**Biotech 101: The Science That Drives Biotech**

**Dr. Collins Jones:** Hi. Welcome to Biotech 101: The first in a series of webinars about biotechnology. In Biotech 101 we’re going to discuss the science that drives biotech. My name is Dr. Collins Jones and I’ll be the presenter for this series of webinars. Let’s get started.

**Objectives**

In Biotech 101 we have a number of objectives that we’d like to accomplish. First, we wish to define biotechnology. Next, we’d like to able to draw a basic cell and label all of the different components inside the cell which are called organelles. Each organelle has a unique function so we want to know what each of those different things does. Finally, we’re going to get down into the structure of DNA (Deoxyribonucleic Acid) which is at the heart of all cellular functions. We want to understand, then, how DNA actually works. And then from there, we’re going to move into how the information contained in DNA is converted into a protein. We then want to understand the correlation between the shape of a protein and how a protein functions. From there, we’re going to talk about DNA mutations, what causes them and what the effect of those mutations are, because ultimately, a number of DNA mutation can lead to a disease. Finally, we want to understand, again, how differences in a person’s DNA can influence diagnosis and treatments of various diseases. And from this, we’re going to drill down into how we can use the information attained in DNA to develop an approach to personalize medicine.

**Biotechnology Is**

Let’s start by defining biotechnology. Biotechnology is the use of cellular and biomolecular processes to solve problems and to make useful products. Cells are at the heart of biotechnology.

**The Basic Unit of Life: The Cell**

So let’s look at what a cell is. A cell is defined as the basic unit of life. A cell is the smallest living thing. Viruses are not alive by this definition, but cells are. So let’s just take a minute here and look at the structure and function of all the different parts of the cell. What you see initially here is a cell membrane. So the cell membrane encloses all the other components of the cell. The cell membrane is composed of what are called lipids, specifically phospholipids. And the definition, the most basic definition for a lipid is that it won’t dissolve in water which, of course, is a very good thing because if you think about it, when you take a shower in the morning or walk in the rain, if you were water soluble, you would dissolve. So lipids prevent us from dissolving in water. Imbedded in the lipid membrane are what are called proteins. So this is a representation of a protein molecule.
Proteins are how the cell communicates with the outside world. And we’ll talk a little bit more about this at the very end of this review of the cell. Contained in what’s called the nucleus of the cell is DNA (Deoxyribo nucleic Acid). This is where all of the information to run the cells is stored in the DNA. And we’re going to spend a little bit of time discussing the structure of DNA once we understand everything there is to know about a cell. We then have what’s called the nuclear membrane which encloses the DNA inside the cell. We have the endoplasmic reticulum. And these little dots here are called ribosomes. And that’s really what we want to focus on. Ribosomes are protein factories. So in a ribosome the information to make a protein that’s contained in the DNA is converted into the protein itself. And proteins are what actually carry out all of the functions inside the cell.

Next, we have mitochondria. Mitochondria is where energy is produced, because it takes a lot of energy to run a cell. This is why you have to eat everyday because the food that you take in is converted in the mitochondria into energy. Mitochondria are becoming important players in a number of diseases. And one way that we can start to diagnose or detect certain diseases is by looking at the changes and the number of mitochondria. If you just take a second and think about it, what tissues in your body do you think have a lots of mitochondria in them? Okay. Time is up. So the two tissues that have lots of mitochondria are one, your heart and the other is your brain. Your brain consumes 20% of the sugar that you take in everyday. And a number of neurodegenerative diseases like Parkinson’s and Alzheimer’s disease, the number of mitochondria is decreased in the patients who have those diseases, has decreased in their brain.

Next, we have what are called microtubules, spindle fibers and cytoskeletal components. These do two things; they give structure to the cell. Believe it or not, each cell in you body has it own little skeleton and that gives it a three dimensional shape. And in addition, these little cytoskeletal components act like little transport pathways to move things around inside the cell. So we can move a molecule from one place to another in the cell by the cytoskeletal.

Next, we have what are called the Golgi. And in case we want to move something outside of the cell, for example, we make a protein on the ribosome using the information from the DNA. And we don’t want that protein to be in the cell. We need to release it to send it to some place else in your body. That protein is then sent to the Golgi apparatus here. And the Golgi packages up and delivers this protein, sends it the cell membrane, where it is then released from the cell. And the release of a protein from the cell is known as secretion. So a secreted protein is a protein that was originally made in the cell, then packaged in the Golgi and sent out for use elsewhere in the body.

Finally, we want to come back to this concept of proteins in the cell membrane because the cells in your body don’t act alone, they act in concert with one another. And this sort of is the whole concept behind systems biology, that you are a giant system of cells and all of these cells have to interact with one another in order for you to function. And if you think about that, that’s a pretty remarkable fact. Human beings have roughly 12 to
14 trillion cells as an adult. Two hundred different kinds of cells and they all function in your body for about 80 years.

So in the membrane there are what are called receptors. Receptors are a special class of proteins that receive information. The information is handed to the receptor or given to the receptor by the small molecule called a signal molecule, which is also sometimes called a ligand or ligand. And so what happens here is the ligand binds to the receptor. And this is analogous to the lock in key model. So the ligand is sort of like the key and the receptor is the lock. Once the ligand has occupied the receptor, the receptor then sends a message through the cell, transmits the message given by ligand to the nucleus. And that tells the DNA inside the nucleus which proteins to make. And accordingly, once those proteins are made, the cell then begins to respond and do what it needs to do.

**All Cells Contain DNA**

All cells contain DNA. So it’s very important to the function of a cell. Let’s talk about what DNA is and what it does. Any living cell has DNA in it. The only cells that don’t have DNA are red blood cells. And you might say, “How come not red blood cells?” Because if you think about it, what the red blood cells do, they transport oxygen. And oxygen is highly reactive. So we don’t want the molecule of life, the molecule that contains all the information around the cell, right next to highly reactive oxygen because bad things could happen. And we’ll talk about that a little bit later when we talk about mutations.

**Life’s Building Blocks**

One important thing that we’re going to emphasize in the next slide is the fact that DNA is just a chemical. And amazingly, every living organism has exactly the same chemical structure of DNA. So whether we talk about a lowly bacterium like E. coli, a plant, like a tulip or a human being, the chemistry of DNA is exactly the same in every one of those living organisms. The implication of that, as we’ll see in our next webinar, the implication is that DNA from one organism can be read and understood by another organism.

