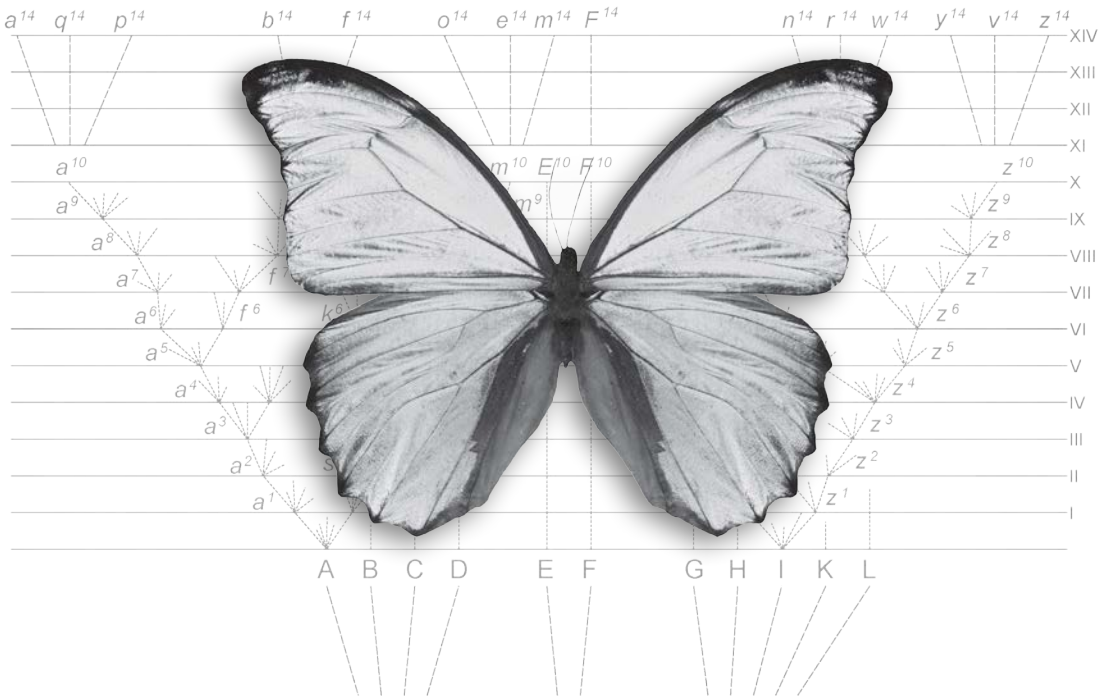
REPLACING DARWIN

The New *Origin of Species*

Nathaniel T. Jeanson

REPLACING DARWIN

The New *Origin of Species*



Nathaniel T. Jeanson

First printing: October 2017

Master Books®, P.O. Box 726, Green Forest, AR 72638
Master Books® is a division of the New Leaf Publishing Group, Inc.

ISBN: 978-1-68344-075-8
Library of Congress Number: 2017913476

Cover by Intelligent Designs Creative

Please consider requesting that a copy of this volume be purchased by your local library system.

Printed in the United States of America

Please visit our website for other great titles:
www.masterbooks.com

For information regarding author interviews,
please contact the publicity department at (870) 438-5288.

A portion of the proceeds from this book funds
continuing research on the origin of species.

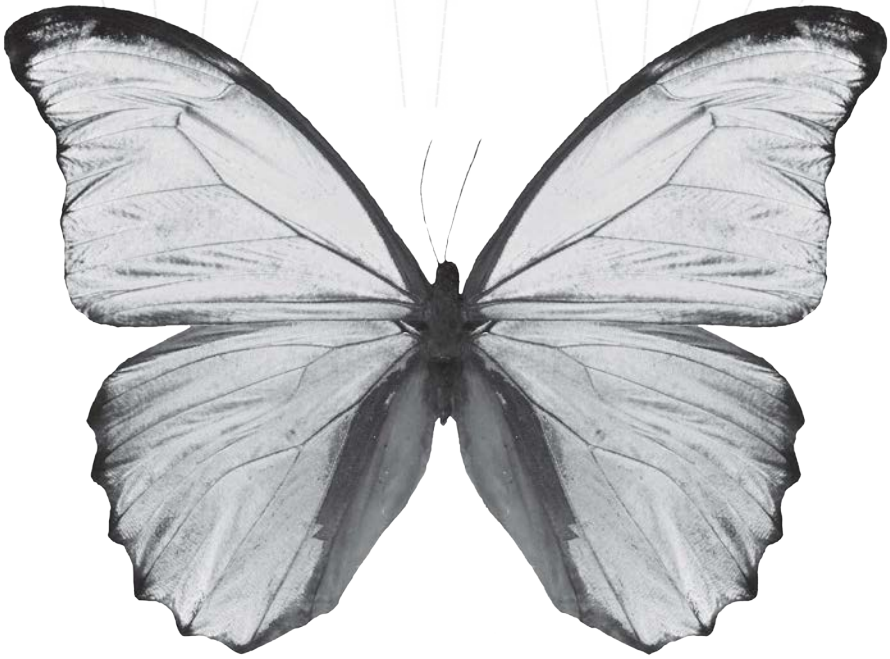
A Division of New Leaf Publishing Group
www.masterbooks.com

To
Patrick, Jason, and José
And many more like them

Contents

Introduction

Why Now?



Chapter 1

Inevitable

This book makes a bold claim — that the events of the last 130 years have rewritten the history of life on this planet. On the surface, this statement might seem outrageous. A development this provocative would create at least as much upheaval as Darwin's publication did.

Yet, upon deeper reflection, both revolutions become less surprising. In fact, in the wider context of the history of biology, you could argue that both paradigm shifts were inevitable.

To see how, an analogy is helpful. If we think of the origin of species* as a scientific jigsaw puzzle, each species is one piece of the bigger puzzle. Clues to their origins represent additional pieces. Paradigm shifts are major revisions in how the puzzle is put together.

This latter concept requires some explanation. For jigsaw puzzles that come in a box, paradigm shifts are rare. The box cover guides the progress, and the total number of pieces constrains the possible arrangements. Even if a puzzle is large, these two factors streamline and smooth the assembly process.

Unlike typical jigsaw puzzles, the puzzle of the origin of species does not come in a box. No cover exists. The final number of pieces are unknown. In fact, nearly all pieces must be actively sought. Consequently, with each new discovery, the potential for massive overhaul lurks in the background.

---

* Unless otherwise noted in this book, when I use the term *species*, I am using it in the biological sense — in other words, as a formal unit of classification.

For example, consider the state of the puzzle prior to Darwin's day. Just a century before Darwin published *On the Origin of Species*, the first pieces were discovered. In the late 1700s, Carolus Linnaeus entered the modern concept of *species* into our scientific vocabulary.[1] Since Linnaeus began the modern discipline of *taxonomy* (the identification and classification of life), the pieces of Linnaeus' puzzle were a fraction of the total known today. As an illustration, of the more than 5,400 mammal species that exist today, Linnaeus documented less than 200.[2] Linnaeus and his contemporaries had very little data with which to work.

Nevertheless, despite this dearth, the available pieces suggested themselves into a plausible puzzle image. As a modern illustration of this old image, consider the polar bear species (Color Plate 1). It occupies a snowy environment along with its fellow Arctic residents, the Arctic fox (Color Plate 2) and Arctic hare (Color Plate 3). In all three species, white coats[3] camouflage them against the stark Arctic environment.

If you move south to warmer climates, bears, foxes, and rabbits lose their white coats in favor of something more suitable.[4] In the North American forests and mountain ranges, black bears lumber much more secretively than if their coats where white (Color Plate 4). In the grasslands and forests of Eurasia, red foxes dart more clandestinely than the Arctic fox would (Color Plate 5). In the blazing heat of the desert, jackrabbits don't need white coats; instead, they need to stay cool. Unlike arctic hares, jackrabbits possess large ears — efficient thermoregulators (Color Plate 6). The challenges of each environment are matched by the unique features of each species.

This pattern is true globally. From the frigid ice floes of the Arctic to the tropical lagoons of the Great Barrier Reef; from the expansive Serengeti savanna to the dense Black Forest of Germany; from the airless slopes of the Himalayas to the soundless depths of the Pacific; from the deathly dry Sahara to the humid and lush jungles of the Amazon, species seem to have been made for the environments in which they reside.

The natural implications of these observations are clear. Going back to William Paley[5] in 1802, scholars recognized that purpose implies design. Design implies a designer. With an analogy to the familiar realm of human design, Paley illustrated his reasoning. For example, if a watch were discovered lying on the ground, no one would explain its origin by the action of wind and rain over millions of years. Rather, they would observe the intricate fit of the parts to one another and conclude the obvious — that a designer made the watch for a specific purpose. Similarly, the match between species and their environments suggested deliberate purpose — as if, by the purposeful action of a Designer, species were designed for the environments in which they reside.

In technical terms, this view led to very specific conclusions in two arenas. Since species fit their locales so well, it would seem that they were made for their individual habitats — that is, that they were created separate and distinct from other species. This implied that species do not become other species — a view known as *species' fixity*. Conversely, since species appear to have

been made for their individual habitats, it would seem that they have always been in their current locations. Thus, in the arena of geography, the design argument implied the fixity of species' locations.

In jigsaw puzzle terms, it was as if each species formed its own isolated puzzle. Environmental clues might decorate the outer edges of each species' puzzle. But, rather than connect different species' puzzles to one another, these clues seemed to separate the puzzles.

With few pieces in hand, this view was easy to maintain. The absence of obvious connecting pieces between puzzles produced a convincing set of isolated images. Nevertheless, due to the fact that many pieces were still waiting to be discovered, the potential for massive overhaul lingered.

By 1859, Darwin and his contemporaries had discovered many new pieces to these puzzles. Drawing on the growing knowledge in the fields of biogeography (i.e., the geographic distribution of species around the globe), anatomy, physiology, embryology, geology, and paleontology, Darwin began to see connections where prior investigators saw only empty space.[6] Eventually, Darwin proposed that all species evolved from one or a few common ancestors — a massive paradigm shift.

Because of the unusual nature of the puzzle of the origin of species, paradigm shifts are inevitable.

Like the 18th century, the scope of species diversity in Darwin's day was a fraction of today's variety. In 1859, the scientific community had no knowledge of the majority of species we have now documented. With over 1.6 million[7] plant, animal, fungal, and bacterial species currently known, hundreds of thousands of pieces were missing from Darwin's puzzle.[8]

With Darwin barely 100 years removed from Linnaeus' foundational work, this fact shouldn't be surprising.

Darwin didn't compensate for this ignorance of species diversity with any special abilities. His lifespan wasn't any longer than the average lifespan today. He observed the world for 73 years. And then he died. Furthermore, in those 73 years, he was subject to the technology of the 1800s. He couldn't travel the globe as easily as we do. Without the information exchange facilitated by the Internet, he couldn't benefit as easily from the travels and discoveries of others. Yet Darwin tried to tackle one of the biggest questions in biology.

