Darwin in the Genome

Molecular Strategies in Biological Evolution

Lynn Helena Caporale

McGraw-Hill
New York Chicago San Francisco
Lisbon London Madrid Mexico City Milan
New Delhi San Juan Seoul Singapore
Sydney Toronto

Contents

Preface

I was in a new world, and . . . could not help speculating on what my wanderings there . . . might bring to light.

—Alfred Russel Wallace[1]

Long ago, consciousness began to emerge within life on Earth. Now, at the very moment when we are alive, consciousness is starting to comprehend the extraordinary information that each of us carries inside on the thin ribbons of our genome. What will we find, now that we can peer so deep within ourselves?

Just four hundred years ago, lenses were ground, placed together in a tube, and turned toward the night sky. Galileo's revelations did not require the see-to-the-edge-of-time spacecraft-mounted telescopes of our day. He looked through lenses with the strength of the 12× binoculars that today we carry so casually to examine the markings on a bird overhead. By turning such simple lenses to the sky, he showed us something that directly contradicted the overwhelming evidence of our senses. Contrary to what seemed so obvious every day from the sunrise, to the movement of the sun across the sky, to its setting on the opposite horizon, and its predictable rise the next day, we learned that the sun does not revolve around the Earth. We were certain that the sun circled around us, but in fact it does not.

When we learned that we are not at the very center of all that there is, our view of who we are and our place in the universe was dramatically changed (although there was, and in many ways still is, resistance to this change in perspective). Yet this revolution in our comprehension of the world and our place in it was, I believe, small, very small, compared to the change in

our comprehension of the world and our place in it that is about to occur as we look inside at our own genome, and at the genomes of other creatures.

The initial impulse for this book came when I organized and cochaired a conference that may well represent a landmark in discussions of evolution. This gathering was filled with enthusiastic discussions among people whose work varies from sitting at computers to studying the life that lives within beautiful shells along the beach. I was convinced that the research we discussed, and, since then, the new discoveries that continue to capture our imagination, could be shared with a broad range of readers. For nonscientists, I worked hard to break through the barrier of technical language to share the heart of these ideas and discoveries with you, but if you find that I have failed to do this in some spots, I also worked hard to make it possible to skip some sections and pick up the train of thought. There is a glossary in the back to help out, too. For readers who are professional scientists, I ask of course for patience as I explain some things that you may already have learned. (For everyone, I have inserted a few brief word games, which a friend of mine dislikes intensely and has characterized as "typos on steroids"; if you agree with him, jump over them.) I ask all readers to share with me in reflecting on what the implications of these new discoveries might be.

On the shelves of a bookstore, imaginative writings are thought to be confined to the fiction section, but in fact imagination drives scientists as much as it does novelists. Indeed, much as a novel is enhanced when its locations are described in factual detail and its characters made to seem real, so nonfiction can be more inspiring when the reader is taken into the imagination of the scientist. In such writing, as in the mind of a good scientist, the boundaries between data and imagination must be made very clear. In this book, I have worked to weave a wide range of research into a coherent vision, while describing much that remains the subject of active current investigation and even controversy. Some sections, which it is my job to indicate clearly, step into the realm of scientists' imagination, suggesting where new discoveries may lead.

As one of the founders of the field of molecular biology, Max Delbrück, pointed out, "Any living cell carries within it the results of billions of years of experimentation by its ancestors." The steps of this long journey are inscribed in our DNA, and in the DNA of all the living creatures that share our time on Earth. There are messages in our genome for each of us; they come from our diverse ancestors, and they outline our path across past ages, in other forms, to here. If we strain to listen, will our genome help us to comprehend how mutation and selection ever could have led to us—to

me, sitting here and writing, and to you, sitting there and reading? I have written this book to explore these very questions.

Our genome encodes many types of information. Perhaps the best studied is the information that encodes our proteins. Other information controls which proteins are made in a cell in our heart, or in our brain, or in our thumbs. But it is clear that there is room in our genome for other types of information that we are only now beginning to perceive. This information can be found in the spaces between genes, and even doubled up along the stretches of DNA that encode our proteins. Early examples of such multiple messages involved information encoding proteins that ran together along a piece of DNA in different reading frames, something like you find **tthhee itrw ol emtetsesrasg easl tmeirxneadt iunpg** here. In fact, it was this ability to transmit more than one message through the same stretch of DNA that first led to my writing about evolution, as described in Chapter 13.[2] As will be explained in this book, I have proposed that this extra coding potential has been harnessed[3] by natural selection to improve the probability of survival. Some scientists may disagree with my interpretation of the data before us, leading to interesting discussions about data; in studies of the genome, data are coming very quickly now. Some people do seek data that show that their ideas are right, and discard data that are not consistent with their preconceived ideas, but that is not science.

Because this book mentions evolution, it may attract religious debate, but faith is, of course, something entirely different. This book addresses scientific research, not faith. But knowledge does have implications. Rather than view evolution as a no-holds-barred fight to the death, I find that the discoveries described in this book harmonize with the teachings of many great religions: that people are brothers and sisters; that we should respect, welcome, and share with others; and that reverence should extend to other forms of life. Indeed the intellectual, practical, political, and philosophical implications of being able to read the information within genomes reach into everyone's lives and emphasize the importance of diversity for survival. This includes biodiversity and, especially, should lead us to treasure the diversity of our own species.

