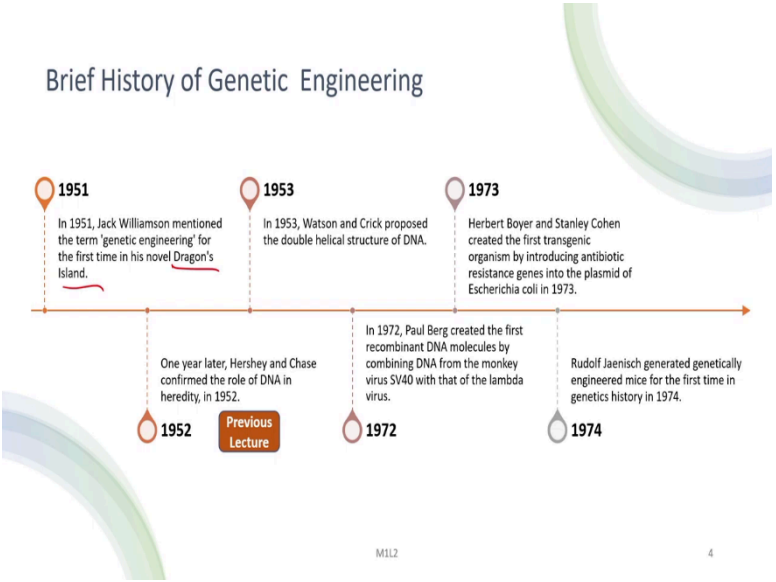


Let us first begin with a definition of genetic engineering which is rather diverse and it means a several things to several people, I have picked up a standard definition offered by the National Human Genome Research Institute, which tells that the word genetic engineering or genetic modification is a process that uses laboratory-based technologies to alter the DNA makeup of an organism. This may involve changing a single base pair or deleting a region of the DNA or adding a new segment of DNA.

Genetic engineering may involve adding a gene from one species to an organism from a different species to produce a desired trait. Having wide applications in research and industry, genetic engineering has been applied to the production of cancer therapies, brewing yeast, genetically modified plants, livestock, and many more.

We are going to discuss some of the things that are required for changing of single base pairs or deleting or adding a new segment of DNA.

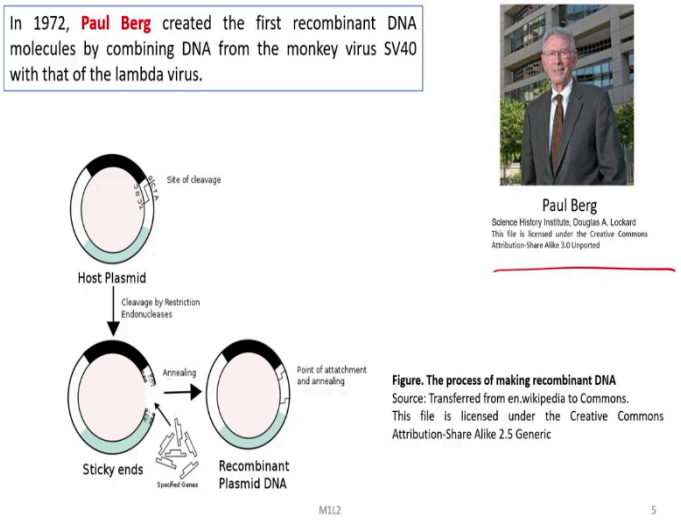
(Refer Slide Time: 03:37)



This is a brief history of the subject Genetic Engineering. So, in 1951, Jack Williamson mentioned the term ‘genetic engineering’ for the first time in his novel ‘Dragon’s Island’. In following year, Hershey and Chase confirmed the role of DNA in heredity in 1952. We had a discussion on these previously. In 1953, Watson and Crick proposed the double helical structure of DNA that is also known to you in last lecture. In 1972, one important development took place; Paul Berg created the first recombinant DNA molecule by combining DNA from the monkey virus SV40 with that of the lambda virus.

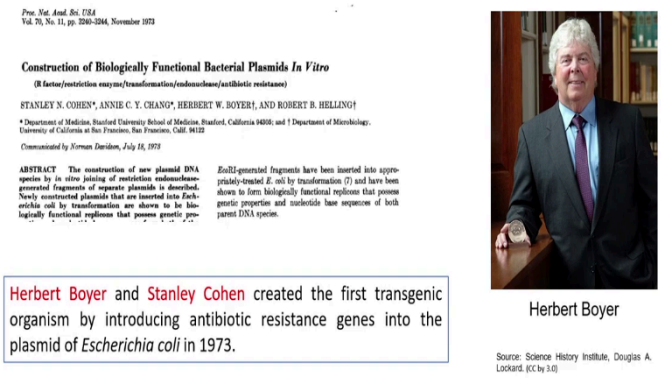
And the following year, Herbert Boyer and Stanley Cohen created the first organism transgenic organism by introducing antibiotic resistant genes in the plasmid of E. coli. In 1974, Rudolf Jaenisch generated genetically engineered mice for the first time in the history of genetics. So, this is in brief the rapid development and as you can see some of the developments took place with a very narrow span of time starting from 1972 to 1974. So, in a way, these rapid developments revolutionized biology in a very different way.

(Refer Slide Time: 05:04)



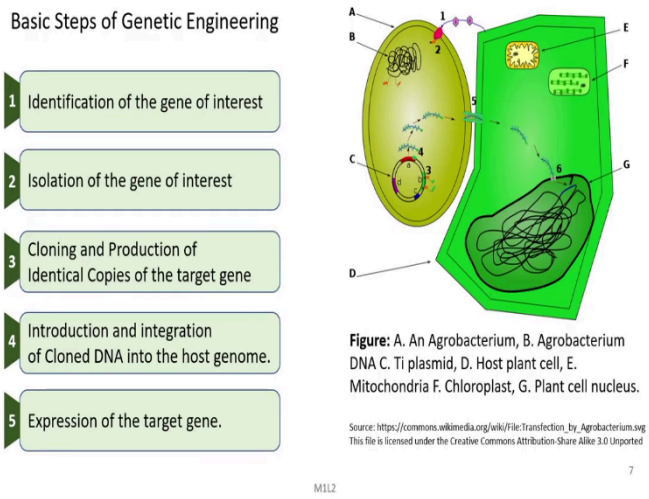
So, this is the contribution of Paul Berg who created the first recombinant DNA molecule, combining DNA from monkey virus SV40 with that of the lambda virus and you can see the photograph of Paul berg on the right side. So, we will be discussing some of these methods how the DNA is cut and joined in the process of genetic engineering.

(Refer Slide Time: 05:33)



So, this is the paper published in 1973 by Cohen and his colleagues in PNAS who created the first transgenic organism by introducing antibiotic resistant genes into the plasmid E. coli in 1973.

(Refer Slide Time: 05:55)

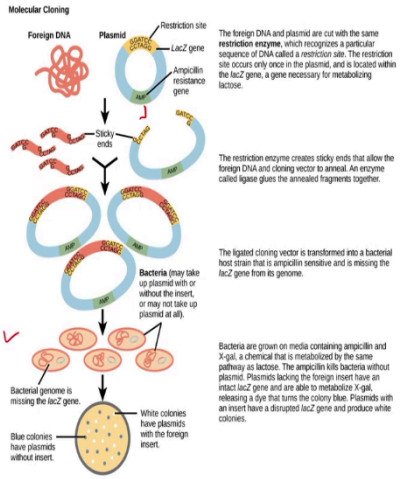


What are the basic steps in genetic engineering or the critical steps in genetic engineering? Although very advanced technique, it has some of the very simple steps in a way. The first step is the identification of a gene of interest. We have to know what is the trait that we want to be expressed in the host organism. So, you need a target gene for this purpose and that target gene has to be identified across species. Once we identify the gene, we need to isolate the gene of interest and then, we move on to the cloning and production stage which is the production of identical copies of the target gene.

Once we do that, we introduce and integrate that cloned DNA into the host genome and finally, the last stage is the expression of the target genome. So, certain steps in plant genetic engineering and animal genetic engineering are little bit different due to the nature of the organisms.

But more or less, all these five steps are followed in each and every case. In this figure, you can see the transfer of genetic character with the help of a vector to another host organism, which is basically a plant cell and we are using agrobacterium plasmid over here. We will discuss this in one of the sections later on.

(Refer Slide Time: 07:34)



So, let us look into the molecular technique process in a little bit further detail. As I already told you, you need a target gene or identified gene. So, we have these foreign DNA or foreign gene, that we have identified and then, we need a carrier. For example, here we are taking a plasmid, which we call as a vector into which we load this foreign DNA and then, make it hybrid plasmid or a mutated plasmid which carries these targeted gene and then, transfer these into the host cells and then, we screen the host cells for the mutants.

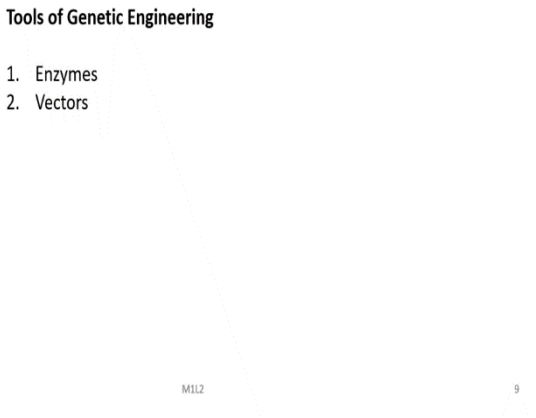
So, for doing this, we need to cut open this plasmid. So, that is being done with the help of restriction enzymes. There are certain things which are going to discuss later, when we discuss about the features of a plasmid that it has certain antibiotic resistance genes which helps us in the selection of the mutants. So, the foreign DNA and plasmid are cut with the same restriction enzyme, which recognize a particular sequence of DNA called a restriction site.

The restriction occurs only once in the plasmid and is located within the lacZ gene. It is necessary for metabolizing lactose; we will be discussing the role of these lactose gene, lacZ gene in the expression later on.

The restriction enzyme creates sticky ends that allows the foreign DNA and cloning vector to anneal. An enzyme called ligase glues the annealed fragments together. The ligated cloning vector is transformed into a bacterial host strain which is ampicillin sensitive and is missing the lacZ gene from its genome. So, finally, the bacteria are grown on media containing

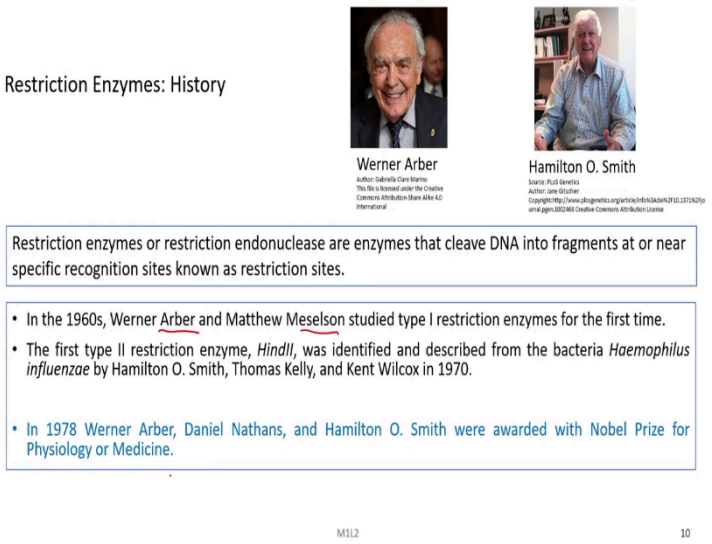
ampicillin and X-gal, if there is no any gene transfer or the transfer of the plasmid containing the foreign DNA and which has an ampicillin resistance gene, the bacterial cells will die because they are ampicillin sensitive. Only those bacterial cells will survive into which the plasmid has been transferred successfully. So, this is the basic principle of the molecular cloning and it is selection process over here. This we will discuss in later slides as well.

