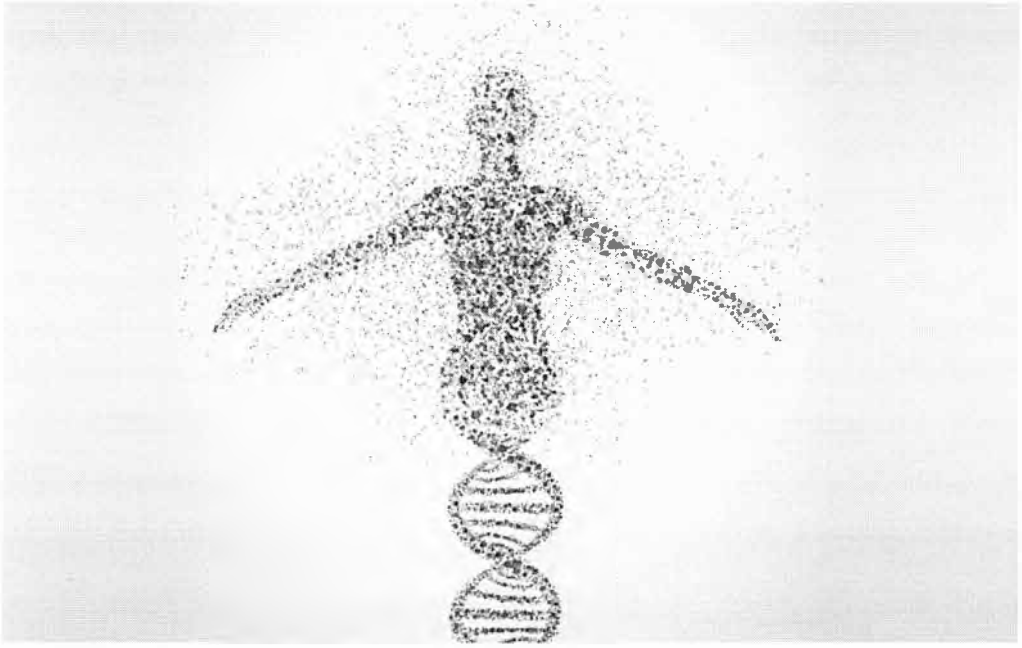
Complexity of Life



[M2U3]

Name _____

Period _____ Date _____

PBS Nova Documentary Collection: Cracking the Code Of Life

Viewing Guide

1. How long is the instructions for a human being?

2. What small, very powerful molecule is this video about?

3. How similar are humans to a banana?

4. What are 2 things that we have in common with a banana that makes our DNA similar?

5. What is one of the goals of the Human Genome Project?

6. What are some if the current issues associated with knowing our genetic code?

7. What does DNA look like?

8. How much DNA do we have compared to fruit flies?

9. What are the four different kinds of parts that DNA is made of?

10. What is the shape of DNA called?

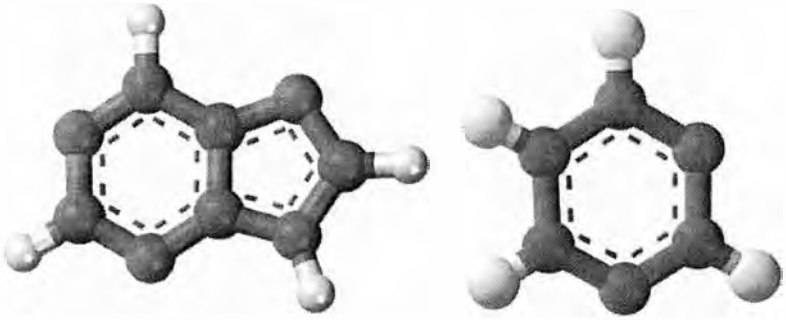
Purines vs. Pyrimidines

Purines and Pyrimidines are nitrogenous bases that make up the two different kinds of nucleotide bases in DNA and RNA. The two-carbon nitrogen ring bases (adenine and guanine) are purines, while the one-carbon nitrogen ring bases (thymine and cytosine) are pyrimidines.

Comparison chart

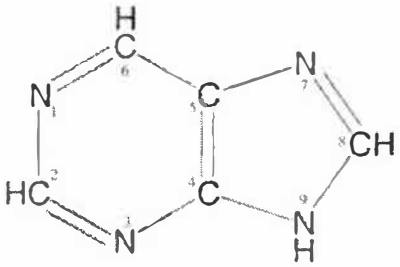
	Purines	Pyrimidines
Introduction (from Wikipedia)	A purine is a heterocyclic aromatic organic compound, consisting of a pyrimidine ring fused to an imidazole ring.	Pyrimidine is a heterocyclic aromatic organic compound similar to benzene and pyridine, containing two nitrogen atoms at positions 1 and 3 of the six-member ring. It is isomeric with two other forms of diazine.
Function	Production of RNA and DNA, proteins and starches, the regulation of enzymes and cell signaling.	Production of RNA and DNA, proteins and starches, the regulation of enzymes and cell signaling
Nucleobases	Adenine and guanine	Cytosine, thymine, uracil
Structure	A pyrimidine ring fused to a imidazole ring. Contains two carbon-nitrogen rings and four nitrogen atoms.	Contains one carbon-nitrogen ring and two nitrogen atoms.
Melting Point	214 °C, 487 K, 417 °F	20–22 °C
Type of Compound	Heterocyclic aromatic organic compound	Heterocyclic aromatic organic compound
Molecular formula	$C_5H_4N_4$	$C_4H_4N_2$
Molar mass	120.11 g/mol	80.88 g/mol
MeSH	purine	pyrimidine
SMILES	c1c2c(nc[nH]2)ncn1	C1=CN=CN=C1
CAS number	120-73-0	289-95-2Y
PubChem	1044	9260
Synthesis in Lab	Traube Purine Sythesis	Biginelli Reaction

Structure



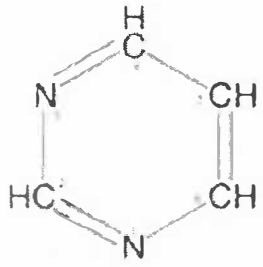
Purine (L) and Pyrimidine (R) molecules, where Black= Carbon, White=Hydrogen, Blue=Nitrogen

A purine is a heterocyclic aromatic organic compound containing 4 nitrogen atoms. It contains two carbon rings, and is made of a pyrimidine ring fused to an imidazole ring.



Structure of a purine

A pyrimidine is a heterocyclic aromatic organic compound containing 2 nitrogen atoms. It contains only one carbon ring.



Structure of a pyrimidine

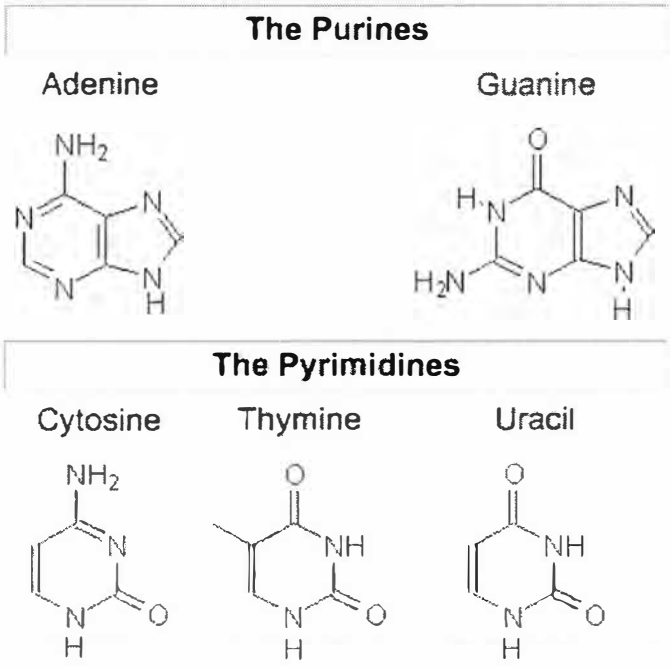
Function

Both purines and pyrimidines have the same function: they serve as a form of energy for cells, and are essential for production of DNA and RNA, proteins, starch, regularion of enzymes, cell signaling.

Nucleobases

Purines make up two of the four nucleobases in DNA and RNA: adenine and guanine. Pyrimidines make up the other bases in DNA and RNA: cytosine, thymine (in DNA) and uracil (in RNA). Useful mnemonics to remember these bases are:

- "CUT the Py": CUT: Cytosine, Uracil, Thymine; Py (pyrimidines)
- "Pure as Gold (Pur AG): purines are Adenine, Guanine



The chemical structure of all purines (adenine, guanine) and pyrimidines (cytosine, thymine, uracil).

Meet the Decoders

Dr. Eric Lander

Dr. Eric Lander

Krulwich: Let me see, let me just start with the gene itself. First of all, we're all familiar with this thing. This shape is very familiar.

Lander: Double helix, yes.

Krulwich: Double helix. First of all, this is my version of a DNA molecule, but I'm just curious how small is it in real life? Like the distance--is this by the way, what it looks like?

Lander: Well, give or take, a cartoon version. Yes. That's right.

Krulwich: This is made of what, these sort of walls?

Lander: This chain here that runs along the outside is made of sugar molecules and negatively charged phosphate molecules. Then it goes: sugar/phosphate/sugar/phosphate/sugar/phosphate, running all along here, boring and utterly repetitive.

Krulwich: So if you lick it, it would be a little bit sweet this part?

Lander: Well, no. It wouldn't because they're all a patch. You see, it only tastes sweet if it comes off and activates your taste receptors. So these long chains are not very good.

Krulwich: Okay.

Lander: I mean wood is also made out of these sugars and things, and it doesn't taste very sweet, because it's all sown up.

Krulwich: Okay.

Lander: But if you broke it all up, there is a little sugar in there.

Krulwich: The main thing about this is the ladder, the steps of this ladder.

Lander: So the news is along these steps, right? The outside is the same all along; the inside sticking out of these sugars comes one of four different things, one of four bases that we call A, T, C, or G. They correspond to different funny molecular shapes with little rings and things.

Krulwich: And when you say A, T, C -- I'm just going to set this thing up here.

Lander: Okay. So let's say that this red guy here is an A and that's a T.

Krulwich: Why do you call them A and T?

Lander: Oh, it's adenine and thymine.

Krulwich: So those are their initials?

Lander: But it's a fast-moving field. You can't say adenine all the time and thymine and cytosine all the time and guanine all the time, so you say, A, T, C, and G.

Krulwich: So there are in almost every cell in your body, if you look deep enough, you will find this chain here?

Lander: Oh, yes. Stuck to the nucleus of your cell there is I guess 23 pairs of these chains called chromosomes.

Krulwich: Right.

Lander: Each chromosome is a very long chain of about anywhere 100 million rungs of the ladder.

Krulwich: In every cell?

Lander: Every cell.

Krulwich: Over and over. How many of these steps do we have in a typical cell in me?

Lander: Well, you've got six billion, of which three billion came from your mom and three came from your dad. Those two copies are pretty similar, so every cell has a genome of three billion, and it actually has about--it has two copies of it; one from mom and one from dad.

Krulwich: Let's just think for a moment. Let's just look at the whole landscape. The human being has 3.1 billion of these letters, not all of them are genes, in fact only 1 percent of them are genes. So, where are they? Is there any pattern?

Lander: The genome is very lumpy.

Krulwich: Very?

Lander: Very lumpy, very uneven. You might think, if we have 30,000 genes, they're kind of distributed uniformly across the chromosomes. Not so. They're distributed like people are distributed in America: they're all bunched up in some places, and then you have vast plains that don't have a lot of people in them. It's like that with the genes. There are really gene dense regions, that might have 15 times the density of genes, sort of New York City over here. And there are other regions that might go for two million letters and there's not a gene to be found in there.

Krulwich: So there's this whole conversation going on that we didn't know about, like, ten years ago.

Lander: The genome is, it's a fossil record; the genome is a landscape; the genome is a whole geography of distributions. The thing about the human genome that most surprised me was how many amazing stories there were in it. That you might think the genome's just a boring string of letters, like reading the ones and zeros on your hard disk. The genome is a storybook that's been edited for a couple of billion years, and you could take it to bed, like *A Thousand and One Arabian Nights*, and read a different story, in the genome, every night.

Krulwich: If I could read each of the individual letters, I might find the picture of what?

Lander: Well, of your children. This is what you pass to your children. You know, people have known for 2,000 years that your kids look a lot like you. Well, it's because you must pass them something, some instructions that give them the eyes they have and the hair color they have and the nose shape they do. And the only way you pass it to them is in these sentences. That's it! It is hereditary.

Krulwich: So if I decode the two parts of each of these ladders, I would know many things about my children, things to be.

Lander: Well, in principle you would, the same way that if you read all the ones and zeroes on your hard disk, you would in principle know the Mozart symphony or the Shakespeare play that you put on your hard disk. Now, whether you could actually take the ones and zeroes and sing the Mozart symphony from that, is another question. So getting the letters out is one thing. Extracting their meaning is another. The Human Genome Project is about sitting there and getting the letters out. The next century is about extracting all of the meaning out of the text, but it's a pretty good text. It's worth the effort.

Krulwich: Getting the letters out has been described as finding the blueprint of a human being, finding a manual for a human being, finding the code of a human being. What's your metaphor?

Lander: Oh, golly, gee. I mean -- I think this is very much like in chemistry, the way the chemists describe all of matter in terms of elements that they put into a periodic table.

Krulwich: Right.

Lander: I think this is kind of biology's periodic table. Everything that gets made in your body, whether it's, you know, carotene in your hair or collagen in your skin or hemoglobin in your blood, is specified by an instruction here. This is basically a parts list. Blueprints and all these fancy -- it's just a parts list. It's a parts list with a lot of parts.

If you take an airplane, a Boeing 777, I think it has like 100,000 parts. If I gave you a parts list for the Boeing 777, in one sense you'd know a lot. You'd know 100,000 components that have got to be there, screws and wires and the rudders and things like that. On the other hand, I bet you wouldn't know how to put it together. And I bet you wouldn't know why it flies.

Well, we're in the same boat. We now have a parts list. That's what the Human Genome Project is about is getting a parts list. If you want to understand the plane, you have to have the parts list, but that's not enough to understand why it flies. Of course, you'd be crazy not to start with the parts list. So we figured that for the next century of medical work, we'd better get the parts list and so everybody rolled up their sleeves and decided we could work together and get a parts list.

Krulwich: Before we go to how we do that, let me just give you what I guess is one advantage. If I discover that some human being is sick, and I suspect that the sickness comes from some genetic trait, I guess I could ask and ask and ask this molecule, "What's wrong with my kid? What's wrong with my kid? What's wrong with my kid?" But the fact that it's a molecule with a beginning and an end means that whatever the mistake is it's somewhere in this molecule.

Lander: That's what's so important about this being like a periodic table. When the chemists try to go figure something out, they know there's only 100 elements or so, and whatever you're trying to explain about matter, you've got to explain in terms of those 100 elements -- oxygen, hydrogen, etc.

What's now happening in biology is we know that whatever you want to explain about heredity it's got to be in the three billion letters specifying maybe 30,000 genes or so. In the past you could always say, "Well, the cause is something we haven't discovered yet." Not anymore. There isn't a something we haven't discovered yet in genome form, or there won't be very soon, and so it's a sea change in science to know that you're explaining cancer or you're

explaining brain degeneration that's inherited, whatever you're explaining, you only have 30,000 explanations to use and maybe their combinations.

Everybody looks for the genome and thinks, "Ah! We're going to cure everything." No. It doesn't mean we're going to cure everything. It means for the first time we stand a fighting chance of explaining the causes of things and getting understanding about how it works. And then sometimes we'll be able to cure it by the understanding. It's just that up to now we've been trying to do it in utter ignorance about most of the parts.

Krulwich: Most people think that this project has to do with getting sick, that is, you're going to discover the gene that causes blindness and the gene that causes deafness and the genes that cause cancer and the genes that cause these terrible things. But aren't we also trying to figure out the gene that make you just a bit different from me, so we could discover the chemistry of difference? Or is that not the issue?

Lander: We want to understand what makes a human being tick. Sometimes we study differences between people, and the best way to find out what's wrong common to all of us. By figuring out why I'm taller than someone who is shorter, we can figure out the common mechanism that controls bone growth, for example. And that mechanism is useful to everybody, so in some sense all this focus on difference is not really about classifying people by their differences, it's that differences are our best way in to the biological mechanisms that are the same in all of us.

We can only study those mechanisms when we see a difference in them, because that's our clue to know that the mechanism is there. If we were all exactly the same height, we wouldn't think much about the mechanism that controls height, but because we see a lot of difference in height, we're clued in to the fact that there are mechanisms at work that make some of us taller than others, and we can study those.

Krulwich: How will we know the difference that genes make and the difference that environment makes? For example, if we're looking for cancer-causing agents, and we go to Finland, we might say after rigorous study of the people of Finland that everybody who gets cancer there seems to have a problem, here, here, and here. But then when we go to Tanzania, we may discover that everybody who gets cancer in Tanzania has a problem here the same, but now here and here. Then what have we learned?

Lander: Well, the first thing to say is that the differences around the world are much less than you think, even though we're six billion people scattered around the whole world. In point of fact, the entire human species traces back a mere

5,000 generations ago, to a small population in Africa. All of us descend from a population of 10- or 20- or 30,000 people in Africa about 100,000 years ago. And it means that, because 5,000 generations is such a short time from the point of view of genetic evolution, the variance present in Africa and the variance present in Asia and the variance present in northern Europe are largely the same. There's not that much variation.

Krulwich: How close are we? I know that we are very close to chimpanzees for example.

Lander: Any two chimps in Africa are five times more different than any two humans on this globe.

Krulwich: Any two chimps?

Lander: Any two chimps in Africa have five times more genetic variation than two random humans on this globe. You think the chimps all look the same; believe me, they think we all look the same.

Krulwich: And on paper, anyway, we are more of the same to each other?

Lander: We are much more similar to each other than chimps in Africa are.

Krulwich: Because we are such a young species?

Lander: We are a small species that has grown large in the blink of an eye. We tend to forget that; we like to focus on our differences, but our differences are a pretty trivial portion of our genome. And the distribution of those differences across the globe is actually much flatter than we think about, much more uniform. If we take any village anywhere in the world, it has roughly 80 percent of all the genetic variation found anywhere in the world, just within that village. We just like to have wars and things about how different we all look, but compared to most species, we're not very different.

Krulwich: If we go down that list of 3.1 billion chemical constituents.

Lander: Yup.

Krulwich: Do all of them manufacture something that makes us?

Lander: No. No. Actually most of the DNA letters in your genome don't manufacture any proteins that do anything in your body. Indeed, only a few percent are probably needed to manufacture all the proteins in your body.

Krulwich: Why do you mention proteins?

Lander: Genes contain instructions for making things, and almost always the things that they contain the instructions for are proteins like carotene in your hair or collagen in your skin or hemoglobin in your blood.

Krulwich: Huh! So genes create proteins and proteins build us.

Lander: Proteins do the work. Genes are the instruction sets for those proteins. Actually only a few percent of all of your DNA is devoted to writing down the instructions for those proteins.

Krulwich: Say if five percent makes the proteins, what does the other 95 percent do?

Lander: Well, at least 50 percent of it consists of selfish DNA elements.

Krulwich: Of what?

Lander: Selfish DNA elements, that is to say chunks of DNA that know how to reproduce themselves. And they just live in your chromosomes, so you think it's your genome, but, in fact, more than half of your genome consists of elements that know how to copy and move themselves around or fossils of such elements.

Krulwich: Are they my predecessors up the genealogical tree?

Lander: No. No.

Krulwich: Are they a snake that I formally was?

Lander: Not at all. They couldn't care less about the snake or anything like that. They've been around for about one and a half billion years, and they're little DNA segments. There's one that's about 6,000 letters long. It knows how to copy itself into RNA and move itself to someplace else in your DNA, and when it gets there, it knows how to copy itself and move itself someplace else. It doesn't do anything useful for you, and, thank you, it isn't interested in doing anything useful for you.

Krulwich: Is this some billiard ball that I've got wandering around inside my genome?

Lander: Well, it's something like that, yes. You've got these things hopping around your genome that get called "transposable elements," and they look out for themselves. They're like little parasites in your genome.

Krulwich: Okay. So I've got letters that do stuff that makes me.

Lander: Right.

Krulwich: I've got wandering around stuff that doesn't do anything. It just seems to have the habit of living in me.

Lander: It does a lot for itself. It largely thinks your genes are just there to help propagate it.

Krulwich: So I've got five percent that's making me, a lot of stuff that's jumping around and then trash or junk?

Lander: It's junk. You keep it around, because it turns out that occasionally the junk in your genome turns out to be very handy. Every once in a while a piece of junk that lands in your genome serves a useful purpose and gets used by the body. A great example is your immune system. Your immune system does a very cute trick where it can rearrange pieces of DNA to make antibodies to recognize anything. You know where it learns how to do that?

Krulwich: Where?

Lander: It's a piece of junk that landed in the genome about 400 million years ago that carried with it an enzyme for hopping around. Your body took over that enzyme and used it to rearrange genes to make antibodies. So every once in a while we get quite a nice gift from this junk.

Krulwich: So you make no apologies for spending all those long machine hours writing down the names and letters of all that stuff.

Lander: Indeed, many people argued in the 1980's that we should just spend the money, sequencing the important stuff. The problem is the important stuff is scattered about in the sequence. We could have spent 20 times as much money, just finding the important stuff and sequencing it, as sequencing everything.

It would be like, you know, going through a library and, you know, reading the books or something and trying to -- "I'm just going to read the interesting sentences." Well, it's really tough when you've got a novel to just read the interesting sentences or the interesting paragraphs without reading the novel.

Now somebody will invent a very clever way to do that, but nobody has yet, and so the most efficient way was to sit there and read the whole thing. It turned out to be not that expensive, and every one of those letters helps us understand human biology. We can actually learn a lot from the junk. You want a great example?

Krulwich: Yeah.

Lander: The gene that causes Duchenne muscular

dystrophy when it's mutated, that gene is spread out over two million letters of DNA, of which only 16,000 matter.

Krulwich: Wow! So the useful part is scattered in the junky part?

Lander: The cell copies and copies and copies two million letters, and then it splices together, throwing out almost all two million letters to save only 16,000 letters to make that protein.

