# Figure and Ground

## Translating the Genome

by Tom Sgouros, Jr.

May 8, 2001

> Life is not an illogicality; yet it is a trap for logicians. It looks just a little more mathematical and regular than it is; its exactitude is obvious, but its inexactitude is hidden; its wildness lies in wait. —G.K. Chesterton, *Orthodoxy* (1908)

## 1 Patterns and Languages

Above the table where I write, there is a still life oil painting. The rendering is brusque, almost telegraphic, and the painting readily dissolves back and forth between an image — a coffee pot, a bottle and some garlic — and a meaningless pattern of brushstrokes. The uneasy balance between meaning and mess is delicious, and it's what I enjoy most about the painting.

This book is also a pattern, made up of 26 letters and punctuation. It, too, is potentially meaningless. Like the painting, the interpretation of the meaning described by the pattern of these letters is a fragile thing, liable to be impossible except under certain special conditions.

Meaning is granted to this arrangement of letters by the reader's command of the English in which it was written. The author's intentions are not irrelevant, but they are only a part of the story. I do have a meaning in mind when I arrange these letters, but that meaning is only conveyed to readers who share my language.

My language is not just English, but English informed by the time and place in which I live. It is not simply a set of dictionary definitions, but a set of definitions augmented by the formal and informal rules of usage in effect right now, where I sit. If I write of someone that his elevator doesn't go all the way to his top floor, I divide the speakers of English into a group that understands the meaning and a hopefully much smaller group that does not. My use of the word "hopefully" in the previous sentence further divides the world into a group that disapproves of this usage as a corruption of the adverb, and another that considers it an unremarkable part of everyday English.

But what is often overlooked is that the context in which the pattern of this book is most accurately interpreted is tremendously, absurdly, unimaginably, more complex than the book itself. As we all have private fancies with which we arm ourselves against the cruel world, I imagine I have thoughts about life that others may find interesting. But in the deepest embrace of this fancy, in the wildest ravings I've ever suffered, I've never once imagined I could alone build an intellectual edifice to rival the richness and complexity of the accumulated store of the American English language in the 21st Century. I may occasionally wander in fields of self-delusion, but I know the way home.

This book can be translated into another language, allowing people who don't speak English to share its meaning. But some translations are more challenging than others. France, for example, shares with America a great deal of intellectual tradition. The French have elevators, for one thing. Between these two languages, there are relatively few utterly untranslatable cultural references, perhaps excluding a few very current ones. Consequently, a French-English dictionary and some basic grammar is all you really need to do an adequate translation. Of course, the more sensitive you are to the nuances of each language, the better the translation, but the point is about the conveyance of meaning, not the elements of style.

In contrast, a translation into one of the many indigenous languages of the interior of New Guinea—where elevators are scarce—may be more challenging. To make a real translation into a language like Bwaidoka or Enga, I must make explicit much of what I can leave implicit to those who share my culture.

The cultural common ground on which the reader and the author stand is often slighted in the same way it is easy to overlook the air we breathe. But, like the air, close examination reveals normally unseen whorls of complexity.

So far, so obvious. So what?

The point is that somehow, this common sense analysis leaves us when we consider the complexity of DNA. Like this book, a molecule of DNA contains a pattern made up of a small number of "letters" arranged in slightly larger groups which could be considered "words" which are themselves arranged into groups fancy could call sentences or paragraphs. The whole sequence, across all the separate chromosomes, makes an organism's genome. The comparison to a written human language is nearly unavoidable, and few people—or at least few writers—avoid it.

## 2 The Language of Our Genes

One regularly reads about genes containing language; sometimes the "book of life" others would have it the "program" for the development of an organism. Other times you read about the DNA being a "blueprint" for an animal.

In a ceremony at the White House, on June 26, 2000, to celebrate the initial successes of the Human Genome Project, President Clinton said, "Today we are learning the language in which God created life."[1]

Kevin Davies, the author of a history of the Human Genome Project, the decade-long effort to "read" the sequence of the human genome, writes: "We are the first species with the intelligence to be able to read the text of life."[2] Later, he refers to DNA explicitly as a computer program: "DNA is essentially digital information, a 3-billion-year-old Fortran code."[3] Matt Ridley, the author of a popular recent book on genetics, thinks it's more than just a comparison:

> The idea of the genome as a book is not, strictly speaking, even a metaphor. It is literally true. A book is a piece of digital information written in linear, one-dimensional and one-directional form and defined by a code that transliterates a small alphabet of signs into a large lexicon of meanings through the order of their groupings. So is a genome.[4]

Any alert reader of popular genetics literature can come up with dozens of similar references. But what are we to make of this comparison? What is it that these writers mean by their comparison? Ridley goes on to elaborate the analogy:

> The filament of DNA is information, a message written in a code of chemicals, one chemical for each letter. It is almost too good to be true, but the code turns out to be written in a way that we can understand.[5]

As Ridley has said, this is not meant as an analogy. We are meant to understand this as a scientific fact. Ridley even includes a short history of information theory to make the point. None of this is meant to be his own contribution; he is simply representing the consensus view of the subject, shared by every writer who has let slip the unqualifed metaphor.

A characteristic of scientific facts is that we can infer conclusions from them. What kind of conclusions do we infer from the equation between DNA and information?

Here's a passage from a recent introductory genetics text:

> Our bodies contain billions of cells. In each of those cells is a nucleus that contains all the information required to make a complete human being. The information exists in the form of 50,000 to 100,000 structures called *genes*. Each gene possesses the ability to encode one *protein*...[6]

The authors here have concluded from somewhere that *all* the information necessary to make an organism is contained in the DNA. They present it as uncontested; this is the opening paragraph of the book. One might complain that this was not inferred from the metaphor, but from data about DNA. But in fact, the assertion that one could duplicate some animal relying solely on the information in its DNA is nothing more than speculation, unsupported by data. No one has ever proven it to be true, and there exists a wealth of data to contradict it, much of which has been available since the dawn of the age of the double helix. So where did this idea come from that *all* the information needed to create an organism is in the "language" of DNA?

It could have come from Nobel Laureates. Here's what Walter Gilbert, who recieved the Nobel Prize in 1980 for his work in uncovering the function of DNA, said back at the inception of the Human Genome Project:

> Three billion bases of sequence can be put on a single compact disc (CD), and one will be able to pull a CD out of one's pocket and say, 'Here is a human being; it's me!'[7]

Though few molecular biologists today will admit to having shared in it, this kind of hubris was not at all uncommon in the recent past. In an interview during the 1970's, Jacques Monod, who helped discover the regulatory function of DNA said:

> The secret of life? But this is in large part known—in principle if not in all details. For a

> simple living creature to be synthesized, in my opinion, there is no further principle that would need to be discovered.[8]

The Human Genome Project, recently past the first hurdle of completing the human DNA sequence, has put a merciful end to this sort of talk. This was, of course, not the intent of the planners, who, like Gilbert, apparently expected to see the inner workings of life itself unfolded on the computer screens in front of them. But as with any field, the more you learn, the more you realize how much is left to learn. We've learned a lot from the HGP, and a lot of it is about the dimensions of our ignorance. We've learned, for example, that there are many fewer genes encoded in our DNA than we'd thought, that rearrangements of genetic elements play a much more important role than previously thought, and that duplication of genetic elements is far more common in humans than in yeast, fly, or worm genomes.[9] In other words, we've learned a lot of facts whose proper interpretation is far from obvious. That is, we've learned that we are not about to find the Holy Grail, and that though it is possible we may now be in the same fairy land as the Grail, it is a far larger domain than we once could have imagined. (With a more spiteful Faerie Queen, too.) And so, scientists have learned to use a new spice in their gene talk: humility.