So let’s look at what the chemicals are that make up DNA. The building block of DNA is called a nucleotide. So DNA is simply composed of nucleotides. And let’s say, again, before we go into the structure of a nucleotide, what DNA stands for. DNA stands for Deoxyribonucleic Acid. Big name here, all right? So let’s break that down. The first part of a DNA molecule is a sugar and that sugar is called ribose. That’s where the ribo comes from. The deoxy means without oxygen. And if you notice on the sugar, there’s an OH group here. And normally on the sugar ribose there’s another OH group here. But in DNA, nature has removed the second OH group and just put hydrogen here, so we have deoxyribose, that’s a ribose without one oxygen. So every single DNA molecule, no matter what cell it’s in, has this sugar deoxyribose.
In order to be able to locate different positions on the sugar, scientists assign a number to each relative position. And you’ll notice there’s a little apostrophe here. This is read as prime. So this is the one prime position, the two prime, the three prime, the four prime and the five prime. What’s important to us is the three prime and the five prime because this is how one nucleotide connects to the next in a DNA molecule. And we’ll see that on the next slide.

The next component of a nucleotide is the group called phosphate. And so what you notice here is there’s a phosphorous in middle of this. Phosphorous is the energy atom inside a cell. So what phosphates do is provide the energy to allow one nucleotide that connects to the next. And if you think about this for a second, what I’m implying here is that DNA is just a string of nucleotides. And we’ll illustrate that on the next slide.

But let’s just summarize for a second here, we have sugar deoxyribose and a phosphate. Every single nucleotide has these two components and they’re exactly the same in every nucleotide. Every nucleotide has deoxyribose. Every nucleotide has a phosphate. And that’s every nucleotide whether it’s in a bacteria cell, a plant cell or an animal cell. So what becomes different here is what’s called the base. The base is really where the information is contained in a DNA molecule. There are four bases. There are only four bases. And these are guanine, adenine, thymine and cytosine.

Here, we have the chemical structures of each of the base, okay, for each base, and hopefully you won’t have to memorize these for your place of work. But what we want you think about here is that when we see a sequence of DNA which is usually represented by letters G, A, T or C, what these letters really represent is this chemical structure, a base. And again, the chemical structure of these bases is exactly the same whether it’s a bacteria, a plant or a human being. So in the structure, in the nucleotide, what changes is the base. We have the deoxyribose and the phosphate, but what makes one nucleotide different from another is which base is located or attached to the one prime position, whether it’s a G, an A, a T or a C.

Now, this is a lot of chemistry here and scientists don’t want to have to throw all these structures out all the time so we begin to abbreviate them. And our first layer of abbreviations here is indicated by this schematic picture down here at the bottom. So the pentagon here is the deoxyribose. The circle is the phosphate. And then the base, instead of writing the term base in this box here, we’re going to put one of these letters; G, A, T or C.

**Putting the Building Blocks Together**

So how do we connect nucleotides? Remember they’re connected, the phosphate to the sugar. So here we have our first nucleotide and notice the base is thymine. Then we have cytosine, adenine, guanine and so forth. And so what we notice is that we’re connecting five prime to three prime here, and the phosphate is the intermediate. So the sequence here of sugar phosphate, sugar phosphate, sugar phosphate, that’s called the
backbone of the DNA. And you notice it’s identical for every nucleotide. What changes, if you’ll notice, is the order of the bases. So we have a different base with each nucleotide. And then you might ask yourself, “Well, how does this convert into information?” It turns out that it’s the sequence of the bases where the information is stored in DNA.

So if we just focus here on the bottom three nucleotides A, C, T, we’re going to pretend that this is actually a word in English. So if we look at A, C, T, what does that spell? This isn’t a trick question, just think about what that spells. It spells ACT, all right? If we change the sequence, that is if we change the order of the letters and we make this C, A, T instead, that spells what? CAT. Two completely different things. So hopefully we’re able to illustrate to you that it’s the sequence, the order of the bases that defines the information contained in DNA.

Now, it turns out though that DNA doesn’t occur as a single strand, that’s what this is called, a single strand of DNA. DNA actually has a matching or complementary strand in which guanine always binds to cytosine, that is G always matches with the C and the T always matches with an A. And this is called base pairing. And T to A or G to C are called complementary base pairs. The implication of this is very profound because it means that if we know the sequence of one side of the DNA by complementary base pairing, we know the sequence of the other side of the DNA molecule. So we only need the sequence of one side of the DNA in order to know the whole structure.

This is what Watson and Crick deduced in the 1950’s. The fact that DNA is double stranded and it forms a helix and this is idea of complementary based pairing that nature designed chemically the fact that G’s can only fit with C’s, and T’s can only fit with A’s. And you might be asking what happens if we make a mistake, what happens if a T fits with a G? Well, the apparatus in your cell’s able to detect that and in some cases, it corrects it the right way and in some cases, it makes a mistake when it tries to correct it and that’s where we get a mutation.

**Genetic Flow**

So the next question is – hopefully, we now understand what a DNA sequence is, it’s just the order of the nucleotides. How do we take this information and convert that information into a protein? Well, let’s use an analogy here. A specific sequence of DNA that is a specific order of letters, G, A, T, C, is called a gene. And a gene is really like a recipe or a set of directions in order to make a protein. So if you think about this, your DNA is actually subdivided into a number of different genes. Human beings have roughly 25,000 different genes. That means that there are at least 25,000 different proteins that could be made from each set of directions. So basically what we do in your cell is we make a copy of the set of directions, and that copy of a specific sequence of DNA known as a gene is called RNA, and this is called a transcript because we’ve made a copy of this.
So let’s just sort of go back and review this again. A gene is a specific sequence of DNA that contains the information or the directions to make a protein. You want to be really clear about this. A gene is not a protein; it’s just the directions to make a protein.

Now, what your cells do is they don’t use all 25,000 genes at once. They just use select genes based on the information to get from the ligand that is connected to the receptor and the cell membrane. So cells don’t use all 25,000 genes at once, they use specific genes depending on the information they get, depending on what the ligand tells the nucleus which pieces of DNA the nucleus was told to use.

So a gene is then copied into what’s called RNA, because we don’t want to deal with all those 25,000 genes at once, we just want to deal with the genes that we need to. So we copy specific genes into what’s called RNA. And I want you to notice here that RNA is a copy of a gene but that it is single stranded because we only need one side, one half of the information in order to make a protein. The RNA is made in the nucleus of the cell; it’s then shipped out and taken to the ribosome. And the ribosome is kind of like the decoder ring. The ribosome reads the RNA and then it assembles the protein. Proteins are made of molecules called amino acids. So essentially what the sequence of the DNA does that then becomes the RNA transcript, essentially what that does is it tells the ribosome which amino acids to connect one to the next.