Since 1859, we've had time to reevaluate his picture — much more time than he had to propose and appraise it. We've also had more space. Today, travel is virtually unrestricted. Few corners of the world have remained recalcitrant to scientific exploration. Furthermore, the Internet makes information sharing faster than ever before. A global community of millions[9] of scientists can pool their resources and build on one another's work. Though lifespans have changed little, the cumulative observations of these scientists have built an unprecedented body of knowledge on the diversity and operation of life.

Consequently, the puzzle image has changed.

Three developments have led the way. First, after Darwin wrote *On the Origin of Species*, an entire field of science was born — and then matured. Unlike any other field of science, this field directly constrains and guides the answer to the origin of species. Consequently, it's the most relevant field to our question. You could say that the edge pieces of the puzzle have finally been found.

Second, premature conclusions were corrected. Anyone who has gotten stuck trying to put together a puzzle without a box cover and without edge pieces would do what Darwin did — they would test piece after piece until they found a plausible connection. However, without the constraints of a box cover image and of edge pieces, it's easy to link pieces where no link exists. Connections that initially appear plausible eventually give way to the correct links, once additional pieces are found. Darwin made many such premature links. Since 1859, we've been able to unlock some of the connections that he erroneously made, while cementing correct ones.

Third, in the last few years, the critical corner pieces were found. Several remarkable scientific discoveries were made — ones entirely unanticipated by the trajectory of discoveries prior. With these pieces in hand, the framework of the puzzle and the existing connections among pieces have been reoriented.

Individually, each of these developments carried minor significance. By analogy, if you were trying to put a puzzle together, the discovery of a few edge pieces would be helpful. But it wouldn't be earth-shattering. Conversely, if you found a corner piece, this discovery would be fantastic. But if the remaining pieces have been forced together in clumsy and incorrect ways, the corner would do little good. Finally, if all you did was unlock a few misconnected pieces, you would rejoice in the removal of barriers to progress. But the happiness of this success would soon be outweighed by the intimidating scope of the remaining task. In isolation, these discoveries would do little to reveal the final image.

Together, they produced a compelling picture of how species came to be.

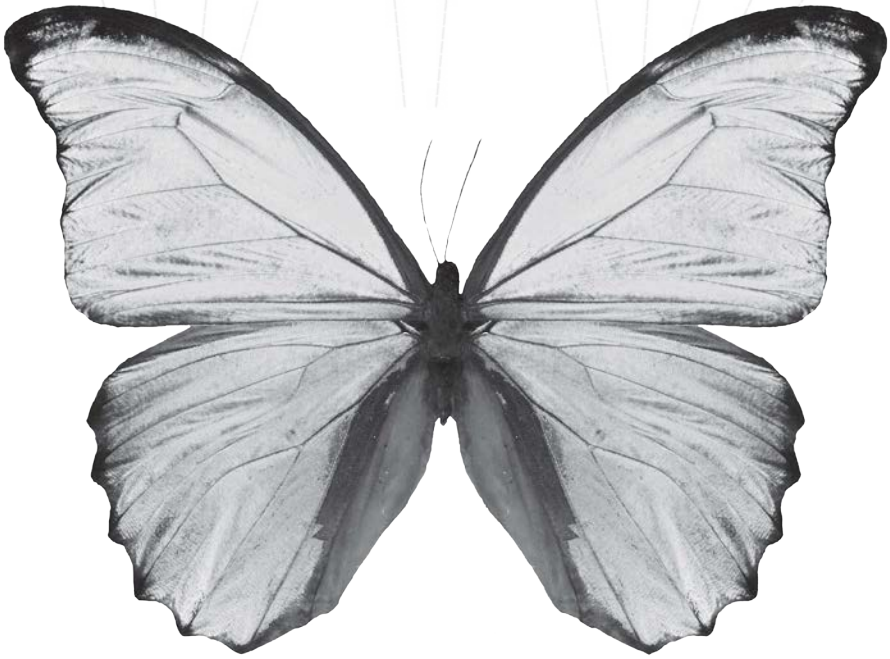
To be sure, large chunks of the puzzle still need to be filled in. Having the corner pieces, edge pieces, and a couple of correctly connected center pieces is a huge step forward. But significant holes in the puzzle remain.[10] Explaining in detail the origin of every species that ever lived is a monstrous undertaking. Much work remains to be done.

Nevertheless, the puzzle picture that we possess today is far different from the one that Darwin created. And it is far superior. It puts the far reaches of the globe — and the species that they contain — into an image that is as captivating as it is surprising.

This book tells the story of how this picture came to be.

Part One

A Field Is Born



Chapter 2

The Secret of Life

Sweating over a large, unassembled jigsaw puzzle, I take deliberate steps to simplify the vexing challenge. First, I search for the edge pieces. Once I've found them all, it's a fairly simple task of trial and error to connect them. Far fewer possible connections exist among these pieces than among the center pieces. Furthermore, once connected, they provide an enormously helpful framework in which to connect the rest of the puzzle.

If my puzzle came without a box cover, the edges would be doubly useful. Edges limit the amount of possible horizontal and vertical connections among the center pieces. This saves me the enormous frustration that follows endless trial and error of unassembled center pieces. Edges also naturally suggest how the final image will look. Even though each piece contains a tiny part of the whole image, I can sense the final subject matter just by looking at the edge pieces. In the assembly of a jigsaw puzzle, the identification of edge pieces is a major step forward.

Since species are not literal jigsaw puzzles, the biological analogy for edge pieces might not be obvious. The parallel becomes clearer upon brief reflection. Consider species with which we are familiar. We recognize zebras by their stripes, elephants by their trunks, giraffes by their long necks, bald eagles by the color of the feathers on their heads, and monarch butterflies by the patterns on their gossamer wings. Species are defined by their traits.

This is true across all life. Mammals, reptiles, birds, amphibians, fish, starfish, sea urchins, crustaceans, arachnids, insects, worms of all sorts, shellfish, octopi, snails, corals, jellyfish, sponges, mosses, ferns, conifers, grasses, orchids, fruit trees, fungi, algae, bacteria, and all the

other life forms on earth possess unique combinations of traits. Though some features require a microscope to visualize, traits define species.

Therefore, the question of the origin of species is a question of the origin of traits. If you want to know the origin of zebras, you need to discover the origin of stripes. The origin of elephants is bound up in the origin of trunks. Giraffe origins are inextricable from the origin of long necks. The origin of eagles goes hand in hand with the origin of white feathers. Butterfly history is read off the history of butterfly wing patterns. The origin of the rest of the species is found in the origin of the traits that define them.

The solution to the origin of traits represents the hard constraints on the origin of species — the edge pieces of the puzzle.

In 1859, zero edge pieces were known. With careful reflection, the reason for this is easy to grasp. For example, if you wanted to discover the origin of these traits, you could begin by watching how they behave each generation. However, unlike humans, species don't keep written records of their family trees. Thus, as a first step, you might start by investigating human family trees.

The simplest place to start is your own family tree. Perhaps you recognize your father's chin in your jaw. You might investigate how far back on your family tree you can trace this chin shape. However, if you're like me, your family tree is probably small. Going back further than a few generations, I don't know who my relatives are. If your tree is like mine, your ability to examine trait behavior is severely limited by your ignorance of your ancestors.

Your attempts to track traits might encounter a second hurdle. If your family tree is small, you might compensate by tracing additional family trees. In doing so, you'd probably have to follow more traits than chin shape. For example, you might document the behavior of red hair and freckles. If you did, you might observe that, on occasion, the trait disappears from a family tree. A red-headed parent might have no red-headed offspring. Or a great-grandparent might have red hair, but several generations of descendants might not. As the scope of your investigation expands, you might find several traits that behave in odd and inexplicable ways.

These idiosyncrasies apply to both living and extinct species. In fact, when fossils are part of the equation, the problems multiply. Unlike recorded family trees, fossils have no explicit genealogical connection to anything alive today. Ancestral relationships have to first be inferred from indirect data before trait behavior can be tracked. Even more troubling, the placement of fossils on family trees requires implicit assumptions about how traits behave. Assuming a mode of inheritance to prove a mode of inheritance is circular reasoning. In other words, fossils cannot inform how traits behave.

If you had access to a microscope, you'd discover the most inexplicable behavior of all. *All* traits are erased each generation — and then rebuilt. When sperm meets egg, the visible features that define multicellular species are not present. Instead, these characteristics arise via the process of development.

In summary, if you rigorously tracked the behavior of visible traits, you'd discover an intimidating number of paradoxes. These paradoxes would raise a host of perplexing questions. Do traits form spontaneously? Can they be destroyed? Can they be changed? If so, how much can they be changed? Are traits blended? Particulate? Inherited as a whole? Separated into units? Independent? Interdependent?[1]

Consider the ramifications of this uncertainty. If traits can appear and disappear, how could you trace species' ancestry? What markers would you use to fill in the family tree? Furthermore, if all traits are rebuilt every generation, can any species become any other species? Do any constraints on change exist? Might a fish spontaneously spawn a spider? Without an answer to the mystery of heredity, the origin of species would be an enigma.

When Darwin wrote *On the Origin of Species*, paradoxes — not edge pieces of the puzzle — were all that the scientific community possessed.

The first steps toward resolving these paradoxes were taken in 1865 — six years after Darwin's seminal publication. An Austrian monk, Gregor Mendel (Figure 2.1), solved the paradoxes of family trees. Like nearly every other species, his subject of study — pea plants — did not keep written records of inheritance. So Mendel did it for them.

Figure 2.1 — Gregor Mendel

Mendel watched and documented inheritance in pea plants over several generations. His tedious labors involved counting pea plant offspring and recording the traits that appeared in each generation. Over the span of nearly a decade, Mendel made hundreds of crosses and counted thousands of peas. The mathematical precision with which Mendel documented his results laid the groundwork for several revolutionary inferences.[2]

One of the first discoveries that Mendel made was the discrete nature of genetic information. For example, Mendel crossed pea plants with pure-breeding* yellow seeds to pea plants with pure-breeding green seeds. All of the offspring (i.e., the F_1 generation) of this union bore yellow seeds (Color Plate 7). He didn't observe yellowish-green seeds or some other blend between the two colors. Instead, the traits remained distinct. Mendel's experiments demonstrated the fact of *particulate* inheritance rather than *blended* inheritance.

Mendel called these particulate units of inheritance *unit factors.*

Applied more broadly to species, Mendel's discoveries were a tremendous step forward. When traits appear and disappear on family trees, it must have something to do with the unit

* In other words, they always bred true for a particular trait. For example, pea plants that are pure-breeding for yellow seeds produce offspring that always have yellow seeds.

factors behind these traits. For example, red hair and freckles might be encoded by unit factors. Your chin shape might be as well. Consequently, if we can identify and track unit factors, we might be able to reconstruct the family trees for each species.