One of the two papers presenting the completion of the first sequencing of the human genome contained this admission:

> The modest number of human genes means that we must look elsewhere for the mechanisms that generate the complexities inherent in human development.[10]

So if the idea that all our genetic information is in our DNA didn't come from the data, where did it come from?

## 3 Whose Idea Was This?

The root of modern molecular genetics grew from the 1944 findings by Avery, MacLeod, and McCarty and the 1952 experiment of Hershey and Chase (sometimes called the "Waring Blender experiment" because it involved using a blender to purée bacteria cultures) that demonstrated that DNA is genetic material. Shortly after that, Watson and Crick determined the structure of DNA, and over the course of the next decade it was demonstrated that: DNA contains descriptions of protein sequences; it is used to create RNA as an intermediary; gene expression is regulated by proteins that interact with regulatory sites on the DNA; and other varieties of RNA, along with ribosomes, are what catalyze the final translation to proteins. All these parts of the basic, high-school-biology picture of DNA, along with the interpretation of the protein-encoding code were in place by the middle of the 1960's. Together it all makes a consistent and appealing picture of how development works, including how genetic information is transferred from one generation to another.

In their landmark 1961 paper describing the apparatus by which gene expression is regulated, François Jacob and Jacques Monod suggested that the combination of protein encodings and regulatory apparatus would make an organism's DNA comparable to a "program" which would produce a cell or an organism, when executed. Evelyn Fox Keller writes that this was the first use of the idea in the scientific literature.[11] Though talk of the "code" was by then common, this was the first serious theoretical hypothesis about it. Keller points out that because the regulatory agents (proteins) are also genetic products, it was easy to conclude that the whole is a self-contained process, with the regulatory products interspersed with the regulators themselves. Though the protein code was not known at the time, its elucidation was imminent, and it was easy to make the leap into the realm of the abstract.

Most of the early molecular geneticists avoided much theorizing beyond the biochemical domain in which they worked. The great advances of the 1950's and 1960's were almost all framed as problems in biochemistry: not "How does inheritance work?", but "What are the steps in protein synthesis?" Not "What is the language of life?", but "How does DNA code for proteins?" The landmark achievements were accomplished with years of painstaking experiment. What few purely theoretical efforts had been made had largely led to dead ends, like George Gamow's various code attempts, and Francis Crick's development of the comma-free code, an elegant, but mistaken, hypothesis.[12] Indeed, several of the vanguard rebuffed attempts at assistance from mathematicians and information theorists.[13] But what the first generation found easy to resist, the next generation found irresistible, and though the first generation resisted, their casual use of language encouraged those who followed. In several papers, published and unpublished, and in correspondence, the leading lights of molecular biology spoke about "codes," "cybernetics," and "information." For example, in the follow-up to their famous paper elucidating the structure of DNA, Watson and Crick said:

> The phosphate-sugar backbone of our model is completely regular, but any sequence of the pairs of bases can fit into the structure. It follows that in a long molecule many differ-

> ent permutations are possible, and it therefore seems likely that the precise sequence of the bases is the code which carries the genetical information.[14]

Abstracted as a language or a code, genetic "information" can exist in its own abstract domain, available for theoretical analysis on its own terms, and dissociated from the dull considerations of mere "context"— energy budgets, coenzymes, and the like. But in fact, this context in which DNA exists—the nucleus of a cell—is itself filled with implicit information and information-bearers. Considering genetic information apart from its context, and expecting to be able to do anything with it, is an error roughly comparable to imagining that a dictionary would be all one would need for a useful translation of Huckleberry Finn into Bwaidoka.

To translate that book, or this one, into a culture unlike ours, you have to add what would be understood here. In a place where elevators are scarce you might have to add information about elevators. To readers who won't know about cells, you need to add something about them. To translate the DNA into a human-readable form, you have to add information about the context in which the DNA is to be read. You have to make *explicit* information that is *implicit* in the construction of the cell from which the DNA came.

But what information is that?

## 4 What A Difference Context Makes

First, to clear up a persistent myth, there is nothing "autocatalytic" or "self-reproducing" about DNA. DNA alone in a test tube stays alone. It doesn't make copies of itself, and it doesn't make anything else either. It just sits there. Actually, DNA is about the most inert biochemical substance around, which is why people can find useful scrapings of it from ancient mummies, old bones, or crime scenes. It doesn't degrade easily. In order to get the DNA to do anything, like reproduce itself, you need to add some enzymes that do the work, and some more nucleotide bases for them to work with. You need helicase to unwind the double helix, an assortment of binding proteins to keep the strands from winding back up together, DNA polymerase to make the copy on one side, and DNA primase, DNA ligase, and DNA polymerase to make the copy on the other side. Each of these enzymes requires a handful of helper enzymes and coenzymes to do its work.

In other words, to get our DNA to replicate, we've had to add information to the test tube, in the form of specific quantities of enzymes, coenzymes, nucleic acid bases, and sugars: a recipe *not* contained in the DNA itself. This information is present in the nucleus of a cell, but in an implicit form, and therefore not easily recognizable as information.

What else can we make with our DNA strand? Well, perhaps it codes for a protein. We could try to make that protein. First, of course, we have to transcribe the DNA onto a messenger-RNA (mRNA) template, which would then be used to make the protein. This process is roughly comparable to replicating the DNA. You need to make sure there are enough bases, and you need an enzyme called RNA polymerase, along with an assortment of eight or ten different "general transcription factors," molecules whose jobs vary from marking the start point of the transcription to providing the energy the polymerase needs to do its job. A new recipe; more information

Now we have an RNA template we can use to make a protein. But before we can make anything useful out of it, we have to edit it. DNA consists of protein-encoding parts called "exons" interspersed with parts whose function is still mysterious: "introns." There are enzymes whose job is to remove the introns, and splice the exons together. These are the "spliceosomes." Now the mRNA we've edited must be applied to some ribosomes, supplied with amino acids, dozens more enzymes, and a supply of transfer RNA (tRNA) to translate the base code of the mRNA into the protein's peptide chain, a linked train of amino acids.