On a practical level, the knowledge that we gain from looking into our genome will facilitate medical research and also may teach us ways to design novel coding systems for our computers. By bringing our consciousness into the analysis of microbe and tumor genomes, our new skills should strengthen our hand in the age-old battle against these adversaries, as discussed in this book. Some private messages in each person's genome may

warn of susceptibility to certain diseases, but other information will reassure people that there are other diseases that they are less likely to suffer. Many observations lead to caution in the face of proposals to fix "errors" in the human genome.

When I was a little girl sitting at my desk in school, I was asked to draw a circle around a picture of a ball and to connect it, with a line, to the word *ball*. I assume you too had workbooks in which you connected pictures and words, one at a time. In retrospect, it was a lot like that when we first began to look at our DNA, connecting one gene to one protein. But now that we finally can see the genome as a whole, we are learning new ways to read DNA that go beyond letters and words, to understand decision-making networks and hierarchies.[4] Physically, a genome may be a string of letters along a strand of the DNA double helix, but functionally it is an interconnected, highly cross-referenced system. In the more than 3 billion letters of our genome, there are concepts to pull out, ideas in one place that are connected to and built up from ideas in other places. All of biology, including our journey to this place, is waiting within us to be pieced together into an integrated living, whole. I feel fortunate to have become a biochemist at this extraordinary moment in history, and I invite you to join me on this journey of discovery into our genome, to explore what our "wanderings there might bring to light."

Prologue
Chance Favors the Prepared Genome

Delicate, elusive . . . is that mysterious principle known as "Organization," which leaves all other mysteries concerned with life stale and insignificant by comparison.

—Loren Eiseley, *The Immense Journey*

There was a moment in time when the dust itself edged, in slow motion, over a boundary into life. It entered onto a path strewn with dangers, uncertainty, and creativity. It spread its growing skill across the Earth until it learned how to fly at will and how to sit still here, and to discuss its own evolution. Those who do not believe that we have evolved from life forms that are invisible to the naked eye—and even those who do—find it hard to *conceive* of how this journey, relying only on random mutation and survival of the "fittest," could have succeeded. It seems almost inconceivable that there has been enough time for mere molecules to organize themselves into a being that could compose music, travel to the moon and back, and indeed analyze its own genome. How did we get this far, even once, in only billions of years? How could we have happened so randomly?

In the greatest achievement in human intellectual history, the information in our genome, the product of billions of years of evolution, is now opening before us. What answers does our genome hold to the Big Questions that whisper to us all? What does it say of the immense journey it has taken as it has passed through uncounted life forms to be carried within us as we sit here, discussing our origins and our fate?

It has been well over a century since Charles Darwin and Alfred Russel Wallace first proposed that from among the natural variations in a population, the "most fit" individuals would tend to survive in greater numbers, passing their selected variations on to their progeny. But when Darwin and Wallace proposed that evolution happens through variation and then selection, they did not know the mechanism by which the varied traits were inherited. Science at that time was still was about a century away from uncovering the chemistry of heredity. And, although Darwin and Wallace had each traveled from England to the Pacific Ocean, they had not traveled to Gregor Mendel's garden at Brno and so had never talked to their contemporary about how traits might be transmitted between generations. (What a conversation that would have been!)

For his part, Mendel went beyond the general understanding that we resemble our parents and showed that traits, such as whether the peas borne by the plants in his garden were wrinkled or smooth in form or yellow or green in color, were inherited independently of each other in a way that was *predictable*.[1] Mendel used careful statistics to work out the "laws" of heredity. He could anticipate the mix of smooth and wrinkled, yellow and green peas borne by tall and dwarf plants from generation to generation, but he could not explain how this worked. We now know that the predictability of the peas comes from the fact that the inherited variations of the peas are encoded in separately assorting stretches of the DNA of pea chromosomes.

Of course, the naturalists and the statistician/monk did not have a chance to travel forward in time to kick ideas around with researchers who studied the mechanisms of mutation, but some decades after Darwin, Mendel, and Wallace were gone, their ideas were incorporated into the then more recent work to form our current understanding of evolution: Variations among organisms are due to variations in genes; variation is due to different "versions" of genes (giving, for example, green peas or yellow peas, green eyes or blue eyes) and mutation of the DNA that makes up our genes. From this variation, selection picks the most "fitted."

When we say "fit" in our daily lives, this usually leads into a conversation about workouts, but for the moment I would like to talk about personality quirks. Imagine that a developer built new homes in an isolated canyon. The homes were very desirable, but the developer had a prohibitive personality quirk: He would allow only certain cars through the gate on the road that led to the canyon. No one understood why their car was allowed through the gate or turned away, but I can tell you: The imaginary developer let families get through the gate only if at least one child in the

car was wearing a blue T-shirt. The natural consequence was that among the people who settled in the valley, there was a higher proportion of children who tended to wear blue T-shirts than there were in the general population. After a couple of generations, the isolated canyon became crowded, and the developer's grandchildren built homes in the next canyon over. The developer's grandchildren inherited the developer's quirky attention to T-shirts, but they were so used to blue, from its enrichment in the first canyon, where they too lived, that they did not favor blue at the new barrier. They would admit only families that had a child wearing a yellow T-shirt.

