**Eukaryotic Gene Expression: Basics and Benefits**

**Prof. P. N. Rangarajan**

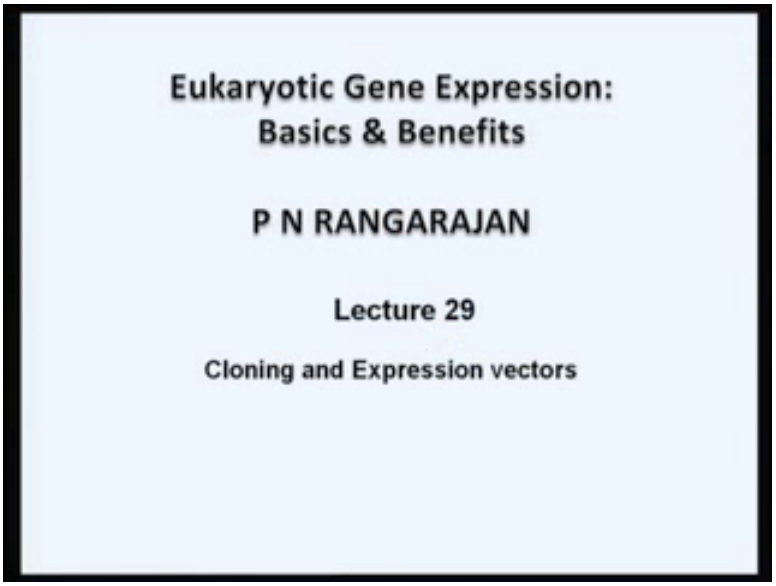
**Department of Biochemistry**

**Indian Institute of Science, Bangalore**

**Lecture No. # 29**

**Cloning and Expression Vectors**

(Refer Slide Time: 00:23)




Welcome to this lecture series on eukaryotic gene expression basics and benefits. Today, we are going to discuss about cloning and expression vectors, because the purpose of this lecture course is not only to explain some of the major concepts in the area of eukaryotic gene expression, and make you understand how gene expression is regulated. We also, time and again, keep emphasizing how this knowledge that we have obtained from this basic research has been, or is being used for the benefit of mankind.

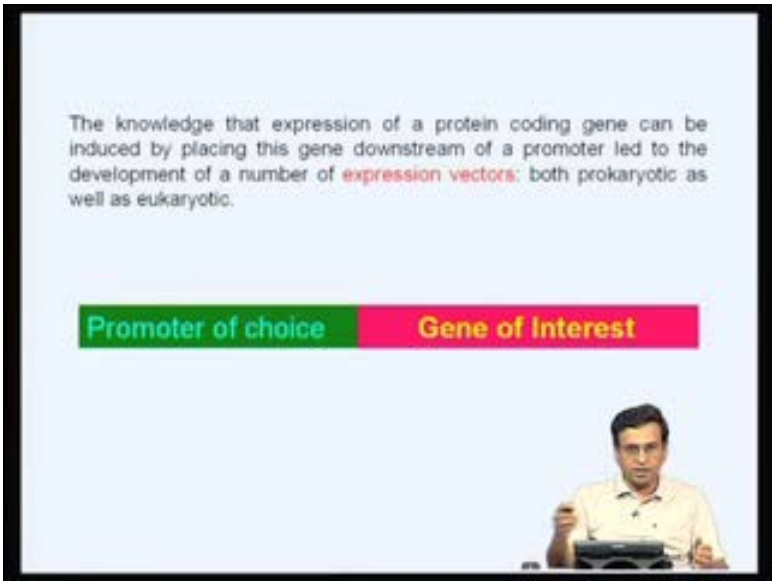
So, one of the very important aspects of regulation of gene expression that has really benefited mankind is our ability to produce recombinant proteins. A number of recombinant proteins like insulin, growth hormone, interferons, many other proteins are now being produced in large amounts, both in microbial as well as mammalian expression systems, and this has largely benefited mankind in a big way.

So, it is very important for us to understand how one designs or one constructs expression vectors; what kind of expression vectors are now being used, and tomorrow, if you want to now express a protein using a particular promoter, what kind of expression vector you would design? Would you express it in a prokaryotic vector? Would you express in eukaryotic vector? All this information is very, very essential.

So, the way I have designed this lecture series is not only you understand some of the basic concepts of eukaryotic gene expression, but time and again, you also apply this knowledge and see how you can, or how this knowledge has been used for the benefit of mankind. So, one of the important aspects in the area of biotechnology, which has really made a huge difference is our ability to express proteins from any species, in any other species.

So, today, we can easily take a gene that codes for a protein from a plant or from a human protein, and you can put it in either in a plant, in E.coli system, or you can express in a yeast vector, or you can express in a plant, and so on and so forth. So, let us now try to understand how these expression vectors, or how this cloning vectors are designed and how you can actually clone genes into these vectors, and sometimes, express these genes, so that you can get your protein of your interest.

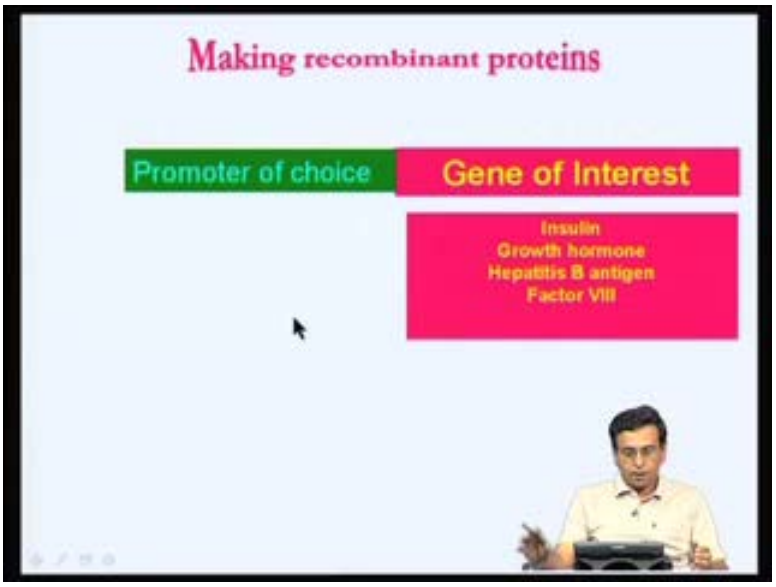
(Refer Slide Time: 02:43)




So, let us try to understand about cloning and expression vectors. So, the knowledge that expression of protein coding gene can be induced by placing this gene downstream of a

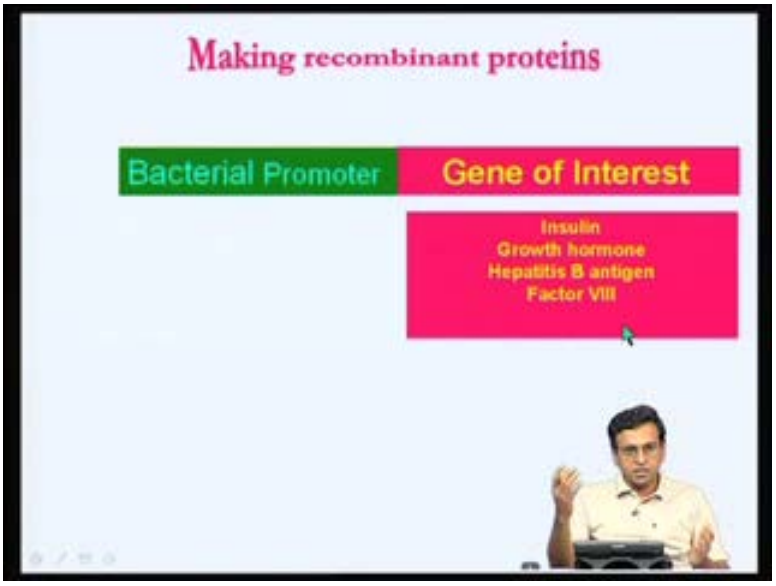
promoter has led to the development of a number of expression vectors, both prokaryotic as well as eukaryotic. So, today, it is possible if we want, for example, any gene that you are interested in, whether this gene can code for insulin or growth hormone or any protein that is of interest to you, you can take this gene, put it downstream of promoter of your choice, and you can make this protein in the particular organism of your choice.

(Refer Slide Time: 03:14)

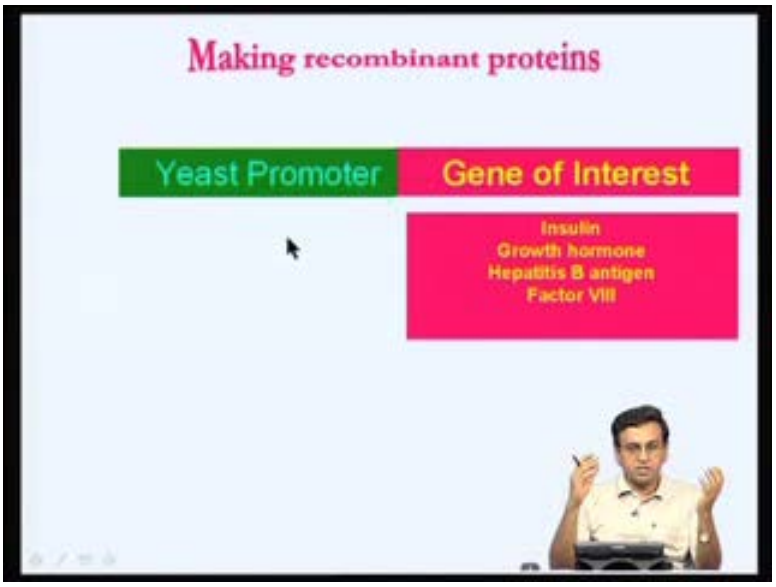



This is explained much better in the next slide. For example, suppose you want to express whether insulin or growth hormone or hepatitis B surface antigen or a clotting factor, you simply take these genes which code for this respective proteins and clone into a promoter of your choice

(Refer Slide Time: 03:29)

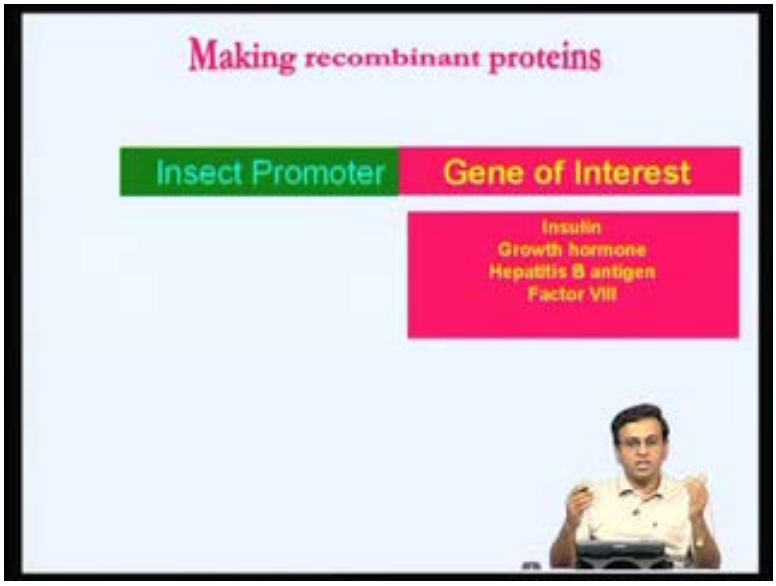



(Refer Slide Time: 03:39)

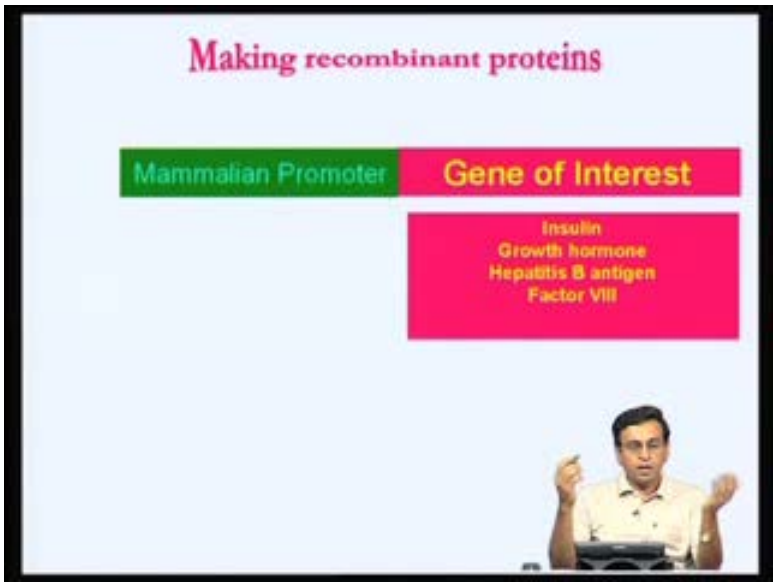



For example, if you want to express this in a bacterial system, you put this gene in front of a bacterial promoter, and you have to construct what is called as a bacterial expression plasmid. On the other hand, if you want to express this gene in yeast cells, you have to put this gene in front of yeast promoter and construct what is called as yeast expression vector.

(Refer Slide Time: 03:48)

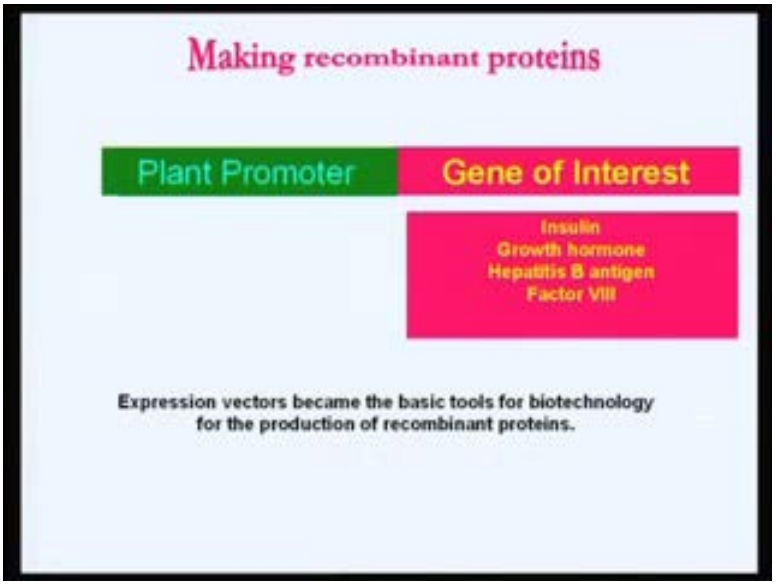



(Refer Slide Time: 03:58)




In the same way, if you want to express your gene in insect cells, you have to express what is called as insect expression vectors, and you have to use a promoter that works in insect cells, and similarly, if you want to express in mammalian cells, including human cells, you have to clone this gene in front of a mammalian promoter, and make what is called as mammalian expression plasmid and introduce them into mammalian cells. And mammalian cells will now start expressing your gene of interest and your protein will be made in mammalian cells.