Krulwich: But how do you know of all these rungs which is the stuff that's doing stuff for you?

Lander: That's what we're engaged in right now. That's called interpreting the sequence, annotating the sequence. You've got all the letters. It actually doesn't come color-coded; it doesn't come with little special marks saying "This is important." You have to sit there with all the A's, T's, C's and G's and figure out which bits matter. And so we have elaborate computer programs to try to figure out which fits have the potential to code for a protein in your body.

Krulwich: I get the sense that everybody is getting out of the gene business and suddenly going into this new business I hear about called the protein business. There's even a new name, instead of the genome, I'm hearing this other name, which I don't--

Lander: The proteome.

Krulwich: The proteome, what is that?

Lander: Well, the genome is the collection of all your genes and DNA, the proteome is the collection of all of your proteins. See what's happening is we're realizing -- I think we always realized that if we wanted to understand life, we had to start systematically at the bottom and get all the building blocks. The first building blocks are the DNA letters, from them we can infer the genes. From the genes, we can infer the protein products that they make to do all the work of your cell.

Then we've got to understand what each of those proteins does, what its shape is. How they interact with each other, and how they make kind of circuits and connections with each other. So in some sense, this is just the beginning of a very comprehensive systematic program to understand all the components and how they all connect with each other.

Krulwich: How many proteins do we have?

Lander: Well, it's very interesting. One gene can make multiple proteins. It turns out that the gene makes a

message, but the message can be spliced up in different ways, and so a gene might make three proteins or four proteins, and then that protein can get modified. There could be other proteins that stick some phosphate group on it or two phosphate groups.

And, in fact, all of these modifications to the proteins could make them function differently. So, while you might only have say, 30,000 genes, you could have 100,000 distinct proteins, and when you're done putting all the different modifications on them, there might be a million of them. It's scary talk.

Krulwich: The proteome project could then last a lot longer than the genome project?

Lander: Oh, no. I doubt it, because, I mean, the difference between 30,000 and 100,000 is only a factor of three. The great thing about this explosive period of science is that these big numbers, they sound scary, and then within five years some bright student figures out a way to say, "Oh, yeah, well, what's a factor of five, and what's a factor of 10 here and there?" It wasn't very long ago that sequencing a couple hundred letters of DNA was enough to get a whole Ph.D. thesis.

Krulwich: Now a robot does it?

Lander: A robot does an awful lot more than that in the course of a day. There's just this incredibly explosive learning curve about how to work with this material.

Krulwich: Before I leave the subject of proteins, one last question.

Lander: Yeah.

Krulwich: When I think about proteins, if I'm trying to think strategically -- I think if I get sick that may mean that my proteins are causing me harm, because the bad gene created a bad protein?

Lander: Yup.

Krulwich: So are pharmaceutical companies now going to be trying to figure out how to maybe not get down and rearrange the genes, but maybe just deal with the protein?

Lander: Most drugs on the market today affect the protein, not the gene. The gene is the instruction for making this component. The drug you take, the pill you pop, usually interacts with the protein, to slow down its function, speed up its function. That's what's really going on. Very few drugs interact with genes.

Krulwich: So if I had a horrible genetic disease, like

Huntington's, maybe someone will figure out a way to cure the protein part of my disease?

Lander: For Huntington's disease we know that the problem is the gene is defective in a way that it makes a protein that is a big long extra chunk of it that's very sticky, and it sticks to other stuff and gums up the works. The Holy Grail for Huntington's disease is to make a drug that would block this sticky protein from interacting with other things.

Krulwich: How about cystic fibrosis?

Lander: For cystic fibrosis, it's a little trickier. There you have a protein that can't perform its job. It can't transport chloride across itself. You've got to make a drug there that somehow would restore its function. That's not very easy. When you have a broken part, it's hard to add a gene that will restore its function. So there people are talking about somehow spraying a virus into your lungs that will carry with it a new good copy of the cystic fibrosis gene, for example. Those are the kind of ideas people have.

Krulwich: And, last, I get the sense that the genome project was largely a way of discovering a long chain of things? When I think about the proteins and the proteome project, for some reason I think about origami. I don't know why, but when I read things, I hear that you're folding and unfolding and bending and twisting. What is that about?

Lander: The genome project was a piece of cake compared to most other things, because genetic information is linear. It goes in a simple line up and down the chromosome. Once you start talking about the three-dimensional shapes into which protein changed and can fold, and how they can speak to each other in many different ways to do things, or the ways in which cells can interact like wiring up in your brain, you're not in a one-dimensional problem anymore. You're not in Kansas anymore. You're way off.

Krulwich: You need geometry--

Lander: You need complicated geometry and diffusion. That's why it's not a quick shot to just go from the genome to curing something. The work of the next century will be to take the simple linear sequence and see if we can build models that explain these complex three-dimensional interactions and the kind of conceptual circuits of A tickles B, tickles C, tickles D at the molecular level and figure out, you know, which proteins to goose to fix some disease.

Krulwich: Here's one of the big astonishments, to me. We are related, more closely related, apparently, to creatures, than I had ever imagined. You too?

Lander: We are all incredibly related to creatures.

Krulwich: But a chimpanzee, that I could figure. A big ape, that I could figure. Eighty to eighty-five per cent of the genes in a mouse are in one.

Lander: More than that.

Krulwich: More than that.

Lander: More than that. I'd say, it's 98 percent of the genes in the mouse, you can find clear matches in a human. There's not a lot of difference between you and a mouse.

Krulwich: What does that mean exactly? If you took off the skin of a mouse, went right in and looked at the cell, would you just see the same exact stuff, in the same exact order that you see in me?

Lander: If you looked inside the body of a mouse, you see lungs, you see hearts, you see pancreases.

Krulwich: But let's get cellular.

Lander: But wait a second. All of the genes that have to direct the arrangement of those organs have to be the same. The only difference is, a few genes might control some size. There might be a little variation of the exact shape of things. But those are small compared to building all these complex structures. Pretty much every structure we have a mouse has. And so most of the genes you're going to need to build it are the same. It's like two kinds of airplanes.

Krulwich: But we are so much bigger, smarter. We make music, they squeak. We walk around, we go to school. There has, as far as I know, never been a mouse school.

Lander: Yeah.

Krulwich: Therefore, how could they have as many genes as we do and yet we be so much, excuse the expression, better.

Lander: Well, it's a somewhat self-centered view to think we're better, but we're certainly different, right? I think the main take-home message from this is that the differences in the human from many other organisms may actually be small differences of degree that then have gotten amplified.

See, we do all sorts of things that a chimp doesn't do. And that's not because we invented a lot more genes. It may be for something as simple as the little genetic controls that cause the number of nerve endings made them double in the human, for example. That could lead to lots more connections without lots more genes. It might be just tweaking the dials a little bit ends up producing quite qualitatively different behavior.

Krulwich: So you could look deep into the eyes of a great ape or a chimpanzee and there is something vaguely familiar about that.

Lander: Vaguely familiar? It's virtual identical. To any other organism on this Earth, they'd have a hard time telling you apart from the great ape.

Krulwich: And yet what's the difference between us may be simply that our genes turned on some dial so we got ten times more brain cells than they got, or something like that?

Lander: Or, probably more likely, that brain development slowed down in humans so that more of it could be affected by environment and experience and childhood. In many ways, it said that the human brain development is more like a baby chimpanzee, a neonatal chimpanzee, and that we are therefore open to more learning and more shaping by environment for much longer in our development. That may be the secret to it, that we actually fossilize our brains into the adult form much more slowly.

Krulwich: Which means that the difference between us and the big ape is not the genes that we don't have together, it's... What is it, then?

Lander: It may be the controls on these genes. It may be simply how soon the gene is turned up or off or what level it's turned on to. It might be two letters of DNA that affect how actively the gene is transcribed in the cell. Those are the sort of small variations on the theme that could make a huge difference in how an organism turns out.

Krulwich: Let me get back to relationships. If we are almost identical to mice and 80 percent of the genes in the cow are in us, and 61 percent of the genes in a fruit fly are in us, and -- this is the one that gets me -- 50 percent of the genes in a banana are in us?

Lander: How different are you from a banana?

Krulwich: I feel -- and I feel I can this with some authority -- very different from a banana.

Lander: You may feel different from a banana--

Krulwich: I eat a banana, but I have never--

Lander: Look, you've got cells, you've got to make those cells divide. All the machinery for replicating your DNA, all the machinery for controlling the cell cycle, the cell surface, for making nutrients -- all that's the same in you and a banana.

Krulwich: Wait. You honestly feel, when you came to this

subject, a genetic relationship to the banana that was this strong?

Lander: Take baker's yeast. Baker's yeast we're related to, one and a half billion years ago. You can take genes out of a baker's yeast that controlled the basic cell processes like cell division, and you can put them into a human cell, and they function. And, in fact, they're the very genes that cause cancer in us, because they affect cell division and they affect cell division in yeast.

Krulwich: Wait. In other words, if you had a person who was sick in a particular way -- let's say his cells weren't breaking down properly or multiplying properly. You could take a healthy gene from a yeast, stick it in the person, and make the human feel better, fix the human?

Lander: It would be unethical to do it that way, but we do the reverse all the time. You can take the gene out of the human, stick it into the yeast, to make it play the same role as its evolutionary relative, one and a half billion years ago. And it works. Presumably, it works in the other direction, too, although we don't do it on patients.

Krulwich: So you were saying, then, that you could take a yeast that wasn't feeling too good and then go into my cell and take a piece of me, a healthy chunk of me, and fix the yeast?

Lander: Well, if there was a defect in how the yeast's cell division occurred, your gene would work just fine in the yeast, and people do that experiment all the time. We can do gene therapy to cure a sick yeast using a human gene. Presumably, it works the other way, too, although we don't do it, because it's not ethical. But even after one and a half billion years of evolutionary separation, the parts are were still interchangeable for lots of these genes.

Krulwich: Now, does that mean--I just want to make sure if I understand this right. Does that mean when you look through those things that all the C's and the A's and the T's and the G's, are you seeing the exact same letter sequences in the exact same alignment? When you look at the yeast and you look at the person, is it C-C-C-A-T-T?

Lander: It's eerie. The gene sequence is almost identical. There are some genes, like ubiquition, that's 97 percent identical between humans and yeast, even after a billion years of evolution.

Krulwich: Well, with a name like that, it's got to be.

Lander: Well, yeah. But then there are other genes where, when you look at other genes, you can still see, for example, that out of the protein sequence, oh, 60 percent of the amino acids in the protein are exactly the same, and

maybe 20 percent more are at least similar chemically. And there's no doubt it's the same object, even though it's varied some over one and a half billion years.

Krulwich: Now, is this why you can go to a pig, a big fat pig, and get the pig insulin out of the pig and give it to a human diabetic? I mean, are we talking about the same thing?

Lander: Pig insulin, human insulin, it's doing the same thing; it's interacting with the same receptor on the surface, an insulin receptor making the same kind of molecular signal. It's the same thing. Now, there're slight variations in pig insulin and human insulin that could lead to problems of immunological cross-reaction, but that's usually about details, small details in the insulin. The function of the pig insulin is virtually identical to the function of the human insulin.

Krulwich: So when you say interchangeable parts, that's actually become a business: People take a product from a pig and give it to a sick human, and the sick human uses the product and gets better, feels better.

Lander: Although these days we do it by cloning techniques where we can actually get a bacteria to produce the human product, and that's even safer still. But you've got to understand that, deep down, the fundamental mechanisms of life were worked out only once on this planet and have gotten reused in every organism on the planet. Evolution doesn't go reinvent something when it doesn't have to.

Krulwich: Well, is it a headline that we are more closely related to each-- By the way, before I ask that, how about plants? Let's suppose I had like a mustard seed, I don't know, a fern -- I'm trying to think as distant from myself as I can. So I walk up to a fern and I say Hello. What part of the fern has already got more or less pieces of me in it?

Lander: Oh, golly. Mostly what you share in common with a plant are the basic genes that run a cell. Because while you look very different from a plant, standing back, the closer and closer you get to a cell, the more you see bag with stuff in it and a nucleus. And most of those basic functions are the same. I suppose if you look at two different automobiles driving on the street, they may look very different, but if you look under the hood the engines look remarkably similar. And I think that's true for most cells, whether it's in plants or animals -- that, under the hood, the engine is pretty similar.

Krulwich: And that seems to be a major discovery of this project, so far.

Lander: Well, this has been a discovery, I think, that's been dawning on us with genome sequencing over the last couple of years, and the human sequence has made very clear that we are not separate and different in any way there. We are

very much partaking of that same bag of tricks that evolution's been using to make organisms all over this planet.

Krulwich: Let me go to the next headline. Fruit flies, as much as you might admire them, are small and foreign like creatures to moi. Yet I discover that the number of genes in a fruit fly is only about half the number of genes in me.

Lander: Yeah. That's really bothersome to many people, that we only have about twice as many genes as a fruit fly. Because we really like to think of ourselves as a lot more than twice as complex as a fruit fly.

Krulwich: Don't you?

Lander: I certainly like to think of myself that way. And so it raises two questions: Are we really more complex? Well, I think so. I think we've got lots more different tissue types, cell types, interactions in our body.

Krulwich: Well, let's let things speak for themselves. You show me the fruit fly that can compose like Mozart, and then I'll obviously--

Lander: Show you the human that can fly, right?

Krulwich: All right.

Lander: We all have our talents, right?

Krulwich: Yes. Anyway, but you feel somewhat diminished, don't you, a little bit, at least at first blush, by this?

Lander: Well, at first blush the reaction is to be a little insulted that you only have twice as many genes as a fruit fly. But when you begin to think about it, you realize that complexity may not just be in the number of genes. Twice as many genes may translate into four times or eight times or 16 times as much complexity. And we can see from the genome sequences -- this is really one of the most exciting things from the genome sequences -- ways in which our own genes are more complex.

Krulwich: So, the difference, then, the possible difference, between a fruit fly and a human is that the number of different splices and restitchings in a human is just higher?

Lander: It seems we've got twice as many. So one of the levels at which we're more complex than a fly is that a typical gene might get spliced and diced up into twice as many products as in a fruit fly. That's one level of complexity.

Krulwich: So twice as many proteins are being produced by the same gene in a human than in a fruit fly.

Lander: That's one level of complexity. But there's more to it. When we actually look at the proteins in the human, proteins are made out of building blocks that we call modules. Humans haven't invented a lot of new modules I've got to say. Most of our modules are the same as in the fruit flies. But we put them together in more combinations. We call those architectures. We have more different architectures for our proteins than a fruit fly does. And I think what that does is it lets our proteins interact in lots more ways. We can make lots more complex tinker toys out of these parts, because we've combined those modules in twice or three times as many combinations.

Krulwich: So starting with the same raw ingredients, the fruit fly goes... But the human, by somehow or other being able to arrange all the parts in many different ways, can produce melodies perhaps.

Lander: Yes, although we're not that good at hearing the melodies yet. One of the exciting things about reading the genome sequence now is we're getting a glimpse at that complexity of the parts and how it's a symphony rather than a simple tune. But it's not that easy to just read the sheet music there and hear the symphony that's coming out of it.

Krulwich: But is a metaphor okay? Should we think of the fruit fly somewhat as a creature that's learning to play chopsticks, and think of the human as a creature whose genes have learned to play Mozart?

Lander: Yeah. That's probably a bit extreme. I mean, it's pretty fancy making a fruit fly, too. I'd say it's way past chopsticks, and we may not quite be Mozart, but there's no doubt that we have more lines of music coming in together, in a much more complex melody that's coming out of it.

Krulwich: And so the bottom line question is, What is it that makes a human being human? Since if you're sitting next to a fruit fly and he seems to have almost the same number of ingredients that you do, the difference, the thing that makes us human, is?

Lander: Well, the one thing that distinguishes human beings from all other organisms on this planet is they're the only organisms that worry about what distinguishes them from all other organisms on this planet. Other than that--

Krulwich: I was hoping for some more architectural response.

Lander: I understand, but every other attempt that people have to say why humans are distinctive usually fails when you look really hard. But we're the only ones that worry about that question.

Krulwich: But, if I hadn't forced you, you'd say it had something to do with the human being's genes able to produce more proteins and more variation among the proteins.

Lander: Now, be careful. This isn't what makes humans different. This is what makes vertebrates different. Because whatever you're claiming for the human at the level of its genomic complexity, you've got to claim for the mouse and the dolphin. So this is not distinctively human. If you ask, "What's distinctively human about the human genome, compared to, say, a chimpanzee?" We haven't a clue. In fact, if I gave you the three billion letters of the chimp, and the three billion letters of the human, and I didn't tell you which was which, without cheating by peeking at the right answer, no scientist on Earth could tell you which was human and which was chimp.

Krulwich: But why? Isn't there somewhere you look for "hair all over" or "big face?"

Lander: Of course, somewhere it says, "Hair all over the body" in the chimp genome. But we don't know how to read that. It's there, but we don't know how to know that somewhere it said, turn off the hair on most of the skin of the body. That's our level of ignorance about this.

This is what the next century is about. See, we have the text, for the first time. It's incredibly exciting. We can see sentences and nouns and verbs all over the place. We haven't got the plot from all that. We've just barely got the text, we know bits of the language. We don't yet know, subtly, how -- it's embarrassing. I give you a cat and a dog, the two genomes. It wouldn't be easy to tell which is which. We can't tell.

Krulwich: Let me ask you about another important difference. Imagine, if you will, a Sumo wrestler on the one hand, and one of those *Sports Illustrated* bathing suit beauties on the other -- two humans of strikingly different physique. What is the genetic difference between the fat guy and the skinny girl?

Lander: The genetic difference between any two people, whether it's a Sumo wrestler or a *Sports Illustrated* bathing suit model -- one-tenth of a percent. Those two, and any two people on this planet, are 99.9 percent identical at the DNA level. It's only one letter in 1,000 difference.

Krulwich: So that one letter in 1,000 makes him look this way and her look that way?

Lander: Well, of course, him and her has a little to do with X and Y chromosomes, so we've go to take that out of the equation. But the picture I always like to show is Wilt Chamberlain, basketball player, and Willie Shoemaker,

jockey. Both men, 99.9 percent identical, but one of them is almost twice as tall as the other. What's going on? Well, it tells us that, first, as a species, we are very, very closely related, because any two humans being 99.9 percent identical means that we're much more closely related than any two chimpanzees in Africa.

Krulwich: But if you look at an Icelander, that Icelander has such a different physique than an African.

Lander: How many genes does it take to change a physique?

Krulwich: How many genes *does* it take to change a physique?

Lander: Well, we don't know, but we know there's not a lot of gene differences. So maybe it only takes four or five genes to affect skin color. That's certainly the sort of estimates people have made. And height. Oh, okay, maybe there'll be some tweaking of 15 or 20 genes or something. But when you add it all up, it's not that much.

Now, of course, the implications of this is small differences in the genetic code can have big consequences for the appearance or the disease risk. I mean, there's a single letter change, out of three billion letters, that can increase your risk of Alzheimer's disease by 40-fold. So, little differences -- spelling counts.

Krulwich: I don't quite understand. If you are looking at a black woman and you're looking at a white woman, there's a drastic difference to your eye. Are you saying that there is--

Lander: That's accounted for by a smallish number of genes, but that most of the genetic variations that are going on within those populations are common to both populations.

Krulwich: But from the gene's point of view, the two people who look so different to your eye are remarkably similar, as far as the genes are concerned.

Lander: Remarkably similar, and they both might be suffering from the same predisposition to diabetes due to the same genetic variation, for example. They might be suffering from other kinds of conditions that are exactly the same DNA spelling difference. And it's only if you want to focus on those things that are just skin deep that you can say, "Aha, the genes distribute differently in the two populations." Otherwise, for most other things, the genes are continuously present across all these populations.

So race is not a very helpful category to a geneticist, because it's focusing on a fairly small number of genes that describe appearance. But if we're talking about physiology, if

we're talking about the 30,000 genes that run the human symphony, that's a tapestry that weaves through every population. That's why geneticists really don't think race is a terribly helpful concept.

Krulwich: And what about you and Arthur, your brother? What's the difference between you and him?

Lander: He's a better cook, far better cook.

Krulwich: We're talking genetically.

Lander: Oh, genetics.

Krulwich: You come from the same parents, you're clearly not the same people.

Lander: Well, 50 percent of our genes are identical, by descent. So the other 50 percent are as different as any two people on this planet, so I suppose we're twice as close as any other two people on this planet, because we're brothers.

Krulwich: So that's what we think of it -- that, as the closer you are in relationship, the more you have in common, but there's this other whole category, called everything else, which can vary.

Lander: Half of our chromosomes, we got exactly the same copy for that region. The other half, it's as different as you and me.

Krulwich: And identical twins?

Lander: Identical twins, it's 100 percent of the chromosomes are the same. They are genetically identical.

Krulwich: But they don't feel identical at all.

Lander: I didn't say they're identical. They're genetically identical. Whether or not you're identical, well, that depends on your whole life experience, your whole history. They're not identical people. But, at the genetic level, the proteins are the same, the controls are the same. And it just reminds us the outcomes can be quite different.

Krulwich: I want to do one thing before we go to the last category, about sex. Men have something to celebrate, apparently. There is a headline that's good for guys. Tell me the headline.

Lander: One of the cool things you can find by reading the genome sequence is that the rate of mutation is different in sperm than in eggs. You can actually see this by reading the genome story, by looking at how quickly certain things are mutating if they happen to be on the Y chromosome versus on the X chromosome. And by measuring that over 40

million years, we can find out that about twice as much mutation is happening in sperm -- that is, in males passing on their genes -- than in females. And that leads to a wonderful battle of the sexes. The guys I know say, "Aha, men are responsible for two-thirds of all evolutionary progress."

But, of course, women looking at that picture are equally well entitled to say the guys are responsible for two-thirds of all the misery of genetic disease that occurs through those mutations, too. Me, I'm going to be a conscientious abstainer on this question.

Krulwich: Now when your group, the public group, decided to survey all the DNA of a human being, which human beings did you survey?

Lander: So the genome that we sequenced, in fact, comes from about a dozen different people.

Krulwich: A dozen? Were they fat ones, skinny ones, black ones, white ones?

Lander: We don't know who they come from, because, in fact, what happened was--

Krulwich: You don't know who they come from?