At each step of the way, the recipe for the new protein becomes a little more intricate and a little more subtle. Subtle enough that as of this writing, the last couple of steps in this chain are not so readily reduced to information. Scientists can make short protein-like chains, linking together short chains of amino acids, but have difficulty reliably making many real proteins, which often contain hundreds of amino acids, and may contain multiple chains. To turn DNA into protein, biologists typically enlist the aid of some living organism: a native speaker. To make a protein, they splice the DNA into the genes of some bacteria, and let it do the dirty work. We know more or less how it all works, but somehow the detail—the translation into a language we understand—remains elusive. Even the steps we can master *in vitro* are generally much more easily done with the help of engineered bacteria.

Much talk about genetic information seems to depend on the image of the cell as little more than a growing medium for "information." But cells contain a sophisticated mixture of molecular components necessary to express DNA—as well as the mechanisms for creating and maintaining those components—and characterizing them as neutral sites, where DNA expression "happens" hardly seems right.

The mixture is actually even more sophisticated than this outline implies. At this point, we've created a protein from a strand of DNA. We've recreated what Francis Crick dubbed the "central dogma of biology": DNA makes RNA makes proteins. But he formulated this dogma in the 1950's. What have we learned since about how these processes really work inside growing cells? What other details are there to cloud the crystalline information of the DNA strand?

## 5 More Complications

Actually, the editing step, where spliceosomes separate mRNA exons from introns, wasn't apparent in the 1950's. But let's look at it for a moment. It's not yet entirely clear how the enzymes that do the editing do what they do, but it appears that the simple picture of cutting out the introns isn't quite adequate, either. It seems instead that these enzymes play a sophisticated role in constructing proteins out of assembled pieces, irregardless of the order in which they appeared in the DNA sequence. With the help of these enzymes, one stetch of DNA can be responsible for hundreds of different proteins.

A recent study of chick development estimated that a single gene can make over 500 different proteins for "tuning" the hairs in the chick's cochlea.[15] In a cochlea, part of the inner ear, each hair is sensitive to a different range of acoustic vibrations, and the sensitivity appears to be largely determined by the form of the protein produced by the chick's *cSLO* gene. For normal hearing, you want your cochlea to contain a collection of hairs sensitive to a wide variety of ranges, and that's apparently what normal development produces. However, there is no good explanation of why any particular cell chooses one *cSLO* variant over another. Perhaps it is just random; a chicken might be able to hear normally if all the cochlear hairs are tuned randomly, but we don't know that. If it's not random, what is the mechanism by which these hair cells organize themselves? No one knows. One possibility is that the cells actually respond to the sound waves that hit the cochlea, that the tuning of the ear is done *after* the ear is constructed. But this is just speculation.

A very similar mechanism works in our immune system, which needs variety to work. We need lots of different kinds of antibodies to throw at whatever antigens appear in our blood. Antibody variety is created by enzymes in your marrow who maniacally reorder the genes in a certain stretches of immune-system DNA to create an astronomically large number of possible antibodies.[16] Lymphocytes armed with these new proteins go out into battle against these intruders. The successful ones—that find and engage the enemy—reproduce, which is how your immune responses become sensitized over time.

Even the "just-sitting-there" stage of DNA turns out to be more intricate than originally thought. DNA is inert, but it's not perfectly so, nor is it ultra-stable, like a crystal. It decays and falls apart from time to time, and copies are often imperfect. To fulfill its function as genetic material requires an assortment of repair genes to mend tears, excise ruined sections, edit poor copies and undo occasional mutations. Much of DNA's stability, then, is not static, but a kind of dynamic balance maintained by a crew of molecular repair technicians.

These repair technicians do more than just repair things. Sometimes, they deliberately screw them up. It appears, for example, that you can artifically breed bacteria to improve the fidelity of their DNA copies. Naturally occurring bacteria, the product of billions of years of evolution, haven't done so themselves, so the implication is that there may be some adaptive value in messiness. This may be simply because of an energy trade-off between fidelity and efficiency (that is, higher fidelity requires more energy, and if you've got repair technicians around, why bother?), or it may be that there is value in leaving the field open to occasional mutations. There is evidence that some bacteria produce messy versions of DNA polymerase when they are under environmental stress. These enzymes, dubbed "mutases," act like normal DNA polymerase, but the resulting copies have more mistakes than copies made with regular polymerase. That is, they seem to promote mutation among the bacterial offspring.

Promoting mutation may not, in fact, be a bad survival strategy: if you're having trouble, at least there's a chance that your daughters may have a better time of it. There even exist clues that the induced mutations may not be entirely random, but at this writing, what evidence there is remains inchoate.[17]

Human genomes contain codes for enzymes that resemble these mutases, though no one now knows what we do with them. It does appear that some lymphocytes, the agents of our immune system, use directed mutations, "somatic hypermutation," to fine-tune the immune response to antigens.[18] The mutases may be part of that system, or they may be used in some other way entirely.

Barbara McClintock noted, back in the 1940's, that attentive analysis of classical (non-molecular) genetics results show that genes have some interesting ways of dealing with environmental stress. Among many other effects, she pointed out that chromosomes can combine and re-combine, and that genetic elements are capable of simply moving from one chromosome to another, changing their effects with their position. The mechanisms for detecting stress, and the procedures by which the genome is rearranged, are not yet understood, but some of the effects are quite clear.[19]

The transcription step, making RNA from DNA, also has its intricacies. Regulatory sites have been known about since Jacob and Monod, but the picture has become successively more interesting. The history of what is known about a single gene is worth retelling. The *lac* gene, implicated in the processing of milk sugars by *E. Coli* bacteria, was discovered in 1947, and determined to be a single point on a chromosome. Fourteen years later, when Jacob and Monod published their discovery of gene expression controls, it was about the *lac* gene. With their new analytical tools, and an increased resolution view of the chromosome site, Jacob and Monod found that the gene corresponded to three different enzymes, and an operator site that controlled their expression. The operator that binds to the site only does so in the absence of lactose, and it acts to repress expression of the gene. Therefore, the gene products, used in digesting milk sugars, are only expressed in the presence of those sugars. It seemed a tidy picture. The whole complex was dubbed an "operon" to imply that it was a unit, albeit a complex one.

Thirty years later, increased scrutiny has filled out the picture. It's now known that the "operator site" found by Jacob and Monod is actually a composite of three repressor and three promoter sites that overlap each other (at least one of which overlaps the enzyme-encoding part of the gene) along with two sites activated by repressors from some entirely different gene group.[20] The picture that emerges is one of a much more sophisticated mechanism than had been expected: not a simple switch, but a collection of interdependent switches, dependent on each other and on a host of different external stimuli.

# 6 How Do The Complications Add Up?

The possibilities available with a sophisticated network of regulatory genes have not been wasted on multicellular organisms. Our genes are riddled with intricate cross-dependencies, where the product of one gene is used to regulate the product of another. The simplest of triggers can produce results of stunning complexity. For example, it is possible to manipulate a single controlling gene to turn on the formation of eyes in fruit flies. One gene controls the expression of the 2,500 other genes necessary to build an eye.[21] Genes are generally triggered by some protein made by some other gene. But here is something interesting: the regulatory proteins controlling an animal's development do not always come from the that animal. Sometimes they come from a parent. And sometimes they come from the environment.