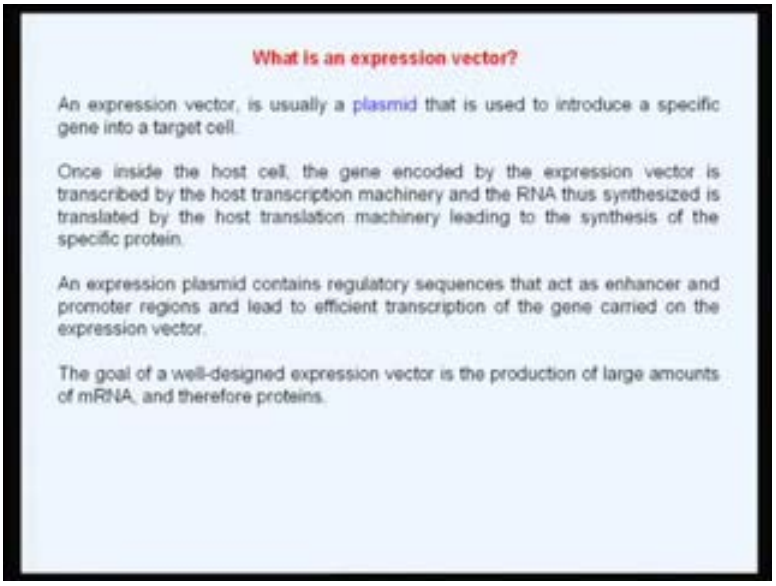
(Refer Slide Time: 04:15)




And last, but not least, you can also now express your gene of interest in plant cells. All that you have to do is you put a plant promoter in front of this gene and make a plant expression plasmid, introduce into the plant cells. Now, plant cells will make your protein of your interest. So, you can see, the idea or the knowledge that promoters, which contain binding for transcription factor sites and RNA polymerase, they can be exploited for you express. The expression of any downstream gene has now let to the development of what are called as a recombinant DNA technology and production of recombinant proteins of your choice.

And it is one of the major areas in the area of biotechnology, where a number of industries and number of companies are expressing a number of therapeutic proteins using this kind of a cloning technology. So, let us spend some time to understand how do you design a cloning vector; how do you design a expression vector; and what kind of expression vectors and cloning vectors are being used. So, expression vectors became the basic tools for biotechnology for the production of recombinant proteins– this is going to be the basics of today's lecture.

(Refer Slide Time: 05:18)




So, what is an expression vector? An expression vector is usually a plasmid that is used to introduce a specific gene into a target cell, and express the protein that is coded for by the gene. So, once inside the host cell, the gene encoded by the expression vector is transcribed by the host transcription machinery; that is, the host transcription factor, and host RNA polymerase will transcribe the gene, and the RNA that is synthesized is then translated by the host translation machinery, leading to the synthesis of a particular protein of your interest.

So, if the gene has to be expressed inside the host cells, you need to contain…, the expression plasmid should contain regulatory sequences that act as a either enhancer and promoter regions, and lead to efficient transcription of genes carried by the expression vector. So, if you want to express a gene of your interest in a particular system, either bacteria or yeast or mammalian system, the regulator region that you have chosen must contain a promoter and powerful enhancers, so that powerful transcription factors can go and bind to this sequences and a large amount of mRNA can be synthesized, which in turn, gets translated into a protein, and you can then make a protein of your interest in large amounts.

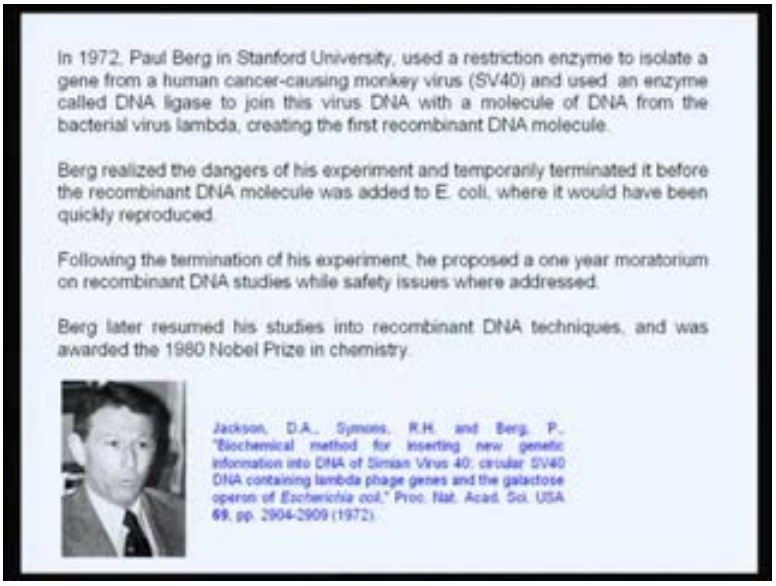
So, the goal of a well-designed expression vector is, therefore, production of large amounts of mRNA, and therefore, large amounts of proteins. So, this is what is the rationale behind constructing an expression vector

(Refer Slide Time: 06:37)


So, the design and development of expression vectors and their use in biotechnology for the benefit of mankind is very closely into discoveries in two major areas, namely, recombinant DNA technology, also known as genetic engineering. See, a lot of development took place in the period between 1970s to 1990s. Advances in cell biology, recombinant DNA technology, cloning technology, and so on so forth, and it is these technologies, together with our knowledge that promoters and enhancers are important for the expression of genes, has led to the development of an entire area of biotechnology, leading to a new field, wherein you can express any protein of your interest and make what are called as the recombinant proteins.

So, let us spend some time to understand what are the important advances that took place in the area of recombinant DNA technology or genetic engineering, especially in the early 70s and late 70s, and how these advances has led to the development of expression vectors and a billion dollar biotech industry, leading to the expression of recombinant proteins.

(Refer Slide Time: 07:38)




Now, in 1972 a researcher known as Paul Berg in Stanford University, California, used certain enzymes called restriction enzymes. These are enzymes, which can specifically cut DNA by recognizing the specific sequences. Today, we know there are number of such restriction enzymes. For example, you have an enzyme called EcoR1, which cuts a specific sequence called GAATTC. So, if you have a DNA and if you have the sequence GAATTC, and if you add EcoRI to this DNA preparation, it will cut this DNA wherever GAATTC is there.

So, like that, you have a number of restriction enzymes; now the number is now goes in thousands. These restriction enzymes are usually present in bacterial cells; these enzymes have been purified and recombinant enzymes are now available. So, using these enzymes, you can precisely cut DNA at specific regions. So, what Paul Berg did, in 1972, is he used such a restriction enzyme and isolated a gene from human cancer causing monkey virus, called as the simian virus 40 or SV40, and used the enzyme called DNA ligase to join this virus DNA with a molecule of DNA from a bacteriophage or bacterial virus called lambda.

So, he cut a monkey virus DNA with restriction enzyme, and took this DNA and joined to a DNA, which is present in the lambda phage or the lambda DNA, by using DNA ligase. So, restrictions and cuts the DNA. You take this DNA and you can ligate or attach to another DNA molecule from the bacteriophage region using a ligase. This is the first

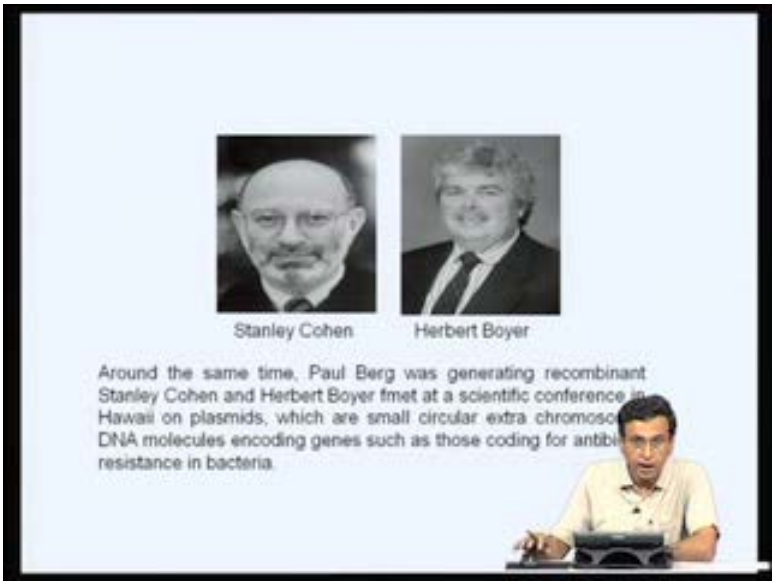
example of creation of a recombinant DNA. Today, recombinant DNA technology or genetic engineering is an household name.

But this Paul Berg who, actually, demonstrated for the first time that using restriction enzymes you can cut a DNA from a mammalian virus and then put it into a or ligate it to or attach to a DNA from a bacteriophage. Berg realized that this experiment of making chimera, that is, you can take DNA from one species and ligate to the DNA of another species, may have very tremendous advantages, and if it is not used properly, can lead to lot of disadvantages or dangers to the mankind. And therefore, he suggested that the regulatory agencies should come forth and design proper guidelines, so that recombinant DNA technology– whoever wants to use this recombinant technology– follows these guidelines, and some kind of a overseeing body is there to make sure that this recombinant technology is not used improperly.

So, he did his experiment and he proposed that for at least one another one year, nobody does these recombinant DNA experiments till proper guidelines are framed. Then, he later resumed, and his effort– this paper– he was published in PNAS in 1972– biochemical method for inserting new genetic information into DNA of simian virus 40 circular SV40 DNA containing lambda phage genes and the galactose of operon of E.coli.

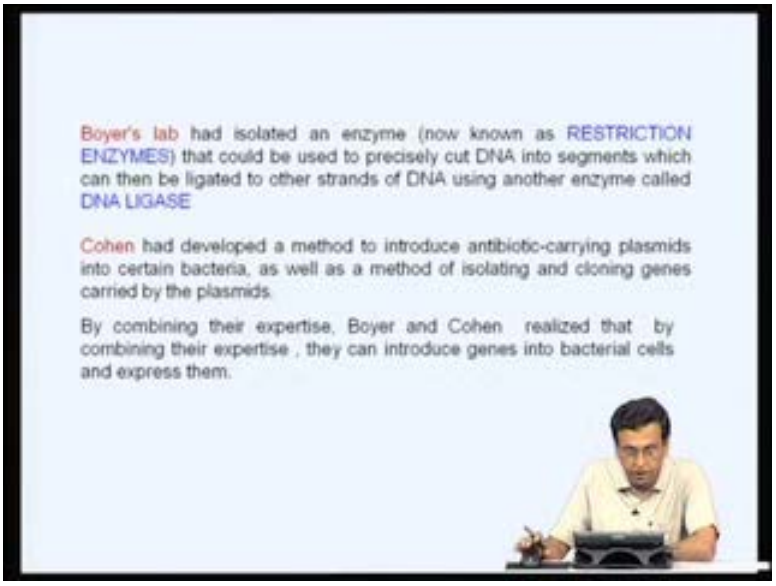
So, the landmark paper, which actually discusses the generation of a first recombinant DNA molecule, and for which Paul Berg was awarded Nobel prize in the year 1980. So, Paul Berg got the credit for generating the first recombinant molecule, where he took a DNA from a monkey virus and inserted into a lambda or a bacteriophage virus, and demonstrated that it is possible to stitch two different kinds of DNA together using recombinant technology.

(Refer Slide Time: 11:02)




Now, of about the same time Paul Berg was doing these experiments, two researchers– Stanley Cohen and Herbert Boyer– one person was in University of California and another person is in Stanford university– they, actually, were also interested in generating this kind of recombinant DNA molecules, and they actually met in a scientific conference in Hawaii, which was discussing on plasmids. Plasmids are nothing but circular extra chromosomal DNA molecules, which are present in bacterial cells. We will discuss little bit later exactly what plasmids are. So, this conference was actually discussing about plasmids, because those were the times these plasmids are actually being generated, and researchers were working on this plasmids in a big way.

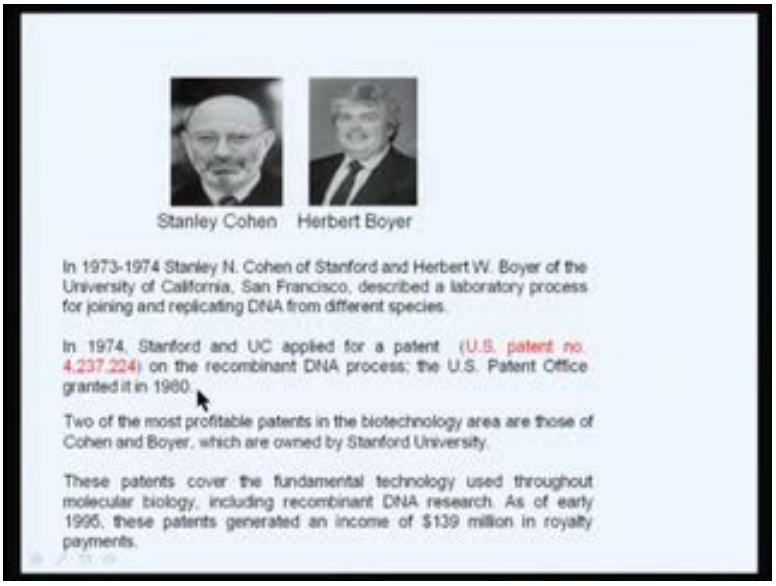
(Refer Slide Time: 11:50)




So, they were attending a conference, which was primarily discussing about plasmids and how these plasmids confer antibiotic resistance in bacteria, and they met over a cup of coffee. And Boyer's lab were, actually, had isolated enzyme known as the restriction enzymes, which actually, Paul Berg had actually used to generate a recombinant DNA, and these restriction enzymes, as I told earlier, can be used precisely to cut DNA into segments, and then using enzymes called DNA ligases, you can ligate into another DNA molecule. This is what Paul Berg also did. So, Boyer's lab is the one who actually isolated the first restriction enzyme.

Stanley Cohen had actually developed a method to introduce antibiotic-carrying plasmids into certain bacteria, as well as a method of isolating and cloning gene carried by the plasmids. So, here is one person who discovered enzymes that can precisely cut DNA, and here is another person who was trying to characterize plasmids and how to introduce plasmids into E.coli cells. They soon realized, while discussing over a cup of coffee, that if they can combine their expertise, together, they can actually generate lot of money because this technology has tremendous potential. So, they can introduce any gene into bacterial cells, and you can express these genes and this can have tremendous implications.

(Refer Slide Time: 12:53)



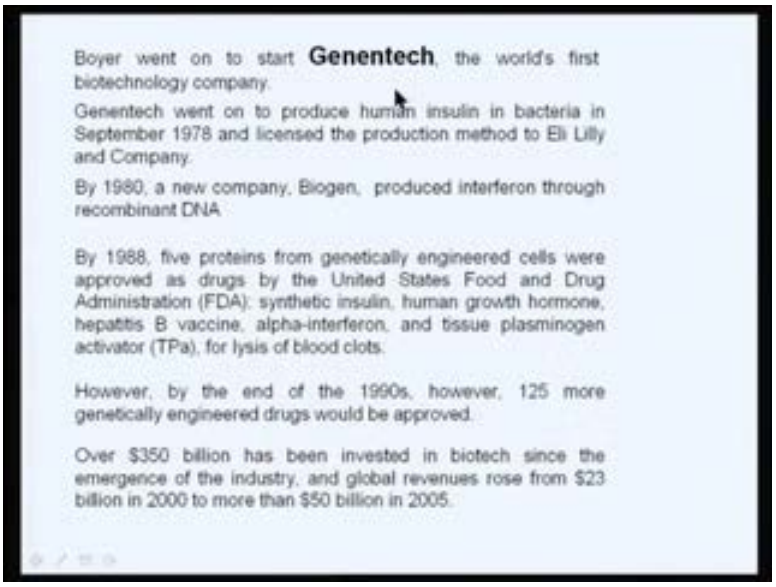

So, what Stanley Cohen and Herbert Boyer did is they published the paper between 1973 and 1974. Before publishing the paper, they also filed two patents– one in 1974 and 1975, about how this process of joining two DNA molecules can have tremendous importance, and how proteins can be made– synthesized– in organisms like E.coli or other prokaryotic species. So, this US patent number 4237224 is one of the landmark patents in the area of biotechnology. They filed a patent around 1973 and 1974, and the US patent office actually granted this patent in 1980.