Lander: No. Part of the rules was that we don't know and they don't know. See, ads were put in the papers, inviting people to sign up and donate DNA for the Human Genome Project. But we got five or six times as many donors as we were going to use.

Krulwich: What kind of ad? What sort of ad was it? "Please come downtown and give us your blood," or "give us your cheek"--

Lander: Yes, yes. "There's an opportunity to donate DNA for a biomedical research project." And then people had to receive a complete description of the project and sign an informed consent document. The blood sample was taken. And then--

Krulwich: What if a busload of very fat people came up and all nominated themselves, and you had no skinnies? Then you would have the human genome of the fat people.

Lander: Which differs by one tenth of a percent from anybody else.

Krulwich: Well, you don't know that going in.

Lander: Oh, we do. We have a really good sense of genetic variation. And it's not a lot.

Krulwich: Now wait, I'm suspicious here. How do you-- You said you got a mix, 10 or 12 people.

Lander: Yeah.

Krulwich: Were they a mix of-- did they come from, some of them from Borneo and some from America and some from--

Lander: So the genes from which most of the work was done come from Buffalo, New York.

Krulwich: From Buffalo, New York?

Lander: Yes. It's mostly a guy from Buffalo and a woman from Buffalo. But that's because the laboratory that was making--

Krulwich: An anonymous couple from Buffalo?

Lander: They're not a couple. They've never met. And we don't know who they are. So we got a lot of the samples--

Krulwich: And they don't know who they are, except that they--

Lander: The laboratory that prepared the large DNA libraries that were used was a laboratory in Buffalo. And so they put an ad in the Buffalo newspapers, and they got random volunteers from Buffalo, and they got about 20 of them. They then erased all the labels and chose at random this sample and that sample and that sample. So nobody knows who they are. We don't have any links back to who they are, and that's deliberate.

But then to define all the human variation on top of it, we sequenced millions and millions of DNA segments from a worldwide population of 24 people: Pacific Islanders, Asians, Africans, Americans. And that defined more than one-and-a-half million sites of genetic variation in the chromosomes.

Krulwich: So you feel you have a pretty good sample of humanity?

Lander: Oh, I think, about half of the common genetic variation that exists in the human population is already represented in that data base of DNA differences that we found amongst people. And what it does is, it reinforces the message of how these differences are interwoven through the whole planet, through the whole population.

Krulwich: And how many other creatures are we sequencing?

Lander: Oh, goodness, there's so much sequencing going

on today. In terms of large beasts, wow: vertebrates, mice are being sequenced right now, rats are being sequenced. Two different kinds of puffer fish, one zebra fish, various different types of worms, flies, a couple dozen fungi, bacteria, a couple more plants going on. I mean, sequencing is such a great way to get a first look at an organism. You'd be crazy not to do a little sequencing first.

Krulwich: One of the sort of happy futures that we're rolling towards is customization. Tell me a little bit about what might be around the corner in that way. Three people have the same mistake, they go take the same medicine. Two people get well, one doesn't. So what do you do for the third guy?

Lander: Well, what we're coming to understand is that three people who supposedly have the same disease really may have it because of very different reasons. You know, we call it all asthma, but it might be Asthma Type 1 and Asthma Type 2 and Asthma Type 3. And the medicines might work very differently in those people, because you're trying to fix something that's not broken in the third person, and not fix the thing that is broken. Or you're giving someone a medicine that's going to be broken down by their body in a different way, maybe to something that's a toxic side product. Or maybe it'll just get broken down four times as fast, and can't do its job.

We have different physiologies. And understanding those subtle differences in physiology, whether it's the cause of the disease that's different or the reaction to the medicine that's different, I think will be important in matching the right drug to the right patient.

Krulwich: Well, in the same sense as there used to be Pan-American and TWA and Eastern Airlines. They'd fly you everywhere. And there was Grand Union, and there was the A&P, and they would sell you all the food. And there was NBC, CBS, and ABC -- those were the broadcasters. Now medicine can narrow-cast.

Lander: Well, the hope is that medicine can narrow-cast. Of course, you can't make it too narrow, because it's pretty expensive to develop drugs. And if we make it so narrow that you can't actually afford to run the clinical trial, we're not going to get medicines, either. I think the hope is that the broad category of asthma will turn into four types of asthma, and that we'll be able to develop drugs that have a much higher therapeutic index, much higher ability to cure the disease in each of those four cases than before when we were just trying to treat the general symptoms that were in common, but not the causes that were different.

Krulwich: So Hayden [a boy profiled in the NOVA program "Cracking the Code of Life" who has Tay Sachs, a fatal disease caused by a single-letter mutation], like all human beings, has three point something billion chemicals inside

the DNA molecule, and one of the three billion--

Lander: One letter.

Krulwich: One letter.

Lander: One letter. It's I suppose understandable if you think about screws in an airplane. There's an awful lot of screws in an airplane that don't matter if you don't have them in the right place. Maybe your seat back will be a little loose or something. But if it happens to be a screw on the propeller, it can matter a lot. There are some positions that encode the information for unique proteins, which you absolutely have to have. And if they're misspelled you don't have this essential function, and you'll die.

Krulwich: Now that we have this sort of microscopic view, do you have the sense that this is more unfair than we had previously realized? It seems amazing to me that something so slight could result in such a huge difference.

Lander: Genetics is largely random. Every time your DNA is copied usually the right letter is put in, but there's some very small probability the wrong letter gets stuck in, and it's utterly random. Every time a baby is born, about 30 new mutations occur in the sperm and in the egg, giving rise to that baby. It usually makes no difference because these 30 new mutations occur in places that don't matter. But sometimes they occur in places that do matter, and there's no rhyme or reason and no fairness in it. It's chance where they happen to occur. And so, no, it's not fair.

Krulwich: The family says--and we asked them, you know, "What should have happened in your case?" They say quite simply, "Well, we should have tested." But they weren't Jewish, or they weren't obviously in the risk population. So they say, "We really should have tested." In fact, everybody should test for everything, because then we won't have situations like this. What are the consequences of that logic?

Lander: Ideally we should understand everything that could go wrong and be able to offer people the chance to test for whatever they want to. We're nowhere near that today. We don't understand all the roles of all the genes. We don't understand which misspellings really matter. And we also don't understand how to explain to people enough about these choices for them to be able to make intelligent decisions.

In some cases it's pretty clear. If you could know that you're at a very high risk for having a baby with Tay Sachs, I think almost everybody would choose to know. What if you're offered the chance to find out whether your baby might be of high risk for Alzheimer's disease when he or she is 70? Would you want to know? What if that risk is only 50

percent? What if you could be offered a whole range of tests for whether or not you might have not the usual one percent risk for some kind of a cancer but a four percent risk for that kind of cancer?

I don't know quite what to do with it. It might be relevant information. It might be the case that it's sensible to screen somebody twice as often if they've got a four percent than a one percent risk. The challenge is going to be two-fold. Understanding it as science and then understanding it as people. How to deal with not certainties but probabilities, relative risks. We have got a tremendous amount of work to do both in the laboratory and with the public to get to the point where we know how to deal with this information and offer this information in a way that's useful.

As a parent I just imagine what it would have been like if I could have been offered a menu of 10,000 genetic tests to be done. See, there are changes in many, many, many genes, and many of them don't mean anything, they're just irrelevant changes. But you could get back a menu with all of them. Some of them might matter and we can't tell. There will be tremendous anxiety in knowing and there will be tremendous anxiety in not knowing. You might feel irresponsible if you said, "I'm not going to find out." And you would feel tremendous regret if something happened. But you might also feel just tremendous pressure if you found out and doctors couldn't explain to you whether it mattered or not.

We have opened a box here that has got a huge amount of valuable information. It is the key for understanding disease and in the long run to curing disease. But having opened it, we're also going to be very uncomfortable with that information for some time to come. It would be immoral not to get this information, because in it is the secret of cures. But it's going to be very uncomfortable dealing with some of this information for the course of the next century.

We obviously want to know about those things because by knowing the mechanism for breast cancer or colon cancer or Alzheimer's, pharmaceutical companies can try to work on cures. They can try to work on a drug that will slow down the problem or prevent the problem. So we want that knowledge to fashion a cure, but with it comes a certain cloud of uncertainty that can hang, and every person has got to decide, "Do I want to find out about this uncertainty and that uncertainty?"

Right now maybe we only have 10 or 12 uncertainties we can bestow upon people. But fairly soon we may be able to offer hundreds of uncertainties, half which you can do something about, have of which you can't.

The really great challenge for society is how we're going to stand by each other, how we're going to say we all face

exactly that same problem of having potential knowledge about our genomes, but the potential knowledge is different for each of us. Will we unite around the idea that we all are largely facing the same set of issues? Or will we split apart into those people who have this risk, that risk?

I worry a great deal about whether we're going to go down a path where we use this genetic information to classify people, insure them differently, separate people, or instead to unite people by saying, contrary to what they used to say, "There's no perfect genome. None off us is particularly better or more normal. We all have a wide range of variations. Those variations make us spectacularly rich as a species, as in a population. But they also give us each our own particular problems." I think the biggest choice ahead is how we're all going to come to understand the ways in which we're genetically all the same and genetically all different.

Krulwich: Last on this one, I guess I mentioned Justin, who is the son [who also figures in the NOVA program, and whose genetic heritage predisposes him to potentially developing a certain disease later in life)]. His way of handling it is to say, "I just don't want to think about it right now." "I just don't want to think about it right now" seems to me a perfectly acceptable position, but it might be one of the hardest ones to have.

Lander: Oh, yeah. Saying that you don't want to think about a problem right now is a perfectly valid response. If you really were mindful of all the potential problems you could be completely paralyzed by worrying about everything.

Sometimes the happiest life might be lived by just saying, "Whatever happens will happen. I'm going to get everything I can out of life and I'm not going to worry about it." The hard part is where there are tests and diagnostics and therapies that would let you change your risk. If we're talking about something that you can't affect -- well, I don't go out and find out my risk for Alzheimer's disease. I don't want to know because there's nothing I can do, it will just make me worry or it will make me very depressed if I were to get bad news.

But if I were to find out I was at special risk for colon cancer I would really want to know because I really could do something about cutting that risk by twofold or fourfold or tenfold. That's the hard part, is where sticking your head in the sand really does come back to hurt you. At the moment there are a limited number of cases where genetic knowledge can be translated into medical action, but that will probably increase with time.

Krulwich: And my second last question for this whole shebang has to do with the power of genes. I think a lot of people, particularly in a program that is going one to two hours, get the impression that genes are our destiny to

some degree. You live with them and study them all the time. But how fragile are all these predictions? Well, you have this gene, then consequently--

Lander: People often have this sense of genetic determinism -- that they're nothing but their genes, that their genes contain their destiny written in the DNA. This is nonsense. We know that there is a tremendous influence of environment, of society, on outcome.

If you just think about the fact that the human gene pool hasn't significantly changed in the last 5,000 years, and you notice how much society has changed and opportunity has changed, you realize that our genes really can't determine things. They can be influences, but for the longest time people said, "Well, women are incapable of doing something biologically," or "Eastern Europeans aren't smart enough for et cetera, et cetera." People looked to the genome to justify their prejudices about what people could or couldn't do.

But when we look at it there is a huge range of what human societies have done, how they have been organized, what opportunities people have had. That whole range is encompassed within our genome. If that's the amount of limitation in our genome that's okay; I'll live with that.

Krulwich: Do you have any sense, any spiritual sense, that what you've studied reflects anything that isn't on the page?

Lander: Evolution is a pretty mysterious process. It works by randomness and then selection. And I think we tremendously underestimate the power of random experimentation. You think, oh, if I vary things, how could I ever get anything sensible out of it? And when you look at the genome and you see the evidence of churning and mixing and variation and duplication, you realize that this completely undirected evolutionary force of variation generation is incredibly powerful at invention, probably more powerful than design.

No committee of engineers getting together, no matter how smart, would have come up with a human being. And yet evolution, by tinkering and getting it wrong most of the time, but occasionally right, came up with a human being. It's a pretty awesome process. We don't fully understand how it works. But it is really humbling to look at a genome and see what that slow and steady process of incremental improvement has wrought.

Krulwich: So if God is designing into this system, He may or She may or It may have designed the evolutionary process, but there's nothing that you have seen beyond the evolutionary process from your recent work?

Lander: The genome is awe-inspiring, but what the basis of that awe is, we can't tell from that. We can see a remarkable

history of experimentation, of variation. But a lot of it looks like it's random churning and just saving a couple of good things that worked. We can't see more than that in that. One of the great things that--

Krulwich: You save a couple of good things and you save a couple of good things, and it all adds up.

Lander: That's what they tell you, right? A little savings here, a little savings there -- it sort of adds up. Well, there must be greater themes than that, because evolution seems to have found ways to invent efficiently. See, when it wants to make a new gene, it doesn't start from scratch. Maybe back in the primordial ooze it would start from scratch, but then you had a lot of time to work things out.

It seems, by looking at the human genome, that when evolution needs a new gene, it waits for some existing gene to get randomly copied and then to slightly work on it. You build out of existing parts. I think the secret of evolution is that the genetic churning mechanism lets us make theme and variation and theme and variation. And it turns out to be much more powerful than we ever imagined to just make small variations on the theme. You get from a couple little circuits to us.

Krulwich: So your recent emersion in this stuff just makes you think that Darwin was smarter even than he even imagined?

Lander: This gives you a tremendous respect for life. It gives you respect for the complexity of life, the innovation of life, and the tremendous connectivity amongst all life on the planet. You come away from reading the genome recognizing that we are so similar to everything else on the planet.

My take-home message from reading this genome is that we are such a piece of every other living thing on this planet and every innovation in us, we didn't really invent it. These were all things inherited from our ancestors, some of which were yeasts, some of which were slugs. And we're walking around with all these inventions that we got from our forebears, and making good use of it. But we shouldn't be too proud about how wonderful the human is. Very little that's human was invented in humans.

Krulwich: What is the intellectual excitement here [in studying the genome]? By your description, once the DNA sequencing machines are invented, I just go whoop, whishh, pishh! And then like I can take a nap, and then out comes the answers.

Lander: You wish. It's just that, of course, since one is on the cutting edge -- I guess in electronics they always call it "the bleeding edge" -- they know what they're talking about.

because nothing really is working as you expect. All of the stuff we're doing will be working perfectly as soon as we're ready to junk it. So we get the tremendous excitement about working with new ideas, that are almost completely functional. And so you're constantly figuring out why they're not completely functional. It's sort of like flying a very large plane and repairing it while you're flying. It's a challenge, right?

The exhilarating thing about the genome project is we set ourselves the challenge of trying to get this done really fast. If this wasn't important stuff, you'd sit back and say, "Let's take a couple of years, and let's get it all worked out just perfectly. We'll make sure it's humming. We'll work out everything, and it will be flawless."

Of course, the information is too interesting. The insatiable hunger for this information is pressing, and so you say, "Let's get the plane into the air, and we'll work out the rest, 10 percent of the bugs, while we're on our ascent." And it's an interesting challenge.

It attracts intellectual people, because it takes a lot of thinking to figure out what's wrong. It attracts people who sort of love the exhilaration of working together with other people. It attracts people who love the terror of not knowing whether this is all going to fall together. And it also attracts people who realize how much incredible information about medicine and evolution is pouring out the back end, and that if we can only get it to flow perfectly, we have a lifetime's worth of information to pour over, if we can just get it to work.

Krulwich: And you didn't mention money.

Lander: Well, in fact, it flows from our computers straight onto the Internet every 24 hours without patents, without any restrictions whatsoever. The one commitment we made was that taking public money to do this, this was a kind of basic knowledge that 10,000 medical scientists around the world could use every day, 2,000 biotech companies, 30 pharmaceutical companies, and that there's no way that we could do as much with this information all alone as the whole world could do together.

And so the folks who did the genome project decided very early on that we would just put our data on the Web, the information we got every single day, 24 hours a day, and that way, anybody who is working on some brain degeneration or some bowel disease could look at it, and see if they found something really useful.

So when we're writing up the paper about the human genome sequences, we're doing right now, we have a tremendous pleasure. Most scientific papers you speculate on how this stuff is going to be useful. We can write a whole

section about dozens and dozens and dozens of papers on different human diseases that all have worked because of the sequences put out there.

I mean the folks at the genome centers around the world are the tremendous heroes of this project, because everybody did this with no personal gain. Everybody did this because they knew that they were doing the single most important project they could be involved in in their life. And that they would be proud to tell their grandchildren what they spent these years doing. And the way we did it was incompatible with people seeking personal gain out of particular genes. It wouldn't have worked unless everybody rode together.

Krulwich: This doesn't mean, yourself included, that people couldn't go across the street and work for a private company?

Lander: You can go work for a private company. It means you can use this information later, once it's been freely available to everybody to make important discoveries. I have no objection to patents. I have no objection to companies. Our feeling is that the basic information that the letters in the cells were such fundamental infrastructure, such a foundation that that had to be made available without any restrictions, and then companies can go off and use our information, file patents, make millions of dollars. Scientists can do that. Anybody at the genome center can go off and do research on it. But for starters, the basic information has to be available to everybody. And then the value can be added to it. I think that's the right solution for the public.

Krulwich: Well, first of all, did it ever strike you as odd that there you could patent something that we all have in almost all of our cells? Something that's so natural as DNA, that it could become somebody's private property and not mine?

Lander: Well, the patent office says, of course, that you can't patent the gene in the human body. All you can do is get a patent on taking it out of the body and putting it into a bacteria to produce a protein or using the DNA for a diagnostic test. So they profess that you're not really patenting the human gene.

Krulwich: You're taking the stuff out of us and putting it like a museum-like pristine place, and then when it's isolated and pure --

Lander: You might use it to make a protein.

Krulwich: Yes. But it's still me.

Lander: Well, of course it's still you. But the patent office thinks once it's outside of you, it's an invention, it's not you.

Krulwich: That's an odd idea.

Lander: Well, it's an odd idea, but when you get down to it, it's not completely nutty to think about allowing some patents, because some of the proteins in the body can be used as drugs as pharmaceuticals. If we didn't allow some patenting, then a pharmaceutical company, thinking about spending $100 million in order to develop this protein as a drug would say, "It's a stupid idea. Somebody else could come along after we have a successful clinical trial, not have invested the money, and just make the protein." So society says, "All right. We'll grant the monopolies to these inventors."

Krulwich: Which lasts how long?

Lander: Twenty years. It used to be 17 years; it's now 20 years. "And because of that, we're going to try to incent people to make inventions." It's not a bad idea in principle. What bothers me about the patenting system is over the last 10 years, we've been giving away patents for very trivial amounts of work.

I don't object to giving somebody that limited-time monopoly when they've really invented a cure for a disease, some really important therapy. I do object, because I think it's a crummy bargain, for society to giving a monopoly when somebody has simply described a couple hundred letters of a gene, has no idea what it does. Has no idea what use you could have in medicine. Because what's going to happen is you've given away that precious monopoly to somebody who's done a little bit of work, and then the people who want to come along, and do a lot of work, to turn it into a therapy well they've got to go pay the person who already owns it. I think it's a bad deal for society.

Krulwich: I want to make sure I understand this. So if this were back in the 1880's, and it was the mining [industry] -- I see a mountain. I think maybe there's gold up there. I could go file a claim. I don't even have to go up the hill. I don't have to dig a hole, just do it.

Lander: We went through this with the Homestead Act in the middle part of the 1800's. The government said, "We'll give you a big tract of land, if you will work the land for three years." That was a great bargain, because we gave land and people developed its economic uses. If the Homestead Act instead said, "We'll give you land that all you have to do is walk the boundary," we would have accomplished nothing. That's the difference.

If you've got to work your claim, if you've really got to by the sweat of your brow add value for society, I don't object to these limited-time monopolies. But if we're giving away the store for effectively no work, I think society is getting the short end of the bargain.

Krulwich: Let me see what you get. If you haven't done a whole lot, I guess you want the people to do some serious science, some research, some benchmark, actually manipulate these things?

Lander: I want, in order to get a patent, that someone do something that we would all really call useful. Not just that a patent examiner might say is technically satisfying the condition of utility, but that a person on the street would acknowledge, "That is useful."

Krulwich: Were companies patenting segment after segment after segment, doing almost nothing at all to get the patent?

Lander: Oh, yeah. One could write automated programs to file patents on the automated sequences you generated.

Krulwich: No!

Lander: Oh, yes. Sure. Why not?

Krulwich: You'd have a robot do the science--

Lander: You'd have a robot write the patent.

Krulwich: A robot do the research?

Lander: Yeah.

Krulwich: And what would you do; you'd just become a landlord?

Lander: You got to put the stamp on it.

Krulwich: So that's stopped, right? I mean the patent office changed its mind.

Lander: Well, the patent office has been slowly ratcheting up the standards. I honestly don't know if there really is an appropriate level right now. They're certainly imposing more and more of an obligation on a would-be inventor to say that they've done something really meaningful...

Krulwich: Let's say now I have to worry, though, that if I work this particular set that some landlord is going to call me up and say, "I own this. You have to pay me to work here." Isn't that what the landlord does?

Lander: Well, you mostly have to worry if you're a scientist in a company, who is trying to make some important therapeutic, some cure, based on that segment. If you're a research scientist, just doing research, in principle the other company could call you up and tell you to cease and desist. In practice, of course, they're not going to bother. In fact, they're happy you're doing it, because you're adding value.

to their property.

Krulwich: Oh, I got it. Okay.

Lander: You know, you're putting a hotel on their square.

Krulwich: So let's change the argument. I'm a company trying to do work on this, this, and this rung of the ladder, because I think that I can maybe develop a cure for cancer right here, just for the sake of argument. But, of course, I have to worry that somebody owns this thing.

Lander: Oh, you have to worry a lot that this region here that you're working on that might cure cancer has already been patented by somebody else and that patent filing is not public. And so you're living with the shadow that all of your work may go for naught.

Krulwich: Because one day the phone rings, and says, "Sorry, you can't work here. Get off my territory."

Lander: That's right.

Krulwich: Or "You can work here, but I'm going to charge $100,000 a week," or "You can work here, and I'll charge you a nickel, but I want 50 percent of whatever you discover."

Lander: And the problem here is even worse, because many companies don't start the work whenever there's a cloud over who owns that. If there's uncertainty, companies would rather be working some place where they don't have uncertainty, and, therefore, I think work doesn't get done, because of the confusion over who owns stuff.