When the offspring of Mendel's cross (i.e., the plants bearing strictly yellow seeds) were self-fertilized, about 75 percent of this second generation (i.e., the F_2 generation) bore yellow seeds. Twenty-five percent bore green seeds — a 3:1 ratio (Color Plate 7). In other words, in Mendel's experiments, green seed color disappeared in the offspring of the cross between pure-breeding plants. Yet it reappeared in subsequent generations. Thus, the instructions for green didn't get eliminated. They were just hidden by yellow for a single generation. Mendel referred to the behavior of the yellow seeds as *dominant* since the instructions for yellow dominated over the instructions for green. He referred to the behavior of the green seeds as *recessive*.

The mathematical proportions of these offspring yielded another discovery about the nature of inheritance. Mendel self-fertilized the second-generation plants (Color Plate 7). In the offspring (the third generation, or F_3), he again observed consistent mathematical ratios in the seed color trait. If second-generation (F_2) green seed plants were self-fertilized, they always bore green seeds. When second generation (F_2) yellow seed plants were self-fertilized, one-third of the offspring (F_3) bore strictly yellow seeds, while the other two-thirds bore yellow seeds and green seeds in a 3:1 ratio (Color Plate 7). Without going into the details of the math, Mendel was able to infer the behavior of a single trait. In this case, the differing versions of the seed color trait *segregated* from one another each generation. The paternal instructions and the maternal instructions were not combined; they stayed distinct.

These observations begin to explain why red hair appears and disappears irregularly on family trees. Red hair behaves in a recessive manner. When two red-haired individuals bear children, the offspring all possess red hair. But if one of the parents lacks red hair, the offspring are mixed — some might have red hair, some might not. In fact, if the parent without red hair has no red-headed ancestors, none of the children might possess red hair. If both parents lack red hair — but if both have a red-headed ancestor — a red-headed offspring might still arise. But the red-haired offspring will likely represent a minority of the children.

Take zebras as another example. Though zebras, horses, and donkeys are all separate species, they can hybridize to produce sterile offspring. When striped zebras are crossed to unstriped donkeys, the offspring bear stripes (Color Plate 8). They aren't spotted or strictly solid color throughout. Instead, the striping pattern is still distinct. The same phenomenon occurs when zebras and horses hybridize (Color Plate 8). In short, it seems that the unit factor for stripes is dominant.

However, in the offspring of these crosses, the striping pattern varies. Sometimes stripes occur primarily on the legs of the hybrids (Color Plate 9), rather than both on the legs and on the sides of the torso (Color Plate 8). Perhaps several unit factors control coat color patterning. The stripes trait might dominate only one aspect of the process of creating the coat color patterning

in the adult. Regardless, stripes in zebras behave largely according to the principles of inheritance that Mendel uncovered in plants.

Together, the discovery of unit factors, of dominant and recessive traits, and of the segregation of genetic information began to define clear rules for the ways in which traits behave each generation.

Mendel performed even more complex crosses, which revealed yet another fundamental principle of inheritance. In addition to seed color, Mendel crossed plants bearing variety in other traits, such as seed form. For example, some of his plants bore seeds that were wrinkled while others bore seeds that were round (Color Plate 10). Just like with seed color, one form was dominant (e.g., round), and the other recessive (e.g., wrinkled).

In one series of experiments, Mendel crossed plants that differed in seed color and seed form. The first cross involved pure-breeding yellow-colored, round form seed plants. These were crossed with pure-breeding green-colored, wrinkled form seed plants. As expected, all offspring (F_1 generation) had yellow seeds that were round. When these second-generation plants (F_1) were self-fertilized, every possible combination appeared in their offspring (F_2 generation). Some F_2 plants had yellow and round seeds (about 9/16th of the offspring); some had yellow and wrinkled seeds (about 3/16th of the offspring); some had green and round seeds (about 3/16th of the offspring); and some had green and wrinkled seeds (about 1/16th of the offspring) (Color Plate 10). In other words, the behavior of one trait was not tied to the behavior of another. If the seed color trait behaved one way, it had no effect on the behavior of the seed shape trait.

Mathematically, the proportions of the offspring corresponded to what might be predicted from a model of two traits inherited independent of one another. Diagrammatically, all of these probabilities can be derived with a *Punnett square* (Color Plate 10). Since we're dealing with two traits that each have dominant and recessive forms, the predicted frequencies are slightly skewed. Due to their recessive nature, recessive forms are predicted to appear less frequently than the dominant ones. But by assuming that traits are independent of one another, and that the differing versions of each trait segregate from one another (i.e., this is what the Punnett square diagram does in Color Plate 10), we can make mathematical sense of the results.

In contrast, if the behavior of one trait was dependent upon the behavior of another, then the offspring would show ratios that did *not* agree with the probabilities derived from the Punnett square.

Since the ratios agreed, Mendel inferred that the instructions for different traits *sorted independently of one another.*

This discovery opened up a whole new world of possibilities. Think of all the traits that define your external features. Think of all the traits that define each species. Unit factors must exist for each of these traits. Some might be dominant, while others are recessive. Since the

dominant and recessive versions of unit factors segregate from one another, traits can appear and disappear each generation. In a sense, these genetic discoveries define some of the limits to trait behavior. Conversely, since unit factors sort independently of one another, an overwhelming number of potential combinations exist. Together, Mendel's principles uncovered both the boundaries and the enormous opportunities for variety that exist within each species.

In the years following Mendel's work, other principles of inheritance have been discovered. In other words, exceptions to Mendel's rules were found. For example, some unit factors are linked and do not sort independently of one another. As another example, other unit factors do indeed act in concert to produce a blended trait outcome. Nevertheless, Mendel's findings laid a major foundation for our modern understanding of trait behavior.

For reasons unknown, Darwin appears to have been unaware of Mendel's work. Conversely, for equally unknown reasons, when Mendel died in 1884, his discoveries died with him, not to be resurrected until the turn of the century.

Despite the strength and rigor of Mendel's conclusions, his results never answered the question of *why* unit factors behave in the way that they do. In other words, Mendel's rules successfully predicted how traits interact and combine. Yet these predictions didn't reveal whether the rules could change — or whether they have changed in the past.

If the rules could change, then traits might behave in entirely unexpected ways. Perhaps a fish could become a spider. Maybe it did in the distant past. Or maybe this change is impossible. Without the *why* of the rules of inheritance, these speculations would exist unfettered by reality. Without these fetters, the origin of traits would remain a mystery. Without an explanation for the origin of traits, the origin of species would remain an unsolved puzzle.

Several decades after Mendel's death, meticulous observation of cells during cell division put Mendel's principles in more concrete, subcellular, and potentially mechanistic terms. Under the microscope, *somatic cells* (i.e., non-reproductive cells) were observed to divide through a process of cell division termed *mitosis*. In both animals and plants, before the nucleus breaks down, structures that look like flexible noodles — *chromosomes* — appear in a period of time termed *prophase* (Figure 2.2). By *prometaphase*, the membrane surrounding the nucleus breaks down (Figure 2.2).

During *metaphase* (Figure 2.2), the chromosomes line up in the center of the cell. These chromosomes appear as x-like structures because they represented two identical (replicated) chromosomes that are still partially joined. In *anaphase* (Figure 2.2), the replicated chromosomes are separated from one another and are pulled toward opposite ends of the cell. By *telophase* (Figure 2.2), the cell begins to split into two cells, each cell containing a chromosome content identical to the other cell, and the nuclear membrane begins to reappear. *Cytokinesis* completes the splitting process.

Leaving Mendel's discoveries aside for the moment, consider the ramifications of what we just observed. The fact that the process of nuclear division was so complex suggested a functional role for chromosomes. If chromosomes were inert and irrelevant to heredity, why would cells take such care to pass them on via such a convoluted cycle?

The behavior of somatic cells was just the beginning of the answer to the question of inheritance. More relevant was the behavior of *germ cells* or *gametes* — sperm and egg. Chromosomes are also distributed among these cells during normal development in a cell division process termed *meiosis*. This process bears strong resemblance to mitosis, but also differs in key ways. In both mitosis and meiosis, the chromosome number and behavior is very predictable. In the early 1900s, the American scientist Walter Sutton provided the documentation. During the process of meiosis, Sutton observed that the chromosomes occurred in recognizable *pairs* (Figure 2.3). For example, in Figure 2.3, for each long x-like structure, there exists another long x-like structure. For each short x-like structure, there exists another short x-like structure. Technically, each x-like structure consists of two chromosomes — one a replicated copy of another. Furthermore, these pairs were separated during meiosis, and the individual chromosomes in each x-like structure were distributed among the various products of *spermatogenesis* and *oogenesis* — the processes which give rise to sperm and egg (Figure 2.3). Sutton suggested that one member of each chromosome pair was ultimately maternal in origin and the other, paternal (Figure 2.3). In other words, an individual inherits one

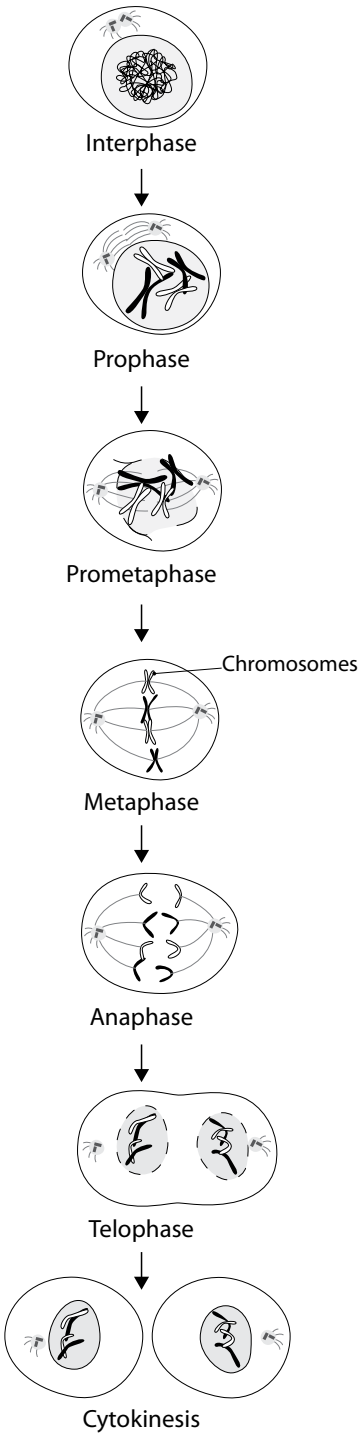
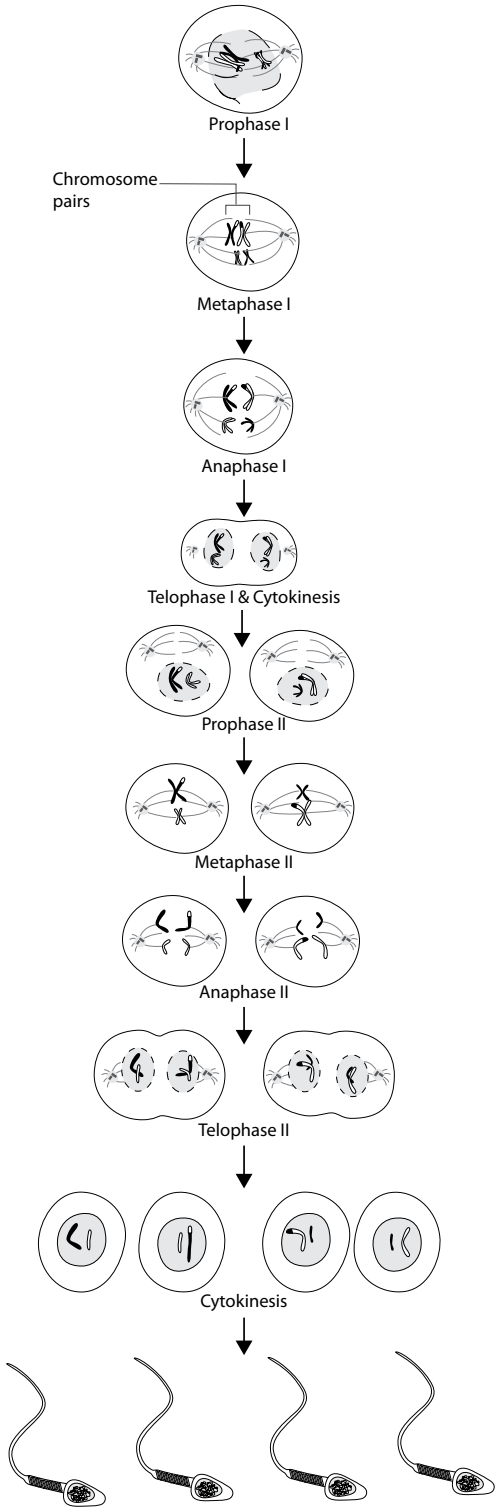