Now, as I speak today, today, if you ask me which of the most profitable biotechnology patents in the area of biotechnology, the Cohen and Boyer patent is one of them. You can see the two patents of generating recombinant DNA, which was filed by Stanford and University of California, called as the Cohen-Boyer patent. These patents cover the fundamental technology used throughout molecular biology, including recombinant DNA research, and from 1980 till 1995 about 15 years time, these patents generated an income of 139 million dollars.

So, whoever generates the DNA molecule, and which has been used for making a recombinant protein, has to pay royalties to University of California and Stanford, because they own the intellectual property for this. So, you can see how this knowledge that you can use restriction enzymes to create a recombinant DNA molecule, and these recombinant DNA molecules are in the form of plasmids can be introduced into bacterial

cells and you can make a protein of your interest, has generated a huge amount of money, and has now led to a billion dollar biotechnology industry. So, expression of genes using appropriate promoters and inserting these gene into appropriate plasmid vectors and making bacteria make proteins a large amounts has had a huge impact in the area biotechnology.

(Refer Slide Time: 15:01)




Boyer went on to found what is called as a biotech company called Genentech. This is the world's first biotechnology company. Today, we have a number of such companies, so the credit for finding the first biotechnology company goes to Boyer, along with another venture capitalist, and Genentech then went on to produce human insulin bacteria. So, using this recombinant DNA technology, what Boyer did is they chemically synthesized the insulin gene and put this insulin gene in a bacterial expression plasmid, and demonstrated that bacterial cells can now make insulin.

So, this is the first demonstration of a production of a recombinant protein using recombinant DNA technology. So, Genentech, in September 1978, actually demonstrated that a human protein can be produced in bacteria, and they licensed this technology to another biotech company called Eli Lilly, and by 1980, this marked the start of the biotech industry. So, you can see, it is the Stanley Cohen and Boyer who are actually responsible for the creation of a huge biotech industry, that is now running to billions and billions of dollars, now.

So, Genentech was the first biotech company to be formed, followed by Biogen in 1980, which actually produced another very important molecule called the interferon using the same recombinant DNA technology, and by 1988, within about 8 years time, five proteins were produced using that real cells, and were actually approved by the United States Food and Drug Administration for use in humans. These include synthetic insulin, human growth hormone, hepatitis B vaccine, alpha interferon, and a blood clotting protein called a protein, which is involved in resolving blood clots, called tissue plasminogen activator.

So, five important therapeutic proteins were actually synthesized using this recombinant DNA technology patented by Stanley Cohen and Herbert Boyer. So, by the end of 90s, at least 125 more genetically engineered drugs were approved. So, you can see what was the impact of this demonstration of a recombinant DNA technology by Cohen and Boyer. Today, over 350 billion dollars has been invested in biotech industry since the emergence of this industry, starting from Genentech, and global revenues rose from 23 billion dollars in 2000 to more than 50 billion dollars in 2005.

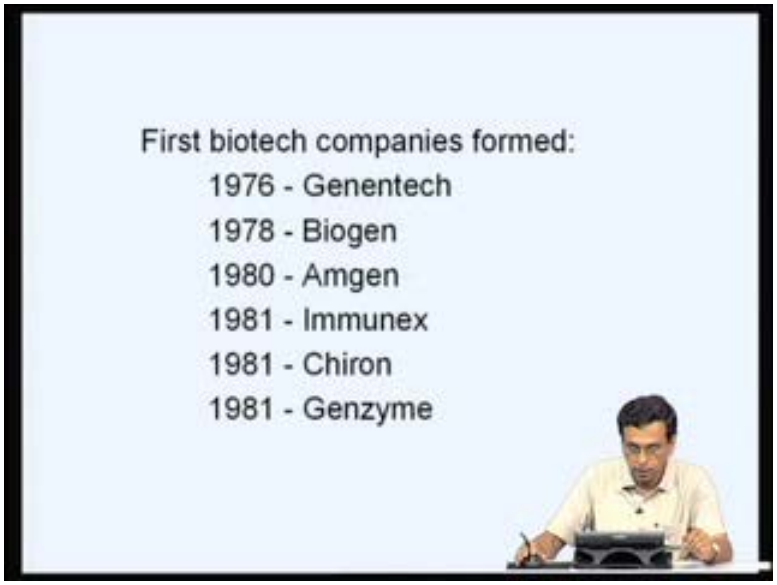
So, you can see, the roots or the fundamental– the foundation for the biotech industry that is prospering today was actually laid by Stanley Cohen and Herbert Boyer way back in 1970s, when they created the first recombinant DNA molecule and one of them started to establish the first biotech company– Genentech.

(Refer Slide Time: 17:43)

So, the growth of recombinant DNA technology, in addition to these discoveries, another important discovery also made sure the commercialization of this fundamental knowledge became possible. That is another important ruling by US Supreme Court, which actually, in 1980, ruled that another Indian who is residing in United States at the time– Anand Chakrabarty– was granted a patent, when he was working in a company called General Electric, that you can actually patent a genetically modified living organism.

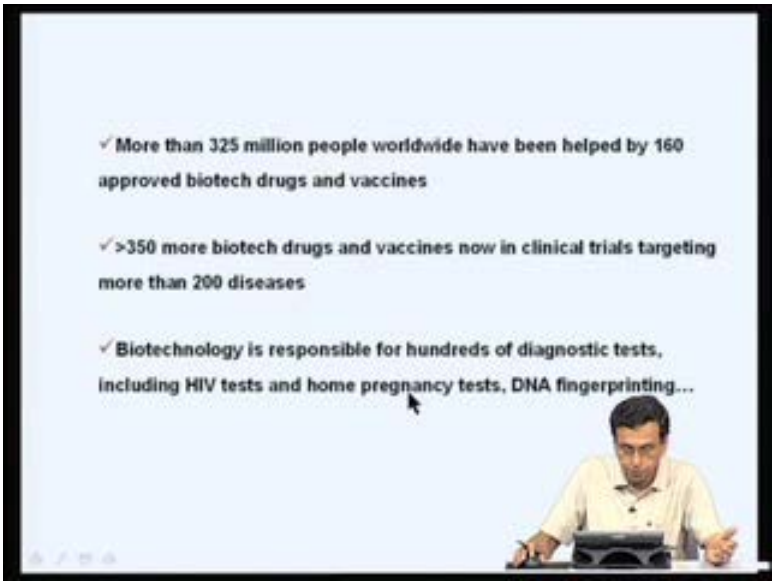
So, the patenting of creation of recombinant DNA molecules and the rule– US Supreme Court ruling– that you can actually patent genetically modified microorganisms, these were the two very important event that took place in the early 70s, and between 1970s and 1980s, that actually are responsible for the growth of a billion dollar biotechnology industry today.

(Refer Slide Time: 18:27)




So, starting form Genentech in 1976, we had a number of other biotech companies which were formed in United States– Biogen, Amgen, Immunex, Chiron, Genzyme, and so on and so forth, and all these companies expanded this knowledge and made a number of recombinant proteins that is being used by human beings all through.

(Refer Slide Time: 18:45)




So, as I speak today, more than 325 million people worldwide have been helped by 160 approved biotech drugs and vaccines. So, you can see, this technology has led to exports of so many recombinant proteins, and so many recombinant vaccines are being used or made by this recombinant DNA technology. 350 more biotech drugs and vaccines are now in various stages of clinical trials that, if many of them are successful, you can, actually, can cure a number of diseases, which are at least about 200 of them.

And biotechnology is, today, responsible for hundreds of diagnostic tests– HIV test, pregnancy test, DNA fingerprinting, and so on and so forth. So, the recombinant DNA production or recombinant protein production, which was started with Herbert and Boyer, is now prospering as a very, very biotech industry today.

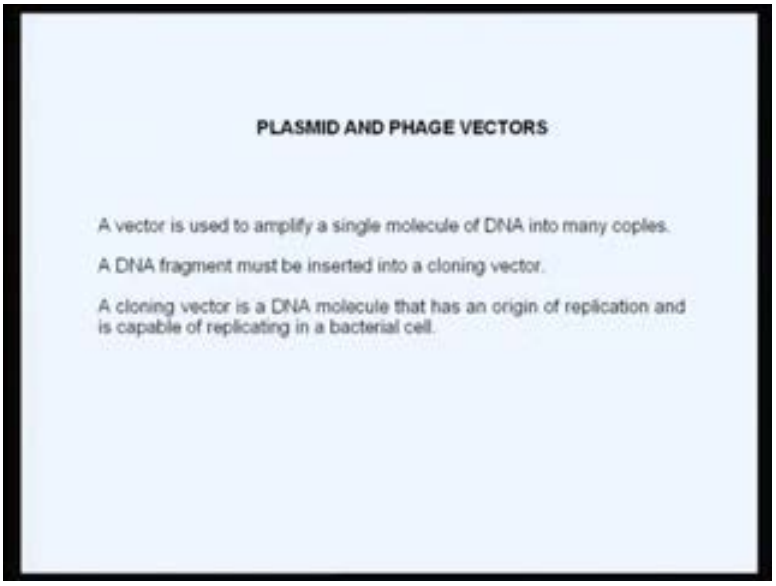
(Refer Slide Time: 19:31)


So, having said how important it was to demonstrate or express clone genes into bacterial cells, express these genes successfully, let us now spend some time to understand how does one design vectors for cloning genes, as well as for making recombinant proteins? So, I gave this introduction so that you understand the importance of this knowledge of generating or making recombinant proteins. So, let us now try to understand some signs of it– how exactly one would go out and construct a cloning vector; how would you go out and construct a expression vector, which is actually the basis for the making a recombinant proteins in a number of organisms.

So, I am going to start explaining to you with what are called as plasmid vectors and phage vectors– these are actually prokaryotic vectors. Although this course is on eukaryotic gene expression, because the entire protein expression system started with prokaryotic vectors, and a number of eukaryotic proteins have been expressed using prokaryotic expression vectors, you must first understand how prokaryotic vectors were designed.

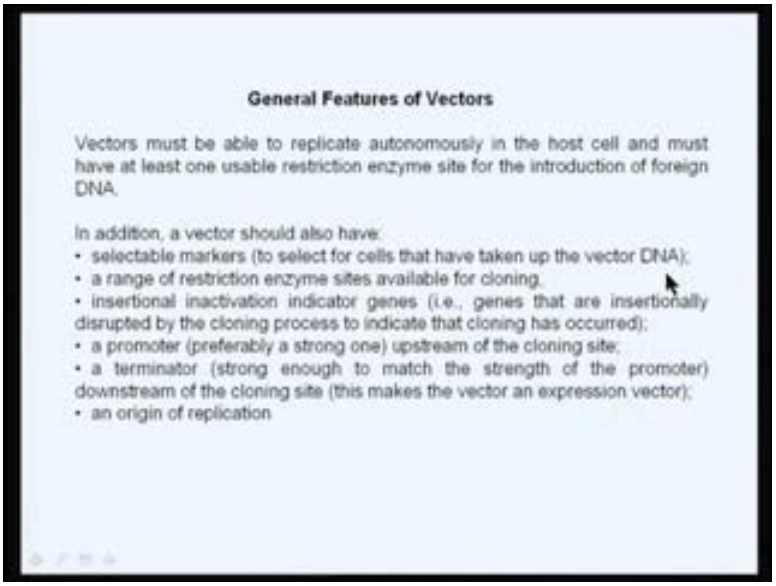
And it is the design and development of the prokaryotic vectors which then paved the way for the development of eukaryotic vectors. So, in this class, let us spend some time to understand the history or the development of various vectors that were actually used in prokaryotes, and in the next class, we will discuss how this knowledge was used for the development of various eukaryotic vectors.

(Refer Slide Time: 20:57)




So, what is a vector? A vector is used to amplify a single molecule of DNA into many copies, and then a DNA map fragment is actually inserted into this plasmid vector, and in addition to this recombinant DNA molecule, this cloning vector should also have what is called as an origin of replication, so that once you introduce in the bacterial cells, these vectors should be able to replicate, so they can make large amounts of these plasmid vectors. So, these are, basically, what is called as a plasmid vector.

(Refer Slide Time: 21:24)



So, the general features of these vectors– suppose you want to now clone a gene into a vector, what kind of a vector you would use. A vector must be able to replicate autonomously in a host cell. If you want to make a bacterial cloning vector, that vector should be able to replicate autonomously bacterial vectors; or if you want to put it in yeast cells, it should be able to express or it should be able to replicate in yeast cells; or mammalian cells should have a mammalian origin of replication, so you can replicate in mammalian cells.

So, vectors must be able to replicate autonomously in the host cell, and must have suitable restriction enzyme sites for the introduction of foreign DNA, because you can precisely cut the vector at a specific site and introduce your gene of interest into those sites. So, the presence of origin of replication, the presence of convenient restriction sites is a must for a vector. In addition, the vector should also have a selectable marker. Usually, genes coding for antibiotic resistance are used as selectable markers, because you have to distinguish cells which have taken up the DNA and cells which have not taken up the DNA.

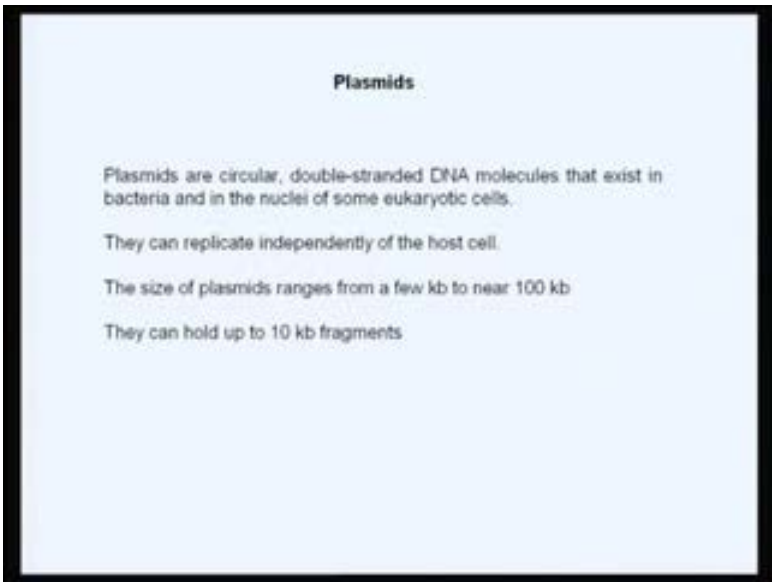
Remember, this procedure called as a transformation of bacterial cells is a very, very inefficient process, and only 1 in 1000 or 1 in 100,000, or 1 in a million cells actually take up this DNA, depending upon the efficiency of the transmission procedure, so majority of the cells do not take up your DNA.

So, you need to have a mechanism by which you have distinguished cells which have taken up the DNA and those which have not taken up the DNA. That is usually done using appropriate selectable markers. It should also have the restriction enzyme sites, so that you can clone your gene of your interest. It should also have insertional inactivation indicator genes, that is genes that are insertionally disrupted by the cloning process, to indicate that cloning has occurred.

It should also have a promoter. Usually, if you want to interest in expressing protein of your interest, it should have a strong promoter upstream of the cloning site, so that when your clone your gene in the multiple cloning site, the upstream promoter, the transcription factors, and RNA polymerase should bind to the promoter and transcribe the gene of your interest.