So let’s then assign some formal names to this process. The process of copying a specific gene into a corresponding piece of RNA is called transcription because we’re copying. So the process of reading the information contained in the RNA transcript and converting that information to a protein which occurs in the ribosome is called translation. So we have transcription and translation. DNA has the gene on it, RNA is called the RNA transcript and the RNA transcript is sometimes also known as MRNA or messenger RNA. And then we have the protein and the protein is made up of amino acids.

To illustrate this, then, using actual sequences, let’s go this slide. So here we have a gene, a specific sequence of DNA. Notice that we have our nucleotides. We copy one side of the DNA and we get our messenger RNA transcripts. There are a couple of subtle little points that we want to make about our messenger RNA here. First, you notice that there are no T’s in that messenger RNA. Instead of a T we have a U, that’s how the cell knows that we’re dealing with RNA. U is our fifth base uridine, but U only occurs in the messenger RNA. So every place there’s a T in the DNA sequence, in the corresponding RNA sequence there will be U.

The second thing you notice is that in the RNA we’ve actually broken our transcript up into sequences of three bases. We have AUG, that’s one sequence of three. We have GAU, that’s another sequence of three and so forth. Three bases are known as a codon. And each codon codes for or tells us which amino acid we’re going to use when we make the protein. So AUG is the codon for the amino acid methionine. GAU is the codon for the amino acid of asparagines.
So the ribosome reads each codon, brings in the correct amino acid and connects it to the preceding amino acid. And then we get a sequence of amino acids which becomes a protein. How does it know when we’re done? There are stop codons. So UAA is sort of like the period at the end of the instruction, it says, “Stop we’re done.” And at that point, the ribosome disassembles, the completed protein is released into the cell and then it can go and do its job.

**Proteins**

So what are proteins? Once again, proteins are made up of amino acids and they carry out all the functions in the cell. There are several classes of proteins. Some of the most important classes include enzymes. Enzymes carry out chemical reactions. Your cell performs thousands of chemical reactions every second. So what do we mean by chemical reactions? Well, one chemical reaction is actually copying the DNA into RNA, that’s carried out by an enzyme called a polymerase. And one hint to know that you’re dealing with an enzyme is that all enzymes end with a suffix ASE. Enzymes break down the food that you eat, meats, sugars, lipids and convert it into small molecules that, in turn, can either be converted into energy or reassembled into other molecules in your cell. So enzymes do lots and lots of things in your cell and they often become important targets for small molecule drugs that we’ll talk about in our third webinar here.

Receptors, you’re already experts in receptors, right? Receptors are proteins generally imbedded in the cell membrane that receive information from outside the cell and then relay that information inside the cell. So, receptors have to interact with a signal molecule which is often known as a ligand. Examples of receptors are neureceptors, growth factor receptors. There are hundreds and hundreds of receptors. Structural proteins give shape and infrastructure to the cell. These are what make up the cytoskeleton. Structural proteins that you’ve probably heard of, one might be actin, the other one is actually a biotech product, collagen.

And then antibodies which are a very important protein not only in your body but as a product of a biotech industry are antibodies. Antibodies are proteins that recognize and bind to other proteins. And in so doing, they send a message to your immune system that this protein or this virus or bacteria doesn’t belong here, and targets for destruction.

**Introducing Mutations**

Now, we’ve made a big deal about saying how important the sequence of the DNA is. I want to reemphasize the fact that the DNA sequence, whether it’s ACT or CAT, each of those determines a different amino acid. So hopefully it makes sense to you that if we change the sequence of the DNA, we then change a codon sequence which potentially could change the amino acid. And if we change an amino acid in a protein that can alter the function of the protein and that could lead to a disease. So, a change in the sequence of the DNA is called a mutation. And there are a number of ways that mutations can arise. One way that a mutation can arise is when we have replication errors.
Mutations Caused By the Environment

The next type of way that we can create mutations besides replication mistakes are through radiation or through exposure to chemicals or toxins. So if we look at each of these in turn, replication errors are simply recopying our DNA. Your DNA is composed of 3 billion nucleotides, A’s, G’s, C’s and T’s. Your cells take roughly 20 hours to divide. So if you just think about this, if I ask you to go home tonight and copy 3 billion letters in the correct sequence and come back tomorrow morning, think you’d make a few mistakes? And the answer is probably yes. So occasionally, your cells make mistakes as well but they have an editing system and they can correct it, but sometimes they don’t get to correct it.

So the types of replication errors we have are substitutions, instead of putting an A in there, we put a T or instead of putting an A in there we put a G. And then on the corresponding complementary strand, if we’ve accidentally put a G in where an A should have been, then we get a C. So we now have a GC where there should have been an AT. So that’s a substitution error.

Deletions, just forgot to copy some of the DNA sequence, that could be one or two bases or it could be several thousand bases.

And insertions, we’ve copied the same sequence of DNA more than once. So anyone of those can cause a problem. But sometimes it doesn’t cause a problem at all because it’s in a part of the DNA that’s not important to the functioning of the cell.

Radiation, you’re simply getting exposed to energy and that energy can chemically alter the DNA, can cause breaks in it, can cause erroneous repair of mechanism to occur. So radiation is another source of mutations. And then being exposed to chemicals, these actually react with your DNA. Similar to radiation, they can cause damage to the DNA and that damage can sometimes be repaired correctly and sometimes not. But for any one of these mechanisms, whether it’s a replication error, exposure to radiation, exposure to a chemical or toxin, the end result could be that the DNA sequence is altered, we have a mutation. And that in turn, alters the amino acid sequence of the protein and now the protein won’t be able to function properly.

The Nature of Mutation

So let’s give some examples here of these different types of mutations. If we look at substitution errors, deletion errors and insertion errors, we’re changing the base sequence of the DNA. So we start of here at the top sequence, this is what we call the normal or wild type sequence. This is a sequence that’s correct whether there isn’t mutation, we’re going to have a disease, and there should be no ill effects if you have this sequence. You’ll see in a few slides that there is no such thing as normal but this is the sequence that works properly.
So in our first type of mutation we have what’s called the substitution error. And here you notice that’s what happened is we’ve accidentally, when we’re copying the DNA, we’ve accidentally substituted a C in place of the T. So now instead of having the codon ATG, we have the codon ACG. And if we look in the codon table, the little decoder key, we could see that that may change what the amino acid is and that will change the function of the protein.

