Penn State Environmental Law Review

Volume 3 | Number 3 Article 4

8-1-1994

A Primer on Genetic Engineering II: Gene Cloning

Michael P. Roberts Ph.D.

Follow this and additional works at: https://elibrary.law.psu.edu/pselr

Recommended Citation
Michael P. Roberts Ph.D., *A Primer on Genetic Engineering II: Gene Cloning*, 3 *Penn St. Envtl. L. Rev.* 11 (1994).

This Article is brought to you for free and open access by the Law Reviews and Journals at Penn State Law eLibrary. It has been accepted for inclusion in Penn State Environmental Law Review by an authorized editor of Penn State Law eLibrary. For more information, please contact ram6023@psu.edu.

A PRIMER OF GENETIC ENGINEERING II: GENE CLONING.

*Michael P. Roberts, Ph.D.**

Good morning. I'd also like to thank the organizers for inviting me to speak today. Now you know all there is to know about DNA structure and function; we can forge ahead.

I teach genetics at Dickinson College and take about a month of my course to cover the topics that were just discussed by Dr. Bardale in fifteen minutes, but let's keep going. I'm going to talk about three things today: Gene cloning, amplification of genes, and gene transfer.

Let's start with gene cloning. I want to begin by defining some terms, and a term that is often misused in the popular press is cloning. When a molecular biologist talks about cloning, he or she is generally referring to "gene" cloning. Gene cloning is the isolation of a gene or a DNA fragment and the clonal propagation of that fragment as a recombinant DNA molecule.

This is not the same as cloning organisms. So there's a distinction to be made between cloning organisms, which you will hear more about in a minute, and cloning genes. In gene cloning, the goal is the production of genetically identical recombinant DNA molecules.

What is a recombinant DNA molecule? Recombinant DNA molecules are novel DNA molecules that consist of a combination of a DNA insert and a DNA vector. The vector DNA simply carries the insert DNA, and the insert DNA is the cloned gene. The other special property of the vector is that it allows the cloned gene to be replicated.

One type of vector DNA that is used to carry cloned genes is a naturally occurring bacterial DNA molecule that's called a plasmid. Plasmids are small, circular DNA molecules that are found within bacterial cells and have the special property of self-replication. That is, they are capable of reproducing themselves to produce many, genetically identical copies.

The typical plasmid is about 4,000 base pairs in length. That is a fairly small molecule. And let's contrast that to the bacterial chromosome that has a genome of about four million base pairs. So the plasmid is significantly smaller than the chromosome of the bacterium which is the repository for most of the bacterial genes.

Since the plasmid can reproduce itself or replicate independently of the bacterial chromosome, which replicates only when the bacterium undergoes cell division, it can exist in multiple copies within the bacterial cell.

Because of the size difference, plasmid DNA can be easily isolated from bacterial cells. All right, now I'd like to talk about just how we can create recombinant DNA molecules.

Lets say, for example, that the gene we want to clone is a human gene. The human genome, which is that collection of all the genetic material or genes present in each cell that define us as human organisms, is significantly larger than the bacterial genome. It consists of about three billion base pairs. That's reassuring, because we'd like to think that we're more complicated than bacteria.

---

*B.S., Colgate University, 1977; M.S., Miami University of Ohio, 1979; Ph.D., Yale University, 1988.

The first step in cloning then is to take that 3 billion base pairs of human DNA and cut it into smaller pieces which we can actually clone into the vector molecule. There are special enzymes which do that cutting called restriction enzymes. Restriction enzymes are molecular scissors. They read the DNA, and at specific locations within the DNA they make a cut.

The DNA fragment that contains the gene of interest is then inserted into the plasmid vector, which has also been cut with a restriction enzyme, and is ligated into place using another enzyme that seals the DNA and reforms a circle. So, we've started with three billion base pairs of DNA. We've cut it into smaller gene-containing fragments which are ligated into the small self-replicating molecule called the plasmid.