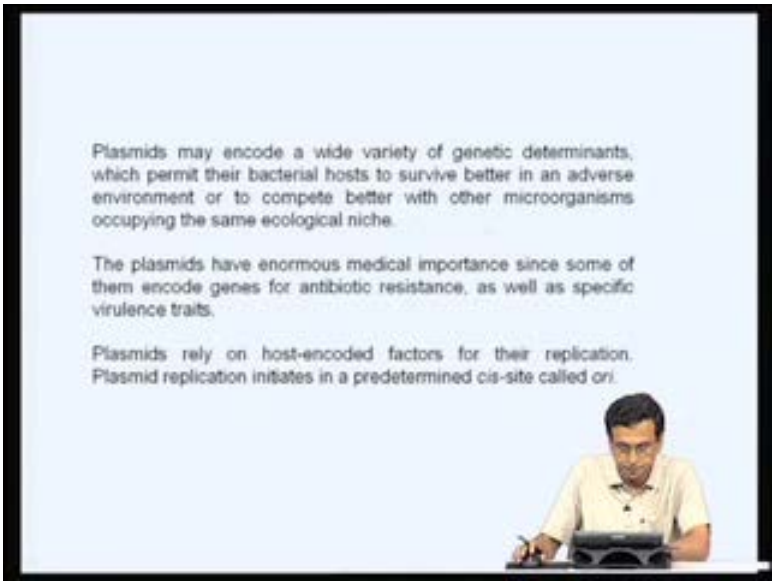
You should also have a terminator at the end of the gene, so that after transcription, the transcription is terminated. And of course, as I said, it should have origin of replication, so that when you put it in organism, you can, organisms can make a multiple copies of this vector. So, you can make large amounts of this vector. So, these are some of the general features of a what is called as a vector.

(Refer Slide Time: 23:54)




Plasmids, as I already mentioned, plasmids are nothing but circular double-stranded DNA molecules that exist in bacteria and in the nuclei of some the eukaryotic cells, and these cells also have some plasmids; they usually replicate independently of the host cell. They are called as the autonomous replicating elements. The size of the plasmids vary from few kb to 100 kb. These are called as mega plasmids. Bacteria like Pseudomonas have many plasmids called mega plasmids ranging up to 100 kb, and these have, actually, genes, which can metabolize number of exotic compounds, and plasmids usually can DNA up to 10 kilo bases can be easily inserted into the plasmids.

(Refer Slide Time: 24:34)



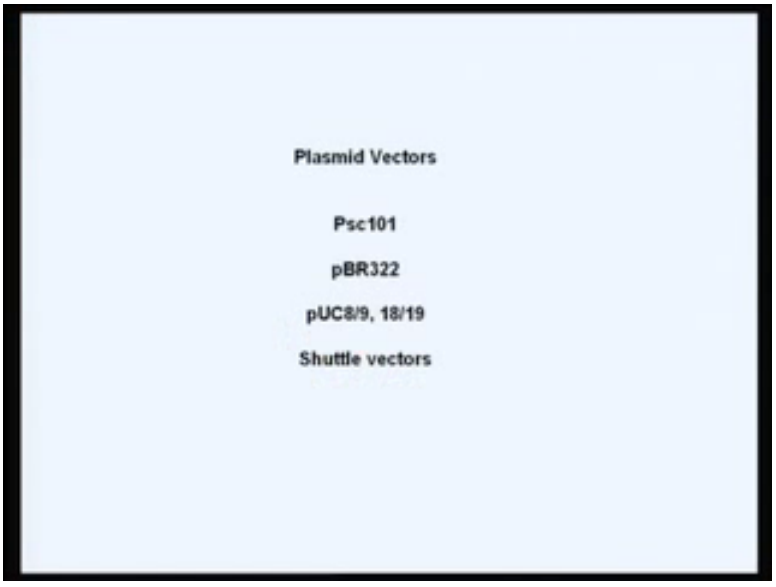

Plasmids may encode a wide variety of genetic determinants, which permit the bacterial hosts to survive better in an adverse environment, or to compete that better with other microorganisms occupying the same ecological niche. So, bacteria which have plasmids have several advantages over those which do not have plasmids– one of the most important genes that these plasmids carry are genes coding for antibiotic resistance. So, when bacteria have this antibiotic resistance on plasmids, they are resistant to antibiotics, whereas those bacteria which do not have this kind of a gene, they are susceptible by bacteria.

In fact, today, one of the major problems in the area of biomedical research is the emergence of what are called as a drug resistance bacteria, and plasmids play a very, very important role in this. So, plasmids have enormous medical importance, since some of them encode antibiotic resistance as well as specific virulence traits. So, many genes, which are responsible for virulence of the certain pathogenic bacteria, are also encoded by the plasmids.

Plasmids rely on host-encoded factors for their replication, and plasmid replication initiates at a predetermined site called ori or origin of replication. So, the plasmids should contain a region called as a ori or the origin of replication, from which the origin of replication can start. It should have convenient restriction sites; it should have an

appropriate promoter for your expression of gene; it should have certain selectable markers, usually antibiotic resistance genes, for selection.

(Refer Slide Time: 26:03)

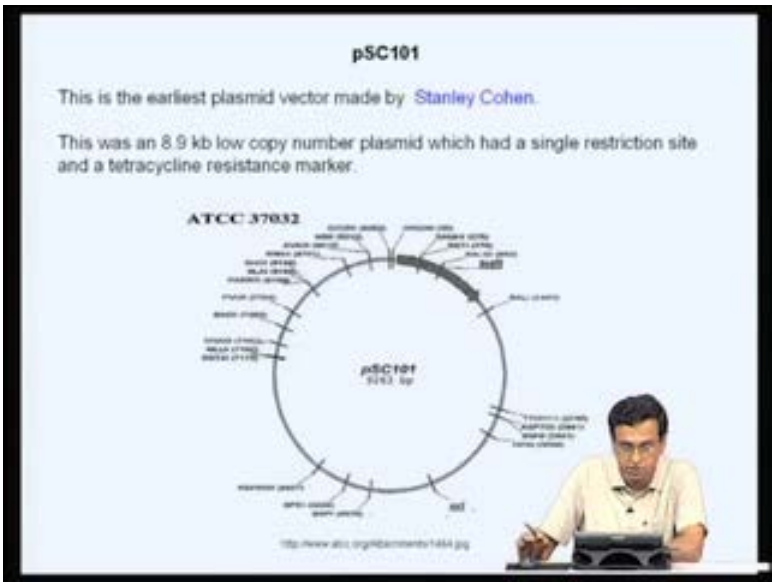



Now, let us see what kind of plasmid vectors people have been using for last years. I just listed some of the plasmid vectors here: Psc101, pBR322, pUC series 8, 9, 18, 19, and so on so forth, and something called as shuttle vectors. So, let us discuss some of these vectors and see what are these vectors, and how these vectors are actually used for making, cloning genes, and expressing genes.

(Refer Slide Time: 26:21)

So, I am going to discuss the first generation plasmid vectors. The reason why I told you the story of Stanley Cohen Herbert and Boyer is the plasmid vector, which Stanley Cohen designed, is actually belongs to the first generation plasmid vector, which was, actually, responsible for generating a huge amount of money, and led to the birth of the first biotech company and a billion dollar biotech industry.

(Refer Slide Time: 26:43)



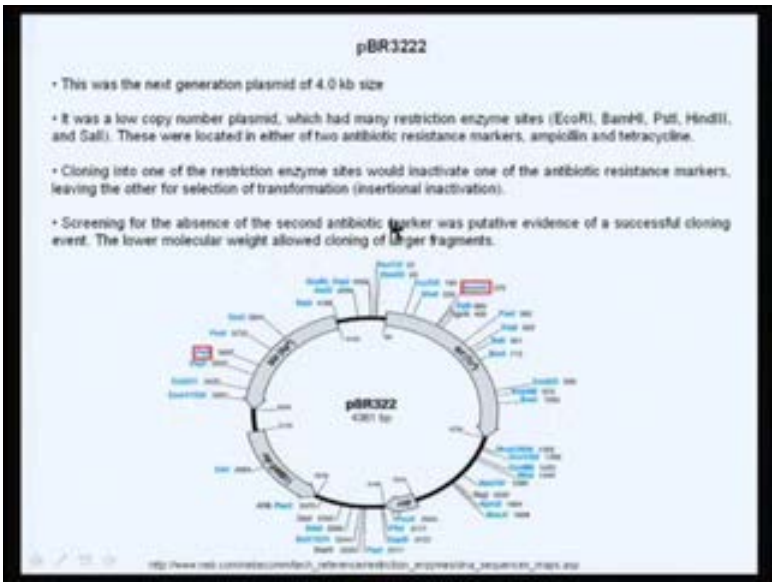

This is the plasmid which Stanley Cohen actually made, and which he patented and made a huge amount of money. That is why this called as pSC101– S is stands for Stanley Cohen. Now, this plasmid vector, as you can see, it is loaded from an ATCC web site; he actually deposited this vector in American Type Culture Collection. Anybody can go and buy this vector, any time you want, from the web site here.

So, this is the vector, here, and as you can see here, it contains an antibiotic resistant gene called tetR. So, if E.coli contains this plasmid, they became resistant to tetracycline. It also has a site for a restriction enzyme called as BamHI, and many other restriction sites. So, if you now cut this plasmid with BamHI, the antibiotic resistant gene is disrupted and you can now clone another gene, which contained BamHI ends, into this, and such bacteria carrying this plasmid now cannot grow on tetracycline, whereas bacteria in which the foreign genes is not inserted can, happily, grow on tetracycline.

So, you can see, you can easily distinguish between two recombinant bacteria– bacteria which contain the unmodified plasmids and bacteria containing a recombinant plasmid,

in which a foreign gene has been inserted into this antibiotic system gene. So, by based on the sensitivity to tetracycline, you can distinguish organisms which contain a recombinant plasmid, organisms which contain only the pSC101. This is the plasmid Stanley Cohen actually designated to demonstrate that it is possible to generate the recombinant DNA molecules and propagate them in bacteria.

(Refer Slide Time: 28:22)




Following this pSC101, a number of second generation vectors were, actually, developed. One of them, which was very popular in the 1980s and 1990s, is a vector called pBR322. This plasmid is about 4-kilo base in size. It is a low copy number plasmid. Plasmids are of two types: one is called as the low copy number, another, high copy number. There are some plasmids, which are present in more than 100 to 200 copies per cell. These are called as high copy plasmid numbers, whereas there are other plasmids which are present only in ones and tens, and this is called as low copy number plasmids.

It also had restriction sites for enzymes like EcoRI, BamHI, PsIII, HindIII, etcetera, and they were located on two antibiotic resistance markers– ampicillin and tetracycline. pSC101, which Herbert Boyer used, Stanley Cohen used, had only one antibiotic resistance marker, that is tetracycline. pBR322 has two antibiotic resistance markers– one for resistance for tetracycline, and another for ampicillin. So, you can see, you can actually clone a gene into the PstI site here, which will disturb the ampicillin resistant

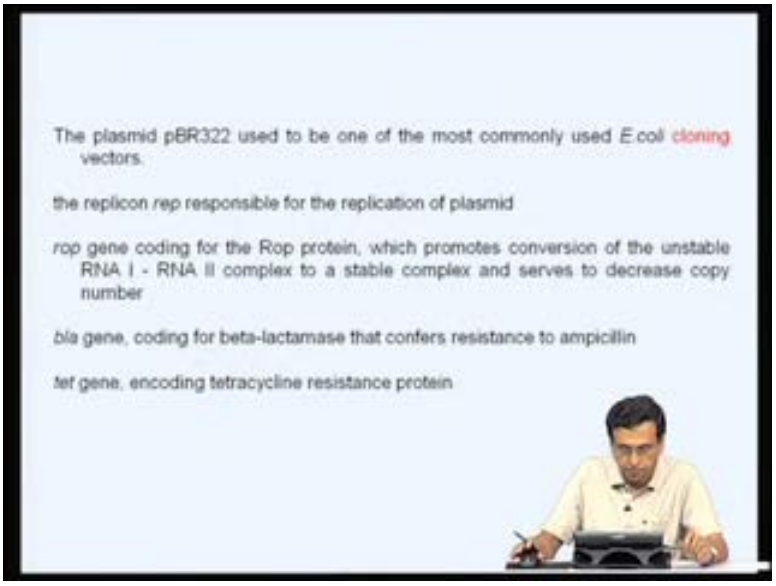
gene, and therefore, if you clone your gene of your resistance the PstI ampicillin resistance gene, and cells that which harbor such kind of a recombinant plasmids will now be resistant to tetracycline, but sensitive to ampicillin.

On the other hand, if you clone your foreign gene, or gene of you interest, in the BamHI site and disturb the tetracycline gene, and E.coli cells harboring such recombinant DNA molecules will be resistant to ampicillin, but sensitive to tetracycline. So, depending upon which antibiotic resistance gene you are introducing your gene in, you can either go for tetracycline resistance or ampicillin resistance.

So, E.coli cells harboring only the pBR322 will be resistant for both tetracycline and ampicillin, whereas if you clone your gene into the ampicillin gene, cells harboring this will be resistant to tetracycline. Whereas we clone your gene into tetracycline gene, those cells will be resistant for ampicillin. So, using this kind of definition, antibiotic selection marker, you can distinguish cells which harbor your recombinant DNA molecules.

So, cloning into one of the restriction sites, just like what I mentioned now, would activate one of the antibiotic resistance markers, leaving the other for the selection of the transformation. So, insertion of your foreign gene would inactivate the antibiotic resistance marker, and that can be taken the advantage for selection of those cells harboring this particular recombinant DNA. Screening for the absence of the second antibiotic marker was the putative evidence of a successful cloning event, and lower molecular weight of the pBR322 also allowed cloning of larger fragments. So, this was the first very popular cloning vector, which was designed after Stanley Cohen’s pSC101.

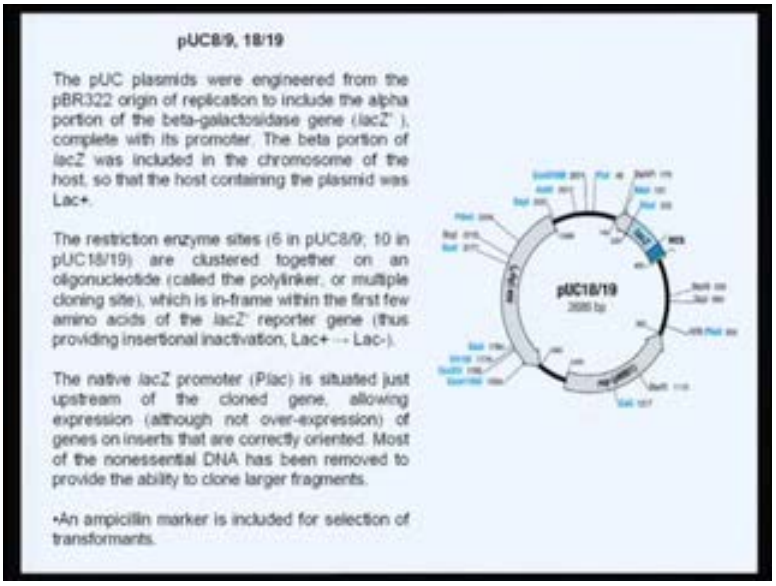
(Refer Slide Time: 30:56)




The pBR322 was one of the most commonly used E.coli cloning vectors, especially in the late 1980s and late 1990s. It had a replication called rep responsible for the replication of plasmid; it also had a rop gene, which codes for a Rop protein, which promotes conversion of a unstable RNA I-RNA II complex to a stable complex, and serves to decrease the copy number. So, pBR322 is very low copy number plasmid.