(Refer Slide Time: 10:22)



Now, let us focus on what the tools of this technique, genetic engineering. This is basically a kind of a technique which you can tell as a molecular carpentry, where lot of cutting and joining work is involved as in the case of making furniture. So, for cutting, we require enzymes as well as for modification of the DNA ends or fragments and then, we need the vectors as we have seen in the earlier slide for carrying these the targeted gene into the host organism.

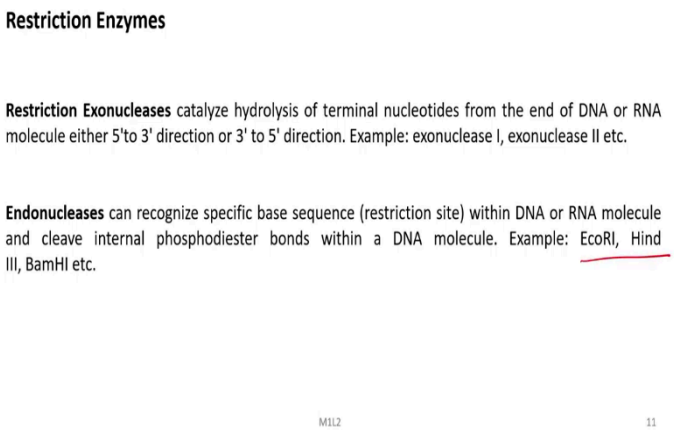
(Refer Slide Time: 11:00)



So, we will discuss about these enzymes a little bit in detail and the first enzymes that are important for this technique is obviously, the restriction enzymes which cut open the DNA molecule or the plasmid vectors as well as which helps us in cutting out the desired gene from the species in which it is located. These enzymes cleave DNA into fragments at or near specific recognition sites known as restriction sites.

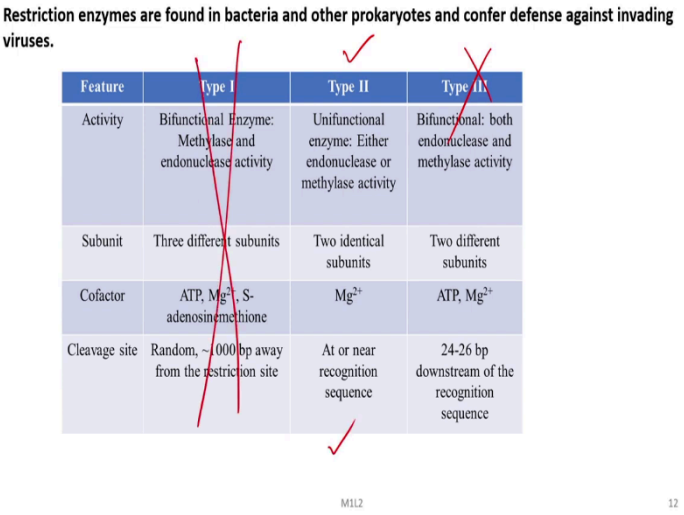
So, it was in 1960, two persons Werner Arber and Matthew Meselson, who studied type I restriction enzymes for the first time. The first type II restriction enzyme, HindII, was identified and described from the bacteria Haemophilus influenzae by Hamilton O. Smith, Thomas Kelly and Kent Wilcox after a decade of these in 1970. In 1978, Werner Arber, Daniel Nathans and they were awarded the Nobel Prize for physiology and medicine for their discovery.

(Refer Slide Time: 12:13)



Basically, what these restriction enzymes does? They fall into two classes the Restriction Exonucleases and the Restriction Endonucleases. The Exonucleases catalyzes hydrolysis of terminal nucleotide from the end of DNA or RNA molecule either 5' or 3' direction or 3' to 5' direction. You have examples like exonuclease I, II etcetera. While endonucleases recognize a specific big sequence, which we call as restriction sites within DNA or RNA molecule and cleave internal phosphodiester bonds within a DNA molecule and examples are EcoR1, Hind III, BamH1 etcetera.

(Refer Slide Time: 13:00)



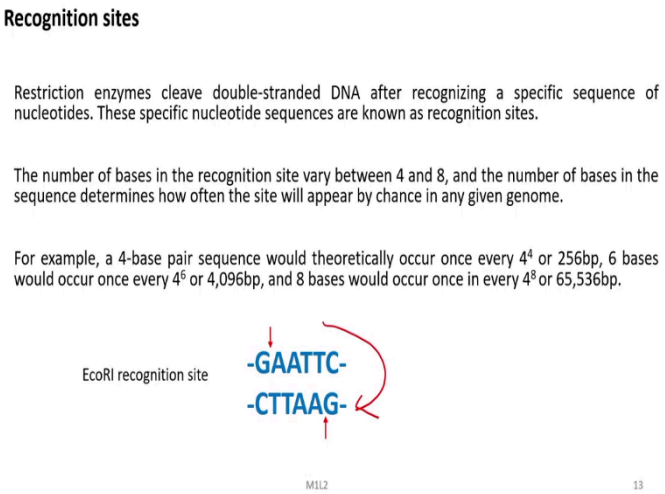
The restriction enzymes are found in bacteria and other prokaryotes and confer defense against invading viruses. There are several types; type I, II and III and they have different kind of structure. For example, type I has three different subunits; type II has two identical subunits; type III has two different subunits and they require certain co factors to act and you can see that type I requires ATP, Mg^{2+}, S-adenosyl methionine. Type II requires Mg^{2+}. Type III requires ATP and Mg^{2+} and the cleavage pattern is different in all of these enzymes.

Type I the cleavage pattern is random; it is 1000 base pairs away from the restriction site. In type III, the restriction again is random; but it is not that far away from the restriction site, it is about 24 to 26 base pairs downstream of the recognition site. Type II is something very interesting, where the restriction site and recognition site are located or co-located. So, they are at or near the recognition sequence.

So, in the case of type I enzyme, we know the binding site for both type I and type III, but we do not know where it is going to cleave. We know the distance at which it is going to cleave; but we do not know the sequence at which it is going to cleave. So, they are very unspecific digestors or cutters in a way. But in the case of type II, we know exactly where it is going to bind and exactly where it is going to cut.

For genetic engineering, we cannot rely on unspecific restriction enzymes like type I and type III. As regards with the cutting, we need specific cutting for our work and for this purpose, only type II enzyme is used for a genetic engineering or cloning.

(Refer Slide Time: 15:10)



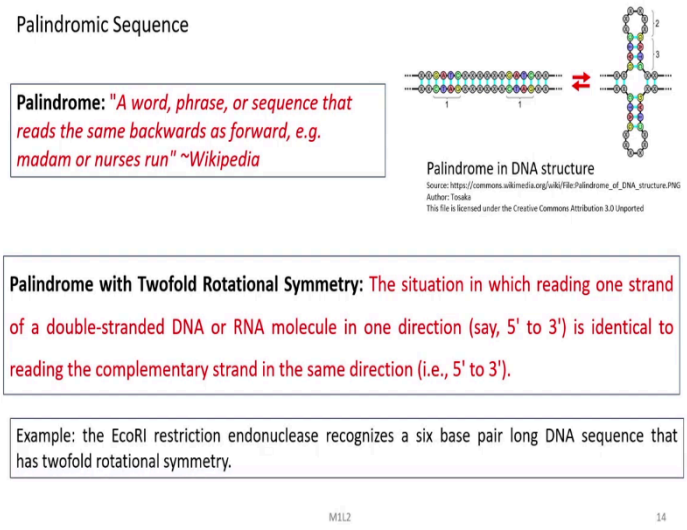
Now, let us discuss what are the recognition sites that we are referring to all the time. The restriction enzymes cleave double stranded DNA after recognizing a specific sequence of nucleotides. These specific nucleotide sequences are what we call as recognition sites.

The number of bases in the recognition sites vary between 4 and 8 and the number of bases in the sequence determines how often the site will appear by chance in any given genome. For example, a 4-base pair sequence would theoretically occur every 4^4; 6 base paired occur every 4^6 and 8 base paired occur every 4^8 base pairs.

So, a 4-base pair cutter appears more frequently in the genome than an 8 base pair cutter or a sequence. So, if we want only partial digestion of a given genomic DNA, we will use the 8 base pair cutter. And this is very important that for mapping, we will use basically the rare cutter which is for example, the 8 base pair cutter, we will refer to as a rare cutter because the occurrence of those sequences are comparatively rare in comparison to the 4-base pair cutter.

Now, this is the recognition site of the typical enzyme EcoR1. You can see over here, marked by the red arrows. So, similarly, other enzymes also have their specific recognition sites.

(Refer Slide Time: 16:58)

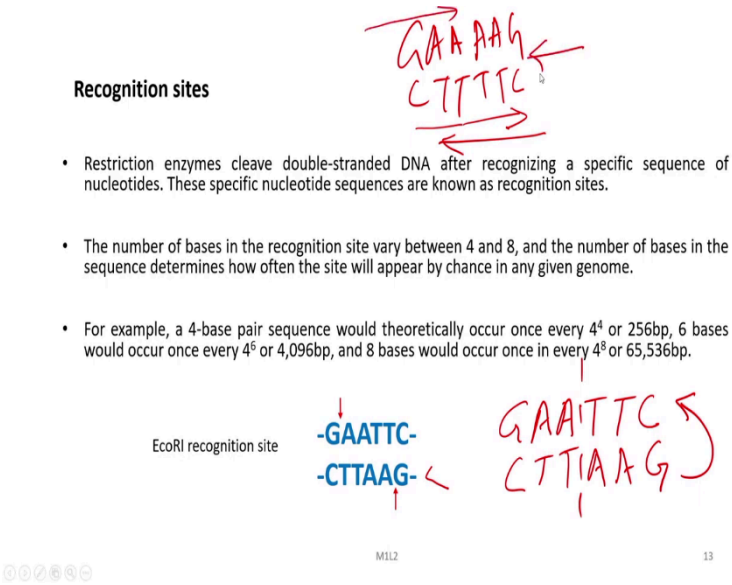


Another concept that is important in this regard is the Palindromes or Palindromic Sequences. So, what are DNA palindromes? Palindrome is a word or phrase or sequence, they are read the same backward and forward and these are the examples like ‘madam’ or ‘nurses run’. Similarly, ‘Malayalam’ is also a word which reads the same from both the ends. In DNA also,

we have this kind of symmetry, but this symmetry is twofold rotational symmetry. So, we have to turn the double stranded DNA to get a palindrome. So, that is why we call it as a twofold Rotational Symmetry.

So, this is the situation in which reading one strand of a double strand DNA or RNA molecule in one direction say 5' or 3' is identical to reading the complementary strand in the same direction or from the opposite direction. And for example, in the EcoRI restriction nuclease, as we have shown in the earlier case, if you can read it GAATTC, GAATTC. So, you have to continue the reading from the forward strand to the reverse strand. So, you will get and if you read it from the other side, CTTAAG, CTTAAG, so they will form a double stranded rotational symmetry.