For example, DNA does not encode the location of a fruit fly's head or tail. The location of the embryonic cells destined to become the head or the tail of a fly is determined by an asymmetry to the egg cell engineered by its mother, and operating on what are called "homeobox," or *HOX*, genes. In flies and other invertebrates, there are "nurse cells" to do this, sitting next to the egg developing within the mother. The nurse cells inject some RNA into the egg cell, and the RNA makes proteins and those proteins bind onto various regulatory regions of the HOX genes in the developing embryo's DNA, and determine how it is to be read, and therefore how regions of the egg differentiate into various body parts. The construction of the nurse cells is, of course, partly determined by DNA, but these nurse cells will help create the sons and daughters of the animal that contains them, not the animal itself. Fruit fly eggs require 15 different nurse cells apiece to add information not already present in the egg's own DNA.[22]

In vertebrates, there are similar mechanisms to do similar things, though the processes seem more subtle. "Follicle cells" in humans surround the developing egg, adding protein layers and injecting information to be used by the developing embryo. Whatever the process, the result is the egg: a complicated little device, on whose structure an animal's development depends. The "structure" in this case is the variety of building blocks an egg contains — RNA, proteins, amino acids, nutrients, and the remaining undiscovered what-have-you — and the relative concentrations of these chemicals in different microscopic regions of that egg. We can think of cell interiors as a soup of cytoplasm, but it's a soup where the noodles and vegetables are all arranged in some pattern and it matters which noodle is where.

And since it's the real world we're talking about, the complications don't end there. The part of the *fertilized and already developing* fruit fly egg cell that will become the underbelly of the fly is selected by maternal cells next to the embryo.[23] These cells release a signalling protein that diffuses across the embryo, and the part with the highest concentration of that protein becomes the new fly's belly. A developing egg doesn't just get nutrition from its mother; it also gets important directions that control its development. There is no "switch" turned on at the moment of fertilization, after which development proceeds autonomously. Fertilization is simply one step among many in the interaction between genes and their environment that eventually produce a new organism. The unsettling part—the aspect that mocks attempts to analyze separate steps of the process—is that in the early stages of development, half of the genes controlling the environment are the same as the ones developing in the egg. The genes control development, but they do it *both* from within the egg, and from within the parent. The egg needs its genes to develop, but it also needs its mother.

These unexpected features of DNA—mutases, *HOX* genes, spliceosomes—barely touch the surface of what is known about how cells and DNA behave. We have not mentioned transposable DNA elements, nor horizontal gene transfer, nor the commonly-found overlap between functional units, nor the existence of multiple genetic codes. The framework of the central dogma can hardly support the weight of all its emendations, but there isn't a successor dogma in sight. In a survey of the quandary of molecular genetics, James Shapiro wrote:

> Four decades of dissecting genome function at the molecular level have brought many insights that were not anticipated in 1953. Two of the most far reaching are: 1. Many different genetic codes exist in addition to the triplet code for amino acids. These codes affect many diverse aspects of genome function,, such as replication, transcription, recombination, DNA packaging and chromatin organization, imprinting, RNA and protein processing, and chromosome localization, pairing, and movement. 2. There do *not* exist fundamental genomic units larger than the individual codons in the various functional codes.[emphasis in original][24]

That is, it is not useful to talk about the "letters" of DNA making up "words," since many of the words overlap, or are interpreted in different ways by different mechanisms, or if they appear in different places. Under these circumstances, words such as "operon," "coding region" and even "gene" do not indicate specific functional units of DNA, but rather overlapping and shifting assemblies of base pairs. They are, that is, purely conceptual entities, abstractions that help us organize our thoughts about genetics, but without corresponding directly to any underlying physical reality. They are, that is, metaphors.

## 7 Who Needs Genes?

One might object that the implicit information mentioned here so far is qualitatively different from genetic information simply because only genetic information is stable across generations; it can be inherited. But this is not so, either. Implicit, or "epigenetic" information can indeed be heritable, through a variety of mechanisms. Some forms of epigenetic inheritance are unarguable and unproblematic. Nerve cells beget other nerve cells and kidney cells beget other kidney cells, even though they both have precisely the same set of DNA. Whatever makes a nerve cell into a nerve cell can indeed be passed along to its offspring, and no one thinks that impossible, though the details remain mysterious.

Well before the discovery of the nature of DNA, Max Delbrück wrote in 1948 that biological homeostasis was a possible form of inheritance, though he added that it would work in cooperation with some kind of "definite series of genes."[25] According to his idea, the interaction between several different interdependent chemical reactions could produce many different stable states. Each state would be stable enough to propagate forward through generations of organisms.

Some years later, the discovery of gene regulation made possible such a system, and it wasn't many years more before one such system was observed in bacteria, in the *lac* gene. Part of the *lac* gene codes for an enzyme that allows lactose through the cell wall, a "permease." In a bath containing a low concentration of lactose, some bacteria will, just by random fluctuation, happen to switch on the *lac* gene, and make this enzyme. Once the enzyme is made, lactose comes into the cell, disables the repressor, and keeps the gene operating. When the cell divides, both children will continue to let lactose in, and digest it, as will their children and so on. The cells that don't switch on the gene will make no permease, and let in no lactose, and neither will their children.[26]

This sort of inheritance is not as stable as inheritance through the genes; researchers observe that they can last for hundreds of generations, but tend eventually to decay. But this is inheritance nonetheless, and of an acquired character to boot, albeit one corresponding to a genetic predisposition.

There are also forms of structural inheritance, some of which have been observed in experiments with single-celled *Paramecium*. It is possible, through micro-surgery, to create physical modifications of *Paramecia* that are transmitted to the succeeding generations, showing both that acquired traits can be passed along, *and* that animals with identical genes need not have identical bodies.[27] It is not clear how relevant these effects are to multi-cellular organisms, but, to debunk yet another myth, cells do not build themselves from scratch. They divide, producing new cells with some of the same materials from the parent. That is, part of the reason a nerve cell may beget another nerve cell is the complex of genes that happen to be repressed or activated at the moment of division, and another part may simply be that the parent *was* a nerve cell.[28]

Another form of epigenetic inheritance is regularly seen in sex-linked traits. Placental mammal genomes contain a variety of genes that are typically "imprinted." That is, their activity seems to depend on whether the gene was inherited from the mother or the father. [29] Much of the development of the placenta and the fetus seem to proceed from genes imprinted with their parental origin.

DNA comes in chromosomes in the nucleus of each cell. To make a chromosome, you coil the DNA up tight, and wrap it around little protein plugs. The combination is called "chromatin." In addition to the kinds of activity recorded on the DNA—repressors and activators and the like—chromatin can apparently receive different kinds of "marks." One well known mark involves additional chemical groups that may be attached to some DNA bases, acting to suppress the activity of some regions of the DNA. The additional chemical group is often a methyl group, a carbon atom surrounded by three hydrogen atoms, and scientists refer to some stretch of DNA as "methylated" if it is disabled in this way. Sex-linked imprinting is thought to work this way.