In our second example here, we’ve deleted the base. So we just forgot to copy the T. And the implication of that is very far reaching because here we had the codon ATG, now we’ve omitted the T so the next set of three instead of being ATG, it’s going AGT and every codon after that is going to be altered. So instead of altering just one amino acid, we’ve now altered all of the amino acids that follow our deletion. And that’s a huge mistake that can sometimes be a fatal error to the cell.

And then in our last example we’ve just accidentally inserted another base. And here we’ve actually inserted two bases, a G and a C. And again, in terms of our idea of a codon, now, instead of having ATG we’ve put in AGC, a completely different codon. And again, every codon that follows has also been altered so we’re now going to get a completely different amino acid sequence which makes probably, in most cases, for a non-functional protein.

So when we talk about mutations, so far we’ve indicated that they can be deleterious, that is they damage gene function, we make the wrong protein, we get a disease, this is a bad thing. But there are also a number of mutations that are neutral and this could be because the change in the DNA sequence occurred somewhere in the whole sequence of DNA and the whole 3 billion letter that occurred somewhere were there is not gene. It’s just, if you will a nonsense sequence of DNA or sequence of DNA that doesn’t contribute to the function of the cell. So if you get a mutation there, who cares? Nothing really happens. And some mutations can actually be adaptive. They may actually confer an advantage. This may be a mutation that makes you more resistant to this year’s flu virus. This could be a mutation, for example, between lots of mammals, some mammals see in color, other mammals only see in black and white. So mammals that need to function during the day time need to see in color. That’s the result of a mutation.

**Genetic Variation**

So now that we understand what mutations are and the potential consequences, the idea here is that mutations lead to what we call genetic variation, that is how one person is slightly different from the next due differences in their DNA sequence. And it turns out that through sequencing multiple human beings’ DNA, scientists have predicted that there’s only 0.1% difference in the DNA sequence from one person to the next. So that means that essentially we are all 99.9% identical when it comes to our DNA. But that 0.1% difference is enough to make us different in appearance from one another and also in some of our other physiological properties, which we’ll discuss in a second. Most of
the difference between individuals is what are called single nucleotide polymorphisms or SNPs. So SNP is pronounced S-N-I-P.

So now that you understand the nature of what a mutation is and the consequences of it, let’s look at this in a little more detail. It turns out that if you sequence multiple DNA strands from human beings that there’s only a 0.1% difference in the DNA sequence between individuals. So that means that we’re 99.9% the same when it comes to our DNA. Just 0.1% of our DNA sequence differs from one person to the next. But these differences can have profound effects as we’ll see in a minute.

Most of the sequence differences between human beings are classified as what are known as single nucleotide polymorphisms or SNPs, that’s how they’re pronounced. So if we just sort of dissect the word here, single means one, nucleotide is our A, G, T or C, and polymorphism just means many different shapes. Because if we put this in a context of using the information in DNA to make a protein, what happens here is if we change a single letter in our DNA sequence, an A to a T or an A to a G, that changes the amino acid sequence of the protein. And by changing the amino acid sequence in our protein, the shape of the protein can change, and that’s where we get the polymorphism from.

And that’s what we’re illustrating here with the mister potato heads. I really actually kind of like this part of the slide. But this is where we’re illustrating with Mr. Potato Head that these slight differences in shapes are due to single base changes. So what is the effect of that physiologically on how you function? Well, the obvious effects are things like differences in eye color, blue eyes versus brown eyes, hair color and so forth. But from a healthcare point of view, the differences are even more profound and even more important. And so we’re looking at things like your susceptibility to disease, how you respond to environmental factors, how you would respond to certain drug treatments or to vaccines.

And if I could just sort of give an example of this, let’s look at disease susceptibility. We all know that we differ in our susceptibility, for example, to the seasonal flu virus from year to year because each year the flu virus is a slightly different version of the previous years’ flu. And how does this arise? Well, the flu also has a DNA sequence, actually it’s RNA in the flu, but the idea is the same. The sequence of the flu virus changes giving the flu virus a slightly different protein shape and then in order for the flu to infect you, the virus has to land on a receptor on your cells. Believe it or not, your throat cells have a receptor protein in their membranes that the flu virus recognizes and attaches to. And so the idea here is that if you have a specific amino acid sequence for the flu virus receptor in your cells, then the flu virus can bind to that and infect you. If the sequence is a little different, then the virus may not be able to bind as well and then you’re less susceptible to the flu.

So for example, if this year you don’t get the flu, the sequence of your flu virus receptor on your cells won’t recognize that and so the flu virus can’t attach and infect you. But if the person next to you gets the flu, then their sequence, they have a SNP, for example,
they have a single nucleotide polymorphism, a one base change in their DNA that’s changed slightly the shape of the protein on their cell’s surface that the flu virus is able to bind to and they get sick. But next year it might be the reverse. Of course, there are other factors that influence your susceptibility to getting the flu, but that’s one of them and it’s attributable to these changes, these one base changes in specific sequences of DNA. So it would be of interest to begin to sort of identify these sequences and know how important they are.

**Genetic Basis of Disease**

To illustrate this a little further and maybe in a little more depth, let’s talk about the genetic bases of disease. And what we’re really talking about here are diseases that are inherited, that are passed down from one generation to the next. So we classify genetic diseases into two broad categories. The first are called monogenic diseases where one gene causes the diseases. And the second category is polygenic diseases where many genes contribute to the disease.

So in a monogenic disease the example we have are Sickle Cell Anemia where we substitute an A for a T. Cystic Fibrosis where there aren’t substitutions, but in this case we simply delete three bases, we delete three of the nucleotides, the A’s, G’s, C’s and T’s. And that deletion dramatically changes the ability of a particular protein in your lung cells, called the chloride ion channel, to function. And so when that chloride ion channel doesn’t have the right shape because it’s missing one amino acid, then what happens is the lung cells begin to die, as they die they break up and release DNA into your cell and that DNA clogs up your lungs and essentially you begin to suffocate. And then we have Huntington’s disease where we have what’s called an insertion, we’re actually adding extra bases, extra nucleotides. And in this case, it’s a very specific insertion called a trinucleotide repeat. And so tri means three, our nucleotides are A’s, G’s, T’s and C’s and repeat means it’s just done over and over again. And so in Huntington’s diseases, unique to this disease is what defines it, is this GAC trinucleotide repeat. So there’s a guanine, adenine, cytosine nucleotides that are being repeated over and over again. So what that means is in a particular gene we see GAC, GAC, GAC, this happens in multiple times. And it turns out that if you have 30 or more GAC’s in a row, then you absolutely will get Huntington’s disease and that’s a disease, unfortunately, that is always fatal. So we know the cause of the disease but it’s just now that people are beginning to understand how this repeated sequence affects the proteins in the cell and then how that ultimately leads to damage to the brain and how people then die from that.