After generations of families getting through gates controlled by generations of quirky developers, the tenth canyon was inhabited by families that, without thinking about it, each day dressed their several children in different-colored T-shirts. These colorful families were the ones that had the best chance of having one of their children wearing the "right" color T-shirt and thus being able to get across barrier after barrier, from the canyon enriched in blue to the canyon enriched in yellow and beyond.

As genomes travel from generation to generation across evolutionary time, we face something far more serious than quirky developers. Pathogens are among the life-and-death challenges that allow only some children to pass into the next generation. Just as the imaginary developers let families who varied the colors of their children's T-shirts get through, pathogens let genomes get through that varied their progeny enough to avoid the pathogens' biochemical tricks.

You can be too careful. True, a genome must be conserved as it is passed from generation to generation. To reproduce a genome, it is necessary to be careful in copying DNA and to repair errors. But having all progeny be exactly the same may not be the safest strategy. In fact, Darwin and Wallace and their contemporaries were impressed with the tremendous variation they observed within each species, from birds to beetles. Now, the genome, which had been hidden from their view, is becoming a landscape for a new kind of naturalist to explore. This naturalist views the variations within each species using a different kind of binoculars. Far from being carried on a strap, these "binoculars" involve laboratory infrastructure and computers to read and analyze each species' DNA. The variation that these naturalists study is not limited to that between feathers and limbs; rather, what captures our imagination is the variation among worm or flower or human, or bacterial genomes.

What has made evolution so hard for many to accept is the assumption that it depends upon random mutation for the generation of new variations. Momentarily sloppy, the gene-copying mechanism drops something, messes

up, and passes on a mistake to a probably unfortunate member of the next generation. Through sheer luck, the change in the DNA, the accident, may turn out OK, so that the child who inherits it survives and passes it on. Rarely, through even greater luck, the change may turn out to be for the better; with those rare lucky accidents, the random mistake makes a fitter child (or fawn or tadpole or sprout or bacterium), one that is favored by natural selection. Slowly, over unfathomable lengths of time, from one rare, lucky mutation after another, these rare fitter children in turn give birth rarely and accidentally to even fitter children; and so, at its stumbling pace, evolution proceeds, selecting any advantage in a wing or a protein, one by one.

That variation comes from random mutation of DNA was not, of course, Darwin's proposal. But just as Darwin and Wallace could not incorporate into their theories what they did not know about genes, when our current understanding was developed, there was a lot that we did not know about genomes. In this book I will propose that it is time to incorporate our new discoveries into our understanding of evolution. As Baldomero Olivera, whose work will be discussed in Chapter 3, pointed out when describing his observations, "Unconventional hypotheses for these unusual data merit serious consideration."

The work described in this book has led me to the conclusion that natural selection must work not just on each individual mutation, but also on the very *mechanisms* that generate genetic variation—as it does on all biological functions. The research discussed in this book leads to the conclusion that mutations are not all accidents and that mutations are not always random. Our genomes, and those of other life forms, have evolved mechanisms that *create* different kinds of mutations in their DNA, and they reuse and adapt useful pieces of DNA, even to the point that there are genomic "interchangeable parts." Biochemical mechanisms can arise that tend to focus genetic variation, resulting in "hot spots" of genetic change at certain places in the genome. The probability of genetic change at any given point in the genome is dependent upon the surrounding sequence of the DNA, the environment, and the proteins that are present in the cell that interact with the DNA; for example, specific types of mutation can be increased in our immune system.

Evolution may not have been reaching for the goal of two eyes and a brain and two arms and two legs, but it didn't just stumble onto us through clumsy wandering. Randomness fades in a world that rewards each step of getting better at finding food, avoiding predators, or adapting to recurring challenges. As the dust organized, it faced selection. Over time, there emerged something that, viewing the effects now, we might call *strategies*—

such as the ability to actively generate diversity—that enabled life to emerge from the darkness of random wandering. Because the mechanisms that change the genome fall under selective pressure, I propose, based on the new observations discussed in this book, that information can flow back from survival to the places in the genome that affect the generation of the diversity that we see around us, and that this will make genomes become more efficient at adapting and evolving. If one of the predictable characteristics of the world is that it changes over the course of generations, natural selection will lead to organisms that are more efficient at adapting to an environment that may change.

These discoveries do not refute the theory of natural selection developed by Darwin and Wallace, but instead provide a deeper understanding of how natural selection leads to organisms that are better adapted to their world. Natural selection acts on all biological properties. That means that natural selection acts not only on fins and wings, but also on the mechanisms that change a genome. With time, it turns out, the "fittest" genomes, the "successful" genomes—the ones that survive—are the genomes that evolve what here I will call mutation strategies. Some readers may disagree with this use of the word *strategies*, as I am, after all, discussing groups of molecules. But I use this word to emphasize that the molecular mechanisms I will describe in this book have the effect of anticipating and responding to challenges and opportunities that continue to emerge in the environment.