Through the action of the enzyme "methyltransferase," methylation appears to be generally conserved across replication of DNA, though at lower fidelity than the DNA base pattern itself. That is, methylated patterns on a DNA copy will correspond to similar patterns on the original, providing an easy way for the suppression of a gene in a parent to be communicated to daughter cells. It appears that much of the methylation pattern on a gene is stripped off during the early development of eggs, but apparently not all, and it may be that patterns of methylation are behind the sex-linked imprinting of genes.

So not only is there implicit information available in the context in which DNA exists, but it is uncontroversial that this context can be modified to carry information from one generation to the next. There may, of course, be many more methods awaiting discovery than are outlined here. Genetic inheritance is still, of course, predominantly due to the effects of DNA and its various coding regimes, but there is sophistication enough in the context of the cell to carry a substantial amount of information from one generation to the next. A few specific mechanisms of this sort have now been identified in complex animals, such as mice,[30] as well as in plants,[31] and there is no principled reason why many more shouldn't exist.

One difficulty with the information and the information-carriers outlined here is that their information does not reside in "patterns" instantly comprehensible to some external observer. The information, such as it is, is carried by systems of activity, relative concentrations of chemicals in solution, and the complex shape of biological actors. These forms of information are not readily reducible to the kind of information easily analyzed by classic information theory. That theory was developed for a theoretical framework in which to analyze the sending and receiving of a message. But that is simply not the way a cell works.[32]

Erwin Chargaff, one of the pioneers of molecular biology, put it this way, in 1968:

> In the living tissue, many events take place simultaneously; precursors, intermediates, and end products are formed in close propinquity; everything happens on top of each other, apparently without getting into each other's way; proteins and nucleic acids, lipids and polysaccharides are assembled and deposited where they belong: all presumably under the supervision of the genome which at the same time is quite busy reproducing itself.[33]

We usually think of the message of DNA as being read by the cell, or by a nucleus. But our knowledge of the process is at once more detailed than that, and more confusing. If we think of a reader as an instance of the reading apparatus—the enzymes, the messenger RNA, the ribosomes—that makes DNA into a protein, then the message encoded in DNA is read by hundreds or thousands of different entities at the same time. These entities appear, do a little reading, interact with other entities engaged in the same kind of work, and disappear. What they certainly don't do is start from the "beginning" and work their way to the "end," as one might read a novel. Or a popular book on genetics.

One group of researchers has proposed that DNA is more aptly compared to data to be operated on by a program of a massively parallel architecture, "embedded in the global geometrical and biochemical structure of the cell." But the authors go on to say that even this comparison is marred by the complexity and interdependence of cell processes, not to mention the impoverishment of currently available parallel processing paradigms.[34]

## 8 We've Known This All Along

Of all this, little is really news. A randomly chosen recent issue of *Nature* will probably contain some equally devastating counter-argument to simplistic claims about

DNA information. Nor are these all recent findings. The history of molecular biology contains a long list of dissenters from the DNA orthodoxy. Barbara McClintock, for one, saw early that there were regulatory functions to DNA, and that the picture was considerably cloudier than the one painted by the mainstream[35]; David Nanney presented findings about non-genetic inheritance in *Paramecia* at some of the same conferences at which DNA function was being hammered out[36]; Robert Rosen pointed out logical inconsistencies in the use of information in the context of genetics in papers published in the early 1960's[37]; J. Chayen, Alfred Hershey and several others insisted, during the 1950's, that the available evidence showed only that DNA could carry genetic information, not that it was the only such possibility[38]; D. Wilkie published an extensive monograph in 1964 outlining all the evidence that cytoplasm could influence heredity[39]; and the British embryologist C.H. Waddington, throughout his career, provided both examples of development unexplained by the standard interpretation[40], and a compelling synthesis demonstrating how little help conceptions of "information" could be in solving the real problems of biology.[41] There were many others.

As the field progressed, and the standard interpretation became elaborated, many, though far from all, of these objections eventually came to be explained by known features of gene function. Some, of course, took longer than others. McClintock's deduced the existence of genetic "transposable elements," which are now well known and commonly referred to as "transposons." Many of Wilkie's and Nanney's observations have been explained by the discovery that cellular elements (such as mitochondria) can have their own DNA, and therefore their own genetic effects. But many of these observations remain untouched by the current understanding: the organization of McClintock's transposons, and the genomic reactions to environmental stress; the inheritance of acquired traits noted by Nanney; and the "canalization" described by Waddington all remain firmly beyond the explanatory ability of current genetics.

Though data was available at the time that can now be seen as troubling to the central dogma, it was ignored by those eager to embrace the new, simplifying, theory of the role of DNA. But even the objectors could see the appeal of the new ideas, as well as the lay of the land. Alfred Hershey, whose own experiments with Martha Chase had done so much to promote the idea that DNA carries genetic information, said in 1953 that none of the available evidence "forms a sufficient basis for scientific judgement concerning the genetic function of DNA."[42] He followed up with these words:

> The evidence for this statement is that biologists (all of whom, being human, have an opinion) are about equally divided pro and con. My own guess is that DNA will not prove to be a unique determiner of genetic specificity, but that contributions to the question will be made in the near future only by persons willing to entertain the contrary view.[43]

Fifteen years later, Chargaff put it this way: "DNA, a hundred years ago a humble molecule in Miescher's hands, has been hypostasized into one of the symbols of the ever-increasing divorce from reality that characterize our living and thinking."[44] That is, thirty-three years ago, he was already tired of hearing about organisms being nothing more than abstract "information."

## 9 The Meaning: Mysterious Proteins

If DNA is a language, it is a language where the word meanings depend on their location, where words overlap one another, and where the meaning of words depends on long strings of modifiers that may not happen to be near the modified word. But human minds have decoded messy codes before. Perhaps we will make progress on this one, too.

Science has learned how to tease out some of the meanings in the language of your genes. We can examine some of the proteins certain genes build. We can correlate some genetic changes with changes in an organism. But is this the same as decoding the language? Suppose you spoke no English, and could not read this book. You show the book to a friend who can read English, and ask her opinion of it. Now you start changing single words here and there and asking her to reread the whole thing and tell you about the changes. Some changes she might not notice, while others would be more obvious. Some words are important to the sense of the book, but some aren't. (Like this one: "purple.")

When you change a word and it has an effect on your friend, you do know something, but maybe not much. For one thing, you have no way really to know whether the word you've removed was the important one, or perhaps it was the one you've added. You could change words in parentheses to anything you want, and not affect the overall meaning. What could you conclude from that? I could even write sentences that only make sense with nonsense words: "'Bfstplk' is unpronouncable." What will you learn when you experiment with substitutes for *that* word?

Faced with the task of translating a book in a foreign language, finding the letters and the words is the easy part. The hard part is deciphering the meaning behind the words. And it's even harder when the only method at hand is to substitute other words and see what happens. But that's roughly the way scientists are proceeding to "decode" the meaning of our genes. You might find the protein produced by some stretch of DNA, but proteins don't come with labels for their use. They are big ugly complicated molecules, and the only way to figure out their purpose is somehow to catch them in action. Even then, you don't always know if the function you've observed is the only thing they're good for. People spend entire careers studying a small handful of proteins, and there are hundreds of thousands of them to figure out.