At the end of this procedure we have what is referred to as a recombinant DNA molecule; a foreign DNA fragment or gene, cloned into a plasmid vector. As the bacteria divide, more and more of the cloned gene is produced. So within the bacterial cell, we have two naturally occurring forms of amplification. The plasmid, now containing the foreign gene, can replicate itself, so it exists in multiple copies. And then as the bacterial cells divide, the plasmids also are reproduced.

Why is that important that we have multiple copies of the cloned gene? In order to analyze and manipulate cloned DNA, we need thousands or even millions of molecules. So it isn't sufficient to have a single cloned piece of DNA in order to analyze it and manipulate it in the methods that will be talked about next. Recently, a new technique has been developed to produce multiple identical copies of a segment of DNA without gene cloning and this process is referred to as the polymerase chain reaction or PCR for short.

PCR is a way to copy as little as a single molecule of DNA into a million or more molecules in just a few hours in the laboratory. This represents a way to circumvent the lengthy cloning procedure and directly amplify the DNA of interest *in vitro* or in the test tube.

DNA molecules can be melted with heat to produce two mirror image single strands. These single strands then can serve as templates for the synthesis of new DNA, and a mirror image of the template is produced during DNA synthesis. So that one DNA molecule has now been copied into two identical DNA molecules.

The amplification that takes place during PCR is a process that increases logarithmically. As we cycle through replicate rounds of that procedure, the two molecules would be reproduced into four, four into eight, and so on. If you do the math, you'll realize that thirty cycles of this procedure would result in over a million copies of the same DNA molecule. So, starting with as little as a single molecule, we can produce millions of copies, identical copies of that molecule in the laboratory.

Applications of this technique, which, by the way, earned the discoverer a Nobel Prize this year, will be discussed by Dr. Howett in a few minutes. Once enough of a gene is obtained, either through PCR amplification, as I just described, or through cloning, it can be analyzed and manipulated. In order to study the function of genes, it's important not only to be able to isolate genes from cells and manipulate them in vitro, but also to be able to return them to cells.

Returning cloned genes to cells is called gene transfer and there are a number of different methods that have been developed to persuade cells to take up DNA and integrate

this foreign DNA into their own genomes. This is despite the fact that millions of years of evolution have established barriers to foreign DNA uptake. Has anyone seen the movie *The Fly*? I know at least a few people have, because I show it to my genetics class and I see a few of my students in the audience.

If you've seen *The Fly*, you know that the evolutionary barriers that prevent the mixing of genes of one organism with another are necessary. I'm not suggesting that the creation of a fly-man is possible using recombinant DNA techniques, but I am simply saying that cells have evolved ways to protect their genes from invasion by foreign DNA. One interesting exception to this occurs during viral infection. Therefore, we have had to develop molecular tricks in order to get cells to take foreign DNA back into their nucleus and use it as if it was one of their own genes.

The first, microinjection, involves the direct injection of DNA back into the nucleus of a cell using microglass pipettes. Calcium phosphate precipitation is a technique which presents DNA to cells in a form that makes the cells think the DNA is "edible." I'll talk more about this in a second. The DNA is taken up or engulfed by the cells.

Liposome-mediated transfer involves making the cell think that the DNA is actually part of the plasma membrane, and it's taken up in that way. Electroporation is a technique that involves electric shock of cells that temporarily creates holes in the plasma membrane through which the DNA can enter the cell.

And finally, viruses have been used to transfer genes. This technique involves the utilization of genetically engineered viruses that have incorporated foreign genes. The virus infects the host cell and carries the foreign DNA into the host cell nucleus.

Let's look a little more closely at one of these gene transfer procedures. Here is a DNA-mediated gene transfer using the calcium phosphate precipitation method. We start with the cloned foreign gene in a plasmid. The other unique feature of the plasmid is that there is an additional gene contained in it that serves as a selectable marker, which just means that it's a gene that we can select for by some scheme.

The DNA is then put in a test tube with calcium phosphate and a chemical complex forms between these ions and the DNA. This complex is then added to the media of cells that are growing in culture and the DNA complex sticks to the cell surface.