The first strategy for survival clearly is to generate diversity. The long-term survival, or fitness, of a genome often depends upon the diversity of its descendants. Genomes have evolved biochemical mechanisms that actively diversify themselves. The more diverse the progeny, the better the chance that at least some progeny will be different in a way that allows them to survive or even thrive, whether they are in an isolated canyon, a salt cave, a hot spring in Yellowstone National Park, or an irradiated can of meat—for life can survive in all of these places.

Miroslav Radman, whose work is discussed in Chapter 8, described it this way: "The generation of a large repertoire of biological diversity [is] the evolutionary equivalent of buying a large number of lottery tickets."[2] The lottery winners are those who survive natural selection. You don't want everyone in your family to buy a ticket that has the same number.

We must have sunlight to construct vitamin D in our skin, but too much sun will burn us. With dark skin to protect against sunburn, a child can survive at the Equator. With lighter skin to let in more sunlight, a child can be wrapped in warmer and warmer coverings and begin a journey away from the Equator. The world inhabited by these children can be very different

from the world that best suited their parents. If the world may become different, it is an advantage for some children to be different too.

If the ability to generate diversity is a useful skill, the fact that genomes can generate genetic diversity in more than one way provides an even greater advantage.[3] To prepare for various levels of selection, the genome too can change just a little, a little more, or still more between generations. Maybe one quirky developer admits applicants to the new canyon not on the basis of the color of their shirts, but on the style of their top shirt buttons. Another issue in getting into some canyons may be the size of the car (too big?) or the strength of its engine (too weak?). A genome's ability to grow and to explore new organizational structures would be severely constrained if its options were limited to changes in the molecular equivalent of the top shirt button. A single letter along a strand of the DNA double helix can change to another letter, but also a patch of letters may expand, be replaced, or be removed, or some pieces of the genome might be rearranged. By now a diverse set of biochemical mechanisms of change has emerged, each mechanism generating a different type, rate, and extent of diversity.

A second strategy that has emerged in genomes, and that will be discussed starting with Chapter 7 in this book, is the reuse of useful pieces of genetic information. This is exemplified by the spread among bacteria of information that encodes resistance to antibiotics. Other useful genetic information, such as how to digest a new food source, also can come into a bacterial genome from outside—in other words, from a genome that was not its parent. As in a port city in a nation of immigrants, within bacteria new genetic ideas arrive, are put together, survive, prosper, and can thrive. Useful genetic information that is already within a genome also can be adapted for a new job, molded by making an extra copy of a piece of DNA, moving the copy around, and tinkering. Shuffling DNA around within our genome can have risks and do damage, leading, in people, to "birth defects." But, as the roof falls, albeit ever so slowly, in, sitting still also has its risks. In the high-stakes game of evolution and survival, genomes don't take time to reinvent the wheel. They network, copy, vary, and explore the potential of the information they already hold inside them.

As is described in several chapters, genetic change is not something that strikes all parts of a genome evenly. It also has become clear that, as illustrated in Chapter 8 by Evelyn Witkin's work with sunburned bacteria, the likelihood and type of genetic change, or mutation, can vary depending upon which molecules a cell contains. In other words, rather than being purely passive, the genetic change that a cell experiences can become somewhat conditional on, for example, which proteins the cell itself makes.

Under some conditions, one bacterium may become more open to new ideas, more likely to swallow DNA; a neighbor might feed it a gene that encodes a recipe for destroying an antibiotic.

While a genome evolves a balance between faithful copying of itself and exploration through mutation, this is a difficult balance to get right. Perhaps it can never really be "right" because the right balance between fidelity and exploration may change as threats and opportunities in the environment change. As Nobel laureate Barbara McClintock, whose work will be described in Chapter 15, said, "In the future attention . . . will be centered on the genome . . . sensing the unusual and unexpected events, and responding to them."

It is becoming clearer and clearer that some classes of nonrandom mutations are very appropriate to the needs of the organism. In this book, I will discuss recent information that supports this new understanding of evolution, which I first proposed in the technical literature in the early 1980s.[4] These ideas attract controversy, but the evidence coming from sources as diverse as the bacteria that cause Lyme disease and our own immune system is growing strong. The work of Richard Moxon and others demonstrates that mutations can become more likely at the very spots in pathogen genomes that speed their race to get a grip on us and to survive. We too are the survivors of many battles with pathogen across the ages, and so those parts of our genome that encode our immune response are creative sites of focused mutation.

To reject purely random mutation as the current substrate of genome evolution is not to reject Darwin and Wallace. Indeed, while Darwin's name may be connected to the phrase "survival of the fittest," this book emphasizes these words of his: "I have called this principle, by which each slight variation, if useful, is preserved, by the term Natural Selection."[5] Among the variations that I propose are preserved, when useful, are intrinsic variations in the probability of mutation along the genome, as described, for example, in Chapter 4.