If we go to polygenic diseases, poly, again, means many. So these are diseases where many different genes contribute to the disease. Examples of these include cancer, heart disease, Parkinson’s disease. And I can remember back when I worked at the Cancer Institute, for example, in the 80’s, the hypothesis at the time was that people thought maybe two, three genes cause any particular type of cancer. Identify those genes and you
know what to develop your therapeutics against. But, of course, nowadays that’s changed. And it turns out that a number of cancers anywhere from 50 to 200 or more genes have been mutated as the cancer evolves. And so now there are a lot of different targets and the challenge becomes which of those genes are really the ones that you want to target for your therapeutic and which are just sort of minor contributors to the disease, same thing for heat disease, same thing for Parkinson’s disease which is a neurodegenerative disease.

But the other implication of this is because many genes have to be mutated at the same time in order for you to obtain the disease, knowing one gene that’s mutated doesn’t necessarily give us the whole answer. But we can use that mutated gene for its predictive value. And so what we say is, is that we have certain genes that indicate your susceptibility or the probability, the likelihood that you are to get a disease in terms of a statistical probability. So for example, the BRCA gene, there are several versions of that, and if you have a particular mutation in the BRCA gene, then the likelihood that you will get breast cancer increases. Just because you have that mutation doesn’t necessarily mean that you’ll get breast cancer but it means the risk factor is increased for it. So the corollary to this then is that a mutation in the BRCA gene alone is not sufficient to cause breast cancer. There has to be at least one or more genes that also have to be mutated at the same time before you get the full blown cancer. So, you know, if you go in and you get a biopsy and they say you have the BRCA mutation, it means that you’re likelihood of getting cancer is increased. It doesn’t guarantee that you’ll get cancer.

So from that it means that what we have to do is identify what the other genes are. That’s where the research opportunities come in. That’s what scientists are trying to do. What is the whole compliment of genes that gives rise to a particular cancer because it maybe in five or ten years when we find those other two or three genes, you go into the doctor’s office and all three genes have to be mutated at the same time in order for you to get breast cancer. So you go in and the doctor checks the DNA sequence of the BRCA gene and the other two genes and says, “Well, your BRCA gene has mutated but the other two genes have the normal DNA sequence so your risk is actually only 10% of getting breast cancer.” Whereas today if you went into the doctor’s office, all they can look at is the BRCA gene mutation and so the estimate is that if you this particular mutation your risk may go up 20% to 50% depending on the position of the mutation in the DNA sequence that codes for the BRCA protein. So it just points to the fact that we don’t have all the answers yet and we need to do lots more research in order to get down to what’s really important and what’s not important.

**Monogenic Disease: Sickle Cell Anemia**

Let’s just, then, expand a little bit on the idea of how a change in the DNA sequence can relate to a particular disease state. And the classic example here is the monogenic disease, sickle cell anemia. So in sickle cell anemia a mutation in a single gene gives rise to the disease. And the gene is the hemoglobin gene. So just to summarize again, the hemoglobin gene contains a sequence of nucleotides, A’s, G’s, T’s and C’s, that contains
the information to make the protein hemoglobin. And so what you see here is in the top part of the slide, we have the so called normal sequence for the hemoglobin protein, the DNA that codes for the hemoglobin protein. And so correspondingly, this DNA sequence is then transcribed into the messenger RNA because, remember, it’s the messenger RNA that actually goes to the ribosome and the messenger RNA sequence which was a copy of the DNA sequences then are translated into ribosome to make the protein hemoglobin. And so here at the bottom we have the amino acid of a particular sequence. Now, this is not the entire DNA sequence, RNA sequence or amino acid sequence for hemoglobin, we’re simply illustrating the part of the gene where the mutation occurs.

If we look at the sequence that corresponds to a person with a sickle cell anemia you’ll notice that we’ve changed a T to an A. This is a single nucleotide polymorphism. One base has been altered, the T for an A. And by changing this T to an A the amino acid is changed. So we go from a glutamate to a valine. These are two different amino acids. And when valine is added to the amino acid’s sequence of place of glutamate, it causes the protein hemoglobin to fold differently. It has a different shape to it.

Now, you have to remember that your red blood cell doesn’t have just one hemoglobin protein molecule in it. It has millions of hemoglobin protein molecules in it. And so now all of these proteins, all of these hemoglobin protein molecules are mis-folded, they have the wrong shape. And what happens then is because they have the wrong shape they begin to clump together, this is called aggregation. And as they begin to aggregate or clump together in the red blood cell that causes the red blood cell to change its shape and it assumes this characteristic shape of a sickle. So this is why it’s called sickle cell anemia because the red blood cell kind of resembles, under the microscope, a sickle.

Because these cells no longer have the nice round spherical shape as they begin to travel through your blood vessels, they start to catch on things because they’re not nice and round and they don’t bounce off and roll away. And particularly at the branch points, these sickle cells get hung up, they crash in to each other, the red blood cells then begin to break open, come apart, the hemoglobin comes out. Another major component of hemoglobin is the metal iron. Iron is actually what carries the oxygen around. So now the iron has been released, that gets into your blood stream and begins to damage the cells and then you’ll get all of the classic effects or sickle cell anemia and it’s also quite painful.

So the idea here though is that we know exactly which gene is responsible for the disease, we even know which nucleotide has been changed. But despite all of that knowledge, it’s still difficult to create an effective therapeutic, although some are now being tested clinically to remedy sickle cell anemia.

**Polygenic Disease: Cancer**
A more complex disease is cancer. So cancer, again, is a polygenic disease. And as cancer begins to develop in a person, multiple genes become altered or mutated. And so we’re just sort of going to look on this slide at the evolution, in general terms, of a tumor. So we start off with normal tissue, we then get what’s called an initiating event. This could be exposure to a cancer causing virus like the papillomavirus, this could be exposure to a chemical, this could even be exposure to too much sunlight, ultraviolet light, and you get a multiple sun burns.

And so what happens at random, at least that’s what scientists believe, is that there’s an initial mutation in one of your cells. So just one cell gets altered, the DNA sequence is changed and that then confers on the cell a growth advantage. And all that means is this one cell that was altered now begins to divide a little faster than all of its normal neighbors. The consequence of that increased cell division is the fact that we might get a second mutation because, remember now, we’re trying to copy all of this DNA in a much shorter time. So the less time you have to make a copy, the less time you have to proofread and correct the copies. So we accumulate a second mutation. But it may be, in some cases, that a second mutation is actually a good thing because the second mutation may, in fact, cause the cancer cell to die, it may be a mutation in a protein at the cell absolutely needed to survive and now that protein is altered and the cells don’t survive, the tumor basically goes away. You don’t even know that you had the cancer.