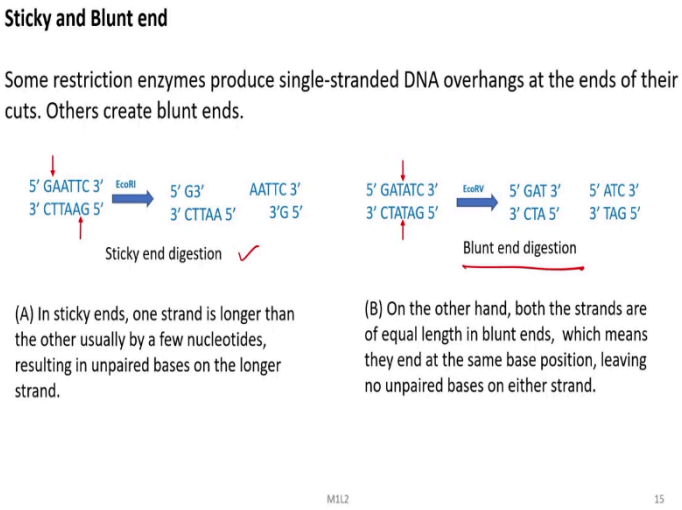
(Refer Slide Time: 18:37)



Another way to understand is if we write this sequence and carry out some operation. For example, we have a line dividing them in this way and then, we rotate this; we rotate this by 180°. So, the sequence will become GAA and below it is CTT and you revert it, it becomes AAG and this will become TTC.

So, now, whether we read CTT TTC from this side or this side and this sequence from this side or this side, we are going to have two different palindromes in this case. I hope this discussion makes it clear what do you mean by two-fold rotational symmetry. This is very important for understanding the recognition and cleavage site of type II restriction enzymes.

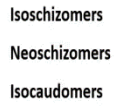
(Refer Slide Time: 19:41)



Now, these restriction enzymes when they recognize and cut a particular sequence, they may yield two type of ends; one end may be sticky and the other end may be blunt end and here are some of the examples, where you can see the different kinds of ends being generated. For example, when EcoRI digests its sequence, the type of ends that it creates its sticky; but EcoRV generates a blunt end digestion.

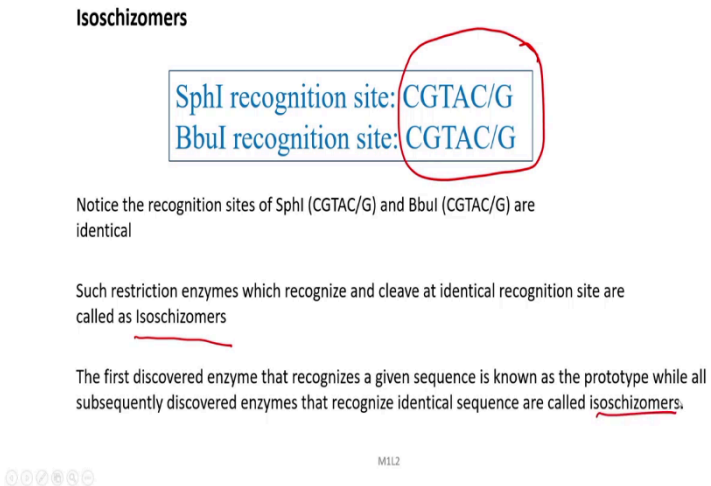
So, in the sticky end, one strand is longer than the other usually by a few nucleotides and we have overhangs over here. But in the blunt end digestion, both the strands are equal and that is blunt, which means they end at the same base position leaving no unpaired bases on either strand.

(Refer Slide Time: 20:36)



Now, in to discuss about Isoschizomers, Neoschizomers and Isocaudomers with respect to the digestion patterns.

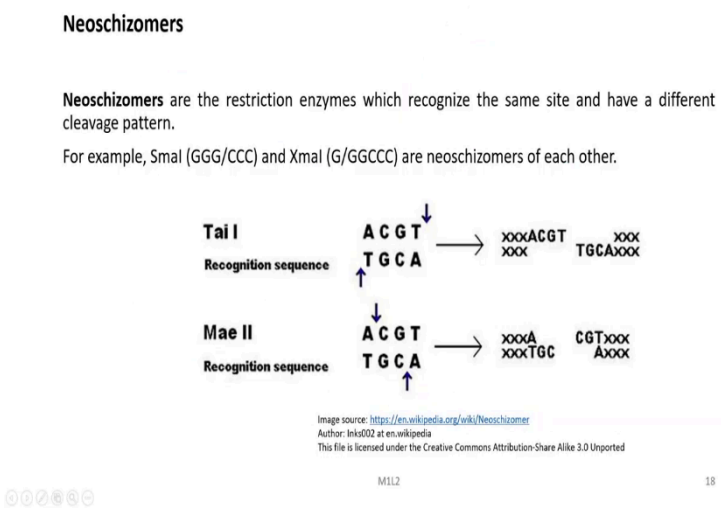
(Refer Slide Time: 20:52)



So, what are isoschizomers? So, here are two restriction enzymes and their recognition sites; SphI and BbuI. If you look keenly into the two recognition sequences over here, they are same CGTACG for both the restriction enzymes. Such recognition enzymes which recognize and cleave at the identical recognition sites are called as isoschizomers.

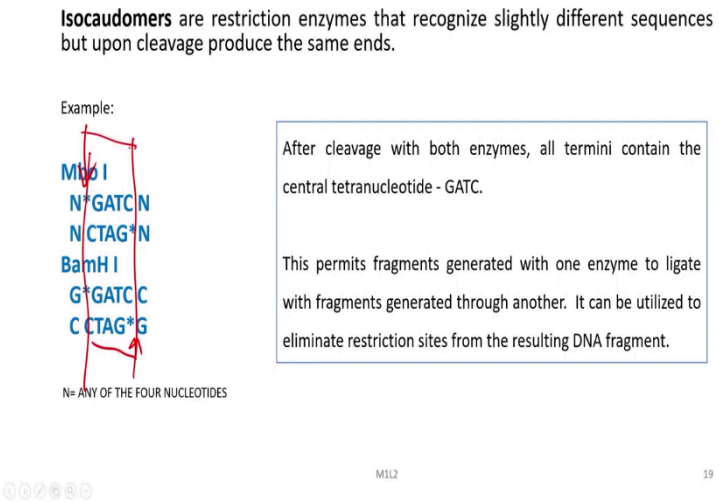
The first discovered enzyme that recognize a given sequence is known as the prototype, while all subsequently discovered enzymes that recognize identical sequences are called as isoschizomers. So, we may have many different kinds of enzymes which identifies the same sequence. The first one will be the prototype and their other following members discovered later will be called as isoschizomers of the prototype.

(Refer Slide Time: 21:57)



Now, there is another concept called neoschizomers. These are the restriction enzymes which recognize the same site and have a different cleavage pattern. So, here you can have two examples of Sma I and Xma I and they there are neoschizomers of each other.

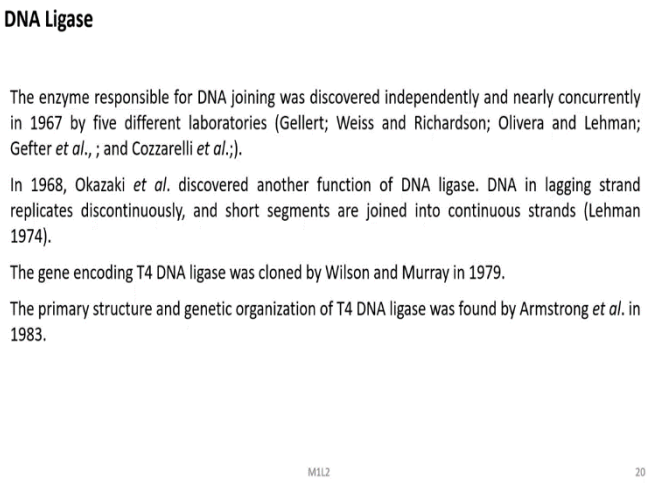
(Refer Slide Time: 22:26)



And here you have Tai I and Mae II which are neoschizomers of one another. What are isocaudomers? These are restriction enzymes that recognize slightly different sequences, but upon cleavage produce the same ends. This is little bit different from the earlier two cases. So, you have two examples here; Mbo I and BamH I and you can see that the recognition sequences are different at the ends.

The central sequence GATC is same; this is same, but the terminal sequences N can be any of the bases are different. But when they will be cleaving, this is the cleavage point you can see over here. So, they generate the same sequence. So, that these kind of enzyme pairs are known as isocaudomers. These isocaudomers permit fragments generated with one enzyme to ligate with fragments generated through another, it can be utilized to eliminate restriction sites from the resulting fragments.

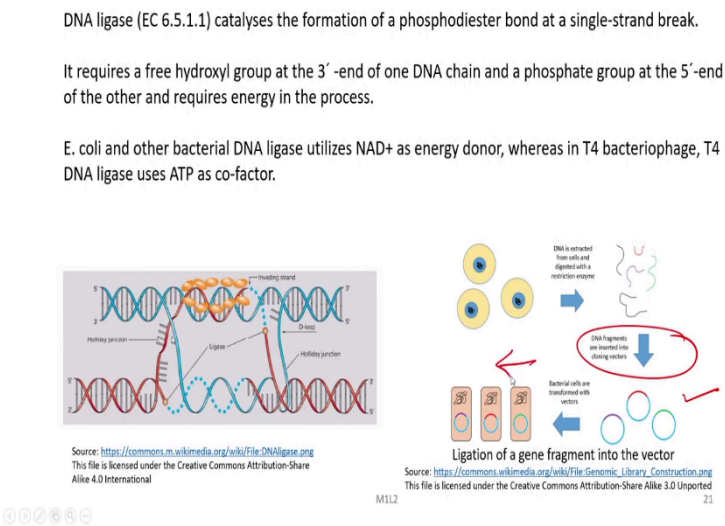
(Refer Slide Time: 23:39)



So, we now know a lot about enzymes which digest DNA and which generates either sticky ends or blunt ends. Now, let us learn about enzymes which do the opposite operation. That is instead of cutting they join DNA fragments. Such enzymes are known as ligases or DNA ligases. So, they are the enzymes responsible for DNA joining and was discovered independently and nearly concurrently by 5 different persons working in 5 different laboratories in 1967.

In 1968, Okazaki et al. discovered another function of the DNA ligase. DNA in the lagging strand replicates discontinuously and the short fragments generated by this process are joined into continuous strands with the help of a DNA ligase. The gene coding T4 DNA ligase was cloned by Wilson and Murray in 1979. The primary structure and genetic organization of the T4 DNA ligase was found by Armstrong et al. in 1983.

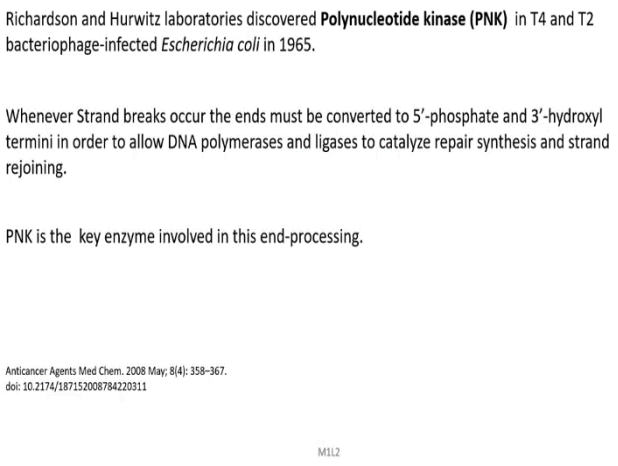
(Refer Slide Time: 24:55)



So, as already discussed DNA joins, DNA ligase joins DNA fragments. So, it basically catalyzes the formation of a phosphodiester bond at a single strand break and it requires free hydroxyl group at the 3' end of one DNA chain and a phosphate group at the 5' end of the other and requires energy in the process. E. coli and other bacterial DNA ligases utilize NAD plus as energy donor; whereas, T4 bacteriophage DNA ligase uses ATP as a co-factor.