Figure 2.2. Cell division in non-reproductive cells. Somatic cells (non-reproductive cells) undergo a very strict series of cell division events in a process termed mitosis. In early phases (e.g., prophase), replicated chromosomes become visible. Until anaphase, the replicated chromosomes remain attached, leading to the familiar x-like shape seen in each of the four chromosomes in the metaphase display shown here. After metaphase, the remaining phases of the cell cycle are dedicated to separating these replicated chromosomes and divvying them up into two daughter cells.


Prophase I
Chromosome pairs
Metaphase I
Anaphase I
Telophase I & Cytokinesis
Prophase II
Metaphase II
Anaphase II
Telophase II
Cytokinesis

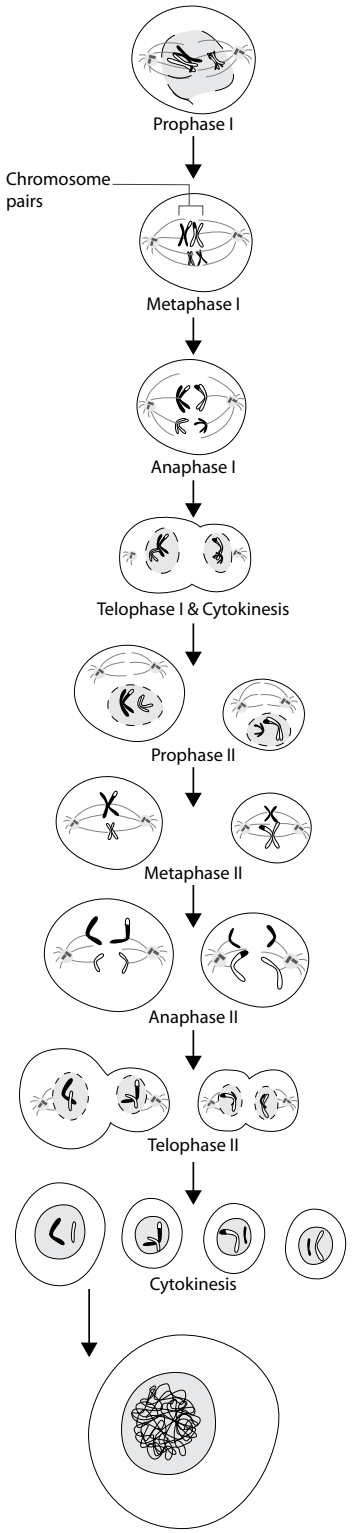

Prophase I
Chromosome pairs
Metaphase I
Anaphase I
Telophase I & Cytokinesis
Prophase II
Metaphase II
Anaphase II
Telophase II
Cytokinesis

Figure 2.3 (previous page). Cell division in sperm and egg. To produce sperm and egg, cells undergo a process of two cell divisions termed meiosis. One of the major differences between mitosis and meiosis is the arrangement of the chromosome pairs in the first cell division. In mitosis, chromosomes line up vertically (i.e., with respect to the display in figure 2.2) during metaphase, and the replicated chromosomes are separated to produce identical chromosome content in the resultant daughter cells. In meiosis, during the first metaphase, chromosome pairs line up next to one another (i.e., with respect to the display in this figure), and pairs — not individual chromosomes — are separated from one another. Then, during the second round of cell division, in which chromosomes line up vertically during metaphase II, the replicated chromosomes are separated from one another. Because the second round of cell division during meiosis proceeds without additional DNA replication, the chromosome content in sperm and egg is half of that in somatic cells (a condition known as haploid; somatic cells have the full — diploid — chromosome content).

Of the four products of meiosis during the formation of egg cells, only one of the four products eventually becomes a mature egg. In contrast, all four products of meiosis during the formation of sperm cells result in mature sperm.

During prophase I of meiosis, a chromosomal diversification process happens ("crossing over"), which is discussed in more detail in chapter 9.

member of each pair from each parent — as implied by the chromosome numbers and arrangement in Figure 2.3.

Immediately, Sutton's discovery suggested a link to Mendel's findings. Since chromosomes physically segregate from one another during the formation of sperm and egg, perhaps chromosomes contain the unit factors which segregate over successive generations.

For example, pea plants contain chromosomes. If two pure-breeding plants were crossed to one another, each pure-breeding parent would possess a pair of chromosomes. However, in the process of meiosis, even though the parents would each contain two chromosomes, they wouldn't produce gametes with *pairs* of chromosomes. They would produce reproductive cells with *individual* chromosomes (Color Plate 11).

If we treat chromosomes as the repositories for unit factors, we can make sense of Mendel's results. For example, a pure-breeding yellow seed color parent would pass on a chromosome containing only yellow seed color instructions. A pure-breeding green seed color parent would pass on a chromosome containing only the green seed color instructions. In the offspring (F_1), the gametes would fuse and restore the chromosomal pairing arrangement, resulting in a chromosome for yellow seed color paired with a chromosome for green seed color. Since yellow dominates green, the seeds of the offspring (F_1) would all be yellow, despite having a chromosome for yellow seed color and a chromosome for green seed color (Color Plate 11).

Self-fertilization of these chromosomally mixed individuals to produce the next generation (F_2) illustrates the chromosome-Mendel link as well (Color Plate 11). Since the process of meiosis occurs again in the F_1 individuals, individual chromosomes — not pairs of chromosomes — would be produced in the gametes of the F_1 individuals. This entails that the chromosome for

yellow seed color and the chromosome for green seed color would segregate from one another. To produce the F_2 individuals, these gametes with their individual chromosomes could be combined in a variety of ways — resulting in both yellow and green seeded individuals in the F_2 generation (Color Plate 11).

Treating chromosomes as the repositories for unit factors also makes sense of Mendel's more complex crosses (Color Plate 12). For example, in the parents, let's postulate that one set of chromosomes had the unit factors for seed color and another set of chromosomes the unit factors for seed shape.[3] Since chromosomes come in pairs, the pure-breeding pea plants would have chromosome pairs containing the same information. Pea plants with smooth yellow seeds would possess a chromosome pair with both copies specifying yellow color. In another chromosome pair, both copies would specify smooth shape (Color Plate 12). The pure-breeding pea plants with rough green seeds would possess a similar state. One chromosome pair would consist of both copies specifying green color. Another chromosome pair would consist of both copies specifying rough shape (Color Plate 12).

Chromosome number		Combinations in germ cells	Combinations in somatic cells
In somatic cells	In germ cells		
2	1	2	4
4	2	4	16
6	3	8	64
8	4	16	256
10	5	32	1,024
12	6	64	4,096
14	7	128	16,384
16	8	256	65,536
18	9	512	262,144
20	10	1,024	1,048,576
22	11	2,048	4,194,304
24	12	4,096	16,777,216
26	13	8,192	67,108,864
28	14	16,384	268,435,456
30	15	32,768	1,073,741,824
32	16	65,536	4,294,967,296
34	17	131,072	17,179,869,184
36	18	262,144	68,719,476,736

Table 2.1. Tremendous potential for combinatorial diversity in species with high chromosome numbers. Adapted and redrawn from W.S. Sutton, "The Chromosomes in Heredity," *Biological Bulletin*, 1903, 4:231–251.

Since each parent would pass on only one member of each chromosome pair during the formation of gametes, the fusion of these gametes to produce the offspring (F_1) would lead to a very predictable outcome (Color Plate 12). In the F_1 individuals, one member of the chromosome pair specifying seed color would have the unit factor for yellow. The other member of the pair would have the unit factor for green. For the chromosome pair specifying seed shape, one member of the pair would have the unit factor for smooth. The other member of the pair would have the unit factor for rough. Since yellow and

smooth dominate green and rough, only pea plants with smooth yellow seeds would be seen.

Chromosome number		Combinations in germ cells	Combinations in somatic cells
In somatic cells	In germ cells		
46	23	8,388,608	70,368,744,177,664

Table 2.2. Tremendous potential for combinatorial diversity in humans.

When these individuals (F_1) were self-fertilized, the offspring (F_2) would each receive a single member of each chromosome pair from their parents. The individual members of each chromosome pair would segregate from each other. The two pairs of chromosomes would sort independently of each other. Thus, all four combinations of traits — due to differing combinations of chromosome pairs — would be visible in the offspring. The mathematical proportions of these traits would follow the predicted probabilities outlined by the way in which chromosomes could be inherited (Color Plate 12).

In sum, predicting and tallying results either by visible appearance or by chromosome distribution yields the same conclusion — the ratios that Mendel documented in 1865.

Along with a number of critical functional tests in the early 1900s, Sutton's observations began to sketch a more comprehensive picture of inheritance.