The flow of information from the biological effects of genetic change to the intrinsic variations in the type, location, and probability of mutation and the mechanisms that generate mutations is not a simple loop, for we are not adjusting an aileron to restore level flight but must survive the unexpected. Still, intrinsic differences in genetic variation would *tend* to focus in classes of places along a genome where these mutations are more likely to be creative, and tend to move away from areas where changes have done more harm than good. This is a tendency, an adjustment, and not an absolute; for still, now, many mutations do damage. This adjustment would

emerge because those genomes that accidentally keep losing important information will have fewer descendants across the generations and thus will be less likely to survive. In contrast, genomes will tend to endure when their most likely mutations create effective responses to their most likely challenges, as is detailed in this book. As the fittest molecular strategies emerge through natural selection, by the survival of the descendants of the genomes that encode them, those who remain in the world tend to be those whose ancestors were lucky enough to keep making more creative mistakes. Any genome that we find today, including our own, has been successful, because it has survived through the molecular equivalent of countless canyon gates—though, of course, it has survived only this far, so far.

A genome can't predict what will happen to the next generation, nor can you or I. But genomes have faced some challenges over and over again, such as in host/pathogen battles, and this has left its mark on the genome. A genome evolves a "worldview" of which types of changes, under what types of circumstances, may yield a new function and are less likely to destroy something essential. While a genome can't predict the future, a genome that has been so prepared by experience is likely to be favored by chance. For me, evolution becomes more conceivable when it is viewed through such a "strategic" biochemical window. The ability to evolve and adapt is an acquired skill, responsive to the environment and acquired through the experience of genomes across generations. Through selection, genome structure emerges from randomness.

We live at an extraordinary moment in human intellectual history. Until now, the information that gives rise to all that we try to study in biology and medicine remained hidden from view, as if our eyes were covered by blindfolds. These blindfolds are becoming transparent, and soon they will fall off. The sequences of entire genomes are opening before us.

Varied genomes, based on similar chemistry, have spread across the Earth, taking advantage of opportunities, establishing themselves in new environments. After the asteroid hit, what emerged was not another *Tyrannosaurus Rex* but instead us. Each of us is, in a way, an experiment, and an example of the life-preserving, creative diversity expressed at our moment in time by the human genome. Indeed, we share with one another, no less than with the majesty of the redwoods and the doves, the fact that each of us is a unique creation of the barely tapped potential immanent in the first genomes on Earth.

1
Diversity or Death

"We are caught in an inescapable network of mutuality, tied in a single garment of destiny. Whatever affects one directly affects all indirectly."

—Rev. Martin Luther King, Jr.,
"Letter from the Birmingham City Jail," 1963

It was a mystery. There were very few of these lucky people, but there were a few. They surely had been exposed to HIV, but they simply did not get sick, not even after ten years. Each day each one faced the terror of knowing that the deadly virus had touched his body. And yet each one awoke, morning after morning, to the tentative joy that there was no sign that the virus had gained a foothold inside him. For each one of these lucky few, doctor after doctor—nurses, nutritionists, researchers, reporters, friends—hoped, hypothesized, and investigated. What was he eating, taking, doing differently? Or was it something in his genes? Yet as each was questioned, poked, and wondered over, it remained a mystery.

This mystery remained unsolved for a full decade after the HIV virus was discovered in people with AIDS. While this mystery remained unsolved, Ed Berger, at the National Institutes of Health near Washington, D.C., was studying a different HIV mystery. Berger was trying to figure out how HIV gets into cells. HIV does serious damage when it invades the T cells of our immune systems, but it can't get into just any cell. To get inside a T cell, HIV needs to find a gate it can crash, a *receptor* on the T-cell surface. The receptor is a protein called CD4. But if we force a cell that does not normally let HIV in to put the receptor protein CD4 on its surface, HIV still cannot get

into that cell. So CD4 is not the whole answer. Berger thought that there must be another molecule, a *coreceptor,* needed along with CD4 to let HIV in. He was right. Berger discovered the coreceptor, called CXCR4.[1]

CXCR4 allows HIV to spread from T cell to T cell. Finding CXCR4 led investigators to another coreceptor, a protein called CCR5 that looks a lot like CXCR4. CCR5 also can let HIV into cells, but it lets HIV into different cells. Unlike CXCR4, CCR5 is not found at the entry to most T cells. HIV needs to get into T cells in order to destroy the immune system, but first it needs to get into our bodies. CCR5 opens the front door to HIV. For most people, CCR5 is what helps the virus get into the very first cells it infects, where it first touches a person. These cells usually are not T cells.

Most HIV particles that attach to CCR5 and infect us cannot get into T cells and thus cannot directly harm our immune system. But, once CCR5 has let HIV inside some of our cells, the deadly invader has penetrated the barrier between being outside of us and being inside us. It has gained a foothold in our bodies; it is living with us, within us.

