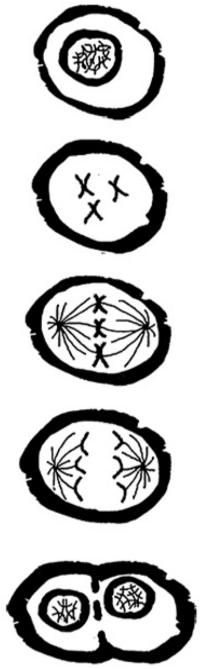


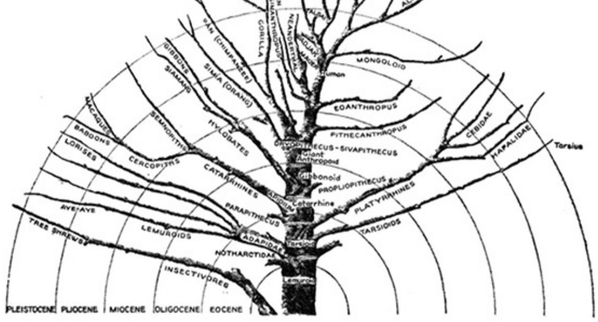
A
CRASH COURSE
IN
DNA

EACH GENE IN OUR DNA CONTAINS THE INSTRUCTIONS TO MAKE ONE TINY PIECE OF WHO WE ARE.

GENETICS 101

FROM CHROMOSOMES AND THE DOUBLE HELIX TO CLONING AND DNA TESTS, EVERYTHING YOU NEED TO KNOW ABOUT GENES

IN THE 1800s, GREGOR MENDEL FIGURED OUT THAT TRAITS OF PEA PLANTS WERE PASSED DOWN FROM PARENT TO CHILD IN A WAY THAT COULD SOMETIMES BE HIDDEN SO THEY APPEARED TO SKIP A GENERATION.



BECAUSE NUCLEOTIDES HAVE A CURVY SHAPE, THEY DON'T STACK UP STRAIGHT LIKE A LADDER BUT INSTEAD TWIST LIKE A SPIRAL STAIRCASE. THE RESULT IS A DOUBLE HELIX.

BETH SKWARECKI

GENETICS 101

FROM CHROMOSOMES AND THE DOUBLE HELIX TO CLONING AND DNA TESTS, EVERYTHING YOU NEED TO KNOW ABOUT GENES

BETH SKWARECKI

Adams Media
New York London Toronto Sydney New Delhi

CONTENTS

INTRODUCTION

Less than two centuries ago, all people really knew about genetics was that children tend to look like their parents and that careful breeding of dogs or horses or crops can result in bigger and better dogs or horses or crops. We've learned a lot since then.

In the 1800s, a monk named Gregor Mendel figured out that traits of pea plants—like whether peas were yellow or green—were passed down from parent to child in a way that could sometimes hide traits so they appeared to skip a generation. He figured out how to predict whether and when a hidden trait would show up next.

Around the same time, naturalist Charles Darwin figured out that species evolve over time. The traits of pets and crops are influenced by a farmer who breeds them, but according to Darwin's theory of evolution by natural selection, it is nature, rather than human judgment, that determines which creatures live long enough to have offspring. Darwin knew the whole idea hinged on some mysterious way that parents can pass down traits to their children, but he had no idea how that might work.

And then, in the 1950s, Rosalind Franklin managed to form DNA into a crystal and take an x-ray photograph that revealed its structure. James Watson and Francis Crick built on her work to deduce that the DNA molecule had the shape of a double helix and that DNA's structure was uniquely suited to pass down traits from one generation to the next. Over the remaining decades, scientists have worked out the details of exactly how DNA makes us who we are—and how we can tinker with it.

This book will explain genetics, which is the study of how living things give their offspring the instructions, or genes, for particular traits. We'll also talk about genomics, which is a related field that studies the totality of all the information contained in your DNA. Along the way, we'll cover other bits of biology as needed. We'll do all this with a focus on you and what's going on in your body, plus a few things you might see in the news.

Along the way, we'll take some detours to visit the genomes of animals, plants, bacteria, and even viruses. You have more in common with all of these creatures than you probably realize.

First, we'll learn about the nuts and bolts of deoxyribonucleic acid—DNA—itself. It's a stringy substance that, on a smaller-than-microscopic scale, is an incredibly long molecule. You have forty-six of these strands stuffed into the nucleus of each cell in your body, and each strand contains instructions for building and maintaining every part of your body. These instructions are in a chemical language that we’ll learn to decode.