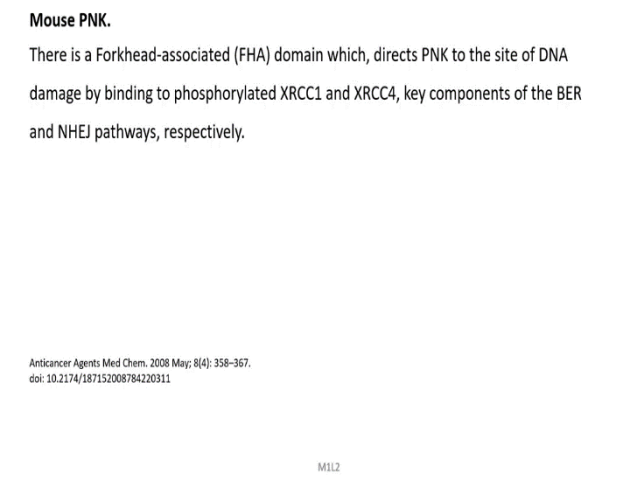
So, here, you can see the role of DNA ligase at the step, where you want to join the DNA fragments into a cloning vector, before we transform the bacterial cells with the vectors loaded with target DNA.

(Refer Slide Time: 25:57)



Another enzyme which is important for genetic engineering is PNK or polynucleotide kinase. So, it was discovered by Richardson and Hurwitz in Escherichia coli in 1965 which was infected by T4 and T2 bacteriophage. Whenever there is a strand break, the ends must be converted into 5' phosphate and 3' hydroxyl termini in order to allow DNA polymerases and ligases to catalyze repair synthesis and strand rejoining.

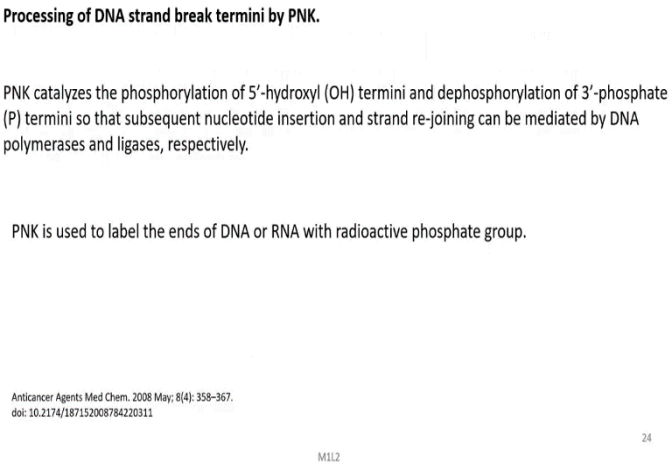
(Refer Slide Time: 26:40)



PNK is the key enzyme involved in this end processing. So, there are various sources of PNK. For example, you have a mouse PNK and they have some conserved structures there is

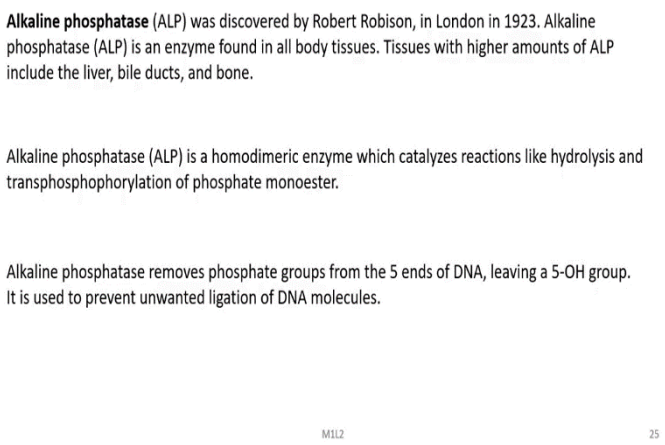
a Forkhead-associated domain which directs the PNK to the site of DNA damage. PNK has to find out the position in the DNA which needs to be repaired and Forkhead associated domain helps PNK in doing this.

(Refer Slide Time: 27:11)



How the DNA strand is processed at the termini by a PNK. So, PNK catalyzes the phosphorylation of 5' hydroxyl termini and dephosphorylation of the 3' phosphate termini so that subsequent nucleotide insertion and strand re-joining can be mediated by DNA polymerases and ligases respectively. Another application of PNK is in the labeling the ends of DNA or RNA with radioactive phosphate group.

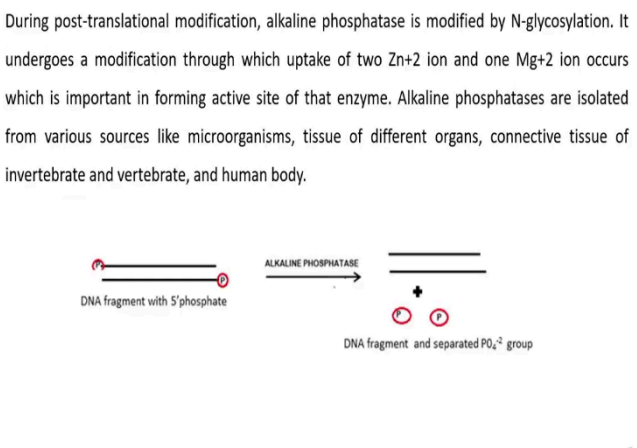
(Refer Slide Time: 27:49)



And this is highly helpful in DNA auto radiography experiments. The next enzyme which is important in the process of recombinant DNA technology is alkaline phosphatase which was discovered by Robert Robinson in London in 1923. Alkaline phosphatase is an enzyme found in all body tissues and some tissues contain very high amounts of this enzyme like the liver, bile ducts and bone.

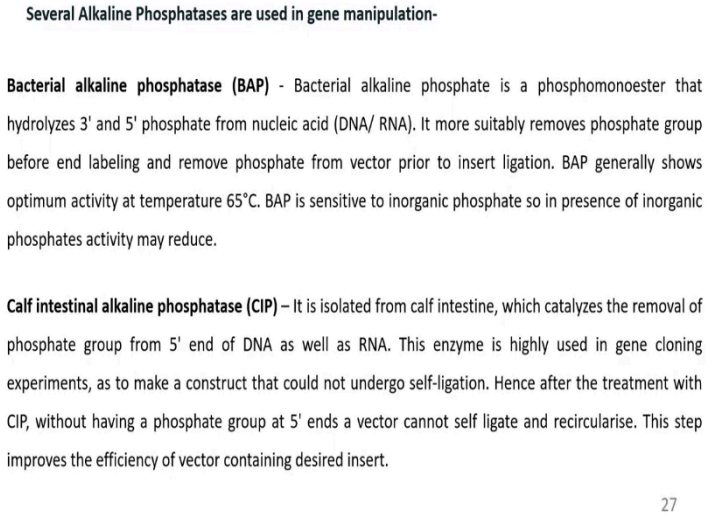
This enzyme is highly useful in the process of genetic engineering. It is a homo dimeric enzyme which catalyzes reactions like hydrolysis and transphosphorylation of phosphate monoester. It removes phosphate groups from the 5 prime ends of DNA leaving a 5' hydroxyl group and is used to prevent unwanted ligation of DNA molecules.

(Refer Slide Time: 28:46)



During post translational modification, alkaline phosphatase is modified by N-glycosylation. It undergoes a modification through which uptake of two zinc ions and one magnesium ions occur which is important in forming active site of that enzyme. Alkaline phosphatases are isolated from various sources and tissues and you can see the reaction over here, how the DNA fragment with 5' phosphate under the catalytic activity of alkaline phosphatase become free of the phosphate groups at the DNA ends.

(Refer Slide Time: 29:29)

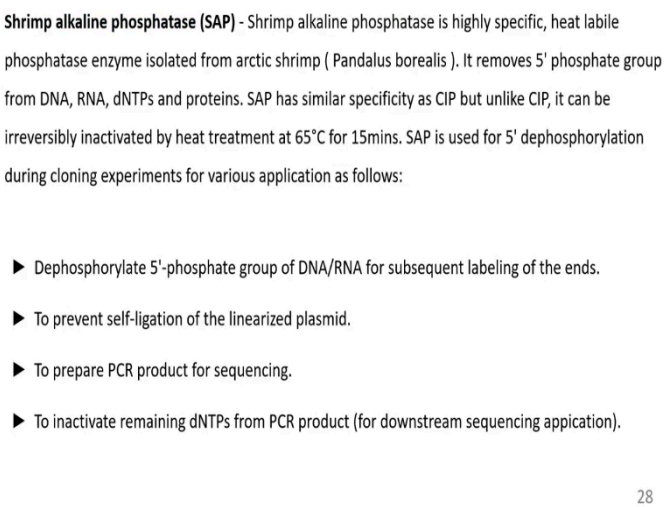


As already told you there are several sources of alkaline phosphatase and they are used for genetic manipulation. We have bacterial alkaline phosphatases, calf intestinal alkaline phosphatases. As the name indicates they are isolated from the respective organisms. BAP is a phosphomonoester and it hydrolyzes 3' to 5' phosphate from nucleic acid. It is more suitable for removing phosphate group before and labeling and remove phosphate from vector prior to insert ligation.

BAP generally shows optimum activity at 65° C. It is sensitive to inorganic phosphates. So, in presence of inorganic phosphates, the activity may be reduced.

The calf intestinal alkaline phosphatase is isolated from calf intestine, catalyzes the removal of the phosphate group from the 5' end of the DNA as well as RNA. This enzyme is highly used in gene cloning experiments to make construct that could not undergo self-ligation. Hence, after the treatment with CIP without having a phosphate group at 5' ends, a vector cannot self-ligate and recircularise. This step improves the efficiency of the vector containing the desired insert.

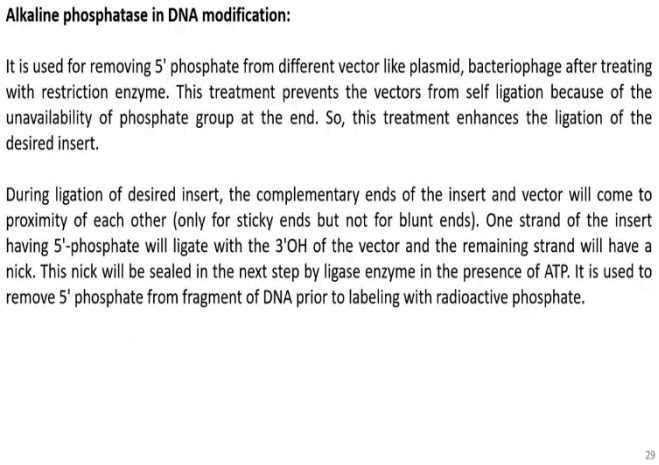
(Refer Slide Time: 30:52)



There is another source from shrimp and we call it shrimp alkaline phosphatase. This is highly specific heat labile phosphatase enzyme and this is isolated from the arctic shrimp. It removes 5' phosphate group from DNA, RNA, dNTPs and proteins. SAP has similar specific activity as SIP; but unlike CIP, it can be irreversibly inactivated by heat treatment at 65° C for 15 minutes.