The synthesis between these chromosomal observations and Mendel's conclusions raised a number of intriguing possibilities. If chromosomes do indeed encode Mendel's unit factors, and if unit factors specify traits, then the number traits encoded by chromosomes is mind-boggling. In 1903, Sutton published a table[4] (adapted in Table 2.1) showing the theoretical number of chromosome combinations[5] (and, by extension, trait combinations) that were possible from varying chromosome numbers. In humans, 46 chromosomes exist, implying a bewildering possibility of trait combinations[6] (Table 2.2). In zebras, 32–46 chromosomes exist, a fact which

Species	Common name	Chromosome number		Combinations in germ cells	Combinations in somatic cells
		In somatic cells	In germ cells		
Caenorhabditis elegans	Roundworm	12	6	64	4,096
Apis mellifera	Honeybee	32	16	65,536	4,294,967,296
Xenopus laevis	African clawed frog	36	18	262,144	68,719,476,736
Anolis carolinensis	Green anole	36	18	262,144	68,719,476,736
Danio rerio	Zebrafish	50	25	33,554,432	1,125,899,906,842,620
Equus caballus	Horse	64	32	4,294,967,296	18,446,744,073,709,600,000
Gallus gallus	Chicken	78	39	549,755,813,888	302,231,454,903,657,000,000,000

Table 2.3. Tremendous potential for combinatorial diversity across the animal kingdom.

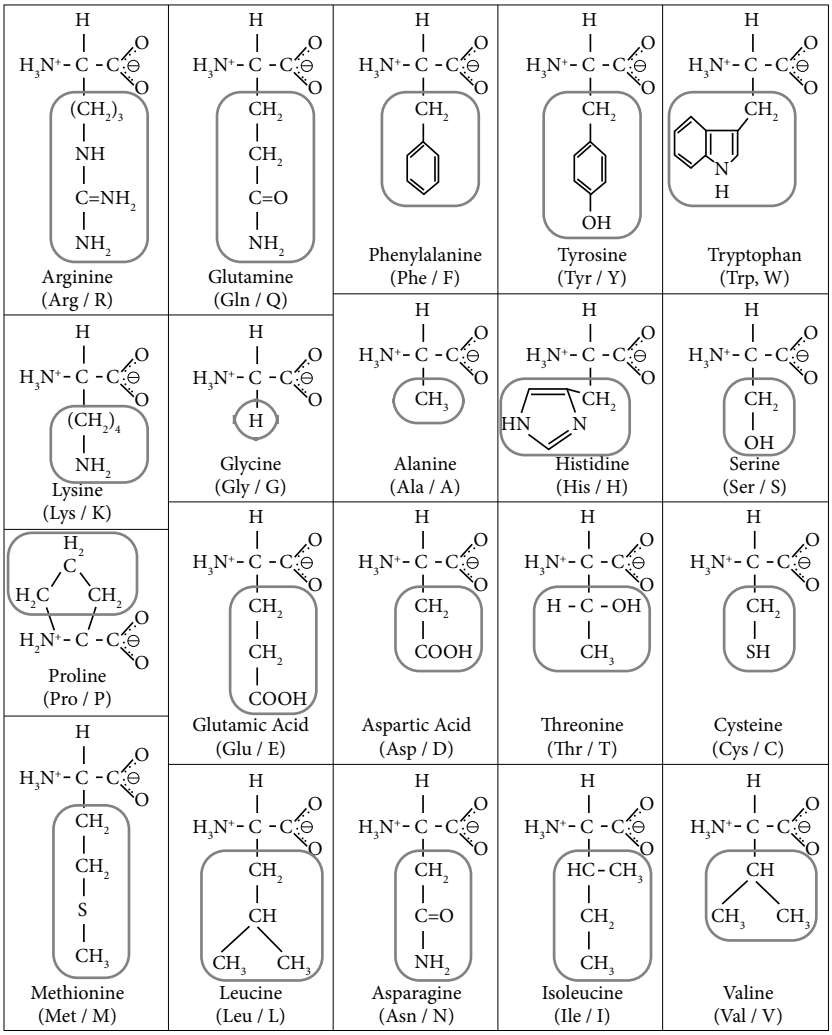


Figure 2.4. Chemical diagrams of 20 standard amino acids. In the chemical diagrams themselves, chemical elements are represented with single letter abbreviations: carbon (C), nitrogen (N), oxygen (O), hydrogen (H), sulfur (S). Five-sided and six-sided chemical structures consist entirely of carbon and hydrogen, unless otherwise indicated. Chemical bonds indicated by lines (single bond by one line; double bond by two lines). Subscripts denote number of atoms of the adjacent element. Below each chemical diagram (i.e., at the base of each box), the name of each amino acid is given three ways — as a full name, as a three-letter abbreviation, and as a single-letter abbreviation (the latter two ways are in parentheses). Rounded gray boxes indicate the various identities of the "R" placeholder in Figure 2.6.

also implies that a tremendous diversity of traits could be encoded on chromosomes. Across the animal kingdom[7] (Table 2.3), chromosome diversity suggested that Sutton's discoveries[8] were applicable to virtually all multicellular life.[9]

Sutton's findings began to explain *why* traits behave in the manner that they do, which moved the scientific community one step closer to understanding the origin of traits — and to discovering the critical edge pieces to the puzzle of the origin of species.

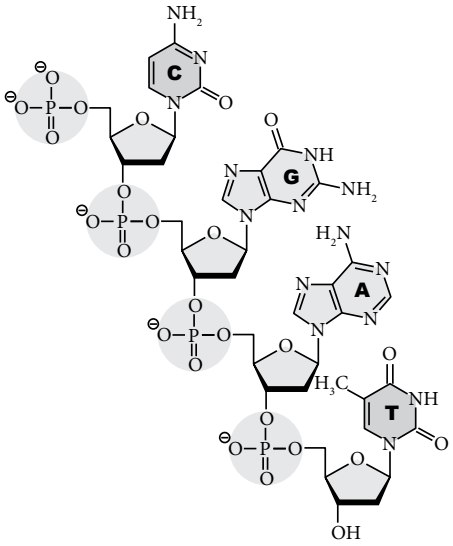


Figure 2.5. Chemical diagram of 4 linked DNA nucleotides. Elements are represented with single letter abbreviations: carbon (C), nitrogen (N), oxygen (O), hydrogen (H), phosphorus (P). Five-sided and six-sided chemical structures consist entirely of carbon and hydrogen, unless otherwise indicated. Chemical bonds indicated by lines (single bond by one line; double bond by two lines). Subscripts denote number of atoms of the adjacent element. The four different nucleotide subunits are designated by the bold, single-letter abbreviations for the defining bases (i.e., the six-sided chemical structures, or the joined five-sided and six-sided chemical structures): Adenine (A), cytosine (C), thymine (T), guanine (G).

Once inheritance was firmly connected to chromosomes, a new question greeted geneticists (scientists who study genetics). At a surface level of explanation, chromosome behavior described trait behavior. But the behavior of chromosomes didn't explain enough detail to be useful. How *specifically* did chromosomes contain the information for traits? How could a microscopic cellular structure possess the instruction manual for building the traits that define each species? How do chromosomes act like a molecular code?[10]

Before these questions could be broached, a much more mundane — but inescapable — problem had to be solved. In multicellular species, chromosomes are chemically composed of at least two major biological molecules, protein and deoxyribonucleic acid (DNA). Proteins are composed of chemical substances termed *amino acids*. Twenty different amino acids occur normally in most species (Figure 2.4). DNA is composed of chemical substances termed *nucleotides*, of which four different kinds occur (Figure 2.5). Which substance — DNA or protein — was the primary carrier of the genetic information? And how did it do so?

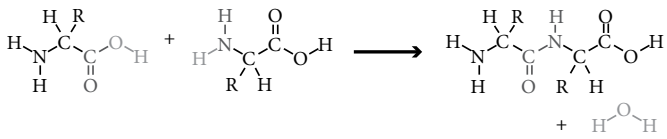


Figure 2.6. Linking individual amino acid subunits into a chain. Elements are represented with single letter abbreviations: carbon (C), nitrogen (N), oxygen (O), hydrogen (H). The single letter R is a placeholder for various amino-acid-specific chemical links (see Figure 2.4 for examples). Chemical bonds indicated by lines (single bond by one line; double bond by two lines).

In a sense, the discovery of the chromosome-trait link was like connecting the content of this book to the pages between the covers — but without identifying whether the paper or the ink contained the information. It would

be impossible to determine the book's contents without knowing whether the fabric of the paper or the arrangement of the ink into letters carried the information.

In the cellular realm, the chemical details naturally suggested a candidate. Both proteins and DNA are composed of chains of their respective subunits, amino acids and nucleotides (Figures 2.5–2.6). Theoretically, proteins made better candidates than DNA for carrying genetic information. With 20 amino acids to choose from (Figure 2.4), many protein chain combinations were possible — many more than for DNA chains of nucleotides.

For example, let's say that the information in chromosomes is contained in the chemical equivalent of three-letter words. If DNA carried the information, only 64 total words would be possible (4 * 4 * 4 = 64). In contrast, if proteins carried the information, 8,000 total words would be possible (20 * 20 * 20 = 8,000). The contrast grows stronger as the word size grows bigger.

At this stage in our discussion, these speculations might seem completely disconnected from reality. How could a chain of chemicals carry information for two eyes, two ears, a nose, and a mouth? Of all the things to hypothesize, these scientific pursuits might appear irrational. Yet this is where prior investigations led. The hypothesis that information was encoded in chains of chemicals was the best guess possible in light of the known facts of the time.

Furthermore, we regularly encode information in chains other than printed Roman letters. Braille is language that uses raised dots. A chain of physical bumps encodes complex information. Chemicals represent a kind of braille — but at a much tinier level than the dots we touch.

The theoretical contrast between the information-carrying capacity of protein and DNA chains prompted investigators in the early 1900s to propose that DNA acted like a scaffold — like the paper on which braille dots are printed. With just four nucleotides to choose from, it seemed straightforward to erect a repeating DNA structure upon which protein information could be hung. Furthermore, researchers measured the amount of the four different nucleotides in cells and found them to be equal. Together, these facts seemed to support a model termed the *tetranucleotide hypothesis* — the view that DNA is a passive, structural edifice in the process of inheritance, not an active repository of heritable information.[11]

Other observations seemed consistent with this idea. In terms of dry weight in cells, proteins represent 50 percent — they are major components. By contrast, early chemical analyses of DNA suggested that it was a small molecule — too small to carry much information. In addition, when nucleotide material was compared across various species, the nucleotide ratios appeared to be the same. If nucleotides contained the heritable information, you might expect their ratios to vary across creatures whose heritable features are diverse. Yet no simple to complex hierarchy was apparent.

In the early 1900s, traits such as red hair, chin shape, zebra stripes, and the elephant's trunk were thought to be encoded by chains of amino acids on chromosomes.