A second possibility is that the mutation could occur in a protein that’s present on the surface of the cell. And now that cell surface protein is different for the cancer cell than all of the proteins in your normal cells and your immune system is able to detect that difference and now your immune cells, you natural killer cells and your cytotoxic T cells combine and say, “Whoa! You look way too different. You must be a cancer cell. I’m going to kill.” And again, the tumor goes away and you won’t even know that you had any cancer.

But in the third scenario the mutation may add an additional growth advantage. It may confer an additional growth advantage to the tumor cell, and now the cell survives even better, it’s less susceptible to being killed, it may divide faster. But in some way, shape or form, this additional mutation gives the cell an additional advantage for survival. And now the cells continue to divide over the course of time, and we’re really talking here anywhere from a year to 20 years. The mutations begin to accumulate where we get 50 to 200 different genes that have had mutations in them, and the cancer cell. And now at this point we’re getting to the point where we have a tumor that you can detect by ultrasound or X-ray or some other imaging methodology.

But what’s interesting is that scientists recently sort of have confirmed is the concept that even though we have a tumor that may started from one original cell, because we’re getting all these mutations and the mutations are apparently occurring relatively at random, that within a particular tumor there are different populations of cells. That is there are cells with different combinations of mutations. So we say that the population of cells in the tumor is heterogeneous. So it’s not necessarily true that every cell in the
tumor is identical. There may be several different types of cells due to different types of mutations. And, of course, that adds to the challenge of treating the tumors because some cells may be susceptible to one treatment but resistant to another.

In any event, over time as this tumor begins to grow, it needs another source of nutrition. It can’t just steal nutrients from the nearby cells because its demands are too great. So another mutation that occurs and lots of the more developed tumors is that they begin to release a signal molecule called a VEGF, and this is a growth factor, and growth factors tells cells to grow and divide. And the VEGF’s specifically tells blood vessel cells to grow and divide. So VEGF is released from the tumors cells, they find their way to the closest blood vessels and then the blood vessel cells begin to divide and grow towards the tumor, and that way the tumor now has its own supply of blood and blood is, of course, what’s supplies nutrients and oxygen to any cells in your body to survive, including the tumor cells. So now the tumor cells have their own sort of, if you will, their own supply chain independent of anything else. And now they are getting nutrients and oxygen and they can begin to grow and divide even better. And this process of new blood vessel growth is called angiogenesis. The “angio” for the blood vessels and “genesis” is just new.

And so this process takes a couple of years to occur. And during that time, these newly grown blood vessels are fairly fragile and they’re apparently fairly easy to disrupt. But once the tumor cells sense, if you will that they’re being supplied with their own supply of blood, then they start to shut down certain genes, and become totally dependent on these newly formed blood vessels. And if you can knock these blood vessels out, if you can cause them to dissolve because they are fairly fragile, it turns out that a lot of times the tumors die because they can’t revert back and find a different way to acquire nutrients. And that type of therapy is called anti-angiogenesis therapy.

Realizing Goals

Now, here’s a challenge though, if you think about this, I mentioned earlier that all of the DNA in a human being is composed of 3 billion, that's billion with a B, 3 billion nucleotides A’s, G’s, T’s and C’s, and scattered amongst those 3 billion nucleotides are 25,000 genes. So the challenge is which genes and then specifically which sequences of nucleotides within the gene are responsible for a particular disease. And you have to admit that’s a pretty challenging task, almost insurmountable with today’s technology, to just fish through all that DNA and try to figure out which sequences are important and which are not. So a scientist came up with a strategy that’s been called Genome-Wide Association Studies or GWAS studies, GWAS, we just pronounce the abbreviation. And the idea here is that it turns out that genes and even mutations tend to be clustered in those 3 billion letters. So in other words, we don’t have to look at all 3 billion letters, we could look at the sequence of 10 million base pairs in to 25 million bases in. And in that particular region, scientists, just by analyzing a few people’s DNA, realize that that’s probably where some important genes are. And so we’re just going to look at that little
stretch of DNA instead of the whole 3 billion letters. That’s what Genome-Wide Association Study is.

And so what we do then is we isolate DNA from people who don’t have a disease and we analyze that particular section of their DNA. And then we select another cohort of people who do have a disease and we analyze that same sequence of DNA. And then we look for differences in the nucleotide sequence. Is it ACT or is it CAT? And mostly what we’re really looking for here are these SNPs, these single nucleotide polymorphisms or these one based differences, the A to T that we saw in sickle cell anemia for example. But this takes a lot of money. And as you might imagine, it’s kind of like looking for needle on a haystack. So getting this funded can sometimes be a challenge. But one such study that’s frequently been sited, one such study has been carried out and published in Nature in 2007 and it’s called the Welcome Trust Study. In this study scientists were looking at seven diseases and they used DNA samples from 14,000 people who had a particular set of diseases and they used 3,000 controlled DNA samples, these are people without the disease, that’s what a control is. So there, the controlled DNA is the so called normal DNA and then the disease samples should have a change or mutation in the DNA sequence that hopefully will correspond to the disease.

So they analyzed over 500,000 SNPs, over 500,000 single based changes. So this is high throughput study. We’re doing lots and lots of DNA sequences. And as a result of this, they identified 24 disease genes. One gene for bipolar disorder, one coronary artery disease, nine Crohn's disease, three rheumatoid arthritis diseases, seven for Type 1 diabetes and three for Type 2 diabetes. Which in one sense is very encouraging but in another sense out of 17,000 total samples, 500,000 analyses, that’s all we came up with. So it clearly points the need that we need to do a lot more samples and hopefully that we can find a more efficient technology for identifying which mutations, which changes in the DNA sequence give rise to a particular disease.