And the bla gene actually codes for a beta-lactamase protein, which actually confers resistance to ampicillin, whereas the tet gene codes for the tetracycline resistance protein. So, these are the basic features of a second generation plasmid vector pBR322, which had two antibiotic selection markers and a origin of replication, and a low copy number property.

(Refer Slide Time: 31:54)



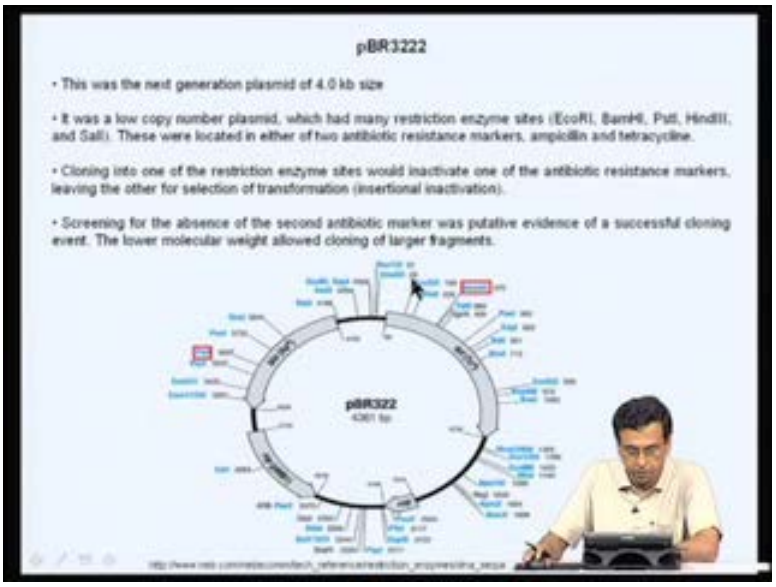

Third generation plasmid vectors– the third generation plasmid vectors are pUC vectors. It started from pUC 1, 2, 3, and so on, and went on up to pUC 18, 19, and so on so forth. These pUC plasmids were, actually, engineered from pBR322 origin of replication to include the alpha portion of beta-galactosidase gene. Beta-galactosidase is enzyme involved in lactose metabolism. So, it has the promoter of the beta galactosidase gene; the beta portion of the lacZ was included in the chromosome of the host. So, the host contains the plasmid was lac plus.

I will explain this little bit later, and more details what is called as alpha complementation. Basically, the pUC also had more restrictions site than the pBR322. For example, the pUC 8 or 9 plasmid had six restriction sites, whereas pUC 18 and 19 had ten restriction sites, and this is actually called as the multiple cloning site or the MCS or a polylinker, and this multiple cloning site is the place where you can insert your foreign genes.

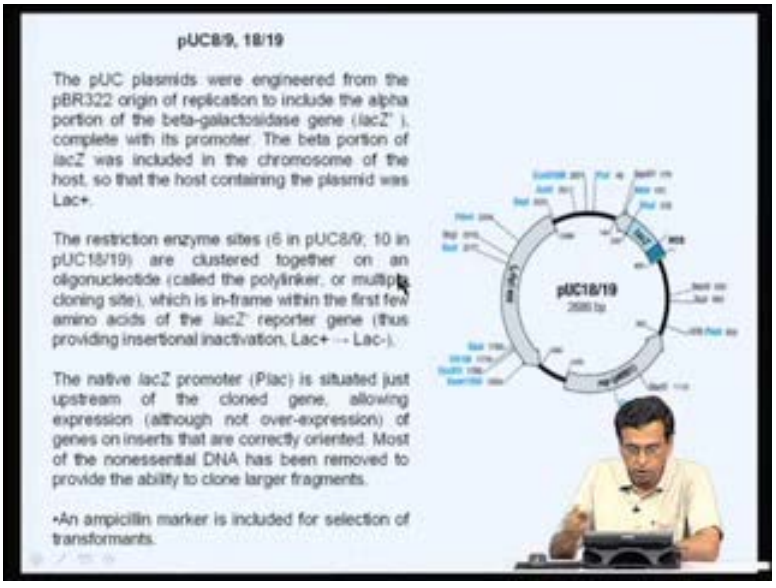
So, when you insert your gene into this multiple cloning site, it would disrupt the lacZ gene, and therefore, the lacZ protein will not be made, and therefore, cells harboring a recombinant pUC plasmid will form white colonies, whereas if you do not insert a gene, the lacZ gene will be properly made, and those cells will, if you now plate this on a plate containing what is called as a chromogenic substrate called X-gal, it will be cleaved. Beta galactose will cleave this X-gal and you get blue colonies.

So, E.coli cells which harbor this plasmid– pUC plasmid– will turn blue if you plate them on a plate containing a X-gal, whereas if you clone a gene into pUC plasmid and then plate them on a X-gal plate, those E.coli colonies will be white in color. So, using simply blue-white selection, you can distinguish cells which have taken up the native plasmid and cells which have contained the recombinant plasmid. The native lacZ promoter is situated just upstream of the cloning gene, allowing the expression of genes on inserts that are correctly oriented, and most of the nonessential DNA has been removed to provide ability to clone larger fragments.

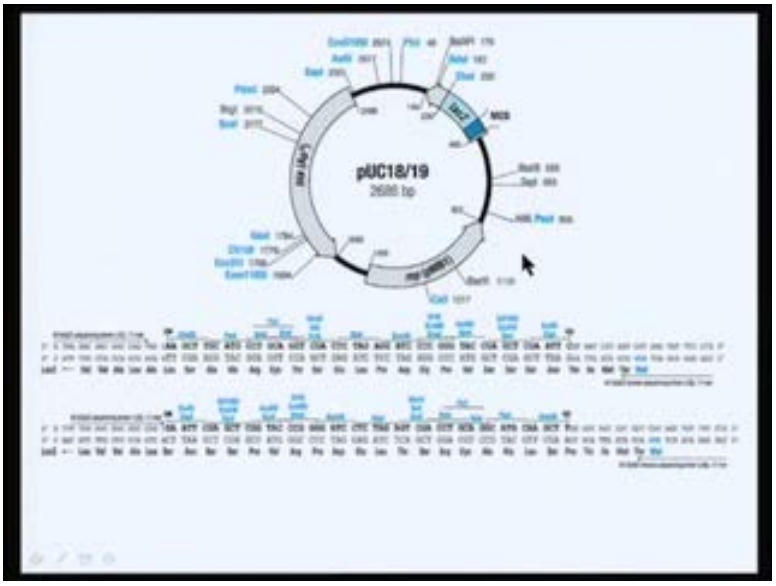
(Refer Slide Time: 33:53)




(Refer Slide Time: 33:57)

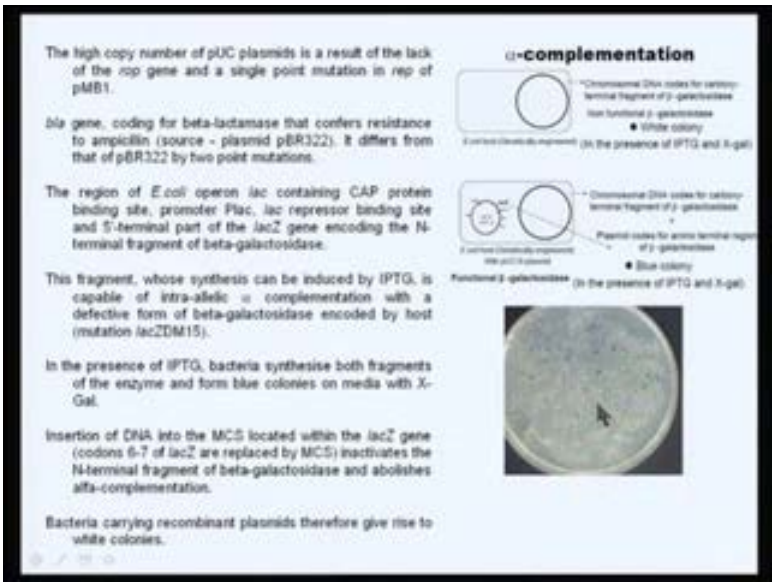


(Refer Slide Time: 34:08)




And you can see, the pBR322 was almost about 4 kb, whereas the pUC 19 is only about 2.6 kb. So, the vector size is much smaller, and therefore, larger plasmids can be cloned into pUC. It had an ampicillin selection marker for selection, for antibiotic resistance. This is the detailed map of the pUC plasmid. As I said, the multiple cloning site contained a number of restriction sites here.

(Refer Slide Time: 34:25)




So, if you insert your gene of interest, you can clone your gene of interest using any of these restriction enzymes, and then, you can do a blue-white selection to select the

recombinant plasmids. This is what explains how exactly do the selection of the cloning genes into pUC plasmids. The pUC plasmid had a high copy number. Remember, the second generation plasmid– pBR322– is a low copy number plasmid, whereas the pUC plasmids are high copy number plasmids. The bla gene, actually, is responsible; it codes for beta-lactamase. Therefore, it confers resistance for ampicillin. The region of the E.coli operon– lac– contains the cap protein. I am sure those of you have studied eukaryotic gene expression would have studied lac operon in detail.

How a lac operon is regulated? You have what is called as an operator, promoter, and then structural genes, and how two proteins– the lac repressor and the cap protein– play a very important role in regulation of lac operon. I will not go to the details. So, basically, E.coli lac operon containing the cap protein binding site, as well as the promoter Plac and the lac repressor binding site, is part of the vector that what constitutes the promoter. So, because you have this lac promoter here, the pUC actually serves an expression vector. So, if you clone your gene downstream of this pUC plasmid, you can actually express your protein of your interest.

So, when you clone your gene downstream of the gene, the gene gets expressed using the lac promoter, and whose expression induced by isopropyl thiogalactose or IPTG. Basically, this vector contains only 5 prime terminal part of the lacZ gene, encoding the N-terminal part of the beta-galactosidase. So, this is where the very unique system comes. This plasmid contains only the 5 prime region of the lacZ gene, so it only codes for the amino terminal amino acids or the lacZ. The C-terminal amines of lacZ, actually, come from the chromosomal bacterial chromosome. So, only if you have both the amino-terminal part and carboxy-terminal part of the lacZ together, then you will get a functional beta-galactosidase or lacZ protein.

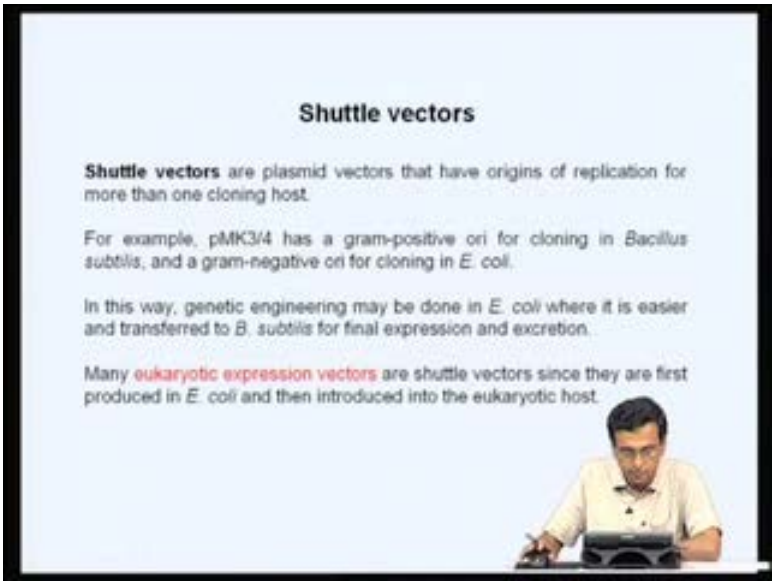
So, if you have only the vector, then you will….. if you have both vector and the appropriate host, you get a functional beta-galactosidase, but if you clone your gene into this multiple cloning site, then a functional amino-terminal part of the lacZ will not be made, and therefore, a functional lacZ gene will not be produced. So, in the presence of IPTG, which is nothing but isopropyl thiogalactose, which is an inducer of the lac operon, they synthesize both the fragments– one coding for the amino-terminal region of lacZ, another coding for the carboxy-terminal region of lacZ, which comes from the bacterial chromosome.

And therefore, you get functional beta-galactosidase, and if you grow such colonies and DNA contain what is called as a X-gal, you get blue colonies. That is what is shown here, because you have blue colonies here. Whereas, if you insert a gene into the multiple cloning site, then the amino-terminal lacZ will not be made, and therefore, you will not get a functional lacZ protein. And such bacteria which contain a recombinant plasmid, in which gene is inserted into the multiple cloning site will become white on the plate; you can see here. So, the blue colonies means, cells which are expressing only the pUC plasmid; the white colonies are those which are expressing recombinant pUC plasmids, in which insert has been inserted.

So, bacteria carrying the recombinant plasmid, therefore, give rise to white colonies. This entire process, known as the alpha complementation, is, actually, described here. You can see, the chromosomal DNA codes for the carboxy terminal fragment of the beta-galactosidase. So, if you do not have plasmid, the E.coli can express only the carboxy terminal part of the lacZ, and therefore, it can form only a white colony in the presence of IPTG and X-gal.

Whereas, if you now introduce the pUC plasmid into these bacterial cells, the C-terminal part will made from the bacterial chromosome. The amino-terminal region of the lacZ will made from the plasmid. Therefore, both complement each other; you get a functional beta-galactosidase, and now, if you plate them on a plate containing IPTG and X-gal, you get blue colonies. So, by doing what is called as a blue-white selection, you can distinguish cells which are harboring only the plasmid, and cells which are harboring the recombinant plasmids. It is a very, very popular method of generating recombinant DNA molecules in laboratories.

(Refer Slide Time: 38:33)



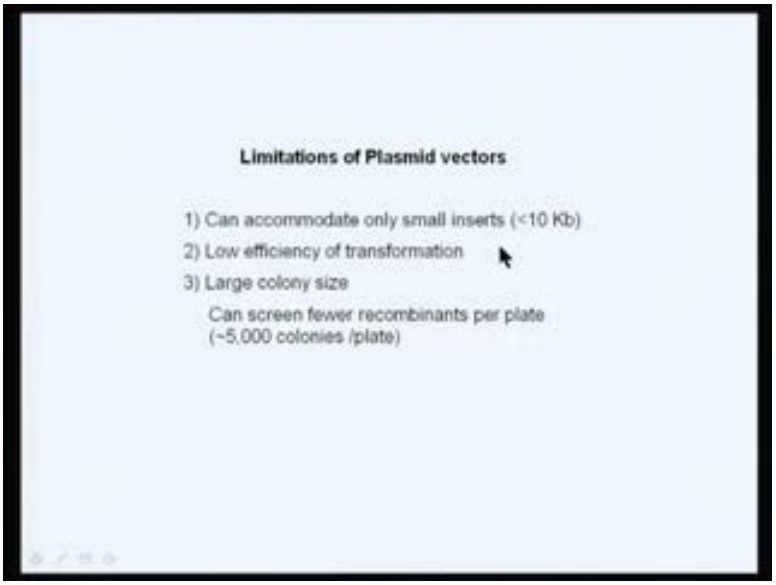

So, we have, so far, discussed about what is called as the first generation plasmid vectors, second generation plasmid vectors, third generation plasmid vectors. Now, there are also very important vectors called as shuttle vectors, that is what described here. Shuttle vectors are plasmid vectors that have origins of replication for more than one cloning host. So, virtually, shuttle vectors can replicate in two different organisms. For example, there is a vector called as pMK3/4, which has a gram-positive origin for cloning in Bacillus subtilis, and gram-negative origin for cloning in E.coli.