We’ll see how your cells read that code and carry out the instructions. Often the instructions tell the cells to build a protein, so we’ll learn what these proteins do too. Some of them give your eyes and hair and skin their distinctive colors. Some help your body to process food and drugs. Some are so important to the way your body functions that if they aren’t built in exactly the right way, you could end up with an increased risk for cancer or other health conditions.

We’ll also learn about how your DNA got to you in the first place: how it was passed down from your parents and what it can tell you about your family tree. And we’ll see what you can learn from personal genomics services that promise to reveal your deepest secrets based on a sample of saliva. Finally, we’ll take a look at what scientists and companies are doing with DNA, from genetically modifying crops to improving treatments for cancer.

YOUR CELLS’ INSTRUCTION MANUAL

What DNA Does

DNA is what makes us who we are. But how does it do that?

To answer that question, we have to zoom in to a level even smaller than what microscopes can see. DNA is a long, stringy molecule whose job is to carry information. To understand that, think for a minute about this book. It’s just letters, one after another, that taken together form words and sentences and chapters. A DNA strand is made up of millions of chemical components that function like letters, spelling out an instruction manual with all the information it takes to build and run a human body. (Or an animal’s body or even a plant or a bacterium. Every living thing has DNA.)

Noncoding DNA

There’s more to our DNA than just recipes, though. Think of the genome as a deluxe cookbook with a ton of extra information, like how to plan a dinner party or suggested menus for a week’s dinners. That information is helpful so you know when to make the recipes. But it’s also a sloppy cookbook: there might be three versions of the same dish, and only one of them is worth making. Perhaps there are even some recipe cards and scraps of paper tucked into the pages, things that you’re not quite sure where they came from but you’re not sure if it’s okay to throw them away. Our DNA has scraps like that too.

There is far more information in DNA than in a book, though. If you printed it out, our **genome**—all the information carried in our DNA—would fill twelve thousand books this size.

Our genome isn’t just one string of DNA; it’s actually split into pieces called **chromosomes**. I like to think of our twenty-three chromosomes as a recipe collection in twenty-three enormous volumes. Like a real cookbook, DNA contains short sets of instructions—think of them as recipes. Each recipe, or **gene**, contains the instructions to make one tiny piece of who we are.

Since we are all different, our recipes vary slightly. My genes include instructions on how to make a brown pigment and put it into my hair follicles. But your hair may be colored differently from mine if your genes encode a recipe for a different pigment. Or perhaps that page in your recipe book is blank, and you don't put any pigment in your hair at all.

You have two copies of this cookbook encyclopedia in each cell of your body. Cells are, in a sense, the kitchens where the recipes are made.

WHAT IS A CELL?

Your body contains over thirty-seven trillion cells. That's a huge number, right? You have more cells in your body than there are dollars in the national debt or stars in the Milky Way.

You have skin cells, muscle cells, fat cells, nerve cells, and bone cells, just to name a few. They're all so small you can only see them with a microscope. Every time you scratch an itch, hundreds of skin cells flake off, and you don't even notice.

Nearly every one of those cells contains all of the DNA we just talked about. Double that, actually, since you keep two copies around—the one you got from your mom and the one you got from your dad.

Mitochondria Have DNA Too

Most of our organelles are pretty boring, but there's a special one called the mitochondrion (plural: mitochondria) that helps turn food into energy. It's so special that it has its own DNA that it doesn't share with us. Scientists think this is because mitochondria used to be free-roaming bacteria that one day got eaten—but not digested—by a larger cell. After millions of years, we're like best friends: inseparable.

Our cells have different compartments, or **organelles**, separated from each other by membranes. We keep those two full sets of our DNA in their own organelle called the **nucleus**. This way they're safe from all the chaos going on in the rest of the cell. (Think of it like a special library for our cookbook collection.)

CHOOSING RECIPES

If all of our cells have the same cookbook collection, why aren't they all following the same instructions all the time? If they did that, all thirty-seven trillion of our cells would look alike.

What actually happens is that skin cells only use the genes that are necessary to do skin cell things. Muscle cells only use the genes that help them do muscle cell things. (Skin cells and muscle cells have plenty of things in common, of course, so some recipes are used by both.)

Even in a single cell type, things change all the time. Brain cells use different genes during the day than at night, for example. Your stomach cells use different genes when you're digesting food than they do during those long stretches between meals. And you'll use a different mix of genes as an adult than when you were a baby.

ATOMS AND MOLECULES

The Building Blocks of DNA

DNA is a huge molecule, made of millions of atoms.