To make matters worse, it turns out that many genes are redundant, or don't seem to matter that much. That is, when scientists try to figure out what is going on by knocking out or mutating some gene, they often find that there is no effect at all. (Or that the effect is not detectable in the adult animal.) It may be that many genes are actually redundant, but it may also be that a lot of genes may not code for a specific thing, and work in combination with many others, creating a kind of effective redundancy. That is, there may be no other source of some protein that a mutation has modified, but that protein may itself not be a particularly crucial ingredient, it may only change things slightly, too subtly to be detected by our experimental design, or it may be readily replaced by some other gene product. Some of these problems may also be due to interference by the DNA-restoring mechanism of the cell.

This is not to belittle the enormous work already done in reading the genetic code, nor the tremendous medical advances made possible by that work. But the fact remains that in the attempt to understand how organisms make themselves, "reading" the genome—figuring out what parts code for what proteins—is but a very small first step. When we know what protein is made by every single part of your DNA, *and* what each one is for, we still won't have the slightest idea how to make another one of you.

## 10 We Live In Our Environment, But It Doesn't Contain Us

We think of an organism as a *thing* of its own. A thing has boundaries to separate it from the rest of the world, and it has a beginning to its existence, and (generally) an end. We want a developing animal to start from a clean slate, read its own instructions and build itself, but that's just not the way it works. Eggs inherit information from their genes (from the nucleus *and* the mitochondria), but also directly from their mothers, in the egg's cytoplasm, and in guidance to the process of development.

The genes guide the embryo's development, and they also moderate the actions of the parents, so they do ultimately control most of the process, but they do it through this strange symbiosis between generations. Looking for a "beginning to life"—trying to pinpoint the moment when a developing organism becomes an independent thing—is an exercise in futility. The beginning of life was billions of years ago. For now, life just is. I am a continuous extension of my mother's life, as she was of her mother's and so on and on right back to the blue-green algae mats floating in shallow Pre-Cambrian seas.[45]

Spatial boundaries have a similar problem. We want a thing to exist, contained in its environment, as a brick might sit in an empty room. But the environment is not a simple container, and no living thing is as docile as a brick. Living organisms interact with their surroundings, often changing them in profound ways by the very act of living. A rabbit eats the grass in its fields; a beaver cuts down the trees in its woods. Some changes are minor, but some are dramatic: a virus might kill its host; Earth is habitable for us only because billions of years of plant photosynthesis have produced enough oxygen for us to breathe.

Life is not a Platonic Form, existing in some airy theoretical vacuum. Life is a process carried out in the constant dialogue between an organism and its habitat (which includes other organisms). Change one and you change the other.

A gene can't be read unless it's in the right sort of nucleus, in the right sort of cell, in the right sort of organism, which itself needs to be in the right sort of environment to thrive. Not only is there a continuity in time, as we trace our beginnings right back to the beginning of life, but there is a continuity in space, as we trace the chain of dependence out from the gene, like spreading ripples on a pond.

Genes did not evolve to satisfy the dictates of Platonic categories, so we shouldn't be surprised to find unexpected relationships between genes and their surroundings. Genes evolved by chance and natural selection. That is, they evolved to work, and nothing more. They work fabulously well, and have changed the planet by so doing, but the structure they have created is riddled with subtle and delicate interdependences, between different parts of the gene, between different cells, and between different organisms.

The nucleus incorporates the gene, and depends on it.

Our environment incorporates us, and depends on us as well. Just as it makes little sense to attempt to search for the moment when life begins, it makes little sense to try to draw lines to separate ourselves from our environment; the threads of interdependence are far too tangled.

Unfortunately, the philosophical trend since Plato is to see ourselves apart from the rest of the world, as a separate (and exalted) thing. But the evidence for that view has always been thin, obtained largely by philosophical navel-gazing rather than accumulation of data. The past century's close scientific observation of how life works has shown this separation is nothing more than a vain illusion.

In a painting, one speaks of a relationship between the "figure" and the "ground," between the subject of the painting, and the space around it. The figure defines the ground, and the ground defines the figure. It makes no sense to talk of one without the context defined by the other, and a successful painting is one where the relationship between the two elements adds to the overall effect.

The study of genes has us on the beginning of an exciting road. But each new step in understanding how life carries on will reinforce this point: it is as impossible to separate any life from its context as it is to analyze a painting by only looking at the figure. It is pointless to analyze a developing embryo without considering the mother surrounding it, or to talk of that mother as separate from the water she drinks and the air she breathes.

We also need to take care. In a painting, it is impossible to make changes to the figure without also changing the ground. When we attempt to engineer life itself, we are tinkering with a system about which we know very little; we can *do* far more than we can *understand.* But we understand this: when you let a little piece of genetic material out into the world, you effect a permanent, possibly cascading, change in our world. For example, we know now that many of our genes were foisted upon us by bacterial invaders, through mechanisms as yet unknown.[46] There is no empirical reason for confidence in bland assurances that the release of genetically engineered animals into the world are minor, or transient, changes. To be sure, reckless modification of our environment predates genetic engineering; ask anyone who's been attacked by the "killer bees" of South America. But the potential for mischief is now greater than ever. One hopes that the newfound humility of the geneticists will prove contagious. We could all benefit.

In the years to come, we will make some progress disentangling the threads of interdependence between organisms and their environment. We will understand much better how life works. But, possibly aside from screening and preventing simple genetic dysfunctions, it will be a long time—possibly forever—before we are able to engineer our children to have desirable behavioral traits, which are fantastically complicated results of the interplay of genes and environment. On the other hand, as we learn more about this interplay, we will come to understand, more deeply than ever before, how important the world around us is to our own lives, and how delicate the balance that maintains us. One hopes that the result of the advances of genetics will be an increasing awareness of the dimensions of our ignorance, and that the popular understanding of this fascinating science will shift from wonderment at its advances to a real appreciation of the the vastly greater complexity—and the precariously teetering structure—of the context in which those genes thrive: nucleus, cell, organism, and environment.

# Notes

[1]Quoted in Davies, 2001, page 6

[2]Davies, 2001, page 7

[3]Davies, 2001, page 9. To his credit, Davies also points out some of the shortcomings of these metaphors, and proposes a comparison to the periodic table of the elements in their stead.

[4]Ridley, 1999, p. 7

[5]Ridley, 1999, p. 13

[6]Hawley & Mori, 1999, p.3

[7]Gilbert, 1992. In Gilbert's defense, the entire article quoted is much more measured than this quote would imply. He does acknowledge there is a lot of work to be done after the sequencing is complete. But his excited rhetoric leaves one with the feeling that this equivocation is only slightly more than a formality.