Once inside, HIV experiments. It floats around, poking and prodding, and exploring. It mutates, and these mutations inevitably produce a small change on its surface that it can use to attach to the coreceptor door. Once HIV's coreceptor binding site mutates to a form that binds to CXCR4, it has evolved the key to let itself into T cells. It continues to mutate and can become an even more unwelcome guest, discovering the keys to additional doors.[2]

If the HIV in other cells didn't mutate into a form that could get into the T cells, we might all be infected with HIV without even knowing that HIV exists. This is not a reassuring thought. In fact, I find this haunting: How many other viruses already have found their way inside me, changing, exploring, trying out new keys, perhaps doing no harm—for now?

Berger had set out to solve the coreceptor mystery, but his work also explained the mystery of the people who remained healthy after they were exposed to HIV. The discovery of the coreceptors CXCR4 and CCR5 not only explained how HIV gets into cells, but also, unexpectedly, revealed why those few lucky people were so resistant to HIV.

The answer to the mystery was that the lucky survivors' CCR5 coreceptor was different. It was damaged—it was missing a piece, and so couldn't let HIV in. HIV could not get a foothold in their bodies; it could not get in the door. Out of all the countless molecules in a human body, this one mutation—one small change in one protein—was enough to keep HIV out of the cells of a few lucky people and so saved their lives.

Like all breakthroughs, the discovery of mutant CCR5 answered one question but also led to so many others. How did these lucky people get the unusual CCR5 protein? Obviously it was in their genes, which came from their parents, but what generated the mutant protein in their parents' genes, or their parents' ancestors' genes? Which ancestors were they? Other primates, indeed other mammals, have CCR5 receptors too. Had the ancestors of the HIV-resistant people already encountered HIV, or perhaps another pathogen that uses the same coreceptor to get in? Perhaps smallpox?[3] If so, how did *their* ancestors prepare for it and survive?

Or, had the genes of their ancestors, and our ancestors, somehow "learned," by surviving infection after infection for generation after generation, that something like HIV, another new pathogen, inevitably would come? To become a survivor, it may simply be enough for an individual to be a little different, in a way that provides no apparent advantage in fitness until a new, never before encountered pathogen appears. For an individual, such a difference is great luck; but at the level of the genome, it may reflect a strategy that has emerged in successful genomes, a way to be prepared for the unexpected.

The mechanism of evolution, natural selection, usually is described as "survival of the fittest." But as we look at genomes that have been handed down from generation to generation, what do we mean by survival? What do we mean by fittest? Survival demonstrates fitness: We gather that the *successful* ancestors, the ones whose genes made it to us, must have had the better genes. But what *is* "better"? "Better" is a moving target. Continents move; climates change; predators, competitors, food, and the atmosphere all evolve. We call mutations "errors," but from the perspective of evolution, the most serious error for a genome is to make no mutations.

Genomes can prepare for the unexpected by being diverse. If every one of us were the same, and a pathogen were to hit us suddenly and hard (especially before we had laboratories that could work more quickly than generations), we could be wiped out, all of us. The human genome would become extinct in the brief time it took for the pathogen to spread through the human community and do its work. Our libraries and our architecture would be left behind for the product of a future genome to decipher. The reptiles and now the mammals having had their day, perhaps the cephalopods would come next.

But, if every one of us were a little different, even in some small way, a few of us might be different in the very way that would protect against this new pathogen. Then a few of us would survive, and with the survivors the

human genome also would survive. It would not become extinct, but would continue to come forward in time, as it came forward to live within us. It would survive in us, walking, through our descendants, into the future.

I think back before protease inhibitors, before antibiotics, before research labs, before we understood how a pathogen spreads (swamp air? drafts?), before we could fight back against pathogens. In that time before worldwide travel, a new pathogen like HIV could sweep through a local human community and kill nearly everyone in it. Indeed, it does not take a journey back in time to see this, just a journey to many places on the planet. HIV cuts huge holes in families, in communities, even in entire countries where there is no access to information about prevention or the right medicine for treatment. As I write this, I am reading a report from Allan Rosenfield, dean of Columbia University's School of Public Health,[4] pointing out that 28 million people in Africa are HIV-positive. There are 12 million HIV orphans; one-third of all adults in sub-Saharan Africa are infected. Parts of Asia are risking the same fate. The human genome will notice this.

In an unprotected community, a new pathogen with the right keys could kill everyone except those few who were just a little different, who had a mutant protein in the right place—like, for example, those with the mutant CCR5 coreceptor. After the pathogen had swept through the unprotected community, if the human genome remained there at all, it would have been touched by this tragedy. It would "remember"; it would have been adjusted. Among the survivors, and their children and their children's grandchildren, the mutant protein would no longer be rare; it would have become the common form of the protein, the one that everyone had. Examining the genome centuries later, we never would guess that it was ever otherwise. We might never be inspired to wonder what protection a now-lost form, or a minor, unnoticed new mutant form, might provide against a different pathogen another day. The pathogen, by removing those without the once-rare mutant protein, would have left behind a changed human genome to be shared by those who survived to live in the future.