Krulwich: Is there any place on the genome that is safe to work on, that's unclouded by legal uncertainty?

Lander: No. We're just not aware necessarily of what anybody has done, and so there's no place concrete that's unclouded.

Krulwich: So one argument could be that this patent system that we have frustrates our very basic desire which is to get medicines to people quickly.

Lander: I agree.

Krulwich: Very briefly, I need you to help me describe what the business of these businesses is? In the case of Celera and some of these other companies, they don't actually seem to be in the landlord business. At least to hear them describe it, they're in some other kind of business. What is that business?

Lander: It's interesting. The companies that are involved in

sequencing the genome, and there are really three that have popped up all through the course of the 1990's, could either be seen to be in the landlord business, holding patents on genes, and waving them around when people want to come settle on that gene, or they could be seen as being in the database business. That is having some computer database, where any researcher who wants to gain access to information about the genome could pay a subscription fee and get the data, plus some tools to study it.

Krulwich: Would it tell you the names of the universities that are working that part of the territory, and the names of the professors and their telephone numbers and the articles they have published? And what--

Lander: It depends whether they put that into the database. In the first instance, they could just tell you the letters. Maybe they might also tell you, "We think there are genes here." Maybe they might also tell you, "We think this gene is expressed in your intestines." And maybe they would tell you, "Here're all the papers that have ever been written about this gene." That's valuable information. There are databases of geography of the United States that will tell you, you know, where the cities are, where the buildings are, who lives in what buildings. These are useful databases, and I have no objection to having databases available.

But the raw data, for example, of who is in the phone book, that should be available to everybody. But if you want to build a commercial database, that adds more value and aggregates together those people who live in the same neighborhood and also, you know, buy books from Amazon.com and all these other -- more power to you, that's great.

Krulwich: Finally, what was the most fun that you have had since this whole thing began? Was it standing there at the White House with all these prime ministers?

Lander: No.

Krulwich: Was it the toast you gave, which we have, which is beautiful, where you're looking at the people who helped you get to the billionth base pair, when you had the sort of daddy look on? Or what was it? Was there some intellectual thing?

Lander: I went into biology because I have the sense that it was one place where you could really share your pleasures, where you could really share what you were doing as part of a whole community. That's what I got out of the genome project.

We pulled together through some really challenging, really difficult times, and we did it. We did what we set out to do,

and it's something we're all going to tell our grandchildren about. And it's pretty wonderful. That feeling of having pulled together with other people and accomplished something that matters, let alone something that I think will matter for hundreds of years, that's a pretty wonderful feeling. I couldn't have asked for anything better in all the time I was wandering around figuring out what to do with my career. It's a tremendous gift.

Krulwich: Was this your Mark Spitz U.S. Hockey Team moment, after which there will never be as high a high?

Lander: Oh, golly, no. I think this was a very special time in its own way. It was also a very stressful time, it was a challenging time. But I think it was special. But, you know, the fun is just going to start now. I think it's a different kind of fun. I think we have finally gotten to the other side of this massive genome. We saw this vast hill, this vast mountain standing in front of us 15 years ago. Now here we are. We've gotten over to the other side. And there's this huge new land to explore.

So it was a tremendous feeling to have been part of an expedition going up this mountain and coming down the other side. And that probably won't be replaced by any other experience. But, boy, the experience of seeing this huge valley now just spread out in front of us! That's going to have years and years of fun ahead. I'm really looking forward to that part, too.

Interviews: Collins | Lander | Venter

Photo: WGBH/NOVA

Watch the Program Here | Our Genetic Future (A Survey)
Manipulating Genes: How Much is Too Much? | Understanding Heredity
Explore a Stretch of Code | Nature vs Nurture Revisited
Sequence for Yourself | Journey into DNA | Meet the Decoders
Resources | Update to Program | Teacher's Guide | Transcript
Site Map | Cracking the Code of Life Home

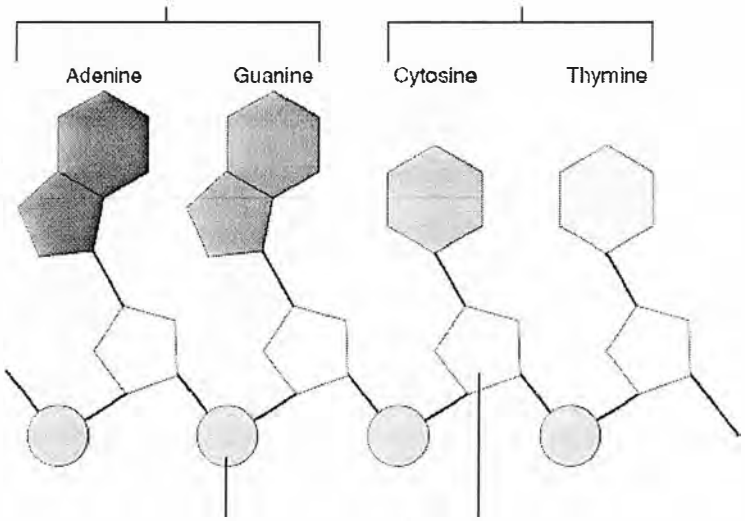
Editor's Picks | Previous Sites | Join Us/E-mail | TV/Web Schedule | About NOVA
Watch NOVAs online | Teachers | Site Map | Shop | Search | To Print
PBS Online | NOVA Online | WGBH

Name _____
Period _____ Date _____

DNA Nucleotides - Component parts of DNA

1. Follow the link to get information on the basic chemical structure of DNA.

2. Label the parts of the diagram below -
(Purine, Pyrimidine, Deoxyribose, Phosphate)



3. Which part is the pentose sugar? (hint: pentose is related to the term "pentagon")

4. What is a "nucleotide"? Name the 3 parts that make up a nucleotide.

5. What part of this polymer would be called the "backbone"?

6. What is a "base pair"?

7. What holds the "base pairs" together?

8. How many hydrogen bonds are there between Adenine and Thymine?

9. How many hydrogen bonds are there between Guanine and Cytosine?

10. Do you think that the numbers of hydrogen bonds matter when DNA forms base pairs? Why or why not?

11. Make a rule that will help you differentiate a Purine from a Pyrimidine.

12. Compare the sizes of the Purines and Pyrimidines. What generalization can you make about their sizes?

13. Do you think that the size of each of the nucleotides matter when they are paired (A-T; C-G) ? Explain how mismatched pairs would result in the wrong spacing between the two backbones when paired together.

ThoughtCo.

What Are the 3 Parts of a Nucleotide? How Are They Connected?

How Nucleotides Are Constructed

by Anne Marie Helmenstine, Ph.D.

Updated August 17, 2018

Nucleotides are the building blocks of the DNA and RNA used as genetic material. Nucleotides also are used for cell signaling and to transport energy throughout cells. You may be asked to name the three parts of a nucleotide and explain how they are connected or bonded to each other. Here's the answer for both DNA and RNA.

Nucleotides in DNA and RNA

Both deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are made up of nucleotides which consist of three parts:

1. **Nitrogenous Base**
   Purines and pyrimidines are the two categories of nitrogenous bases. Adenine and guanine are purines. Cytosine, thymine, and uracil are pyrimidines. In DNA, the bases are adenine (A), thymine (T), guanine (G), and cytosine (C). In RNA, the bases are adenine, thymine, uracil, and cytosine,

2. **Pentose Sugar**
   In DNA, the sugar is 2'-deoxyribose. In RNA, the sugar is ribose. Both ribose and deoxyribose are 5-csrbon sugars. The carbons are numbered sequentially, to help keep track of where groups are attached. The only difference between them is that 2'-deoxyribose has one less oxygen atom attached to the second carbon.

3. **Phosphate Group**
   A single phosphate group is PO_4^{3-}. The phosphorus atom is the central atom. One atom of oxygen is connected to the 5-carbon in the sugar and to the phosphorus atom. When phosphate groups link together to form chains, as in ATP (adenosine

triphosphate), the link looks like O-P-O-P-O-P-O, with two additional oxygen atom attached to each phosphorus, one on either side of the atom.

Although DNA and RNA share some similarities, they are built from slightly different sugars, plus there is a base substitution between them. DNA uses thymine (T), while RNA uses uracil (U). Both thymine and uracil bind to adenine (A).

How Are the Parts of a Nucleotide Connected or Attached?

The base is attached to the primary or first carbon.

The number 5 carbon of the sugar is bonded to the phosphate group. A free nucleotide may have one, two, or three phosphate groups attached as a chain to the 5-carbon of the sugar. When nucleotides connect to form DNA or RNA, the phosphate of one nucleotide attaches via a phosphodiester bond to the 3-carbon of the sugar of the next nucleotide, forming the sugar-phosphate backbone of the nucleic acid.

Instructions for Paper DNA

Students will make a paper model showing how the components of DNA fit together to make a large molecule.

Materials

- Class set of scissors
- School glue
- Printed DNA pieces (Bases, Phosphate, Deoxyribose) on colored paper (6 colors)
- Plain printer paper (one per student)
- Packing tape (optional to tape sections together)

The pieces that make up DNA fit together more or less like a puzzle. All of the shapes need to match to assemble it properly.

1. Print out several of each sheet, each on a different color of pastel paper (you will need about 20 bases per student)
   (Option - print them all out on white paper and have the students color code each part).

2. Each student uses one piece of plain paper to glue as many base pairs as they can on the length of the paper (should be 9 to 10 base pairs). Care should be taken so that each strand has at least one C, G, A, and T.

   Hint: To speed things along, students can cut straight across the notches rather than cutting them out. The tabs should be glued over where the notches would be. You might want to cut and glue a few base pairs together in advance to show students what it might look like.

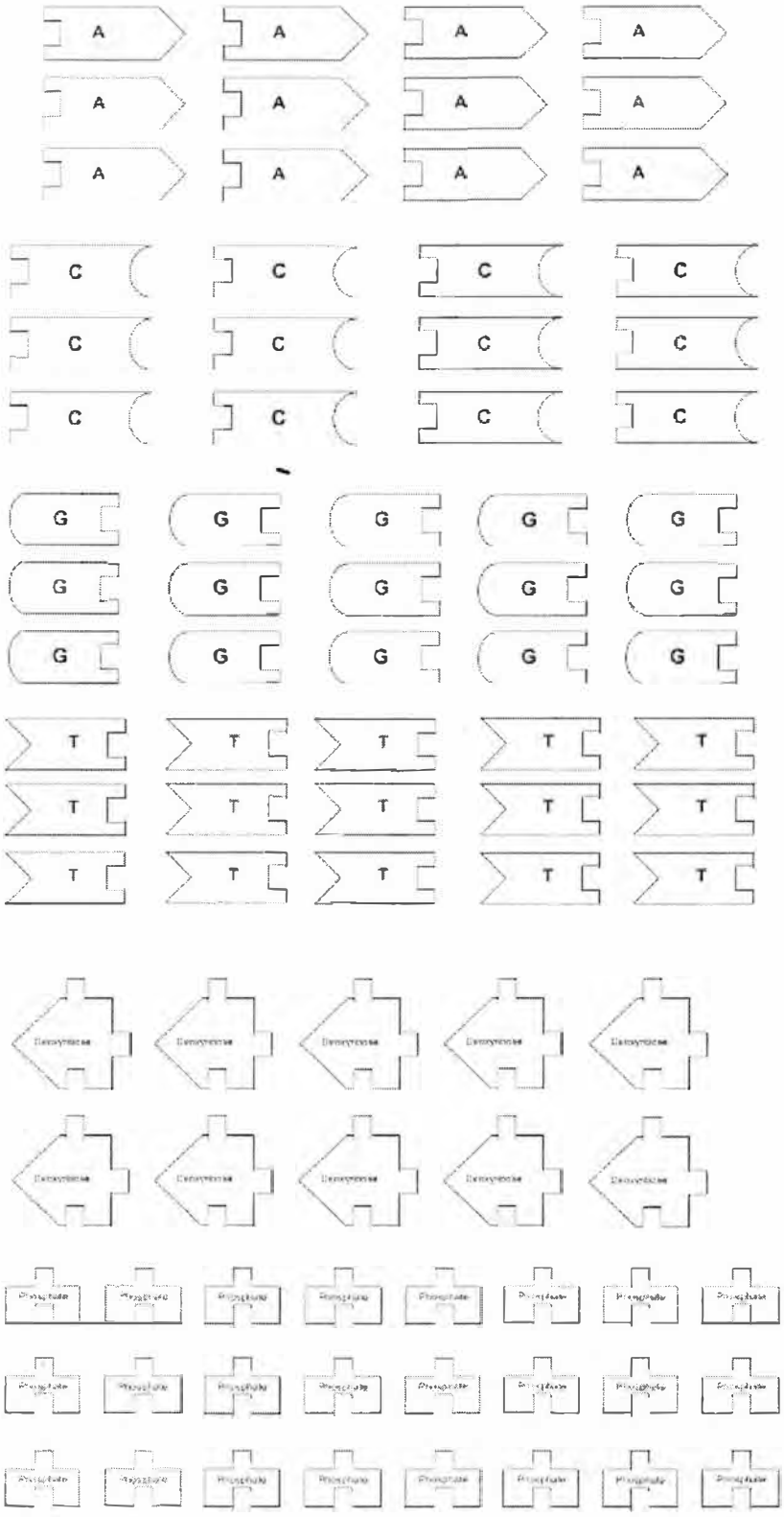
3. When all of the students are done and the glue has dried you can join sections together to make a longer DNA sequence.

Extra Credit - Students can coordinate and copy the DNA sequence for part or all of a human gene.

Assessment -

Structure - Students will write a short report about their mini-project. In it, they will explain in what ways that the model is like DNA (antiparallel, base pairing of purines with pyrimidines, sugar phosphate backbones). They will also explain the ways that the model is different from DNA molecules (mainly, not twisted into a double helix).

Replication - Students will tell how DNA is replicated and explain the characteristics of the molecule that makes errors difficult.


A
C
G
T

Name _____ Period ____ Date ____

Breaking the Code

Find the secret message!

Use the following decoder to decode the hidden message!

Code set	Corresponding letter
111	a
112	b
113	c
114	d
121	e
122	f
123	g
124	h
131	i
132	j
133	k
134	l
141	m
142	n
143	o

Code set	Corresponding letter
144	p
211	q
212	r
213	s
311	t
312	u
313	v
412	w
413	x
414	y
424	z
222	space
223	period
224	question mark
333	exclamation point

311, 124, 121, 212, 121, 222, 111, 212, 121, 222, 311, 124, 212, 121, 121, 222,

114, 131, 123, 131, 311, 213, 222, 144, 121, 212, 222, 113, 121, 212, 222

113, 143, 114, 143, 142, 223, 222, 311, 124, 212, 121, 121, 222,

Breaking the Code

Find the secret message!

Use the following decoder to decode the hidden message!

Code se	Corresponding letter
111	a
112	b
113	c
114	d
121	e
122	f
123	g
124	h
131	i
132	j
133	k
134	l
141	m
142	n
143	o

Code se	Corresponding letter
144	p
211	q
212	r
213	s
311	t
312	u
313	v
412	w
413	x
414	y
424	z
222	space
223	period
224	question mark
333	exclamation point

311, 124, 121, 212, 121, 222, 111, 212, 121, 222, 311, 124, 212, 121, 121, 222,

114, 131, 123, 131, 311, 213, 222, 144, 121, 212, 222, 141, 212, 142, 111, 222

113, 143, 114, 143, 142, 223, 222, 213, 143, 141, 121, 222,

311, 124, 131, 142, 133, 222, 111, 134, 134, 222, 311, 124, 121,

213, 124, 121, 121, 311, 213, 222, 111, 212, 121, 222, 311, 124, 121, 222,

213, 111, 141, 121, 223

Breaking the Code

Find the secret message!

Use the following decoder to decode the hidden message!

Code se	Corresponding letter
111	a
112	b
113	c
114	d
121	e
122	f
123	g
124	h
131	i
132	j
133	k
134	l
141	m
142	n
143	o

Code se	Corresponding letter
144	p
211	q
212	r
213	s
311	t
312	u
313	v
412	w
413	x
414	y
424	z
222	space
223	period
224	question mark
333	exclamation point

311, 124, 121, 212, 121, 222, 111, 212, 121, 222, 311, 124, 212, 121, 121, 222,

114, 131, 123, 131, 311, 213, 222, 144, 121, 212, 222, 141, 212, 142, 111, 222

113, 143, 114, 143, 142, 223, 222, 121, 111,113, 143, 222, 113, 143, 114, 143,

213, 144, 121, 113, 131, 122, 131, 121, 213, 222, 143, 142, 121, 222,

111, 141, 131, 142, 143, 222, 111, 113, 131, 114, 223

Name_____ Date_____

DNA Extraction: Strawberry

Background: The long, thick fibers of DNA store the information for the functioning of the chemistry of life. DNA is present in every cell of plants and animals. The DNA found in strawberry cells can be extracted using common, everyday materials. We will use an extraction buffer containing salt, to break up protein chains that bind around the nucleic acids, and dish soap to dissolve the lipid (fat) part of the strawberry cell wall and nuclear membrane. This extraction buffer will help provide us access to the DNA inside the cells.

Pre-lab questions:

1. What do you think the DNA will look like?

2. Where is DNA found?

Materials:

heavy duty ziploc bag
1 strawberry
10 mL DNA extraction buffer (soapy, salty water)
cheesecloth
funnel
50mL vial / test tube
glass rod, inoculating loop, or popsicle stick
20 mL ethanol

Procedure:

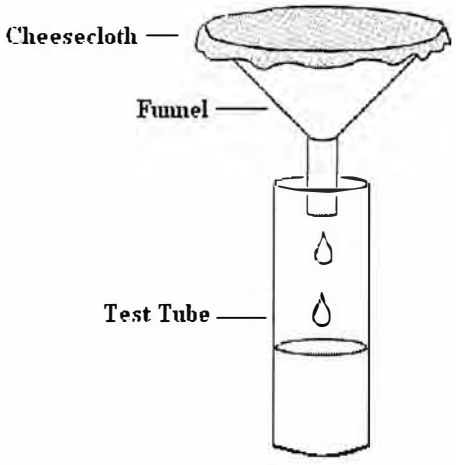
1. Place one strawberry in a Ziploc bag.

2. Smash/grind up the strawberry using your fist and fingers for 2 minutes. *Careful not to break the bag!!*

3. Add the provided 10mL of extraction buffer (salt and soap solution) to the bag.

4. Kneed/mush the strawberry in the bag again for 1 minute.

5. Assemble your filtration apparatus as shown to the right.



6. Pour the strawberry slurry into the filtration apparatus and let it drip directly into your test tube.

7. Slowly pour cold ethanol into the tube. OBSERVE ☺

8. Dip the loop or glass rod into the tube where the strawberry extract and ethanol layers come into contact with each other. OBSERVE ☺

Conclusions and Analysis

1. It is important that you understand the steps in the extraction procedure and why each step was necessary. Each step in the procedure aided in isolating the DNA from other cellular materials. Match the procedure with its function:

PROCEDURE	FUNCTION
A. Filter strawberry slurry through cheesecloth	___ To precipitate DNA from solution
B. Mush strawberry with salty/soapy solution	___ Separate components of the cell
C. Initial smashing and grinding of strawberry	___ Break open the cells
D. Addition of ethanol to filtered extract	___ Break up proteins and dissolve cell membranes

2. What did the DNA look like? Relate what you know about the chemical structure of DNA to what you observed today.

3. Explain what happened in the final step when you added ethanol to your strawberry extract. (*Hint*: *DNA is soluble in water, but not in ethanol*)

4. A person cannot see a single cotton thread 100 feet away, but if you wound thousands of threads together into a rope, it would be visible much further away. Is this statement analogous to our DNA extraction? Explain.

5. Why is it important for scientists to be able to remove DNA from an organism? List two reasons.

6. Is there DNA in your food? _____ How do you know?

Gene Expression

Genes encode proteins and proteins dictate cell function. Therefore, the thousands of genes expressed in a particular cell determine what that cell can do. Moreover, each step in the flow of information from DNA to RNA to protein provides the cell with a potential control point for self-regulating its functions by adjusting the amount and type of proteins it manufactures.

At any given time, the amount of a particular protein in a cell reflects the balance between that protein's synthetic and degradative biochemical pathways. On the synthetic side of this balance, recall that protein production starts at transcription (DNA to RNA) and continues with translation (RNA to protein). Thus, control of these processes plays a critical role in determining what proteins are present in a cell and in what amounts. In addition, the way in which a cell processes its RNA transcripts and newly made proteins also greatly influences protein levels.

How Is Gene Expression Regulated?

The amounts and types of mRNA molecules in a cell reflect the function of that cell. In fact, thousands of transcripts are produced every second in every cell. Given this statistic, it is not surprising that the primary control point for gene expression is usually at the very beginning of the protein production process — the initiation of transcription. RNA transcription makes an efficient control point because many proteins can be made from a single mRNA molecule.

Transcript processing provides an additional level of regulation for eukaryotes, and the presence of a nucleus makes this possible. In prokaryotes, translation of a transcript begins before the transcript is complete, due to the proximity of ribosomes to the new

mRNA molecules. In eukaryotes, however, transcripts are modified in the nucleus before they are exported to the cytoplasm for translation.

Eukaryotic transcripts are also more complex than prokaryotic transcripts. For instance, the primary transcripts synthesized by RNA polymerase contain sequences that will not be part of the mature RNA. These intervening sequences are called **introns**, and they are removed before the mature mRNA leaves the nucleus. The remaining regions of the transcript, which include the protein-coding regions, are called **exons**, and they are spliced together to produce the mature mRNA. Eukaryotic transcripts are also modified at their ends, which affects their stability and translation.

Of course, there are many cases in which cells must respond quickly to changing environmental conditions. In these situations, the regulatory control point may come well after transcription. For example, early development in most animals relies on translational control because very little transcription occurs during the first few cell divisions after fertilization. Eggs therefore contain many maternally originated mRNA transcripts as a ready reserve for translation after fertilization (Figure 1).