[8]Quoted in Judson, 1979, p.216

[9]See IHGSC, 2001, p.860, or Venter *et al.*, 2001

[10]Venter *et al.*, 2001, p.1346

[11]Keller, 2000, p. 80. She also credits the great evolutionary biologist Ernst Mayr with coming up with the idea around the same time, presenting it in a contemporaneous, but unpublished, talk.

[12]

[13]See, for example, Judson, 1979, p. 244, or Kay, 2000, p.128ff.

[14]Watson & Crick, 1953, p.966. See also Ephrussi *et al.*, 1953

[15]The actual number is at least 576, if you must know. See Black, 1998

[16]Tonegawa, 1983

[17]Radman *et al.*, 1999 contains a review of evidence for, and possible mechanisms of "adaptive mutability," and Radman, 1999 describes some of the known mutases.

[18]Delves & Roitt, 2000, p.45

[19]McClintock, 1984

[20]Shapiro, 1999, p.24 contains a summary and an entertaining figure describing the progress of understanding the *lac* operon.

[21]Halder *et al.*, 1995. Gehring, 1998 contains quite a bit of context for the discovery, as well as further findings about homeobox genes.

[22]Coen, 1999, p.147ff.

[23]Coen, 1999, p.250ff

[24]Shapiro, 1999, p.24

[25]Delbruck, 1949

[26]Jablonka & Lamb, 1995, p.82ff

[27]Nanney, 1985, Jablonka & Lamb, 1995, page 86

[28]Nanney, 1985

[29]See, for example, Ridley, 1999, pp.209–218, Jablonka & Lamb, 1995, pp.112–118

[30]Morgan *et al.*, 1999, Sutherland *et al.*, 2000

[31]Cubas *et al.*, 1999, Scheid & Paszkowski, 2000

[32]There is an aside to be made here about the inadequacy of classic "information theory" to DNA. Yes, the sequence of DNA base pairs can be construed as information. However, the classic definition of information leaves out detailed analysis of *who* it is that understands that information. Information is not universally accessible. Given the long and fascinating history of that culture, I regret that I cannot read Chinese. Nonetheless, despite my regret, a page of written Chinese is of no use to me. The way that information theory acknowledges this is to say that we append to the information the program that will allow our universal Turing machine to understand the message. But this presupposes that the system in question can be adequately modeled by such a machine, which is, in the case of a living cell, a highly suspect claim.

Such things as the recipe for cytoplasm and the timing of RNA injection into a developing cell are information in a certain sense, but it is a sense that is unavailable for capture by the classic paradigm of information theory.

[33]Chargaff, 1968, p. 327

[34]Atlan & Koppel, 1990

[35]McClintock, 1961, also see Keller, 1983, a delightful scientific biography of McClintock.

[36]Nanney, 1957, Nanney, 1985

[37]Rosen, 1960 and Rosen, 1961

[38]e.g. Chayen, 1958, Hershey, 1953

[39]Wilkie, 1964

[40]Waddington, 1953, and Waddington, 1942

[41]Waddington, 1968

[42]Hershey, 1953, p.138

[43]Ibid., p.138

[44]Chargaff, 1968 p.330

[45]The ongoing arguments about when, if ever, is the appropriate moment to deny a woman an abortion don't illuminate much, but they do illustrate this point. The reason no one can agree on a single point is that all such choices are arbitrary. The process is continuous; there are no clear intervals. One point on the continuum is as scientifically defensible as any other.

[46]IHGSC, 2001, p.860

# References

Atlan, Henri, & Koppel, Moshe. 1990. The cellular computer DNA: program or data. *Bulletin of Mathematical Biology*, **52**(3), 335–348. Suggests that DNA is better viewed as data, but that the program is somewhere else entirely. Contains a proposal for an information-theory definition of "sophistication", which is interesting.

Black, Douglas L. 1998. Splicing in the inner ear: a familiar tune, but what are the instruments? *Neuron*, **20**(February), 165–168. Proteins that "tune" hairs on chick cochlear cells are generated by recombination of the cSLO gene. The problem is that it isn't really clear who controls how the recombination is done, nor is it clear how it is decided which cell gets which recombination.

Chargaff, Erwin. 1968. What Really Is DNA? Remarks on the changing aspects of a scientific concept. *Progress in Nucleic Acid Research and Molecular Biology*, **8**, 297–333. Very realistic statement of the state of affairs circa 1968. No hype. See esp. the "Modest Glimpse of the Future" section, and the Epilogue.

Chayen, J. 1958. The Quantitative Cytochemistry Of DNA And Its Significance in Cell Physiology and Heredity. *Experimental Cell Research Supplement*, **6**, 115–131. Considers and criticizes the three most significant sources (at the time) for empirical confirmation of DNA being the sole source of genetic information.

Coen, Enrico. 1999. *The Art of Genes: How Organisms Make Themselves*. XX: Oxford University Press. Good book, with a decent and clear explanation of homeotic genes, but the efforts to extend the analogy between painting and ontogeny were eventually tiring.

Cubas, Pilar, Vincent, Coral, & Coen, Enrico. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature*, **401**(9 September), 157–161. The is the same Coen as Coen, 1999, which uses a lot of floral examples.

Davies, Kevin. 2001. *Cracking the Genome: Inside the Race to Unlock Human DNA*. XX: Free Press.

Delbruck, Max. 1949. Génétique du bactériophage. *Pages 91–103 of: Unités Biologiques Douées de Continuité Génétique*. Colloques internationaux du Centre national de la recherche scientifique, vol. 8. CNRS.

Delves, Peter J., & Roitt, Ivan M. 2000. The Immune System. *New England Journal of Medicine*, **343**(1), 37–49. Review article. First of two parts.

Ephrussi, B.U., Leopold, U., Watson, James D., & Weigle, J.J. 1953. Terminology in Bacterial Genetics. *Nature*, **171**(18 April), 701. A short note proposing to clear up some misunderstandings by adopting the word "information" when talking about bacteria genetics.

Gehring, Walter J. 1998. *Master Control Genes in Development and Evolution: The Homeobox Story*. New Haven, CT: Yale University Press.

Gilbert, Walter. 1992. Visions of the Grail. *Pages 83–97 of:* Kevles, D.J., & Hood, L. (eds), *The Code of Codes*. Harvard University Press.

Halder, Georg, Callaerts, Patrick, & Gehring, Walter J. 1995. Induction of ectopic eyes by targeted expression of the *eyeless* gene in *Drosophila*. *Science*, **267**(24 March), 1788–1792. This is the paper where they induced the growth of eyes on the legs and antenna of fruit flies, by switching on a control gene. It was also found that a homologous mouse gene would work the same function on the fruit fly.

Hawley, R. Scott, & Mori, Catherine A. 1999. *The Human Genome: A User's Guide*. New York: Academic Press. A triumphalist version of the current state of genetics. Meant to be either an introductory text or a deeper popular account.

Hershey, A.D. 1953. Functional differentiation within particles of bacteriophage T2. *Cold Spring Harbor Symp. Quant. Biol.*, **18**, 135–139. A review of what was known about DNA at the time of writing.