So, it can replicate in both the host both in Bacillus as well as in E.coli. So, in this way, genetic engineering may be done in E.coli, because E.coli is much more easy organism to manipulate, and once you do all the basic cloning techniques in E.coli and generate recombinant plasmid, then you can take this recombinant plasmid and put it in Bacillus subtilis for expression. In fact, many eukaryotic expression vectors, which we are going to discuss extensively in next class, they are called as shuttle vectors because they contain an eukaryotic promoter, but they will contain a bacterial origin of replication as well as a mammalian origin of replication.

So, you can first do all the cloning into this vector and put them in bacteria, and make this plasmid in large amounts, and then introduce them into an eukaryotic cell. Then, the eukaryotic promoter will work in eukaryotic cells and your protein can be expressed

there. So, shuttle vectors are very, very important for making recombinant proteins and expressing genes of your interest.

(Refer Slide Time: 40:10)




Now, what are the limitations of plasmid vectors? They can accommodate only small inserts, so if you have gene, which is more than 10 kb, you cannot clone them into bacterial vectors, I mean, plasmid vectors, because they cannot accommodate more than 10 kb in insert size. The efficiency of transformation is also very low, so if there are probably 10000 or 1000 cells of bacteria, only 1 or 2 percent of them actually take up the plasmid, but although there are more efficient methods of transformation, like electroporation and so on so forth, the efficiency of transformation still high.

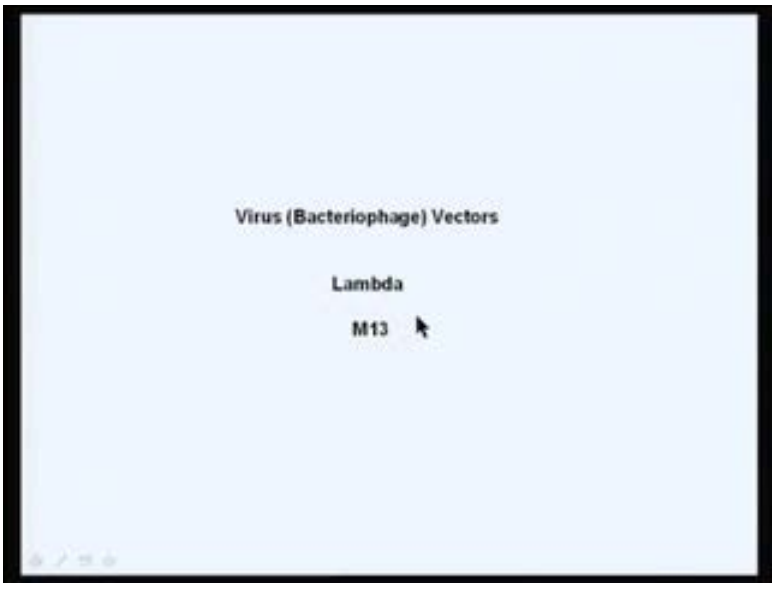
Therefore, it becomes very, very important to have an appropriate selection marker, so that you can easily kill the cells which have not taken up the recombinant DNA. That is why the blue-white selection, antibiotic resistance, all become very, very important. Because of the low efficiency of transformation, you can eliminate all those cells, which have not taken up the recombinant DNA. The colony sizes are usually large, and if the E.coli colonies are very large, and therefore, you cannot screen too many colonies.

This is a big drawback, because if you want to now make genomic libraries, where you want a large, huge number of colonies on a plate, then the bacteria are not the ideal ones. That is where you have to do what is called as the phage vectors, which will discuss in the next few minutes. So, you can only screen few recombinants per plate, because the

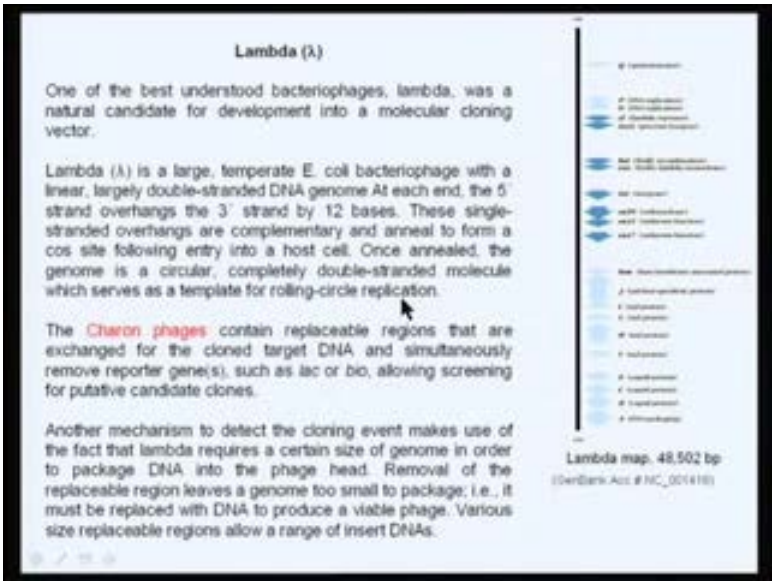
colonies are very large. So, you cannot plate this into very high cell density. So, that is one of the major drawbacks of plasmid vectors.

So, because of these drawbacks in plasmid vectors, that is smaller size of cloning and you cannot screen too many colonies, and the third one is low efficiency of transmission, people went on to develop what are called as the phage vectors. So, what is a phage? Phage is a virus that infects a bacterial cell. So, this is, basically, a bacterial virus. So, there are number of viruses which routinely infect bacteria and lyse them. So, people took advantage of this and see can you actually eliminate some of the nonessential regions of this phage genome where you can now put your foreign DNA into this genome and develop what are called as the phage vectors.

(Refer Slide Time: 42:05)



(Refer Slide Time: 42:10)




The two popular phage vectors, which routinely used are called as the lambda phage vectors and M13 phage vectors. Now, let us see what these are. Lambda– now, lambda is a linear double-stranded bacteriophage, which is one of the most well studied bacteriophages, and in fact, very important information on regulation of gene expression in respect to prokaryotes came from the understanding of the bacteriophage life cycle, especially the lambda life cycle.

So, lambda is a large temperate E.coli bacteriophage with a linearly large double-stranded DNA genome. At each end of the genome, the 5 prime strand overhangs the 3 prime strand by 12 bases, and these single strand overhangs are complementary and anneal to form a cos site, following entry into a host, and once annealed, the genome is circular and completed a double-stranded molecule, which serves template for rolling-circle replication.

So, the two ends of the genome contain what are called as the cos sites, and once they enter into the E.coli cell, these cos sites anneal with each other and generate a circular DNA, so that this phage genome can be replicated by using a rolling circle model. Some of the phages, which have been extensively used for making this kind of a phage vectors called Charon phages, they actually contain a repressible regions that are exchanged for the clone target DNA, and simultaneously remove reporter genes such as lac or bio, allowing screening for putative candidate clones.

So, another mechanism to detect cloning event makes use of the fact that lambda requires a certain size of the genome in order to package DNA into the phage head. Removal of the replaceable region leaves the genome too small to package, that is it must be replaced with DNA to produce a viable phage, and various size replaceable regions allow a range of insert DNA. What this actually ensures that in order for the phage DNA to be packaged inside the phage head, you require a certain size. So, if you have a phage without a foreign DNA, then the phage genome is too small, and therefore, such phage genomes will not be packaged into the bacteriophage.

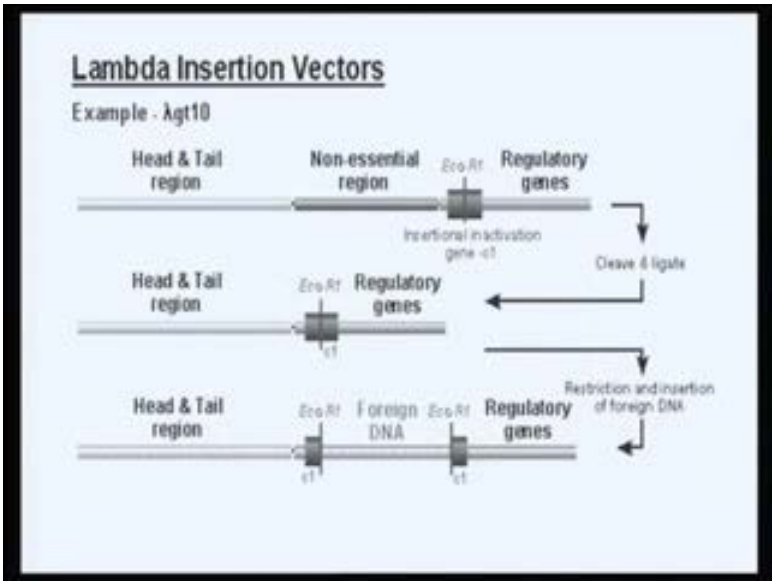
So, this is one good, very good way of screening. So, if you want to eliminate bacteriophages which have taken only the small phage DNA, such phage DNA, which have not taken up the inserts, will never get packaged into bacteria. So, you can easily eliminate those phage DNA molecules which are not recombinant. So, only the recombinant phage DNA in which the foreign DNA has been inserted into the phage is of appropriate size, and only they will be the packaged into the phage head, and therefore, you will get colonies.

(Refer Slide Time: 44:31)


Many lambda-based vectors have been developed by companies such as New England Biolabs, Clontech, Stratagene, etcetera, and they are all commercially available now. So, if you want to clone any large insert, which is more than 10 kb, or you want to make genomic libraries, the preferred vector is lambda vector and not the plasmid vector. But,

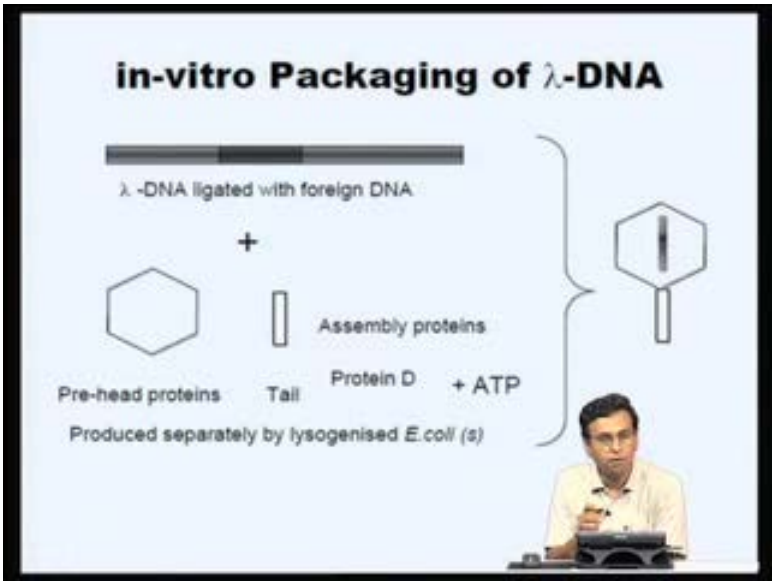
if you want to make a small gene, like 3 or 4 kb, and you want to just express them in bacteria, plasmid vectors are the appropriate choice.

(Refer Slide Time: 44:58)




This just (( )) mentions one of the popular lambda vectors called as lambda gt10, which is extensively used for a making genomic libraries in the late 1980s and late 1990s, and for example, contains– these are the phage genomes; it contains the head and tail region. The phage genome also contains what is called as a non-essential region, and it contains the regulatory genes, and basically, you create an EcoRI site here, which inactivates the gene called c1. So, c1 protein will not be made if you clone your gene of your interest in the EcoRI site.

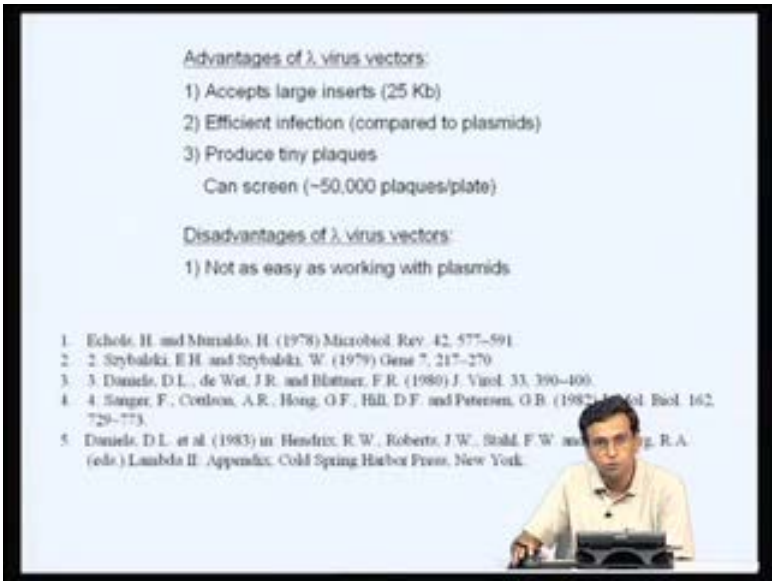
(Refer Slide Time: 45:50)




So, basically, what you do is you cut this phage DNA with the EcoRI, and then put your gene of you interest– clone your foreign DNA– into the c1 region, and this recombinant DNA now can be put inside is done what is called as a…. Once you put your foreign DNA between the head and tail region and regulatory regions, you take this recombinant DNA, and then use what is called as the packaging extract; that is, these are the proteins and other things which are required for packaging the lambda DNA into the bacteriophage.

And once you have this packaging extract containing all these assembly proteins and ATP and so on so forth, this DNA will get packaged inside phage head, and now, if you now add to the E.coli cells, this phage will be taken up by the E.coli, and you will get… the bacteriophage will successfully replicate in E.coli and you get a recombinant phage. So, you can generate a recombinant phage containing a foreign DNA by using this kind of a mechanism. So, instead of cloning your DNA into plasmid DNA, you clone your foreign DNA into a phage DNA, and then package this into a recombinant phage, and then infect the E.coli cells with this phage. So, you get large amount of this recombinant phages.

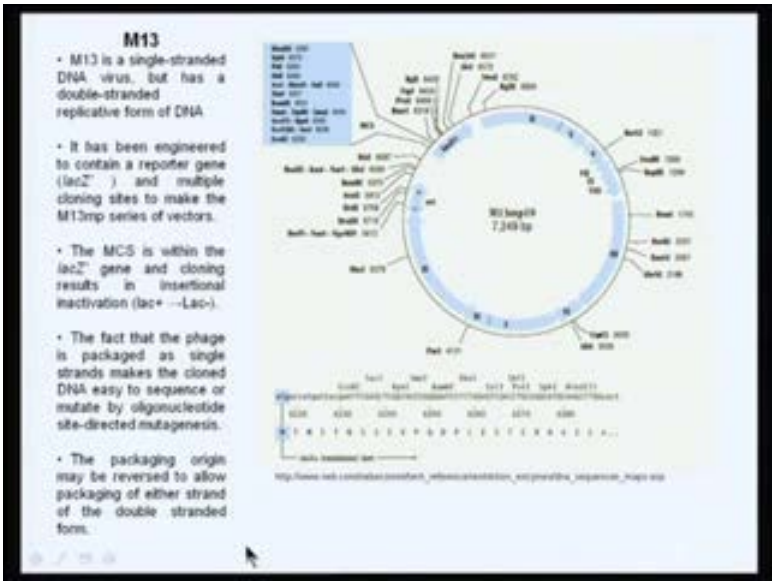
(Refer Slide Time: 46:35)




What are the advantages of these lambda vectors? It can accept inserts up to 25 kilo bases, whereas the plasmid vectors, which we discussed just now, can accept only a up to 10 kilo base DNA, and the efficiency of infection is very high compared to lower efficiency of transmission of the plasmid DNA into E.coli. They produce very tiny plaques, so you can screen much larger number of colonies.

