2 Genes Are the Instructions for Life: AIDS and the Uncommon Man

Just one small change in one gene might have given the world more books like *Pebble in the Sky*, *The Stars Like Dust*, *The Foundation Trilogy*, and *I, Robot*. To science fiction enthusiasts of a certain age, the publication of a new Isaac Asimov novel or short story was cause for celebration. Before iPods and instant messaging, before YouTube and Facebook, before Xboxes and PlayStations, young fans would curl up under the bedcovers with one of Asimov's intergalactic tales and read late into the night. For a period in the 1960s and 1970s, Asimov's large black glasses and mutton-chop sideburns made him one of the world's most recognizable authors. His books sold in the millions. Millions more might have flown off booksellers' shelves had Asimov inherited a personal DNA code with one small change.

Asimov penned more than just science fiction; he wrote on almost any topic. He explained mathematics and astronomy, chemistry and biology, as well as the Bible and Shakespeare, American history and the Roman empire, along with Gilbert and Sullivan and *Paradise Lost*, and limericks, and Egyptian history. Asimov wrote almost nonstop, averaging about a thousand words a day, every day, for fifty years. "Being a prolific writer has its disadvantages, of course," Asimov commented in one of his three autobiographies. "It complicates the writer's social and family life, for a prolific writer has to be self-absorbed . . . and has not time for anything else." This obsession for writing—and the accompanying unwillingness to do much else—had unwelcome consequences, including the breakup of his first marriage. And "a prolific writer . . . has to *love* his own writing," Asimov noted. He certainly loved writing: he wrote over four hundred books!

Asimov's many popular nonfiction works include *The Chemicals of Life* (1954), and *The Genetic Code* (1963). In the latter work he attempted to

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022

explain to a lay audience what a gene was at a time when even the research biologists working on genetics barely understood what it was.

What is a gene? Surely those research biologists must have figured it out in the forty-seven years that have intervened since Asimov weighed in on the subject. And the answer should be common knowledge to a society accustomed to headlines proclaiming "Scientists identify gene for schizophrenia," or for obesity, or for colon cancer, or for any number of other diseases and conditions.

Apparently not. None of the people we queried who weren't biologists came close to providing a definition of the gene that would pass muster in a high school biology classroom. Many associated the gene with physical traits or emotional characteristics, a reflection of the realization that genes come from our parents and grandparents, and the notion that they explain Johnny's big ears and Mary's quick temper. The gene was occasionally linked to concepts such as DNA or chromosomes, although those terms were fairly fuzzy to those more than two years out of tenth grade biology.

Most commonly, our requests to explain what a gene is were met with the same look that might surface in response to a question about what the World Bank does, or why tort reform is needed, or how tornados start. But unlike those technical details, which we can safely leave to economists, lawyers, and meteorologists, genes are too important to be relegated to the safely obscure. We fail to understand them at our peril, because they influence crucial aspects of our daily life: our health, our lifespan, our mood, our insurability, and our food supply, to name a few.

But before we deal with *what* genes are, let's ask an even simpler question: *where* in your body are genes found? This was the query posed to Americans of varied ethnic and educational backgrounds by Angela D. Lanie, a scientist at the University of Michigan. Nearly a quarter of the respondents said genes are "in the brain"; about one eighth said "in the blood"; a few said in the reproductive system, or heart, or bones, or lymph nodes or various other locations.

Only about one third of Lanie's respondents gave the correct answer: genes are present all over the body—everywhere, in virtually every cell. Genes are found in your skin and your stomach cells, your lung and your liver cells, your brain and your bone cells. They are in every part of your body. And not just in your body: they are in every cell of broccoli and beets, and chicken and cows, and apples and apricots. If you stopped eating

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022

food that contains genes, you'd have made the unpalatable (and unsustainable) choice to dine on not much more than sugar and water.

An understanding of genes is within easy reach for all of us, because the principles that govern the operation of genes are simple. What is a gene? It is a stretch of DNA that contains the instructions for the cell to manufacture a protein. Once the terms "DNA" and "protein" and "cell" have been explained, this definition is surprisingly satisfying to most nonscientists (and to most scientists, too, for that matter).

Let's start with the cell. Imagine your body as an enormous hotel composed of about 100 trillion (1 followed by fourteen zeros) rooms, each a self-contained space enclosed by a set of walls within which sits a bed, dresser, night table, and other furniture. Each cell in your body is a similarly self-contained unit measuring about a fortieth of a millimeter across (about a tenth the width of a human hair), surrounded by a flexible membrane that protects it from the environment. There are compartments in the cell where specific functions are carried out, including maintaining the cell's DNA, burning fuel to provide energy, and transporting material to where it needs to be.