The best way to understand the difference between atoms and molecules is to sit down with a molecular model kit. You can sometimes find these in chemistry classrooms or college bookstores, which makes them seem very serious, but in reality, a molecular model kit is just a very fun toy.

Try Building Alcohol

If water is too boring, you can make ethanol, which is the kind of alcohol that's in beer and wine. Start with a carbon and add three hydrogens. On the fourth toothpick, stick another carbon, and give that carbon two hydrogens. The second carbon should now have three toothpicks in it, so for the fourth toothpick you'll add the hydroxyl group, which is just an oxygen atom that has a hydrogen attached. That -OH group is what makes it an alcohol.

If you don't have one, that's fine! You can play along at home with a bag of gumdrops and a box of toothpicks.

Let's start with a simple molecule: water. You probably know water's chemical formula already: H_2O. That means it has two hydrogen atoms and one oxygen atom. If you're doing the candy version of this exercise, grab a red gumdrop and stick two toothpicks into it. That red gumdrop is your oxygen atom. Take two white gumdrops to represent the two atoms of hydrogen and stick them at the other end of each toothpick. You've just made H_2O.

If you're lucky enough to have a model kit, the hydrogen atoms will be built with just one socket where a connector can fit in. The oxygen atoms will have two sockets. That's because in real life, oxygen can (normally) only make two bonds. Hydrogen makes just one.

The Atoms We'll Be Working With

Most of the molecules in our cells can be made with just six atoms. Think of them as your organic chemistry starter kit:

• Carbon (4 bonds)
• Hydrogen (1 bond)
• Oxygen (2 bonds)
• Nitrogen (3 bonds)
• Phosphorus (it's complicated)
• Sulfur (likewise)

Of these, you only need the first five to build DNA.

Carbon, on the other hand, can make four bonds, so the little black spheres that represent carbon will have four sockets in them. That's the advantage of the model kit: each piece has an appropriate number of sockets. If you're using gumdrops, you have to remember, on your own, how many toothpicks to put into each atom.

WHAT ARE ATOMS?

Have you ever heard of the periodic table of elements? It's that weird-shaped chart with one square for each known element. Some are things you've heard of: carbon, hydrogen, and oxygen, for example. Others are metals, and you've heard of a lot of these too: gold, silver, aluminum, copper. Neon, the gas that fills the tubes in light-up neon signs, is also an element.

These elements are really just the flavors that atoms can come in. What determines the flavor of an atom? It's the number of protons the atom has. Hydrogen has one proton, helium has two, and so on. If you're wondering about some of the elements we've already met, carbon has six protons, nitrogen has seven, and oxygen has eight. Gold has seventy-nine, and uranium has ninety-two.

Protons have a positive charge, so the more protons an atom has, the more negatively charged electrons it can collect.

You don't need to understand protons and electrons (or their neutrally charged buddies, neutrons) to be able to understand this book, so if this sounds like too much chemistry all at once, don't sweat it. We just mention them because the electrons are what determine the number of bonds an atom can make.

GIANT MOLECULES

As you tinker with your gumdrops and toothpicks, you might get carried away and decide to make the biggest molecules you can. And if you have a mega-sized bag of gumdrops, you'll find that molecules can be enormous!

For example, say you look up how to make a molecule of glucose—it forms a ring, so it looks kind of like a spiky crown. Make a bunch of these, and you can start chaining them together to build starch, the carbohydrate that provides most of the calories in foods like bread, pasta, and rice. Molecules like starch that are made from repeats of smaller building blocks are called **polymers**.

DNA is another polymer, but it's a bit more complicated than starch. Instead of one building block that repeats over and over, DNA has four different types of building blocks. Its pieces also come together in a way that makes a unique structure called a **double helix**. We'll learn more in the next sections about how this molecule is put together.

BUILDING BLOCKS

We've done a lot of building today, and we're about to do some more. Here's your cheat sheet for what builds what:

- Atoms are the smallest possible piece of an element. They are the building blocks of molecules.
- Molecules are the smallest possible piece of a compound, such as water or DNA. (Imagine a glass of water; the smallest item in the glass would be a single H_2O molecule.)
- Glucose, a sugar, is the building block of starch.
- Amino acids are the building blocks of proteins.
- Nucleotides are the building blocks of DNA.

NUCLEOTIDES

An Alphabet of Four Letters

Now that you know how atoms can join together to form molecules, we're going to learn how they make DNA. This isn't just so you can build a giant gumdrop-and-toothpick DNA molecule for your next party (although that's a great idea, isn't it?). The **structure** of DNA is crucial to understanding how DNA can carry instructions for the cell. Its structure also determines how parents can copy its information to pass down to their children. So, if we want to understand genetics, we have to zoom in to the molecular level.