Whereas, if you make a library– cDNA library or genomic library– in a plasmid vector, you can only screen about 5000 colonies per plate, whereas if you make a same library in a phage vector, you can screen up to 50000 plaques in a single plate. So, it is huge advantage. I have given some references, which you can use for studying a more about these lambda vectors and so on so forth.

(Refer Slide Time: 47:18)

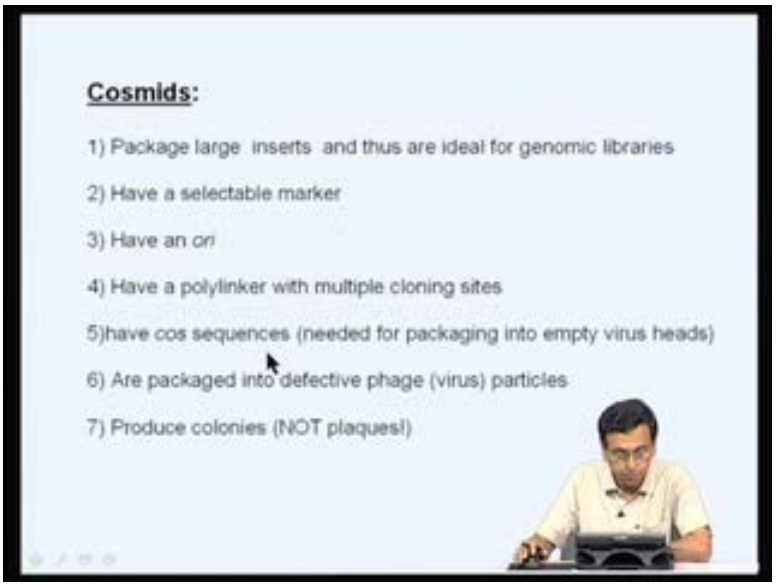



Another very popular bacteriophage vector which is routinely used for generating recombinant DNA molecules is called as M13. The difference between M13 phage and the lambda phage– lambda phage is double-stranded DNA; M13 is a single-stranded DNA, and another important feature is a M13 is that it can exist both in a single-stranded form as well as the double-stranded form. Inside the bacterial cells, it exists as a double-stranded form, and once it comes out of a phage, in the phage it exists as a single-stranded form.

So, I will not go into the details, because these are mostly the prokaryotic vectors. You can, actually, go through some of the details and understand how an M13 phage is constructed. One can always go to the website of New England Biolabs; extensive details about some of these vectors are given in these websites. So, I strongly urge you to visit some of these web sites– NAB; all the details about these vectors– bacteriophage vectors, plasmid vectors, is given in this website.

(Refer Slide Time: 48:16)

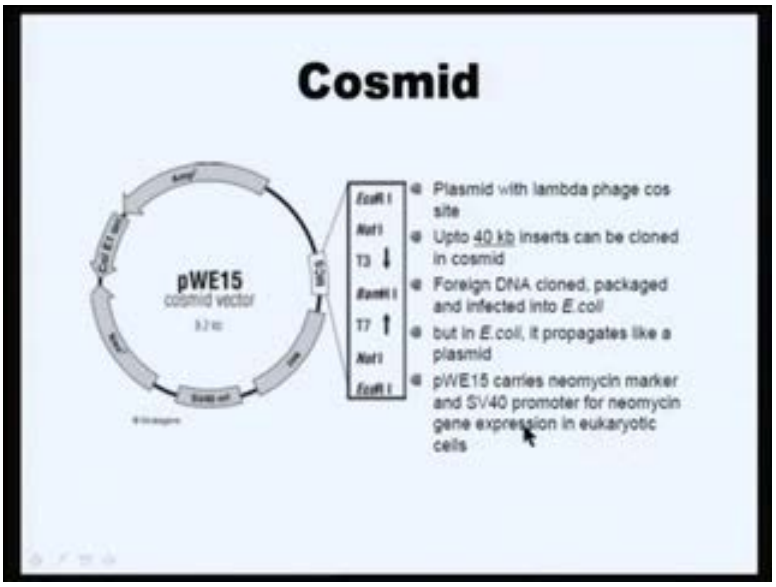

(Refer Slide Time: 48:23)




The other important vectors that are designed were called as cosmid vectors. Now, cosmid vectors are plasmids that contain cohesive ends of bacteriophage lambda, and allow packaging of DNA into the lambda phage heads. So, what are the advantages of cosmids? They package large inserts, and thus, are ideal for genomic libraries. They have, also, a selectable marker, have an origin of replication; they also have a polylinker with multiple cloning sites.

They have the same cos sequences like the lambda. So, this can be easily packaged into virus heads, and they are packed into defective virus particles, they produce…. The difference between a cosmid and a phage is that cosmid produce colonies, whereas phages produce plaques. This is the major difference between the cosmid vector and a phagemid vector or a phage vector.

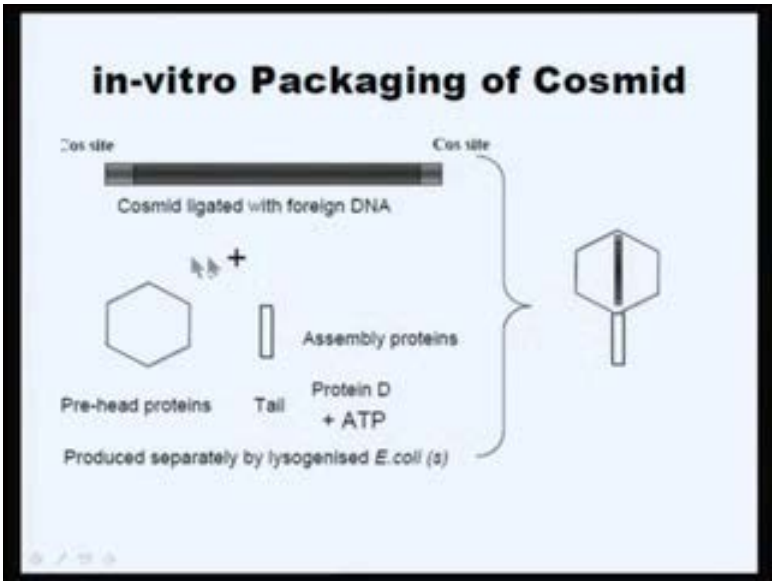
(Refer Slide Time: 48:55)




Again, cosmid vectors are, actually, made from the number of companies. Here is the one of the cosmid vector made by company called as Stratagene, called pWE15. Again, it is a plasmid with lambda phage cos sites, so they can circularize inside bacterial cells. It can take up to 40 kilo base of inserts, whereas the phage can take up only about 25 kilo bases.

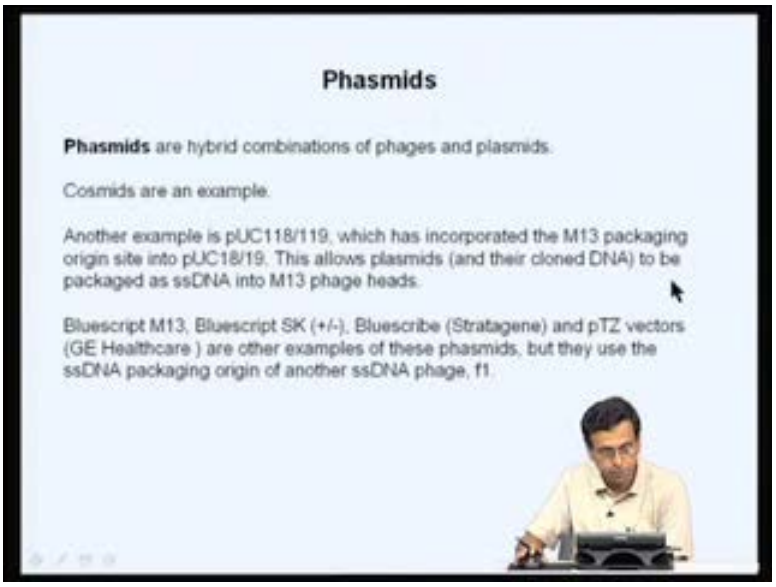
So, cosmids can take up your longer DNA. So, the foreign DNA can be made, easily cloned into these multiple cloning sites of the cosmids, packaged, and infect the E.coli, and in E.coli, it propagates like a plasmid, and it also has what is called as neomycin marker. So, this plasmid can also be introduced eukaryotic cells, and you can also select eukaryotic cells harboring these cosmid molecules.

(Refer Slide Time: 49:38)




Again, you have packaging extracts, so once you have generated recombinant DNA containing the foreign DNA inserted into the cosmid vector, you can package them and you can produce a recombinant phage, and the advantage of this is that once it is inserted in the E.coli cell, this will multiply like a plasmid. That is the differentiation between a cosmid vector and a phage vector.

(Refer Slide Time: 49:57)

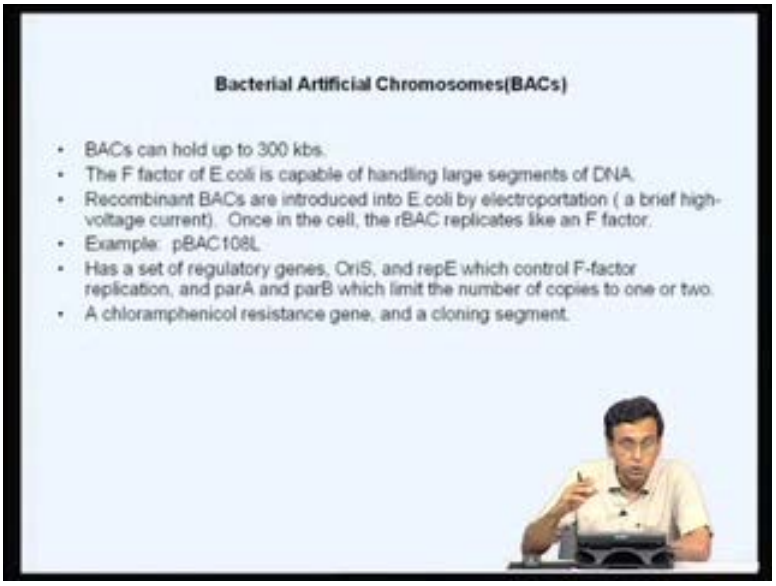



Similarly, you have phasmids which are, again, hybrid combinations of phages, and phasmids are very similar to cosmids. We will not go to the details; some of the very

important or very successful phasmids are, for example, the pUC 118 and 119, which also contains what is called as a M13 packaging origin site, and therefore, it allows the phasmids to be packaged as single-stranded DNA into M13 phage heads.

So, the Bluescript M13, again, is a very popular vector made by Stratagene. Again, you have pTZ vectors made by GE Healthcare; earlier, it used to be called as the Formatia. These are all very, very important called phagemids or phasmids, which has been extensively used in the area of molecular biology and recombinant DNA production.

(Refer Slide Time: 50:38)

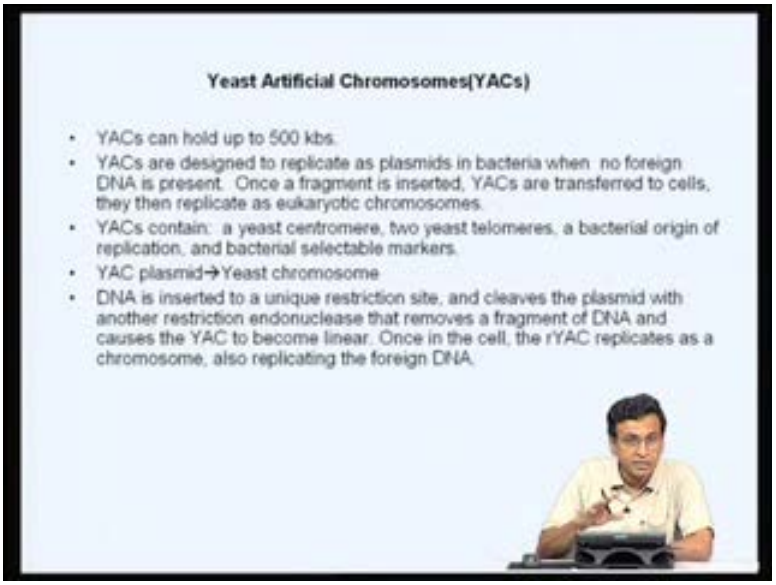



If you want to clone even larger fragments of DNA, you have, now, what are called as bacterial artificial chromosomes or BACs. Now, BACs can hold up to 300 kilo bases of DNA. You can see, it started with the plasmids which can hold up to 10 kb; you went to phage vectors, you know, about 25 kb; you went to cosmids, where you can clone up to 50 kb. Now, you have bacterial artificial chromosomes, which can take up to 300-kilo bases of foreign DNA.

So, you can see, depending upon what size of DNA you want to clone, you can choose any one of these vectors. So, the F factor of E.coli is capable of handling large segments of DNA. The recombinant bacterial artificial chromosomes are introduced in the E.coli by electroporation, and once inside the cell, the recombinant bacterial replicates like an F factor, and one of the very popular bacterial artificial chromosomes used is called as

pBAC108L. I will not, again, go to the details, but these are all very, very popular vectors, in fact, used for introducing large chunks of DNA into bacteria.

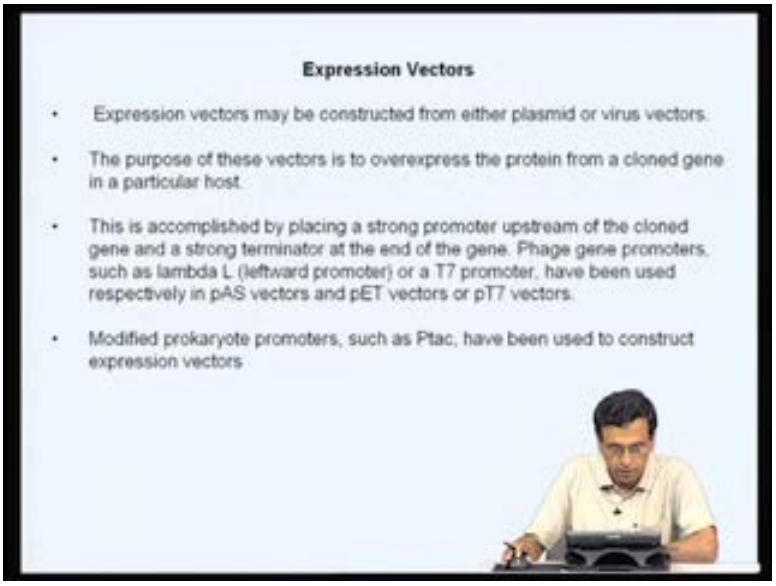
(Refer Slide Time: 51:31)




These BACs and YACs they all played a very, very important role in things like human genome sequence. In all that, when you have to clone huge pieces of human DNA into these human genome projects, mouse genome projects, and such things, these BACs and the YACs made a very, very important difference. So, just like you had bacterial artificial chromosomes, we have yeast artificial chromosomes; they can also hold up to 500 kilo bases of DNA.

And YACs are designed to replicate as plasmids in bacteria. When no foreign DNA is present, and once a fragment is inserted, YACs are transferred into cells and they replicate like eukaryotic chromosomes. So, the YACs contain a yeast centromere, two yeast telomeres, a bacterial origin of replication, and bacterial selection marker, and these plasmid behaves like a yeast chromosome.