Each room in a hotel has plumbing that connects it to a central water supply, a source of electricity to power its appliances, and heating and cooling units to control its temperature. A central processing system housed on the top floor of the hotel—a phone switchboard and a computer with an Internet connection—allows every room to be in contact with the front desk, and with all the other rooms in the hotel—indeed, with the rest of the world. Likewise, each cell connects to and communicates with adjacent cells and with the rest of the body using chemical and electrical signals.

From a distance the many rooms of the hotel look the same, but if we look more closely we see that some have one bed, others two. Some have separate kitchenettes, or sitting areas, or extra bedrooms. Some are so tiny that they barely accommodate one guest; others are suites that accommodate a large family. Most rooms are rectangular, but some special rooms have unusual shapes (like the octagonal honeymoon suite). Cells, too, are specialized to carry out particular functions, differing in their size and shape and internal structures. There are blood cells and brain cells, lung cells and liver cells—cells with long and narrow projections, cells with unusual talents to filter blood or detoxify alcohol.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022

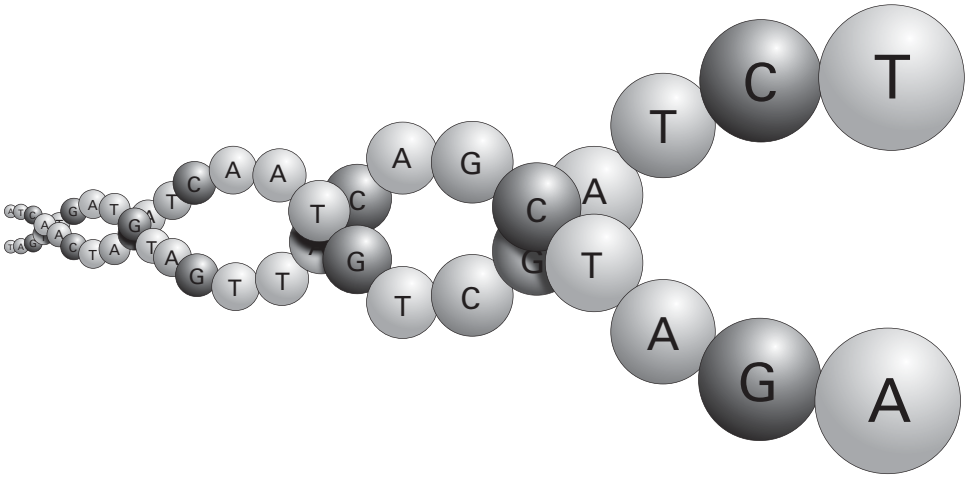
Rooms are not arranged willy-nilly in a hotel but are arrayed in orderly wings of multiple floors. Cells are also arrayed in the body in an orderly fashion; they organize themselves into successively larger units of tissues, organs, and organ systems. Most of these systems are familiar to us. The collection of cells that make up the mouth, esophagus, stomach, small and large intestine, the liver, pancreas, and gall bladder constitutes the digestive system, which processes food. Cells of the heart, arteries, veins, and blood form the circulatory system, which delivers nutrients and oxygen to the far reaches of the body. Cells that make up the bladder and the colon cooperate to manage a storage system that holds waste until it is ready for disposal. Cells of the brain and spinal cord constitute the nervous system that manages it all.

Cells entrust the instructions for their construction and operation to DNA, a chemical found in all living things. How did we come to know that DNA serves this vital function? In 1944, Oswald T. Avery, a physician and scientist working at the Rockefeller Institute for Medical Research (now The Rockefeller University) and his coworkers Colin MacLeod and Maclyn McCarty reported that they could dramatically change the properties of a cell—in their case a cell of the bacterium that causes pneumonia—by changing only its DNA. They concluded that DNA was the long-sought substance of heredity.

Attributing such importance to DNA was a startling result, because DNA was known to be a molecule consisting of a seemingly endless, monotonous string of only a few very similar subunits. How could such a "stupid molecule" (as some then called it) determine what kind of covering enclosed a bacterial cell (the trait that Avery and his colleagues analyzed), much less perform the amazing feat in more complex creatures of specifying the appearance of limbs and lungs and livers in all the right places and of the right size, and the proper number of teeth and toes, and irises and corneas and retinas that form eyes, and much, much more? Surely, thought many biologists, a more complex molecule was needed to accomplish those amazing feats.

It's easy to see why they thought this way, because DNA is indeed a simple molecule. It is composed of only five atoms: carbon, hydrogen, oxygen, phosphorus, and nitrogen—the organic elements from which all living things are built. DNA is a polymer—a long molecule made up of

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022



small units linked together, one after the other, like pearls in a necklace. These smaller units are known as "bases," and they come in only four types, commonly called by the first letters of their names: A, C, G and T (adenine, cytosine, guanine, and thymine).