On the degradative side of the balance, cells can rapidly adjust their protein levels through the enzymatic breakdown of RNA transcripts and existing protein molecules. Both of these actions result in decreased amounts of certain proteins. Often, this breakdown is linked to specific events in the cell. The eukaryotic cell cycle provides a good example of how protein breakdown is linked to cellular events. This cycle is divided into several phases, each of which is characterized by distinct **cyclin** proteins that act as key regulators for that phase. Before a cell can progress from one phase of the cell cycle to the next, it must degrade the cyclin that characterizes that particular phase of the cycle. Failure to degrade a cyclin stops the cycle from continuing.

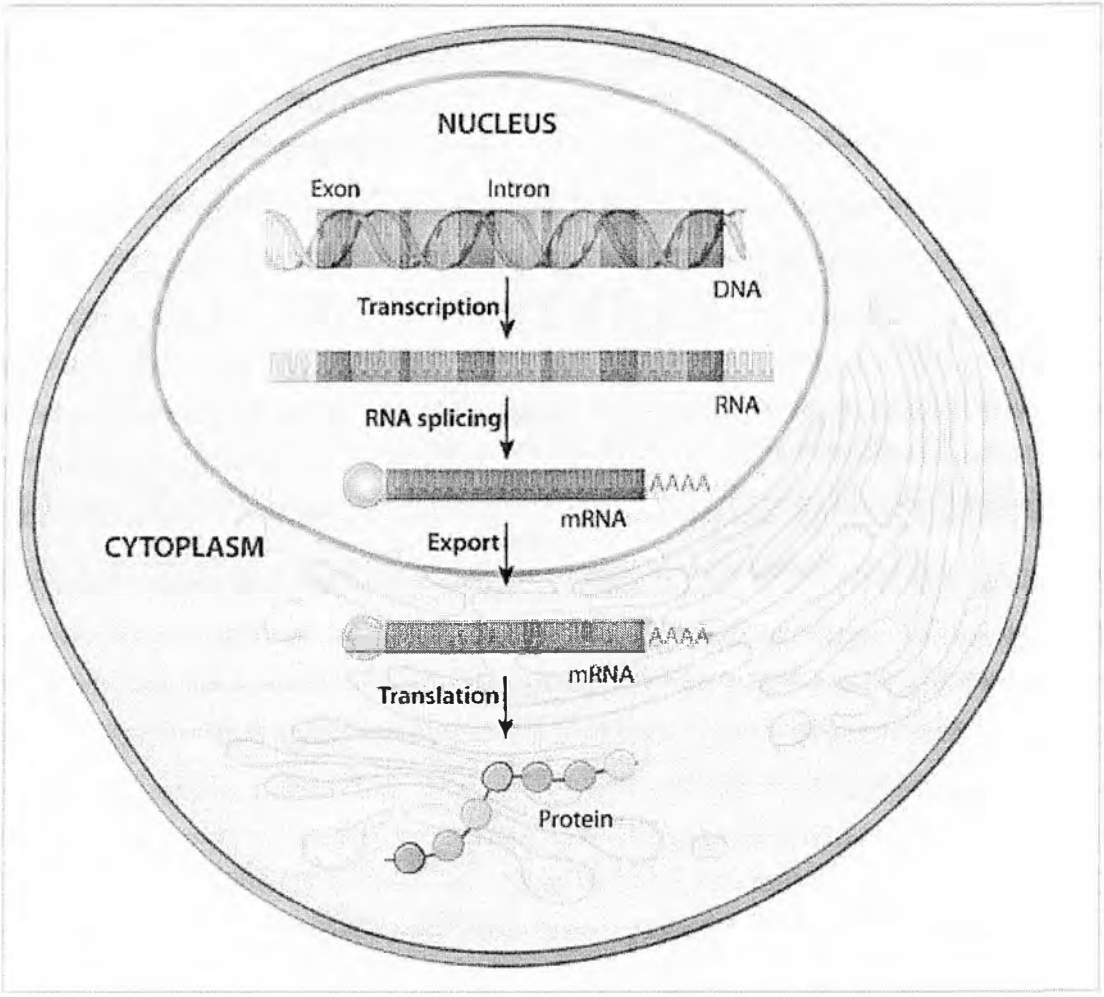


Figure 1: An overview of the flow of information from DNA to protein in a eukaryote

First, both coding and noncoding regions of DNA are transcribed into mRNA. Some regions are removed (introns) during initial mRNA processing. The remaining exons are then spliced together, and the spliced mRNA molecule (red) is prepared for export out of the nucleus through addition of an endcap (sphere) and a polyA tail. Once in the cytoplasm, the mRNA can be used to construct a protein.

Figure Detail

A schematic diagram shows transcription, RNA splicing, nuclear export, and translation within the outline of a eukaryotic cell with a large nucleus. At the top of the diagram, within the nucleus, is a grey DNA double helix. A transparent, rectangular box is drawn on top of most of the double helix. The rectangular box is shaded with two alternating colors; the purple segments represent exons, and the light-green segments represent introns. A black arrow labeled "transcription" points downward from the DNA molecule showing that the boxed region of double-stranded DNA is transcribed into a pre-mRNA molecule. The pre-mRNA molecule is shown as a grey, single-stranded RNA molecule made up of a linear backbone with vertical rectangles arranged along its length. The tops of the rectangles are either pointed, rounded, cupped, or V-shaped to represent different nucleotides. A transparent, rectangular box is drawn over most of the pre-mRNA molecule, with red regions that align with the purple DNA exons and light-green regions that align with the light-green introns in the DNA template. A thin, black arrow labeled RNA splicing points downward from the pre-mRNA molecule to a mature mRNA molecule that contains only red exons. The mature mRNA also has a light peach-colored sphere attached to its left end to represent the 5-prime cap, and four adenosine molecules attached to its right end to represent the poly-A tail. Each adenosine is represented by a red letter "A." Nuclear export is represented by a thin black arrow labeled "export" that points from the processed mRNA molecule inside the nucleus down to a processed mRNA molecule in the cytoplasm. In the background of the cytoplasm, thin black lines show silhouettes of cytoplasmic organelles, including the Golgi apparatus and endoplasmic reticulum. Translation of mRNA into protein occurs in the cytoplasm. A thin black arrow labeled "translation" points from the processed mRNA molecule in the cytoplasm to a protein, which is depicted as a chain of seven differently colored spheres, each one representing a different amino acid.

How Do Different Cells Express the Genes They Need?

Only a fraction of the genes in a cell are expressed at any one time. The variety of gene expression profiles characteristic of different cell types arise because these cells have distinct sets of transcription regulators. Some of these regulators work to increase transcription, whereas others prevent or suppress it.

Normally, transcription begins when an RNA polymerase binds to a so-called **promoter sequence** on the DNA molecule. This sequence is almost always located just upstream from the starting point for transcription (the 5' end of the DNA), though it can be located downstream of the mRNA (3' end). In recent years, researchers have discovered that other DNA sequences, known as **enhancer sequences**, also play an important part in transcription by providing binding sites for regulatory proteins that affect RNA polymerase activity. Binding of regulatory proteins to an enhancer sequence causes a shift in chromatin structure that either promotes or inhibits RNA polymerase and transcription factor binding. A more open chromatin structure is associated with active gene transcription. In contrast, a more compact chromatin structure is associated with transcriptional *in*activity (Figure 2).

Some regulatory proteins affect the transcription of multiple genes. This occurs because multiple copies of the regulatory protein binding sites exist within the genome of a cell. Consequently, regulatory proteins can have different roles for different genes, and this is one mechanism by which cells can coordinate the regulation of many genes at once.

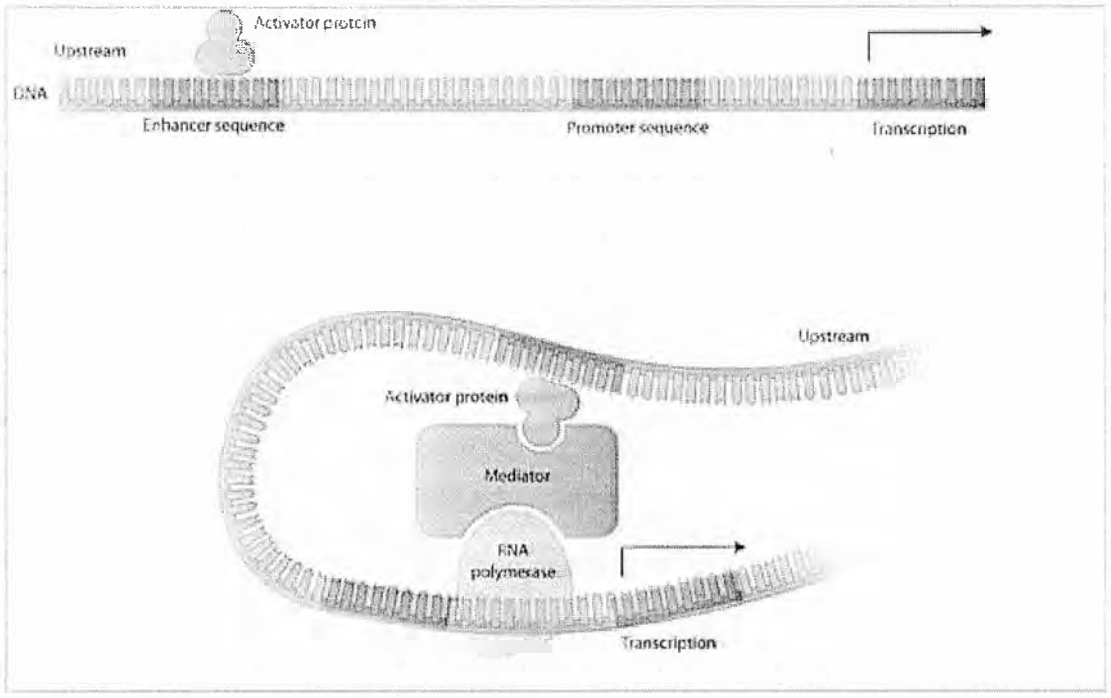


Figure 2: Modulation of transcription

An activator protein bound to DNA at an upstream enhancer sequence can attract proteins to the promoter region that activate RNA polymerase (green) and thus transcription. The DNA can loop around on itself to cause this interaction between an activator protein and other proteins that mediate the activity of RNA polymerase.

Figure Detail

A linear DNA molecule (top) looks like a horizontal bar from which many parallel rectangular nucleotides protrude. The left-hand terminus of the molecule is labeled "upstream." An "enhancer sequence" spanning 9 nucleotides is shaded green and bound to an activator protein. Downstream, a 9-nucleotide "promoter sequence" is also shaded green. Further downstream, the site of transcription spans 9 nucleotides and is shaded orange. An arrow above this orange region points to the right, indicating the direction transcription proceeds. A DNA molecule folded into a loop shape (bottom) shows the enhancer sequence parallel to and nearly aligned with the promoter sequence. RNA polymerase is bound to the DNA between the promoter sequence and the site of transcription. A mediator protein bridges the

gap between the activator protein and RNA polymerase by binding to both of them, allowing an interaction between the activator and RNA polymerase to occur.

How Is Gene Expression Increased or Decreased in Response to Environmental Change?

In prokaryotes, regulatory proteins are often controlled by nutrient availability. This allows organisms such as bacteria to rapidly adjust their transcription patterns in response to environmental conditions. In addition, regulatory sites on prokaryotic DNA are typically located close to transcription promoter sites — and this plays an important part in gene expression.

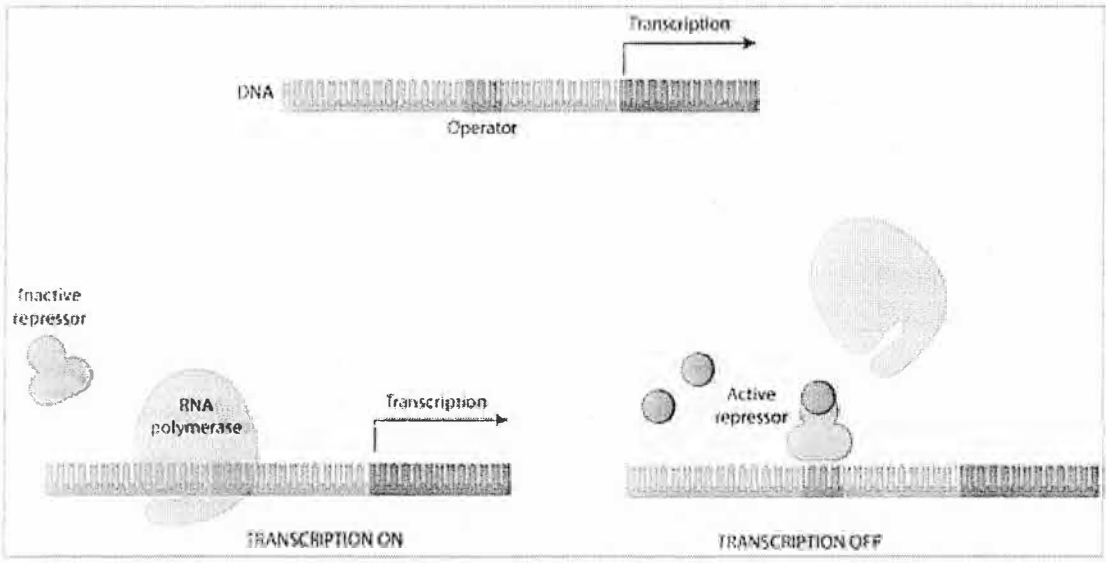


Figure 3: Transcription repression near the promoter region.

Molecules can interfere with RNA polymerase binding. An inactive repressor protein (blue) can become activated by another molecule (red circle). This active repressor can bind to a region near the promoter called an operator (yellow) and thus interfere with RNA polymerase binding to the promoter, effectively preventing transcription.

Figure Detail

Part 1 of this three-part schematic shows the location of the operator and site of transcription on a DNA molecule. In part 2, RNA polymerase is shown bound to an 11 nucleotide-long region of DNA; part of this region is considered the operator. An inactive repressor protein is represented as a globular shape in proximity to the DNA molecule. In part 3, three activator molecules are shown: one is bound to the active repressor, which is oriented above the operator. RNA polymerase is blocked by the repressor molecule from binding in proximity to the operator, and is shown floating unattached above the DNA molecule.

For an example of how this works, imagine a bacterium with a surplus of amino acids that signal the turning "on" of some genes and the turning "off" of others. In this particular example, cells might want to turn "on" genes for proteins that metabolize amino acids and turn "off" genes for proteins that synthesize amino acids. Some of these amino acids would bind to positive regulatory proteins called **activators.** Activator proteins bind to regulatory sites on DNA nearby to promoter regions that act as on/off switches. This binding facilitates RNA polymerase activity and transcription of nearby genes. At the same time, however, other amino acids would bind to negative regulatory proteins called **repressors**, which in turn bind to regulatory sites in the DNA that effectively block RNA polymerase binding (Figure 3).

The control of gene expression in eukaryotes is more complex than that in prokaryotes. In general, a greater number of regulatory proteins are involved, and regulatory binding sites may be located quite far from transcription promoter sites. Also, eukaryotic gene expression is usually regulated by a combination of several regulatory proteins acting together, which allows for greater flexibility in the control of gene expression.

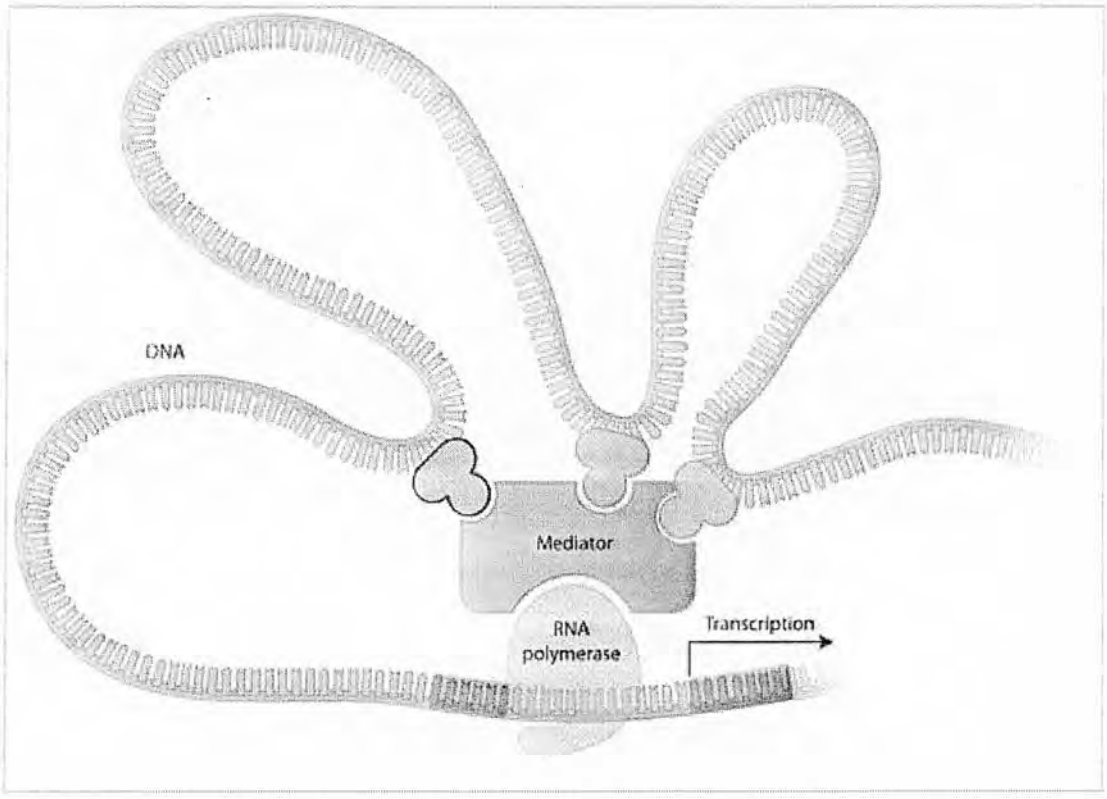


Figure 4: The complexity of multiple regulators

Transcriptional regulators can each have a different role. Combinations of one, two, or three regulators (blue, green, and yellow shapes) can affect transcription in different ways by differentially affecting a mediator complex (orange), which is also composed of proteins. The effect is that the same gene can be transcribed in multiple ways, depending on the combination, presence, or absence of various transcriptional regulator proteins.

As previously mentioned, enhancer sequences are DNA sequences that are bound by an activator protein, and they can be located thousands of base pairs away from a promoter, either upstream or downstream from a gene. Activator protein binding is thought to cause DNA to loop out, bringing the activator protein into physical proximity

with RNA polymerase and the other proteins in the complex that promote the initiation of transcription (Figure 4).

Different cell types express characteristic sets of transcriptional regulators. In fact, as multicellular organisms develop, different sets of cells within these organisms turn specific combinations of regulators on and off. Such developmental patterns are responsible for the variety of cell types present in the mature organism (Figure 5).

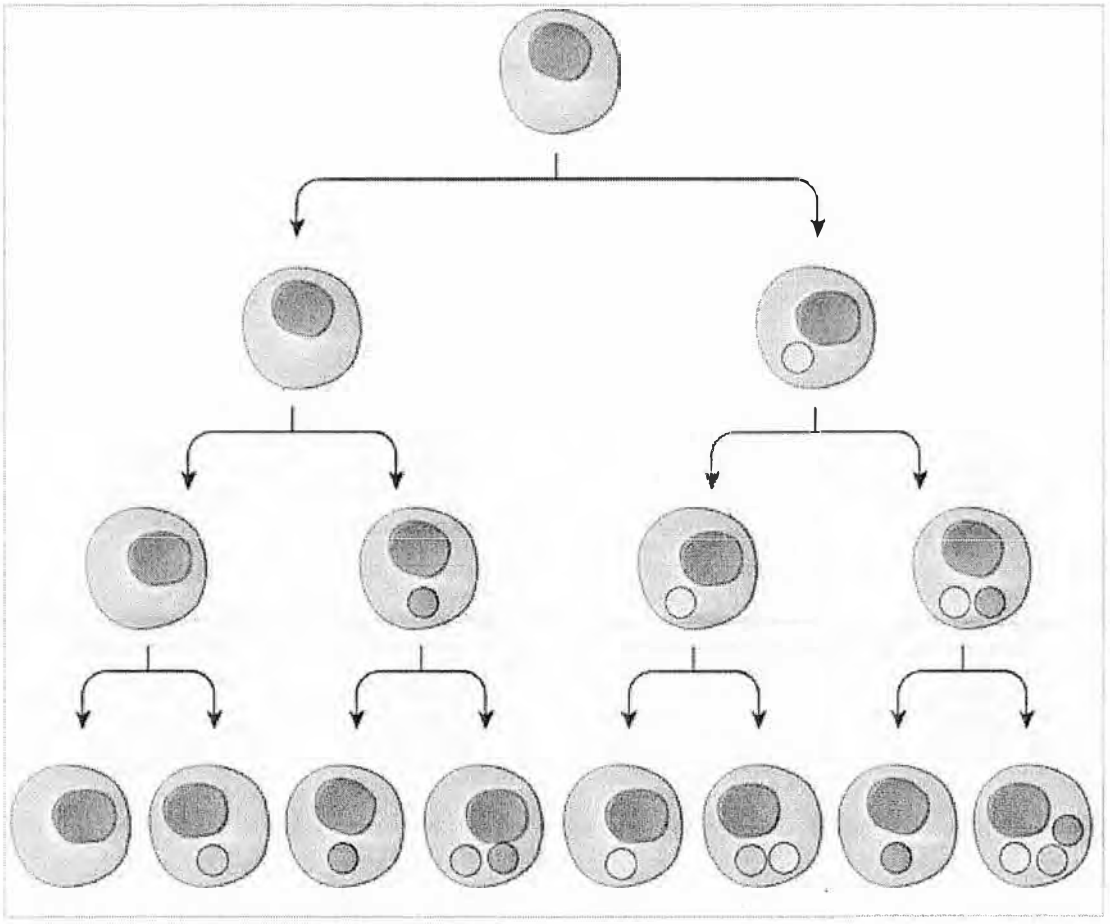


Figure 5: Transcriptional regulators can determine cell types

The wide variety of cell types in a single organism can depend on different transcription factor activity in each cell type. Different transcription factors can turn on at different times during successive generations of cells. As cells mature and go through different stages (arrows), transcription factors (colored balls) can act on gene expression and change the cell in different ways. This change affects the next generation of cells derived from that cell. In subsequent generations, it is the combination of different transcription factors that can ultimately determine cell type.