(Refer Slide Time: 52:20)

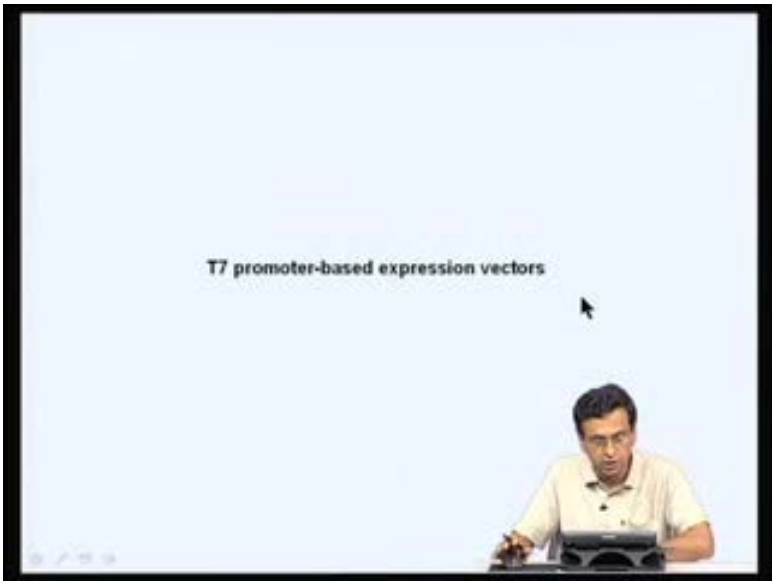



The other important aspect, what we can now discuss, are called as the expression vectors. They can also be constructed from any plasmid or virus vectors. The purpose of this expression vector is to over express the protein from a cloned gene. So, for many of the vectors we have discussed, they are called cloning vectors; we cannot, really, express your proteins in large amounts using this cloning vectors.

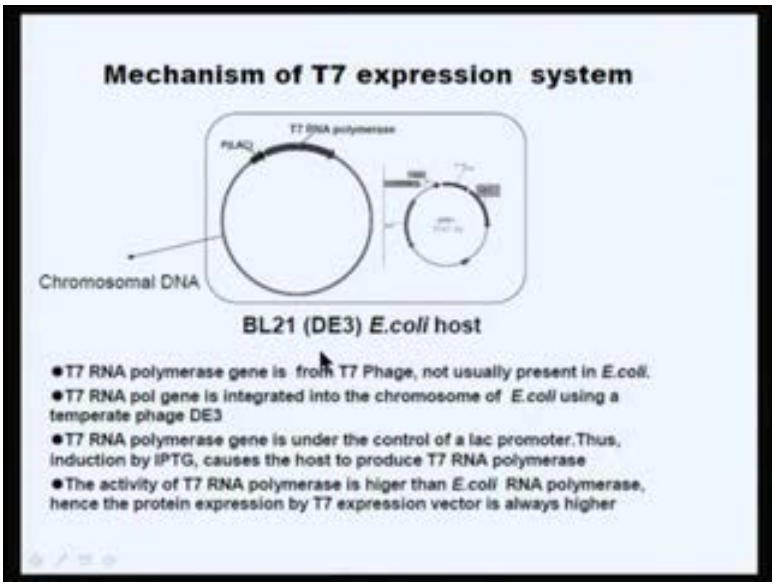
But now, you have expression vectors where you put a powerful promoter and you can express your gene as well. So, these are called as expression vectors. So, the difference between a cloning vector and expression vector is that in cloning vector, you can just clone the genes, whereas the expression vectors, usually, contain very powerful promoters, so you can not only clone a gene, you can also express the protein encoded by the gene in very high amounts.

So, usually, expression vectors contain a very strong promoter upstream of the cloned gene, as well as a strong terminator promoter. Phage gene promoters such as lambda leftward promoter or a T7 promoter are very popularly used in the vectors like pET vectors and pT7 vectors, which are very, very popular expression vectors that people now use for making recombinant proteins.

(Refer Slide Time: 53:20)




(Refer Slide Time: 53:27)




I will just take couple of minutes to explain what are called as the pT7 vectors. T7 promoter-based expression vectors, which are extensively used today in the area of molecular biology and recombinant DNA technology. What is the mechanism of T7 expression system? T7 RNA polymerase is a RNA polymerase coded by, required by the T7 bacteriophage, and this T7 RNA polymerase recognize a very short sequence, about 15 to 20 bases. It is called as a T7 promoter.

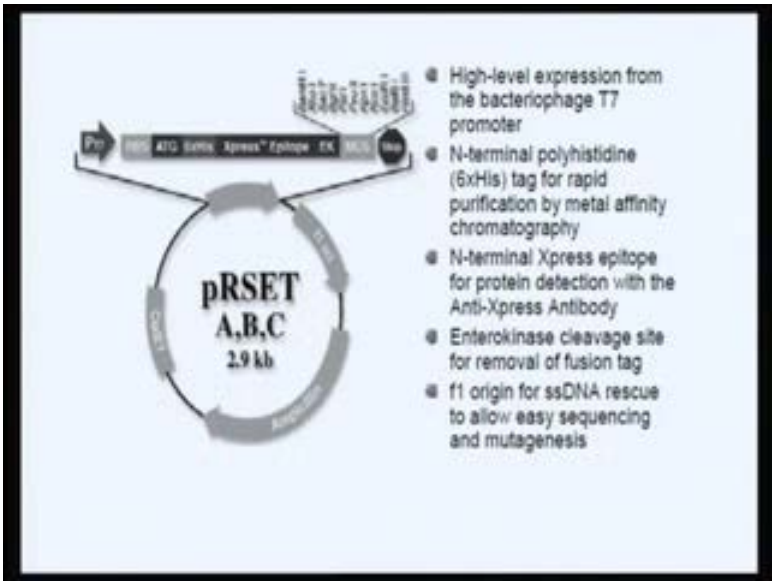
So, what you do is that you place this T7 RNA polymerase gene under the promoter– under a IPTG inducible lac promoter– and put it in a chromosomal DNA of a bacterial cell, and usually, the E.coli strain which harbors such a T7 RNA polymerase in the bacterial chromosome is usually called BL21 DE3 E.coli cells. Then, you introduce your plasmid, which actually contains your gene of interest downstream of a T7 promoter.

So, your foreign gene of interest downstream of the T7 promoter is placed in a plasmid, and then introduce this plasmid into these E.coli cells. Now, if you take the cells and add IPTG, the IPTG will induce the lac promoter. The T7 RNA polymerase enzymes synthesized from the bacterial chromosome– that T7 RNA polymerase will now come and bind to the T7 RNA promoter binding promoter site present in the plasmid, and induce the expression of your foreign gene.

So, you can see, the RNA polymerase comes from the bacterial chromosome and goes and binds to the promoter binding site present in the plasmid, and your protein is expressed in very high amounts. So, the T7 RNA polymerase gene is from T7 phage; it is not present in the E.coli. The T7 RNA polymerase gene is integrated into the chromosome of E.coli using a temperate phage DE3, so DE3 stands for a temperate phage.

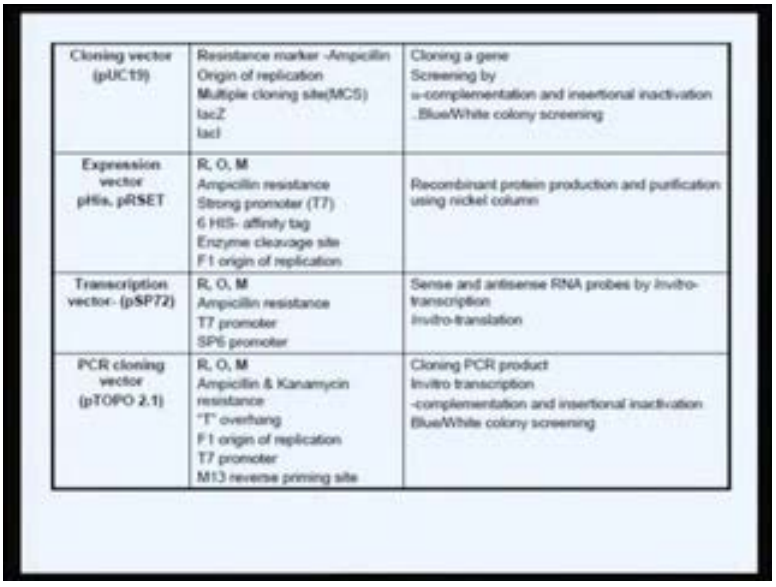
The T7 RNA polymerase is under the control of a lac promoter. Therefore, by adding IPTG, you can induce the host to produce the T7 RNA polymerase, and this T7 RNA polymerase activity is much higher than E.coli RNA polymerase, and therefore, protein expression by T7 expression vector is very high. So, you can express proteins at very high levels using this T7 expression system.

(Refer Slide Time: 55:21)




I have just discussed somehow, described the some very popular T7 based expression vectors that are sold by number of companies. I will not go into the details; one of them called as a pRSET A, B, C, and so on so forth. You can go to this– this is owned by company called Clontech; you can go to the websites and learn more about them.

(Refer Slide Time: 55:36)



| | | |
|---|---|---|
| Cloning vector (pUC19) | Resistance marker -Ampicillin<br>Origin of replication<br>Multiple cloning site(MCS)<br>lacZ<br>lacI | Cloning a gene<br>Screening by<br>α-complementation and insertional inactivation<br>.Blue/White colony screening |
| Expression vector pHis, pRSET | R, O, M<br>Ampicillin resistance<br>Strong promoter (T7)<br>6 HIS- affinity tag<br>Enzyme cleavage site<br>F1 origin of replication | Recombinant protein production and purification using nickel column |
| Transcription vector- (pSP72) | R, O, M<br>Ampicillin resistance<br>T7 promoter<br>SP6 promoter | Sense and antisense RNA probes by invitro-transcription<br>Invitro-translation |
| PCR cloning vector (pTOPO 2.1) | R, O, M<br>Ampicillin & Kanamycin resistance<br>"T" overhang<br>F1 origin of replication<br>T7 promoter<br>M13 reverse priming site | Cloning PCR product<br>Invitro transcription<br>-complementation and insertional inactivation<br>Blue/White colony screening |

So, what I discussed, so far, is something about vectors. I can just mentioned here, we talked about cloning vectors; we talked about expression vectors. There are also what are called as transcription vectors, where you can make a large amounts of RNA using what

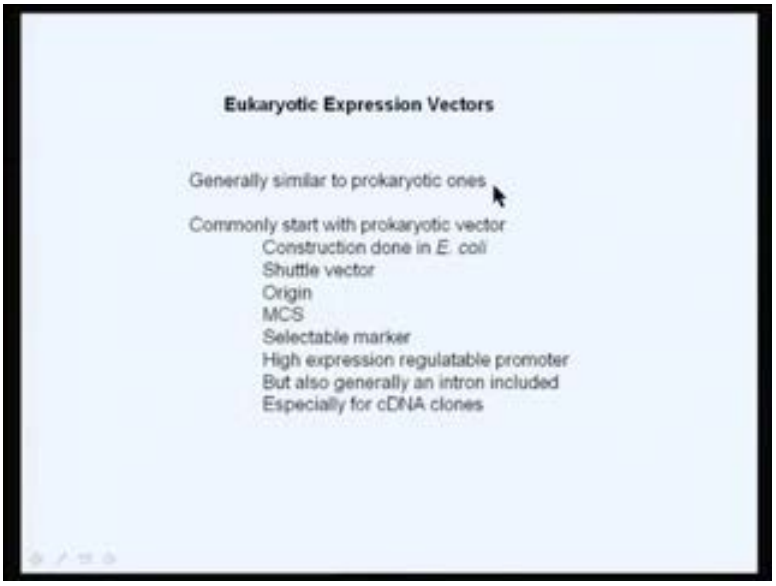
is called as the T7 RNA polymerase and SP6 RNA polymerase. We also have are called as PCR cloning vectors, called pTOPO, and so on so forth, but since they are not directly related to eukaryotic gene expression, will not discuss about those things. One can always go and read more about that.

(Refer Slide Time: 56:03)

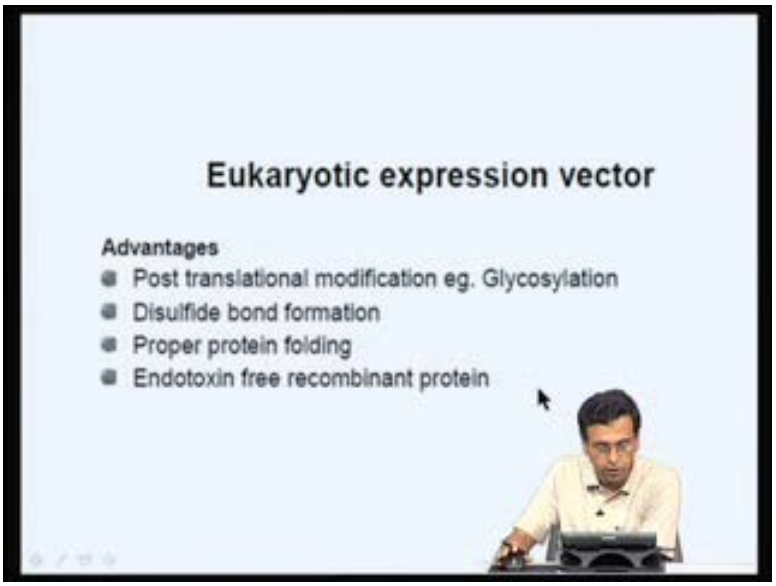
| | | |
|---|---|---|
| Secretion vector<br>pET21 | R, O, M<br>Ampicillin resistance<br>F1 origin of replication<br>T7 promoter<br>**PelB signal peptide**<br>6 HIS- affinity tag<br>Enzyme cleavage site | Recombinant protein production, secretion and purification using nickel column |
| Cosmid<br>pWE15 | R, O, M<br>Cos site<br>(Does not have the lambda phage gene except COS site)<br>Can be package with head and tail | Upto 40 kb inserts can be cloned<br>Used for making gene libraries |
| Lambda insertion vector<br>λgt10 | MCS in the c1 repressor gene<br>Can be package with head and tail | Upto8-23 kb inserts can be cloned<br>Used for making gene libraries |
| Filamentous phage | Single stranded DNA<br>Infect E coli having pilus (F') | useful for sequencing, in-vitro mutagenesis<br>Upto 3 kb inserts can be cloned |
| Phagemid | Phage and plasmid properties<br>Contains M13 initiator and terminator | Used for making gene libraries<br>Converting into plasmids with out subcloning |

You have secretion vectors, where you can secrete your protein of your interest. You have cosmid vectors, lambda vectors, phagemids; one can go and read up, and I have listed here what all the major features of these vectors and what are their applications, and what are the characteristic features of these vectors. One can read up these and get some overall view about what these vectors are. So, what we will discuss in the next class– so far, we discussed primarily about prokaryotic cloning vectors and prokaryotic expression vectors, and with this background of knowledge, in the next class, we will discuss about what are eukaryotic expression vectors.

(Refer Slide Time: 56:41)




(Refer Slide Time: 56:56)




So, the eukaryotic expression vectors are similar to prokaryotic expression vectors. They are also constructed in E.coli; they, usually, are shuttle vectors. So, first clone the genes, propagate the plasmids in E.coli, and then shift them to eukaryotic cells. They contain a again regulatable promoter, usually, they also contain a intron, and usually, cDNAs are used for expressing eukaryotic vectors.

(Refer Slide Time: 57:16)


Why do need an eukaryotic expression vectors? They have many features. I think we will discuss these things in next class in more detail. They contain many features, which are not present in prokaryotic expression vectors, and if you want many of these features, you have to clone your gene into eukaryotic expression vectors. I just mentioned some of the very classical papers, which I think one can go through and understand how the biotechnology industry was started in the 1970s and 1980s, and how lot of people made money using these kinds of a technologies. I think I will stop here.