IHGSC. 2001. Initial sequencing and analysis of the human genome. *Nature*, **409**(15 February), 860–921.

Jablonka, Eva, & Lamb, Marion J. 1995. *Epigenetic Inheritance and Evolution: The Lamarckian Dimension*. Oxford: Oxford University Press. A review of what is currently known about epigenetic inheritance. Fairly technical. Authors are committed Darwinians, but want to make sure that this aspect is not ignored, since it appears to be real. Extensive bibliography.

Judson, Horace Freeland. 1979. *The Eighth Day Of Creation: The Makers Of The Revolution In Biology*. New York: Simon and Schuster. A detailed history of the discovery of the structure and function of DNA, as rendered through interviews with many of the principals a couple of decades after the fact. Most of the history is developed through the interviews, though there is a contribution from the scientific literature as well. Much of the tone is rather smug to those who turned out to be quite right in their misgivings, but this was written in the glory days before the Genome Project screwed it all up.

Kay, Lily E. 2000. *Who Wrote The Book Of Life?* Palo Alto, CA: Stanford University Press. A history of the DNA-is-language metaphor, developed through parallel histories of molecular biology, information theory, and the militarization of scientific research, from 1945 to the present.

Keller, Evelyn Fox. 1983. *A Feeling For The Organism: The Life and Work of Barbara McClintock*. San Francisco, CA: W.H. Freeman.

Keller, Evelyn Fox. 2000. *The Century of the Gene*. Cambridge, MA: Harvard University Press. A very sensible appraisal of what we know and what we don't know about genetics. The author is an active participant in some of the cooler currents in modern genetics.

McClintock, Barbara. 1961. Some parallels between gene control systems in maize and in bacteria. *The American Naturalist*, **95**, 265–277. Points out that the genetic control structure envisioned by Jacob and Monod is similar to that envisioned by McClintock in maize.

McClintock, Barbara. 1984. The Significance of Responses of the Genome to Challenge. *Science*, **226**(16 November), 792–801. Her Nobel address. Prefigures later findings about adaptive mutability.

Morgan, Hugh D., Sutherland, Heidi G.E., Martin, David I.K., & Whitelaw, Emma. 1999. Epigenetic inheritance at the agouti locus in the mouse. *Nature Genetics*, **23**(November), 314–318. Authors induced an epigenetic mutation into mice, and then found that the mutation appears to be inheritable.

Nanney, David L. 1957. The Role of The Cytoplasm In Heredity. *Pages 134–166 of:* McElroy, William D., & Glass, Bentley (eds), *The Chemical Basis of Heredity, A Symposium*. McCollum-Pratt Institute of Johns Hopkins University. Shows results from paramecia that challenge the view that heredity is exclusively a nuclear function. The paramecia are shown to have inherited traits in a non-mendelian fashion, but only when they exchanged cytoplasm.

Nanney, David L. 1985. Heredity without genes: ciliate explorations of clonal heredity. *Trends in Genetics*, **1**(November), 295–298. A review of earlier work. Places some formerly inexplicable results within recent advances in the standard molecular genetic framework, but points out that some findings, with *Paramecia*, among others, remains unexplained by that same framework.

Radman, Miroslav. 1999. Enzymes of evolutionary change. *Nature*, **401**(28 October), 866–869.

Radman, Miroslav, Matic, Ivan, & Taddei, François. 1999. Evolution of evolvability. *Pages 146–155 of:* Caporale, L.H. (ed), *Molecular Strategies in Biological Evolution*. Annals of the New York Academy of Sciences, vol. 870.

Ridley, Matt. 1999. *Genome: The Autobiography Of A Species In 23 Chapters*. XX: HarperCollins. A nice account of the Human Genome Project so far. Omits much mention of the shortcomings, and spends most of the effort on celebration of what's been achieved.

Rosen, Robert. 1960. A Quantum-Theoretic Approach to Genetic Problems. *Bulletin of Mathematical Biophysics*, **22**, 227–255. I don't know about the quantum part, but he expresses some interesting ambivalence about the dogma of DNA as the only important component to heredity.

Rosen, Robert. 1961. On The Role Of Chemical Systems In The Microphysical Aspects Of Primary Genetic Mechanisms. *Bulletin of Mathematical Biophysics*, **23**, 393–403. Talks about whether DNA can be said to carry information in the Shannon sense. Concludes the answer may well be no, unless DNA is not static.

Scheid, Ortrun Mittelsten, & Paszkowski, Jerzy. 2000. Transcriptional gene silencing mutants. *Plant Molecular Biology*, **43**, 235–241.

Shapiro, James A. 1999. Genome system architecture and natural genetic engineering in evolution. *Pages 23–35 of:* Caporale, L.H. (ed), *Molecular Strategies in Biological Evolution*. Annals of the New York Academy of Sciences, vol. 870. Contains a great figure about progress in understanding the *lac* operon.

Sutherland, Heidi G.E., Kearns, Margot, Morgan, Hugh D., Headley, Alexander P., Morris, Christine, Martin, David I.K., & Whitelaw, Emma. 2000. Reactivation of heritably silenced gene expression in mice. *Mammalian Genome*, **11**, 347–355. The authors observed differences among genotypically identical mice that were inherited. The mutation was reversible in later generations.

Tonegawa, Susumu. 1983. Somatic generation of antibody diversity. *Nature*, **302**(14 April), 575–581. An account of genetic splicing to come up with the wide array of antibodies needed to generate a healthy immune system.

Venter, J. Craig, *et al.* 2001. The sequence of the human genome. *Science*, **291**(16 February), 1304–1351.

Waddington, C.H. 1942. Canalization of development and the inheritance of acquired characters. *Nature*, **150**(3811), 563–565. Makes the argument about ostrich callouses, which are formed while still in the egg. His point is that the ability to develop callouses implies a genetic readiness to have a 'switch' turned on. Evolving a situation where the callouses can be formed in ovo is a case of the environmental triggering of the switch getting replaced by the genetic triggering.

Waddington, C.H. 1953. Genetic assimilation of an acquired character. *Evolution*, **7**(June), 118–126. Experimental support of Waddington, 1942. Creates an environment where fruit flies acquire a trait due to temperature shock. He selected (manually) for those strains that most readily acquired the trait, then showed, that after several generations, he no longer needed the temperature shock: the trait was expressed from birth in normal individuals.

Waddington, C.H. 1968. The Basic Ideas of Biology. *Pages 1–32 of:* Waddington, C.H. (ed), *Towards a Theoretical Biology: Prolegomena*. International Union of Biological Sciences Symposium. A brilliant summation of the problems faced by the "central dogma" version of biological analysis.

Watson, James D., & Crick, Francis H.C. 1953. Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature*, **171**(30 May), 964–967. The follow-up to the paper announcing the structure of DNA.

Wilkie, D. 1964. *The Cytoplasm in Heredity*. XX: Methuen. An account of how cytoplasm may affect heredity. Predates the modern accounts of mitochondria and chloroplasts, but still has points to make about the intricacy and hidden aspects of the rest of the cytoplasm.