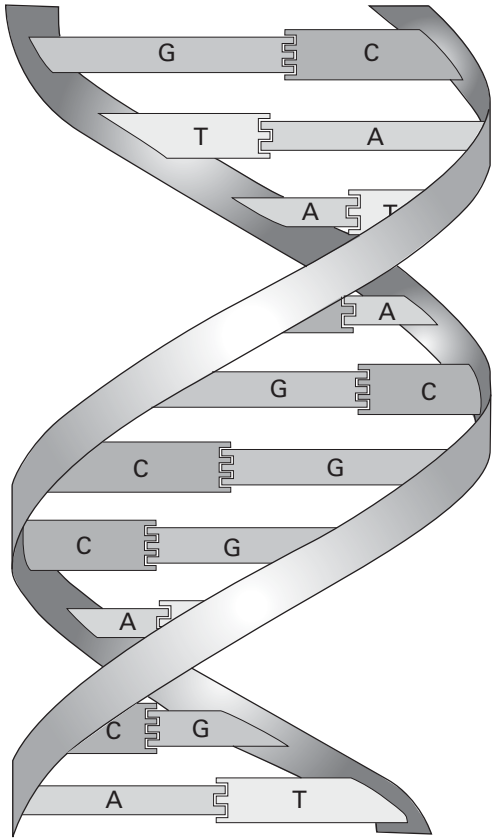
These four bases link one to another to form a very long string—think of an extremely long necklace made up of four different kinds of pearls. Two of these strings of bases wrap around each other—think of a double pearl necklace—to form the iconic double helix structure with two strands of bases spiraling around each other (see figure).

A chromosome is a long, unbroken string of the DNA double helix (along with some packaging material to wrap up the DNA molecule so it fits inside the cell). Each human chromosome is a double string of around 100 million DNA bases. In some creatures the chromosomes may be less than one hundredth that size; in others they can be up to ten times longer.

But DNA turns out to be not so stupid a molecule, because the order of the bases—the exact sequence of the A's, C's, G's, and T's—is the information that specifies the characteristics of an organism. Of all organisms. Of us.

Now comes the most important fact: the sequence of A, C, G, and T bases needs to be specified for only one of the two strands of a DNA molecule because the sequence of bases of one of the DNA strands specifies the sequence of bases of the other. This is a result of the way the two strands of the double-helical structure are held together by interactions between the bases: bases in one strand attract and stick to bases of the other.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022



But the bases don't interact haphazardly: A and T stick to each other but not to G or C; likewise, G and C stick to each other but not to A or T. Think of 110-volt electrical plugs with two prongs and 220-volt plugs with three prongs. The bases A and T are like a matched pair of a two-pronged plug and a two-prong socket; G and C are matched like a three-pronged plug and a three-prong socket (see figure). (The "plugs" and "sockets" of the double helix are really a kind of chemical bond—a quite weak one—called a "hydrogen bond"; G and C have three of them, A and T have two.) As a consequence of these specific matches, each strand of the helix carries the information to specify its partner strand's type.

When a cell divides to form two cells, it copies the two strands by peeling them apart—unplugging the plugs from their sockets—and then uses each strand as a template on which a new strand is synthesized. The result is two identical copies of the chromosome, because every A attracts a T in the

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022

newly made strand, every T attracts an A, every C attracts a G, every G attracts a C. These A-T and C-G combinations are the "base-pairs" of DNA.

On Saturday morning, February 28, 1953, in Cambridge, England, James Watson and Francis Crick realized how the sequence of bases of one strand of the DNA double helix specified the sequence of bases of the other strand. They saw for the first time how the plugs and sockets fit together. When they went to lunch that day at their favorite haunt—the Eagle Pub, not far from their lab—they left the other patrons dumbfounded with their announcement that they had learned the secret of life.

Indeed, Watson and Crick's revelation of the double helix as the structure of DNA, and the realization that the sequence of bases on one strand specifies the sequence of bases on the other, is one of the most important scientific discoveries ever made. Immediately, and very clearly, it explained a major mystery: How does one cell give rise to two identical cells? By revealing the double-helical structure of DNA, Watson and Crick answered a question that had confounded people since before the time of Aristotle: How do organisms replicate themselves? They do it by using one strand of the DNA double helix to specify the sequence of bases of the other strand, producing two identical DNA molecules from one. For their discovery, Watson and Crick were awarded the Nobel Prize in 1962.

With that discovery Watson and Crick ushered in the age of molecular biology, which provided a detailed understanding of the nature of the gene. It culminated in the Human Genome Project, an international effort to determine the sequence of base-pairs in the DNA of every one of our chromosomes. That goal was achieved on the fiftieth anniversary of Watson and Crick's discovery.