In 1928, a British biologist, Fred Griffith, described an intriguing finding in tiny bacterial cells — with large ramifications for the relationship between chromosomes and traits. In his experi-

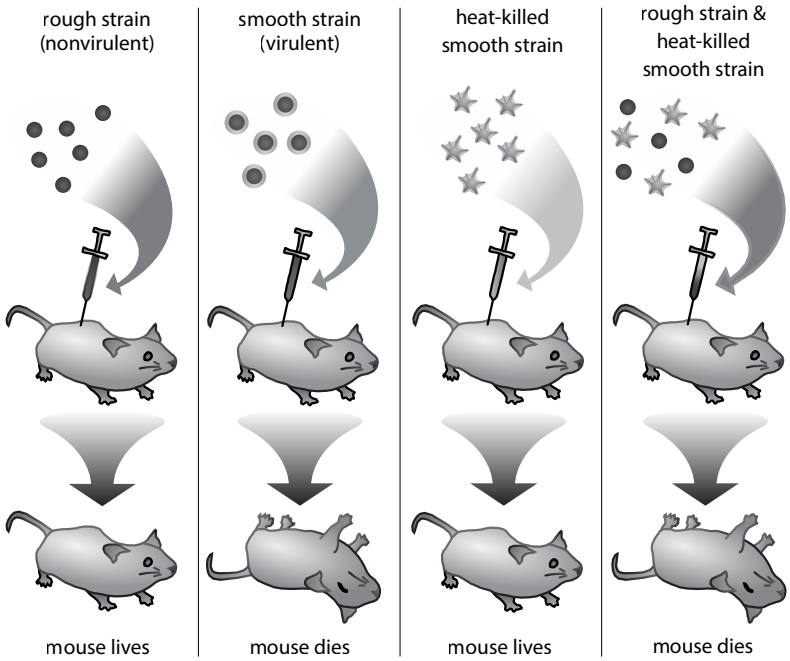


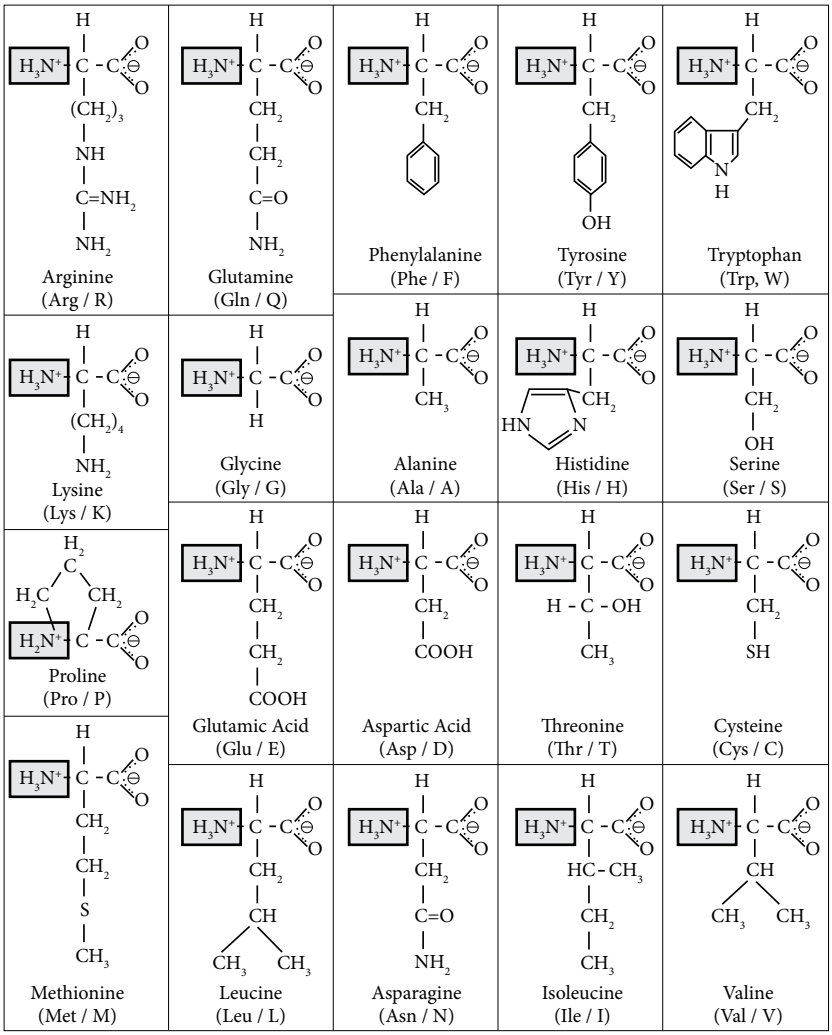
Figure 2.7. Griffith's experiments with rough and smooth bacterial cells. Upon injection in mice, the smooth strain of bacteria killed their hosts, while the recipients of rough strain bacteria lived. When Griffith killed a group of smooth bacteria with heat, they lost their killing ability. When he mixed the dead smooth bacteria with live rough bacteria, and then injected the mixture in mice, the recipients died. It was as if the rough cells became smooth in their virulence.

ments, he could distinguish two major types of cells, rough and smooth. These differed in the shape of the colonies they produced in a petri dish. They also differed in their virulence. Upon injection in mice, the smooth cells killed their hosts, while the recipients of rough cells lived (Figure 2.7).

Remarkably, when Griffith killed a group of smooth cells with heat, and then mixed the dead smooth cells with live rough cells, and then injected the mixture in mice, the recipients died. It was as if the rough cells became smooth in their virulence (Figure 2.7). Griffith never elucidated the reason why.[12]

Before 1944, a group of researchers succeeded in repeating Griffith's *in vivo* results (e.g., in the mouse) strictly *in vitro* (e.g., in the lab rough cells could be transformed into smooth cells without the need for a mouse host). In 1944, at the hospital of the Rockefeller Institute, three investigators (Oswald Avery, Colin MacLeod, and Maclyn McCarty) utilized the *in vitro* system and published an explanation for Griffith's discovery.

They discovered the explanation by purifying the transforming substance from heat-killed smooth cells and then sending it through a battery of tests to determine its identity. For example, Avery and colleagues examined the stability of the purified substance. In pure water, the purified material degraded quickly, while salt solutions preserved it. If the material was kept at high temperatures for an hour, it still transformed rough cells to smooth. When the material was dissolved in a highly acidic solution, the transforming ability disappeared. Consistent with these properties, direct chemical tests for DNA were positive.



Nitrogen present (phosphorus absent)

Figure 2.8. Distinguishing chemistry of amino acids. Amino acids contain copious amounts of nitrogen (select examples highlighted with black boxes) but, unlike nucleotides, do not contain any phosphorus. In the chemical diagrams themselves, chemical elements are represented with single letter abbreviations: carbon (C), nitrogen (N), oxygen (O), hydrogen (H), sulfur (S). The element phosphorus (P) is absent.

But was the rough-to-smooth transformation due to DNA? Or to some other chemical that contaminated the DNA preparation?

Attempts to remove fats did nothing to diminish the transforming ability. Conversely, the chemical constituents of the purified material were highly suggestive of pure DNA. Plenty of phosphorus was present, relative to nitrogen. Since amino acids do not contain phosphorus, but do contain copious amounts of nitrogen (Figure 2.8), this finding suggested that proteins were largely absent from the purified material. Furthermore, direct enzymatic* elimination of proteins from the purified material did nothing to stop transformation. Enzymatic elimination of another

* Enzymes are molecules (often, but not exclusively, proteins) that catalyze particular types of chemical reactions. In this case, the enzymes in this experiment catalyzed the severing of amino acid chains in proteins.

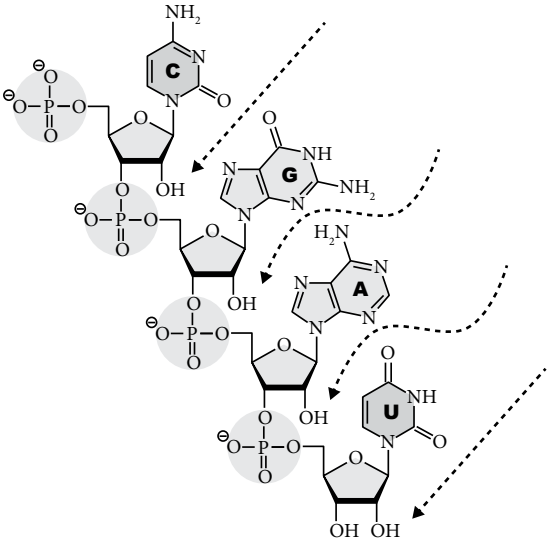


Figure 2.9. Chemical diagram of 4 linked RNA nucleotides. Elements are represented with single letter abbreviations: carbon (C), nitrogen (N), oxygen (O), hydrogen (H), phosphorus (P). Five-sided and six-sided chemical structures consist entirely of carbon and hydrogen, unless otherwise indicated. Chemical bonds indicated by lines (single bond by one line; double bond by two lines). Subscripts denote number of atoms of the adjacent element. The four different nucleotide subunits are designated by the bold, single-letter abbreviations for the defining bases (i.e., the six-sided chemical structures, or the joined five-sided and six-sided chemical structures): Adenine (A), cytosine (C), uracil (U), guanine (G). RNA is distinguished from DNA by the extra "OH" linkage below the five-sided chemical structure (highlighted with dashed arrows).

type of nucleic acid, RNA (Figure 2.9), was equally impotent. Only when enzymes specific for DNA were added to the purified material did the transforming ability vanish.

When the material was examined by other tests for purity, the results consistently showed a concentrated, highly pure substance. Under UV light, the material absorbed in the same wavelengths as DNA. When diluted, even to 1 part in 600,000,000, the material was still effective in transforming rough cells to smooth.

The results of Avery, MacLeod, and McCarty pointed squarely at DNA.[13]

Thus, red hair, chin shape, zebra stripes, the elephant's trunk, the giraffe's neck, the bald eagle's feathers, and the butterfly's wing patterns now appeared to be encoded, not by chains of amino acids, but by chains of nucleotides on chromosomes.

These bacterial transformation results were just the beginning of this shift in scientific opinion. New evidence demonstrated that DNA was, in fact, a very large molecule with much more theoretical capacity to store information. In addition, the chemistry of DNA was reanalyzed. The Austro-Hungarian biochemist Erwin Chargaff found that the different nucleotides did, in fact, vary in at least two ways. First, in species as diverse as bacteria, yeast, and human, the nucleotide ratios were not constant. This was compelling evidence that DNA might, in fact, have the potential to specify the diverse traits in these creatures.

Second, Chargaff found that, within an individual, the amount of the *adenine* nucleotide* ("A") was always approximately the same as the amount of the *thymine* nucleotide ("T"). The amount of the *cytosine* nucleotide ("C") was always approximately the same as the amount of the *guanine* nucleotide ("G"). However, the amount of A or T was not equivalent to the amount

* That is, the amount of the nucleotide containing the adenine nitrogenous base.