SAP is used for 5’ dephosphorylation during cloning experiments for various applications as follows. It dephosphorylates 5' phosphate groups of DNA, RNA for subsequent labeling of the ends. It is used to prevent self-ligation of the linearized plasmid and to prepare PCR products for sequencing, to activate remaining dNTPs from PCR products.

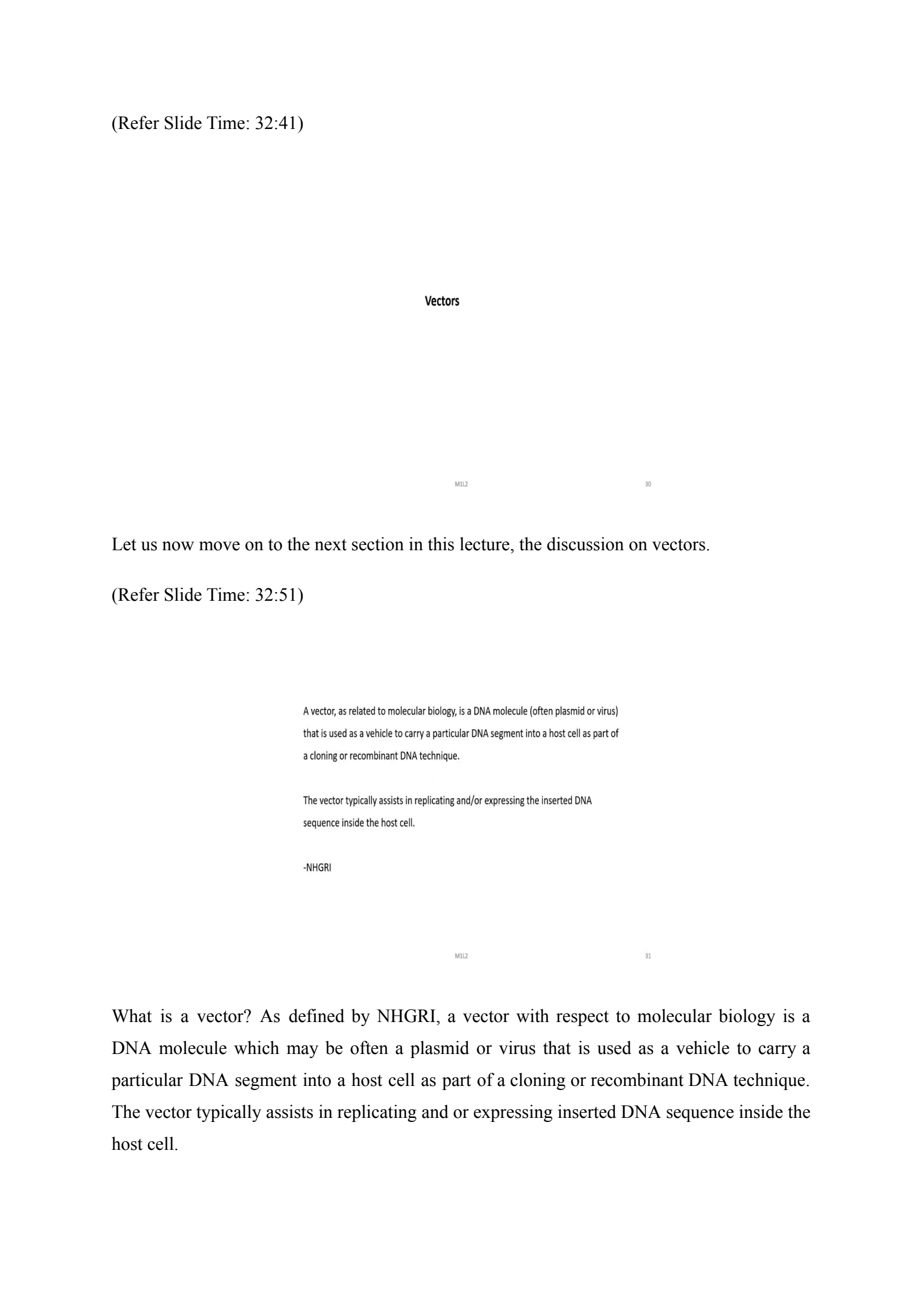
(Refer Slide Time: 31:51)



How alkaline phosphatase help in the DNA modification is more or less now known to you. It is used for removing the 5' phosphate from different vectors like plasmids, bacteriophages after treating with restriction enzymes. This treatment will prevent the vectors from self-ligation because of the unavailability of phosphate group at the end. So, this treatment enhances the ligation of the desired insert.

During ligation of desired insert, the complementary ends of the insert and the vector will come to proximity to each other. One strand of the insert having 5 prime phosphate will ligate with the 3' -OH of the vector and the remaining strand will have a nick. This nick will be sealed in the next step by ligase enzyme in the presence of ATP. It is used to remove 5' phosphate from fragment of DNA prior to labeling with radioactive phosphates.

(Refer Slide Time: 32:41)

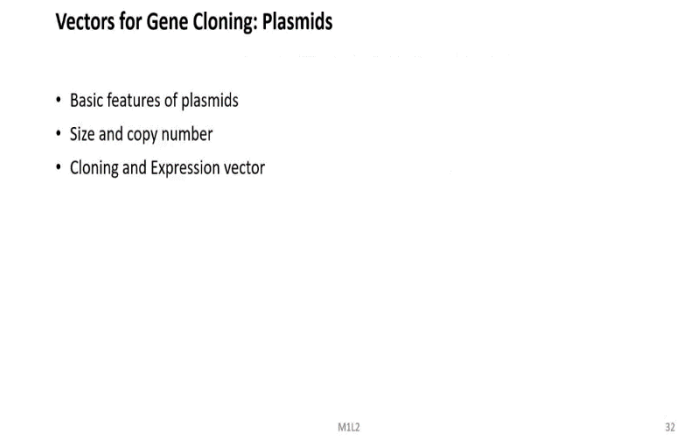


Let us now move on to the next section in this lecture, the discussion on vectors.

(Refer Slide Time: 32:51)

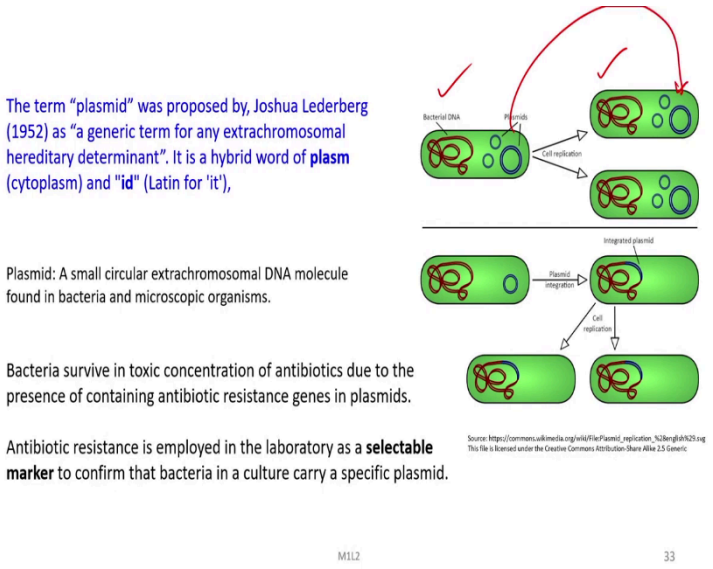
What is a vector? As defined by NHGRI, a vector with respect to molecular biology is a DNA molecule which may be often a plasmid or virus that is used as a vehicle to carry a particular DNA segment into a host cell as part of a cloning or recombinant DNA technique. The vector typically assists in replicating and or expressing inserted DNA sequence inside the host cell.

(Refer Slide Time: 33:19)



There are various kinds of vectors. Let us start the discussion with plasmids and look into the basic features of plasmid vectors and important concepts like size and copy number and different types of vectors like cloning and expression vector.

(Refer Slide Time: 33:36)



This term 'plasmid' was proposed by Joshua Lederberg as a generic term for any extra chromosomal hereditary determinant. It is a hybrid word of plasm coming from cytoplasm and 'id' which is a Latin word for it, which together make the word 'plasmid' which is basically a small circular extra chromosomal DNA molecule found in bacteria and

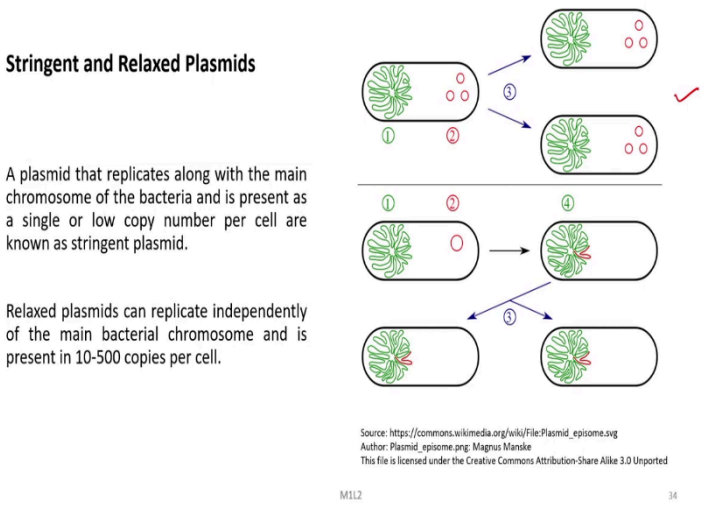
microscopic organism. This is called extrachromosomal DNA because the DNA has a main bacterial chromosome and anything which is in addition to that are called as the plasmid DNA.

Now, with every cell replication, you can see over here both the bacterial DNA and the plasmid DNA are replicated. So, in many cases, the replication of the plasmid may be independent of the replication of the bacterial DNA and sometimes, they may be connected to each other.

The rate of bacterial main chromosome replication may govern the rate of plasmid DNA replication that we will discuss later. In certain cases, the plasmid may get integrated into the main chromosome and in this case, the rate of replication obviously, will be the same because now the plasmid DNA has become a integral part of the main chromosomal DNA.

Now, we know that bacteria is a very hardy organism. It can survive in very difficult environment such as highly toxic amounts of antibiotics. All these fitness characters are due to certain genes present in the extra chromosomal DNA called as plasmids. This antibiotic resistance trait is a boon for the recombinant DNA technology discipline. Here, it is employed as a selectable marker to confirm the bacteria in a culture to carry out a specific plasmid.

(Refer Slide Time: 36:04)

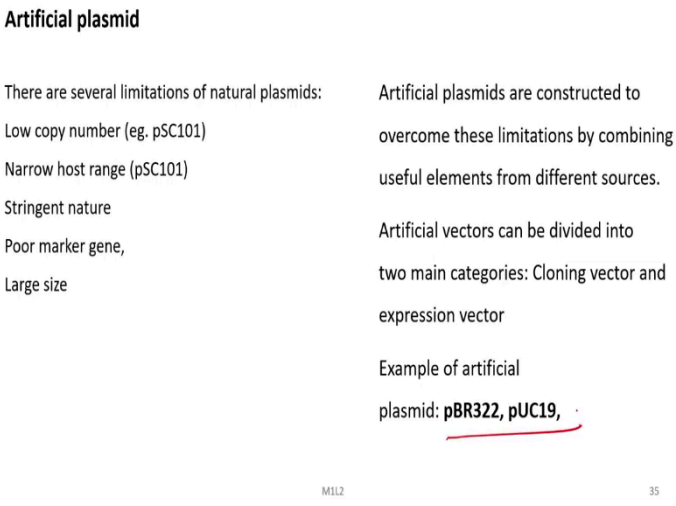


Now, what are stringent and a relaxed plasmid. So, I just discuss about the relation between the main bacterial plasmid and the extra chromosomal bacterial plasmid DNA with respect to their replication.