This has happened before. There is clear evidence that a pathogen has marked our collective genome. This is a pathogen with which we are still doing battle, which kills the equivalent of one to three 747 loads of people every few hours. Most of the passengers on these imaginary 747s are little children, the majority of them under 5 years old. They've been bitten by mosquitoes and are suffering from the anemia of malaria, with two-thirds of their oxygen-carrying red blood cells destroyed. Every day, every week, last week, yesterday, today, it continues.

And for each person who dies, many more are ill; 300 to 500 million people suffer from malaria each year.[5] It is as if every single person in the United States, plus every single person in England, Germany, and Japan, were infected, each one shaking with the chills and sweating with the fever caused by this persistent pathogen. Of course, few of the real victims are lying in beds in the United States, England, Germany, and Japan. Most of malaria's direct victims lie in a belt across the middle of the Earth, a belt containing close to half of the people in the world. Their parents do not anchor our evening news, nor edit our daily newspapers, nor write the advertisements that shout to us on TV; but these children share with us this moment in time, in the continued evolution and sculpting of the human genome—and the human genome takes notice.

With hundreds of millions of victims across the centuries, including today, the human genome has felt the pressure of the malaria parasite. Of a group of children, friends and brothers and sisters, all bitten by infected mosquitoes, one child comes down with a sudden fever, with seizures, and slips into a coma. Left untreated, half of these ill children die; but even without treatment, half of them live. And some children, even though they are attacked by malaria, don't get very sick. In the genomes of these survivors, some genes are different. Just as HIV kills those who lack the mutant CCR5, leaving as survivors those with the rare version, malaria too has sculpted the human genome.

This sculpting of our genome, this battle with malaria, is not without its victims. For these victims, too, the struggle is painful. Tears roll down the cheeks of an infant screaming in pain. She cannot explain it, but it feels as if countless nails are being driven through her body. Her red blood cells, the carriers of desperately needed oxygen, are bent out of shape and are jamming up her blood vessels. Her nerves, sensing the lack of oxygen, are exploding in pain. There were over 300,000 of these children, born last year,[6] in a tropical band across Africa and in families whose ancestors lived in these places, suffering from what we have come to call sickle cell disease. What happened to leave so many children suffering in, of all places, the paradise of our imaginations, the tropics?

Biochemically, at first glance, it is a minor thing that is causing so much suffering for the little girl—a change in a single amino acid in one protein out of many tens of thousands of distinct proteins in her red blood cells. But this tiny change is in hemoglobin, a vital molecule that carries oxygen to all of our tissues. Surely such a mutation, one that messes up a vital molecule, damages the red blood cells, and causes so much pain, should have been removed from

the human genome by natural selection long ago. But it has not been removed. In fact, this painful mutation seems to have thrived and spread.

This child's pain comes from the wounds of a battle the genome has fought against a tiny adversary, invisible to her unless someone were to give her a microscope. The tiny adversary is a protozoan, with its relatives called *Plasmodium*, that causes malaria. Both copies of her hemoglobin gene are altered, and the mutant protein encoded by these genes has sickled her red blood cells out of shape and jammed up her tiny blood vessels; but when only one of the copies of this hemoglobin gene is altered, as in the genome of each of the little girl's parents, it can be a good thing for people who live at sea level in the tropics, for it protects them against malaria. And so, as malaria continues to take down its jumbo-jet loads of people in the tropics every few hours, those who survive the attack are more likely to have a mutant hemoglobin within their red blood cells. In the tropics, the malaria parasite has left its mark on the human genome.

For malaria and mammals, the oxygen-carrying red blood cells have been a major theater of battle. The hemoglobin molecule, the molecule that catches oxygen in our lungs and delivers it through the body, has been a major battleground. Not only sickle cell hemoglobin, but also the thalassemias, a swath of variants in hemoglobin structure and regulation, follow the anopheles mosquito that carries malaria. Surely, after all this time, the mammalian genome should have won; it should have evolved a way to keep the malaria parasite out of our cells. Indeed, we have changed. But so has our adversary. As the malaria parasite has sculpted our human genome, we surely have sculpted its genome, too.

We can change, but there is a limit to how drastically we can change our red blood cells. We can fool around only so much with our life-sustaining oxygen-carrying mechanism. And so the battle with our ancient adversary continues. Perhaps now that we finally are gaining the real-life equivalent of the magical secret decoder ring with the sequencing of the complete malaria genome, we will find a more effective vaccine or therapy and at last overcome malaria's bag of tricks. Perhaps our new level of consciousness will finally bring us complete victory over our ancient adversary. But not yet. For now, malaria continues to take down its 747 load of victims every few hours, last week, yesterday, and today, this afternoon. If you happen to glance at your watch again in a few hours, note that another planeload of children will be gone.

As we look into the forest at night, instinctively we fear snarling fangs and coiled vipers. But it is the once-invisible predators—protozoa such as

malaria; bacteria such as the plague, typhus, and tuberculosis; and viruses such as smallpox—that take the largest toll on our species. Just as they still threaten us today, these tiny predators put selective pressure on our ancient ancestors. As our ancestors escaped from and banded together to capture the large predators, the predators we listen for when we wake up in the moonless forest at night, our genes have been doing battle with the tiny ones. And new ones keep emerging. Where does Ebola hide deep in the forest?