Figure Detail

A schematic diagram shows a pedigree of cells that contain various transcription factors, which are depicted as different colored circles. As the cells divide, they produce new generations of cells. All of the cells from a given generation are shown in the same row. At the top of the diagram, a single cell is shown as a light purple circle with a dark-purple, circular nucleus inside it. This cell does not express any transcription factors. A thin black arrow with two branches points downward from this first cell to show that it divides to form two second-generation cells. The second-generation cell on the left does not express a transcription factor, but the second-generation cell on the right side contains a yellow circle in its cytosol to show that it expresses a transcription factor. Thin black arrows with two branches point down from each of the two second-generation cells to show that each divides to form two third-generation cells. This makes a total of four third-generation cells. The second-generation cell on the left that did not express a transcription factor produces a third-generation cell that does not express a transcription factor and a third-generation cell with a red transcription factor. The second-generation cell on the right that expressed a yellow transcription factor divides to form a third-generation cell with a single, yellow transcription factor and a third-generation cell with both a red and a yellow transcription factor. Thin black arrows with two branches point down from each of the four third-generation cells to show that each divides to form two fourth-generation cells. This makes a total of eight fourth-generation cells. The third-generation cell that does not express any transcription factors divides to produce a fourth-generation cell with no transcription factors and a fourth-generation cell with a green transcription factor. The third generation cell expressing a single, red transcription factor divides to produce a fourth-generation cell with a single, red transcription factor and a fourth-generation cell with both a red and a green transcription

factor. The third-generation cell expressing a single yellow transcription factor produces a fourth-generation cell with a single yellow transcription factor and a fourth-generation cell with both a green and a yellow transcription factor. The third-generation cell expressing both a yellow and a red transcription factor produces a fourth-generation cell with a single, red transcription factor and a fourth-generation cell with a yellow, a green, and a red transcription factor.

Conclusion

To live, cells must be able to respond to changes in their environment. Regulation of the two main steps of protein production — transcription and translation — is critical to this adaptability. Cells can control which genes get transcribed and which transcripts get translated; further, they can biochemically process transcripts and proteins in order to affect their activity. Regulation of transcription and translation occurs in both prokaryotes and eukaryotes, but it is far more complex in eukaryotes.

eBooks
This page appears in the following eBook
Essentials of Cell Biology, Unit 2.3
Cell Biology for Seminars, Unit 2.3

Homeotic Genes and Body Patterns

From Learn Genetics - Genetic Science Learning Center

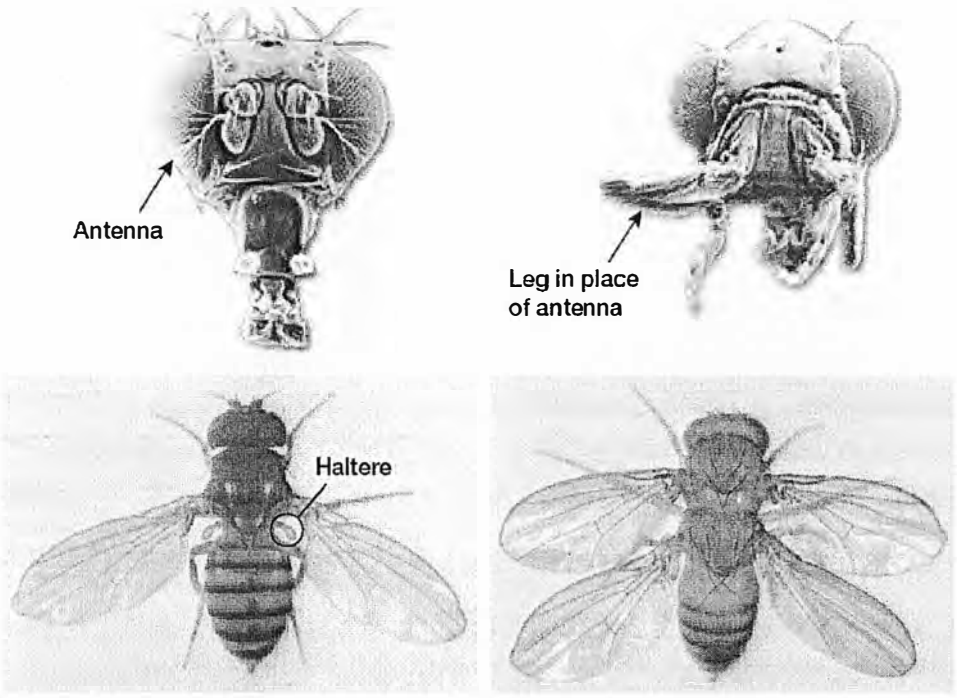
Every organism has a unique body pattern. Although specialized body structures, such as arms and legs, may be similar in makeup (both are made of muscle and bone), their shapes and details are different. While an embryo grows, arms and legs develop differently due to the actions of homeotic genes, which specify how structures develop in different segments of the body.

How did scientists discover genes that determine body pattern?

Scientists discovered homeotic genes by studying strange transformations in fruit flies, including flies that had feet in place of mouthparts, extra pairs of wings, or two pairs of balance organs (called halteres) instead of wings. Some even had legs growing out of their heads in place of antennae!

Scientists called these modifications "homeotic transformations," because one body part seemed to have been replaced by another. Researchers, including a group headed by Ed Lewis at Caltech, discovered that many of these transformations were caused by defects in single genes, which they termed homeotic, or Hox, genes.

This work demonstrated that antennal cells carry all of the information necessary to become leg cells. This is a general principle: every cell in an organism carries, within its DNA, all of the information necessary to build the entire organism.



Top: (Left) Normal fruitfly; (Right) Fruitfly with mutation in antennapedia gene Bottom: (Left) Normal fruitfly; (Right) Fruitfly with a homeotic mutation that gives it two thoraxes. Bottom images courtesy of the Archives, California Institute of Technology.

Shared characteristics

Fruit flies begin life as worm-like creatures made up of repeating units, or segments. Early in development, Hox genes are switched on in different segments. Patterns of Hox gene activity give each segment an identity, telling it where it is in the body and what structures it should grow. For instance, genes that are active in the head direct the growth of mouthparts and antennae, while genes that are active in the thorax direct the growth of legs and wings.

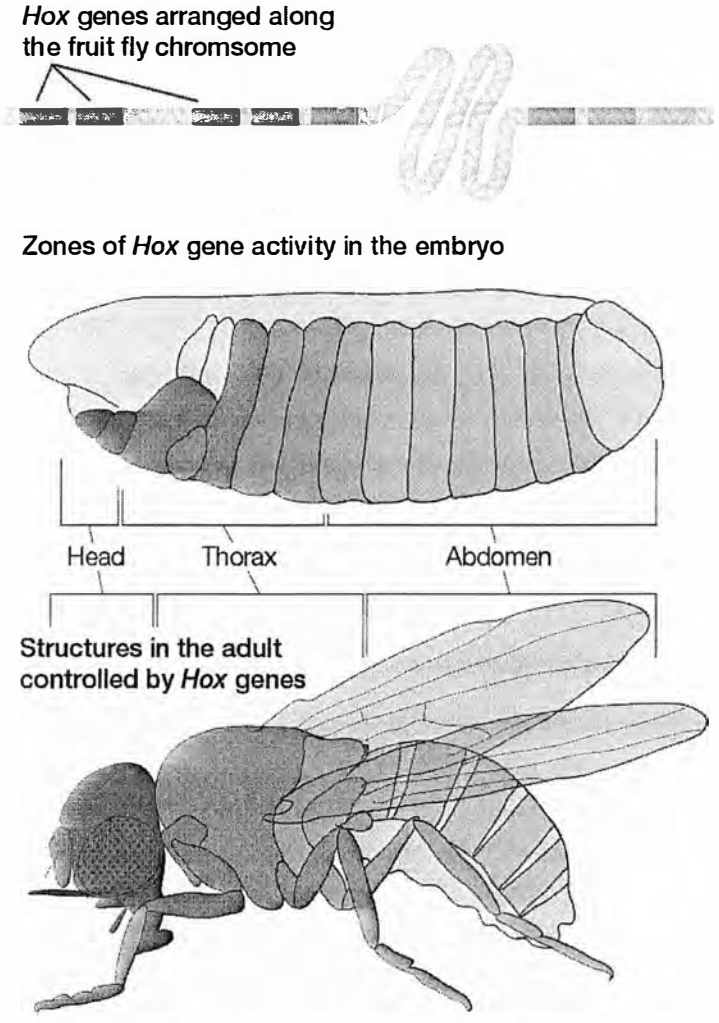
Changes to Hox gene expression change a segment's identity. For example the first segment of the thorax normally grows legs, the second grows legs and wings, and the third grows legs and halteres. When the Hox gene activity in the third segment is made the same as that in the second, both segments grow legs and wings (see photos above).

While studying the DNA sequences of homeotic genes in fruit flies, researchers found that they all shared a similar stretch of about 180 bases; they named this stretch the homeobox. The homeobox is just a portion of each gene. If the words below were

homeotic genes, the capital letters would represent the homeobox: togeTHEr - THEoretical - gaTHEring - boTHEr

Researchers used DNA-sequence similarity to find genes with homeoboxes in other species, including other insects, worms, and even mammals. Together, these genes make up the Hox gene family (Hox is short for homeobox).

Interestingly, Hox genes are arranged in clusters. Typically, their order on the chromosome is the same as the order in which they appear along the body. In other words, the genes on the left control patterning in the head, and the genes on the right control patterning in the tail.



Genes in different organisms that share similar sequence and function are called homologous genes.

Genes hold clues about evolutionary relationships

Nearly every animal that's been tested has homeobox sequences in its DNA, suggesting that Hox genes arose very early in the animal family tree. Genetic sequences maintained over evolutionary time are thought to be especially important to the basic development of even distantly related organisms. The presence of homeotic gene sequences in animals as different as jellyfish, insects, and mammals suggests that these genes have a crucial function in many, and perhaps all, animals.

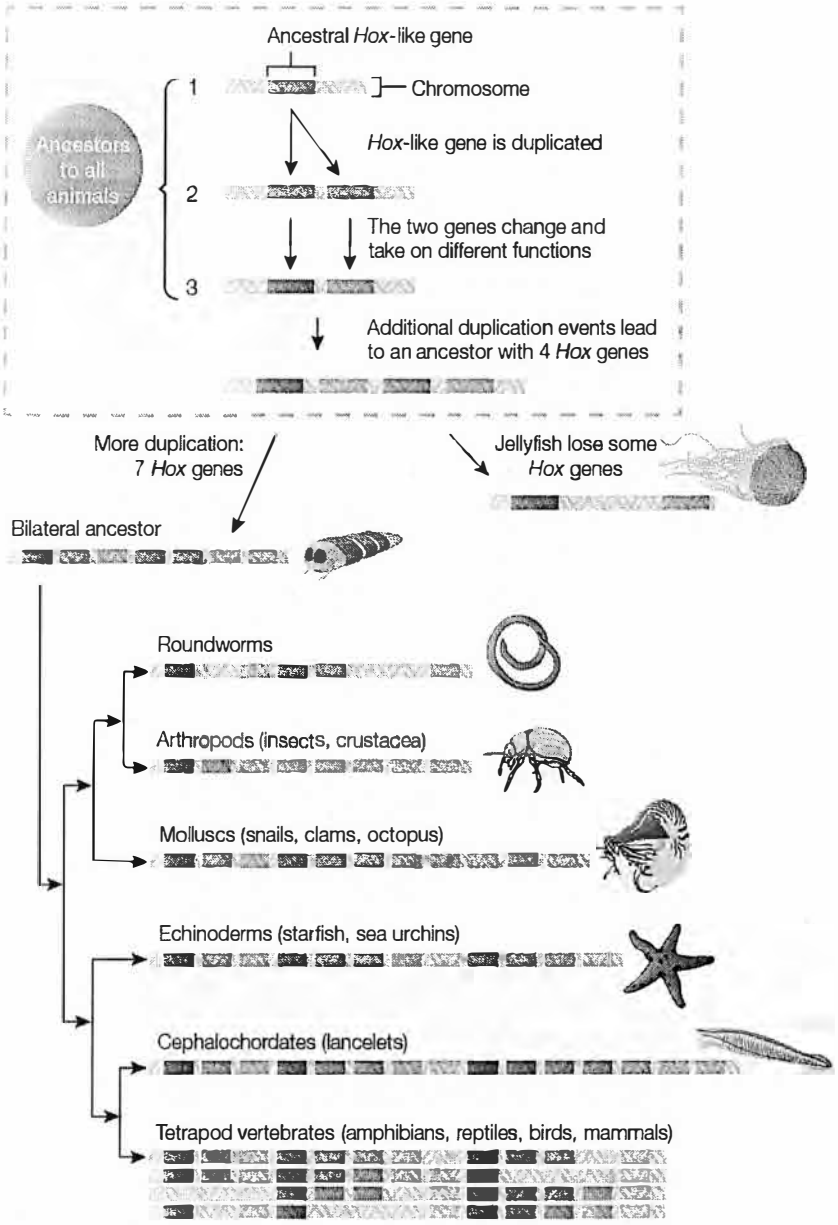
Scientists have studied the genes' DNA sequences, functions, and organization to learn about evolutionary relationships. Hox genes have revealed many clues about the evolution of the animal family tree.

The similarities among Hox genes, especially in the shared homeobox sequence, suggest that they all arose from a single ancestral gene that was duplicated multiple times. After each duplication event, the genes gradually changed, taking on slightly different jobs. This process is known to evolutionary biologists as "duplication and divergence."

The first duplications happened a very long time ago. An animal that lived about a billion years ago, the ancestor to all animals, had at least 4 Hox genes. By 600 million years ago, in the ancestor to all modern animals that have bilateral symmetry, the number grew to at least 7. We know this because animals descended from this ancestor have homologues of these genes.

Additional duplication events happened in some branches of the animal family tree. In insects, for example, a gene near the right end of the cluster was duplicated. In vertebrates, the entire Hox cluster was duplicated—three times in mammals and up to 8 times in some types of fish. The duplicate genes were then free to take on new functions, often leading to more-complex body structures.

Hox genes share not only DNA sequence, they even share functions. Mouse Hox genes can substitute for their homologues in flies. And when they're activated in other segments, the mouse genes can cause homeotic transformations in flies.



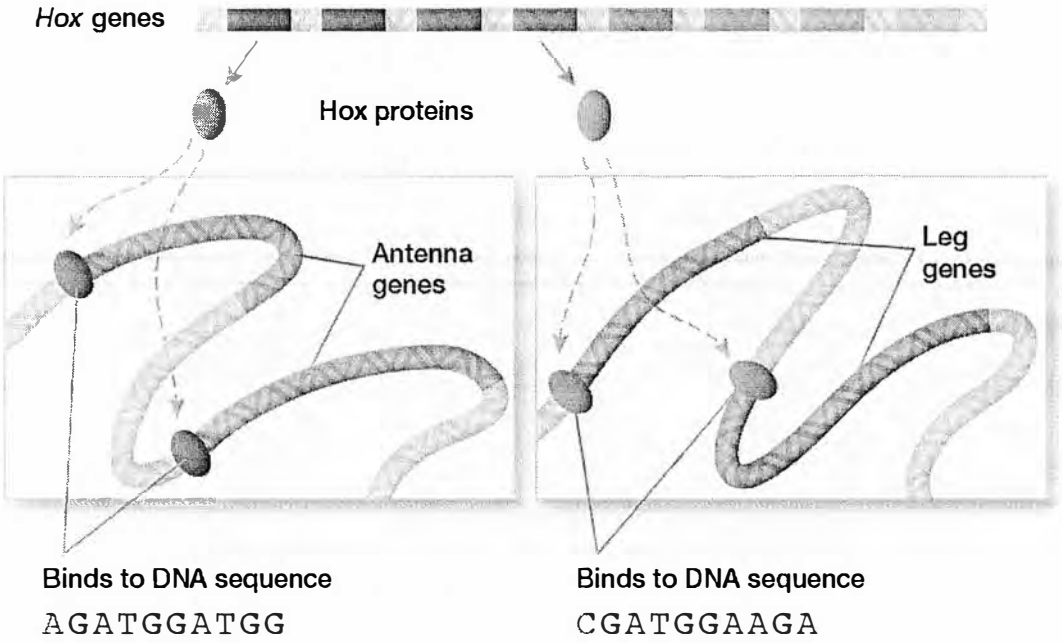
Like other genes, Hox genes are more similar in closely related species and less similar in more distantly related species. By comparing sequence similarity, scientists can determine when in evolutionary history certain duplication events happened, and where some Hox genes were lost along the way (additional gains and losses have happened within individual species in each group).

Illustration based on information from Pascual-Anaya et al (2005), Carroll et al (2005), and Garcia-Fernandez (2013).

Hox proteins regulate other genes

Hox genes code for proteins that attach to molecular switches on DNA, turning other genes on and off. The DNA-binding piece of a Hox protein is called the homeodomain, and it's encoded by the homeobox. The homeodomains in different Hox proteins are similar but not identical—they bind to different DNA sequences. So different Hox proteins regulate different sets of genes, and combinations of Hox proteins working together to regulate still other sets of genes.

As regulators of other genes, Hox proteins are very powerful. A single Hox protein can regulate the activity of many genes. And sets of genes work together to carry out "programs" during embryonic development—programs for building a leg or an antenna, for example—much like computer programs carry out specific tasks.



Homeotic genes and evolutionary change

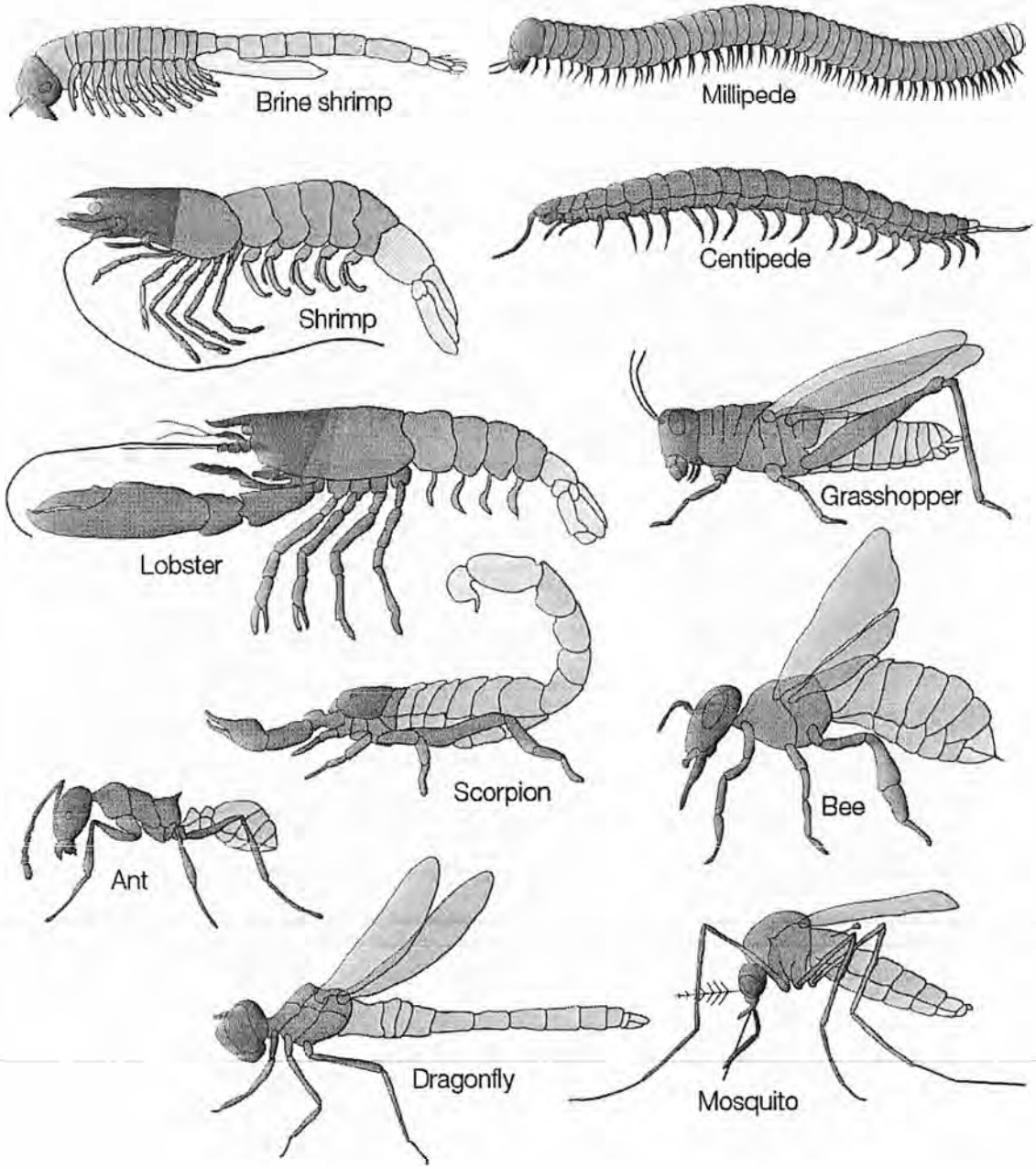
A large amount of animal diversity is built on two simple ideas: bodies made up of repeating units (or segments), and genetic programs for building structures.

Just within arthropods (shown on the right), variations on this theme have given rise to an enormous diversity of body types. And in fact, in many cases, domains of Hox gene activity parallel the different types of structures that grow out of the animals' body segments.

The bands of color highlight body segments that have similar identities; you can think of each color as running a different genetic program: a "leg" program or an "antenna" program, for example. Once a program exists for building a structure, it can be reused elsewhere simply by shifting Hox gene expression. It's easy to see how adding some segments and running the "leg" program in them can build an organism with a few more sets of legs.

And the genetic programs themselves can be modified (through changes in the "leg" or "antenna" genes) to build structures that are a little different. For instance, the "wing" program didn't come about from scratch—it's simply a modified "leg" program.

A genetic change that leads to a change in body shape might allow an organism to capture food more effectively or avoid predators, giving it a reproductive advantage. Its genes may be preferentially passed along to the next generation, thus influencing the course of evolution.