A plasmid that replicates along with the main chromosome of the bacteria and is present as a single or low copy number as per cell are known as a stringent plasmid. While relax plasmids can replicate independently of the main bacterial chromosome and is present in 10 to 5000 copies per cell. So, here you can see the two different cases; here the plasmid number and the main chromosome number is remaining constant.

But in this case, the number in the case of relax plasmid the number will be more because the main chromosomal replication cannot govern the number of the plasmid DNA.

(Refer Slide Time: 37:11)



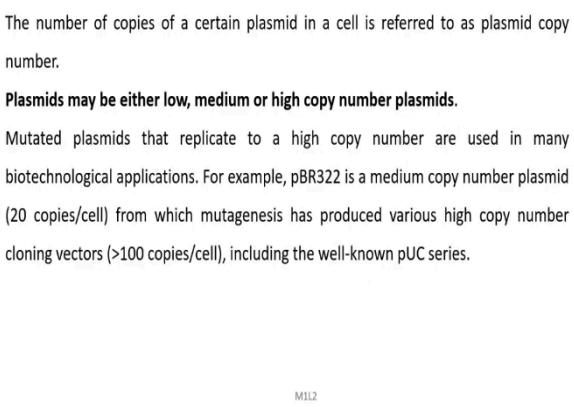
So, that those are natural plasmids that we are discussing. Now, we cannot carry out recombinant DNA technology work with the help of natural plasmids because certain features are missing in those plasmids. So, for our work, we have to develop the tools, develop the vector in such a way that they become amenable to the various demands of these technique. For this, artificial plasmids were created and there are several limitations which are overtaken by these artificial plasmids over the natural plasmids.

For example, artificial natural plasmids may be having low copy number, they may have narrow host range and are stringent nature; then, they have poor marker genes and their size

may be very large. So, a large size vector will not be suitable for cloning because with the insertion of the gene of interest, it will become much larger and then, delivery of these vector into the host will become a challenge. So, smaller vectors having higher loading capacity are the most desirable for gene transfer into the host. Now, let us go by one by one, what are the special things and features that we require in a plasmid vector.

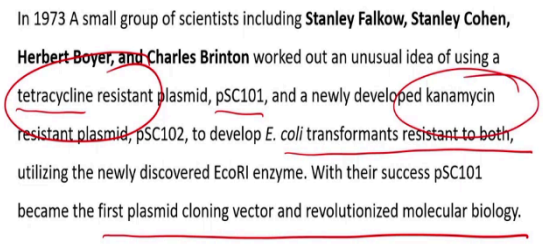
Artificial plasmids are constructed to overcome these limitations by combining useful elements from different sources. You may not be finding the features in a single place. So, we bring in all the suitable features from different plasmids and then, we combine them into one single vector. The artificial vectors can be divided into two main categories; the cloning vector and the expression vector.

(Refer Slide Time: 39:19)



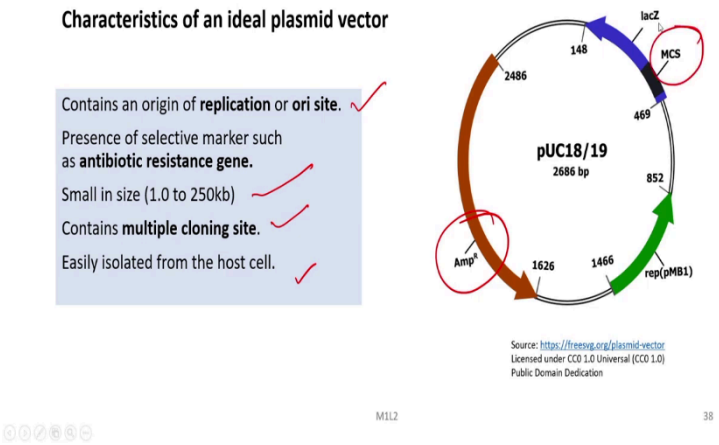
And here some of the artificial plasmid vectors pBR322 or pUC19. The number of copies of a certain plasmid in the cell is referred to as the plasmid copy number which is already discussed in a way. The plasmids may be either low, medium or high copy number plasmids. Mutated plasmids that replicate to a high copy number are used in many biotechnological applications.

(Refer Slide Time: 39:47)



If we want to do cloning, if we have a high copy number vector that is much more advantageous than a low copy number vector. In 1973, a small group of scientists comprising of Stanley Falkow, Stanley Cohen, Herbert Boyer and Charles Brinton worked out an unusual idea of using a tetracycline resistant plasmid pSC101 and a newly developed kanamycin resistant plasmid pSC102 to develop E. coli transformants resistance to both tetracycline and kanamycin, utilizing a newly discovered EcoR1 enzyme which is a restriction enzyme. And you already know the recognition site of this particular enzyme and they become successful in doing so and they created history by creating the first plasmid cloning vector which revolutionized molecular biology. So, as I told you earlier, we combine desirable features, picking them up from various other sources. So, in this case, this particular team picked up the tetracycline resistance sequence and the kanamycin resistance sequence and joined them into one single vector to create this vector with a hybrid character having resistance to both the antibiotics.

(Refer Slide Time: 41:09)

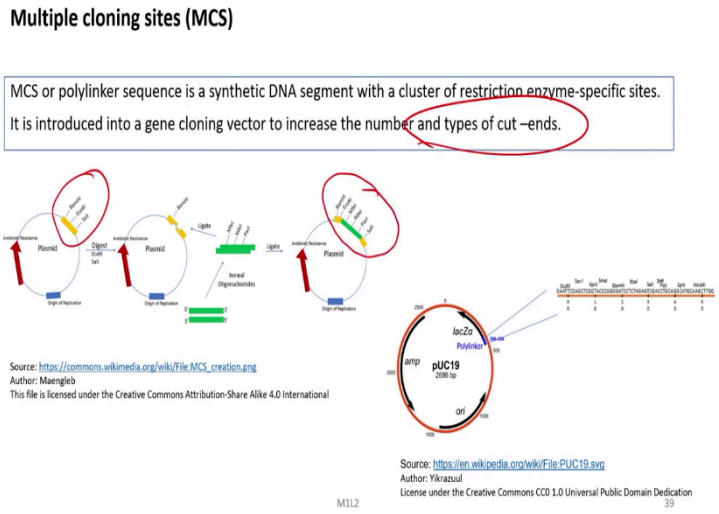


Let us discuss in brief the characteristics of an ideal plasmid vector. A plasmid we are telling should be ideally, vector should be ideally of high copy number. Now, for replication of these particular vector, we always need a replication origin without which the plasmid cannot replicate. So, one of the desired characters of a plasmid vector that it should contain an origin of replication or Ori site.

In the earlier case, we discussed how two different antibiotic marker sequences were joined into one single vector. So, another desirable character is that there should be presence of selective markers such as antibiotic resistant genes. Plasmids should be vectors particularly, should be smaller in size; the larger size plasmids are not suitable to be vectors due to challenges involved in their delivery.

Now, for cloning the desired targeted DNA sequence or gene, we knew they need a cloning site. A single enzyme specific cloning site is not ideal because that limits the choice of the cutting ends generated. So, we need a multiple cloning site which can be digested by different kind of enzymes and the most important thing is that these plasmid vectors should be easily isolated from the host cell.

(Refer Slide Time: 42:56)

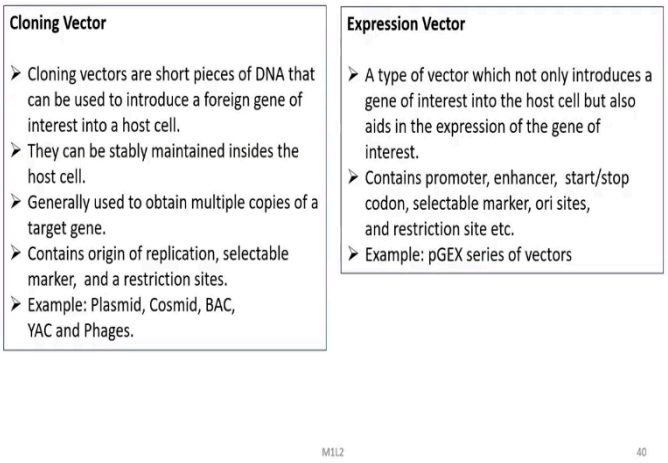


What are multiple cloning sites? The multiple cloning sites or MCS or polylinker sequence is a synthetic DNA segment with a cluster of restriction enzyme specific sites. It is introduced into a gene cloning vector to increase the number of number and types of cut ends as already I have discussed to you and you can see here the various features of these plasmid vector which is having a marker gene, then you have these MCS over multiple cloning site and for example, if you focus on this multiple cloning site, you have three over here and you can see, here it is having more multiple cloning sites.

So, for example, to increase the multiple cloning sites, so this particular plasmid was digested with EcoR1 and then, and as well as Sal I. So, that removes a fragment in between these two sequences. Now, we have another sequence here which has recognition sites for three additional restriction enzymes. Now, when we ligate these into this gap, we get a new plasmid with a more number of multiple cloning sites and you may notice here the origin of replication and the antibiotic resistance gene.

So, through this discussion, you can now understand how we can increase the number of cloning sites in a particular site by the process of cloning itself. So, let us now move on to discuss the two type of vectors; one is the cloning vector and the other is the expression vector. They have many of the features which are common to both of them, but expression vector would have something additional. So, let us start with the cloning vectors.

(Refer Slide Time: 45:16)



So, they are short pieces of DNA that can be used to introduce a foreign gene of interest into a host cell. They can be stably maintained inside the host cell. Generally, used to obtain multiple copies of a target gene and it contains origins of replication, selectable marker and a restriction site or a multiple cloning site and you have various examples for cloning vectors like Plasmid, Cosmid, BAC, YAC and Phages. Expression vectors are a type of vector which not only introduces gene of interest into the host cell, but also aids in the expression of the cell of gene of interest.

So, here the gene will be finally translated into it is a protein product because you have the elements of transcription as well as translation incorporated into the gene construct. So, it contains promoter, enhancer, start stop codon besides the selectable markers, then origin of replication and the restriction cloning site.

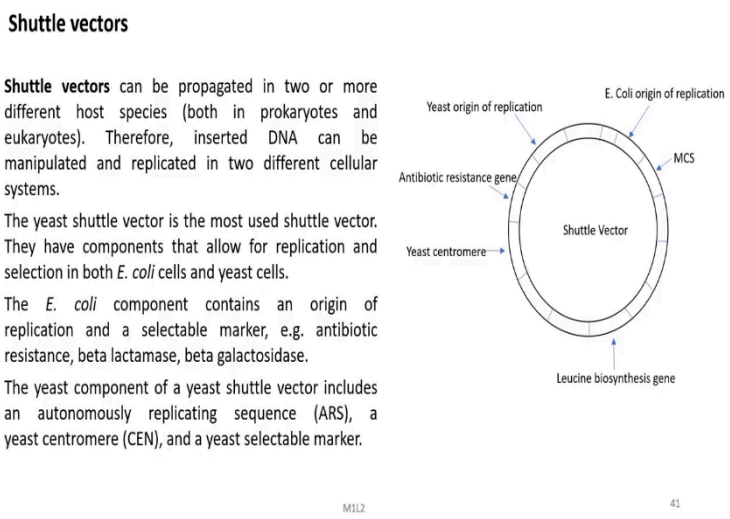
So, this promoter, enhancer, start, stop codon are not available in a cloning vector because their purpose is not for expression of the gene construct into its protein product and therefore, it is unnecessary to have those elements over there. The sole purpose of a cloning vector is to increase the number of the gene construct in very high amounts.