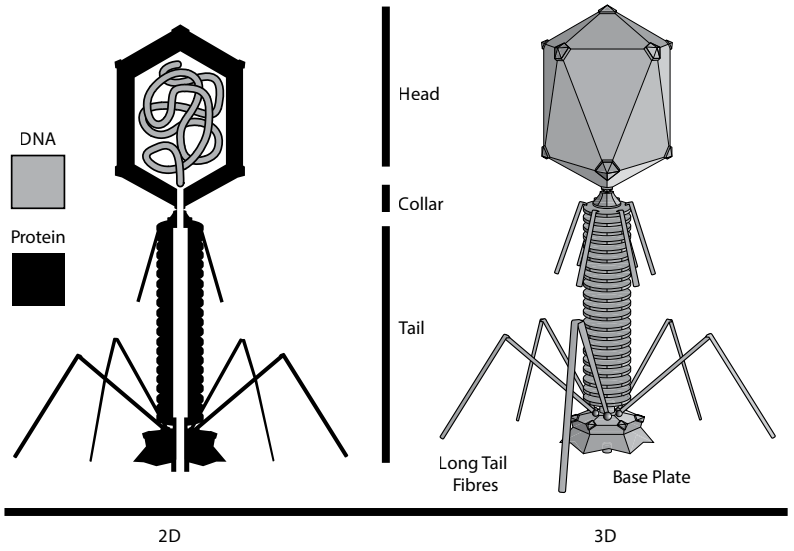


Figure 2.10. Structure of a bacteriophage.

of G or C.[14] Due to poor technology in the early 1900s, the ratios among the four DNA nucleotides were measured in error.

Then in 1952, two American biologists, Alfred Hershey and Martha Chase, reported a stunning find in viruses. Like Avery, MacLeod, and McCarty before them, Hershey and Chase utilized the chemical differences between proteins and DNA to identify which was the heritable material. But unlike Avery, MacLeod, and McCarty, Hershey and Chase utilized a bacterial system in which the primary subject of study was not bacteria. It was viruses — *bacteriophages* — that infect bacteria.

The experimental setup was straightforward. Bacteriophages have a comparatively simple structure. Their viral coats are composed of protein. The internal contents consist of DNA (Figure 2.10). Some of the amino acids in proteins contain sulfur (Figure 2.11), whereas nucleotides do not (Figure 2.12). Conversely, nucleotides contain phosphorus (Figure 2.12), but amino acids do not (Figure 2.11). Thus, the viral coats could be chemically distinguished from the internal contents.

When bacteriophages infect bacteria, they inject a certain material into their hosts, while another part of them remains behind. The injected material induces the bacterial hosts to make more bacteriophages, eventually causing the bacteria to burst open, or *lyse*, and release freshly synthesized bacteriophages. The fresh viruses then go on to infect other cells. For Hershey and Chase, the key question was the identity of the injected material.

To determine which molecule carried the information for making more bacteriophages, Hershey and Chase labeled one experimental group of bacteriophages with radioactive sulfur. Another they labeled with radioactive phosphorus. After letting the bacteriophages infect the bacteria, they isolated the extracellular bacterial contents and the intracellular bacterial contents. Radioactive sulfur showed up outside the bacteria. Radioactive phosphorus appeared in the bacteria cells (Color Plate 13). Therefore, nucleic acids were the heritable material of bacteriophages.[15]

But what about species that have multiple chromosomes? Which chemical carries the information in these creatures? A simple correlation suggested the answer. As we discussed earlier, sperm and egg are formed via the process of meiosis. Since meiosis results in just one member

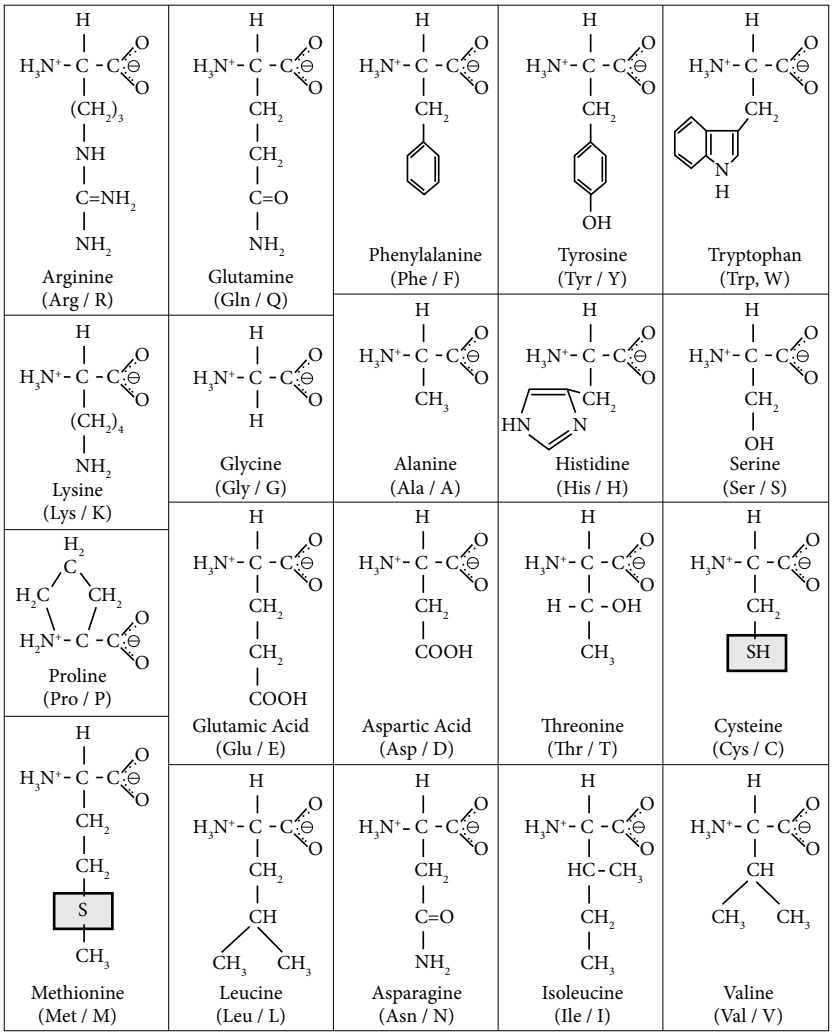


Figure 2.11. Additional distinguishing chemistry of amino acids. Unlike nucleotides, amino acids contain sulfur (examples highlighted with black boxes), but do not contain any phosphorus. In the chemical diagrams themselves, chemical elements are represented with single letter abbreviations: carbon (C), nitrogen (N), oxygen (O), hydrogen (H), sulfur (S). The element phosphorus (P) is absent.

of each chromosome pair in sperm or egg, meiosis reduces the genetic material by half in germ cells. Fertilization restores the full complement of genetic material. If it didn't, the chromosome number would increase each generation and eventually balloon to unmanageable amounts. Consequently, if DNA was the heritable material in species with chromosomes, DNA content should track with the pattern predicted by meiosis — twice as much DNA should be in somatic cells as compared to sperm and egg.

DNA matched this prediction. But the amount of protein did not.

In addition, since mitosis keeps the amount of genetic material constant in somatic cells, DNA content should be constant in these cells as well.

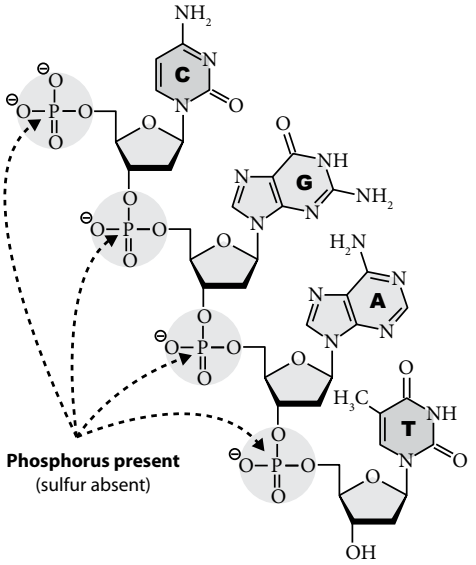


Figure 2.12. Distinguishing chemistry of DNA nucleotides. Elements are represented with single letter abbreviations: carbon (C), nitrogen (N), oxygen (O), hydrogen (H), phosphorus (P). Five-sided and six-sided chemical structures consist entirely of carbon and hydrogen, unless otherwise indicated. Chemical bonds indicated by lines (single bond by one line; double bond by two lines). Subscripts denote number of atoms of the adjacent element. Unlike amino acids, nucleotides possess phosphorus. The element sulfur (S) is absent.

It was.

Combined with the fact that the four nucleotide ratios varied widely among a diversity of species, these observations were highly suggestive of DNA being the physical substance of heredity. Furthermore, in species with chromosomes, DNA was found in the nucleus — the site of hereditary transmission. Proteins were found in both the nucleus and *cytoplasm* (the extra-nuclear space; Figure 2.13). Finally, the fact that UV light was *mutagenic* — UV light altered

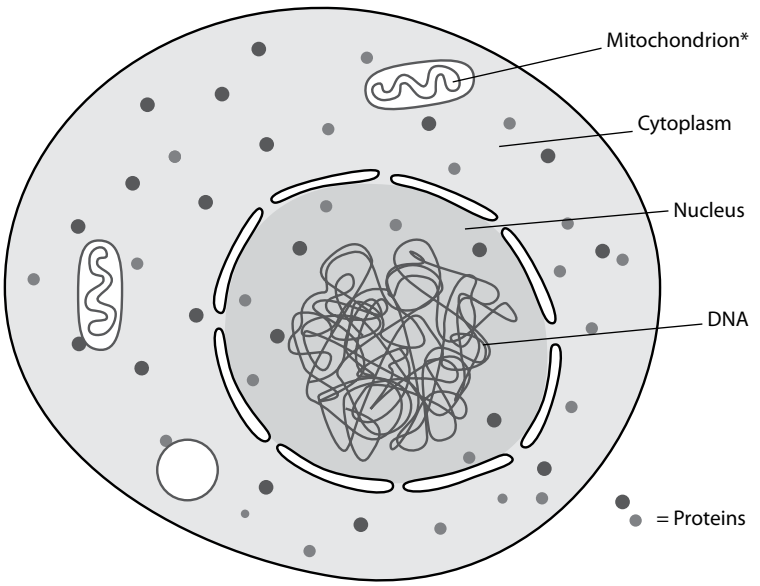


Figure 2.13. Basic elements and partitions of the cell. The two major partitions of the cell are the nucleus and cytoplasm. Proteins are found in both; DNA is found in the nucleus, but not the cytoplasm. *Technically, DNA is also found in another subcellular compartment, the mitochondria — a fact which is discussed at length in later chapters.

genetic material — pointed to DNA. The mutagenic activity was highest in the wavelengths of light in which DNA highly absorbed light, arguing that DNA was the substance of heredity in species with chromosomes.

Thus, traits were traced to the DNA in chromosomes.[16]

All of this evidence was compelling, but not complete. If DNA was linked to traits, how could nucleic acids physically carry the information for red hair, chin shape, zebra stripes, pine needles, and all the other traits that characterize the diversity of species on earth?

The answer to the *how* of DNA function was bound up in the structure of DNA. Many lines of evidence constrained the path to elucidating this structure. First and foremost, any potential structure for DNA would have to demonstrate how it could carry complex hereditary information. If the structure was simply repetitive unchanging units, it would imply a situation much more akin to the discarded tetranucleotide hypothesis than to a bona fide basis for heredity. In addition, any potential structure for DNA would have to satisfy the ratios that Chargaff discovered. In other words, the structure would have to explain why the amount of A and T in a cell were roughly equivalent, and why the amount of G and C in a cell were about the same.