Some of the cells actually take the DNA in as if it were food, and the genes that go in, the cloned genes, are transported to the nucleus, become integrated in the host cell genome and act as if they were host cell genes.

The way we identify those few cells that actually successfully pick up the DNA and use it as if it was their own is established through the selection scheme which will kill cells that haven't picked up the foreign gene on the plasmid with a selectable marker. So what we're left with is a clonal population or a cell line that contains the foreign gene that we're interested in.

One really important extension of the gene transfer techniques that I've just described is in gene therapy, and Dr. Howett will address that issue next. I think I'll stop here and entertain questions.

Member of the Audience: Does the size of the DNA you want to transfer mandate what choice of method you use to transfer?

Dr. Roberts: In part. There are practical limitations, there are technical limitations to the size of the DNA that can be transferred. Plasmids carry fairly small DNA fragments. Viral vectors can use larger segments of DNA. But it's an interesting problem, because, as we already heard, genes vary dramatically in size. And so when we talk about issues of gene therapy, where we wish to replace a mutated gene with a normal gene, there are some limitations in terms of the size of the gene and what can technically be accomplished at this point.

Member of the Audience: We've been recently reading about some development which may speed up the entire genome project, in terms of identifying all the genes. Is anything you've just talked about involved in that? Is PCR involved or what is the development that's going to speed up the process?

Dr. Roberts: I think that the main development has been in the automation of the sequencing technique, the techniques that we use to actually read the DNA alphabet. That whole process which was illustrated with the little gel slide and the different fragments of DNA is actually a fairly time-consuming process and only short stretches of DNA can be read in any given experiment, but the new techniques have been developed to automate that process so that we can continually read DNA sequence for long stretches at a time. PCR has also been an important technique in isolating certain segments of DNA.

Member of the Audience: When you talk about the DNA alphabet, what is it that we're reading? What do those letters represent?

Dr. Roberts: The letters represent the composition of the double-stranded DNA. They're referred to as bases and there are only four. That's why we said that the alphabet of DNA is just four letters long. It's that sequence of those four letters that specify the gene. Ultimately embedded within that alphabet are triplet messages that specify the amino acid subunits that go into the proteins that are produced, but it's that linear array of those four letters that is the DNA sequence that contains the information for each gene.

Member of the Audience: But the letters themselves, for instance, the A. I mean is that identifying a chemical compound or a group of chemicals?

Dr. Roberts: Yes. The A stands for adenine and it's a nitrogenous base. It's part of the structure of DNA.

Member of the Audience: Occasionally in a clinical setting, I hear people talk about the PCR product as though they're going to get different information from it than they would, say, be getting from a Southern Blot. Can you address that?

Because I always looked at PCR as a convenience for amplifying, so that you can look at things more quickly and get results for patients more quickly, but they seem to be using it in a context that you're going to hear something else from the PCR.

Dr. Roberts: Yes. I think that's a good question and I think I'll leave that to Dr. Howett to discuss. There are problems with PCR in terms of the fidelity with which the region is amplified and it's possible that mutations can be incorporated into the amplified DNA. I'll leave that for now.

Member of the Audience: In each DNA molecule, each pair that exists, is there a left-to-right sequencing? How do you know which one you start with?

Dr. Roberts: That's a good question. The genes embedded in the DNA, on average, are in the order of 10,000 base pairs in length I said that there are 3 billion base pairs of DNA in the human genome, and we guess that there are 100,000 to 300,000 genes. That means there's a lot of other DNA in the nucleus of the cell that's not specifying the information for a gene. And there are cues embedded within the DNA sequence that say to the machinery that uses the genes, this is the start of the gene and this is the end of the gene.

The genes themselves are first copied into RNA, and that copying proceeds in an unidirectional way on one of the DNA strands. So, there is a polarity to genes and a strand preference. That's not to say that all the genes in the genome are on the same strand. In fact, they aren't. They're in different directions on different strands, and they're embedded in this matrix of DNA, most of which we don't know what its function is. I'll now turn the discussion over to Dr. Mary K. Howett.