Molecular Structure

If you haven't studied chemistry, you probably think of chemicals as liquids. But zoom in to the individual molecules, and each one has a three-dimensional shape. Molecules can bump into each other. They can wrap around each other. They take up space. We call the three-dimensional shape of a molecule its **structure**.

Since a single strand of DNA is a polymer—made of repeating units—we can't build it until we have the right units to start with. Those repeating units are called **nucleotides**.

There are a few different kinds of nucleotides, but they all have three components:

- A **nitrogenous base**, which can come in one of four different versions, nicknamed A, T, G, and C. We'll learn more about these in a minute.
- A **sugar**, specifically a kind called deoxyribose. This is different from the table sugar you put in your coffee, but in a sense, it's in the same chemical family. Deoxyribose is in the shape of a five-sided ring, and it's connected to the nitrogenous base.

- A **phosphate**, which is a phosphorus atom surrounded by oxygen atoms. This is also attached to the sugar.

To attach one nucleotide to the next, you attach the new nucleotide's phosphate group to the sugar on the previous nucleotide.

In real life, in your cells, there's no pair of hands stringing gumdrops together. Instead, when it's time to make more DNA, special proteins pick up nucleotides that are floating around the cell. They stick them onto the bottom of the growing DNA chain.

ATP

The structure of a nucleotide—base, sugar, phosphate—is something you might recognize if you've studied biology before. ATP, or adenosine triphosphate, has the same structure. ATP is best known as a source of energy for the cell, thanks to the high-energy bonds between its three phosphate groups. The nucleotides that can be incorporated into DNA are built the same way: when they're floating around the cell, they usually have three phosphate groups. It's like carrying their own little batteries to provide the energy needed to attach themselves to the growing DNA strand.

(From here on out, we're going to deal with gumdrops by the millions, so it may be best to stick with imaginary gumdrops rather than the real thing.)

THE FOUR NITROGENOUS BASES

Just like we spell words with the twenty-six letters of our alphabet, the information in DNA is "spelled" in an alphabet of four chemicals. With the right lab equipment, you can read down the strand of DNA and see what it says. Perhaps something like:

. . .

ATCGTCTGACTGACGACTGATCGTAGTCGATCGATGCGTACGAT(

Each of those four letters—A, T, G, or C—represents a different kind of nucleotide. The differences are in the part of the nucleotide called the nitrogenous base. The bases' full names are:

- Adenine
- Thymine
- Guanine

- Cytosine

They are called nitrogenous bases because they are basic—the opposite of acidic—and they contain a lot of nitrogen atoms in their structure. Thymine and cytosine are in the shape of flat rings, while adenine and guanine have a structure with two rings, like a figure eight.

THE DOUBLE HELIX

A Spiral Staircase

By this point in the history of the world, probably everyone reading this book has seen a picture meant to represent DNA. Any company whose work has to do with genetics or biotech has a doodle of DNA in their logo, for example. Any video that mentions genes or cells will include animations of a lumpy, oddly lit DNA strand. But these depictions aren't always accurate, so let's take a look at how DNA is really built.

DNA's True Shape

Real DNA has a wonky, asymmetrical look to it. Imagine that you have a ladder—that's your DNA—but before twisting it into a helix, you also pull the two uprights toward each other, bending the rungs. In real DNA, one side of the ladder (where the uprights are closer together) is called the minor groove, and the other side is the major groove. When you twist it, the spiral won't look neat and symmetrical. DNA, like life, is kind of messy.

BASE PAIRING

Those nitrogenous bases we discussed in the last section are responsible for a very important feature of DNA. They like to stick to other nitrogenous bases, and each has a specific partner.

Take adenine, for example. It's shaped like two flat rings, and as part of DNA it's attached to the sugar-phosphate backbone. But on the side that isn't connected to the backbone, it has some atoms that can make **hydrogen bonds**.

Remember when we were playing with gumdrops and toothpicks? The toothpicks represented **covalent bonds**, which are chemical bonds that keep two atoms pretty tightly connected. Hydrogen bonds aren't like that. They're a less permanent sort of link between two atoms or molecules, resulting from the atoms' electrical charge. It may help to think of a hydrogen bond as being like the static electricity that can keep a sock from the dryer stuck to your pant leg. You can pull it off any time you like, since it's not permanently connected to your pants.