Chemically, a proposed structure for DNA would need to be stable. If DNA was indeed the molecule of heredity, it would need to be stable over many generations. Since elephants produce elephants each generation, and snakes more snakes, and eagles more eagles, the structure of DNA would need to explain this fact. In contrast, a delicate molecule would fail to account for this consistency in traits.

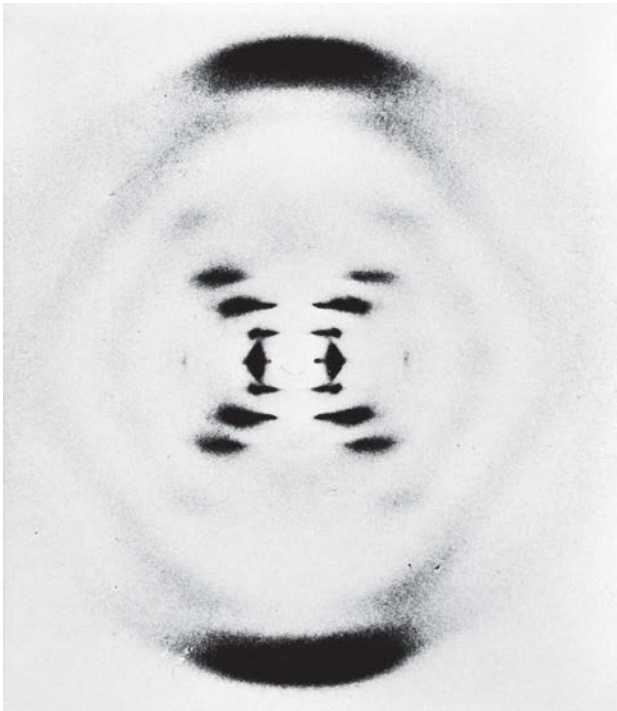


Figure 2.14. X-ray image of DNA.

Finally, the structure of DNA would have to suggest a means by which it could be replicated. Without faithful transmission, hereditary information would be diluted and extinguished in just a few cell divisions.

As science moved into the 1950s, indirect observations on the structure of DNA became available. Because the molecular structure of DNA was too small to be seen with visible light, even

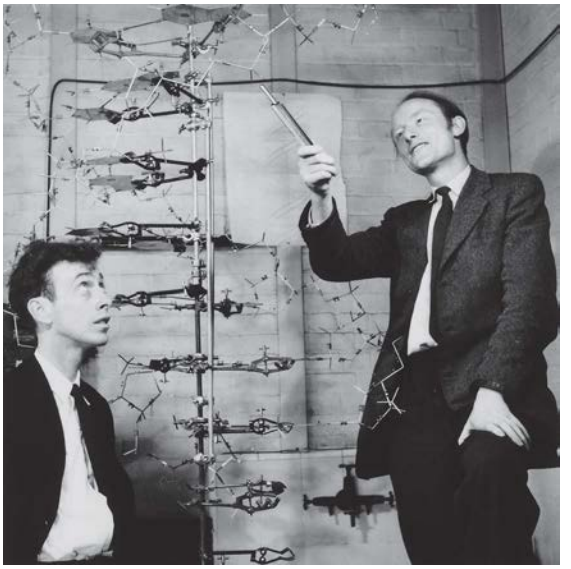


Figure 2.15. Solution to the structure of DNA. James Watson (left) and Francis Crick (right) assembled a physical model of DNA that solved the many constraints on the three-dimensional structure of DNA.

high-powered microscopes could not find the answer. However, x-rays could penetrate the spaces between the parts of the DNA structure. The x-ray image of DNA (Figure 2.14) was critical to orienting the molecule.

The limited components of DNA quickly narrowed the number of possible structures. For example, the carbohydrate and phosphate elements of nucleotides did not possess complexity such that they could carry reams of information. The carbohydrates and phosphates of nucleotides came in only one version. The nitrogenous bases, however, came in four different possibilities (Figure 2.5). Hence, the former suggested themselves as structural components; the latter, as information carriers.

But how would the structure be oriented? If the carbohydrates and phosphates formed the backbone, which way would it face — in or out? Would it be a stiff, straight chain of bases, or would the structure have a twist to it? How many chains would compose the molecule — one? Two? Perhaps even three?

After much toil, in 1953, James Watson and Francis Crick finally put the pieces together — literally. Using cut-outs of the various subunits (Figure 2.15), Watson and Crick put together a structure that arranged the nitrogenous bases in a manner that could carry information and that explained the ratios among them.

In short, DNA formed the now-familiar twisted double helix (Color Plate 14).[17] The nitrogenous bases faced inward and the carbohydrate-phosphate backbone outward. The weak electrostatic attraction between the nitrogenous bases held the structure together. Combined with the fact that the carbohydrate-phosphate backbone was water-soluble, the attraction and backbone produced a very stable structure that was consistent with the known spatial relationships.

However, the structure was not completely stable. Since the attraction between the nitrogenous bases were weak, they could be broken with sufficient force. Put simply, the DNA double helix could be unzipped (Color Plate 14).

This fact immediately suggested a means by which DNA could be replicated during cell division. If DNA was unzipped, each chain could act as a template for synthesis of a new DNA chain. Since A always paired with T, and since G always paired with C, the sequence of

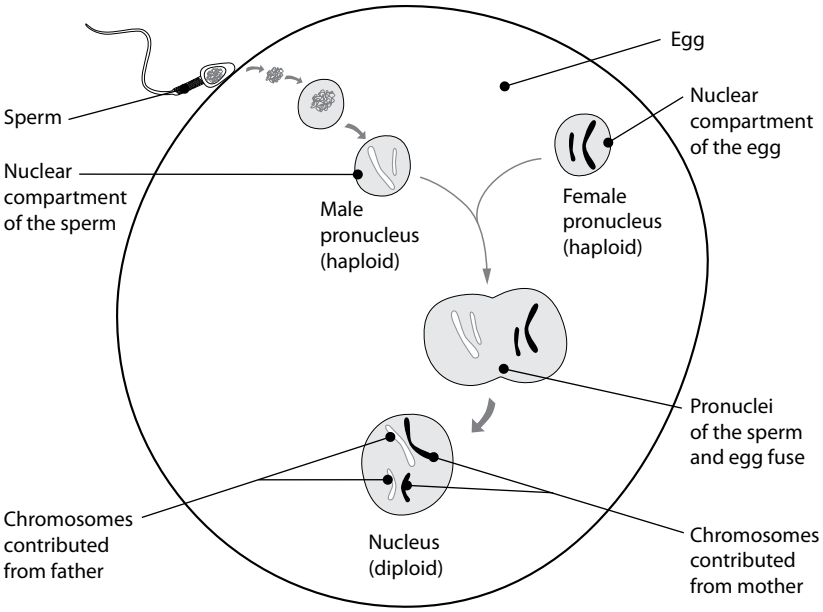


Figure 2.16. Fate of chromosomes during fertilization. When sperm fuses with egg, the sperm chromosomes ("male pronucleus") join the same nuclear compartment as the egg chromosomes (which began in the "female pronucleus"). However, the chromosomes are not melded together. Rather, these paternal and maternal chromosomes exist as individual entities, each carrying information from one of the parents. "Haploid" and "diploid" refer to states in which chromosomes exist as isolated entities or as pairs, respectively.

nitrogenous bases on one strand was sufficient to determine which bases belonged on the other — and vice-versa (Color Plate 14).[18]

The physical basis of heredity — the secret of life — had been solved.

Since the discovery of the double helix, the structural relationship between DNA and chromosomes has been uncovered. DNA doesn't exist in chromosomes as a long, straight stretch of helix. Chromosomes represent densely packed forms of DNA. In fact, to condense the packaging of DNA, cells wrap these helices around proteins in progressively higher levels of compaction to form the familiar chromosome shape (Color Plate 15). Thus, by watching chromosomes during cell division, we can watch the movement of DNA at a very zoomed-out level.

This perspective reveals what happens to DNA each generation. When sperm and egg fuse, the chromosomes of the sperm join the same nuclear compartment as the chromosomes of the egg (Figure 2.16). However, when this happens, the chromosomes are not melded together. The DNA helices along each of these chromosomes do not fuse. Rather, these paternal and maternal chromosomes exist as individual entities, each carrying information from one of the parents. In other words, pairs of chromosomes — and, therefore, of DNA helices — exist. Since both the father and the mother contribute an equal number of chromosomes, both parents make an equal contribution to the features of the newly conceived offspring. Some of the information from each parent might be hidden via the phenomena of dominance and recessiveness. But the information is still physically present.

In summary, the first major insights to mechanisms controlling trait behavior were slowly but steadily uncovered over a span of nearly 100 years. Hidden inside males and females was the microscopic code for visible traits. Males and females pass on this code via sperm and egg. The union of these cells produces the first cell — the *zygote* — which contains the instruction manual, not only for visible traits, but also for everything else that constitutes the individual members of a species.

We observed at the beginning of this chapter that species are ultimately defined by their traits. We then observed that traits are defined by genetics. Therefore, the origin of species is a fundamentally genetic question. Genetics defines the edge pieces of the puzzle of the origin of species.

Yet the physical basis for heredity — the nature of the code of life — was not uncovered until nearly 100 years after Darwin wrote *On the Origin of Species*.

Consider the significance of this fact, especially in light of the puzzle pieces that Darwin did possess in 1859. Fossils aren't inherited in sperm and egg. A miniature adult is not passed on through germ cells. A geographic location is not the substance of heredity. Instead, a set of instructions (encoded in DNA) is. Darwin tried to assemble a jigsaw puzzle of sorts without any edge pieces to guide his progress.

Without this genetic knowledge, could Darwin have speculated intelligently on the origin of species? If he had no idea how traits were coded and inherited each generation, could he have identified the origin of a particular trait? Before the advent of genetics, would his explanation have had any hope of being accurate?

The history of genetics poses a second set of questions to Darwin. Not only was his question a fundamentally genetic one, but his specific answers to the origin of species were also deeply tied to this field. For example, Darwin proposed that all species had one or a few common ancestors. In other words, he said that the vast diversity of life belongs to one or a few family trees. Genealogical relationships are directly recorded in genetics — and nowhere else.

Furthermore, Darwin claimed that new species arose via the process of survival of the fittest, or *natural selection*. Natural selection is useful to evolution if — and only if — the survivors pass on their superior traits to offspring. In other words, the mechanism of evolution is inextricably tied to inheritance. Inheritance is directly recorded only in genetics.

Finally, Darwin placed the origin of species on a very long timescale. However, the process of inheritance also acts like a timekeeper, independently recording the length of time over which species appeared (a concept we'll explore in detail in later chapters). How could Darwin have written *On the Origin of Species* without any genetic data to test his ideas? Since both his question and his hypotheses were deeply tied to inheritance, what prompted him, not only to pen, but also to vigorously argue for his proposal?

When Darwin wrote his most famous work, he took a scientific risk of massive proportions.