**Personalized Medicine: Therapeutics**

Now, how can we apply all of this? We can apply all of this in the rapidly emerging field of what’s known as Personalized Medicine and so we’re going to use a particular example here of breast cancer. And if you read the statement on the slide here, it says that, 25% of all breast cancers are caused by the over expression of a particular growth factor receptor, HER2. So remember, a receptor is a protein, the name of this particular protein is HER2. And in case you’re wondering the HER2 is actually an abbreviation for Human, that’s the H, epidermal which is the E, growth factor which isn’t included in this abbreviation, receptor which is the R so HER2 is Human Epidermal Growth Factor Receptor Number Two. And that implies that there is a HER1, there’s also a HER3 and, I believe, a HER4. So remember, a receptor is a protein that’s in the membrane of the cell and receptors receive information from a ligand. And in this case, the ligand is called epidermal growth factor.
So the epidermal growth factor binds or attaches to the cells just like the key and the lock. The epidermal growth factor binds to the HER2. And once the receptor is now occupied by the ligand, a message is sent to the DNA. And this message then, because it’s the growth factor that tells the cell to divide, growth factors tell cells to undergo division. And that’s all well and good. And this is exactly what supposed to happen under normal conditions. However, in 25% of the breast tumors what happens is – is we over express the HER2. And over expression is just the biotech word for saying we’re making too much of this. So it turns out that an addition to the HER2 gene which contains the nucleotide sequence to tell us how to make HER2, in front of every gene there’s a little on/off switch called the promoter region. And the promoter region determines how much of the protein we actually make. So we have a set of directions and then the promoter region tells us how often we’re going to use that set of directions.

And so in the case of HER2 induced breast cancer, the on/off switch is turned on too much and we make too many copies of the HER2 protein. So making too much of a particular protein is called over expression. We’re over expressing HER2. And so now what happens is, is instead of getting just one message to divide, the breast cancer cells are getting lots of messages to divide so they begin to divide very quickly and they also begin to divide constantly. So the idea here is that therapeutically, we’d like to be able to block that process. In other words, you know, we have the lock and the key, right, the epidermal growth factor, the ligand is the key, and it fits in to the human epidermal growth receptor. If we can prevent that interaction, then no message will be given to the cell to divide. And at the very least, we’ll have a maintenance therapy for our breast cancer.

And so the way that scientists approach this was to say HER2 is in fact a protein and if it’s a protein, we can make an antibody against it. Remember we talked about antibodies, are proteins that recognize and attach to other proteins. So we take the HER2 protein, we make an antibody against it. And you’ll learn a little bit about how to do that in our subsequent webinars. And we then give the patient this Herceptin, that’s the name of the antibody that binds with the HER2 receptor, the Herceptin antibody. And it essentially covers up the lock. It will be just somebody went to your house and put a cover over your front door lock and now you can’t put the key in get into the house. And so now the EGF, the ligand can’t bind to the HER2 receptor and no message can be sent to the nucleus to tell the cell to divide and therefore, the tumor is maintained, it doesn’t divide anymore, it doesn’t grow. At least that was the strategy that the scientists had hoped for.

It turns out that that strategy actually worked but there is an added, if you will, side effect but it was a beneficial side effect. And that is when the breast cancer cell is actually surrounded by antibodies, when any cell in your body is coated with antibodies, completely covered with antibodies, your immune system says, “Wow! If there’s so many antibodies attached to you, something must be really wrong so I might as well just kill you,” because remember we have 12 trillion cells so what’s a 100,000 cells here and there out of 12 trillion. So your immune system comes by, finds these cells that are coated with antibodies and very often it will then kill them. So not only did the Herceptin
treatment stop the tumor from progressing but actually in a number of patients the tumor shrunk because the patient’s own immune system was killing the tumor cells. And, of course, there aren’t any side effects with that as you would see with normal chemotherapeutic or radiation therapies.

**Personalized Medicine: Diagnostics**

So basically what we have shown here is the idea that we’re blocking the ligand from getting into the receptor and therefore, there’s no message for the cell to grow and that’s what the antibody does. But there’s another layer to this and this is really where we start to get in to personalized medicine. Because if you recall, the statistic we gave you is that only 25% of breast tumors are caused by making too much of the HER2 receptor. So we can’t treat everybody with the Herceptin antibody. It’s only going to be useful to treat people who make too much HER2. So the first step is to do a diagnostic. And what that means is we do a biopsy of the breast tumor and then we literally count the number of HER2 receptors. And so if, for example, you have a normal number of HER2 receptors, so let’s think about this, we’ve identified through imaging, ultrasound or X-rays or whatever, that you do, unfortunately, have a breast tumor. We’ve now taken a biopsy of that tumorous tissue and we’re now counting the cancer cells to see how many HER2 receptors are present on the cells. That can be done. This is standardized, fairly straightforward method. And so the idea here is that only 25% of the people are going to be what we call HER2 positive, they’re going to have too many HER2 receptors. 75% are going to be HER2 negative. And what that means is even though they have a breast tumor, those tumor cells don’t make too much HER2, there’s something else causing the cancer. And because something else is causing the cancer it doesn’t make sense, then, therapeutically, to administer Herceptin to them because there’s nothing there really for it to bind to.

However, if when we get the biopsy results back and we discover that you’ve made much more HER2 than normal cells, then we know that you’re indicated for the Herceptin treatment because there’s lot of HER2 protein for the Herceptin antibody to bind to. And when doctors did this, lots of patients responded, went into remission but sometimes it didn’t completely work. So then the doctor said, “How come this works in some people but not others even though they’re HER2 positive? There must be something else.” So they dug a little deeper and discovered that for patients who also express that has made in additional protein called PTEN, if patients were both HER2 positive and PTEN positive, then when they are administered the Herceptin most of them responded and it significantly prolonged their life, if not, eliminated the tumor. So this is what we mean by personalized medicine. Based on what we discover specifically for you we’re going to design a treatment around that instead of just saying, “You have a breast tumor. Let's go in and give you a radiation treatment and we’ll follow it up with chemo.” Now we’re going to say, “What might be responsible for the tumor? If it’s HER2 positive, we’re going to give you a different course of treatment than traditional radiation and chemotherapy.” And as we get more and more knowledge those types of decisions for
treating a particular disease will be based more on defined fine tooth diagnostics that we’re able to create and apply.

**Personalized Medicine: Dosage**

Now, the last thing, then, here is the concept of giving you a particular drug dose based on your DNA sequences. And so it turns out that right now most drug dosages are just simply designed and administered based on per kilogram of body weight, this is why when you go to the doctors you step on the scale and they weight you because if they ever have to administer a drug to you they’ll say, “Okay. You weigh 35 pounds so you get this dose of drug. You weight 800 pounds; you get, of course, a lot more drug.” But there’s more to it than that. And if you really think about it for minute you’ll know that this is true. For example, you probably or may be someone who only needs to take one aspirin to alleviate your headache but someone else you know may need to take three or four aspirins to get the same therapeutic effect. And the reason for that is, is that the cells in your liver contain a number of enzyme proteins called CYP, CYP proteins. And CYP stands for cytochrome p450. Cytochrome p450’s are enzymes that your body has evolved to break down drugs. So any drug that you take aspirin, penicillin, alcohol, no matter what it is, there’s a cytochrome p450 there that tries to break it down so that you can eliminate it. And just as a side, cytochrome p450’s are of major importance in preclinical studies for small molecules. Scientists in the FDA always want to know what cytochrome p450’s do to small molecules to chemical drugs. So it turns out that human beings have lots and lots of different cytochrome p450’s, there’s multiple families that are involved in metabolizing or breaking down drugs, those are families one through four, so there’s a CYP 1 family, a CYP 2 family, a CYP 3 family and a CYP 4 family, and then there are subgroups within each family.