So, example of expression vectors is pGEX series of vectors.

Let us discuss about the shuttle vectors. What are shuttle vectors? Shuttle vectors are the vectors which can be propagated in two or three different host species; both in prokaryotes

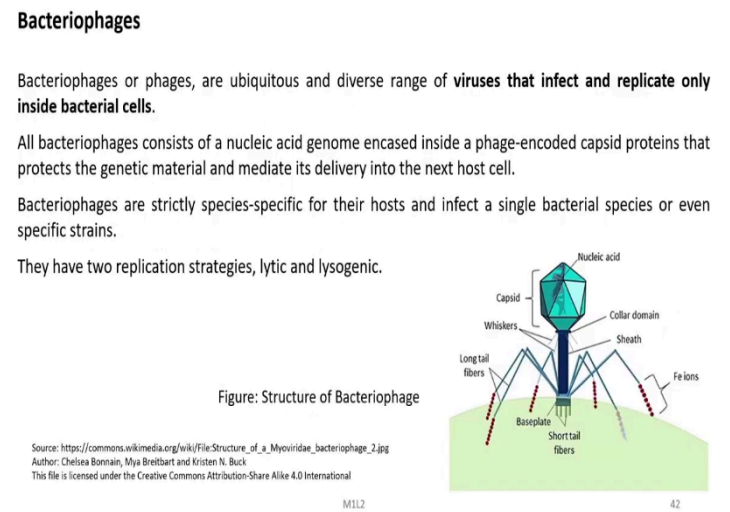
and eukaryotes. Therefore, inserted DNA can be manipulated and replicated in two different cellular systems.

(Refer Slide Time: 47:14)



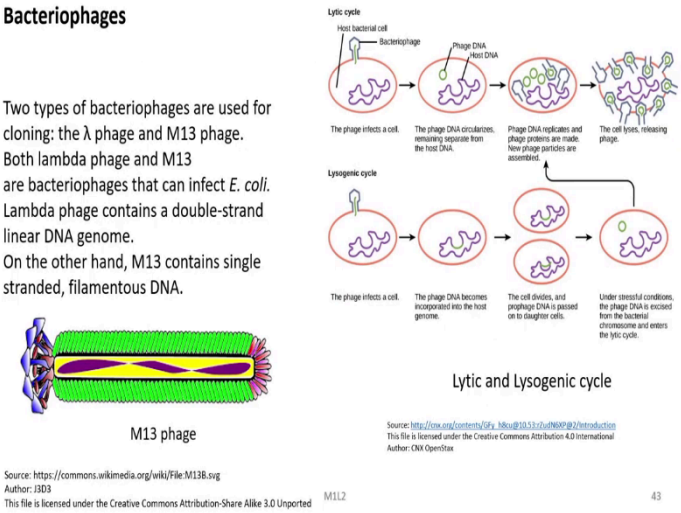
The yeast shuttle vector is the most used shuttle vector. They have components that allow for replication and restriction selection in both E. coli cells and yeast cells. The E. coli component contains an origin of replication and a selectable marker. Example, antibiotic resistance etcetera. The yeast component of the yeast shuttle vector includes an autonomously replicating sequence, a yeast centromere and the yeast selectable marker.

(Refer Slide Time: 47:43)



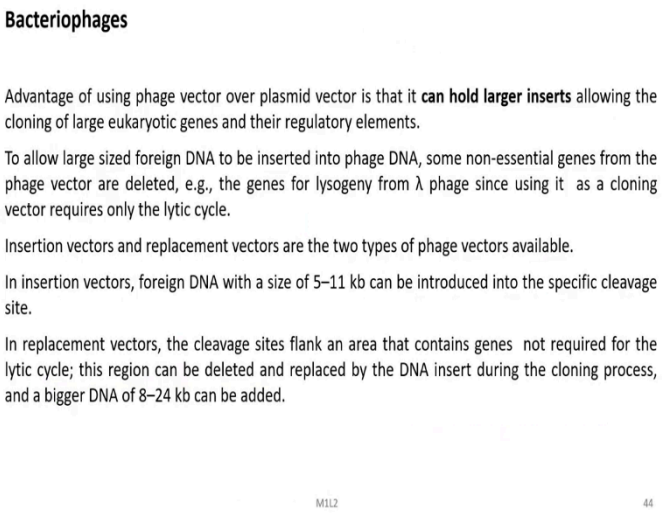
Let us go to the bacteriophages which are ubiquitous and diverse range of viruses that infect and replicate only inside bacterial cells. All bacteriophages consist of a nucleic acid genome encased inside a phage-encoded capsid protein that protects the genetic material and mediates its delivery into the next host cell. Bacteriophages are strictly species specific for their host and in fact, a single bacterial species or specific strains. They have two replication strategies; one is the lytic cycle, another is the lysogenic cycle.

(Refer Slide Time: 48:20)



Two types of bacteriophages are used for cloning; the lambda phage and the M13 phage. Both lambda phage and M13 bacteriophage can infect E. coli. The lambda phage contains a double stranded linear DNA genome, while the M13 contains a single stranded filamentous DNA.

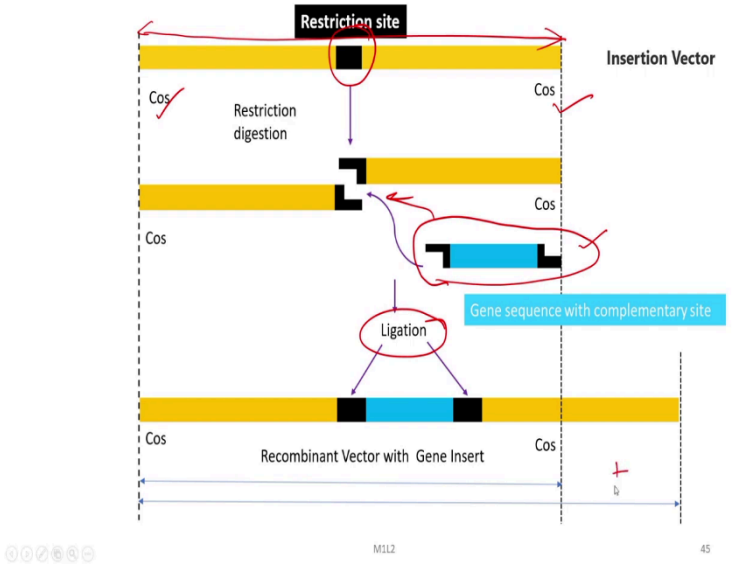
(Refer Slide Time: 48:40)



There are certain advantages of using phage vector over plasmid vector. The phage vectors can hold larger inserts allowing the cloning of large eukaryotic genes and their regulatory elements. To allow large sized foreign DNA to be inserted into phage DNA, some non-essential genes from the phage vector are deleted. Example, the genes for lysogeny from lambda phage since using it as a cloning vector requires only the lytic cycle. Insertion vectors and replacement vectors are the two type of phage vectors which are available and we will have a discussion on them later.

These insertion vectors can carry DNA with a size of 5 to 11 kb and can be introduced into the specific cleavage site. The replacement vectors can carry DNA which are quite bigger and range from 8 to 24 kilo bases.

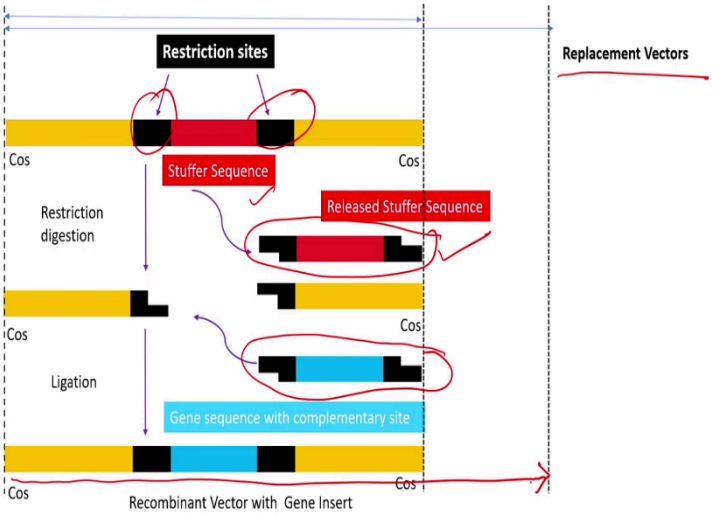
(Refer Slide Time: 49:38)



So, you can see here the diagram of an insertion vector and the size of this vector as can be depicted between these lines say for example, they are ranging from here to here. Now, it has a specific restriction site and then there are two Cos sites on the two ends. Now, when this restriction digestion takes place, it cleaves the vector into two parts and then, if we have a DNA insert with a similar complementary end, it can be inserted into this with the help of ligase molecule.

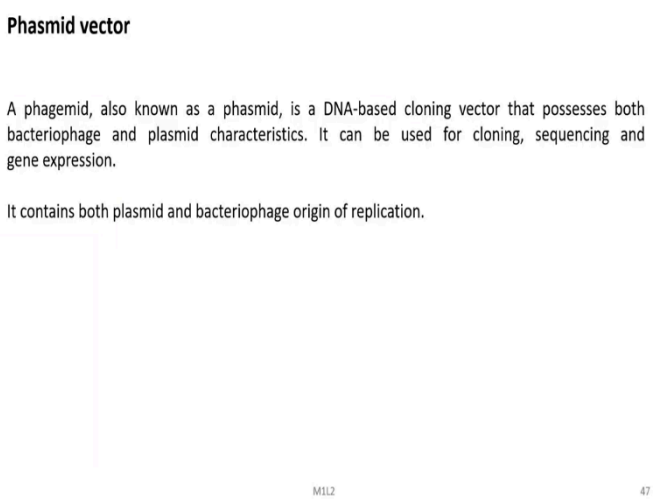
But once we do the insertion of these genetic sequences, the size of the recombinant vector along with the insert will increase because with the addition of the genetic sequence the size will be added up.

(Refer Slide Time: 50:48)



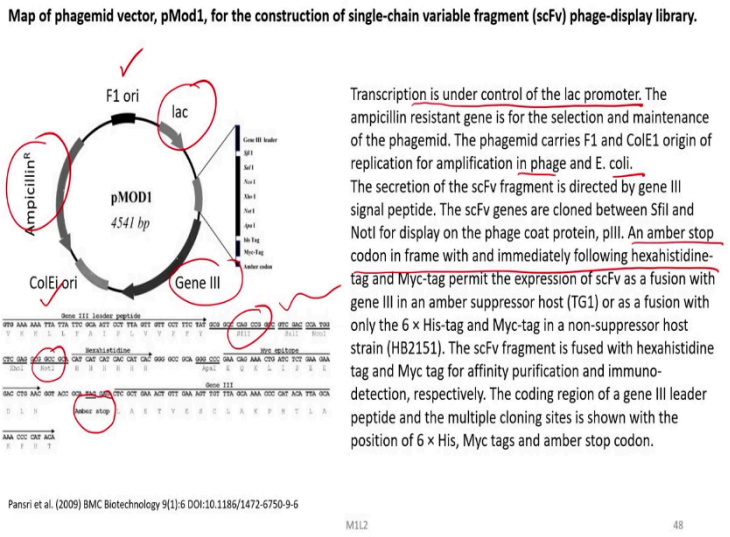
Now, we have other kinds of vectors which are known as replacement vectors. So, here we have two restriction sites and upon restriction digestion, the stuffer sequence which lies in between these two restriction sites are released and into these sequence or gap, we can easily add our gene of interest and this will not add up to the additional length of the DNA recombinant DNA vector sequence. So, therefore, we can load bigger sizes of DNA into the replacement vectors.