Brine shrimp
Millipede
Shrimp
Centipede
Grasshopper
Lobster
Scorpion
Bee
Ant
Dragonfly
Mosquito

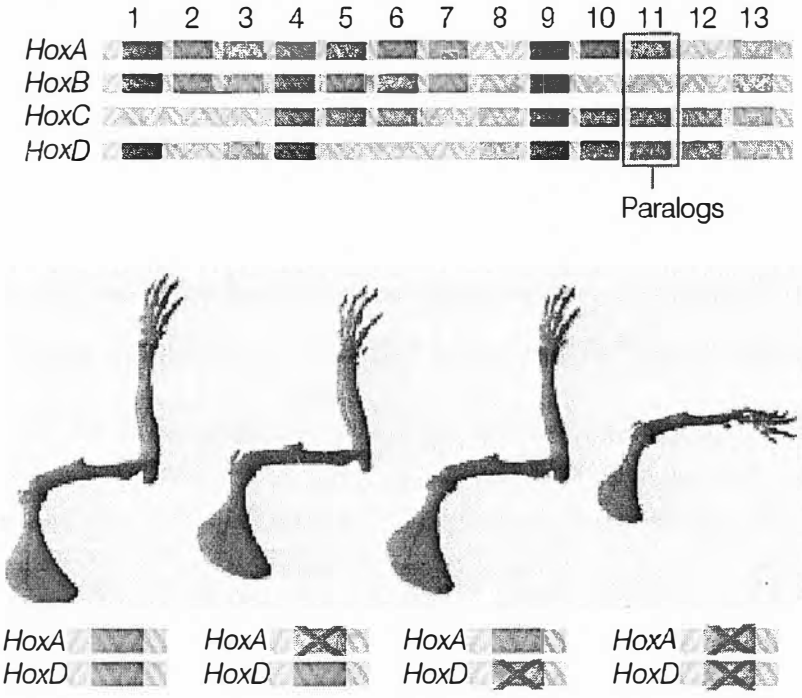
Vertebrate Hox genes

In vertebrates (animals that have backbones), the entire Hox cluster has been duplicated multiple times. Mice and other mammals have four Hox clusters. All four are similar, but each is different. Similar genes in different clusters are called paralogs.

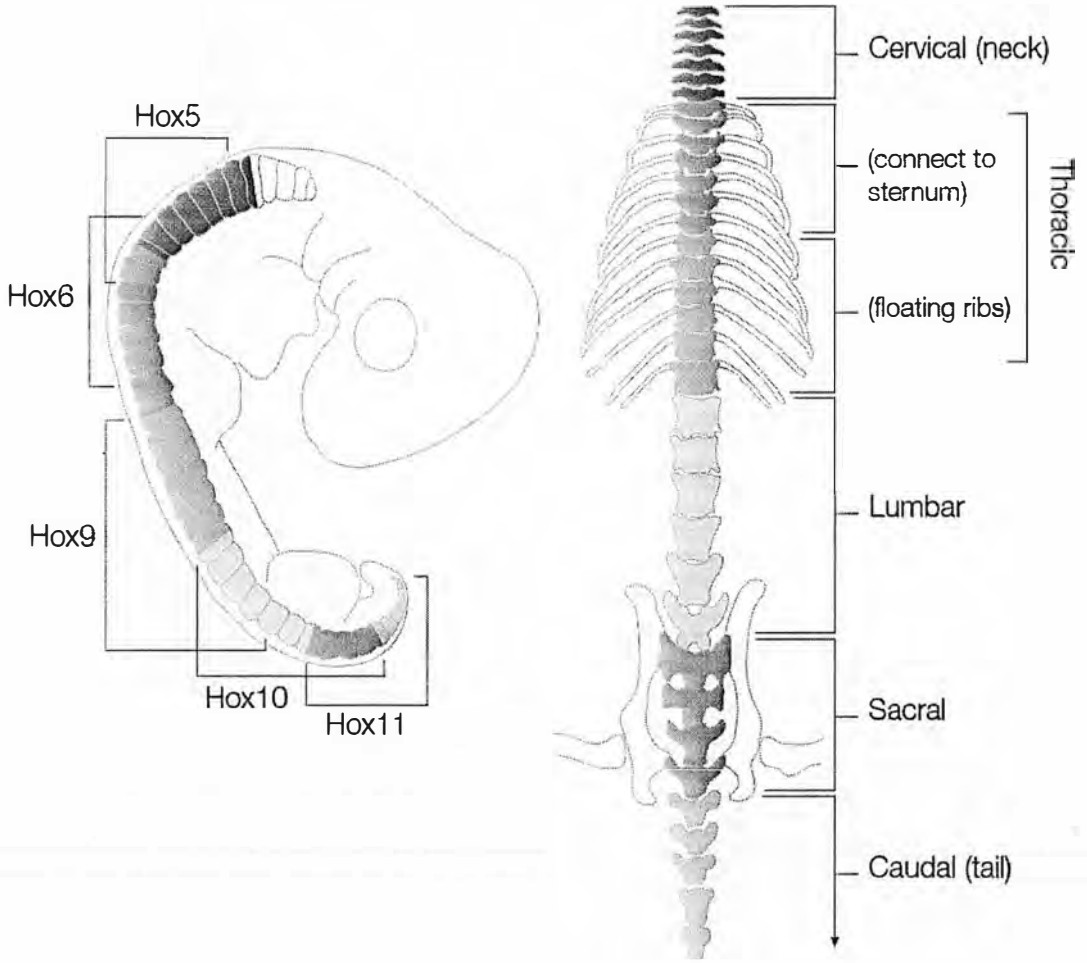
Most paralogs have partially overlapping functions, so figuring out how Hox genes function in vertebrates a challenge: the effects of changing a single gene are often hidden by functioning genes in the same paralogous group. But changing the function of multiple genes in the group can have dramatic effects.

The photos on the left, provided by Mario Capecchi's research group at the University of Utah, show mouse forelegs. Inactivating one paralog or the other has subtle effects (middle two images). But inactivating both makes a dramatically different limb (right). This experiment and others have shown that Hox genes in mice work much the same way as they do in fruit flies.

Mouse Hox genes, located on 4 different chromosomes



Hox genes direct the identity of vertebrae



While mice and other vertebrates are not as obviously segmented as arthropods, certain regions of their bodies actually are. Vertebrae, with all their associated muscles and bones, grow from repeating units in the embryo called somites. Hox genes (often in combination) help define somite identity, directing them to develop differently depending on where they are in the body.

Just like Hox genes in arthropods direct segments to grow legs, wings, or antennae, Hox genes in vertebrates direct segments to grow ribs (or not) or bones that fuse together to form a sacrum.

Experiments in mice show how Hox genes affect vertebra identity. In mouse embryos, the Hox10 genes turn the "rib" program off. The genes are normally active in the lower back, where the vertebrae don't grow ribs, and inactive in the mid-back, where the vertebrae do grow ribs. When the Hox10 paralogs are experimentally inactivated, the vertebrae of the lower back to grow ribs. Something similar may have happened in nature. In snakes, Hox10 genes have lost their rib-blocking ability, which may be why they grow ribs from head to tail.

Hox genes play many more roles in vertebrate development. They help specify the difference between an arm and a leg, as well as a pinky and a thumb. In the nervous system, their expression in segmented embryonic structures called rhombomeres directs the development of different brain regions.

Hox genes are a fascinating example of how a single gene that did something well was copied and re-purposed through evolution to do even more.

References

Carroll, S.B., Grenier, J.K. & Weatherbee, S.D. (2004). From DNA to diversity: molecular genetics and the evolution of animal design (2nd ed.). Malden, MA ; Wiley-Blackwell.

Davis, A.P., Witte, D.P., Hsieh-Li, H.M., Potter, S.S. & Capecchi, M.R. (1995). Absence of radius and olna in mice lacking hoxa-11 and hoxd-11. Nature, 375, 791-795. doi:101038/375791a0

Garcia-Fernandez, J. (2013). The genesis and evolution of homeobox clusters. Nature Reviews Genetics 6, 881-891. doi:10.1038/nrgl723

Heffer, A. & Pick, L. (2013). Conservation and variation in Hox genes: how insect models pioneered the evo-devo field. Annual Reviews in Entomology, 58, 161-179. doi:10.1146/annurev-ento-120811-153601

Lews, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature 276, 565-570.

Mallo, M., Wellik, D.M. & Deschamps, J. (2010). Hox genes and regional patterning of the vertebrate body plan. Developmental Biology, 344, 7-15. doi:10.1016/j.ydbio.2010.04.024

Pascual-Anaya, J., D'Aniello, S., Kuratani, S. & Garcia-Fernandez, J. (2013). Evolution of Hox gene clusters in deuterostomes. BMC Developmental Biology, 13(26). doi:10.1186/1471-213X-13-26

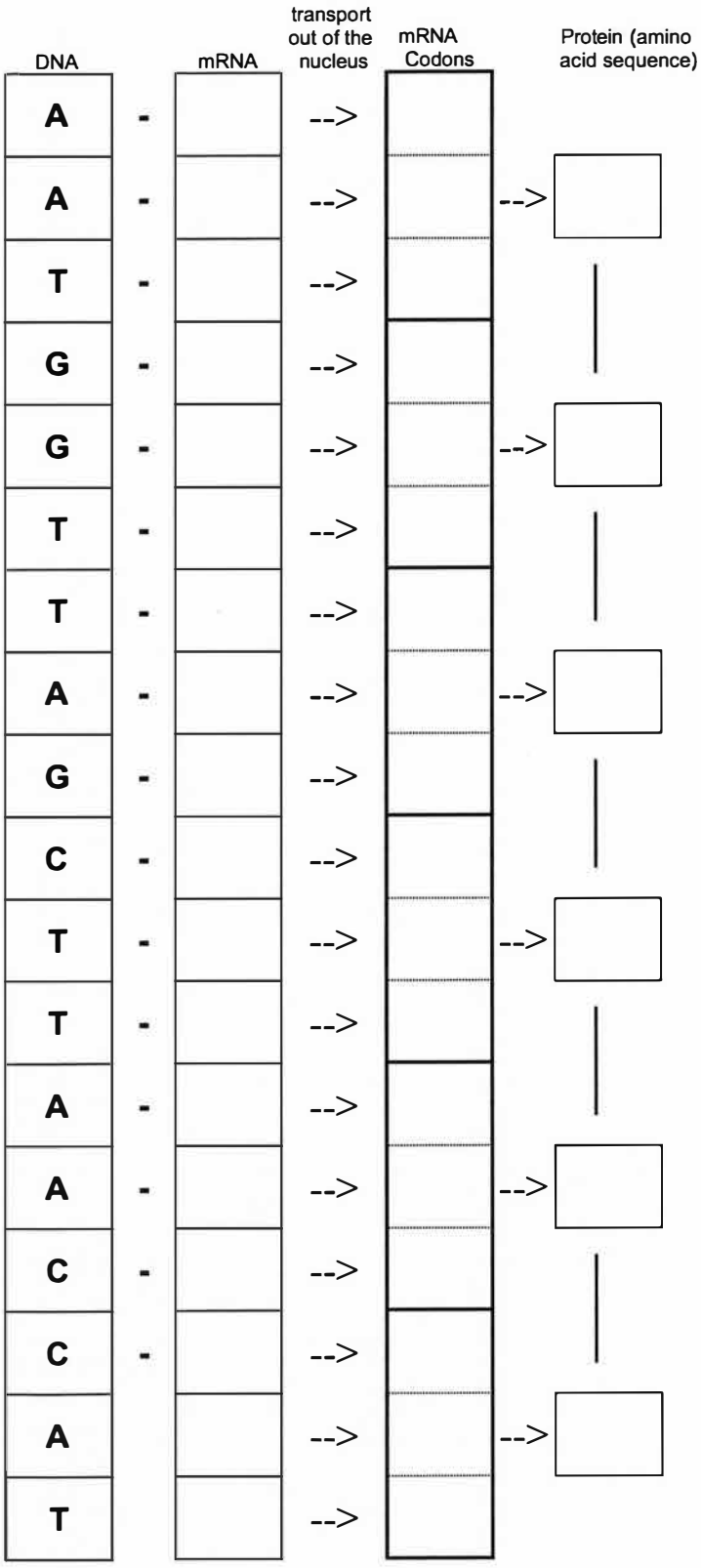
Pearson, J.C., Lemons, D. & McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. Nature Reviews Genetics, 6, 893-903. doi:10.1038/nrgl726

Wellik, D.M. (2007). Hox patterning of the vertebrate axial skeleton. Developmental Dynamics, 236, 2454-2463. doi:10.1002/dvdy.21286

Transcription and Translation

Name _____ Period ____ Date ~~273~~

1. Fill in the mRNA code that is complimentary to the DNA template
2. Transfer the mRNA sequence to the colunms marked "mRNA Codons"
3. Use the table in the book to find what amino acid is specified and write the abbreviation for that amino acid in the Protein column



Transcription and Translation

Name ____________ Period ____ Date ~~274~~

1. Fill in the mRNA code that is complimentary to the DNA template
2. Transfer the mRNA sequence to the colunms marked "mRNA Codons"
3. Use the table in the book to find what amino acid is specified
   and write the abbreviation for that amino acid in the Protein column

DNA		mRNA	transport out of the nucleus	mRNA Codons		Protein (amino acid sequence)
A	-		-->			
A	-		-->		-->	
T	-		-->			
G	-		-->			
C	-		-->		-->	
G	-		-->			
T	-		-->			
T	-		-->		-->	
A	-		-->			
G	-		-->			
C	-		-->		-->	
T	-		-->			
T	-		-->			
A	-		-->		-->	
A	-		-->			
C	-		-->			
C			-->		-->	
A			-->			

Name _____
Period _____ Date _____

Protein synthesis

Determine the amino acid sequence that would be specified by the following sequence.

1. Transcribe the DNA into messenger RNA.
2. Find the Start Codon to establish the reading frame.
3. Mark off the subsequent codons until the Stop Codon is found
4. Translate the mRNA into the proper amino acid sequence

DNA

T C C A A G T A C G G T C T A T A T G T A G T C T T T C C A C C C G C A A C T A C A G A A G A C

mRNA

Amino acid sequence

Anticodons found on tRNA

Name _____
Period _____ Date _____

Protein synthesis

Determine the amino acid sequence that would be specified by the following sequence.

1. Transcribe the DNA into messenger RNA.
2. Find the Start Codon to establish the reading frame.
3. Mark off the subsequent codons until the Stop Codon is found
4. Translate the mRNA into the proper amino acid sequence

DNA

T C C A A G T A C G G C T A T A T G T A G T C T A A T C C A C C C G C A A C T A C A G A A G A C

mRNA

Amino acid sequence

Anticodons found on tRNA

Name _____
Period _____ Date _____

Protein synthesis

Determine the amino acid sequence that would be specified by the following sequence.

1. Transcribe the DNA into messenger RNA.
2. Find the Start Codon to establish the reading frame.
3. Mark off the subsequent codons until the Stop Codon is found
4. Translate the mRNA into the proper amino acid sequence

DNA

T C C A A G T A C G T C T A T A T G T A G T C T T T C C A C A C C G C A A C T A C A G A A G A C

mRNA

Amino acid sequence

Anticodons found on tRNA

Name _____
Period _____ Date _____

Protein synthesis

Determine the amino acid sequence that would be specified by the following sequence.

1. Transcribe the DNA into messenger RNA.
2. Find the Start Codon to establish the reading frame.
3. Mark off the subsequent codons until the Stop Codon is found
4. Translate the mRNA into the proper amino acid sequence

DNA

T C C A A G T A C G T C A T A T G T A G T C T T T C C A C A C C T G C A A C T A C A G A A G A C

mRNA

Amino acid sequence

Anticodons found on tRNA

Name _____
Period _____ Date _____

Protein synthesis

Determine the amino acid sequence that would be specified by the following sequence.

1. Transcribe the DNA into messenger RNA.
2. Find the Start Codon to establish the reading frame.
3. Mark off the subsequent codons until the Stop Codon is found
4. Translate the mRNA into the proper amino acid sequence

DNA

T C C A A G T A C G G C T A T A T G T A G T C T A A T C C A C C C G C A A C T A C A G A A G A C

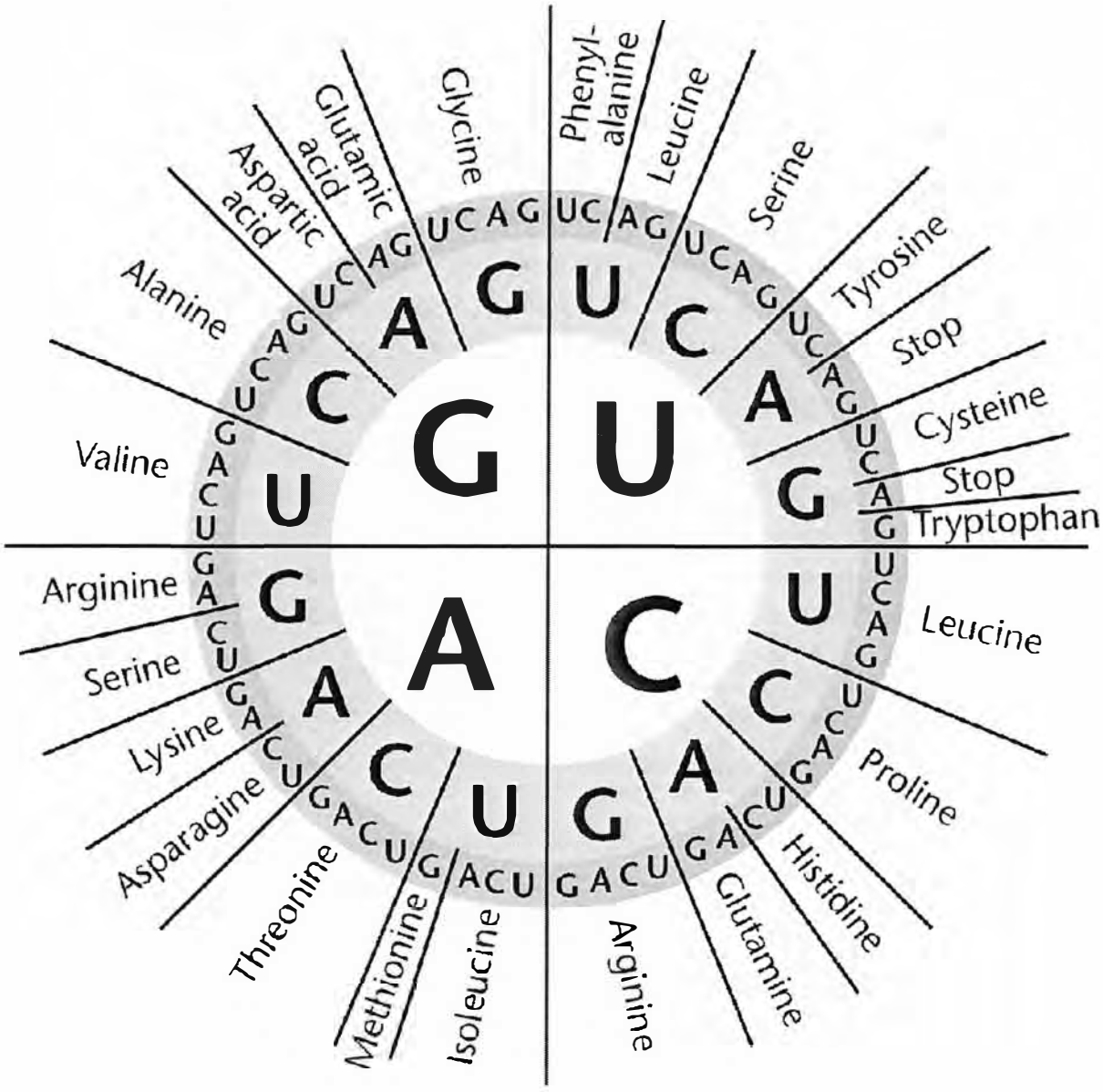
mRNA

Amino acid sequence

Anticodons found on tRNA

Codon Decoder Ring

1. The Codon is the set of 3 nucleotides in the mRNA sequence.
2. Start at the center in the quadrant that matches the first nucleotide in the Codon.
3. Move to the second nucleotide in the codon without crossing any of the solid lines.
4. Move to the third nucleotide in the outer grey ring without crossing any solid line to find the proper amino acid.
5. e.g. If the codon is AUG, start in the center of the lower left quadrant (A), move to the lowermost grey section (U), then move to the outer ring that had the G. The resulting Amino Acid is Methionine.



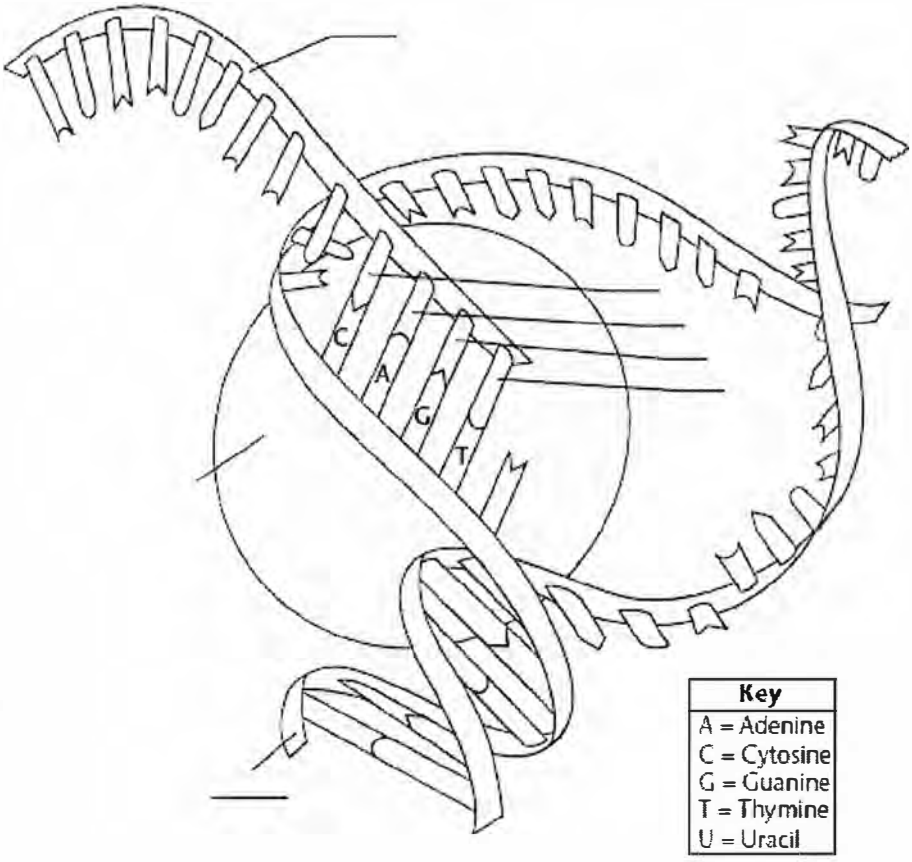
Name _____
Period _____ Date _____

Transcription and Translation

Use complete sentence(s) to answer the following

1. Define **Transcription** :

Use the following diagram to help describe the process of transcription.



