PROFESSOR: Good morning.

AUDIENCE: Good morning.

PROFESSOR: All right. So today-- I haven't seen you in a while. Anyway, today, we're going to turn back to our picture, function gene protein. We filled in genetics. We filled in biochemistry. We've now got the connection between gene and protein through molecular biology. We know gene encodes protein. We know central dogma, DNA goes to the RNA goes to the protein. We know all that in theory.

In fact, people knew this by the middle of the 1960s. People were so excited that they understood the idea of how genes give rise to proteins through transcription and translation, they read the genetic code, that they declared victory. Some of them said, done with the secret of life. Let's go on and do the brain. That was actually the thinking of a lot of the great molecular biologists in the late 1960s. Let's go do the brain.

Why did they say such a thing? Well, because they thought they were done with the problem. They thought that once you knew in principle how a gene gave rise to a protein, you could do it.

But in practice, nobody could read a single gene. Nobody could even identify a single gene. Maybe that's why they went on to say, let's go study the brain, because they actually weren't sure what they could do past that point.

All right. So wait a second, wait a second. I said nobody could even purify a single gene. Didn't we talk about purifying the genetic material? You're supposed to say yes at that point. Yes, right? We talked about that. Avery, McCarty, MacLeod-- we

purified the genetic material. We did it by using this assay of transforming.

So what do I mean by we can't purify a single gene? What I mean is that we can purify the hereditary material away from everything else, but we get all of it together. We don't get individual genes separated from each other. We get the whole mixture of all the genes, all the genetic material.

As a biochemist, how are we going to ever separate the gene encoding-- oh I don't know, ARG1, our favorite gene for arginine biosynthesis-- from the gene encoding ARG2, for example? What kind of biochemistry can we do to separate these two genes? Do they have different biochemical properties? What's so different about ARG1 and ARG2? What's their different biochemical properties?

They're both DNA. They have exactly the same building blocks, slightly different order. You think you have a purification procedure, I'm going to run it over some column and separate it by something that's going to separate ARG1 from ARG2?

No. From the point of view of a pure biochemist, they look exactly the same. All the different genes have the same biochemical properties. How in the world would we ever purify ARG1 from ARG2, or in the human, the gene encoding hemoglobin from the gene encoding collagen from the gene encoding keratin from the gene encoding anything else? Think about it. That's a tough problem.

There is a brilliant solution that arose in the 1970s to how we could purify the individual genes away from each other. But it's like no other piece of biochemistry anybody had ever seen before. It has a totally different principle behind it. Because it isn't just fractionating things according to their biochemical properties. It involves something else. And it's called cloning. It is called cloning. Molecular cloning.

You see, the problem is this. The human genome-- how big is the human genome? How many bases? Three billion bases, three times 10 to the ninth bases, right? How big is a typical human gene? A typical human gene might be 30,000 bases. How big is a typical mutation that we might want to find in a typical gene, like causing sickle cell anemia? One base.

We've got to purify out genes that are one part in 10 to the fifth and mutations that are one part in 10 to the ninth or so. And how are we going to do that?

Well, the trick is this. I'll give you the quick overview, and then we'll spend today looking at it.

Step one is we cut up our DNA. We cut DNA at defined sites, and we then paste the DNA to distinct molecules called vectors. These vectors have a cool property, that when you take a vector and you insert in it a piece of DNA, that vector is able to grow in another organism.

You then transform the DNA-- that's transfer, we use the word transform the DNA-- into something like E. coli, where you get your little vector in there. It grows within E. coli, and as E. coli divides, it makes copies of itself. And then you select those bacteria that have received, that have been transformed, grow them up on a petri plate so that you have little colonies. And then you screen the colonies.

Now, what do I mean? We cut the DNA. We paste the DNA. We transform the DNA. We select the bacteria that have been successfully transformed. And we screen the resulting colonies to find what we're looking for.

Now, notice-- that amazing trick here is when we cut up the DNA into single molecules, lots and lots of single molecules, and we paste them into vectors, and we transform them into bacteria, each one of those bacteria gets exactly one molecule, give or take. It gets one piece of human DNA. We then spread them out on a plate and they grow up, and each one grows up copies for us of individual pieces of DNA.

That is so cool. Because we've just accomplished biochemical purification. It's not based on any different properties of the individual molecules. It's based on the fact that we dilute them, in effect. They're diluted, and one molecule ends up in each bacteria. So they're purified in that sense. And then when that bacteria grows up, everything it grows up is a pure copy, a copy of that single piece of DNA that went into it. That's a different kind of purification.

When I'm done-- and we'll go through this whole process. That's the point of today's lecture, is to go through the whole process. When I'm done, I have bacteria spread out. And right over there, one of these guys has ARG1, and one of these guys has ARG2, and one of these guys has ARG10. Now, admittedly, I don't know which one has which, but I've accomplished the purification. I'll then have to figure out how to screen and find out which one is which, but I have separated the molecules away from each other by this process of cloning, diluting them in a way-- one molecule per bacteria-- and growing them back up.

I could do this for anything. I could dilute proteins down into test tubes that had one protein molecule per test tube, and I could say I've accomplished purification. The problem with it is I have no way to replicate those proteins to get a meaningful amount of it. But when it's DNA, and I've put it back into a bacteria, I have a way to grow it back up. And that's why this trick works-- is because DNA is a molecule that can replicate. No other molecule has that cool property. And so you can pull this off for DNA.

All right. Now we have to dive in to understand how this could possibly work.