(Refer Slide Time: 51:45)



Then, we have other kinds of vectors like the phasmid vectors. A phagemid is also known as a phasmid and it is a DNA-based cloning vector that possesses both bacteriophage and plasmid characteristics. It can be used for cloning, sequencing and gene expression. It contains both plasmid and bacteriophage origin of replication.

(Refer Slide Time: 52:06)

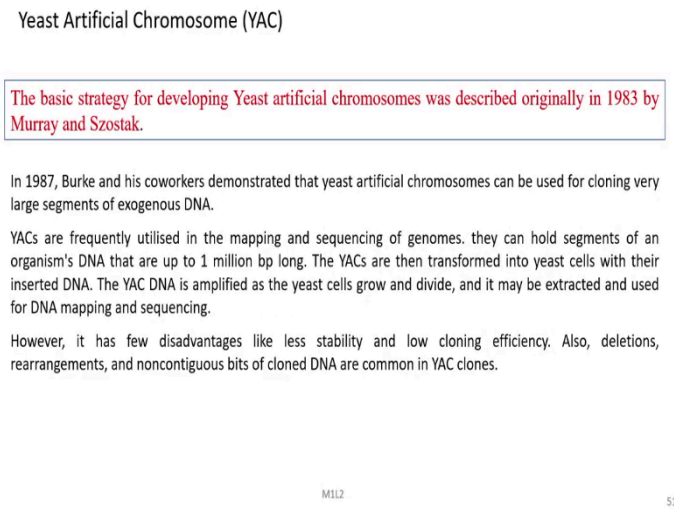


This is a map of a phagemid vector, pModI which has been constructed for single chain variable fragment phage display library. So, here, you can see the transcription is under the control of a lac promoter. Then, there have the ampicillin resistant gene for the selection of the phagemid.

This phagemid carries F1 and ColE 1 origins of replication for application in phage and E. coli respectively. The secretion of the scFv fragment is directed by this particular gene called Gene III signal peptide. The scFv genes are cloned between Sfil and this one the Sfi1 and the not 1 and in between, we clone the scFv sequences.

So, there is an amber stop codon in the frame; here, amber stop codon frame over here. So, these various features make this vector very special and altogether different from the other kinds of vectors that we have discussed till now.

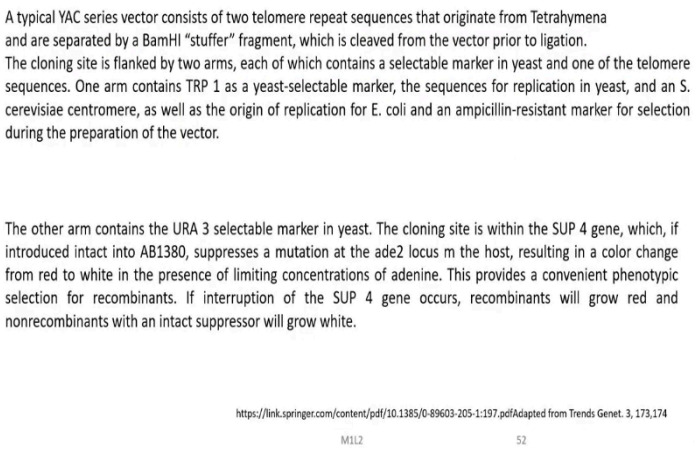
(Refer Slide Time: 53:36)



Let us now discuss about some other kinds of vectors; the yeast artificial chromosome. The basic strategy for developing yeast artificial chromosomes was described originally in 1983 by Murray and Szostak. In 1987, Berg and his coworkers demonstrated that yeast artificial chromosomes can be used for cloning very large segments of exogenous DNA.

YACs are frequently utilized in the mapping and sequencing of genomes. They can hold segments of an organism’s DNA that are up to 1 million base pairs long. The YACs are then transformed into yeast cells with their inserted DNA. The YAC DNA is amplified as the yeast cells grow and divide and it may be extracted and used for DNA mapping and sequencing. However, it has few disadvantages like less stability and low cloning efficiency. Also, deletions, rearrangements and noncontiguous bits of cloned DNA are common in YAC clones.

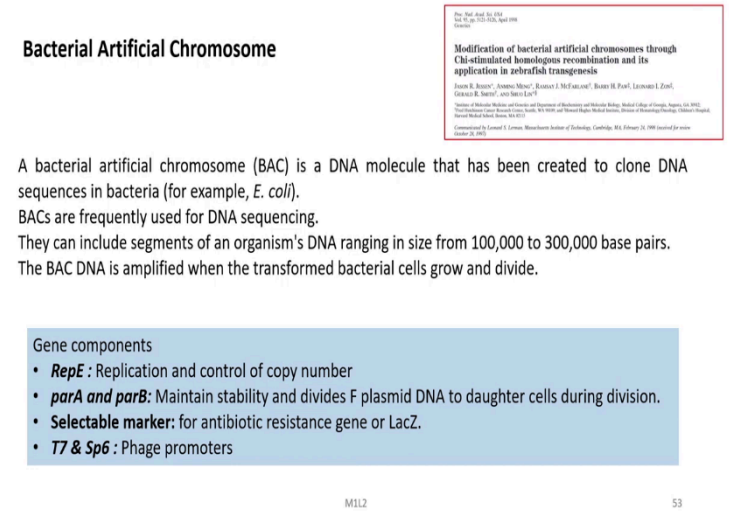
(Refer Slide Time: 54:33)



A typical YAC series vector consists of two telomere repeat sequences that originate from *Tetrahymena* and are separated by a BamHI "stuffer" fragment, which is cleaved from the vector prior to ligation. The cloning site is flanked by two arms, each of which contain a selectable marker in yeast and one of the telomere sequences. One arm contains TRP 1 as a yeast selectable marker, the sequence for replication in yeast and an *S. cerevisiae* centromere, as well as the origin of replication for *E. coli* and an ampicillin-resistance marker for selection during the preparation of the vector.

The other arm contains the URA 3 selectable marker in yeast. The cloning site is within the SUP 4 gene, which if introduced intact into AB1380 suppresses a mutation of the ade2 locus, the host resulting in the host resulting in a color change from red to white in the presence of limiting concentrations of adenine. This provides a convenient phenotypic selection for recombinants. If interruption for the SUP 4 gene occurs, recombinants will grow red and non-recombinants with an intake suppressor will grow white.

(Refer Slide Time: 55:47)



In the same way, we have other kind of artificial chromosomes known as bacterial artificial chromosome. A bacterial artificial chromosome is a DNA molecule which has been created to clone DNA sequences in bacteria. BACs are frequently used for DNA sequencing, they can include segments of an organism's DNA ranging in the size from 100,000 to 300,000 base pairs. The BAC DNA is amplified when the transformed bacterial cells grow and divide. The various gene components in a BAC are the RepE which is replication and which replicate and control the copy number.

Then, parA and parB which maintains stability and divides F-plasmid DNA to daughter cells during division. Then the selectable markers for antibiotic resistance gene or LacZ, then T7 and Sp6 phage promoters.

(Refer Slide Time: 56:42)

Agrobacterium-mediated plant transformation

Agrobacterium species such as *A. tumefaciens* and *A. rhizogenes* infect plants and cause crown gall and hairy-root diseases.

In this process, tumor-inducing (Ti) and root- inducing plasmids play a major role.

Agrobacterium species can transfer a part of their Ti or Ri plasmid, called the transfer-DNA or T-DNA to the host plant cell.

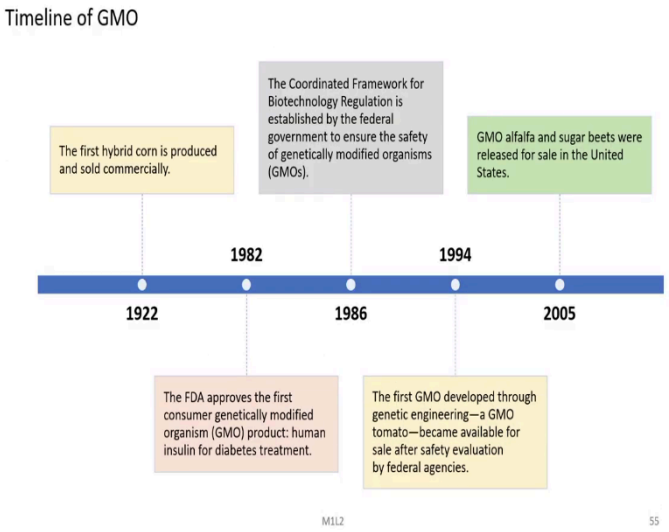
T-DNA carries genes responsible for plant growth hormones. Overproduction of these hormones inside the plant cell causes the formation of tumor or root hair.

Crown Gall Disease

We have a natural genetic engineer which is the agrobacterium. These agrobacterium species like *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes* infect plants and cause crown gall and hairy root diseases. In this process, the tumor inducing Ti and the root inducing plasmids play a major role.

This agrobacterium species can transfer a part of the Ti or Ri plasmid called the transfer DNA or T-DNA to the host plant cells. In the application of these particular plasmids, we re-engineered these plasmids by removing their pathogenic sequences and only retaining the sequences which are useful for mobilizing the insert DNA into the host plant.

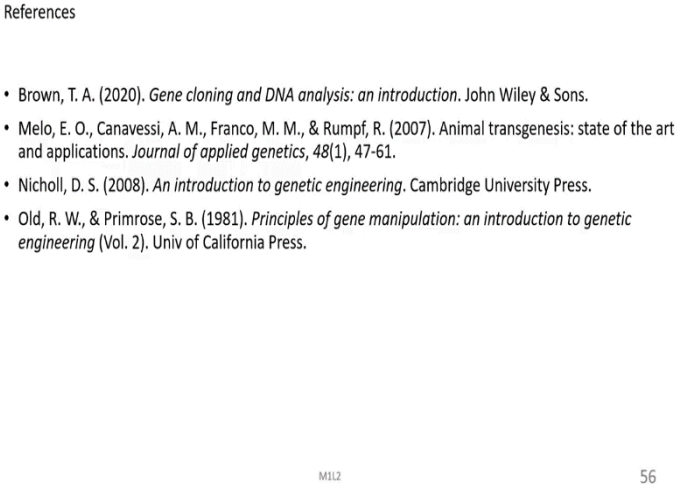
(Refer Slide Time: 57:22)



So, with this, we come to the end of these discussion. This is briefly a timeline of the development of the GMO starting in 1922, when the first hybrid corn was produced and sold commercially. In 1982, the FDA approves the first consumer genetically modified organism which is the human insulin for diabetes treatment.

And the coordinated framework for the biotechnology regulation was established by the Federal government to ensure the safety of genetically modified organism in 1986. In 1984, the first GMO was developed through genetic engineering a GMO tomato and it became available for sale after safety evaluation by Federal agencies. In 2005, GMO alfalfa and sugar beets were released for sale in the United States.

(Refer Slide Time: 58:25)



Thank you.