Apply the following labels to appropriate parts of the above diagram. Label as many parts as you can

- Guanine
- Adenine
- Thymine
- Cytosine
- Uracil
- DNA
- mRNA
- RNA polymerase

Use complete sentences to answer the following

2. Where does the process of transcription take place?

3. Using the diagram, describe what needs to happen before transcription can take place?

4. How does the correct strand get copied? (extra credit)

5. After RNA polymerase attaches to DNA, how is mRNA made?

6. After mRNA has been made, where does it go?

7. What process occurs following the transport of the mRNA?

Name _____
Period _____ Date _____

Transcription and Translation

Use complete sentence(s) to answer the following

1. Define **Translation** :

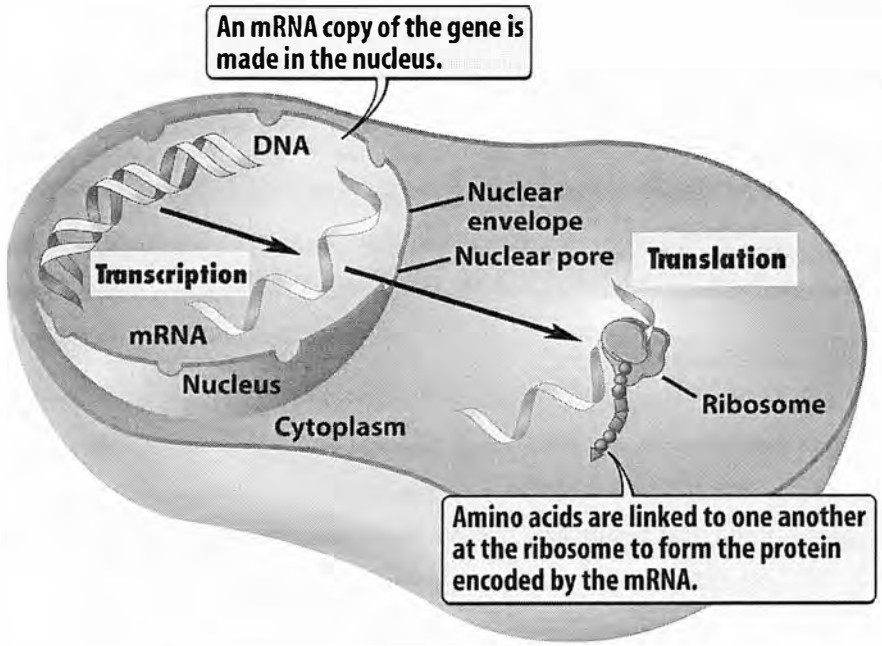


Figure 13-2 Discover Biology 3/e
© 2006 W. W. Norton & Company, Inc.

Use complete sentences to answer the following

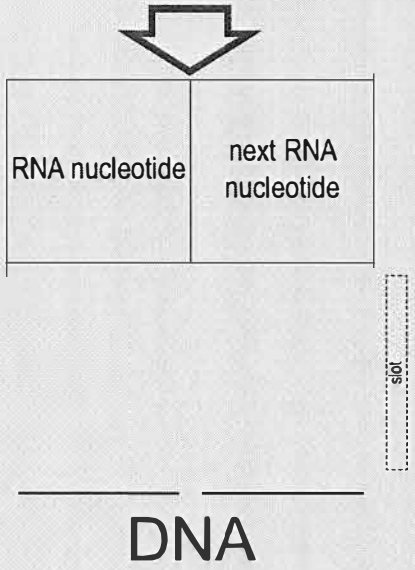
1. Where does the process of translation take place?

2. Using the diagram, summarize the processes of transcription and translation.

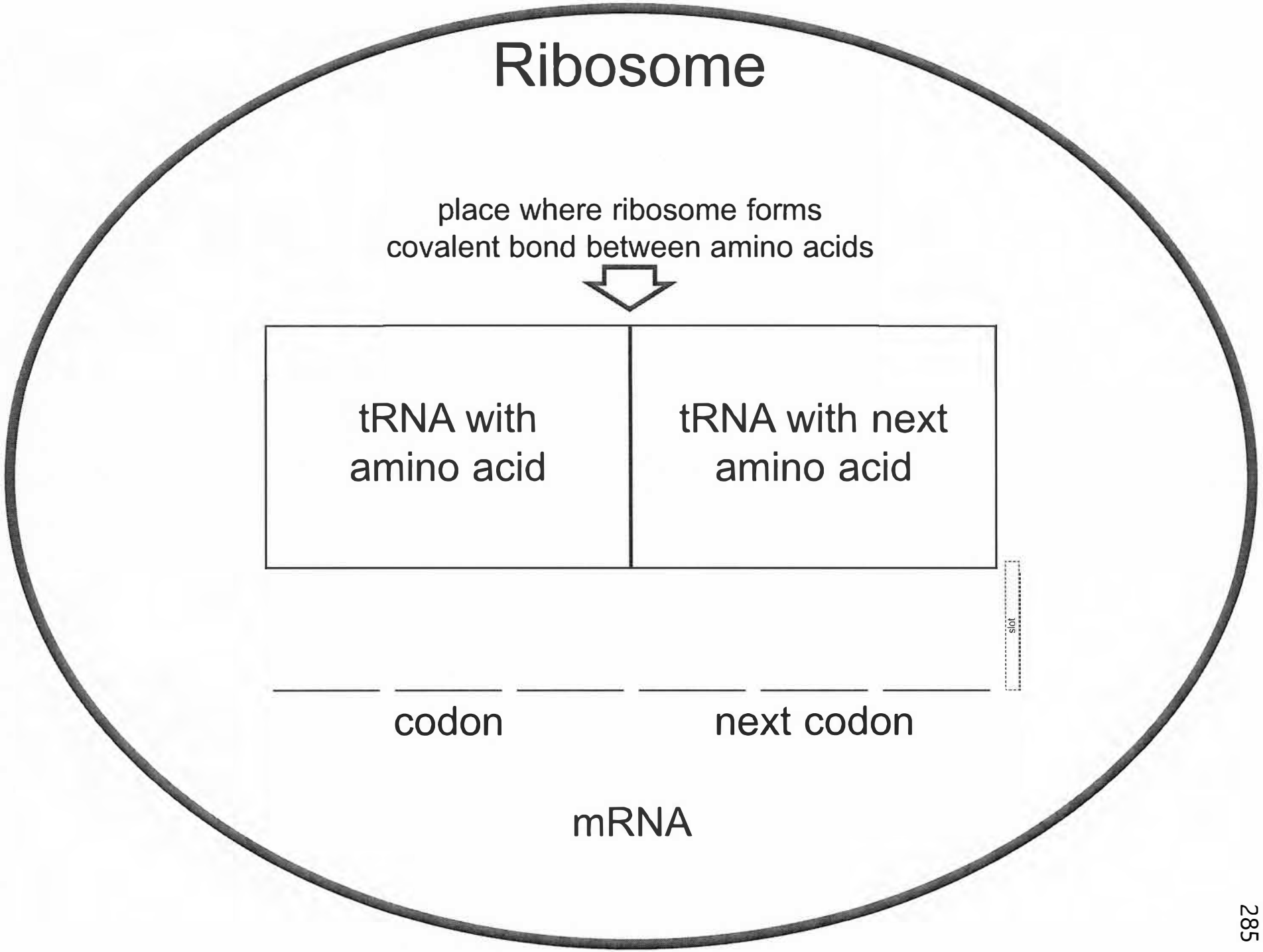
Nucleus

RNA polymerase

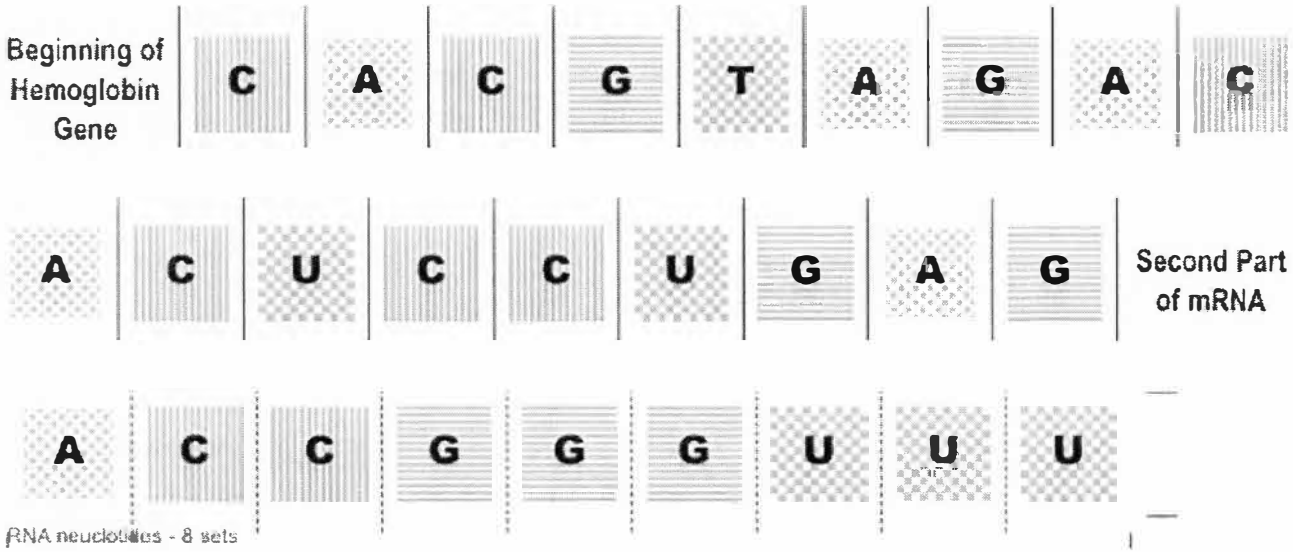
place where enzyme forms
covalent bond between nucleotides



cytoplasm
around the nucleus


Ribosome
place where ribosome forms
covalent bond between amino acids
tRNA with
amino acid
tRNA with next
amino acid
slot
codon
next codon
mRNA

Transcription and Translation Lab - Sample pieces



tRNA molecules - 2 sets

Leucine	Histidine	Valine
Leucine	Histidine	Valine
Leucine	Histidine	Valine
Leucine	Histidine	Valine
Threonine	Proline	Glutamic Acid
Threonine	Proline	Glutamic Acid
Threonine	Proline	Glutamic Acid
Threonine	Proline	Glutamic Acid

Name _____
Period _____ Date _____

Transcription and Translation Viewing Guide

One of the central dogmas of genetics is that the DNA sequence leads to the expression of certain characteristics in organisms. This is done by using the DNA sequence in the nucleus of eukaryotic cells to make an intermediary sequence called Messenger RNA (mRNA). The Messenger RNA sequence is transported out of the nucleus to the ribosomes where it is used as a template to produce an Amino Acid Sequence (Protein, Polypeptide). The Protein becomes a functional unit that produces the characteristic in the organism.

Pre-Video

1. What is DNA?

2. Describe the gross structure of DNA.

3. What is a Gene?

4. Define - Transcription

5. Define - Translation

Watch the Video without pausing

6. What is your impression of the processes involved in gene expression?

Watch the Video with pauses to explain what is happening

7. What happens at the beginning of the gene sequence to start the process of Transcription?

8. Where is the gene in relation the assembly of factors on the DNA?

9. What is the Blue molecule doing once the process starts?

10. What is the Yellow chain snaking out of the Blue molecule?

11. What is the Yellow chain made of?

12. What determines the sequence of the RNA?

13. What is the nucleotide in RNA that is different from DNA? What letter in DNA does it replace?

14. What is the process of making RNA from DNA called?

15. How fast does it occur? Where is this happening, even as we go over this process?

16. Where does the RNA go once it is complete?

17. What molecular machine locks around the RNA? What is this miniature factory called?

18. What is made in the miniature factory using the RNA as a pattern?

19. What brings the Amino Acids to the Ribosome?

20. How are the transfer molecules different?

21. How many letters at a time is the code on the mRNA read?

22. What are the letters in the code on the mRNA matched to?

23. How is the protein chain made?

24. How many different kinds of protein can be made by the ribosome?

25. Which particular protein is being made in this video?

Reflection(s)

26. What impression(s) doe this video leave you that you did not think of before?

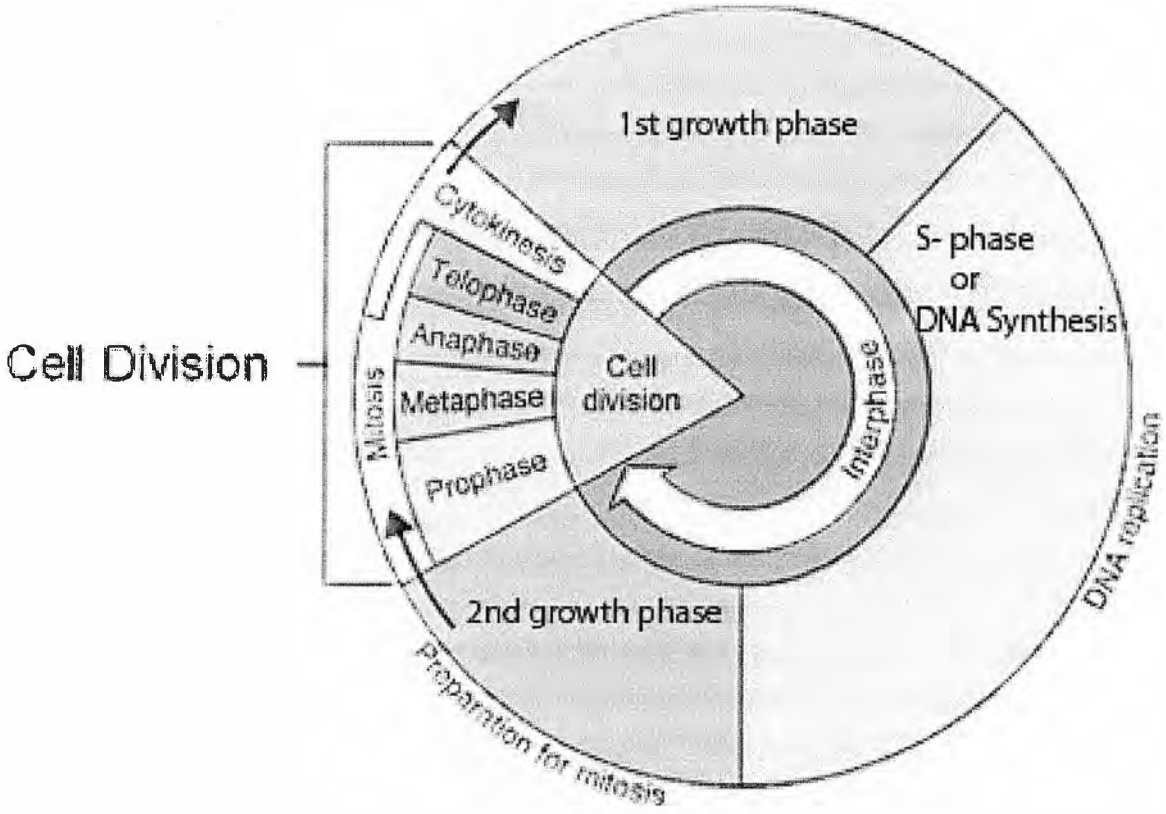
Summary

27. In your own words, write a summary of how protein is made starting with the DNA sequence.

Name _____
Period _____ **Date** _____

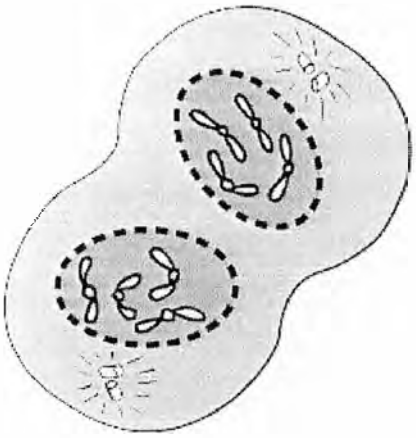
Cell Cycle

1. On the lines below, explain what happens in each of the phases of the cell cycle and mitosis starting with the 1st Growth Phase.

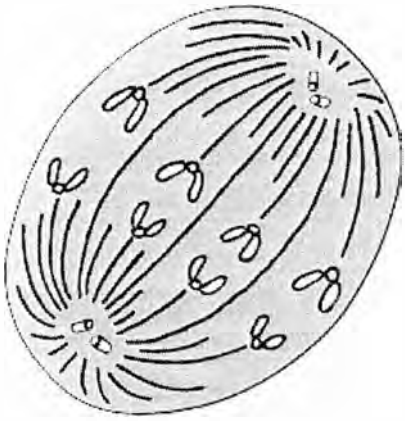


a.	
b.	
c.	
d.	
e.	
f.	
g.	
h.	

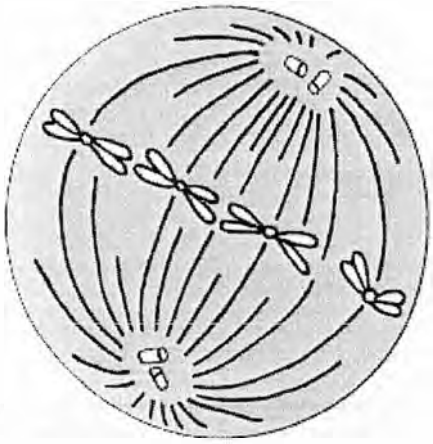
2. Identify the different phases of mitosis pictured below. Describe landmark features in each of the diagrams that helps correctly identify each phase.



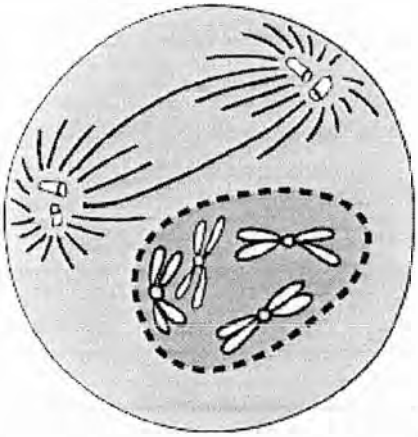
a. _____



b. _____



c. _____



d. _____

DNA Replication

Name _____ Period ____ Date 294

1. Fold this page so that columns 1 and 4 are next to each other
    a. Make a "mountain fold" along the dashes between columns 1 and 2
    b. Make a "valley fold" along the dashes between columns 3 and 4
2. Finish filling in the complimentary nucleotides for column 4
3. Unfold the paper (unzipping the original molecule)
4. Fill in the new complimentary strands in columns 2 and 3

1		2		3		4
A	-		-		-	T
A	-		-		-	T
T	-		-		-	A
G	-		-		-	
C	-		-		-	
G	-		-		-	
T	-		-		-	
T	-		-		-	
A	-		-		-	
G	-		-		-	
C	-		-		-	
T	-		-		-	
T	-		-		-	
A	-		-		-	
A	-		-		-	
C	-		-		-	

Scientific Process

Step in scientific process	Sample sentence frames	Language function
Hypothesis/prediction – If….., then	If _____, then _____	*Cause and effect*
Procedure – Steps: 1. 2. 3.		*Sequencing*
Data/Observations – 1st observation – 2nd observation – Observation on chart	First thing I observed _____ The second, observation is _____	*Sequencing* *Descriptive*
Analysis of Data -	*For example –* *Based on the data observed* _____ *The graphs shows* _____ *The percent error is due to* _____	*Descriptive*
Conclusion - Was your hypothesis correct? What 3 things did you learn? –	My hypothesis was _____, because _____ First, _____ Second thing _____ Finally,	*Completing cause and effect* *Sequencing*

Expository Writing	_____ have _____ known as _____
Compare and Contrast	_____ and _____ are similar/different _____, because _____
Elaboration and Description	A ____of _____ is called _____ One example of _____ is _____ Examples of _____ include _____ and _____ _____ are composed of _____ and _____ _____ can form _____ _____ is an example of a _____

5 sentence Summary writing	
Identify the topic	The _____ titled _____ examines _____ addresses introduces
Idea 1	_____ uses/relies on _____ to explain _____ _____ has/have _____ called _____
Idea 2	Another idea from _____ use/explain _____ Other _____ called _____ have/has _____
example	An example of _____ would be/is _____ _____ and _____ are similar/different _____
Conclusion	_____ enable _____ to _____, because _____ Information/examples _____ support _____, because _____

Proposition and Support - Claim-Evidence-Reasoning (CER)	
Claim – make a proposition	My claim for _____ is _____ My opinion on _____ is _____ I think _____ I think the evidence/data shows _____
Evidence – support your proposition with specific evidence	According to my data _____, _____, _____ _____, _____, _____ supports my claim. [combined evidence and reasoning] _____, _____, _____ supports my claim, because_____
Reasoning – explain how the evidence supports the claim	My data supports my claim because _____ Based on the data, I see _____, because _____

History of Life on Earth

Complexity of Life - Assessment

Review the products that you have produced in this unit. Write an essay explaining to Robert Krulwich why, in spite of the fact that he feels quite different from a banana, half of his DNA is the same as that of a banana.

Rubric

Points	Description
4 points	Student incorporates all of the Main Concepts effectively
3 points	Student uses 5 out of the 7 Main Concepts effectively
2 points	Student uses approximately half of the Main Concepts effectively
1 point	Student uses a few of the Main Concepts effectively
0 points	Student provides no answer or responds with writing that has no bearing on the question