One evening in the mid-nineteenth century, Alfred Russel Wallace lay on his cot at the edge of the rain forest shivering with fever. He had traveled halfway around the world to study nature, and now, on an island in the Moluccas, he was getting a much closer view of it than he had planned. As he lay on his cot for hours each day, alternately shivering and sweating, he later wrote, "I had nothing to do but to think over any subjects then particularly interesting me." One day he began to think about Malthus's book *Principles of Population*, which he "had read about twelve years before." Wallace began to think about Malthus's "clear exposition of 'the positive checks to [population] increase'—disease, accidents, war, and famine." He connected the idea that disease, accidents, war, and food shortages, which would limit the ability of the human population to keep increasing, also could limit the increase of animal populations.[7] Animals breed so quickly, he reasoned, that there must be an enormous loss each year or "the world would long ago have been densely crowded with those that breed most quickly."

As he reflected on "the enormous and constant destruction which this implied," his mind focused on the question, "Why do some die and some live?" This was not a purely abstract question at that moment for Wallace, who was lying with a raging fever in a hut on an island far from home. For him, "the answer was clearly, that on the whole the best fitted live. From the effects of disease the most healthy escaped; from enemies, the strongest, the swiftest, or the most cunning; from famine, the best hunters or those with the best digestion; and so on. . . . That is, the fittest would survive."

Wallace's next thoughts were about dramatic changes in the environment and the great amount of individual variation there is within a species.

> *Then at once I seemed to see the whole effect of this, that when changes of land and sea, or of climate, or of food-supply, or of enemies occurred—and we know that such changes have always been taking place—and considering the amount of individual variation that my experience as a collector had shown me to exist, then it followed that all the changes necessary for the adaptation of the species to the chang-*

ing conditions would be brought about; and as great changes in the environment are always slow, there would be ample time for the change to be effected by the survival of the best fitted in every generation.[8]

With his own life threatened by disease, Wallace had discovered the idea of natural selection, survival of the fittest, all on his own, before hearing the idea from Darwin. While the phrase "survival of the fittest" seems to ring in our ears, there is another, quieter phrase that needs some attention. Wallace said that his realization of the role of natural selection was based on "considering the amount of individual variation that my experience as a collector had shown me to exist." Variation is an essential prerequisite to selection.

Variation comes from mutation: changes in a single genomic letter, changes in a block of letters, and the rearranging of pieces of DNA. The ultimate creative step in variation, for those who have survived to become adults, is the cutting and pasting and mixing and sorting with the genome of another adult that happens when we create a new life together—a child created by mixing pieces of the genomes of all our ancestors.

Survival-of-the-fittest: The words are spoken as one word. But what carries our own specific DNA forward in time? Our DNA survives in a partnership with others, carried like candlelight passed through a branching chain of individuals. As the torch is passed at the border of each new generation, our DNA gets cut up and mixed with DNA from other people. The work of the other genes in the new mixture may increase or decrease each gene's "fitness."

Ten generations from now, an eyeblink in evolution, if your DNA has survived the journey, it will have been diluted to a mere thousandth of your descendants' genomes. It will be mixed in with the genomes of thousands of other people who are now alive, most of whom you do not know. Perhaps a part of your DNA will find itself connected to the DNA of someone in the blue car, in the center lane that you passed yesterday morning, or that person who stepped aside as you left the train in the subway, or the child you saw in the arms of his mother, clinging to a tree, in a flood in a distant land reported on the evening news. Together, at a time in a future that we cannot clearly imagine and that we will not ourselves live to see, all of you may share in the creation of a child.

With luck and fitness, these descendants will survive. Is this how, and when, we tell that you were the fittest? The fittest what? The fittest when? Now, or when future fragments of your genome encounter a new pathogen? What is "better"? What is fit when you are aiming at a moving target? Fitness emerges as a strategy, not a goal; a process, not a place.

The pain and tears of the little girl with sickle cell anemia are her wounds in a struggle that is for us, for the survival of our shared genome. She is suffering because her parents carry important information in their genes, information to be shared, through her parents' children's children, with our children's children, information that protects against malaria. For even today, malaria is not tamed, and it is not caged: In 1999, at a camp about a 1½ hour drive from the Empire State Building, two 11-year old boys got malaria.

To survive, we must absorb this truth, that our genome's ability to change, to explore, to incorporate the discoveries of many individuals—and the diversity that results from this exploration—is a central part of our fitness. It is a lesson incorporated into our genome through billions of years of evolution. If we do not treasure human diversity, we risk an eternally broken chain, the end of our species, the loss of our future. If the human brain and the human heart do not learn fast enough what we have known in our bones, and in our eggs and sperm, for millennia, the genes that have brought us this far, through all of their diversity, will be wiped out. Our hopes, our future, indeed our very survival in a distributed, gene-mixing rebirth, depend upon our connections with one another. We are all, profoundly, siblings in the present, parents of the future.