Here is a portion of the sequence of bases of one strand of a human chromosome:

CCTTCCGTTAAACCAATGGGAAACAAGTCCCTGGGAGTGTCCCGCCCT
GGGGTGAGAACTGCGAACCAATAAAAATTGAAACCTGAGCGGTGGCGC
GGCCAGCTGTGGGTGGAGTCACCCCGCGGACTGGACGGGAACCTGG
CGGGGTCAGGTCCCGTCAAGCAGCCTGGCTCATGGCTGTGTGCGGCC
TGGGGAGCCGTCTTGGCCTGGGGAGCCGTCTTGGCCTGCGCGGGTGC
TTCGGCGCCGCCAGGCTCCTGTATCCCCGTTTCCAGAGCCGCGGCCC
TCAGGGCGTGGAAGACG

This string of three hundred bases is but a tiny fraction of the sequence of one strand of human chromosome 12, which, at over 132 million

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022

base-pairs (actually, 132,349,534 base-pairs), is a medium-sized chromosome. The letters of chromosome 12, if written in the size of the letters of this book, would fill fifty-two thousand pages.

Each human cell carries twenty-three pairs of chromosomes, forty-six chromosomes in all: one set of twenty-three came from Mom; the other set of twenty-three came from Dad. Altogether, we have about 6 billion base-pairs of DNA, 3 billion in each of the two sets of chromosomes, in every one of the approximately 100 trillion cells in our bodies. It's a prodigious amount of DNA: laid out end to end, the DNA from one human would reach to the sun and back more than sixty times. These three billion base-pairs constitute the genetic material that is the human genome, the material that directs the production and maintenance of each of us. But don't let these large numbers intimidate you. Remember, as complex as the complete human genome is, and with biologists still at the earliest stages of deciphering the instructions embedded in its sequence of base-pairs, DNA is just a long chemical, and a simple one at that: just strings of A's, C's, G's, and T's.

Asimov may never have taken time out of his writing schedule to exercise or take vacations, but he would always make time for a good meal. At the age of fifty-seven, and perhaps as the direct result of consuming a giant slice of cheesecake, Asimov suffered a heart attack that hospitalized him for three weeks. He continued to experience angina over the next several years, the pain becoming so severe that even walking became a chore. In November 1983, his doctor advised a triple-bypass operation. Given the choice of waiting until after Christmas or having the operation right away, Asimov chose right away. But he worried that might prevent him from attending the annual banquet of the Baker Street Irregulars, his fellow Sherlock Holmes aficionados, which was to be held on January 6. He had prepared a song for the banquet, and although he expected to be there to sing it, he prepared a taped version that he gave to his wife, just in case.

The evening before his operation, Asimov dreamed that he died on the operating table and that consequently his wife had to play the tape for the Baker Street Irregulars, who stood in tears and applauded for, lo, twenty minutes. But Asimov survived, "and my first thought was that now I wouldn't get the kind of applause I would have gotten if I had been dead. 'Oh—[expletive deleted],' I said in disappointment."

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022

Although Asimov's operation was a success, the blood transfusion that he received was contaminated with the human immunodeficiency virus (HIV), because blood was not then routinely tested for its presence. After suffering numerous medical problems in the years after his surgery, Asimov learned in 1990 that he had AIDS. He died in April 1992 from heart and kidney complications, the true cause of death not being revealed until ten years later when his wife published *It's Been a Good Life,* composed of excerpts from his three autobiographies.

Since you're reading a book about genes, and in particular their role in disease, you may well be wondering why the first disease we mention is AIDS. Surely AIDS, which ranks among the most virulent *infectious* diseases that humankind has faced, is not *genetic* in origin, you may be thinking. AIDS is spread by sexual contact, blood transfusions, contaminated needles, and passage of a fetus through the birth canal of an infected mother. But rare is the disease that escapes the influence of our genes. So we can tell you the following quite confidently: If Isaac Asimov had had a mutation in both copies of his *CCR5* gene—a mutation that resulted in the removal of thirty-two base-pairs of DNA—he would not have contracted AIDS. This gene, identified in the 1990s, specifies a protein that sits on the surface of cells of the immune system, looking for a signal that invaders have breached the lines of defense. The HIV virus uses the CCR5 protein as a landing pad, alighting on it before invading the cell. If Asimov had lacked those 32 base-pairs in his *CCR5* gene his immune cells would not have had the HIV landing pad, causing them to be resistant to the virus. Unfortunately, even though the prevalence of this mutation is higher in the Ashkenazi Jewish population to which he belonged than in most other populations, Asimov was not so lucky. As a consequence, the world got many fewer Asimov books than it might have.

How do we find the gene responsible for a trait such as resistance to the AIDS virus? How does a gene specify a protein? What do proteins do? What does it mean to have a mutation in a gene, and why does the prevalence of different mutations vary in populations? Read on, and you'll see that these questions have straightforward answers.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/1886949/9780262289382_c000100.pdf by guest on 08 August 2022