And I only say this to let you know that it’s a fairly complex field to work with and we’ve, for example, illustrated here CYP-2D6. So it’s in these two families, the D sub family and then six is the sixth member of CYP in the 2D family. And each of these different Cytochrome p450’s breakdown different drugs. Now, to the point, how does this become personal and how does this relate to one versus three aspirins to get the same therapeutic effect? It turns out that due to SNPs changes in our DNA sequence, different p450’s act better of worst. So for example, if you have one particular DNA sequence for CYP-2D6 you may be what’s called a good metabolizer for anti-depressants. And what that means is if I give you a particular antidepressant and you’re a good metabolizer, your liver will break it down and eliminate it very quickly, so you won’t, in effect, get a very high dose. On the other hand, if you have different DNA sequence that codes for the CYP-2D6 enzyme, you may be a poor metabolizer, and in that case, you won’t break the drug down and you’ll actually end up with a higher dose. So the idea here is that if I can identify based on your DNA sequence whether you’re a good metabolizer or a pool metabolizer, I can much more quickly get to the appropriate dosage to administer to you because, you know, for things like blood pressure medication, antidepressant medication, it can take several months to half a year until you finally get it to the right dose. And the idea here is, we could look at you right away and say, “You’re a poor metabolizer.
You’re not going to break this drug down very quickly so let’s start you off right away in a much lower dose.” So it should save the amount of time that we use to get to the particular drug dose. And if you’re a good metabolizer I need to start you off at a higher dose.

So what’s the approach to this? We’re going to use something called a DNA chip or DNA microarray and this allows us to analyze many sequences of DNA at the same time. And on this particular DNA microarray it’s going to be designed to analyze all of our different cytochrome p450 SNPs. And so this is actually called a SNP chip.

The last thing that I want to mention here in the slide is also very important. And this goes to what are termed drug interactions. And it turns out that because so many drugs are metabolized by cytochrome p450, we have to be careful about the combinations of drugs that we take at the same time, and this includes alcohol and homeopathic remedies. And so what we’re illustrating here is, is even grapefruit can affect cytochrome p450-3A4 which is responsible for breaking down about 50% of the drugs that you take. So the idea here is, is that if you’re taking a particular drug that’s supposed to be broken down by the 3A4 cytochrome p450 and you eat grapefruit, the grapefruit will block or inhibit the cytochrome p450-3A4 so the 3A4 can no longer breakdown the drug. That means if I take 100 milligrams of the drug and I normally breakdown 50 milligrams and I only have a remaining 50 milligrams in my blood as a therapeutic dose and I eat grapefruit, now I might only breakdown 10 milligrams. So instead of having the therapeutic 50 milligrams, I have 90 milligrams because the 3A4 was blocked from breaking it down. And 90 milligrams may have serious side effects. For example, if it’s blood pressure medication, this may significantly lower your blood pressure so you become dizzy, you might even faint, you might even die.

And a number of homeopathic medicines, Saint John’s Wort is probably the most notorious for this, interfere with the cytochrome p450-3A4’s. So when you go the doctor and probably more importantly, when you go to the pharmacist you need to fully disclose all of the medications that you’re taking whether they’re prescription medications, over the counter medications or homeopathic remedies, and it’s also why they ask you if you drink and smoke because cigarette smoke and alcohol also affects cytochrome p450’s. So it’s a fairly complex picture and we’re still learning how to account for all of this. But it all comes from our knowledge of DNA, of how cells work, and DNA sequences and how that relates to protein structure. And in protein structure determines how proteins function.

**Summary**

So hopefully, then, in this section, we’ve sort of given you the bases of biotechnology. And the basis is biotechnology really lies in understanding how cells work, because cells are the basic unit of life, they’re the smallest unit or the smallest living thing. Everything is made up of cells. So human beings are multi cellular beings. And what do cells do? Cells grow, they divide, they manufacture proteins and they’re able to communicate with
other cells outside of themselves. The force behind the cell, if you will, the master molecule of the cell is DNA. DNA is really the master recipe book or the master instructional book for the cell. DNA contains all of the directions which are stored in discreet sequences of nucleotides called genes. So gene is just a set of directions but DNA contains all of the directions to run a cell. And we want to emphasize that DNA, though, is just a chemical. DNA is not alive; it’s just a chemical structure. And all cells, whether its bacteria cell, an animal cell or a plant cell, the chemical structure of those cells DNA is exactly the same. What makes one cell, one organism different from another is the sequence of the DNA that is the order of the nucleotides. Cells then use the information in DNA, that information is packaged as a gene to make a protein. So a gene is a sequence of DNA that is the directions to make a protein.

Then slight changes in protein structure caused by mutation or a change in the DNA sequence can have a profound influence on protein function because it changes a mutation and can cause a change in the shape of the protein due to a change in the amino acid. In some cases these mutations, these changes in the DNA sequence can result in a disease because now the protein can’t function properly. And we gave you the example of sickle cell anemia, cystic fibrosis and Parkinson’s disease.

And then finally, by analyzing or characterizing and understanding these differences in gene sequences from one person to the next and correlating those to a particular disease, we can begin to develop a diagnostic by saying, “Do you have a normal sequence or do you have the mutated sequence.” We can begin to make predictions about what disease you have, what your likelihood of getting the disease is, and ultimately, because we know which genes are responsible for the disease, we should be able to apply that knowledge to develop a therapeutic.

So thank you for sticking with us through our first webinar. And there are two more to come. So we’ll talk to you in a bit. Bye now.

**Stacey Franklin:** Hi, my name is Stacey Franklin, CEO and owner of BioTech Primer. This webinar series has been developed and delivered by Biotech Primer. Biotech Primer is a training company that specializes in teaching the basics of biotechnology for the nonscientist. Want to learn more? Please visit our website at [www.biotechprimerinc.com](http://www.biotechprimerinc.com) to see a listing of upcoming classes or contact me to discuss customized in-house classes. Thank you.

[1:12:14]

[Audio Ends]