Adenine and thymine are shaped in such a way that they can make hydrogen bonds with each other. (Adenine has a hydrogen that wants to stick to thymine's oxygen, and thymine has a hydrogen that wants to stick to adenine's nitrogen.) The other two bases, cytosine and guanine, also make a pair, and they actually manage three hydrogen bonds rather than two.

Best Buddies

Here's how the base pairings work:

- Adenine pairs with thymine.
- Guanine pairs with cytosine.

These pairings are specific: if you bring two adenines next to each other, or an adenine and a cytosine, they won't hydrogen-bond very well. But bring an adenine near a thymine, or a cytosine near a guanine, and they'll stick together like socks out of the dryer.

TWO STRANDS

So far, we've described one strand of DNA: a sugar-phosphate backbone with nitrogenous bases just dangling free. But in our cells, DNA doesn't hang around all by itself. It's typically paired with a **complementary strand**.

That's complementary in the sense that the second strand is a perfect counterpart, or complement, to the first one. If one strand is full of adenines, its partner will be full of thymines. In reality, DNA strands have complex, or sometimes random, sequences of nitrogenous bases. The second strand is a reflection of that. If one strand has CTAGGC, for example, the other has to be GATCCG.

Watch Your Spacing

Remember how some of the bases have a single ring shape, and others have a double ring? It just so happens that each matching pair includes a single and a double. That means the backbones of DNA are always the same distance (three rings) apart.

HOW DNA MAKES ITS MATCH

The cell doesn’t create two separate strands and then join them together. Remember how the nitrogenous bases want to stick to other nitrogenous bases? It would be like what happens when a toddler finds a roll of masking tape: you’d get a huge strand stuck to itself in so many places you’d have to throw it out and start over.

Instead, our cells have a very clever way of creating the second strand of DNA. First, let’s think about when and where this happens. When one cell needs to divide into two—say you’re growing more skin cells to help heal a wound—each cell needs its own copy of the DNA.

The cell’s machinery pries open the two strands, so there is a short section of DNA where the bases have nothing to stick to. Fortunately, the liquid inside this part of the cell contains a soup of nucleotides. These “free” nucleotides find their matches on the newly exposed parts of DNA. Special proteins called enzymes link the nucleotides to each other, creating two new DNA double strands. We’ll learn more about this process in the section titled “DNA Replication.”

THE END RESULT

The bases pair with each other, sticking the two strands of DNA together. And because the nucleotides have a somewhat curvy shape, they don’t stack up straight like a ladder; instead they twist, like a spiral staircase.

The resulting shape is called a **double helix**: “double” because it’s made of two strands, and “helix” from the Greek word for a spiral.

There’s one more quirk that you should know: one of the two strands is upside down.

Technically there’s no “up” and “down” in a cell, but the strands do have a direction to them. We usually draw a strand of DNA so the phosphate is on top, with the sugar below. Biochemists like to number the carbon atoms in the sugar, from one to five. The phosphate is attached to the fifth carbon atom, so we call the phosphate end of the DNA strand the **5' end**, pronounced “five prime end.” The other end of the strand, named after the sugar’s other attachment point, is the **3' end**, or “three prime end.” Whenever we’re adding to a DNA strand, the new nucleotides are added to the 3' end.

HOW ALL THAT DNA FITS INTO CELLS

How Big Is DNA, Anyway?

DNA is big and small at the same time. It's a molecule, so it's very, very tiny. You can't see it with a regular microscope, no matter how powerful the lenses. On the other hand, strands of DNA are very, very long.

Measurements of DNA

Here are the important measurements of a strand of DNA:

- **2 nanometers:** the distance from one backbone to another (the width of a "rung" of the ladder).
- **3.6 nanometers:** the length of DNA that it takes to make one full turn of the spiral.
- **0.34 nanometers:** the vertical space between base pairs (between rungs of the ladder). That means a full turn of the spiral includes about ten rungs.

We've already seen that the information in one copy of human DNA is enough to print twelve thousand books the size of this one. Now let's consider what that means in terms of base pairs. We have three billion base pairs in our DNA. If that were all in one long molecule (which it's not; it's divided into twenty-three), we would have a nearly invisible rope about a meter long, or just over three feet.

We actually have two copies of all this information. And when we say that, we're not referring to the two strands of the double helix! Those two strands match each other perfectly, so they don't really count as separate copies of the information. Instead, for each double helix you've got, you also have another complete double helix with almost the same information. You get one from each of your biological parents, so they will differ slightly. If that's confusing, don't worry—we'll learn more in a future section about how parents pass DNA to their children.